Cywilizacja XXI w.
– nowe rozwiązania technologiczne

Cywilizacja XXI w.
– nowe rozwiązania technologiczne

Redakcja:
Monika Maciąg
Kamil Maciąg

Lublin 2017

Wydawnictwo Naukowe TYGIEL składa serdecznie podziękowania dla zespołu Recenzentów za zaangażowanie w dokonane recenzje oraz merytoryczne wskazówki dla Autorów.

Recenzentami niniejszej monografii byli:

- dr hab. Jolanta Jaroszuk-Ściseł
- dr hab. inż. Arkadiusz Parus
- dr hab. Anna Zalewska
- dr inż. Henryk Bąkowski
- dr Barbara Futa
- d inż. Piotr Bogdan Grabowiec
- dr inż. Andrzej Marian Książkiewicz
- dr inż. Tomasz Kujawa
- dr inż. Sebastian Lipa
- dr inż. Agnieszka Makara
- dr inż Konrad Pietrykowski
- dr inż. Wojciech Sawczuk
- dr inż. Tadeusz Sarnowski
- dr inż. Maciej Trochonowicz

Wszystkie opublikowane rozdziały otrzymały pozytywne recenzje.

Skład i łamanie:
Monika Maciąg

Projekt okładki:
Marcin Szklarczyk

© Copyright by Wydawnictwo Naukowe TYGIEL sp. z o.o.

ISBN 978-83-65598-98-1

Wydawca:
Wydawnictwo Naukowe TYGIEL sp. z o.o.
ul. Głowackiego 35/341, 20-060 Lublin
www.wydawnictwo-tygiel.pl

Spis treści

Katarzyna Zdeb
Cyfryzacja w muzeach archeologicznych – wirtualna popularyzacja wiedzy 7

Wacław Brachaczek
Innowacyjne rozwiązania przy osuszaniu ścian budynków zabytkowych 18

Adam Krawiec, Patrycja Stasiak
Tiny House – wielkie pomysły architektoniczne na wykorzystanie małej przestrzeni ... 30

Tomasz Mazurkiewicz, Damian Ferenc
Aspekt bezpieczeństwa we współczesnych urządzeniach elektronicznych 49

Józef Okulewicz, Przemysław Żukowski
Przełamanie bariery rozwoju transportu 65

Magdalena Zawartka
Sterowanie aktywnym zawieszeniem pojazdu z uwzględnieniem regulatora liniowo-kwadratowego oraz algorytmu ślizgowego 80

Marta Paczkowska
Zastosowanie technologii laserowych do zwiększania odporności na zużycie elementów maszyn 102

Krzysztof Foit
Realizacja koncepcji Przemysłu 4.0 w ujęciu zagadnienia programowania robotów przemysłowych 113

Piotr Löschner, Krzysztof Jarosz
Przegląd wybranych programów MES do komputerowej symulacji procesów skrawania 129

Małgorzata Franz
Detekcja materiałów przy użyciu organicznych tranzystorów polowych 138

Małgorzata Skibińska
Zastosowanie nanorurek haloizytowych 157

Bogdan Noga, Marcin Mazur, Agnieszka Wichowska
Nowa technologia wspomagająca eksploatację wód termalnych w ciepłowniach geotermalnych 169

Krystyna Niesiobędzka, Ewelina Skiba
Nowoczesne technologie oczyszczania gleb zanieczyszczonych metalami ciężkimi 184

Kinga Górniak, Monika Tuleja
Indukcja wydajnej ryzogenezy z wybranych eksplantatów *Euphorbia milii* Des Moul. w warunkach kultur in vitro jako alternatywne źródła lateksu 192

Indeks Autorów 205

Katarzyna Zdeb[1]

Cyfryzacja w muzeach archeologicznych – wirtualna popularyzacja wiedzy

1. Wstęp

Rewolucja technologiczna, która nastąpiła w momencie rozpowszechnienia telefonów komórkowych, a później również między innymi tabletów, wymusiła niejako konieczność wzmożonego rozwoju pozostałych dziedzin życia. W tej grupie znalazły się także tzw. kultura wysoka czyli muzealnictwo oraz dziedzina nauk humanistycznych jaką jest archeologia. Wirtualna rzeczywistość oraz możliwość digitalizowania zbiorów, w tym również tworzenie ich kopii w formie 2D i 3D, stało się wyjątkową okazją do uatrakcyjnienia ofert muzealnych pod kątem społecznym i naukowym [19].

Termin muzeum wirtualne pojawił się w latach 90-tych. Zaadaptowanie rozwoju techniki do potrzeb muzealnych traktowane jest jako powrót do idei „muzeum bez ścian". Miało zostać zbudowane na podstawie reprodukcji przygotowanych w formie fotografii eksponatów. A. Malaraux (twórca tego pojęcia) postulował konieczność zmiany podejścia do prezentowanych dotychczas założeń muzealnych. Jego koncepcja miała przyciągnąć nowych odbiorców poprzez zapewnienie nowych doznania wizualnych i intelektualnych [5]. Koncepcja otwartego muzeum została powiązana z globalizacją i popkulturą. Z jednej strony całe kolekcje miały stać się ogólnie dostępne, niezależnie od ich położenia geograficznego. Natomiast z drugiej strony powinny być prezentowane na tyle atrakcyjnie, aby pozyskać nowych odbiorców, podnosząc jakość swoich usług [1, s. 163]. Otwarcie nowego rynku w postaci internetu, muzea stacjonarne zostały niejako sprowadzone do jednego poziomu z firmami konkurującymi w mediach społecznościowych [3].

2. Muzea tradycyjne, interaktywne i wirtualne

Za podstawowe zadania muzeów, oprócz gromadzenia zabytków, uznaje się naukę i edukację, zgodnie z art. 1 i 2 ustawy o muzeach [15]. Do pierwszej z nich należy „powinność wobec nauki i kultury" rozumiana jako szereg czynności związanych bezpośrednio z zabytkami [18]. W przypadku ruchomych zabytków archeologicznych wiąże się to z ich gromadzeniem i przechowywaniem. Ponadto pojedyncze przedmioty, jak i całe zbiory winne być poddawane opracowaniu przy jednoczesnym udostępnianiu wiedzy naukowcom. Warto również zwrócić uwagę, iż samo tworzenie kolekcji muzealnych jest jednocześnie elementem kreowania dziedzictwa kulturowego. Wszystkie wymienione powyżej działania mają na celu również ochronę dziedzictwa archeologicznego.

[1] k.h.zdeb@gmail.com, Instytut Archeologii, Wydział Nauk Historycznych i Społecznych, Uniwersytet Kardynała Stefana Wyszyńskiego

Drugi element działalności tych placówek nawiązuje do działań popularyzujących naukę. Stanowi to możliwość przetwarzania naukowych pojęć w taki sposób, aby stały się one dostępne dla szerszego grona odbiorców. Prezentacja kolekcji w interesujący sposób przyczynia się nie tylko do poszerzania wiedzy ludzi w różnym wieku odwiedzających placówkę. Ponadto stanowi element budowania tożsamości społecznej, który przyczynia się do rozpowszechniania konieczności podejmowania działań związanych z ochroną i zarządzaniem dziedzictwem [4].

Zmiany społeczne zachodzące w świecie w XX wieku spowodowały m.in. odmienne podejścia do muzeum. Instytucja przestała być jedynie budynkiem, w którym spotykają się elity, aby studiować wyeksponowane przedmioty. Wraz z dynamicznym rozwojem techniki, na przełomie XX i XXI wieku muzea musiały zmierzyć się z koniecznością „wyjścia" do zwiedzających i „wejścia" z nimi w pewną interakcję. Stanowi to możliwość do rozwoju sposobu prezentacji zabytków. Muzeum niejako „zaprasza" zwiedzających poprzez prezentację „oferty", której zadaniem jest prezentacja możliwości spędzenia wolnego czasu poprzez aktywny wypoczynek [6, s. 44, 57]. Dodatkową, nową cechą przejętą przez muzea jest łatwość dostępu. Wiąże się to z kulturą popularną, która umożliwia wszystkim otwarty i możliwie ułatwiony dostęp do zasobów, w tym także katalogów zabytków [6, s. 63].

Muzeum tradycyjne koncentrowało swoje działania wokół prezentowania zgromadzonych zbiorów, zapewniając jednocześnie możliwość obcowania z autentyczną substancja zabytkową [6, s. 65]. Na skutek zmian związanych z prowadzaniem nowych technologii w placówkach tych wprowadzono pewnego rodzaju „odświeżenie". Obecnie nawet muzea tradycyjne posiadają swoje witryny internetowe, chociaż w dużej mierze zawierają one jedynie informacje o działalności placówki, a brak w niej miejsca na wirtualny dostęp do kolekcji [17].

Pierwsza ze zmian dotyczyła zauważenia, iż prezentowane obiekty pozostają wyrwane z kontekstu. Dotyczy to przede wszystkim zabytków archeologicznych, które pozyskiwane są głównie z określonych miejsc takich, jak cmentarzyska czy osady. Nierzadko to właśnie kontekst definiuje funkcję odkrytego przedmiotu i staje się przyczynkiem do rozważań nad jego rolą w społeczności.

Definicja muzeum wirtualnego nie jest do końca jasna. Aneta Kamińska rozróżnia muzea tradycyjne od wspomnianych już wirtualnych, za główne kryterium podziału obierając posiadanie własnej siedziby oraz osobowość prawną, przez jednostkę muzealną. Ponadto rozgranicza je ze względu na rodzaj i sposób prezentowanych zbiorów. W muzeum wirtualnym eksponowane są kopie zabytkowych przedmiotów pochodzących z różnych stron świata, a ich opis jest znacznie bardziej rozbudowany. Obcowanie z dziedzictwem możliwe jest tylko poprzez zastosowanie pośredników w postaci ekranów komputerów lub urządzeń mobilnych. W opozycji do tej grupy znajdują się tradycyjne placówki, w których prezentowane są przede wszystkim oryginalne przedmioty [6, s. 214].

Również J. Kowalczyk-Anioł i J. Papińska-Kacperek rozróżniają jedynie dwa typy muzeów, przytoczonych powyżej. Dla tych autorek muzeum wirtualne pozbawione jest swojej siedziby, a zgromadzone zbiory stanowią jedynie elementy zdigitalizowanej bazy danych [7].

Natomiast M. Stefanik i M. Kamel dodatkowo definicję muzeów tradycyjnych uzupełniają o konieczność umiejscowienia ich w konkretnej rzeczywistej przestrzeni. Eksponaty prezentowane są poza zasięgiem zwiedzających. Natomiast muzea wirtualne tworzą możliwość do zapoznania się z zabytkami ruchomymi lub wnętrzami zabytków nieruchomych upublicznionych w internecie. Wirtualne prezentowanie zbiorów może stanowić uzupełnienie dla oferty edukacyjnej muzeum tradycyjnego lub stać się niezależną placówką funkcjonującą w przestrzeni wirtualnej. Do trzeciego rodzaju muzeów należą muzea interaktywne. Zwiedzającym prezentują możliwość skorzystania z technologii multimedialnych wraz z prawdziwym zabytkiem. Wystawy, pomimo iż prezentowane są w sposób cyfrowy, znajdują się w konkretnych budynkach muzealnych. W przypadku muzeów tradycyjnych, jak i interaktywnych, możliwości prezentacji zbiorów ograniczone są kubaturą budynku. Jednakże w muzea interaktywne stwarzają możliwości do szerszego zapoznania się z eksponatami w przestrzeni cyfrowej oraz z zabytkami, które nie zostały wyeksponowane. Takie muzea umożliwiają jednoczesny kontakt z wirtualnymi kopiami, jak i oryginałem umieszczonym obok przygotowanej prezentacji [11].

Wydaje się być zasadne stosowanie rozróżnienia pomiędzy muzeami wirtualnymi a interaktywnymi. Nie tylko ze względu na konieczność posiadania przez tę drugą grupę placówki muzealnej, co raczej ze względu na wejście w interakcję ze zwiedzającym wystawy, oferując jednocześnie możliwość skorzystania z rozszerzonej rzeczywistości. Stanowią one formę przejściową pomiędzy muzeum tradycyjnym a muzeum wirtualnym – całkowicie cyfrowym, dostępnym chociażby dla użytkowników tabletów i telefonów komórkowych z niemalże każdego miejsca na świecie.

Muzea, w których zastosowano nowe technologię, czy to jedynie wprowadzając pewne elementy takie, jak ekrany dotykowe, czy też przekształcając lub tworząc muzea wirtualne, pozwalają przede wszystkim na popularyzację dziedzictwa na szeroką – globalną skalę. Światowe dziedzictwo kulturowe staje się dzięki temu źródłem wiedzy o innych społecznościach i kulturach, jak również elementem edukacji na gruncie tolerancji poprzez budowanie wzorców postaw otwartych dla tego co niecodzienne.

W takiej edukacji muzealnej istnieje możliwość przedstawienia w różnoraki sposób, dostosowany do odpowiednich grup wiekowych, jak i stopnia niepełnosprawności zwiedzających, dziedzictwa kulturowego [16]. Możliwa jest prezentacja rekonstrukcji, która wiernie przedstawia cechy oryginału wraz z prezentacją naturalnego kontekstu i jego w nim funkcjonowania. Druga zaś polega na realizmie – zabytek umieszczany jest w tak odmienionym przez nowe technologie otoczeniu, żeby stworzyć wrażenie realności prezentowanych zdarzeń (np. odtworzenie filmu z eksperymentów, które przedstawiały powstanie przedmiotu lub inscenizacja otoczenia w jakim ów przedmiot funkcjonował). Możliwe jest również z przygotowaniem symulacji, które mają podkreślić historię prezentowanego przedmiotu poprzez wzmacnianie doznań, np. odgłosy ścinania drewna, rozpalania ogniska czy pracy kowala. Natomiast czwarta strategia (ludyczna) polega na upraszczaniu przekazu [6 s. 302]. Po wybraniu jednej lub kilku z powyższych strategii prezentacji eksponatów, należy dostosować do nich odpowiednie narzędzia będące środkami wyrazu planowanych treści edukacyjnych. W przypadku muzeum interaktywnego, jak i wirtualnego istnieje możliwość wykorzystania multimediów takich jak filmy, muzyka, wizualizacje i fotografie.

Natomiast tylko muzeum interaktywne umożliwia zaprezentowanie jednocześnie gablot z oryginałami. Stanowi to szansę na oddziaływanie na zaangażowanie zwiedzających poprzez wszystkie zmysły [6, s.169].

3. Cyfryzacja muzeów

Cyfryzacja muzeów, w tym również archeologicznych, stała się szansą dla ochrony dziedzictwa kulturowego. Nowe technologie umożliwiły zdigitalizowanie zbiorów – nie tylko zdjęć i dokumentów, ale również zabytków takich, jak m.in. naczynia ceramiczne, szklane puchary czy narzędzia krzemienne. Stanowi to niezwykle wygodną formę katalogowania zbiorów, które mogą być w taki właśnie sposób łatwo dostępne dla pracowników muzeum, jak i dla naukowców, korzystających z materiałów do prac naukowych [9].

Formą ochrony zbiorów jest zamieszczenie opisowych informacji o kolekcjach. Obecnie możliwe stało się również odtworzenie całych zabytków w formie 2D lub 3D. Ponadto taka forma zabezpieczenia zbiorów w celu ich zachowania dla przyszłych pokoleń, stała się również szansą na ich upowszechnienie poprzez udostępnianie w sieci [2]. Prezentacja eksponatów możliwa jest poprzez strony internetowe lub aplikacje na urządzenia mobilne, gdzie znajdują się pojedyncze zabytki lub całe kolekcje z odpowiednimi opisami. Możliwe jest zaprezentowanie obiektów lub całych wnętrz dzięki panoramie 360 stopni oraz animacji poklatkowej. Dodatkowy aspekt stanowią również filmy zamieszczane na YouTube, które mogą mieć formę kreskówek dla najmłodszych lub reportaży i wywiadów dla nieco starszych odbiorców. Dzięki temu, iż w przy takiej formie prezentacji nie istnieje ograniczenie przestrzenne, istnieje również możliwość prezentacji materiałów zazwyczaj niedostępnych. Stanowi to niezwykle interesujące uzupełnienie i wzbogacenie materiałów, a co za tym idzie także wiedzy odbiorców [10]. Ciekawą formą edukacji jest stworzenie aplikacji – gry nie tylko dla zwiedzających wystawę, ale także dla zainteresowanych osób, które będą mogły uczestniczyć w niej również, np. siedząc przed komputerem w domu lub idąc na spacer i korzystające z aplikacji pobranej na telefon komórkowy. Ponadto stwarza to również możliwość integracji uczestników lub odwrotnie – pozwoli na spersonalizowanie tras zwiedzania, np. pod względem wieku lub zainteresowań [1, s.160-161]. Dlatego też, niektóre jednostki muzealne decydują się na korzystnie z tzw. rozszerzonej rzeczywistości poprzez możliwość używania okularów do wirtualnej rzeczywistości. Zmysły zwiedzających są poddawane dużej liczbie impulsów zewnętrznych, aby pobudzić zmysły podczas poznawania elementów wystaw.

Natomiast w kwestii samego zwiedzania muzeów interaktywnych to oprócz aplikacji mobilnych, o których była już mowa powyżej, wciąż obecne są audio-przewodniki. Z kolei w muzeach wirtualnych ich rolę zastąpiły spacery wirtualne w formie, np. panoramy 360 stopni. Właśnie dzięki tak rozbudowanym pakietom ofert edukacji muzealnej pojawiła się szansa na przyciągnięcie nowych odbiorców i rozbudowania działalności oświatowej dla pokolenia odbiorców, które przywykło do korzystania z technologii w każdym aspekcie życia [8]. Muzea wirtualne mogą okazać się motywacją do zapoznania się z oryginalnym zabytkiem [1, s. 164]. Cyfryzacja wprowadzona do muzeów tak specyficznych, jakimi są placówki archeologiczne pozwala na przedstawienie w interesujący sposób zabytku ruchomego w jego kontekście przy jednoczesnej prezentacji sposobów jego wytwarzania, a nawet pozyskiwania materiałów do jego powstania.

Ryc. 1 Wirtualne Muzeum Chodlika – strona internetowa (chodlik.edu.pl)

Spośród muzeów archeologicznych do niniejszego rozważań wybrano placówki muzealne, które do swojej oferty wprowadziły nowe technologie. W wyborze kierowano się powyżej przedstawionymi wyróżnikami, uznając podział archeologicznych muzeów na wirtualne – nieposiadające własnej siedziby oraz interaktywne – z placówką, w której ulokowane ekspozycje w dużym stopniu zostały uzupełnione o elementy technologiczne.

Pierwszym z przykładów jest Wirtualne Muzeum Chodlika, które prezentuje rezultaty prac archeologicznych prowadzonych w Dolinie Chodelki (Ryc. 1). Projekt realizowany jest przez Stowarzyszenie Naukowe Archeologów Polskich Oddział w Warszawie we współpracy z Instytutem Archeologii i Etnologii Polskiej Akademii Nauk i Muzeum Nadwiślańskim z Kazimierza Dolnego, finansowany ze środków Ministerstwa Nauki i Szkolnictwa Wyższego. Na stronie internetowej prezentowane jest historia badań wykopaliskowych, które rozpoczęto na tym terenie już w drugiej połowie XIX wieku. Ponadto zaprezentowano współcześnie prowadzone prace, m. in. badania nieinwazyjne na terenie grodziska w Chodliku oraz projekt archeologia lasów Kotliny Chodelskiej. Wirtualne muzeum oferuje możliwość wzięcia udziału w spacerze stworzonym na bazie panoramy 360 stopni. Niezwykle ciekawym elementem projektu jest prezentacja zabytków archeologicznych pozyskanych z prac wykopaliskowych prowadzonych na badanym terenie (Ryc. 2.). Na stronie internetowej znajdują się również materiały filmowe oraz dźwiękowe (zespół Percival użyczył do projektu swoich utworów). Ponadto strona zawiera odnośniki do portalu społecznościowego Facebook i do fotograficznego serwisu społecznościowego Instagram. W tych miejscach zamieszczane są aktualne informacje o trwających pracach wraz z materiałem już opracowanym [12]. Wirtualne Muzeum w Chodliku przedstawia ciekawą ofertę edukacyjną dla osób zainteresowanych archeologią, jednakże całość projektu skierowana jest raczej do odbiorców zaznajomionych z archeologią a niżeli dla laików. Jednakże kontakt z takim muzeum może skłonić zwiedzających do dalszego pogłębiania wiedzy na temat archeologicznego dziedzictwa kulturowego.

Ryc. 2 Topór z badań archeologicznych – Wirtualne Muzeum Chodlika (chodlik.edu.pl)

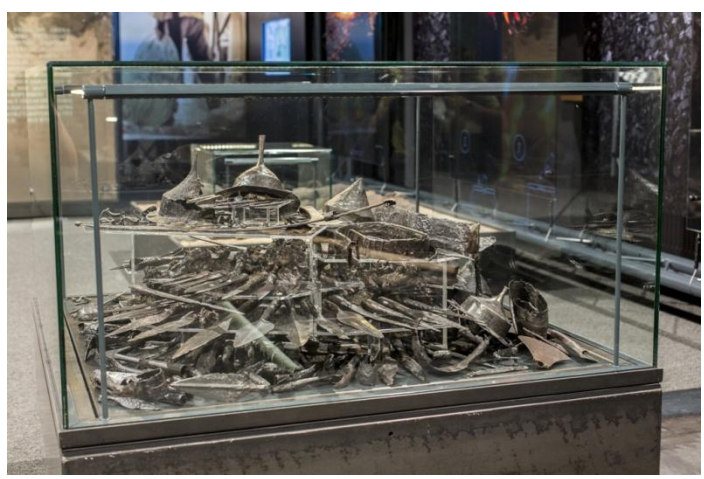
Drugim muzeum, które prowadzi intensywne działa w sferze wirtualnej jest Muzeum Starożytnego Hutnictwa Mazowieckiego im. Stefana Woydy w Pruszkowie. Instytucja ta, w przeciwieństwie do poprzedniej, posiada stałą siedzibę, jednakże oprócz zwykłej działalności edukacyjnej, prowadzą również intensywną kampanię promocyjną w mediach społecznościowych. Jest to przykład muzeum interaktywnego. Działalność tej placówki silnie związana jest z promocją i upowszechnianiem wiedzy archeologicznej. Dzieje się tak poprzez podejmowanie działań edukacyjnych nie tylko w postaci licznych zajęć i festynów, ale także poprzez aktywną obecność w sieci.

Poczynania pracowników muzeum można obserwować na Twitterze, Instagramie i Facebooku (Ryc. 3).

Ryc. 3 Muzeum Starożytengo Hutnictwa Mazowieckiego – strona internetowa (mshm.pl/)

Ponadto podczas przygotowań do festynu „Artefakty" powstały filmiki promujące wydarzenie, zamieszczane na serwisach społecznościowych.

Samo muzeum zostało utworzone na skutek prac archeologicznych prowadzonych w okolicy Pruszkowa. Zostało tam odkryte Mazowieckie Centrum Metalurgiczne. To właśnie na podstawie przeprowadzonych badań oraz na skutek zgromadzenia licznych zabytków. Działalność rozpoczęto 1 czerwca 1975 r. [13]. Wystawa, która obecnie znajduje się w muzeum nosi nazwę „Przedświt – Mazowieckie Centrum Metalurgiczne z przełomu er" (Ryc. 4). Na wystawie prezentowane są zarówno oryginalne eksponaty (w kilku gablotach), jak i kopie, które można wziąć do ręki (kościane igły, miecz, paciorki szklane i inne). Ponadto znajdują się tam również rekonstrukcji fragmentów domów czy pochówku konia. Jednakże największa uwagę zwiedzających przyciągają ekrany dotykowe, które umożliwiają uruchomienie filmu opowiadającego o sposobie wytapiania żelaza, gry i quizy dla najmłodszych oraz film poświęcony składaniu ofiar z koni. Co ciekawe do jego uruchomienia niezbędne jest użycie miecha. Dużym zainteresowaniem cieszy się także studnia znajdując się po środku jednej z sal, która umożliwia wirtualny spacer po gockiej wiosce znajdującej się w Masłomęczu. „Wejście" do studni pozwala również na przyjrzenie się bliżej filmikom nagranym w poszczególnych chatach, na których prezentowana jest różnego rodzaju wytwórczość. Podczas zapoznawania się z ekspozycją, zwiedzającym towarzyszą również odgłosy lasu. Warto również nadmienić, iż na wystawie znajduje się również rekonstrukcja 3D szklanego pucharka, który został odkryty na jednym ze stanowisk. Jest to zabytek unikatowy, a zaprezentowanie go odwiedzającym właśnie w takiej formie jeszcze bardziej podkreśla jego wyjątkowość. Muzeum Starożytnego Hutnictwa Mazowieckiego jest miejscem niezwykłym, nastawionym na interakcję ze zwiedzającymi. Muzeum to ma na celu przede wszystkim popularyzację archeologii przy użyciu najnowszych elementów technik [13].



Ryc. 4 Zabytki metalowe z wystawy znajdującej się w Muzeum Starożytnego Hutnictwa Mazowieckiego (mshm.pl)

Ostatnim spośród wybranych przykładów jest Wirtualne Muzeum Archeologiczne w Herkulanum (MAV) położone we włoskiej Kampaii. Pawilon muzealny mieści się w pewnej odległości od stanowiska archeologicznego, które prezentuje gościom pozostałości dawnych domostw wydobyte z wulkanicznego tufu. Jednakże samo muzeum, poza tematyką, nie jest związane z miejscem dawnych badań archeologicznych i funkcjonuje jako autonomiczna placówka.Podczas zwiedzania możliwy jest spacer czarnym tunelem, który co pewien czas rozświetlany jest prezentacjami, np. animacji patrolu żołnierzy, który spaceruje tuż obok (Ryc. 5). Można zajrzeć do rzymskiego domu, w którym ktoś pisze list na woskowej tabliczce. Istnieje możliwość zapoznania się z najbardziej znanymi zachowanymi freskami.

Ryc. 5 Wnętrze Wirtualnego Muzeum Herkulanum (fot. K. Zdeb)

Niektóre spośród zamieszczonych tam ekranów pozwalają zwiedzającym na zapoznanie się z definicjami haseł. Jednakże wszystkie treści prezentowane są w języku włoskim.

W poszczególnych pomieszczeniach wyświetlane są filmy, na których zaprezentowano uliczki dawnego miasta. Ponadto umożliwiają zapoznanie się z historią starożytnego Rzymu przez pryzmat odsłoniętych pozostałości miast Herkulanum i Pompei (Ryc. 6).

Ryc. 6 Prezentacja w sali poświęconej freskom z Pompei (fot. K. Zdeb)

W niektórych miejscach znajdują się mozaiki podłogowe, które poddano animacji. Wszystkie prezentowane eksponaty to hologramy, jak w przypadku jednego z pucharków lub wyświetlane animacje. Do niektórych spośród nich dodane zostały dźwięki, jak w przypadku patroli żołnierzy rzymskich. Można odnieść wrażenie, iż jest to miejsce przeznaczone jedynie dla dzieci, które będą miały okazję podczas wycieczki szkolnej zwiedzić to muzeum z przewodnikiem. Osoby dorosłe, zwłaszcza te znające się nieco bardziej na archeologii mogą odczuwać niedosyt pod względem dostępu do informacji, jak i do samych zabytków archeologicznych. Warto zwrócić uwagę na filmiki zamieszczone na YouTube, na których prezentowane są wystawy z muzeum [14]. Pomimo, iż samo muzeum może pozostawić pewien niedosyt wiedzy to jego bliskie sąsiedztwo dawnego obszaru prac archeologicznych, będących obecnie trwałą ruiną, stanowi niezwykłe uzupełnienie.

Każda z powyższych instytucji prezentuje zdigitalizowane zbiory w nieco inny sposób. Z punktu widzenia popularyzacji archeologii wszystkie są niezmiernie interesujące i ukierunkowane na nieco inne grupy odbiorców. Cyfryzacja muzeów od modyfikacji już istniejących placówek, jak miało to miejsce w przypadku pruszkowskiego muzeum, czy też tworzenie ich od nowa (Wirtualne Muzeum Chodlika), stanowi szansę ochrony dziedzictwa archeologicznego. Różnorodność prób uwspółcześniania ekspozycji przy jednoczesnym umożliwianiu dostępu do nich w sposób wirtualny zwiększa szansę na edukowanie kolejnych pokoleń pod względem wrażliwości na historię, ale także na różnorodność kulturową, która towarzyszyła człowiekowi od jego początków.

4. Podsumowanie

Wprowadzenie nowych technologii do muzealnictwa przyczyniło się nie tylko do digitalizacji gromadzonych zbiorów w celu zapewnienia im ochrony i możliwości przetrwania dla przyszłych pokoleń. Rozszerzenie oferty edukacyjnej o wirtualne spacery, nowe ekspozycje czy prezentowanie zabytków na ogół niedostępnych w placówkach stanowi punkt wyjścia do zainteresowania i rozbudzenia ciekawości zwiedzających. Wprowadzanie interaktywnych elementów do wnętrza muzeów umożliwia zintensyfikowanie doznań odbiorców, zatrzymanie ich nieco dłużej w placówkach, jak również zachęca do powrotu oraz zwiedzania innych, podobnych miejsc. Cyfryzacja placówek muzealnych związanych z dziedzictwem archeologicznym jest niezwykle wartościowa, również ze względu na fakt, iż same badania archeologiczne są niszczące. Każda ingerencja badacza sprawia, iż warstwy kulturowe zostają naruszone, a tym co pozostaje jest pozyskana informacja naukowa stworzona na podstawie znalezisk (warstw, obiektów, zabytków). Ponadto nowe technologie pozwalają na przybliżenie kontekstów odkryć archeologicznych również poprzez odtworzenie miejsca odkrycia czy użytkowania. Jest to niezwykle cenne, zważywszy iż archeolog spotyka się w swojej pracy z pozostałościami po ludzkiej obecności z nie zawsze kompletnymi informacjami. Wirtualna rzeczywistość stanowi szansę na przetrwanie i umiędzynarodowienie dziedzictwa kulturowego, przy jednoczesnym zapewnieniu mu jak najlepszych form ochrony.

Literatura

1. Bentkowska-Kafel A., *Muzeum wirtualne – muzeum bez granic?*, [w:] Muzealnictwo nr 54, 2013.
2. Berestecka Anna, *Wirtualne muzea Małopolski różne wymiary digitalizacji*, [w:] Muzealnictwo, vol 52, 2015, s. 126.
3. Brodowska K., *Wykorzystanie mediów społecznościowych na potrzeby promocji organizacji czyli twardy grunt pod nogami w świecie wirtualnym*, [w:] Cyfrowy archeolog, 2014, s. 58-60
4. Czopek S., *Muzea i zbiory archeologiczne – uwagi o misji i funkcjonowaniu*, [w:] Czopek S., Górski J., Między nauką a popularyzacją. Muzea i parki archeologiczne, Kraków, 2016.
5. Dziadek A., *Muzeum nowoczesności André Malraux*, [w:] Muzeum sztuki. Antologia, za: Aneta Kamińska (Re)konstrukcje muzeum w kulturze współczesnej. Studium socjopedagogicze, Poznań, 2015, s. 214-215.
6. Kamińska A., *(Re)konstrukcje muzeum w kulturze współczesnej. Studium socjopedagogicze* – praca doktorska, Uniwersytet Adama Mickiewicza w Poznaniu, 2015
7. Kowalczyk-Anioł J., Papińska-Kacperek J., *Wykorzystanie mediów elektronicznych w turystyce kulturowej na przykładzie muzeów i miejskich aplikacji mobilnych*, [w:] Turystyka Kulturowa vol 5, 2015, s. 7
8. Gontar B., *Cyfrowa rewolucja w muzeach*, [w:] Cyfrowa rewolucja w muzeach, 2013, s. 133-141.
9. Grzyb M., *Wirtualne muzeum – przestrzenna dokumentacja obiektów*, [w:] archiwum cyfrowe ASP, por. także Bentkowska-Kafel A., Muzeum wirtualne – muzeum bez granic?, [w:] Muzealnictwo nr 54, 2013, s. 161.
10. Manecki A., Wrzak J., *Projekt wirtualnego interaktywnego muzeum meteorytów z krakowskiego zbioru ING PAN*, s. 171-172.
11. Stefanik M., Kamel M., *Muzea i wystawy interaktywne w Polsce – współczesna atrakcja turystyczna*, [w:] Turystyka Kulturowa vol. 8, 2013, s. 6.
12. http://chodlik.edu.pl (wgląd: 03.08.2017).
13. http://mshm.pl/ (wgląd: 03.08.2017).
14. https://www.youtube.com/watch?v=PdvUacAfMHA (wgląd: 03.08.2017).
15. Ustawa o muzeach z dnia 21 listopada 1996 r., Dz. U. 1996 Nr 5 poz. 24.
16. http://www.polin.pl/pl/zaplanuj-wizyte-muzeum-dostepne/dostepnosc-muzeum-dla-osob-niepelnosprawnych (wgląd: 07.11.2017).
17. por. Chowaniec A., *Wystawy tradycyjne i muzea interaktywne*, [w:] Chowaniec A., Dziedzictwo archeologiczne w Polsce. Formy edukacji i promocji, Instytut Archeologii Uniwersytetu Warszawskiego, 2010.
18. Makuła P., *Kreacja obiektu muzealne na przykłądzie zarysu projketu wystawy Kolekcji Bieżanowskiej*, Muzealnictwo nr 50, 2009, s. 190-201
19. Bogucki, J., *O akcji popularyzatorskiej. Wstęp do wiedzy o sztuce*, Muzealnictwo nr 13, 1966, s. 159-163

Cyfryzacja archeologii – szansa na wirtualną popularyzację wiedzy

Abstrakt
W niniejszym artykule zaprezentowano możliwości wykorzystania nowych technologii w muzeach archeologicznych w celu promocji i popularyzacji wiedzy. Przedstawiono koncepcję „muzeum bez ścian", której odrodzeniem jest stworzenie otwartych muzeów w przestrzeni wirtualnej. Digitalizacja zbiorów stanowi możliwość do przygotowania nowych wystaw i form prezentacji ekspozycji, których nie ogranicza kubatura budynku. Rozwój technologii wpłynął również na przemiany w tzw. muzeach tradycyjnych, gdzie statyczne wystawy składające się z zabytków umieszczonych w gablotach zastąpiono muzeami interaktywnymi. W takich placówkach zwiedzający ma możliwość zapoznania się nie tylko z pojedynczym przedmiotem, ale przede wszystkim z kontekstem, który towarzyszył mu w momencie odkrycia, jak i w czasie, gdy był użytkowany. Cyfryzacja archeologii w placówkach muzealnych przyczynia się nie tylko do popularyzacji wiedzy wśród międzynarodowych odbiorców, ale także ma swój wpływ na ochronę dziedzictwa archeologicznego.
Słowa kluczowe:archeologia, muzea, dziedzictwokulturowe, cyfryzacja, museum wirtualne

Digitization of archeology – a chance for virtual popularization of knowledge

Abstract
This article presents the possibilities of using new technologies in archaeological museums to promote and popularize knowledge. The concept of "museum without walls" is presented, which revitalizes the creation of open museums in virtual space. Digitization of collections is an opportunity to prepare new exhibitions and forms of exhibition presentations, which do not limit the volume of the building. Technological development has also influenced the transformation of the so-called. Traditional museums, where static exhibits made up of monuments in display cases are replaced by interactive museums. In such institutions, visitors have the opportunity not only to familiarize themselves with a single object, but primarily with the context that accompanied it at the time of discovery and at the time when it was used. Digitization of archeology in museums contributes not only to popularizing knowledge among international audiences, but also has an impact on the preservation of archaeological heritage.
Keywords:archeology, museums, cultural heritage, digitization, virtual museum

Wacław Brachaczek[1]

Innowacyjne rozwiązania przy osuszaniu ścian budynków zabytkowych

1. Wprowadzenie

Wszystkie obiekty budowlane z biegiem czasu ulegają degradacji. Istotną rolę odgrywają czynniki zewnętrzne, które podzielić można na procesy atmosferyczne, biologiczne, oddziaływanie wody wraz z rozpuszczonymi w niej solami oraz czynniki mechaniczne. Porowata struktura materiałów, z których wykonane są mury, brak izolacji przeciwwodnych jak też nieprawidłowe odprowadzenie wód opadowych skutkuje zawilgoceniem murów. Jednym z mechanizmów przedostawania się wilgoci do muru jest transport kapilarny wody z gruntu. Wraz z wilgocią do murów przedostają się jony soli mineralnych. Jest to główną przyczyną złuszczania powłok malarskich, destrukcji tynków. Ponadto sól silnie absorbuje wilgoć z powietrza co zwiększa wilgotność murów i sprzyjać może rozwojowi grzybów i pleśni [1]. Zachowanie soli w porowatej strukturze materiału oraz ich oddziaływanie na substancję budowlaną w dużej mierze zależy od zawartości wilgoci w murze, rozpuszczalności soli, temperatury i wilgotności otoczenia [2]. Najwyższy poziom zawilgocenia występuje najczęściej w wewnętrznych częściach muru, z których odparowanie wilgoci jest utrudnione. Sole rozpuszczalne w wodzie transportowane są przez system porów kapilarnych do strefy odparowania na powierzchni muru, gdzie krystalizują w warstwach przypowierzchniowych lub na powierzchni ściany tworząc wykwity. W trakcie krystalizacji sole zwiększają swoją objętość, w wyniku czego w murach generowane są wewnętrzne naprężenia rozciągające obniżające ich wytrzymałość. W konsekwencji mury oraz tynki ulegają stopniowej degradacji. Ważnym czynnikiem wpływającym na trwałość zabytkowych ceglanych budynków jest, jakość cegieł i spajających je zapraw wapiennych [3, 4].

Renowacja obiektów zabytkowych jest procesem złożonym, na który składają się: analiza stanu murów – w szczególności określenie ich zawilgocenia i zasolenia, diagnostyka, dobór odpowiednich materiałów oraz opracowanie metodologii i technologii prac renowacyjnych.

2. Diagnostyka

Jednym z pierwszych podejmowanych działań w renowacji budynków jest wykonanie analizy historycznej opartej na zachowanych dokumentach. Dostarcza ona informacji m.in. o technologii wykonania i rodzaju posadowienia obiektu. Źródłem wiedzy historycznej mogą być zachowane dokumentacje projektowe. Cenny zbiór informacji o obiekcie znaleźć można w karcie ewidencyjnej zabytków architektury i budownictwa, tzw. białej karcie. Poprawnie wykonana biała karta zawiera

[1] wbrachaczek@ath.bielsko.pl, Wydział Inżynierii Materiałów, Budownictwa i Środowiska, Akademia Techniczno-Humanistyczna w Bielsku-Białej.

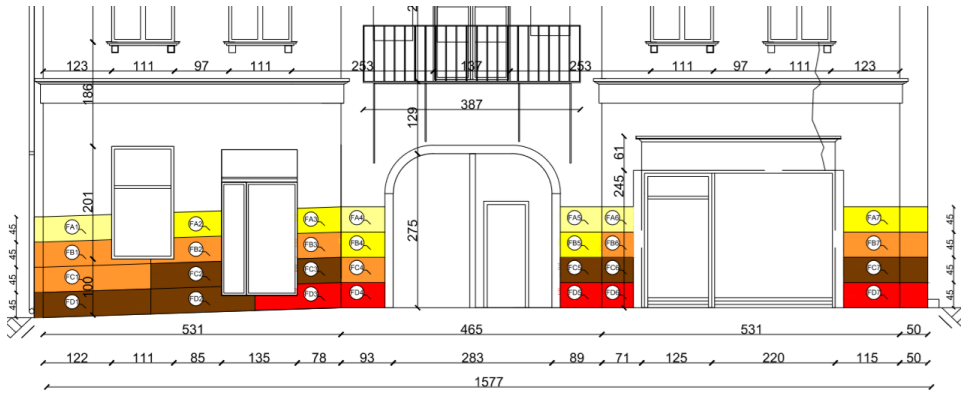
kompleksową informację o obiekcie dostosowaną do potrzeb i zakresu działań konserwatorskich. Do końca lat siedemdziesiątych dla obiektów zabytkowych wykonywano zielone karty, które stanowić mogą cenne archiwalne źródło informacji. Wiedzy na temat ewentualnych spiętrzeń poziomu wód gruntowych, lokalnych uszkodzeń i awarii obiektu, zmian uwidaczniających się podczas opadów deszczu dostarczyć może wywiad z mieszkańcami obiektu [5].

Do kolejnych działań diagnostycznych należą:

- **Badania makroskopowe** polegające na wstępnej obserwacji obiektu bez użycia specjalistycznego sprzętu. Mają one na celu ustalenie stopnia i źródła zawilgocenia obiektu przy uwzględnieniu czynników takich jak warunki gruntowo-wodne i ukształtowanie terenu. Ponadto badania obejmują rozpoznanie rodzaju muru i jego stanu, układu pomieszczeń oraz stanu izolacji jak też elewacji, występowania rys, spękań czy zacieków.
- **Ilościowa i jakościowa analiza szkodliwych soli** wykonywana jest laboratoryjnie na podstawie pobranych próbek. Mogą one służyć także do badania struktury muru, parametrów wytrzymałościowych, określania bilansu wilgoci a także do badań mykologicznych. Najpopularniejszymi metodami pobierania próbek są rdzenie wiertnicze, zwierciny, wykuwanie, wycinanie i zeskrobywanie materiału.
- **Pomiar zawilgocenia murów** dokonywany metodami: bezpośrednimi, do których należą metody suszarkowo-wagowe, półpośrednie – metoda karbidowa lub pośrednimi – np. mierników korzystających z technologii elektrooporowej, dielektrycznej czy mikrofalowej [5, 7, 9].

Uzyskane na podstawie działań diagnostycznych wyniki są podstawą do sporządzenia dokumentacji służącej do opracowania technologii renowacji obiektu zabytkowego. Jednym z takich dokumentów są mapy zawilgoceń, Na rysunku 1 przedstawiono przykład takiej mapy utworzonej w oparciu o wyniki pomiaru metodą suszarkowo/wagową. Jest to rozkład zawilgocenia na długości oraz wysokości przegrody. Wykonuje się je w oparciu o pojedyncze pomiary wilgoci, zwykle przeprowadzane na wysokości 0,5 m, jednakże w bardziej szczegółowych ekspertyzach, gdzie źródło zawilgocenia jest trudne do określenia, zaleca się wykonywać je na wysokości 0,5 m, 1 m oraz 1,5 m nad poziomem posadzki, względnie terenu w odstępach nie większych niż 2 m [8]. Wilgotność zaznacza się za pomocą warstwic, gdzie poszczególne kolory odpowiadają poziomowi wilgotności masowej (rys. 1). W literaturze polskiej przyjmuje się następujący podział ze względu na zawilgocenie murów W_m[%] [13]:

- W_m = 0-3% – ściany o dopuszczalnej wilgotności,
- W_m = 3-5% – ściany o podwyższonej wilgotności,
- W_m = 5-8% – ściany średnio zawilgocone,
- W_m = 8-12% – ściany mocno zawilgocone,
- W_m > 12% – ściany mokre.



Rys. 1. Rozkład zawilgocenia elewacji frontowej budynku [opracowanie własne]

3. Metodologia i technologia prac renowacyjnych zawilgoconych murów

3.1. Odtwarzanie bariery przeciwwilgociowej

Sposobem zahamowania niekorzystnych procesów związanych z obecnością wilgoci i soli jest odcięcie dopływu wody transportującej rozpuszczalne sole. Zabiegi te realizuje się na drodze fizycznej lub chemicznej. Odtworzenie fizyczne polega na mechanicznym przecięciu muru i wprowadzeniu do niego bariery wykonanej z nieprzepuszczalnego materiału, takiego jak metal lub tworzywo sztuczne [9]. Znaczna inwazyjność takiego przedsięwzięcia, gęsta zabudowa miejska w obrębie budynku, duże koszty oraz trudności z uzyskaniem zezwoleń konserwatorskich sprawiają, że metodę tą stosuje się w obiektach zabytkowych rzadko.

Bardziej popularnym rozwiązaniem jest chemiczne odtworzenie bariery przeciwwilgociowej. Polega ono na wykonaniu iniekcji poprzez nawiercenie w przegrodzie otworów, do których wprowadza się preparaty iniekcyjny o hydrofobowym charakterze. Najczęściej stosuje się preparaty na bazie krzemianów metali alkalicznych, alkilometylosiloksanów oraz kompozycji krzemianów alkalicznych i metylosiloksanów.

Ze względu na mechanizm działania preparatów iniekcyjnych wyróżnia się trzy grupy tych preparatów:

- zmniejszanie średnicy kapilarów lub całkowite ich zamknięcie,
- hydrofobizacja kapilar na skutek działania iniektu,
- hydrofobizacja ścian kapilar przy jednoczesnym zwężeniu ich światła.

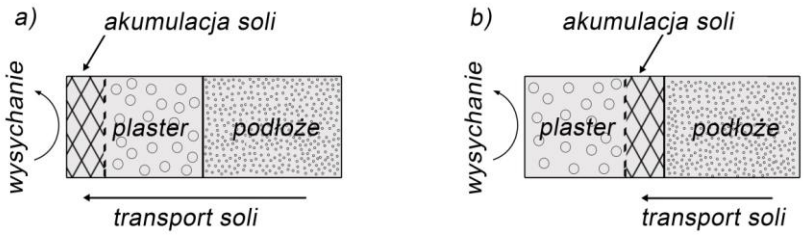
W dotychczasowej praktyce preparaty iniekcyjne dostępne były w postaci ciekłej. Sposób wprowadzania substancji iniekcyjnej zależny jest od wybranego rodzaju iniekcji. Była to albo metoda grawitacyjna, polegająca na bezciśnieniowym wprowadzaniu iniektu do muru bądź też za pomocą niskociśnieniowych pomp iniekcyjnych. Do wykonania iniekcji niskociśnieniowej używane są pompy iniekcyjne wyposażone w pakery jednokrotnego lub wielokrotnego użytku o średnicy ok. 20 mm, ciśnienie robocze oscyluje w zakresie 0,2-0,4 MPa. Zarówno jedna jak i druga metoda najlepiej sprawdza się w murach o jednolitej strukturze, pozbawionych pustek i pęknięć. W przypadku renowacji zawilgoconych budynków zabytkowych mury składają się z różnych materiałów, nierzadko budynki posadowione są na fundamencie

kamiennym. W przeszłości do wykonania fundamentów często wykorzystywano przypadkowe materiały, które następnie obmurowane zostały materiałami ceramicznymi. Konstrukcja takie nazwane są murami żebraczymi. W takich przypadkach bezpośrednie stosowanie ciekłych iniektów często nie przynosi oczekiwanych rezultatów. Dobrym rozwiązaniem odtwarzania bariery przeciwwilgociowej w takich murach jest stosowanie iniekcji wielostopniowej. Operacja składa się najczęściej z dwóch etapów. W pierwszym przeprowadzana jest wstępna iniekcja polegająca na wtłaczaniu, do wykonanych wcześniej otworów, upłynnionej zaprawy systemowej przeznaczonej do wypełniania rys i pustek. Po upływie 24 godzin ponownie nawierca się otwory, odpyla się je i przeprowadza się właściwą iniekcję, najczęściej niskociśnieniową [8, 10].

Innym rozwiązaniem jest stosowanie kremów iniekcyjnych produkowanych na bazie silanów i siloksanów oligomerycznych. Stosowany w tej metodzie specjalny krem iniekcyjny umieszcza się w aplikatorze, a następnie równomiernie wyciska się jego zawartość wewnątrz otworu w kierunku od końca nawiertu do lica ściany, zwracając przy tym szczególną uwagę, by równomierne rozprowadzić substancję i wypełnić nią cały otwór. Po całkowitym wchłonięciu kremu otwór wypełnia się zaprawą systemową lub zaszpachlowuje się szlamem systemowym [8, 10].

3.2. Tynki renowacyjne

Samo wykonanie przepony, niezależnie od sposobu, nie gwarantuje jednak wyschnięcia muru. Jest to jeden ze sposobów odtwarzania izolacji poziomej, w efekcie jej zastosowania mur może wysychać, ale nie musi. Pozostawienie zaś muru samemu sobie po wykonaniu takiej przepony może wręcz przyspieszyć jego degradację. Po utworzeniu bariery przeciwwilgociowej i rozpoczęciu wysychania ściany, powyżej przepony może dojść do wykrystalizowania się soli wewnątrz murów, co wywoła ich dalsze niszczenie. W celu uniknięcia zniszczeń związanych z krystalizacja soli konieczne jest usunięcie soli z murów, które można przeprowadzać za pomocą wyekstrahowania z zastosowaniem kompresów z materiałów absorbujących sole, takich jak kompresy z wapna czy pulp drzewnych. Sole bardzo dobrze migrują też do tynków porowatych [9]. Tynki takie nazywane są tynkami renowacyjnymi. Stosowane są one w kompleksowej renowacji zawilgoconych i zasolonych murów wskutek podciągania kapilarnego wody, w których przewiduje się obok odtworzenia bariery przeciwwilgociowej, uszczelnienie pionowe w obszarze cokołu i fundamentów oraz zastosowanie preparatów do zwalczania korozji biologicznej. Istotną funkcją systemu tynków renowacyjnych jest wspomaganie osuszania i odsalanie ścian [12]. Ze względu na właściwości i strukturę, sam system tynków renowacyjnych nazywa się je również tynkami szerokoporowymi, solochłonnymi, kompresowymi lub traconymi. W zależności od właściwości rozróżnia się dwie podstawowe kategorie tych tynków: transportujące i akumulujące sole oraz hydrofobowe [8, 10, 12]. W tynkach transportujących i akumulujących sole zakres wielkości porów jest szeroki. Obok porów aktywnych kapilarnie, o średnicy od kilku do kilkudziesięciu mikrometrów, występują też pory przerywające przepływ cieczy. Są to tzw. pory powietrzne o średnicy powyżej 70 mikrometrów. W wyniku sił kapilarnych wilgoć z muru przechodzi do tynku transportującego wilgoć. W nim następuje odparowanie wody, a sole wykrystalizowują w porach powietrznych, w wyniku czego nie powstają zniszczenia tynku.



Rysunek 2.Transport i akumulacja soli w tynkach solochłonnych: model transportowy a) i akumulacyjny b) [opracowanie własne]

W tynkach hydrofobowych nie zakłada się transportu wilgoci w formie ciekłej. Cechy te maja bezpośredni związek z krystalizacją soli zawartych w wilgoci. Rysunek 2, przedstawia zasadnicze różnice w akumulacji soli dwóch rodzajów tynków: transportujących wilgoć i hydrofobowych. Jak można zauważyć przy stosowaniu samego tynku renowacyjnego hydrofobowego, sole będą krystalizowały na styku warstwa tynku–podłoże. Rozwiązanie takie, aczkolwiek dopuszczalne tylko przy niewielkim zasoleniu podłoża, w dłuższym czasie ograniczać będzie dyfuzję wilgoci z podłoża do powierzchni zewnętrznej tynku. Sole będą krystalizowały na niewielkiej przestrzeni i, z biegiem czasu, wypełniały będą przestrzeń porów niezbędną dla dyfuzji pary wodnej. Wadą takiego rozwiązania jest ograniczona możliwość odprowadzania wilgoci przy dużym zawilgoceniu lub/i znacznym obciążeniu muru solami.

Dobrą skuteczność w odsalaniu i akumulacji soli uzyskuje się przez połączenie w jeden system tynków transportujących i akumulujących z tynkami hydrofobowymi. Poprzez odpowiedni dobór mikrostruktury porów oraz właściwości sorpcyjnych tych tynków można przyśpieszać szybkość osuszania ścian przy jednoczesnym tworzeniu przestrzeni magazynującej sole wewnątrz porowatej struktury – w taki sposób, aby uniknąć powstania szkód związanych z krystalizacją [8, 10, 12].

O ile zalety wynikające ze stosowania iniekcji chemicznej z systemem tynków renowacyjnych do osuszania murów są znane i wiele budynków zabytkowych jest poddanych renowacji tą technologią, to świat nauki prowadzi dyskusje dotyczące wpływu czynników, na jakość i skuteczność tych zabiegów.

4. Część badawcza

4.1. Cel badania

Celem badań była ocena jakościowa prac renowacyjnych zawilgoconych murów uwzględniających odtworzenie bariery przeciwwilgociowej metodą iniekcji chemicznej z zastosowaniem różnych preparatów. Analizę przeprowadzono na przykładzie konkretnych budynków wpisanych do Ewidencji Zabytków. W ocenie jakości prac posiłkowano się pomiarem zmian wilgotności murów w czasie od zakończenia prac renowacyjnych.

4.2. Obiekt badań i materiały

Badania przeprowadzono na trzech budynkach zlokalizowanych w Bielsku-Białej, w których zastosowana została kompletna technologia osuszania ścian. Do badań wybrano budynki pochodzące z podobnego okresu, o podobnej konstrukcji murów.

Budynek 1

Zabytkowa parterowa kamienica, która wzniesiona została na przełomie XVIII i XIX wieku w stylu klasycystycznym. Obiekt jest częściowo podpiwniczony, fasadę budynku stanowi kompletny wystrój architektoniczny z elementami klasycystycznymi oraz historyzującymi. Główne elementy konstrukcję obiektu to: fundamenty z kamienia naturalnego, ściany piwnic murowane z cegły pełnej i z kamienia naturalnego, ściany nadziemia murowane z cegły pełnej, strop nad piwnicą – sklepienie kolebkowe ceglane, pozostałe stropy o konstrukcji drewnianej.

Budynek 2

Obiekt zlokalizowany jest w centrum miasta przy ul. 11-go listopada, w przejściu ulicy, która stanowi główny deptak oraz historyczne centrum handlowe. Pochodzi on z pierwszej połowy XIX wieku. Jest to budynek jednopiętrowy częściowo podpiwniczony, ze strychem i oficyną w podwórzu, wykonany w technologii tradycyjnej z cegły pełnej, kryty dachem dwuspadowym. Elewacja wykonana jest tynkiem cementowo- wapiennym.

Budynek 3

Budynek przy ulicy Stanisława Dubois. Pochodzi z końca XIX wieku i jest posadowiony na ceglanych ścianach fundamentowych.

Pod względem usytuowania budynki były w zabudowie zwartej, kamienicowej z jednej strony graniczyły z ulicami, za domem natomiast znajdowało się podwórko, często wybetonowane, oddzielone od sąsiadujących kamienic dylatacjami.

4.3. Ocena stanu budynków

Stan techniczny budynków można ocenić, jako dostateczny, jednak ze względu na duże zawilgocenie ścian i niszczenie murów, konieczne jest wykonanie renowacji polegające na odtworzeniu bariery przeciwwilgociowej oraz przeprowadzeniu remontu elewacji. Podczas oględzin nie stwierdzono spękań, ugięć, korozji mechanicznej lub biologicznej elementów konstrukcyjnych budynków.

Pomiar wilgotności murów przeprowadzono metodą suszarkowo-wagową, materiał do badań pobrano za pomocą wybijaka ze środkowej części murów. Pomiary wykonane zostały na wysokości 10 cm, 60 cm, 110 cm oraz 160 cm powyżej poziomu gruntu. Średnie wartości wilgotności dla poszczególnych budynków, uzyskane dla ściany frontowej (z indeksem *f*) oraz od podwórza (z indeksem *z*) zestawione zostały w tabeli 1.

Tabela 1. Uśrednione wyniki pomiarów wilgotności ściany frontowej oznaczonej indeksem *f* i od strony podwórza z indeksem *z*

	Wilgotność masowa [%]					
Nr budynku	1f	1z	2f	2z	3f	3z
160cm	3.2	2.7	2.1	2.7	3.1	1.6
110cm	6.2	5.4	4.1	3.8	4.8	3.9
60cm	10.1	8.7	9.2	7.8	9.1	7.3
10cm	13.8	13.1	11.5	12.4	12.5	11.8

Źródło: Opracowanie własne

Najwyższy poziom wilgotności odnotowano na wysokości 10 cm powyżej poziomu gruntu, wraz ze wzrostem wysokości wilgotność stopniowo malała. Na poziomie 160 cm dla każdego punktu pomiarowego wilgotność masowa jest niska lub bardzo niska. Rozkłady poziomu wilgotności wskazują, zatem na kapilarne pochodzenie wilgoci znajdującej się w ścianach obiektu. W oparciu o powyższe dane zdecydowano o stworzeniu przepony poziomej przerywającej zjawisko kapilarnego podciągania wilgoci z gruntu.

Tabela 2. Uśredniony stopień zasolenia ścian

Stopień zasolenia [%]						
Nr budynku	1f	1z	2f	2z	3f	3z
azotany (NO_3^-)	0.19	0.06	0.24	0.08	0.0	0.0
siarczany (SO_4^{2-})	1.32	0.37	0.42	1.44	0.22	0.13
chlorki (Cl^-)	0.14	0.00	0.35	0.19	0.34	0.2

Źródło: Opracowanie własne

Analizując stopień zasolenia stwierdzono, że w przypadku ścian frontowych zawartość siarczanów i azotanów jest na poziomie średnim, w przypadku chlorków na poziomie zbliżonym do średniej. Natomiast w ścianach od podwórza i od strony przylegających kamienic zawartości siarczanów i azotanów w murach były niższe, a w przypadku chlorków bliskie zeru. Zaobserwowane różnice w przypadku chlorków wytłumaczyć można używaniem soli drogowej. Zaś przyczynami podwyższonego stężenia azotanów i siarczanów jest migracja soli z wodą podciąganą kapilarnie. Stopnie zasolenia ustalono w oparciu o tabelę 3. Pochodzącą z instrukcji WTA Merkblatt 2-9-04 [14].

Tabela 3. Stopnie zasolenia określone na podstawie ilościowej i jakościowej analizy soli [14].

Rodzaj soli	Stopień zasolenia [%]		
Azotany (NO_3^-)	<0,1	0,1-0,3	>0,3
Siarczany (SO_4^{2-})	<0,5	0,5-1,5	>1,5
Chlorki (Cl^-)	<0,2	0,2-0,5	>0,5

4.4. Opis technologii i materiały

Na podstawie zgromadzonych danych dokonano wyboru sposobu przeprowadzenia iniekcji ścian, jak też układu i grubości warstw tynków renowacyjnych. W analizie oparto się na materiałach posiadających dopuszczenie do stosowania w budownictwie, w renowacji obiektów zabytkowych, pozytywnie zaopiniowanych przez Wissenschaftlich-Technische Arbeitsgemeinschaft für Bauwerkserhaltung und Denkmalpflege e. V. – Naukowo-Badawczą Grupę Roboczą ds. Restauracji Zabytków i Ochrony Budowli.

4.4.1. Wykonanie iniekcji

W celu przerwania kapilarnego podciągania wilgoci, utworzona została przepona pozioma, na drodze iniekcji. Na podstawie wstępnych prób stwierdzono, że odtwarzanie bariery przeciwwilgociowej z zastosowaniem preparatów iniekcyjnych

w formie płynnej przeprowadzone powinno być dwustopniowo metodą ciśnieniową. Wprowadzenie iniektu do muru bezpośrednio po nawierceniu otworów nie pozwoliło na uzyskanie założonego ciśnienia aplikacji jak również związane było ze znacznym zużyciem iniektu. Świadczyło to o występowaniu dużej ilości szczelin i pustek w fundamentach budynków. Drugim wariantem, było zastosowanie kremów iniekcyjnych.

Do odtworzenia bariery przeciwwilgociowej w poszczególnych obiektach zastosowano trzy różne preparaty iniekcyjne:

- Budynek 1: Renowator 100 preparat iniekcyjny na bazie alkilometylosilikonów,
- Budynek 2: Renowator 120 preparat iniekcyjny na bazie silanów i siloksanów,
- Budynek 3: Renowator 190 krem iniekcyjny na bazie silanów i siloksanów.

W murach wykonano otwory o średnicy 20 mm, o rozstawie osiowym wynoszącym 10 cm, na głębokości o 5 cm mniejszej od grubości murów. Otwory wykonano w jednym rzędzie, którego wysokość dla części niepodpiwniczonej wynosiła 20 cm od poziomu gruntu, zaś dla części podpiwniczonej na wysokości 10 cm od posadzki.

Do wykonania wstępnej iniekcji wykorzystano zaprawę uszczelniającą do murów Renowator 200, którą aplikowano stosując agregat – pompę iniekcyjną D1 IC firmy INJECTOR.

Wtórne wykonanie otworów przeprowadzono w dwóch wariantach:

- **Wariant 1**: po stwardnieniu zaprawy, (co przeciętnie wynosiło od 3 do 4 dni) przez ponowne nawiercanie,
- **Wariant 2**: przez nakłuwanie wstępnie stwardniałej zaprawy przy pomocy prętów stalowych.

Właściwą iniekcję Renowatorem 100 i Renowatorem 120 przeprowadzono za pomocą pompy iniekcyjnej. W tym celu do przygotowanych otworów zamontowano pakery, które połączone zostały z pompą. Aplikacje prowadzono przez okres 15-20 minut, utrzymując stałe ciśnienie 3-5 barów. Średnie zużycie preparatu iniekcyjnego wynosiło ok. 15 dm^3/m^2 przekroju poprzecznego muru.

W budynku nr 3 w jednej strefie krem iniekcyjny wprowadzono bezpośrednio do otworów bez iniekcji wstępnej za pomocą pompy, zaś w strefie drugiej po iniekcji wstępnej zaprawą uszczelniającą.

Po przeprowadzeniu iniekcji otwory zamknięto Renowatorem 200.

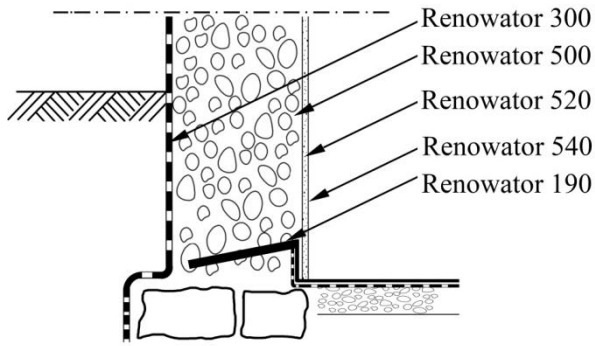
4.4.2. Wykonanie tynków renowacyjnych

W przypadku elewacji frontowej stare tynki zewnętrzne zostały usunięte, fugi pogłębione do 2 cm, a następnie całość muru została odpylona i oczyszczona. Na tak przygotowane podłoże zastosowano system tynków renowacyjnych o następującym układzie warstw:

- preparat do wzmacniania podłoży Renowator 350,
- obrzutka półkryjąca Renowator 500 o grubości 5mm,
- podkładowy tynk Renowator 520 o grubości 15 mm,
- tynk renowacyjny Renowator 540 o grubości 15 mm.

W celu zabezpieczenia warstw cokołowych budynków przed wodą rozbryzgową, elewację od strony ulicy, do wysokości 50 cm od powierzchni gruntu zabezpieczono dwoma warstwami elastycznej cementowej mikrozaprawy uszczelniającej Renowator 300.

Wykonanie pozostałych zabiegów renowacyjnych fundamentów budynku schematycznie przedstawiono na rysunku 3.



Rysunek 3. Schemat przeprowadzenia zabezpieczenia fundamentów przed wilgocią gruntową [opracowanie własne]

4.5. Wyniki i ich omówienie

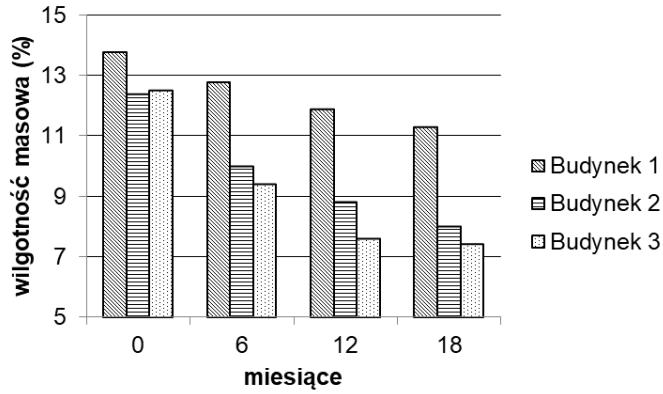
Na rysunkach 4a i 4b przedstawiono uśrednione wartości zmian wilgotności murów należących do budynków po renowacji, w funkcji czasu. Były to mury od strony podwórza niepodpiwniczonych części budynków. Ze względu na inwazyjny sposób przeprowadzenia pomiaru wilgotności (metoda suszarkowo/wagowa) zrezygnowano z pomiaru wilgotności murów od strony ulicy.

Na rysunku 4a przedstawiono wyniki dla budynków 1 i 2 – wariant 1 z wtórnym wykonywaniem otworów przez nawiercanie, zestawione z wynikami uzyskanymi dla budynku 3 ze strefy, w której preparat iniekcyjny wprowadzono metodą jednostopniową – bez wstępnego uszczelnienia muru.

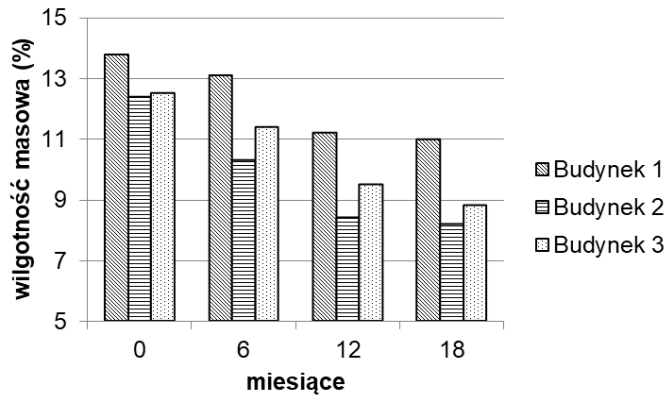
Na rysunku 4b przedstawiono wyniki dla budynków 1 i 2 – wariant 2 z wtórnym wykonywaniem otworów przez nakłuwanie wstępnie związanej zaprawy, zestawione z wynikami uzyskanymi dla budynku 3 ze strefy, w której preparat iniekcyjny wprowadzono metodą dwustopniową. Pomiary wykonane na wysokości 10 cm od poziomu utworzenia bariery przeciwwodnej.

Przedstawione wyniki wskazują, że najskuteczniejszym rozwiązaniem w odtwarzaniu bariery przeciwwilgociowej, mocno zawilgoconych murów budynków zabytkowych jest dwustopniowa iniekcja z wykorzystaniem preparatów na bazie silanów i siloksanów. W tym przypadku już po 6 miesiącach od wykonania iniekcji spadek wilgotności masowej murów wynosił ponad 20 % w stosunku do wilgotności sprzed renowacji. Dobre efekty uzyskano również stosując kremy iniekcyjne na bazie silanów i siloksanów. Również w tym przypadku, po 6 miesiącach, odnotowano spadek wilgotności masowej muru o ponad 20%.

a)



b)



Rysunek 4. Uśrednione zmiany wilgotności murów po renowacji, w funkcji czasu [opracowanie własne]

Najmniej skutecznym rozwiązaniem dla mocno zawilgoconych murów był preparat na bazie krzemianów. Mury wysychały bardzo powoli, po 18 miesiącach od renowacji spadek wilgotności wynosił 15%. Wyjaśnieniem może być fakt, że w skład tego preparatu wchodziły rozpuszczalne w wodzie sole potasowe kwasu metylokrzemowego, które po wprowadzeniu do przegrody stopniowo reagują ze znajdującym się w powietrzu dwutlenkiem węgla. W wyniku tej reakcji wydzielają się hydrofobizujący kwas polimetylokrzemowy oraz węglany alkaliczne. Niska skuteczność tego preparatu związana może być z utrudnionym dostępem dwutlenku węgla do mocno zawilgoconych murów, co czyni go mało skutecznym w takich przypadkach.

Stwierdzono ponadto, że sposób wykonywania wtórnych otworów w murze służących do przeprowadzenia iniekcji nie wywiera większego wpływu, na jakość utworzonej bariery. Wykonywanie otworów przez nakłuwanie wstępnie związanej zaprawy daje natomiast większą pewność, że iniekt wprowadzony zostanie do wypełnionej przestrzeni muru, co pozwoli uzyskać większą szczelność takiej przepony. W przypadku wtórnego wykonywania otworów w murach o znacznych grubościach, może się zdarzyć, że niewielkie zmiany położenia wiertła skutkowały będą powstawaniu kolejnych otworów w niewypełnionej części muru.

W przypadku stosowania kremów iniekcyjnych nieznacznie większa szczelność bariery przeciwwilgociowej (rys. 4a) związana była ze znacznym zużyciem kremu do iniekcji. Stosowanie wstępnej iniekcji pozwoliło na ograniczenie zużycia preparatu iniekcyjnego o ponad 50%, w stosunku do iniekcji jednostopniowej. Wyznaczone zużycie iniektów oraz szacunkowe koszty tych preparatów zamieszczono w tabeli 4.

Tabela 4. Symulacja kosztów preparatów iniekcyjnych do wykonania przepony poziome

Nazwa produktu	Typ preparatu	Cena (zł/litr)	Zużycie (litr/m2)	Wartość (zł)
Renowator 100	krzemianowy	15.88	15	238
Renowator 120	siloksanowy	271	1.2	325
Renowator 190 z iniekcją wstępną	krem iniekcyjny	192	1.3	249
Renowator 190 bez iniekcji wstępnej			3.0	576

5. Podsumowanie i wnioski

Dostępna literatura oraz badania wskazują, że iniekty na bazie silanów i siloksanów mogą być dobrym rozwiązaniem do osuszania mocno zawilgoconych murów.

Ze względu na konsystencję, korzystniej cenowo wychodzi krem iniekcyjny. Uzyskane wyniki dowodzą, że wpływ konsystencji jest niewielki. Dobre efekty uzyskać można tak przez zastosowanie tego typu środków w postaci roztworów, jak i w postaci kremów.

Przy stosowaniu kremów dobre efekty ekonomiczne można uzyskać stosując dwustopniowa iniekcję. Zaobserwowany niewielki spadek skuteczności utworzonej bariery rekompensuje znaczne obniżenie kosztów przeprowadzonej iniekcji, porównywalny do iniekcji z zastosowaniem iniektów na bazie krzemianów.

Stosowanie do odtwarzania barier przeciwwilgociowych w mocno zawilgoconych murów preparatów na bazie krzemianów jest najmniej skutecznym rozwiązaniem.

Przedstawione wyniki w dużej mierze mają charakter rozpoznawczy. Konieczne są dalsze, bardziej szczegółowe badania, uwzględniające wpływ różnego rodzaju murów, z których wykonane są fundamenty, ich grubość,

Literatura

1. Collepardi M. *Degradation and restoration of masonry walls of historical buildings*, Materials and structures., 23.2 (1990), s. 81-102.
2. Charola A.E. *Salts in the deterioration of porous materials: An overview*, Journal of the American Institute for Conservation 39(3) (2000) s. 327-343.
3. Dettmering T., Kollmann, H., *Putze in Bausanierung und Denkmalpflege*, DIN e. V., (2012).
4. Gorecki P., Wyrwał J. *Proces niszczenia murów ceglanych w zabytkowych budynkach i obiektach przemysłowych*., Roczniki Inżynierii Budowlanej 10 (2010).
5. Róziewicz M., *Ewidencja zabytków architektury i budownictw,* Kurier konserwatorski 1 (2008) s. 13-18.
6. *Prüfung eines Mörtels zum Entfernen von Salzen aus Ziegelmauerwerk*, Restoration of Buildings and Monuments (7.1) (2001) 79-90.

7. Barnat-Hunek D., Suchorab Z., Sobczuk H.: *Pomiary wilgotności murów z kazimierskiej opoki wapnistej przy zastosowaniu metody TDR,* Budownictwo i Architektura 2 (2008), s. 125-140.
8. Brachaczek W.: *Jak osuszyć budynek z wilgoci kapilarnej*, Inżynier budownictwa, 2 (2015), s. 75-82
9. Larsen, P. K. *The Development and Testing of a Salt Extracting Mortar,* Restoration of Buildings and Monuments 7.1 (2001) s. 79-90.
10. Magott C., *Odtwarzanie izolacji poziomej i pionowej podczas renowacji obiektów zabytkowych,* Izolacje, 4 (2013),s. 50-57.
11. Brachaczek W., Siemiński W., *Wpływ rodzaju preparatu iniekcyjnego na szczelność przepon przerywających kapilarne podciąganie wilgoci*, Materiały budowlane. 11, 2015, s. 172-174.
12. Lubelli B., van Hees R.P.J., Groot C.W.P., *Investigation on the behaviour of a restoration plaster applied on heavy salt loaded masonry*, Construction and Building Materials 20(9) (2006) s. 691-699.
13. Jasieńko J, Matkowski Z., *Zasolenie i zawilgocenie murów ceglanych w obiektach zabytkowych – diagnostyka, metodyka badań, techniki rehabilitacji*, Wiadomości Konserwatorskie, 14/2003.
14. Instrukcja WTA. Markblatt 2-9-04/D Sanierputzsysteme.

Innowacyjne rozwiązania osuszania ścian budynków zabytkowych

Abstrakt
Główną przesłanka opracowania jest pogłębienie wiedzy w zakresie poszukiwania skuteczniejszych technologii i materiałów do przeprowadzania renowacji zawilgoconych murów budynków zabytkowych. W rozdziale przedstawiono nowoczesne technologie osuszania murów budynków zabytkowych. Na przykładzie trzech budynków wpisanych do Ewidencji Zabytków, porównano skuteczność przeprowadzenia renowacji polegającej na odtworzeniu bariery przeciwwilgociowej na drodze iniekcji chemicznej z zastosowanym trzech różnych preparatów iniekcyjnych. Były to preparaty produkowane na bazie krzemianów i metylosilikonów w postaci roztworu oraz na bazie silanów i siloksanów w postaci roztworu i o konsystencji kremu. Największą skuteczność uzyskano stosując iniekty na bazie silanów i siloksanów, przy czym sposób aplikacji nie wywierał większego wpływu.
Słowa kluczowe: osuszanie ścian, iniekcja chemiczna, tynki renowacyjne, renowacja budynków.

Innovative solutions for draining walls of historic buildings

Abstract
The main objective of the study is to deepen the knowledge in seeking more efficient technologies and materials for restoration of damp walls of historic buildings. This chapter presents modern technologies for drying walls of historic buildings. The effectiveness of renovation consisting in restoring the moisture barrier by chemical injection with the use of three different injection preparations was compared using the example of three buildings listed in the Register of Monuments. The preparations were based on silicates and methyl silicones in the form of a solution and on silanes and siloxanes in the form of a solution and with a consistency of cream. The highest efficiency was achieved using silanes and siloxane based injections, with the application method being no major impact.
Keywords: wall drying, chemical injection, renovating plasters, renovation of buildings.

Adam Krawiec[1], Patrycja Stasiak[2]

Tiny House – wielkie pomysły architektoniczne na wykorzystanie małej przestrzeni

1. Wprowadzenie

Celem pracy jest przedstawienie i przeanalizowanie istniejących rozwiązań Tiny House oraz przeprowadzenie badań.

Tiny House to pojęcie stosunkowo nowe i rzadko używane w polskim środowisku. Jest to koncepcja wspierająca aktywne życie, brak trwałych zobowiązań, wolność finansową, niższe koszty życia i codzienności w pełnym komforcie, połączonym harmonijnie z minimalizmem. Osoby decydujące się na zakup mieszkania zobowiązują się ponieść największy wydatek, jaki przyjdzie ponieść w życiu. Niestabilna koniunktura rynków bankowych, rosnące koszty kredytów i potrzeba posiadania większego wkładu własnego skutecznie zniechęcają do zakupu własnego lokum. Termin „Tiny House" można rozumieć jako mały dom, często zwany także „domem w ruchu". Jest to opis architektonicznego oraz społecznego ruchu grupy ludzi, która mieszka w małych budynkach. Nie zdefiniowano jeszcze dokładnie, co stanowi tą małą przestrzeń, wiadomo jednak, że taki dom ma poniżej 50 m^2 i powszechnie jest uważany za małą przestrzeń.

Tiny House (przynajmniej jak na razie) nie doczekał się polskiej nazwy. Najłatwiej jednak opisać go jako „mini" dom.

W tym typie zabudowy mieści się wszystko – od sypialni po łazienkę, szafy, kuchnie i lodówkę. Budynki te mają działać w trzech obszarach, oferując domy dla osób, które szukają własnego lokum oraz chciałyby zainwestować w dom pod wynajem lub kupić lokum. Ten ostatni wariant pozwala na nabycie Tiny House, postawienie go w turystycznej lokalizacji i czerpanie korzyści z wynajmu. Jednak dopiero kupienie takiego miniaturowego obiektu na własność daje prawdziwe poczucie wolności. Gdy tylko przyjdzie nam ochota, wystarczy „podpiąć" dom niczym przyczepę kempingową do samochodu i ruszyć w drogę. Ogromne znaczenie ma też cena. Urządzony, gotowy do zamieszkania. Tiny House kosztuje tyle co średniej klasy samochód, kupiony w salonie. Cena może wahać się w zależności od zastosowania wyposażenia oraz materiałów. W wyniku użytych surowców wartość takiego obiektu może się różnić.

Zabudowa ta nie jest tworzona według jednego, typowego projektu lecz dostosowywana do potrzeb użytkowników. Tiny House, pomimo niewielkich rozmiarów jest znacznie tańszy niż kawalerka o podobnej wielkości. Budynki tego typu w zależności od lokalizacji i metrażu kosztują od 100 do nawet 500 tysięcy złotych.

Osoby decydujące się zamieszkanie w miniaturowym domu muszą liczyć się z pewnymi minusami takiego kwaterunku. Tiny House, przynajmniej od strony

[1] krawiecadam94@gmail.com, Politechnika Łódzka, Wydział Budownictwa, Architektury i Inżynierii Środowiska, www.bais.p.lodz.pl
[2] patusias2@interia.pl, Politechnika Łódzka, Wydział Budownictwa, Architektury i Inżynierii Środowiska, www.bais.p.lodz.pl

prawnej, nie jest domem. Traktowany jest jak przyczepa kempingowa czy camper, a to oznacza, że nie można go rozstawić tam, gdzie nam się podoba. Chociaż (jak przyznaje założyciel Tiny House), wszystko zależy od przepisów gminy na terenie, której postanowimy się rozstawić.

2. Micro Compact Home

Pierwszym przykładem przeanalizowanego Tiny House jest Mikro-kompaktowy dom o powierzchni 7m^2. Budynek ten jest wykonywany z materiałów o wysokim standardzie i przeznaczony na pobyt dla jednej lub dwóch osób. Ma on formę sześcianu 2.66 m x 2.66 m x 2.66 m . Wysokość wewnątrz w świetle wynosi 1,95 cm. Budynek ten jest wykonany w konstrukcji szkieletowej drewnianej. Na elewacjach zastosowano płyty aluminiowe, dzięki czemu zabudowa ta pozwala na dowolne spersonalizowanie, np. umieszczając grafikę lub zmieniając barwę paneli. Zastosowano szlachetne materiały co czyni budynek odpornym na działania czynników atmosferycznych.

Wewnątrz konstrukcji zastosowano płyty z pianki poliuretanowej PUR Nature- dzięki tak dobranym materiałom budynek ten spełnia wszystkie normy izolacyjności termicznej. Można go w pełni dostosowywać do różnych miejsc i okoliczności, a funkcjonalne przestrzenie do spania, pracy / jadalni, gotowania i higieny sprawiają, że nadaje się do codziennego użytku. W kuchni znajduje się kuchenka indukcyjna, mikrofalowa, piekarnik oraz niewielka lodówka.

Budynek ten posiada również dwa łóżka – jedno pojedyncze oraz podwójne. Micro Compact Home został opracowany przez zespół badaczy i projektantów z siedzibą w Londynie (głównym pomysłodawcą jest architekt Richard Horden), jako odpowiedź na rosnące zapotrzebowanie na krótkoterminowe zakwaterowanie dla studentów, przedsiębiorców, itd. Obecnie staje się on coraz bardziej popularny i łatwo dostępny w całej Europie. Jednak ze względu na swoją wysoką cenę (38,000 €), która nie zawiera w sobie montażu, podłączenia do usług oraz kosztów podatku wiele osób rezygnuje z tego typu rozwiązania.

Rysunek 1.. Bryła Micro Compact Home, Źródło: [8]

Rysunek 2.. Bryła Micro Compact Home, Źródło: [8]

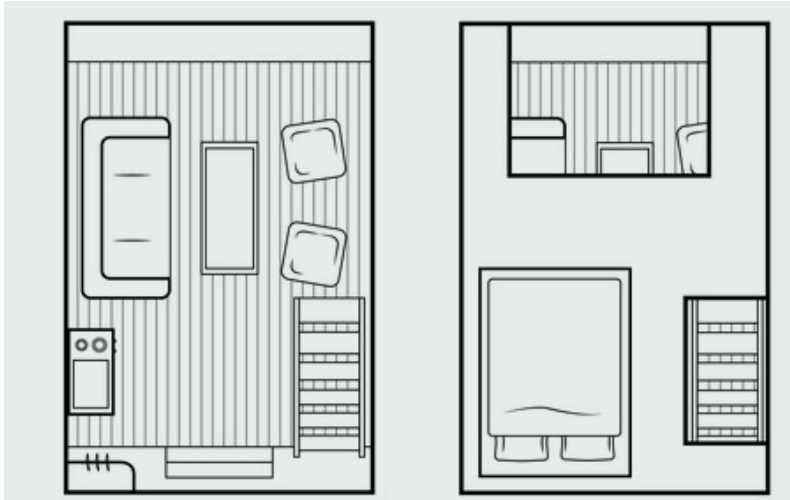
Rysunek 3.. Bryła Micro Compact Home, Źródło: [8]

Rysunek 4. Bryła Micro Compact Home, Źródło: [8]

3. Nido House

Kolejnym przykładem „małego domu" jest Nido House.

Jego powierzchnia wynosi zaledwie 10 m^2. Projektantem i twórcą tego obiektu jest Robin`a Falck`a z Finlandii. Obiekt ten nie wymaga żadnych pozwoleń na budowę. W tłumaczeniu z języka włoskiego „Nido" oznacza gniazdo ptaka. Użytkownik dzięki zastosowaniu dużych przeszkleń może poczuć się jakby nie znajdował się wewnątrz własnego budynku ale wśród natury. Przy budowie tego budynku zostały wykorzystane materiały pochodzące z recyklingu. Sam obiekt został wykonany w konstrukcji szkieletowej drewnianej. Elewacje zbudowano z okładziny drewnianej, dzięki której budynek w łatwy sposób komponuje się ze środowiskiem naturalnym. Na uwagę zasługuje także fakt, iż kąt nachylenia okien pozwala na swobodne przenikanie większej ilości promieni słonecznych do wnętrza obiektu. Dzięki takiemu „zabiegowi" wnętrze zyskuje na doświetleniu oraz przejrzystości. Również w nocy powierzchnia transparentna pozwala na obserwowanie i podziwianie gwiaździstego nieba. Sam rzut Nido House nie jest skomplikowany – oparty na kształcie prostokąta. Na parterze znajduje się salon oraz mikro-kuchnia wyposażona w niezbędne urządzenia np. kuchnię indukcyjną, piekarnik, lodówkę itp. Na piętrze zaprojektowano przestrzeń przeznaczoną do spania i wypoczynku. Na zewnątrz projektant przewidział sporą przestrzeń wypoczynkową w postaci dużego tarasu. Jednakże, mimo szeregu zalet jakie posiada to rozwiązanie, jego wadą jest forma, która może budzić kontrowersje wśród odbiorców.



Rysunek 5. Rzut Nido House, Źródło: [9]

Rysunek 6. Bryła Nido House, Źródło: [9]

Rysunek 7. Wnętrze Nido House, Źródło: [9]

4. Biodomek

Biodomek jest polską realizacją. Jego twórcami są Magdalena Górska (Architekt), Ewa Giczela (Sprzedaż), Moritz Reichert (Generalny Wykonawca), Mariusz Sarnicki (Budowniczy). Wszyscy twórcy należą do Ogólnopolskiego Stowarzyszenia Budownictwa Naturalnego, stowarzyszenia zrzeszającego profesjonalistów związanych z naturalnym budowaniem. Budynek ten został stworzony jako zdrowy, modułowy domek, który w przyszłości można modyfikować w bardzo łatwy sposób (dostawiając kolejne moduły). Został on wykonany z naturalnych materiałów. Podstawowy moduł ma wymiary 2,8 m x 6,44 m i wysokość około 3 m. Wymiary te zostały dobrane w taki sposób, aby dom ten mógł być w całości przewożony na przyczepie standardowej ciężarówki. Budynek może być modyfikowany dzięki wcześniej wspomnianej możliwości dostawiania kolejnych modułów – wzdłuż osi wschód-zachód lub na dachu, czyli w górę. Dom ten wznoszony jest na fundamentach punktowych dzięki czemu po czasie można go w łatwy sposób przetransportować w inne miejsce. Wspomniane moduły wykonane są w konstrukcji szkieletowej

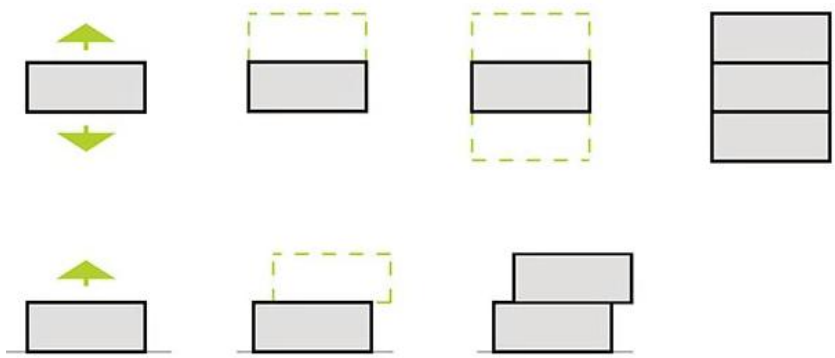
z drewna klejonego KVH 16cm, od środka wykończona płytami OSB EKO (klejonymi bez udziału formaldehydów) a od zewnątrz płytą drzewną (pełniącą funkcję wiatroizolacji) otwartą dyfuzyjnie DWD. Ocieplenie ścian, podłogi i dachu ekologiczną izolacją wdmuchiwana z celulozy (certyfikat Nature Plus), deklarowany współczynnik przewodzenia ciepła λD = 0,037 W/mK, U ścian = 0,19 [W/m²·K], U Dachu i podłogi = 0,18 [W/m²·K] – dzięki temu BioDomek spełnia obecne wymagania odnośnie izolacyjności cieplnej przegród dla budynków. Wentylowana elewacje została wykończona naturalnymi deskami z modrzewia olejowanego naturalnymi olejami. Drzwi zewnętrzne zaprojektowano jako drewniane z okuciami antywłamaniowymi i klamką. Witryny szklane wyposażono w pakiet dwuszybowy wypełniony argonem z ramką ciepłą, od zewnątrz szyba hartowana, bezpieczna. Cały system używa w jak największym stopniu materiałów naturalnych i biodegradowalnych. W zależności od ilości segmentów, obiekt ten może być użytkowany przez różną liczbę osób (jeden segment został zaprojektowany dla dwóch osób). Cena pojedynczego modułu wynosi 18,500 € – jest to cena podana bez podatku oraz montażu.

Rysunek 8. Pojedynczy segment Biodomek, Źródło: [1]

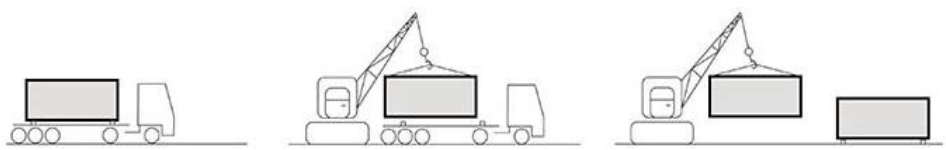
Rysunek 9. Podwójny segment Biodomek, Źródło: [1]

Rysunek 10. Montaż Biodomek, Źródło: [1]

Rysunek 11. Wnętrze Biodomek, Źródło: [1]



Rysunek 12. Ustawienie segmentów Biodomek, Źródło: [1]



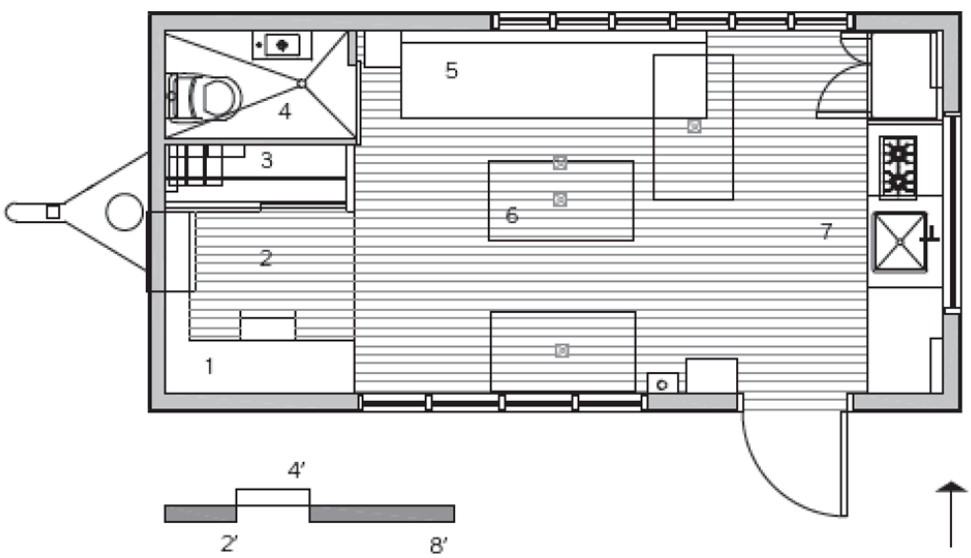
Rysunek 13. Transport budynków, Źródło: [1]

5. Minim Micro Home

Ostatnim przeanalizowanym przykładem mikro-domku jest Minim Micro Home. Obiekt ten ma powierzchnię 20 m^2 i stanowi największy przykład przedstawionych rozwiązań. Sama bryła budynku jest skromna, ale także elegancka. Mamy tu do czynienia z przykładem budynku modernistycznego z dwoma dużym przeszklenia. Obiekt ten jest wykonany w konstrukcji szkieletowej, na elewacjach zastosowano okładzinę drewnianą a od wewnątrz płyty kartonowo-gipsowe. Zastosowano tu również ekologiczne systemy odnawialnych źródeł energii m.in. panele fotowoltaiczne

zlokalizowane na dachu oraz biologiczną oczyszczalnię ścieków. Wewnątrz znajduje się pełnowymiarowa kuchnia z dużą przestrzenią do przygotowywania posiłków oraz niewielka przestrzeń przeznaczona do pracy.

Zaprojektowano tu także sporą przestrzeń dzienną, w której zastosowano „inteligentne" rozwiązania funkcjonalne (pozwalają one dostosować wnętrze do aktualnych wymagań użytkownika). Do rozwiązań tych można też zaliczyć m.in. łóżko wysuwane z podestu pod „gabinetem", mobilny stół/stolik – dzięki specjalnie umieszczonym zamocowaniom w podłodze może służyć jako stolik kawowy lub stół obiadowy itp. Część dzienna może w łatwy sposób zostać przekształcona w salę kinową, dzięki automatycznej rolecie, która całkowicie zasłania okno stając się jednocześnie ekranem kinowym. Do ciekawych rozwiązań warto zaliczyć także pomieszczenie toalety, która jednocześnie pełni funkcję łazienki. Aby „zaoszczędzić" miejsca zastosowano tu odpływ prysznicowy w podłodze. Budynek ten jest przeznaczony na pobyt 2-4 osób. Projektanci skupili się na stworzeniu przestrzeni dla młodych, aktywnych osób, które poszukują swojej przestrzeni do życia.



Rysunek 14. Rzut Minim Micro Home, Źródło: [9]

Rysunek 15. Bryła Minim Micro Home, Źródło: [9]

Rysunek 16. Wnętrze Minim Micro Home, Źródło: [9]

Rysunek 17. Wnętrze Minim Micro Home, Źródło: [9]

Rysunek 18. Wnętrze Minim Micro Home, Źródło: [9]

6. Metoda badawcza oraz jej wyniki

6.1. Cel ankiety

Dążąc do lepszego zrozumienia poruszanego zagadnienia, a także poznania opinii społeczeństwa zdecydowaliśmy się na przeprowadzenie ankiety środowiskowej.

Ankieta składała się z sześciu pytań mających na celu uzyskanie odpowiedzi od respondentów na temat:

- Przedziału wiekowego, w którym się znajdują,
- Wyboru rodzaju lokalu, w którym mieliby możliwość zamieszkania – Tiny House czy mieszkanie w centrum miasta,

- Opinii o rozwiązaniach typu Tiny House,
- Teoretycznej decyzji kupna lub wynajęcia małego domu,
- Wskazania trzech cech, jakimi powinien charakteryzować się taki obiekt,
- Badania czy dana grupa wiekowa spotkała się ze zjawiskiem Tiny House.

6.2. Omówienie wyników ankiety

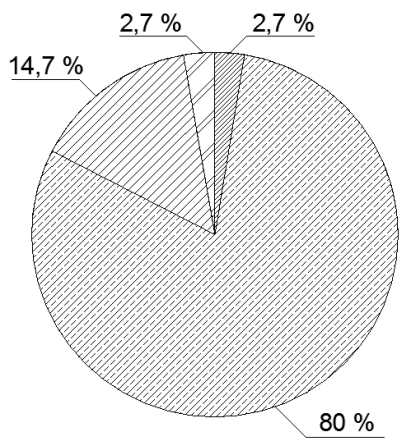
W badaniu wzięło udział 75 osób, które kolejno odesłały wypełnione ankiety przez portal internetowy.

Wyniki, które otrzymaliśmy przez e-mail posłużyły nam jako opracowanie poniższych wyników zachowując kolejność zadanych pytań oraz uzyskanych odpowiedzi.

Wyniki zostały przedstawione w formie kołowych diagramów, powyżej nich zapisano treść pytań a poniżej komentarz.

Pytanie numer 1

„W jakim wieku jesteś?"



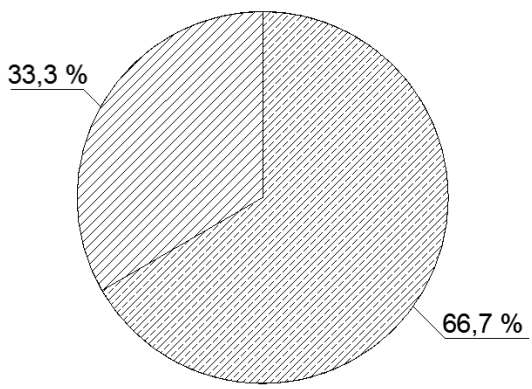
Odpowiedź	Odpowiedzi	Udział
Poniżej 19 lat	2	2,7 %
19- 25 lat	60	80 %
26- 32 lata	11	14,7 %
33 lata i więcej	2	2,7 %

Rysunek 19. Wyniki ankiety, Opracowanie własne

Z załączonego diagramu wynika, że większość respondentów odpowiadających na pytania znajdowało się w przedziale wiekowym 19-25 lata, stanowią oni 80% ankietowanych. Pozostałe przedziały wiekowe, w których znaleźli się odbiorcy badania to kolejno: 26-32 lata (14,7% badanych) oraz poniżej 19 lat i więcej niż 33 lata – po 2,7% dla każdego z przedziału.

Pytanie numer 2

„Gdybyś miał/a do dyspozycji kwotę, za którą możesz kupić sobie mieszkanie w centrum miasta (2-pokojowe) lub dom o porównywalnym metrażu i układzie funkcjonalnym z własnym ogródkiem na osiedlu położonym kilka kilometrów od centrum – którą opcję byś wybrał/a ?"



Odpowiedź	Odpowiedzi	Udział
Dom	50	66,7%
Mieszkanie	25	33,3%

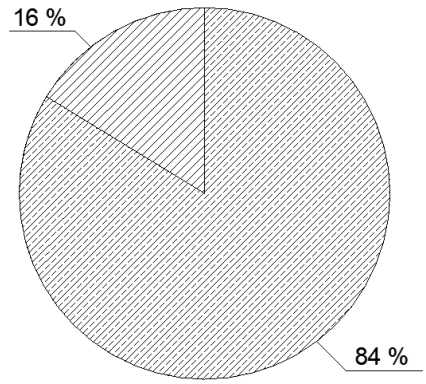
Rysunek 20. Wyniki ankiety, Opracowanie własne

Respondenci zostali poproszeni o zdecydowanie, w którym rodzaju budynku woleliby zamieszkać – w mieszkaniu czy Tiny House.

Spośród badanych, aż 66,7% wyraziłoby chęć posiadania na własność mniejszego domu zaś jedynie 33,3% osób wolałoby mieć na własność większe mieszkanie w porównywalnych kosztach.

Pytanie numer 3

„Czy podoba Ci się idea małego domu, który zachowuje wszystkie funkcje potrzebne do mieszkania, cechuje się małym metrażem, jego koszty utrzymania są porównywalne z wydatkami na mieszkanie w centrum miasta, posiada własny, niewielki ogródek"



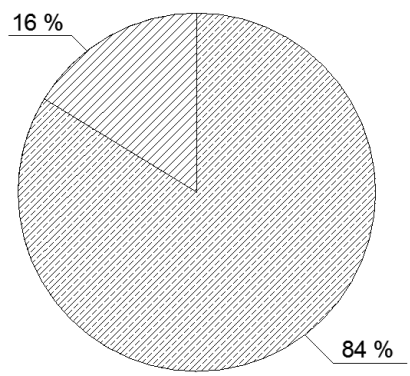
Odpowiedź	Odpowiedzi	Udział
Tak	63	84%
Nie	12	16%

Rysunek 21. Wyniki ankiety, Opracowanie własne

Spośród ankietowanych, sześćdziesiąt trzy osoby (84%) oceniło, że podoba im się idea Tiny House, jedynie dwanaście osób (16%) było przeciwnych temu typowi rozwiązania.

Pytanie numer 4

„Jeśli chciałbyś/-ałabyś mieszkać w takim domu, wolałbyś/-ałabyś kupić go na własność czy wynająć ?”



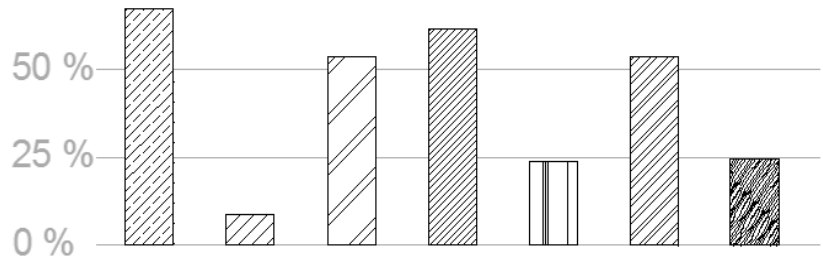
Odpowiedź	Odpowiedzi	Udział
Kupić	63	84%
Wynająć	12	16%

Rysunek 22. Wyniki ankiety, Opracowanie własne

Respondenci w swoich odpowiedziach w większości wykazali większą chęć kupienia na własność Tiny House – 63 osoby (84%), pozostali zaś (12 osób, co stanowi 16% ankietowanych) zaznaczyło, że chcieliby taki obiekt wynajmować.

Pytanie numer 5

„Zaznacz TRZY CECHY z niżej wymienionych, jakie powinien mieć taki dom"



Odpowiedź	Odpowiedzi	Udział
Parking położony na terenie osiedla lub podjazd przy każdym obiekcie	50	66,7%
Pralnia dostępna dla mieszkańców osiedla	7	9,3%
Wnętrze domu możliwe do modyfikacji (m.in. przesuwne ściany)	40	53,3%
Taras	46	61,3%
Dom powinien być parterowy	18	24%
Obiekt powinien mieć funkcję rozbudowy	40	53,3%
Salon osiedlowy (możliwość przyjmowania gości, organizacji imprez itd.)	19	25,3%

Rysunek 23. Wyniki ankiety, Opracowanie własne

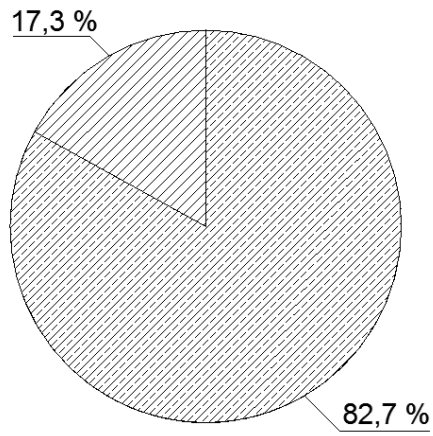
Ankietowani zdecydowali, że najważniejsze cechy jakie powinien posiadać Tiny House to:

- Parking zlokalizowany na terenie osiedla lub zaopatrzenie obiektu w podjazd – 50 odpowiedzi (66,7% badanych),
- Własny taras – 46 osób (61,3%),
- Wnętrze możliwe do modyfikacji oraz posiadania funkcji rozbudowy – po 40 odpowiedzi (co daje po 53,3% badanych),
- Przestrzenie wspólne na terenie osiedla domów – 19 osób (25,3% badanych),

- 18 osób zdecydowało, że obiekt powinien być parterowy (24% wśród ankietowanych),
- Ogólnodostępna pralnia w celu zaoszczędzenia miejsca we własnym domu – 7 osób (9,3%).

Pytanie numer 6

„Czy spotkałeś/-aś się z terminem „Tiny House" ? Jeśli tak, to gdzie (internet, prasa, tv, inne) ?”



Odpowiedź	Odpowiedzi	Udział
Nie	62	82,7%
Tak	13	17,3%

Rysunek 24. Wyniki ankiety, Opracowanie własne

Większość z ankietowanych – 62 osoby (82,7%) nie spotkało się wcześniej z pojęciem „Tiny House”, pozostali zaś – 13 respondentów (17,3%) miało styczność z takim terminem.

6.3. Analiza wyników

- Większość z respondentów (60 osób, co daje 80%) zadeklarowało, że znajdują się w przedziale wiekowym 19-25 lat.
- Ankietowani w większości stwierdzili, że woleliby posiadać mniejszy, mobilny dom z własną przestrzenią wokół niż mieszkanie w centrum o podobnym metrażu.
- Oceniając ideę „Tiny House” zdecydowana większość wśród badanych wyraziłoby chęć zamieszkania w tego rodzaju domu.
- Odpowiadając na pytanie nr 4 ankietowani odpowiedzieli, że preferowaliby wykupić taki dom na własność niż go wynajmować.
- Wydaje się być uzasadniony wybór najważniejszych cech, jakie powinien posiadać wymarzony Tiny House (pytanie nr 5). Najwięcej badanych zdecydowało, że na terenie osiedla powinien znajdować się parking dla samochodów a każdy dom mieć

zapewniony podjazd. Kolejno, badani zdecydowali że obiekt posiadać takie przestrzenie, jak m.in. taras , wnętrze możliwe do modyfikacji, czy też posiadać funkcję rozbudowy.

- Zdecydowana większość respondentów (63 osoby – 82,7%) nie spotkało się z terminem „Tiny House", pozostali (13 osób, co stanowi 17,3% ankietowanych) miało styczność z tego typu pojęciem.

7. Podsumowanie

Informacje uzyskane od respondentów w wyniku wyżej omówionej ankiety były pomocne przy pracy nad dokonaną analizą. Przeprowadzone badania pozwoliły ocenić czy owe zjawisko jest znane wśród społeczeństwa

w Polsce. Dzięki uprzejmości respondentów mogliśmy m.in. dowiedzieć się czy owa idea zostałaby przez nich przyjęta oraz jakimi cechami powinny się charakteryzować aby spełniać oczekiwania użytkowników.

8. Wnioski

Tiny House to idealne rozwiązanie dla osób młodych. Takie domy można w łatwy sposób modyfikować, a dzięki „mobilnym" meblom/rozwiązaniom w ciągu dnia wnętrze jest przestronne, a w razie potrzeby można wykorzystać m.in. wysuwane z podestu łóżko. Jednak w każdym z prezentowanych przykładów architekci wykonali ogromną pracę implementując funkcję całego domu (o powierzchni przynajmniej pięćdziesięciu metrów kwadratowych) w przestrzeń trzy razy mniejszą.

Duże znaczenie w tych działaniach miało zastosowanie nowych technologii które pozwoliły zminimalizować potrzebne miejsce np. nowoczesne systemy biologicznych oczyszczalni ścieków, czy też systemy ogrzewania wody. Sami twórcy starali się tworzyć przestrzenie bardzo funkcjonalne jednak niewielka przestrzeni spowodowała wykluczenie np. spiżarni, czy dodatkowych magazynków. Pomimo faktu, iż na świecie istnieje wiele ciekawych rozwiązań tego typu projektów, to żaden z nich nie jest pozbawiony wad. W większości przypadków kluczową rolę odgrywa wysoka cena zakupu pojedynczego modułu domku w cenę którego nie są wliczone koszty np. montażu czy podatku VAT. We wszystkich przeanalizowanych przykładach występuje konstrukcja szkieletowa.

Jest to jedno z lepszych i najpopularniejszych rozwiązań ponieważ pozwala ona na łatwą rozbudowę domu oraz jego transport. Wypełnienie ścian również zostało w inteligent sposób rozwiązane, ponieważ zastosowano płyty z pianki poliuretanowej, która jest ze względu na swoje właściwości bardzo lekka i ma bardzo dobre współczynniki termiczne.

Te wyjątkowe domy przyciągają uwagę swoim designem. Zastosowanie prostej bryły, dużych przeszkleń i zazwyczaj płaskiego dachu są definicją nowoczesnego stylu.

Z zewnątrz są starannie wykończone naturalnymi materiałami. Charakterystyczną cechą Tiny House są mobilność oraz kompaktowość. Zazwyczaj dla zaoszczędzenia stosuje się elewacje wykonane z drewna to rozwiązanie nie jest najtrwalsze ale najtańsze. Takie rozwiązania umożliwiają dokonały kontakt z przyrodą oraz otoczeniem. Dzięki swojej mobilności dom można dostosować pod indywidualne potrzeby, można mieszkać tam, gdzie aktualnie się chce (dzięki transporcie domku).

Domki te są też bardzo „zgrabne" o zwartej konstrukcji a przy tym niezwykle wygodne i funkcjonalne, tanie w utrzymaniu, ekologiczne oraz energooszczędne. Chociaż zagadnienie Tiny House nie jest jeszcze rozpowszechnione w Polsce, to coraz więcej osób interesuje się tym tematem ponieważ zależy im na jak najtańszym zakupie swojego „miejsca pobytu" i zaoszczędzeniu powierzchni.

Wszystkie przytoczone przykłady są idealnym dowodem na to, iż nawet najmniejsze powierzchnie, które są w dobry sposób zaprojektowane mogą stać się idealnym miejscem do codziennego życia. Tiny House to projekt, który na świecie zaskarbił sobie sympatię użytkowników, których grono stale rośnie. Inwestorzy zdają sobie sprawę z tego, że nasze życie nie powinno zamieniać się w egzystencję, która skupia się na kolejnej racie kredytu i przywiązania do miejsca, w którym nie chce się już mieszkać. Właśnie dlatego społeczeństwo tak docenia wolność, swobodę oraz brak zobowiązań, który niesie za sobą Tiny House.

Tabela 1. Zestawienie systematyki Tiny House

	Micro Compact Home	Nido House	BioDomek	Minim Micro Home
Powierzchnia	7 m^2	10 m^2	12 m^2	20 m^2
Ilość kondygnacji	1	2	1-2	1
Maksymalna liczba osób mogąca mieszkać w budynku	3	4	2 (w zależności od ilości segmentów)	4
Możliwość rozbudowy	Nie	Nie	Tak	Nie
Konstrukcja	Szkieletowa, wykończenie panelami aluminiowymi	Szkieletowa, wykończenie drewnem	Szkieletowa, wykończenie drewnem	Szkieletowa, wykończenie drewnem
Cena	38,000 €	10,500 €	18, 500 €	18, 500 €

Źródło: Opracowanie własne

Literatura

1. http://www.biodomek.pl/ (03.04.2017).
2. http://www.budujemydom.pl/kotly-i-podgrzewacze/254-podgrzewacze-pojemnosciowe-i-przeplywowe (03.04.2017).
3. http://czysteogrzewanie.pl/czym-ogrzewac/promienniki-podczerwieni-jak-dzialaja-i-ile-kosztuja/ (02.04.2017).
4. http://drewiks.pl/wybudowane-domy-szkieletowe/album/37-dom-szkieletowy-motycz-w-trakcie-budowy (03.04.2017).
5. http://www.instsani.webd.pl/ (02.04.2017).
6. http://www.karbon.com.pl/index.php/archiwum/157--ogrzewanie-ogrzewacz-przepywowy-czy-bojler (01.04.2017).
7. http://kominy-online.pl/nawietrzaki-scienne/581-nawietrzak-szpaletowy-z-grzalka.html (01.04.2017).
8. http://www.microcompacthome.com/ (01.04.2017).
9. http://minimhomes.com/ (01.04.2017).

10. http://robinfalck.com/#/nido/ (03.04.2017).
11. http://ogrzewanie.drewnozamiastbenzyny.pl/jak-ogrzac-przyczepe-kempingowa/ (01.04.2017).
12. http://www.planergia.pl/technologie/item/2252-czy-warto-i-kiedy-stosowa%C4%87-rekuperator (03.04.2017).
13. http://pl.pl.allconstructions.com/portal/categories/306/1/0/1/article/9413/budowa-domu-szkieletowego (02.04.2017).
14. http://www.rynekinstalacyjny.pl/.
15. http://sf.curbed.com/2015/2/27/9986526/tiny-house-villages-seek-tiny-plots-of-land-near-san-francisco (01.04.2017).
16. http://www.shareable.net/blog/11-tiny-house-villages-redefining-home (20.03.2017).
17. http://www.wentylacja.org.pl/pages-69.html (01.04.2017).
18. http://wroclaw.wyborcza.pl/wroclaw/51,95327,15566040.html?i=16 (01.04.2017).
19. Mitchell R., *Ideas for Building and Living Well in Less Than 400 Square Feet*, Paperback, 2014.
20. Brown A., *The Very Small Home: Japanese Ideas for Living Well in Limited Space*, Hardcover, 2012.
21. Janzen M., *Tiny House Floor Plans: Over 200 Interior Designs for Tiny Houses*, Paperback, 2012.
22. Kondo M., *The Life-Changing Magic of Tidying Up: The Japanese Art of Decluttering and Organizing*, Ebury Publishing, 2015.
23. McKeown G., *Essentialism: The Disciplined Pursuit of Less*, MT Biznes, 2014
24. Peterson J., *Shipping Container Homes*, Paperback, 2016.
25. Williams D., *The Big Tiny: A Built-It-Myself Memoir*, Blue rider press, 2014.

Tiny House – wielkie pomysły architektoniczne na wykorzystanie małej przestrzeni

Abstrakt
Tiny House to pojęcie stosunkowo nowe i rzadko używane w środowisku polskim. Jest to koncepcja wspierająca aktywne życie, brak trwałych zobowiązań, kredytów na dziesięć lat, wolność finansową- niższe koszty życia i codzienności w pełnym komforcie, połączonym harmonijnie ze zgrabnym minimalizmem. Termin ten można rozumieć jako mały dom, często zwany także „domem w ruchu". Jest to opis architektonicznego oraz społecznego ruchu grupy ludzi, która mieszka w małych budynkach. Nie zdefiniowano jeszcze dokładnie, co stanowi tę małą przestrzeń, wiadomo jednak, że taki dom ma poniżej 50 m^2 i powszechnie jest uważany za małą przestrzeń. Tiny house, przynajmniej jak na razie, nie doczekał się jeszcze Polskiej nazwy. Najłatwiej jednak opisać go jako "mini" dom. W malutkim domu mieści się wszystko, od sypialni po łazienkę, szafy, kuchnie i lodówkę.
Słowa kluczowe: Tiny House, mini-budynek, wielkie pomysły

Tiny House – great architectural ideas for the use of small space

Abstract
Tiny House is a relatively new concept and rarely used in the Polish community. It is a concept that promotes active living, lack of permanent commitments, loans for ten years, financial freedom – lower living and daily expenses in complete comfort, combined in harmony with neat minimalism. This term can be understood as a small house, often also called a "moving house". This is a description of the architectural and social movement of a group of people who live in small buildings. It is not yet defined exactly what constitutes this small space, but it is known that such a house is less than 50 m^2 and is generally considered to be a small space. Tiny house, at least so far, has not yet reached the Polish name. It is easiest to describe it as a "mini" home. In a tiny home everything from bedroom to bathroom, closets, kitchens and refrigerators.
Keywords: Tiny House, mini-building, great ideas

Tomasz Mazurkiewicz [1], Damian Ferenc[2]

Aspekt bezpieczeństwa we współczesnych urządzeniach elektronicznych

1. Wstęp

W ciągu ostatnich dwóch dekad wynalazki, które jeszcze niedawno wydawały się być fikcją literacką, stały się dostępne dla każdego. Ekspres, który sam zaparza ulubioną kawę, lodówka informująca o braku jedzenia, a nawet sama zamawiająca brakujące produkty czy samoparkujące samochody, które dodatkowo doradzą drogę pozwalającą uniknąć korków. Wszystkie te urządzenia, podłączone do sieci i realizujące za nas część obowiązków, stają się nieodłączną częścią codzienności. W 1999 roku Kevin Ashton [1] nazwał sieć takich urządzeń Internetem Rzeczy (ang. Internet of Things, IoT). Jednak dokładne zdefiniowanie tego pojęcia jest dość trudne. Według Smitha i in. [2] IoT jest określany jako globalna infrastruktura sieciowa łącząca ze sobą fizyczne i wirtualne obiekty dzięki możliwości komunikacji i eksploracji danych. Infrastruktura ta obejmuje istniejące i powstające sieci. Zapewnia ona identyfikację obiektów, jak i również wykorzystanie możliwości wynikających z podłączenia do niej sensorów i łączności między nimi. Tworzący się w ten sposób podstawy do samodzielnych, współpracujących ze sobą usług i aplikacji. Według Hallera i in. [3] pojęcie Internetu Rzeczy zostało przedstawione jako świat, gdzie fizyczne obiekty są podłączone do sieci informacyjnej i mogą aktywnie uczestniczyć w procesach biznesowych. Usługi mają natomiast możliwość interakcji z tymi „inteligentnymi" urządzeniami poprzez Internet, mogąc je odpytywać i zmieniać ich stan, biorąc pod uwagę aspekty bezpieczeństwa i prywatności. Definicja zaprezentowana w pracy Duce i in. [4] przedstawia IoT jako sieć stworzoną przez obiekty/rzeczy posiadające tożsamość i wirtualną osobowość. Działają one w inteligentnych przestrzeniach, które dzięki inteligentnym interfejsom łączą się i komunikują z kontekstem socjalnym, środowiskowym i użytkownika.

Analizując powyższe definicje można zauważyć, że istotnymi składowymi Internetu Rzeczy są „inteligentne" urządzenia wymieniające dane między sobą, bez potrzeby angażowania użytkownika. Mediami do wymiany danych mogą być zarówno istniejące, jak i nowo powstające sieci. Komunikacja między urządzeniami może przebiegać zarówno za pomocą protokołu TCP/IP, jak i dedykowanych protokołów w zależności od możliwości poszczególnych urządzeń i sensorów. Warto podkreślić, że urządzenia działające w ramach IoT mogą, na podstawie analizy zgromadzonych danych, same podejmować różnego rodzaju interakcje między sobą lub z użytkownikiem. Świadczą one w ten sposób różnego rodzaju usługi lub realizują przypisane im zadania.

Urządzenia IoT stosowane są nie tylko przez osoby prywatne, ale także w przemyśle, biznesie czy medycynie. Istotnym aspektem, który niestety jest często

[1] tomasz.mazurkiewicz@wat.edu.pl, Instytut Matematyki i Kryptologii, Wydział Cybernetyki, Wojskowa Akademia Techniczna, www.imk.wat.edu.pl
[2] Ministerstwo Obrony Narodowej, www.mon.gov.pl

pomijany przez producentów urządzeń IoT, jest bezpieczeństwo. W Internecie Rzeczy powinien funkcjonować system zabezpieczeń, który zapewni podstawowe funkcje bezpieczeństwa, tj. integralność, poufność, autentyczność i niezaprzeczalność. Bardzo ważne jest, by każdy węzeł sieci w IoT był uwierzytelniony, a dane były przesyłane kanałem szyfrowanym. Brak tych podstawowych mechanizmów może spowodować to, że „inteligentne" urządzenia staną się podatne na wiele ataków. Każde nieodpowiednio zabezpieczone urządzenie może stanowić bezpośrednie zagrożenie dla jego użytkowników, np. w sytuacji, gdy osoba niepowołana sprawi, że wszczepiony pacjentowi rozrusznik serca nagle przestanie pracować. Co więcej, urządzenia te mogą zostać połączone w sieć (ang. botnet) i wykorzystane do przeprowadzania ataków typu DDoS (ang. Distributed Denial of Service) na serwery lub strony internetowe, najczęściej bez wiedzy właściciela.

W dalszej części pracy opisano możliwości zastosowania urządzeń IoT w życiu codziennym. Przedstawiono również praktyczne przykłady wykorzystania braków w zabezpieczeniach urządzeń IoT do przeprowadzenia ataków. W rozdziale 4 przeanalizowano z kolei możliwość zabezpieczenia urządzeń za pomocą algorytmów i protokołów kryptograficznych, a także z wykorzystaniem specjalizowanych układów z wbudowanymi zabezpieczeniami kryptograficznymi.

2. Możliwości zastosowania Internetu Rzeczy

Liczba urządzeń IoT z każdym rokiem znacznie rośnie. Przewiduje się, że w 2020 roku na świecie będzie ich między 20 [5] a 50 [6] miliardów, podczas gdy w 2013 roku liczba urządzeń wynosiła ok. 3 miliardy. Internet Rzeczy cieszy się coraz większą popularnością również w Polsce, a perspektywy jego dalszego rozwoju w naszym kraju przedstawiono m.in. w pracy Grodnera i in. [7].

Jedną z istotnych przyczyn wzrostu popularności Internetu Rzeczy jest ciągły spadek kosztu zapewnienia określonej mocy obliczeniowej. Zgodnie z prawem Moore'a, koszt ten maleje o połowę co 18-24 miesiące [8]. W praktyce przekłada się to na spadek kosztu produkcji urządzeń charakteryzujących się niewielką mocą obliczeniową i ograniczoną pamięcią, a takimi zazwyczaj są urządzenia IoT.

Drugą przyczyną jest znaczna różnorodność urządzeń IoT. Dzięki temu liczba możliwości wykorzystania koncepcji Internetu Rzeczy jest bardzo duża. Wśród przykładowych obszarów zastosowania wymienić można [9]:

- miasta,
- środowisko,
- gospodarka wodna,
- transport,
- infrastruktura (domy, budynki),
- przemysł i produkcja,
- logistyka,
- żywność,
- energia,
- zdrowie.

Miasta są istotnym obszarem zastosowań urządzeń IoT. Wiele z procesów zachodzących w nich na co dzień może zostać zoptymalizowanych dzięki zastosowaniu Internetu Rzeczy. Jednym z najważniejszych aspektów jest poprawa bezpieczeństwa w miastach. Urządzenia IoT, m.in. różnego rodzaju czujniki, rozlokowane w strategicznych miejscach mogą posłużyć do kontroli przeciw-

pożarowej, automatycznego informowania odpowiednich służb w razie wystąpienia pożaru lub wypadku czy monitorowania miasta z algorytmami rozpoznawania zachowań. Internet Rzeczy może także służyć do zarządzania gospodarką odpadami przez kontrolę poziomu zapełnienia kontenerów czy ogólnego stanu czystości miasta. Kolejnym zastosowaniem jest optymalizacja organizacji ruchu pieszych i pojazdów, np. poprzez informowanie o utrudnieniach komunikacyjnych, a także wolnych miejscach parkingowych czy inteligentne sterowanie sygnalizacją świetlną.

Równie istotnym, choć często niedostrzeganym na co dzień, obszarem zastosowania jest środowisko. Urządzenia IoT są w stanie monitorować m.in. stan zanieczyszczenia powietrza i wody, a także poziom rzek, występowanie wstrząsów sejsmicznych i wilgotności gleby w okresie susz. Na podstawie zebranych danych możliwe jest określenie poziomu zagrożenia klęską żywiołową, a w przypadku wystąpienia zagrożenia – natychmiastowe informowanie odpowiednich służb oraz ludzi pozostających na zagrożonym terenie.

W obszarze gospodarki wodnej Internet Rzeczy stosowany może być do kontroli stanu wód czy monitorowania stanów dróg wodnych. Urządzenia IoT mogą też wspomóc zarządzanie procesem dostaw wody – rozpoczynając od inspekcji zdatności wody do spożycia, przez monitorowanie sposobu przechowywania wody i szczelności wodociągów, a kończąc na analizie poziomu zużycia wody przez poszczególnych użytkowników końcowych.

Kolejnym obszarem zastosowań jest transport. Oprócz wspomnianej już optymalizacji ruchu pieszych i pojazdów w miastach, urządzenia IoT mogą być wykorzystane do kontroli położenia pojazdów, ich prędkości, panujących warunków atmosferycznych, a nawet stanu technicznego. W przypadku wypadku lub zdiagnozowania awarii możliwe jest automatyczne wezwanie pomocy drogowej czy też rezerwacja wizyty w pobliskim warsztacie samochodowym. Wymienione dane mogą być zastosowane również do zwiększenia efektywności zarządzania flotą samochodów firmowych i bezpieczeństwa transportu materiałów niebezpiecznych. Za pomocą urządzeń IoT możliwe jest uproszczenie pobierania opłat drogowych, np. podczas wjazdu na autostradę.

Internet Rzeczy coraz częściej znajduje zastosowanie również w naszych domach i mieszkaniach. Co więcej, przytaczany już raport [7] wskazuje, że to właśnie ten obszar zastosowania ma największy potencjał rozwoju w naszym kraju. Urządzenia IoT mogą być wykorzystywane w szeroko rozumianej automatyce budynkowej, zarówno do zapewniania dodatkowych funkcjonalności ułatwiających życie, jak i zwiększenia bezpieczeństwa mieszkańców. Jako przykład można tu przedstawić rolety okienne automatycznie zamykane w zależności od godziny i poziomu nasłonecznienia, termostat inteligentnie dostosowujący się do optymalnej dla użytkowników temperatury w pomieszczeniach czy też instalacje nawadniające trawnik w przypadku zbyt niskiego poziomu wilgotności. Co więcej, inteligentne sensory połączone w ramach sieci mogą wykryć zalanie domu, obecność dymu, a także włamanie. Informacja o wszelkich nieprawidłowościach może zostać przesłana do domowników oraz odpowiednich służb, skracając czas ich reakcji. Dodatkowo mogą zostać podjęte działania mające na celu przeciwdziałanie zagrożeniu, np. uruchomienie alarmu czy też systemu przeciwpożarowego lub odcięcie dopływu wody.

Urządzenia IoT są stosowane coraz częściej również w przemyśle i produkcji. Jednym z ich zastosowań jest ciągły monitoring i kontrola procesu produkcyjnego. Internet Rzeczy pozwala optymalizować wykorzystanie hal magazynowych

i monitorować stany zasobów, np. zbiorniki z wodą, paliwem lub gazem. Istotnym zastosowaniem jest również diagnostyka maszyn i systemy wykrywania i raportowania awarii. W przypadku produkcji rolnej urządzenia IoT mogą być wykorzystane do kontroli wilgotności, temperatury i nasłonecznienia w celu przeciwdziałania suszy, czy też powstawaniu różnego rodzaju groźnych dla upraw grzybów, jak również do monitorowania warunków życia i stanu prowadzonej hodowli zwierząt.

Internet Rzeczy można wykorzystać także w obszarze zarządzania energią. Urządzenia IoT mogą zostać zastosowane do monitorowania zużycia energii zarówno w zakładach przemysłowych, jak i w gospodarstwach domowych pozwalając optymalizować jej zużycie, a co za tym idzie ponoszone koszty. Innym zastosowaniem jest kontrola procesów wytwarzania energii, szczególnie w elektrowniach korzystających z odnawialnych źródeł energii. Sieć urządzeń może optymalizować konfigurację poszczególnych komponentów (np. położenie paneli słonecznych) na podstawie zebranych danych na temat warunków atmosferycznych i maksymalizować tym samym wydajność elektrowni.

Internet Rzeczy znajduje zastosowanie również w służbie zdrowia, gdzie może być wykorzystany m.in. do monitorowania stanu pacjentów zarówno w szpitalach, jak i domach. Urządzenia IoT mogą na bieżąco analizować stan pacjenta po zabiegu lub poddanego długotrwałemu leczeniu i informować lekarza w przypadku wykrycia niepokojących symptomów. Lekarz otrzymuje dostęp do zbieranych na bieżąco danych takich, jak np. tętno i ciśnienie pacjenta. Pozwala to dokładniej diagnozować pacjenta i dopasowywać kurację. W trakcie terapii farmakologicznej takie rozwiązanie pozwala na bieżąco monitorować czy zażywane przez pacjentów leki przynoszą zakładane rezultaty i nie wywołują niepożądanych skutków ubocznych. Urządzenia IoT mogą zostać również zastosowane do kontroli jakości i czasu snu.

3. Analiza zagrożeń w Internecie Rzeczy

3.1. Wprowadzenie

Szerokie możliwości zastosowania Internetu Rzeczy w opisanych we wcześniejszym rozdziale dziedzinach życia codziennego mogą nieść za sobą poważne zagrożenia, jeżeli podczas budowy urządzeń producenci pominą bardzo ważny aspekt, jakim jest bezpieczeństwo. Ze względu na dynamikę rozwoju urządzeń IoT i innowacyjność wprowadzanych rozwiązań technologicznych, wielu problemów związanych z ich wykorzystaniem wciąż nie jesteśmy świadomi. Niewiedza użytkowników jest często kolejnym czynnikiem negatywnie wpływającym na bezpieczeństwo Internetu Rzeczy.

Urządzenia IoT mogą ułatwiać życie w wielu jego dziedzinach, jednak w niektórych sytuacjach mogą okazać się poważnym zagrożeniem. Należy uświadomić sobie, że bardzo często urządzenia te przetwarzają i przesyłają istotne dane o swoich właścicielach lub też na podstawie tych danych podejmują różnego typu akcje. Nieodpowiednie zabezpieczenie urządzeń i komunikacji między nimi może spowodować wyciek danych albo umożliwić dostęp do urządzenia przez osoby nieuprawnione. Może to doprowadzić do powstania szkody materialnej, zniszczenia mienia, a nawet zagrożenia życia użytkownika.

Alarmujące dane zamieszczone zostały w raporcie PWC z 2011 roku [10]. Zgodnie z przedstawionymi informacjami, aż 70% podłączonych do sieci urządzeń IoT nie spełnia fundamentalnych wymogów bezpieczeństwa. Uwarunkowane jest to głównie

niewielką mocą obliczeniową i ograniczeniami związanymi z wykorzystaniem zasilania bateryjnego. Często zdarza się również, że nawet urządzenia wyposażone w mocne procesory i zasilanie sieciowe (np. inteligentne telewizory) są słabo zabezpieczane. Jest to zazwyczaj wynikiem braku odpowiedniej świadomości producentów tego typu urządzeń i użytkowników końcowych.

Urządzenia IoT i wszystkie interfejsy komunikacyjne między nimi powinny być zabezpieczane z uwzględnieniem podstawowych funkcji bezpieczeństwa, takich jak poufność, integralność, autentyczność, niezaprzeczalność i rozliczalność [11]. Pominięcie którejkolwiek z nich może mieć fatalne skutki dla bezpieczeństwa całej „inteligentnej" sieci, czego przykłady można było wielokrotnie zaobserwować [12].

3.2. Wybrane znane ataki na urządzenia IoT

Za jeden z wielu przykładów może posłużyć bezprzewodowa klawiatura komputerowa. W 2015 roku przedstawiono [13] urządzenie przechwytujące dane przesyłane między klawiaturą bezprzewodową a komputerem, które można wykonać samodzielnie. Jego koszt to zaledwie 10-80$. Urządzenie to jest sieciową ładowarką USB, która śledzi wszystko, co zostanie napisane przez użytkownika na klawiaturze bezprzewodowej. Ładowarka ta może działać na kilka sposobów, np. gromadzić dane w wbudowanej pamięci lub wysyłać wszystkie przechwycone dane natychmiast na telefon atakującego dzięki modułowi GSM.

Rysunek 1 przedstawia zdjęcie urządzenia. Urządzenie jest niewielkie i ze względu na wygląd zbliżony do standardowej ładowarki od telefonu, nie przyciąga ono uwagi. Atak z pomocą takiej ładowarki, nazwanej przez twórcę Keysweeper, jest bardzo prosty. Wystarczy wpiąć ją do gniazdka w publicznym miejscu, takim jak biblioteka, dworzec, lotnisko czy biuro. Od tego momentu można zbierać informację ze wszystkich bezprzewodowych klawiatur w pobliżu. Nasuwa się jednak pytanie, co dalej z przechwyconymi danymi, które są zaszyfrowane. Niestety, w wielu klawiaturach bezprzewodowych szyfrowanie odbywa się za pomocą prostej operacji XOR na danych i cyklicznie powtarzanym stałym kluczu, co daje jedynie pozorne bezpieczeństwo. Skuteczną kryptoanalizę tak zaszyfrowanych danych można przeprowadzić z wykorzystaniem prostej analizy częstotliwościowej.

Rysunek 1. Urządzenie Keysweeper do przechwytywania danych z klawiatury bezprzewodowej [13]

Takie urządzenie można również w łatwy sposób wykorzystać do zbierania danych z innych urządzeń IoT, które komunikują się bezprzewodowo i bardzo często nie używają żadnego szyfrowania albo wykorzystują bardzo słabe mechanizmy do zapewnienia poufności transmisji. Co więcej, funkcjonalność urządzenia może zostać

rozszerzona. Zamiast tylko nasłuchiwać, atakujący może zacząć również przechwytywać transmisję, następnie w odpowiedni sposób według potrzeby ją modyfikować i przesyłać do docelowego odbiorcy (atak typu man-in-the-middle).

Bardzo często do przeprowadzenia skutecznego ataku nie potrzeba wykorzystywać specjalistycznego lub dedykowanego sprzętu. Biorąc pod uwagę, że spora część urządzeń IoT wykorzystuje do komunikacji sieci WiFi lub technologię Bluetooth, aby je zaatakować wystarczy zwykły laptop wyposażony w bezprzewodową kartę sieciową i moduł Bluetooth czy też zwykły smartphone. Przykładem tego może być opisany w pracy [14] atak na inteligentną opaskę firmy Nike. Do przeprowadzenia tego ataku wystarczy telefon z systemem Android i odpowiednio przygotowana aplikacja.

Komunikacja ze wspomnianą inteligentną opaską, jak i wieloma podobnymi urządzeniami IoT, np. zegarkami, odbywa się za pomocą modułu komunikacyjnego działającego w standardzie BLE (ang. *Bluetooth Low Energy*). W celu komunikacji z opaską konieczne jest uwierzytelnienie się użytkownika. Proces uwierzytelnienia przebiega następująco:

- użytkownik otrzymuje swój unikalny pin,
- na podstawie kodu pin generowany jest token – 6 pierwszych bajtów z obliczonego za pomocą funkcji MD5 skrótu kodu pin użytkownika,
- aplikacja w celu komunikacji z opaską wysyła do niej odpowiednią komendę,
- opaska w odpowiedzi wysyła do aplikacji tymczasowy token sesji,
- w celu uwierzytelnienia aplikacja wysyła do opaski dwa pierwsze bajty z wyliczonej funkcji CRC32 tymczasowego tokenu sesji i swojego tokenu.

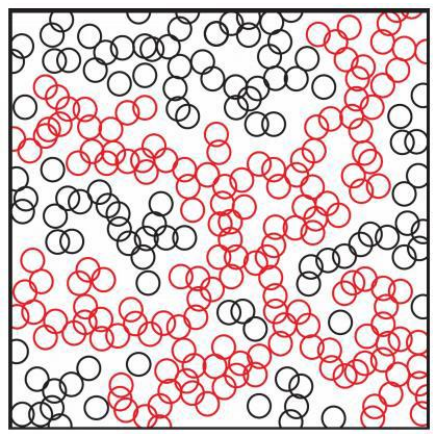
Teoretycznie opaska zabezpieczona jest w sposób prawidłowy. Atakujący musi uzyskać fizyczny dostęp do telefonu właściciela opaski z zainstalowaną aplikacją lub zdobyć jego kod pin, aby być w stanie przejąć kontrolę nad urządzeniem. Niestety, w rzeczywistości przebieg procesu uwierzytelniania użytkownika różni się od tego, który został zaimplementowany w opasce. Przede wszystkim, opaska rozgłasza token użytkownika podczas przedstawiania się. Co gorsza, token ten nie jest wykorzystywany w procesie uwierzytelniania. Został on zastąpiony stałą, ustaloną wartością. Oznacza to, że każdy posiadacz telefonu z system Android, który posiada moduł komunikacyjny zgodny ze standardem BLE jest w stanie poprawnie uwierzytelnić się i komunikować się z opaską.

Analizując opisany powyżej atak należy podkreślić kilka rzeczy. Atak ten był możliwy dzięki pominięciu przez deweloperów kwestii bezpieczeństwa w opracowywanym rozwiązaniu, mimo teoretycznego istnienia bezpiecznego procesu uwierzytelniania w implementowanym standardzie. Co więcej, przygotowana implementacja w żaden sposób nie została zabezpieczona przed inżynierią wsteczną. Niewykonanie obfuskacji (zaciemnienia) kodu ułatwiło odkrycie atakującemu wszystkie podatności przy niewielkim nakładzie pracy. Dodatkowo pozwoliło uzyskać dostęp do gotowych klas i metod stanowiących idealną bazę, którą po pewnych modyfikacjach można wykorzystać do stworzenia aplikacji potrzebnej do przeprowadzenia ataku.

Ciekawy atak opisano w pracy autorstwa Ronena i in. [15]. Autorzy skupili się w nim na inteligentnych żarówkach. O ile powyżej opisane przykłady ataków można potraktować jako ciekawostkę, o tyle ten atak może już poważnie zagrozić zdrowiu i życiu użytkownika. Pozornie przejęcie kontroli nad takim urządzeniem może doprowadzić do irytacji użytkownika, któremu nagle zdalnie ktoś włącza i wyłącza oświetlenie. Jednak po przejęciu żarówki może zostać ona wykorzystana także m.in.

do wywołania ataku epilepsji u użytkownika [16] czy zwiększenia odczuwanego dyskomfortu [17].

Autorzy przeprowadzili atak poprzez modyfikację oprogramowania w jednej z żarówek. Modyfikacja ta nie wymaga bezpośredniego dostępu do urządzenia i może być przeprowadzona z odległości kilkuset metrów. Co więcej, zaatakowane urządzenie będzie rozpropagowywać zainfekowane oprogramowanie do wszystkich takich urządzeń znajdujących się w zasięgu sieci, w której działa. Oznacza to, że zainfekowanie jednego urządzenia wywoła skutki na szeroką skalę. Autorzy określili, że wystarczy 15000 urządzeń w losowych miejscach w mieście wielkości Paryża, aby atak dosięgnął wszystkie takie urządzenia w mieście. Rysunek 2 przedstawia zobrazowanie propagacji zainfekowanego oprogramowania w ramach sieci. Czerwone okręgi reprezentują zasięg pojedynczych urządzeń, które zostały przejęte ze względu na połączenie w ramach jednej sieci z zainfekowanym pierwotnie urządzeniem. Pozostałe urządzenia (kolor czarny) nie zostały przejęte.



Rysunek 2. Propagacja zainfekowanego oprogramowania między sąsiednimi urządzeniami poprzez sieć ZLL – okręgi czerwone symbolizują urządzenia przejęte przez atakującego [15]

Do przeprowadzenia ataku autorzy wykorzystali błąd w implementacji protokołu ZigBee Light Link, który powinien zapobiegać komunikacji z urządzeniem z odległości większej niż ok. 1 m. Mimo że podczas pisania kodu jego twórcy brali pod uwagę aspekt bezpieczeństwa okazało się, że ich implementacja posiada podatność. Autorzy ataku odkryli, że możliwe jest wysłanie do urządzenia komendy (np. reset) z ID transmisji i odpowiedzi równym zero, a komenda taka zostanie przetworzona. Było to spowodowane tym, że test ID odbywa się jedynie w przypadku komendy Scan Request. Co więcej, autorzy przeprowadzili atak side-channel na posiadane urządzenie i byli w stanie odzyskać klucz wykorzystywany przez producenta do szyfrowania i uwierzytelniania nowej wersji oprogramowania. Dzięki temu autorzy mogli zdalnie (z odległości ok. 400m) zresetować i odłączyć pojedyncze urządzenie od istniejącej sieci i wymusić aktualizację oprogramowania. W ramach aktualizacji wgrywane było zainfekowane oprogramowanie.

Kolejny atak mogący poważnie zagrozić zdrowiu i życiu użytkownika przedstawiono w [18]. Jego autorzy pokazali, że możliwe jest przejęcie zdalnej kontroli nad „inteligentnym" samochodem, a dokładnie nad Jeepem Cherokee z 2014 roku. Pierwszym podejściem do ataku było wykorzystanie sieci Wi-Fi, którą samochód

udostępnia dla pasażerów. By połączyć się ze wspomnianą siecią należało jednak znać hasło. Okazuje się jednak, że hasłem do sieci dla każdego samochodu była data jego pierwszego uruchomienia z dokładnością do sekund. Co za tym idzie, znając rok i miesiąc produkcji samochodu hasło można złamać atakiem brutalnym w około godzinę. Jest to możliwe, jednak dość problematyczne, gdyż wymaga poruszania się przez godzinę za samochodem, który chcemy atakować. Okazuje się, że data pierwszego uruchomienia to tak naprawdę domyślna data systemu plus kilka sekund potrzebnych na uruchomienie jednostki sterującej samochodu, co upraszcza atak. Po uzyskaniu dostępu do sieci Wi-Fi, atakujący przejęli kontrolę nad systemem multimedialnym działającym na systemie Linux, a następnie nad całym systemem jednostki centralnej. W ten sposób byli w stanie w pełni kontrolować system audio (zmieniać stację radiową, kontrolować głośność) oraz śledzić położenie samochodu dzięki wbudowanemu modułowi GPS. Jak zostało już wspomniane wcześniej, by wykonać atak w ten sposób należy cały czas być w pobliżu atakowanego samochodu. Co więcej, usługa Wi-Fi dla pasażerów jest płatna, więc nie wszyscy użytkownicy mają ją aktywowaną.

Twórcy zastosowali więc inne podejście i przeprowadzili atak przez sieć komórkową, do której atakowany model był podłączony domyślnie. Z wykorzystaniem kompaktowej komórkowej stacji bazowej, którą można zakupić na portalach aukcyjnych autorzy byli w stanie skanować adresy IP w ramach sieci komórkowej. Na podstawie wcześniej przeprowadzanych ataków na samochodową sieć Wi-Fi potrafili określić, jakie połączenia należy wyszukiwać. To pozwoliło odfiltrować wszystkie samochody grupy Chrysler wyposażone w jednostki sterujące takie, jak w atakowanym Jeepie. Okazało się jednak, że łatwiej jest zaatakować wszystkie odnalezione samochody lub jeden losowy niż konkretny samochód. Jednak nie jest to niemożliwe. Znając jego położenie możemy odnaleźć go wykorzystując opisaną wcześniej możliwość raportowania lokalizacji. Po odnalezieniu wybranego samochód, można sterować systemem multimedialnym, tak jak zostało opisane to wcześniej, oraz śledzić samochód. Autorzy znaleźli również sposób na uzyskanie dostępu do magistrali CAN poprzez system multimedialny, mimo że nie istnieje bezpośrednie połączenie między nimi. W tym celu wykorzystany został kontroler V850, który może komunikować się zarówno z magistralą CAN, jak i systemem multimedialnym. Kontroler ten standardowo może jedynie nasłuchiwać komunikaty przepływające przez magistralę CAN, jednak autorom udało się zmodyfikować kod wykonywalny kontrolera. Jego zamiana nie wymagała żadnego uwierzytelnienia. Po aktualizacji kontrolera i uzyskaniu dostępu do magistrali CAN atakujący byli w stanie przejąć kontrolę nad większością układów, w tym nad kierownicą, silnikiem, skrzynią biegów, hamulcami, zamkami w drzwiach, klimatyzacją i wycieraczkami. Cały atak przeprowadzany jest zdalnie. Jedynym wymogiem jest konieczność połączenia się do tej samej sieci komórkowej, co atakowany samochód.

Wiele ataków pojawia się również na urządzenia, których celem miała być ochrona zdrowia i życia pacjenta, np. pompy insulinowe [19] oraz urządzenia kardiologiczne [18, 20].

W pracy [19] autor, który sam jest diabetykiem, przedstawił atak na wykorzystywaną przez niego pompę insulinową. Zaprezentowany atak może spowodować, że urządzenie wyłączy się albo dostarczy użytkownikowi śmiertelną dawkę insuliny. Co gorsza, do jego przeprowadzenia wymagany jest jedynie numer seryjny urządzenia oraz sprzęt wart niecałe 100$. Do przeprowadzania skutecznej analizy wstecznej

oprogramowania wystarczyła zmiana poziomu raportowania logów z *NONE* na *HIGH*. Pozwoliło to autorowi na rozpoznanie protokołu bezprzewodowej komunikacji z pompą, gdyż dostarczane przez producenta oprogramowanie nie było w żaden sposób zabezpieczone.

Przykład ataku na urządzenia kardiologiczne przedstawiono w [20]. Autorzy wykonali inżynierię wsteczną protokołu komunikacyjnego między programatorem, a urządzeniem kardiologicznym. Urządzenie to wykorzystuje komunikację radiową o zasięgu między 2 a 5 metrów. Analiza autorów wykazała kilka podatności w stosowanym protokole i jego implementacji. Co więcej, zademonstrowali oni kilka innych ataków (m.in. DDoS), na które podatne jest min. 10 typów urządzeń kardiologicznych. Wszystkie te ataki mogą zostać przeprowadzone bez konieczności znajdowania się w pobliżu atakowanego użytkownika.

W pracy [21] przedstawiono z kolei ponad 8000 podatności wykrytych w urządzeniach medycznych takich, jak np. sztuczne rozruszniki serca, wszczepialne kardiowertery-defibrylatory serca czy urządzenia monitorujące. Wiele wykrytych podatności to luki w stosowanym przez producentów urządzeń medycznych oprogramowaniu dostarczanym przez firmy trzecie.

Aspekt bezpieczeństwa w urządzeniach medycznych został przedstawiony w raporcie [22]. Autorzy przeprowadzili ankietę wśród osób biorących udział w tworzeniu urządzeń medycznych oraz pracowników firm dostarczających usługi medyczne. Spośród pierwszej grupy ankietowanych zaledwie jedna trzecia przyznała, że transfer danych pomiędzy urządzeniami IoT ich firmy jest szyfrowany. Z kolei aż 40% osób przyznało, że nie są podejmowane żadne działania w celu zapobiegnięcia atakom na produkowane urządzenia medyczne. Co gorsza, urządzenia wypuszczane na rynek bardzo często nie są testowane pod kątem ewentualnych podatności.

Przytoczony raport potwierdza, że aspekt bezpieczeństwa jest często marginalizowany przez producentów urządzeń, nawet w tak ważnym obszarze zastosowania Internetu Rzeczy jak zdrowie.

3.3. Wybrane znane ataki z wykorzystaniem urządzeń IoT

Przedstawione do tej pory przykłady ataków na urządzenia IoT pokazują, jak dużą rolę powinien odgrywać aspekt bezpieczeństwa na etapie projektowania i testowania urządzeń przez producentów. Wiele zagrożeń związanych z Internetem Rzeczy związanych jest jednak z ignorowaniem tego aspektu przez użytkowników końcowych. Po pierwsze, kryterium ceny jest niestety bardzo często kluczowym podczas wyboru kupowanego urządzenia IoT. Wiąże się to zazwyczaj z zakupem urządzenia, którego poziomu zabezpieczeń nie jest wysoki. Po drugie, użytkownicy często nie wykorzystują lub wykorzystują niepoprawnie zaimplementowane w urządzeniach mechanizmy bezpieczeństwa.

Idealny przykład wykorzystania podatności jaką są sami użytkownicy urządzeń IoT to powstające botnety. Nazywamy tak grupy zainfekowanych urządzeń połączonych w sieć, które mogą być zdalnie kontrolowane przez atakującego. Użytkownik najczęściej nie ma pojęcia, że jego urządzenie jest częścią takiej struktury. Najbardziej znanym botnetem wykorzystującym urządzenia IoT takie, jak kamery CCTV, jest botnet Mirai. Wykorzystany on był w atakach DDoS na rekordowo dużą skalę, m.in. na stronę internetową dziennikarza zajmującego się tematyką bezpieczeństwa, Briana Krebsa.

Kod źródłowy botnetu Mirai został udostępniony w Internecie na zasadzie otwartego oprogramowania (ang. open source). W oparciu o udostępniony kod

powstały kolejne sieci zainfekowanych urządzeń. Przykładem takiego botnetu jest Bashlight, który zainfekował już przeszło milion urządzeń.

Analiza kodu wykazała, że Mirai składa się z następujących komponentów [23]:

- centrum kontroli, które zawiera bazę zainfekowanych urządzeń i kontroluje pośrednie serwery dystrybucji poleceń,
- komponentu gromadzącego wyniki działania każdego zainfekowanego urządzenia (bota) i dystrybuującego zainfekowane oprogramowanie do urządzeń z odkrytymi podatnościami,
- komponentu dostarczającego zainfekowane oprogramowanie na podatne urządzenie,
- narzędzia, które po uruchomieniu na zainfekowanym urządzeniu skanuje określony zakres adresów IP w celu wykrycia kolejnych urządzeń podatnych na ataki.

Po znalezieniu nowych podatnych urządzeń, informacja o wynikach skanowania przesyłana jest z zainfekowanego urządzenia do komponentu gromadzącego wyniki. Na podstawie tych informacji tzw. dystrybutor pobiera zainfekowane oprogramowanie na znalezione urządzenia, które zostają tym samym włączone do botnetu.

Botnet Mirai podczas skanowania wykorzystuje słownik par login-hasło. W tabeli 1 przedstawiono listę przykładowych 60 kombinacji wykorzystywanych przez botnet. Większość z przedstawionych haseł to domyślne hasła do konta administratora urządzeń. Niestety, użytkownicy bardzo często nie zmieniają haseł, ani nie przeprowadzają procesu aktualizacji oprogramowania po zakupie urządzenia. W ten sposób atakujący był w stanie w łatwy sposób zainfekować ponad pół miliona urządzeń.

Na podobnej zasadzie co botnet Mirai działa również BrockerBot. Podstawową różnicą jest cel infekowania urządzeń. Tak jak opisano wcześniej, urządzenia przejmowane przez botnet Mirai dołączane są do sieci i wykorzystywane do przeprowadzania ataków DDoS. Z kolei BrockerBot [25] po przejęciu urządzenia dąży do jego zniszczenia. Taki rodzaj ataku określany jest mianem PDoS (ang. Permanent Denial of Service).

4. Możliwości zwiększenia bezpieczeństwa Internetu Rzeczy

4.1. Wprowadzenie

Bezpieczeństwo Internetu Rzeczy może być zwiększone na różne sposoby. W projektowanych rozwiązaniach możliwe jest zaimplementowanie uważanych za bezpieczne algorytmów i protokołów kryptograficznych w oparciu o mikrokontroler lub układy FPGA (ang. Field-Programmable Gate Array). Można również z korzystać z dostępnych na rynku układów, które posiadają zaimplementowane mechanizmy bezpieczeństwa. Układy takie coraz liczniej pojawiają się na rynku, a najwięksi producenci układów scalonych starają się wykorzystać niszę rynkową.

Należy podkreślić, że już poprawna implementacja jednego algorytmu kryptograficznego może znacznie podnieść bezpieczeństwo urządzenia IoT, znacznie ograniczając podatność na atak. Kluczowe wydaje się zwłaszcza zabezpieczenie komunikacji z urządzeniem. Opracowane rozwiązania powinny zostać zabezpieczone, m.in. poprzez zaciemnianie kodu, również przed inżynierią wsteczną. Wiele ataków okazało się możliwych dzięki wykorzystaniu przekonania producentów, że kodu źródłowego w żaden sposób nie uda się odzyskać.

Tabela 1. Lista danych logowania wykorzystywanych przez botnet Mirai

Login	Hasło	Login	Hasło	Login	Hasło
666666	666666	guest	guest	root	jvbzd
888888	888888	mother	fucker	root	klv123
admin	(none)	root	(none)	root	klv1234
admin	1111	root	00000000	root	pass
admin	1111111	root	1111	root	password
admin	1234	root	1234	root	realtek
admin	12345	root	12345	root	root
admin	123456	root	123456	root	system
admin	54321	root	54321	root	user
admin	7ujMko0admin	root	666666	root	vizxv
admin	admin	root	7ujMko0admin	root	xc3511
admin	admin1234	root	7ujMko0vizxv	root	xmhdipc
admin	meinsm	root	888888	root	zlxx.
admin	pass	root	admin	root	Zte521
admin	password	root	anko	service	service
admin	smcadmin	root	default	supervisor	supervisor
admin1	password	root	dreambox	support	support
Administra-tor	1234	root	hi3518	tech	tech
Admini-strator	admin	root	ikwb	ubnt	ubnt
guest	12345	root	juantech	user	user

Źródło: [24]

4.2. Układy dostępne na rynku komercyjnym

Przykładem układu dedykowanego do stosowania w urządzeniach IoT z zaimplementowanymi mechanizmami bezpieczeństwa mogą być układy STSAFE-TPM firmy STMicroeletronics. Układy te przeszły proces certyfikacji Common Criteria EAL na poziomie 4+. Dostarczane z oprogramowaniem Axio-OS i Axio-RA firmy Security Platform mają stworzyć zintegrowane środowisko dla projektantów urządzeń IoT [26]. W rozwiązaniu tym zaimplementowano takie mechanizmy zwiększające bezpieczeństwo jak:

- uwierzytelnianie użytkownika oraz urządzenia,
- bezpieczne przechowywanie danych kluczowych,
- zabezpieczenie przed ingerencją w układ,
- zabezpieczenie przed podglądem danych w układzie,
- możliwość sprawdzenia naruszenia integralności dzięki informacjom na osobnym serwerze,
- ochrona przed klonowaniem i fałszerstwem urządzenia,
- bezpieczne aktualizacje.

Innym przykładem układu tej firmy jest również gotowy moduł Wi-Fi SPWF04 [27]. Rozwiązanie to ma na celu przyśpieszyć rozwój różnego rodzaju urządzeń IoT

i komunikacji M2M (ang. Machine-to-Machine). Komunikacja zabezpieczona jest za pomocą następujących funkcjonalności:
- obsługa protokołu WPA2 – Enterprise,
- obsługa protokołu TLS,
- możliwość bezpiecznej bezprzewodowej aktualizacji.

Firma Renesans posiada w swojej ofercie mikrokontrolery z rodziny Synergy, które dedykowane są do rozwiązań Internetu Rzeczy i wyposażone w gotowe sprzętowe mechanizmy bezpieczeństwa. Przykładowo mikrokontroler Synergy S3 [28] został wyposażony w następujące mechanizmy:
- sprzętową implementację algorytmu szyfrowania AES (klucz długości 128 albo 256 bitów),
- sprzętową implementację funkcji skrótu,
- sprzętowy generator liczb losowych.

Układami zwiększającymi bezpieczeństwo urządzeń IoT i komunikacji między nimi są również dostarczane przez firmę Infineon moduły z rodziny OPTIGA Trust P [29]. Układy te przeszły proces certyfikacji Common Criteria EAL na poziomie 5. Mogą być one wykorzystane do zabezpieczania transmisji oraz wspomagają realizację usługi uwierzytelniania. Wyposażone są one w następujące komponenty sprzętowe:
- algorytm szyfrowania AES z kluczem do długości 256b,
- algorytmy asymetryczne oparte o krzywe eliptyczne, z kluczem do długości 521b,
- algorytm RSA z kluczem o długości do 2048b.

Wart uwagi jest również układ ATECC508A [30] firmy Microchip. Rozwiązanie to jest dedykowane do zapewnienia bezpiecznego połączenia typu end-to-end między urządzeniem IoT, a chmurą firmy Amazon. Układ ten wyposażony jest w sprzętowy moduł PUF (ang. Physical Uncloneable Function) służący do generacji unikalnych kluczy, których nie można odczytać z zewnątrz. Ponadto układ ten posiada następujące mechanizmy zwiększające bezpieczeństwo:
- wewnętrzne szyfrowanie całej pamięci,
- sprzętowy generator liczb losowych,
- sprzętową implementację SHA-256,
- obsługę protokołów ECDH i ECDSA,
- sprzętową implementację algorytmu AES.

Układ ten został również uodporniony na ataki typu side-channel m.in. poprzez zapewnienie fizycznej ochrony komórek pamięci oraz komórek logicznych czy zastosowanie izolowanej szyny zasilania.

4.3. Własna implementacja algorytmów kryptograficznych

Możliwości zwiększenia bezpieczeństwa urządzeń IoT z wykorzystaniem autorskich sprzętowych implementacji algorytmów kryptograficznych dla układów FPGA przeanalizowano w [31]. Autorzy skupili się na analizie znanych algorytmów lightweight, które mogłyby zostać wykorzystane do implementacji koprocesora kryptograficznego. Algorytmy te, potocznie zwane „lekkimi", idealnie nadają się do zastosowania w urządzeniach charakteryzujących się niewielką mocą obliczeniową, ograniczoną pojemnością pamięci oraz wykorzystujących zazwyczaj zasilanie bateryjne. Ze względu na odpowiedni dobór parametrów funkcjonalnych algorytmów (m.in. krótsza długość stosowanego klucza) czy projektowanie rozwiązań z założenia

zorientowanych na działanie w takich urządzeniach, możliwe jest osiągnięcie kompromisu między gwarantowanym poziomem bezpieczeństwa, a wydajnością i zajętością koprocesora kryptograficznego. Pod tym względem przeanalizowano cztery najważniejsze rodzaje algorytmów kryptograficznych:

- blokowe algorytmy symetryczne,
- strumieniowe algorytmy symetryczne,
- algorytmy asymetryczne,
- funkcje skrótu.

W wyniku przeprowadzonej analizy autorzy wskazali grupę algorytmów, które mogą zostać wykorzystane w urządzeniach IoT. Wiele blokowych algorytmów symetrycznych może zostać wykorzystane do zabezpieczenia Internetu Rzeczy, a najważniejszymi z nich jest algorytm PRESENT. Został on zaprojektowany w taki sposób, aby jego sprzętowa implementacja wykorzystywała jak najmniej zasobów logicznych. W 2012 roku został ogłoszony standardem w normie ISO/IEC 29192-2. Innym algorytmem, o którym warto wspomnieć jest CLEFIA. Algorytm ten gwarantuje większe bezpieczeństwo, kosztem nawet 3-krotnie większego wykorzystania zasobów.

W przypadku pozostałych typów algorytmów kryptograficznych pojawia się większy problem. Algorytm AES działający w trybie CTR okazuje się jedynym algorytmem strumieniowym, który uznawany jest za bezpieczny, a jego implementacja nie wymaga dużych zasobów logicznych.

Zgodnie z obecnym stanem wiedzy, wciąż brakuje efektywnych rozwiązań asymetrycznych dla urządzeń IoT. Nawet zastosowanie koprocesorów krzywych eliptycznych, wykorzystujących znacznie krótsze klucze niż algorytm RSA czy Diffiego-Hellmana, wiąże się z dużym wykorzystaniem zasobów logicznych. Z tego powodu rozwiązania te uznawane są niepraktyczne dla tego typu urządzeń. Ze względu na ograniczenia istniejących algorytmów, coraz większą popularność zyskują alternatywne koncepcje realizacji infrastruktury klucza publicznego. Największą uwagę przyciągają rozwiązania oparte o teorię krat, m.in. algorytm NTRU. Algorytm ten gwarantuje znacznie większą wydajność i mniejsze zużycie zasobów logicznych niż koprocesor krzywych eliptycznych gwarantujący to samo bezpieczeństwo [32]. Pojawiają się również koncepcje realizacji infrastruktury klucza publicznego za pomocą protokołów kryptograficznych, które wykorzystują algorytmy symetryczne.

W przypadku funkcji skrótu, konkurs na nowy standard SHA (ang. Secure Hash Algorithm) spowodował powstanie kilku algorytmów, które konstruowane były pod kątem ich wykorzystania w urządzeniach o ograniczonych zasobach logicznych. Najważniejszym z nich jest algorytm SPONGENT. Zwycięzca konkursu, czyli algorytm Keccak, przy zastosowaniu go z małymi wartościami parametrów bezpieczeństwa i wydajności również idealnie nadaje się do wykorzystania w urządzeniach IoT.

4.4. Inne metody

Interesującym zagadnieniem jest wykorzystanie istniejących podatności w urządzeniach IoT do zwiększenia ich bezpieczeństwa. Koncepcję takiego rozwiązania, które zapobiegać ma atakom typu DDoS, opisali w pracy De Donno i in. [33]. Idea polega na wykorzystaniu podatności wykorzystywanych przez botnet Mirai,

do wykrycia niewłaściwie zabezpieczonych urządzeń. Następnie na urządzenia takiego wgrywane jest oprogramowanie, które usuwa wgrane na nie złośliwe oprogramowanie. Dodatkowo następuje próba poinformowania użytkownika o występujących podatnościach. Jeżeli użytkownik nie podejmie żadnych czynności mających na celu usunięcie podatności, program sam zastosuje szereg czynności mających na celu zwiększenie bezpieczeństwa. Odbywa się to kosztem zewnętrznej ingerencji w urządzenie, czyli de facto przeprowadzeniem na nie ataku i zmodyfikowanie oprogramowania, czasem nawet bez świadomości użytkownika.

Warto zauważyć, że sami autorzy ataków na urządzenia IoT lub raportów dotyczących ich bezpieczeństwa [15, 19, 20, 25] opisują i analizują sposoby zwiększenia bezpieczeństwa Internetu Rzeczy. Co więcej, często współpracują oni również z producentami urządzeń nad sposobem usunięcia wykrytych podatności przed publikacją swoich prac.

Ostatnim, bardzo istotną metodą zwiększenia bezpieczeństwa Internetu Rzeczy jest zwiększenie wiedzy użytkowników końcowych na temat ważności tego zagadnienia. Wiele podatności urządzeń IoT może zostać wyeliminowanych poprzez prawidłową konfigurację urządzeń i zaimplementowanych w nich mechanizmów bezpieczeństwa.

5. Podsumowanie

Aspekt bezpieczeństwa urządzeń IoT był do tej pory bardzo często ignorowany zarówno przez producentów urządzeń, jak i użytkowników końcowych. Wynikiem tego jest mnogość ataków i wykrytych podatności w ostatnich latach. Kolejne głośne ataki, również na urządzenia krytyczne, które odpowiadają za zdrowie i życie użytkowników, spowodował wzrost świadomości wśród użytkowników i przedstawicieli biznesu. Zaowocowało to coraz większą liczbą układów dedykowanych do urządzeń IoT, które posiadają zaimplementowane mechanizmy bezpieczeństwa takie, jak np. algorytm szyfrowania AES czy sprzętowy generator liczb losowych. Rosnąca popularność oraz liczba urządzeń Internetu Rzeczy każe przypuszczać, że w najbliższych latach będziemy świadkami fascynującego pojedynku między atakującymi, a projektantami urządzeń IoT.

Literatura

1. Ashton K., *That ‘Internet of Things’ Thing*, RFID Journal, 2009.
2. Smith I., Sakamura K., Furness A., Ma R., Kim Y.-W., Walk E., Harmong G., Chartier P., Guillemin P., Armstrong D, *RFID and inclusive Model of Internet of Things, Final Report*, CASAGRAS, EU Project Number 216803, 2009.
3. Haller S., Karnouskos S., Schroth C., *The Internet of Things in an Enterprise Context*, Future Internet – FIS: First Future Internet Symposium, 2008, s. 14-28.
4. Duce H., Santucci G., Lange S., *Internet of Things in 2020: Roadmap for the future*, EC EPoSS Workshop, 2008.
5. Gartner, *Forecast: Internet of Things – Endpoints and Associated Services*, 2016.
6. Evans D., *The Internet of Things: How the Next Evolution of the Internet is Changing Everything*, CISCO White Paper, 2011.
7. Grodner M., Kokot W., Kolenda P., Krejtz K., Legoń A., Rytel P., Wierzbiński R., *Internet Rzeczy w Polsce*, IAB Expert Panel, 2015.
8. Eisenbarth T., S. Kumar, C. Paar, A. Poschmann, L. Uhsadel, *A Survey of Lightweight-Cryptography Implementations*, IEEE Design & Test of Computers, 2007, s. 522-533.

9. Vermesan O., Friess P., *Internet of Things – From Research and Innovation to Market Deployment*, River Publishers, 2014.
10. PWC, *Managing cyber risks in an interconnected world*, 2015, www.pwc.com.
11. PN-SIO/IEC-17799:2005, *Technika informatyczna. Praktyczne zasady zarządzania bezpieczeństwem informacji*, PKN, 2007.
12. Smith S., *The Internet of Risky Things. Trusting the Devices That Surround Us*, O'Reilly Media, 2017.
13. Kamkar S., *Keysweeper*, 2015, www.samy.pl.
14. Margaritelli S., *Nike+ FuelBand SE BLE Protocol Reversed*, 2015, www.evilsocket.net.
15. Ronen E., Shamir A., Weingarten A.-O., O'Flynn C., *IoT Goes Nuclear: Creating a ZigBee Chain Reaction*, Cryptology ePrint Archive: Report 2016/1047.
16. Ronen E., Shamir A., *Extended functionality attacks on IoT devices: The case of smart lights*, IEEE European Symposium on Security and Privacy EuroS&P, 2016, s. 3-12.
17. Wilkins A., Veitch J., Lehman B., *Led lighting flicker and potential health concern: IEEE standard par1789 update*, IEEE Energy Conversion Congress and Exposition, 2010, s. 171-178.
18. Miller C., Valasek C., *Remote Exploitation of an Unaltered Passenger Vehicle*, Black Hat USA, 2015.
19. Radcliffe J., *Hacking Medical Devices for Fun and Insulin: Breaking the Human SCADA System*, Black Hat USA, 2011.
20. Marin E., Singelee D., Garcia F.D., Chothia T., Willems R., Preneel B., *On the (in)security of the Latest Generation Implantable Cardiac Defibrillators and How to Secure Them*, Proc. of Annual Conference on Computer Security Applications ACSAC, 2016, s. 226-236.
21. Rios B., Butts J., *Security Evaluation of the Implantable Cardiac Device Ecosystem Architecture and Implementation Interdependencies*, WhiteScope, 2017, www.whitescope.io.
22. Ponemon Institute, *Medical Device Security: An Industry Under Attack and Unprepared to Defend*, Ponemon Institute Research Report, 2017, www.synopsys.com.
23. Makruszin D., *Czy Mirai jest naprawdę taki straszny, jak go malują?*, SecureList, 2016, www.securelist.pl.
24. Cluley G., *These 60 dumb passwords can hijack over 500000 IoT devices into the Mirai botnet*, 2016, www.grahamcluley.com.
25. Radware, *"BrickerBot" Results in Permanent Denial-Of-Service*, ERT Threat Alert, 2017.
26. STMicroelectronics, *STMicroelectronics and Security Platform Inc. to Combine Security Expertise for Safe, Trusted and Easy-to-Use Iot Devices*, 2017, www.st.com.
27. STMicroelectronics, *Wireless connectivity for IoT applications*, 2015.
28. Renesans Electronics, *S3A3 Microcontroller Datasheet*, 2017.
29. Infineon Technologies AG, *OPTIGA Trust P SLJ 52ACA: Programmable authentication and device security for a wide range of functions*, 2016.
30. Atmel Corporation, *ATECC508A Atmel CryptoAuthentication Device*, 2015.
31. Mazurkiewicz T., Rawski M., *Przegląd algorytmów kryptograficznych pod względem realizacji sprzętowego koprocesora do zastosowań mobilnych*, Przegląd Telekomunikacyjny i Wiadomości Telekomunikacyjne, 11 (2016), s. 1293-1298 .
32. Manifavas C., G. Hatzivasilis, K. Fysarakis, K. Rantos, *Lightweight Cryptography for Embedded Systems – A comparative Analysis*, Data Privacy Management and Autonomous Spontaneous Security, Springer, 2014, s. 333-349.
33. De Donno M., Dragoni N., Giaretta A., Mazzara M., *AntibIoTic: Protecting IoT Devices Against DDos Attacks*, arXiv eprint, 2017.

Aspekt bezpieczeństwa we współczesnych urządzeniach elektronicznych

Abstrakt
W ciągu ostatnich dwóch dekad można zaobserwować znaczny wzrost liczby urządzeń Internetu rzeczy (ang. Internet of Things, IoT) używanych przez nas w życiu codziennym. Wynalazki, które jeszcze niedawno wydawały się być fikcją literacką, teraz dostępne są dla każdego. Ekspres, który sam zaparza ulubioną kawę, lodówka informująca o braku jedzenia, a nawet sama zamawiająca brakujące produkty czy samoparkujące samochody, które dodatkowo doradzą drogę pozwalającą uniknąć korków. Wszystkie te urządzenia, podłączone do sieci i realizujące za nas część obowiązków, stają się nieodłączną częścią codzienności. Stosowane są nie tylko przez osoby prywatne, ale także w przemyśle, biznesie czy medycynie. Istotnym aspektem, który niestety jest często pomijany przez producentów urządzeń IoT, jest bezpieczeństwo. Każde nieodpowiednio zabezpieczone urządzenie może stanowić poważne zagrożenie dla jego użytkowników, np. w sytuacji, gdy osoba niepowołana sprawi, że wszczepiony pacjentowi rozrusznik serca nagle przestanie pracować. Co więcej, urządzenia te mogą zostać połączone w sieć (ang. botnet) i wykorzystane do przeprowadzania ataków typu DDoS na serwery lub strony internetowe, najczęściej bez wiedzy właściciela. W ramach niniejszej pracy przedstawiono znane przykłady ataków na urządzenia IoT i przeprowadzonych z ich wykorzystaniem. Przeanalizowano również możliwość zabezpieczenia urządzeń za pomocą algorytmów i protokołów kryptograficznych, a także z wykorzystaniem specjalizowanych układów z wbudowanymi zabezpieczeniami kryptograficznymi.
Słowa kluczowe: bezpieczeństwo informacji, Internet Rzeczy, kryptografia lightweight

Security aspect in modern electronic devices

Abstract
Over the past decade, there has been a noticeable increase in the number of the Internet of Things (IoT) devices that we use in our daily lives. Inventions that have appeared to be just a literary fiction recently are now available to everyone. An espresso machine that brews our favorite coffee on its own, a fridge that informs about the lack of food, and even orders missing products or self-parking cars, which additionally advise the way to avoid traffic jams. All of these devices, connected to the network and performing part of our duties for us, become an indispensable part of everyday life. They are used not only by individuals, but also in industry, business or medicine. An important aspect, which unfortunately is often overlooked by manufacturers of IoT devices is security. Any improperly secured device can pose a serious risk to its users, for example if a person who is not properly trained will cause the pacemaker to suddenly stop working. Moreover, these devices can be networked (botnet) and used to carry out DDoS attacks on servers or websites, most often without the owner's knowledge. This paper presents some well-known examples of attacks on IoT devices and attacks where they were used. The possibility of securing devices using cryptographic algorithms and protocols as well as using specialized systems with built-in cryptographic security is also analyzed.
Keywords: information security, Internet of Things, lightweight cryptography

Józef Okulewicz[1], Przemysław Żukowski[2]

Przełamanie bariery rozwoju transportu

1. Wprowadzenie

Prognozowanie przyszłego rozwoju transportu – podobnie jak i każdej innej dziedziny życia – zależy od stopnia zrozumienia stanu, w jakim znajduje się współczesna technika w danej dziedzinie. Jak jest to trudne można przekonać się na podstawie wystawy we Francuskiej Bibliotece Narodowej [1, 2]. M.in. można obejrzeć tam wizję artystyczną transportu w wieku XXI sporządzoną na początku XX w.

Obecnie można ją skonfrontować ze stanem faktycznym, jaki możemy już obserwować na co dzień. Okazuje się, że była to w dużym stopniu wizja oczekiwań, bez uwzględnienia uwarunkowań technicznych, jakich artysta ze zrozumiałych względów nie mógł być świadomy. Mamy na wystawie np. wizję kolei Paryż-Pekin, realizowaną przez elektryczny pociąg bez trakcji, lecz na bardzo małych kółeczkach i z otwartą kabiną maszynisty oraz osobnymi drzwiami do każdego przedziału w wagonie. Oznacza to, że do ówczesnych pociągów, jakie ukształtowały się po 80 latach rozwoju od powstania pierwszej publicznej towarowej linii kolejowej[3] pomiędzy Liverpool i Manchester [3], dodano jedynie napęd elektryczny.

Mamy też wizję komunikacji lotniczej, choć pierwszy lot braci Wright[4] odbył się zaledwie 7 lat wcześniej [4]. Po kolejnych 100 latach ludzie mieli poruszać się małymi samolotami, którymi można byłoby zatrzymywać się przy domowym tarasie lub częstować się kawą z tacy kelnera. Tej wizji nie udało się literalnie spełnić ze względu na nierozwiązanie problemu z napędem i zasilaniem takich samolotów, choć z Paryża do Pekinu raczej leci się samolotem zamiast jechać pociągiem.

Przykład ten oznacza, że artysta w swej wizji był mocno związany z ówczesnym stanem transportu, jaki można było obserwować na co dzień. W dziedzinie transportu, jakiej jeszcze nie było, poszybował daleko na skrzydłach wyobraźni, sugerując wynalezienie jakiegoś rewolucyjnego napędu, który jednak nie wyeliminuje kolei żelaznej.

Równie trudno wyobrazić sobie dziś postać jakiejś dziedziny życia za kilkadziesiąt, lat nie wiedząc dokładnie, co jest możliwe do osiągnięcia za pomocą współczesnej techniki i jakie są trendy jej rozwoju. Tego zaś nie można wiedzieć gdyż to, co widzi się na co dzień odbiega od tego, nad czym już pracują biura projektowe producentów rozmaitych urządzeń. Nie formułuje się nawet potrzeb, jakie będą zaspakajane przez projektowane obecnie urządzenia, bo to jest podporządkowane przyszłej kampanii marketingowej.

[1] okulewicz@poczta.onet.pl, Wydział Transportu, Politechnika Warszawska
[2] zukowski8@gmail.com, Decorum Architekci
[3] 15 września 1830 r.
[4] 17 grudnia 1903 r.

W niewielkim stopniu w prognozowaniu rozwoju jakiejś dziedziny bierze się pod uwagę to, czy aktualnie spełnia ona wszystkie oczekiwania wobec niej formułowane. Przykładem takiej dziedziny pozornie spełniającej wszelkie oczekiwania jest informatyka. Współczesne osiągnięcia w tej dziedzinie zapierają dech i tylko czas jest potrzebny do realizacji najbardziej fantastycznych pomysłów.

Choć wg filmowej wizji, komputer miał już w 2001 r. kierować wyprawą statku międzygalaktycznego, to wciąż nie widać przesłanek by mógł być do tego użyty. Zapomina się bowiem, że sercem wszelkich informatycznych przedsięwzięć jest procesor będący nieudolną namiastką maszyny matematycznej, której miał być ucieleśnieniem. Za tym idą konsekwencje, które przesłania się kolejnymi wersjami systemów operacyjnych i języków programowania, nie usuwając zasadniczych ułomności tego urządzenia. Od samego zarania są nimi ograniczone zasoby i rozciągłość działania w czasie. Bez ustosunkowania się nauki i praktyki do tych „wad wrodzonych" nie można wiarygodnie prognozować dalszego rozwoju informatyki. Zamiast bowiem obiecywanych ułatwień w życiu, użytkownicy będą nieustannie instalować coraz to nowsze systemy operacyjne i walczyć ze złośliwym oprogramowaniem.

Podobna sytuacja – równie słabo uświadomiona – dotyczy transportu, którego zadania po dwustu latach rozwoju wydają się być oczywiste, a czas ma tylko doprowadzić do pełnej ich realizacji. Tymczasem transport stoi przed zasadniczą barierą w swym rozwoju, której przejawem może być zanegowanie funkcji, dla jakiej transport istnieje, czyli przemieszczanie ludzi i towarów.

Wydaje się, że brak odpowiedniego napędu do pojazdów lądowych stanowi obecnie zasadniczą przeszkodę w formułowaniu wizji transportu w dowolnym przedziale czasowym. Albo taki napęd jest już wynaleziony i z jego użyciem tylko czeka się na zwrot z poczynionych wcześniej inwestycji, albo zostanie wynaleziony za lat kilka, kilkanaście czy kilkadziesiąt albo nigdy. Wiadomo tylko, że jest potrzebny, bo bez niego nie da się zrealizować dowolnej wizji transportu.

Zapowiadana rewolucja w transporcie indywidualnym[5], jaką miał spowodować pojazd z dwoma bocznymi kołami [5], nie dokonała się z prozaicznego powodu, jakim jest brak odpowiednio pojemnej baterii. Poza tym pojazd w codziennym użytkowaniu okazał się zbyt drogi, aby można go było zostawiać przed sklepem czy przed biurem, a za ciężki, by wszędzie zabierać go ze sobą. Po 15 latach istnienia na rynku sprawdza się na zamkniętych terenach lub pod ścisłym nadzorem.

Nie mając wglądu do tajnych laboratoriów przemysłowych czy wojskowych, trzeba więc wniknąć w istotę kłopotów, z jakimi boryka się współczesny transport, by na tej podstawie sformułować kierunki jego dalszego rozwoju. W tym celu trzeba zastanowić się, jaki wpływ na transport indywidualny miały uwarunkowania z początków rozwoju pojazdów mechanicznych, co z nich zostało do dzisiaj i czy nadal wywierają wpływ na rozwój transportu.

[5] Segway wprowadzono na rynek w 2002 r. po długiej i burzliwej kampanii marketingowej.

2. Kryzys transportu miejskiego

Na przełomie XVIII i XIX w. powstała możliwość napędzania pojazdów silnikiem mechanicznym, najpierw parowym, a potem elektrycznym lub spalinowym. Umożliwiło to przewożenie ładunków na znaczne odległości przy stosunkowo niskim koszcie. W krótkim czasie samochód stał się też dostępny dla przeciętnego człowieka. W efekcie nieco ponad dwustuletniego okresu burzliwego rozwoju środków transportu z napędem mechanicznym w znaczący sposób zmienił styl życia ludzi. Współcześnie wytwarza się ogromną liczbę pojazdów różnych rodzajów. Towarzyszy temu rozwój niezbędnej infrastruktury, czyli dróg, stacji paliw, itp.

Jednak podróżowanie na terenach aglomeracji miejskich, w których żyje coraz więcej ludzi, staje wobec narastającego kryzysu. Jego przejawem jest wzrastające zatłoczenie ulic dużymi pojazdami, wiozącymi pojedyncze osoby oraz rosnące zanieczyszczenie powietrza spowodowane spalinowymi silnikami pojazdów.

Jako jedyną receptę na ten kryzys promuje się transport zbiorowy, czy to w postaci kolei podziemnej, tramwajów czy autobusów. Ma on w niedalekiej przyszłości wyprzeć z miast, a przynajmniej z ich centrów, samochodowy transport indywidualny Jednak w praktyce oznacza to, że poza godzinami szczytu pojedyncze osoby byłyby wożone pojazdami większymi niż tzw. pojazdy indywidualne. Do tego wzrósłby czas oczekiwania na przejazd oraz wydłużyłaby się droga, jaką mieszkańcy przebywaliby na piechotę.

Dlatego coraz bardziej popularny staje się rower, jako indywidualny środek transportu, idealny do poruszania się w mieście. Jest mały, lekki i nie zajmuje wiele miejsca na drodze. Jest też łatwo dostępny dla użytkowników. Na potrzeby transportu rowerowego w miastach powstają rozległe sieci ścieżek rowerowych. Problemem jest jednak ekspozycja pasażera na niekorzystne warunki środowiskowe.

Do jazdy w większym komforcie ale bez zanieczyszczania środowiska, próbuje się zamienić silnik spalinowy na elektryczny zarówno w pojazdach indywidualnych jak i w autobusach. Jednak jako samochód elektryczny proponuje się coś, co przypomina dzisiejszy samochód z tą różnicą, że w środku ma być silnik elektryczny. Reszta, czyli przekładnia, układ kierowniczy pozostałyby bez zmian. Jeśli w przypadku autobusów byłoby to do przyjęcia, to jest dziwne w odniesieniu do pojazdów indywidualnych. Taki pojazd mogący mieć oddzielny silnik elektryczny w każdym kole, mógłby być sterowany tak, jak czołg, czyli przez zmianę prędkości kół po obu stronach pojazdu. Zatem wystarczyłyby dwa takie koła z przodu, a z tyłu dwa (albo jedno) koła wleczone. Jednak przeszkodą mogą być dotychczasowe przyzwyczajenia i wyobrażenia użytkowników odnośnie pojazdów indywidualnych.

Tę zamianę możnaby potraktować jako kolejny krok w rozwoju pojazdów indywidualnych, gdyby nie wątpliwość czy rozwój ten idzie w dobrym kierunku. Aby jednak wskazać alternatywny kierunek rozwoju transportu warto zrozumieć, dlaczego ukształtował się on w taki a nie inny sposób.

Jak wiadomo pojazdy z napędem mechanicznym pojawiły się wkrótce po wynalezieniu maszyny parowej[6]. Jednak początkowo były one tak ciężkie, że nie

[6] Pierwszy pojazd o napędzie parowym zbudował Nicolac Cugnot we Francji w 1769 r.

nadawały się na ówczesne drogi [6], przygotowane pod znacznie lżejsze zaprzęgi konne. Dlatego napęd parowy zamontowano w pojazdach jeżdżących po szynach. Ze względu na ich wielkość, wagę i koszt (nie tylko pojazdów ale i torowiska dla nich) nie opłacało się używać ich do wożenia pojedynczych pasażerów. Dlatego pojazdy te musiały zabierać dużą liczbę pasażerów.

Jednak postęp techniczny doprowadził do zmniejszenia silnika parowego i już w 1815 r. Józef Bożek pokazał jeżdżący parowy samochód osobowy [7], który jednak nie był produkowany. Dopiero ważny impuls rozwojowy pojawił się wraz z wynalezieniem silnika spalinowego. Był on na tyle mały i lekki w stosunku do silnika parowego, że Karl Benz mógł go wykorzystać w 1885 r. do budowy pojazdu osobowego [8].

W tamtych, czasach takim pojazdem była kareta lub powóz, zabierający od jednej do kilku osób. Ponieważ pojazd był ciągniony przez konie po wyboistych drogach, to dla złagodzenia wstrząsów na nierównych nawierzchniach dróg i zwiększenia wygody jazdy stosowano resory i amortyzatory. Zaś w trosce o ciągnące pojazdy konie, w pojazdach stosowano koła o dużej średnicy. W efekcie podłoga była umiejscowiona wysoko i do pojazdu trzeba było wchodzić po schodkach.

Do takiego właśnie pojazdu został zamontowany silnik spalinowy[7]. Jednak wkrótce okazało się, że taki silnik mógł mieć na tyle dużą moc, że nie potrzeba było wysokich kół by móc się szybko i wygodnie poruszać. W związku z tym można było zmniejszyć koła[8] i obniżyć podłogę pojazdu.

W efekcie tej technicznej ewolucji ukształtował się kompromis pomiędzy mocą, wagą i średnicą kół, czego wynik można obserwować na przykładzie współczesnych samochodów. Wysokość podłogi jest przy tym odpowiednio dobierana do typu pojazdu, umożliwiając użytkownikom wygodne wsiadanie i wysiadanie.

Można zatem przyjąć, że misja pionierów motoryzacji została wypełniona. Uwolnili oni pojazdy od koni, których utrzymanie było bardzo kosztowne. Dzięki temu jeżdżenie zmodyfikowanymi w ten sposób karetami stało się bardziej dostępne.

Jednak z takiego ukształtowania pojazdów do komunikacji lądowej wynikły trwające do dzisiaj konsekwencje. Albowiem efektem tego rozwoju jest pojazd kilkuosobowy, który zazwyczaj jest wykorzystywany przez jedną lub dwie osoby. To zaś powoduje nadmierne zatłoczenie miast samochodami utrudniając korzystanie z tych pojazdów. Ujawniające się w taki m.in. sposób wewnętrzne sprzeczności transportu powodują, że coraz słabiej wykonuje on swą funkcję przewożenia ludzi i ładunków.

Dlatego z inicjatywy urbanistów, William Lark zaproponował w 2005 r. [9] wiele koncepcji pojazdów miejskich do transportu indywidualnego. Jak można było się spodziewać, były to małe, krótkie pojazdy, których wspólną cechą były cztery małe kołeczka jezdne. Przyjęło się bowiem, że pojazdy – tak jak przedwieczne karety – muszą jeździć na czterech kołach. Do jazdy na takich małych kółeczkach potrzebny

[7] Dokonywali tego na przełomie XIX i XX w. tacy pionierzy motoryzacji jak Giovanni Agnelli, Karl Benz, André-Gustave Citroen, Gottlieb Daimler, Henry Ford, Wilhelm Maybach, Armand Peugeot, Louis Renault oraz wielu może mniej dziś znanych konstruktorów.

[8] Dodatkowo zredukowano drgania pojazdu, używając opon nałożonych na koła.

był silnik elektryczny o odpowiednio dużej mocy oraz pojemny akumulator. Z kompromisu pomiędzy wagą zasilania i zasięgiem pojazdu wynikło tyle, że pojazd taki mógłby przez godzinę jechać z prędkością 60 km/h. Po takim rajdzie czekało pasażera wielogodzinne ładowanie akumulatorów.

Dopiero w 2010 r. jego promotor William John Mitchel ze współautorami [10], dodał do tej kolekcji dwukołowy pojazd PUMA [11] opracowany w 2009 r. przez firmę Segway oraz podobne dwukołowce opracowane w General Motors na światową wystawę w Szanghaju [12]. Zasięg i prędkość tych pojazdów pozostały bez zmian, gdyż elektryczny silnik musiał dodatkowo utrzymywać równowagę pojazdu, mając nad osią koła cały ciężar pasażera.

Prawdopodobnie autorzy zdawali sobie sprawę z tego, że pojazdy te nie spełniają kryterium szyku[9]*, jakie zaproponowali do oceny propozycji przyszłych pojazdów miejskich*[10]. Pomimo to byłyby one do zaakceptowania ze względu na uwarunkowania transportu w mieście, gdyby nie krótki czas działania. To dyskwalifikowało je jako alternatywę transportu zbiorowego[11].

W takiej sytuacji upowszechnienie znacznie dłuższego, tradycyjnego samochodu z podmienionym na elektryczny silnikiem nie spowoduje nowej rewolucji w transporcie, bo nie zmieni to istoty pojazdu. Energia do napędzania takiego samochodu jest pobierana z wiezionego akumulatora, jest obecnie ciężki i wymaga długiego czasu odnawiania[12].

Choć po zamianie napędu samochód będzie mniej – a może i wcale – zanieczyszczał środowisko, będzie też bezpieczniejszy ze względu na wzajemne informatyczne oddziaływanie na siebie samochodów podczas jazdy, to nadal będzie to kareta na czterech kołach z silnikiem. Tym razem elektrycznym.

Dlatego, niestety, los samochodów indywidualnych w mieście w postaci, w jakiej dotrwały do czasów współczesnych, wydaje się być przesądzony. Podmiana rodzaju napędu nieco przedłuży tylko ich agonię.

Można więc sądzić, że ten alternatywny kierunek rozwoju transportu indywidualnego zamarł i pozostała tylko komunikacja zbiorowa, gdyż proponowany krótki pojazd nie może konkurować z masowym transportem publicznym. Gdyby nie to, że w poszukiwaniu innowacyjnych metod rozwiązania problemów komunikacyjnych w miastach, nie zwraca się uwagi na to, że paradoksalnie, praprzyczyną współczesnych kłopotów z transportem lądowym jest właśnie zastosowane w pojazdach koło.

Choć w ciągu minionej historii pojazdów samochodowych wiele się zmieniło, to jedno pozostało bez zmian. Od samego początku używa się kół z centryczną osią. Jest to skutkiem przekonania, że w kwestii tej podstawy transportu, czyli koła, które

[9] Przyjęto kryteria: wygodny (*convenient*), miejski (*urban*), sieciowy (*networked*), szyk (*chic*) i adaptacyjny (*adaptive*).

[10] Niestety, ale na bazie współczesnej techniki, aby pojazd był krótki, to musi mieć malutkie 4 kółeczka. Współczesną wersję takiego pojazdu firma Honda pokazała na Tokijskim Motor Show 2015, wstydliwie zakrywając czwórkę kółeczek [13].

[11] Do dzisiaj nie powstał drugi egzemplarz PUMA, a w GM pod zarządem komisarycznym zlikwidowano dział, w którym opracowano prototypowe pojazdy.

[12] Nieporównanie dłuższego od czasu tankowania paliwa. Z tego względu przewiduje się, że np. powierzchnia stacji zasilania musiałaby być 10-krotnie większa w porównaniu do dzisiejszych stacji paliwowych.

umożliwiło powstanie i rozwój transportu lądowego nie zmieniło się nic od 4 i pół tysiąca lat. Jednak z faktu, że pojazd może toczyć się na takich kołach, nie wynika jeszcze, że właśnie tak należy realizować transport.

3. Kłopotliwe koło

Bezwładny obiekt można przemieszczać po poziomym lądzie na dwa sposoby. Albo ślizgając albo tocząc na kole. W obu przypadkach potrzebna jest siła przyłożona do takiego obiektu. Koło zapewnia mniejsze opory ruchu niż ślizg, a co za tym idzie do toczenia pojazdu jest potrzebna mniejsza siła napędowa. W praktyce i w teorii przyjęto za oczywisty fakt, że koło pojazdu lądowego bazuje na geometrycznej figurze opisanej przez zbiór punktów rozmieszczonych równomiernie wokół osi obrotu.

Naturalną przeszkodą w toczeniu się koła jest opór wytworzony przez obszar styku koła z podłożem. W stanie spoczynku obszar ten rozkłada się wokół pionowej osi przechodzącej przez środek geometryczny koła, który jednocześnie jest środkiem przyłożenia siły odpowiadającej ciężarowi pojazdu. Pokonywanie oporu toczenia dokonuje się za pomocą zewnętrznej siły przyłożonej do osi koła. Przyjmuje się, że jej wartość jest wprost proporcjonalna do ciężaru działającego na koło i współczynnika oporu toczenia a odwrotnie proporcjonalna do wielkości koła [14].

W przypadku pojazdów mechanicznych do uzyskania tej siły potrzebny jest silnik o odpowiedniej mocy, którego moment obrotowy działa na oś koła. Ze względu na jego przeciętne rozmiary i wagę, opłacalne są dopiero pojazdy kilkuosobowe. Z kolei w przypadku pojazdów bez napędu mechanicznego, takich jak np. rower, minimalizuje się ich wagę, wystawiając kierującego na przykre działanie środowiska i skutki kolizji z innymi pojazdami. Do tego dochodzi konieczność budowania dróg i torów, aby można było opory toczenia zmniejszać w czasie jazdy.

Konsekwencje takiego ujęcia zagadnienia toczenia koła są daleko idące i nieoczekiwane. W rezultacie dopasowywania pojazdów do oporów toczenia, w ruchu drogowym uczestniczą pojazdy o dużej rozpiętości zarówno rozmiarów jak i wagi. Skutkiem tego są zgubne następstwa kolizji pomiędzy pojazdami. Kolizje te traktowane jako nieuniknione zjawiska naturalne, przynoszą drastycznie różne obrażenia dla ich uczestników. Przy tym ci najbardziej narażeni nie są w żaden sposób chronieni, aby nie obciążać ich dodatkowo ciężkim wyposażeniem podczas jazdy.

Aby minimalizować ryzyko kolizji pojazdów o nieporównywalnych rozmiarach i masach postuluje się odrębne pasy jezdni dla każdego ze stosowanych środków transportu. W efekcie miasta stają się obszarem rywalizacji różnych środków transportu o ograniczony zasób powierzchni dróg.

Tak więc stosowanie koła centrycznego w transporcie lądowym prowadzi do rozmaitych wynaturzeń niweczących ideę wspomagania mobilności człowieka. Co więcej, używając w pojazdach koła z centralną osią nie można zrobić kolejnego kroku w rozwoju transportu.

4. Innowacyjne koło

Powszechność stosowania centrycznego koła w transporcie jest podtrzymywana przekonaniem o braku dla niego alternatywy. Nietrudno jednak zauważyć, że można zastosować inne podejście do toczenia koła. W tym celu wystarczy przesuwać w poziomie środek ciężkości koła poza jego środek geometryczny. Może to wystąpić

przypadkowo na skutek wadliwego wykonania koła. Wtedy jest to zjawisko niekorzystne dla toczenia się koła i przywraca się centralne położenie środka ciężkości przez odpowiednie wyważanie koła. Jednak może być ono wynikiem celowego działania, gdy w kole ma nastąpić zrównoważenie innej siły działającej na koło, jak np. w kole parowozowym.

Świadome niewyważenie koła było wykorzystywane od czasów starożytnych aż do Renesansu do napędzania młynów, dźwigów czy wyrzutni pocisków. Źródłem siły obracającej koło był człowiek idący wewnątrz koła po jego obrzeżu. Był on przesunięty względem środka koła i przez to mógł swoim ciężarem wprawiać je w ruch [15]. Pozostałością tej metody napędzania koła są współczesne kołowrotki dla zwierząt, choć często zbyt małe w stosunku do ich rozmiarów[13].

Oryginalny sposób przesuwania środka ciężkości koła – lecz bazujący w swej istocie na przesunięciu środka ciężkości w napędach starożytnych dźwigów – wynalazł polski filozof Józef Hoene Wroński [16]. Wykorzystał on zjawisko występujące podczas synchronicznego obrotu dwóch kół o różnych promieniach. Gdy środki tych kół są przesunięte względem siebie, to punkt na obwodzie mniejszego koła wyznacza okrąg na tle większego z kół.

Gdy w kilku takich punktach zamocuje się rolki i umożliwi ich toczenie się po okręgach wyznaczonych na tle dużego koła, to powstanie koło, w którym punkt przyłożenia ciężaru może wysuwać się poziomo przed środek geometryczny dużego koła, toczącego się po podłożu.

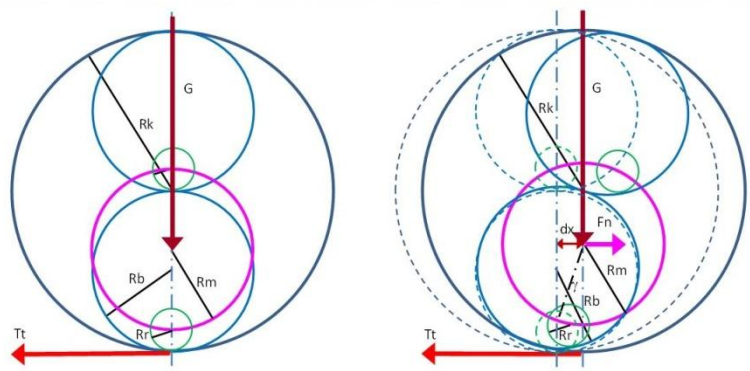
Dzięki temu poziomą siłą działającą na oś mniejszego koła można przesuwać środek ciężkości koła względem jego środka geometrycznego. Zatem można dynamicznie wytwarzać koło mimośrodowe pokonując jedynie opór toczenia rolek po okrągłych bieżniach utworzonych w dużym kole. Po wytworzeniu takiego mimośrodu siła ciężkości dąży do jego zniwelowania, poprzez toczenie dużego koła w kierunku wysuniętego środka ciężkości.

Jeśli środek ciężkości koła jest przesunięty na stałe względem jego środka geometrycznego, to w czasie ruchu koła następuje opadanie środka ciężkości aż do osiągnięcia osi przechodzącej przez środek geometryczny. W przeciwieństwie jednak do statycznego koła mimośrodowego, w kole o przesuwnej osi napędowej obrót dużego koła powoduje także obrót bieżni dla rolek i tym samym obrót małego koła. Skutkiem tego jest utrzymanie środka ciężkości koła na tej samej wysokości.

Zasadę działania koła zilustrowano na przykładowym kole z dwoma bieżniami (rys. 1). W praktyce należałoby stosować więcej bieżni aby zwiększyć sztywność poprzeczną koła[14].

[13] Taki kołowrotek dopasowany do rozmiarów człowieka można wypróbować w Centrum Nauki „Kopernik" w Warszawie.

[14] Wynalazca pozostawił modele pojazdów z kołami o 6 i 10 bieżniach. Są one przechowywane w Bibliotece Kórnickiej PAN, w Kórniku.



a) b)

Rysunek 1. Schemat dwubieżniowego koła o przesuwnej osi napędowej:
a) w stanie spoczynkowym, b) po przesunięciu [źródło: opracowanie własne]

Można więc zmodyfikować model zjawiska przezwyciężania oporu toczenia. Jego istotą jest to, że w czasie przesuwania środka ciężkości koła ustawia się ono pod wpływem siły ciężkości tak, by jego środek geometryczny znajdował się na pionowej osi razem ze środkiem ciężkości. Wraz ze środkiem geometrycznym przesuwa się obszar styku koła z podłożem tak, by znajdował się on pod środkiem ciężkości koła.

Występują przy tym trzy przypadki. Jeden, gdy koło mimośrodowe, w którym mimośród jest stały i większy od zera. Drugi, gdy koło jest centryczne, w którym mimośród jest równy zero. Trzeci przypadek dotyczy koła z przesuwną osią napędową, w którym mimośród jest zmienny. Takie koło jest tożsame z kołem centrycznym, jeśli promienie wewnętrznych bieżni są równe zeru.

Koło mimośrodowe w zastosowaniach transportowych powinno być wyważone. Bez tego toczy się samoistnie, jeśli zostanie ustawione tak, by środek ciężkości był odległy w poziomie od środka geometrycznego. Moment obrotowy powodujący toczenie koła jest proporcjonalny do ciężaru działającego na to koło i odległości w poziomie środka ciężkości od pionowej osi symetrii koła. Ustaje on gdy środek ciężkości znajdzie się pod środkiem geometrycznym koła.

W pozostałych przypadkach, zewnętrzną siłę potrzebną do pokonania oporów toczenia koła centrycznego ($R_b = 0$) lub koła z przesuwną osią napędową ($R_b > 0$) wyraża równanie (1):

$$F_n = \begin{cases} \frac{G \cdot \mu_k}{R_k} gdy\ R_b = 0, \\ \frac{G \cdot \mu_r}{R_r} gdy\ R_b > 0, \end{cases} \qquad (1)$$

gdzie F_n – siła napędowa, G – ciężar przypisany do koła, μ_k – współczynnik oporu toczenia koła po podłożu, μ_r – współczynnik oporu toczenia rolki po bieżni, R_k – promień koła, R_b – promień bieżni, R_r – promień rolki toczącej się po bieżni.

W celu toczenia koła centrycznego siła pozioma przyłożona do jego osi przesuwa jednocześnie środek ciężkości wraz ze środkiem geometrycznym. Jest ona proporcjonalna do ciężaru działającego na to koło i współczynnika oporu toczenia koła po podłożu.

W kole z przesuwną osią ta pozioma siła przyłożona do jego osi przesuwa tylko środek ciężkości przed środek geometryczny. Jest ona proporcjonalna do ciężaru działającego na to koło i współczynnika oporu toczenia rolek po bieżniach. Na ramieniu wytworzonym na skutek takiego przesunięcia działa siła przyłożonego do koła ciężaru, która powoduje przesuwanie środka geometrycznego w kierunku środka ciężkości koła.

Zatem działając na rolki siłą poziomą potrzebną do pokonania oporu toczenia tych rolek po gładkich bieżniach, powoduje się toczenie tego koła po podłożu. Stosunek tej siły poziomej do siły potrzebnej do pokonania oporu toczenia koła przedstawia wzór (2):

$$k = \frac{F_r}{F_k} = \frac{\mu_r \cdot R_k}{\mu_k \cdot R_r} \quad (2)$$

gdzie F_r – pozioma siła napędowa, potrzebna do przezwyciężenia oporu toczenia rolek po bieżniach, F_k – siła potrzebna do przezwyciężenia oporu toczenia koła po podłożu.

W literaturze dotyczącej tarcia tocznego są podawane wartości współczynnika oporu tocznego dla różnych materiałów [17]. Nie uwzględniają one ani promienia koła[15] ani zakrzywienia powierzchni, po jakiej koło się toczy. Przy tym ograniczeniu, przykładowo dla μ_r = 0,5 mm (toczenie stali po stali) i μ_k = 25 mm (toczenie gumy po betonie) oraz koła o r_k = 300 mm i rolkach o r_r = 30 mm, wartość tego stosunku wynosi k = 0,2. Zatem wykorzystując różnice oporu toczenia dla różnych materiałów, siła potrzebna do poruszenia koła o przesuwnej osi napędowej może być znacznie mniejsza od siły potrzebnej do toczenia analogicznego koła centrycznego.

Ponieważ mniejsze koło służy tylko do umocowania rolek, to może być ono zastąpione dowolną figurą odpowiadającą liczbie rolek.

Koła z czterema bieżniami wykorzystujące tę zasadę zamontowano w rowerze (rys. 2). Dodano do niego przystawkę umożliwiającą przesuwanie się osi napędowej tylnego koła[16]. Próbne jazdy wykonane przez autorów z uwzględnieniem przesuwania tej osi i po jej zablokowaniu, potwierdzają oczekiwane własności koła.

Rysunek 2. Rower wykorzystujący koła z przesuwną osią napędową [źródło: fot. Marek Antoszkiewicz]

[15] Oznaczałoby to, że koła o różnych promieniach wytwarzałyby jednakowy obszar styku koła z podłożem.
[16] Oś przedniego koła może przesuwać się razem z rowerem.

Przez niemal 180 lat od tego wynalazku uważano, że był to zapomniany praprototyp czołgowej gąsienicy [18]. Jednak nie ma on z gąsienicą nic wspólnego [19]. W napędzie gąsienicowym koła toczą się po otaczającej je sztucznej drodze. Ta swego rodzaju „ruchoma droga" umożliwia pojazdowi poruszanie się po zbyt miękkim podłożu. W przypadku koła o przesuwnej osi napędowej też mamy do czynienia z tzw. „ruchomą drogą", lecz tę funkcję pełnią dla rolek bieżnie w kole. Umożliwiają one wysuwanie się środka ciężkości koła przed jego środek geometryczny. To zjawisko nie występuje w napędzie gąsienicowym.

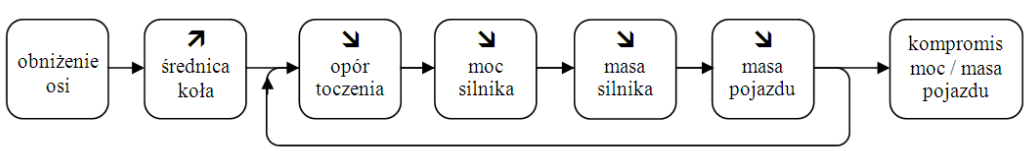
Wykorzystując w pojazdach koło z przesuwną osią napędową można uzyskać znaczące zmniejszenie potrzebnej mocy silnika napędzającego pojazd. Stopień tego zmniejszenia będzie zależał od konstrukcji koła jak i całego pojazdu oraz od stopnia zwiększenia oporów wewnętrznych w stosunku do koła centrycznego. Spowoduje to znaczące zmiany w sposobie wykorzystania pojazdów, a skutkiem tego będą zmiany warunków i stylu życia ludzi. Będą one co najmniej porównywalne do zmian jakie nastąpiły po upowszechnieniu się indywidualnych pojazdów samochodowych. Można tylko przypuszczać, że tempo tych zmian będzie o wiele szybsze.

5. Nowy pojazd indywidualny

Koło z przesuwną osią napędową wynalezione na potrzeby transportu przed 180 laty – mimo wysiłków wynalazcy – nie zostało zauważone i jego istnienie nie wpłynęło na rozwój środków transportu. Za to obecnie może być lekarstwem na bolączki współczesnego transportu. Przede wszystkim wymaga mniejszej siły do toczenia, co daje znaczące oszczędności energii, zarówno w przypadku silnika spalinowego jak i elektrycznego.

Do tego możliwość obniżenia osi napędowej istotnie wpływa na konstrukcję pojazdu. Jeden ze skutków to możliwość obniżenia podłogi pojazdu, do czego współcześnie wykorzystuje się kosztowne przekładnie planetarne. Drugi – to możliwość powiększenia średnicy koła bez zmiany położenia podłogi pojazdu, co znacznie zmniejsza opory toczenia.

To zaś daje możliwość dalszego obniżenia mocy silnika potrzebnej do napędzania pojazdu, a tym samym obniżenie jego ciężaru. To z kolei skutkuje możliwością dalszego obniżenia potrzebnej mocy silnika, itd. Prowadzi to do ukształtowania się nowego kompromisu pomiędzy mocą, wagą i średnicą kół pojazdu, odmiennego niż ten sprzed stu laty (rys. 3).



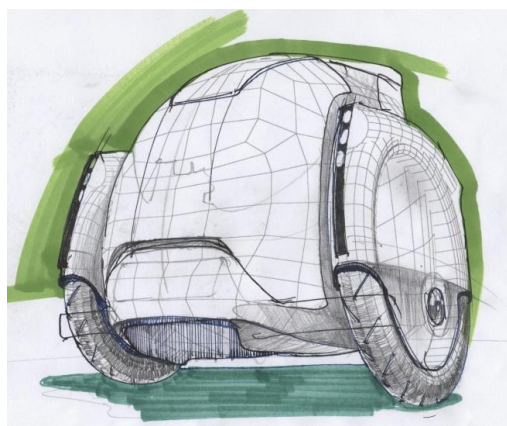
Rysunek 3. Ilustracja skutków obniżenia osi w kole [źródło: opracowanie własne]

W celu wyznaczenia zależności pomiędzy średnicą kół pojazdu a jego masą i mocą silnika można wykorzystać dane o mocy i wadze pojazdów obecnie produkowanych. Efektem będzie statystyczna zależność pomiędzy wagą a mocą pojazdu. Na podstawie

tej funkcji można ekstrapolować wagę pojazdu w zależności od stopnia obniżenia osi napędowej koła albo powiększenia jego średnicy.

Ponieważ maksymalne przesunięcie osi nie może przekroczyć promienia bieżni w kole, to w przypadku dwóch bieżni, przesunięcie nie może przekroczyć promienia współczesnego koła samochodu osobowego. W takim przypadku zmniejszyłaby się o połowę potrzebna moc, a średnica kół mogłaby być zwiększona dwukrotnie.

Pojazd wykorzystujący koło z obniżoną osią napędową, nie musiałby już być czterokołowy, gdyż w zakresie jego długości zabrakłoby miejsca na tak duże koła. Ponadto, jeśli koło jest duże a oś przesuwna, to łatwiej utrzymać stabilizację wzdłużną. Biorąc to pod uwagę pojazd mógłby mieć dwa koła umieszczone po bokach. To zaś dodatkowo zmniejszyłoby opór toczenia. Zatem silnik napędowy mógłby być jeszcze mniejszy, a pojazd jeszcze lżejszy. W efekcie długość pojazdu mogłaby nie przekraczać jednego czy dwóch metrów (rys. 4). To zaś może spowodować, że byłby on opłacalny dla przewożenia nawet jednej osoby. Już obecnie można przewidzieć jego wyposażenie ułatwiające użytkowanie, jak np. rozpoznawanie kierowcy, sterowanie za pomocą telefonu czy wyświetlanie na szybach informacji o parametrach jazdy jak i o możliwych zagrożeniach w otoczeniu pojazdu.



Rysunek 4. Szkic pojazdu dwukołowego [źródło: Adam Adamczyk, 2014].

Zatem sednem zbliżającej się rewolucji w transporcie jest powstanie nowych pojazdów, których do tej pory nie mogło być, ze względu na stosowane w pojazdach centryczne koło. Dotychczasowy środek do transportu indywidualnego, czyli kilkuosobowy samochód czterokołowy, zostałby zastąpiony przez dwuosobowy pojazd z dwoma dużymi bocznymi kołami. Byłoby to istotne wsparcie indywidualnej mobilności człowieka. Taki pojazd ze względu na swe niewielkie rozmiary mógłby towarzyszyć człowiekowi na całej trasie od miejsca zamieszkania do miejsca docelowego podróży.

6. Konsekwencje techniczne i społeczne

Zastosowanie w pojazdach nowego koła zmieni nie tylko dotychczasowe pojazdy, ale spowoduje powstawanie nowych konstrukcji pojazdów lądowych, które dziś trudno nawet sobie wyobrazić.

Konsekwencje upowszechnienia takich pojazdów będą przede wszystkim estetyczne. Nowy pojazd indywidualny na dużych kołach po bokach, prezentowałby

się okazale tak z przodu jak i z tyłu, zadając przy tym odpowiedniego szyku. Ponadto nadawałby się nie tylko na gładkie i płaskie miejskie drogi, ale także na pozamiejskie wertepy. Mógłby przy tym poruszać się po dowolnym podłożu, nie wymagając rozbudowy sieci dróg ani tym bardziej torowisk. A jak już zostałby przełamany mentalny opór przed dużymi kołami, to możnaby zacząć podróżować w cieniu jeszcze większych kół bocznych.

Zmieni się też rola transportu w mieście. W aspekcie technicznym zastosowanie krótszych pojazdów zmniejszyłoby zatłoczenie ulic. Do tego zwiększyłaby się przepustowość skrzyżowań i zmniejszyłyby się ewentualne kolejki przed skrzyżowaniami. Pojazdy mogłyby parkować przodem do krawędzi ulicy, zwiększając liczbę miejsc parkingowych. Gdyby jeszcze miały drzwi wejściowe z przodu, to mogłyby parkować blisko siebie.

W przypadku stosowania napędu spalinowego[17], nastąpiłoby zmniejszenie emisji spalin ze względu na zmniejszenie mocy silników. W przypadku zaś stosowania silnika elektrycznego[18] mniejsze byłoby zapotrzebowanie na liczbę punktów do ładowania akumulatorów, ze względu na większy zasięg pojazdów.

Przede wszystkim jednak taki jedno lub dwuosobowy pojazd na dwóch kołach mógłby już w niedalekiej przyszłości stać się podstawowym modułem mobilności, zastępując wieloosobowy pojazd indywidualny. Dla przewiezienia większej liczby osób, moduły łączyłoby się w zestawy, fizycznie lub elektronicznie. Na czas wspólnej podróży po mieście czy między miastami, pojazdy mogłyby jechać razem.

Początkowo taki moduł mógłby jeździć z kierowcą, lecz w dalszej przyszłości moduły jeździłyby automatycznie bez kierowcy, korzystając z infrastruktury miejskiej do orientowania się w przestrzeni. W takiej grupie sterowany mógłby być tylko pierwszy pojazd. Pozostałe podążałyby jego śladem w sposób automatyczny. Ich kierowcy mogliby zajmować się w tym czasie innymi czynnościami niż sterowanie. Po przejechaniu wspólnego odcinka, pojazdy rozłączałyby się i jechałyby samodzielnie w różnych kierunkach.

Wymagałoby to oczywiście istotnego wsparcia transportu odpowiednio zaawansowanym systemem informatycznym[19], zarówno w pojeździe jak i w ogólnym centrum sterowania ruchem. Jego głównym zadaniem byłoby wzajemne sprzęgnięcie zarówno pojazdów i infrastruktury jak i ich użytkowników w jeden system transportowy[20]. Nadzorowałby on ruch wszystkich pojazdów, reagując na bieżąco na możliwe zakłócenia. Ułatwiałby on dostępność w miejscu rozpoczynania podróży, czy to przez autonomiczny dojazd czy wskazanie pasażerowi drogi do pojazdu. Dzięki zaś wzajemnemu komunikowaniu się pojazdów i ich oddziaływaniu na siebie możliwe byłoby zapobieganie kolizjom i ułatwianie podejmowania decyzji w ruchu.

[17] Który ciągle jest bardziej atrakcyjny od elektrycznego, ze względu na krótki czas tankowania paliwa.
[18] Który z kolei umożliwia lepsze zagospodarowanie wnętrza pojazdu.
[19] Systemy informatyczne przyszłości powinny wykorzystywać komputery zbudowane na innej zasadzie niż te współczesne, które nieudolnie imitują teoretyczną maszynę matematyczną.
[20] Obecnie jest to realizowane lokalnie przez poszczególnych producentów samochodów [20]

Możnaby też zaprzestać rozwijania transportu zbiorowego w obecnym rozumieniu. Pozostałaby na pewno kolej podziemna, ale autobusy i tramwaje stałyby się zbędne. Do przewiezienia dowolnej liczby osób na danej trasie dobierałoby się odpowiednią liczbę modułów. Trasę możnaby dobierać na bieżąco uwzględniając potrzeby jadących pasażerów.

Na skutek zmniejszenia zapotrzebowania na moc silników napędowych pojazdów zmniejszyłyby się znacznie także koszty transportu towarów. Przy tym zmieniłby się także sposób ich przewożenia. Zbędne stałyby się duże samochody ciężarowe kursujące pomiędzy punktami gromadzenia ładunków, a zastąpiłyby je małe pojazdy rozwożące ładunki bezpośrednio od dostawcy do odbiorcy.

Mimo braku dużych pojazdów komunikacji zbiorowej ich torowiska mogłyby być wykorzystywane przez nowe pojazdy indywidualne. Choć może rozległe tereny rozjazdów, bocznic i mijanek kolejowych można byłoby przeznaczyć na inne cele.

Zarazem można przyjąć, że nowe pojazdy nie wyeliminują dotychczasowych środków komunikacji indywidualnej, takich jak motocykle, rowery, deskorolki, koła elektryczne, itp. Choć niewykluczone, że po zastosowaniu w nich nowego koła zmieniłby się ich kształt i wyposażenie.

W konsekwencji zmian w zakresie technicznych środków transportu następowałyby też zmiany w innych dziedzinach życia. Przede wszystkim wzrośnie mobilność ludzi, zmieniając przy tym swe formy. Zdolność do poruszania się będzie mogła być wykorzystywana indywidualnie przez poszczególnych ludzi bez konieczności zabierania przestrzeni innym ludziom dużymi pojazdami indywidualnymi i bez potrzeby dobierani chętnych do wspólnej jazdy dużymi pojazdami komunikacji zbiorowej. Łatwość i taniość indywidualnego poruszania się zmieni styl życia, pracy i wypoczynku.

Dzięki temu, że pojazdy będą mogły mieć niewielkie rozmiary, spełnią się oczekiwania urbanistów. Tym samym nie trzeba będzie szukać alternatywy dla transportu indywidualnego w dużych pojazdach transportu zbiorowego, zniekształcających strukturę miasta. W następstwie tego nowy środek lokomocji może spowodować inne podejście do projektowania miast i przestrzeni. Przewymiarowane obecnie – pod potrzeby dużych pojazdów zarówno komunikacji indywidualnej jak i zbiorowej – ulice można będzie przeobrazić w ciągi pieszo jezdne o mniejszym przekroju poprzecznym i uzupełnić o nową zabudowę. Umożliwiłoby to zagęszczenie zabudowy i zmniejszenie powierzchni tras komunikacyjnych w mieście. Spowoduje to konieczność zmiany obecnego prawa, w zakresie nasłonecznienia i zacieniania budynków w mieście, jak i wymagań odnoście parametrów infrastruktury miejskiej.

Przyczyni się do wprowadzenia bardziej zwartej zabudowy, mniejszych przekrojów ulic, dopasowanie przestrzeni do ludzkiej skali a nie do skali środków transportu. Można będzie zmniejszyć odległości pomiędzy budynkami, tworząc w miastach przyjazne warunki do mieszkania, pracy i wypoczynku. Dzięki rozwojowi indywidualnej komunikacji małymi pojazdami, miasta przestaną być miejscem gromadzenia pojazdów na zatłoczonych ulicach. Staną się przez to bardziej kameralne, dostosowane do potrzeb ludzi.

7. Podsumowanie

Przy wykorzystaniu współczesnych technologii połączonych z zapomnianym XIX wiecznym wynalazkiem koła, można będzie stworzyć optymalny środek transportu przeznaczony do poruszania się w miastach. Umożliwi to przełamanie coraz bardziej uświadamianej bariery rozwoju miejskiego transportu, jaką jest zatłoczenie zbyt dużymi pojazdami indywidualnymi. Zamiast projektowania miejskich pojazdów elektrycznych z ciężkimi akumulatorami należałoby podjąć badania nowej konstrukcji koła i warunków jego stosowania w pojazdach. Tylko w ten sposób można będzie utworzyć postulowane przez urbanistów krótkie pojazdy, likwidując w ten sposób zatłoczenie miast, oraz zmniejszając zanieczyszczenie środowiska dzięki zmniejszonemu zużyciu energii.

Literatura

1. Villemard, 1910, *Chromolithographie Paris*, BNF, Estampes. URL: http://expositions.bnf.fr/utopie/feuill/index.htm .
2. URL: https://pursuitist.com/french-postcards-from-1910-show-technology-visions-of-2000/.
3. URL: http://www.kulturaihistoria.umcs.lublin.pl/archives/1473.
4. URL: http://www.smartage.pl/bracia-wright-i-pierwszy-lot-samolotu/.
5. URL: http://pl-pl.segway.com .
6. URL: https://www.britannica.com/biography/Nicolas-Joseph-Cugnot.
7. URL: https://www.facebook.com/polski.samochod.ma.200.lat/?hc_ref= ARRciwmHqe0MpImLhlZg6z2Vv9P0UtI1u4bHGKDH-FJLFDBDSxWHGj fMlZncUmEnpqQ&fref=nf .
8. URL: https://www.mercedes-benz.com/en/mercedes-benz/classic/history/carl-benz/.
9. Lark W., *City-Car: Optimizing wehicle and Urban efficiencies through a shared adaptive platform*, MIT Press, 2005.
10. Mitchell W.J., Borroni-Bird Ch. i Burns L.D., *Reinventing the Automobile: Personal Urban Mobility for the 21st Century*, MIT Press, 2010.
11. URL: https://www.theguardian.com/technology/blog/2009/apr/07/puma-segway-electric-car , 7.04.2009.
12. URL: http://media.gm.com/autoshows/Shanghai/2010/public/cn/en/env/news .html.
13. URL: http://sputniknews.com/photo/20151028/1029215381/tokyo-motor-show-photos.html#ixzz3pyEV9tV3.
14. Morecki A., Oderfeld J., *Teoria maszyn i mechanizmów*, PWN Warszawa 1987.
15. Rossi C., Russo F., *Ancient Engineers' Inventions: Precursors of the Present*, Springer 2009, 2017.
16. Hoene Wroński J., *Rails Mobiles + Plansze*, Paryż 1836 (rękopis w Bibliotece Kórnickiej PAN, Kórnik).
17. URL: http://www.roymech.co.uk/Useful_Tables/Tribology/co_of_frict.htm.
18. Chomicz P., *Hoene-Wroński w Polsce i zagranicą*, wyd. F. Hoesick, Warszawa 1929.
19. Okulewicz J., *Koło Wrońskiego nie jest gąsienicą*, Przegląd techniczny, 7-8 (2014), s. 28.
20. URL: http://swiatoze.pl/wizjonerstwo-utopia-nissana/.

Przełamanie bariery rozwoju transportu

Abstrakt
Po dwustuletnim okresie rozwoju środków transportu z napędem mechanicznym ujawnia się bariera jego rozwoju, która wymusza rezygnację z mobilności indywidualnej na rzecz transportu zbiorowego. Tym samym neguje się ideę, jak inspirowała pionierów motoryzacji. Proponowana przez autorów recepta na przełamanie tej bariery polega na wykorzystaniu zapomnianego polskiego wynalazku z XIX w., czyli koła z przesuwną osią. Umożliwiłoby to konstruowanie nowatorskich pojazdów, w szczególności do transportu miejskiego, cechujących się mniejszymi rozmiarami i mniejszym zużyciem energii. Upowszechnienie takich pojazdów zmieniłoby nie tylko styl podróżowania ale także sposób kształtowania miejskiej zabudowy. Zatem to nie rodzaj napędu lecz sposób poruszania się pojazdów umożliwi dalszy rozwój transportu w XXI w.
Słowa kluczowe: mimośród dynamiczny, mobilność, pojazd indywidualna

The overcoming a barrier in a transport progress

Abstract
After two century lasting developing means of transport with mechanical drive, a barrier reveals in their further progress. It forces abandoning an individual mobility in favour a mass transportation, especially in cities. This way an idea inspiring automobiles pioneers. A recipe for overcoming this barrier proposed by authors is application of XIX century polish invention of a wheel with moving axle. This could enable developing brand new innovative vehicles with two big wheels aside. They could be smaller and less polluting than today individual cars. Developing such vehicles will change not only ways of travelling but also will have an impact on urban structures. Thus it isn't an electric drive but a way of vehicles moving that changes the transportation in the XXI century.
Keywords: off-center wheel, mobility, individual vehicle

Magdalena Zawartka[1]

Sterowanie aktywnym zawieszeniem pojazdu z uwzględnieniem regulatora liniowo-kwadratowego oraz algorytmu ślizgowego

1. Wstęp

W niniejszym rozdziale wprowadzono trzy różne algorytmy sterowania: sterowanie ślizgowe z wykorzystaniem modelu referencyjnego (MRSMC) sky-hook, MRSMC ground-hook i regulator liniowo-kwadratowy (LQR) Zbadano wpływ zmiany rodzaju sterowania na efektywność układu wibroizolacji modelu ćwiartkowego zawieszenia. Każdy algorytm był testowany pod kątem trzech kryteriów: przemieszczeniowa funkcja przenoszenia drgań, kryterium ugięcia opony ('tire deflection') oraz zapotrzebowanie na energię zewnętrzną. Wygenerowano i przeanalizowano charakterystyki częstotliwościowe zawieszenia sterowanego dla wybranych wartości tłumienia $\boldsymbol{c_{sky}}, \boldsymbol{c_{ground}}$. Porównano wpływ zmian parametrów zawieszenia na przyjęte kryteria oceny wszystkich rozważanych regulatorów. Realizacja algorytmu sterowania ślizgowego (SMC) przy użyciu modelu odniesienia pozwala na uzyskanie porównywalnych charakterystyk częstotliwościowych do tych uzyskiwanych przez układy sterowane algorytmami sky-hook lub ground-hook. Co więcej użyty algorytm sterowania pozwala na obliczenie wartości generowanej siły przez element aktywny. Wygenerowana siła pozwala na osiągnięcie przez obiekt trajektorii fazowej wyznaczonej przez zastosowany model referencyjny. Trzecim rozważanym algorytmem jest sterowanie LQR. Regulator liniowo-kwadratowy został wyznaczony dla modelu zlinearyzowanego a badania zostały przeprowadzone dla modelu nieliniowego. Zaproponowany wskaźnik jakości uwzględnia zarówno poprawę bezpieczeństwa ('road-holding') oraz komfort (przemieszczeniowa funkcja przenoszenia drgań). Pozwala on również na ograniczenie mocy sygnału sterującego. Ponadto wszystkie parametry modelu zawieszenia pojazdu odpowiadają parametrom samochodu typu SUV Isuzu D-MAX.

Jednym z problemów występujących w praktycznej implementacji algorytmów sterowania są założenia uproszczające przyjęte przy modelowaniu. Efektem tego może być pogorszenie jakości sterowania. Mając na uwadze występujące w rozważanym układzie redukcji drgań uproszczenia modelu tłumika oraz niestacjonarność jego parametrów (np. zależność od temperatury) celowe jest zastosowanie metody sterowania odpornego. Regulator ślizgowy (Sliding Mode Control – SMC) znajduje zastosowanie zarówno w układach liniowych jak i nieliniowych. Cechuje go wysoka odporność na niedokładności modelu oraz zakłócenia. Głównym zadaniem sterowania ślizgowego jest sprowadzenie oraz utrzymanie trajektorii obiektu na wyznaczonej hiperpłaszczyźnie zwanej powierzchnią ślizgową. Inną popularną metodą sterowania

[1] zawartka@agh.edu.pl Akademia Górniczo-Hutnicza w Krakowie, WIMiR, KAP

jest zastosowanie regulatora odpornego np. H_∞. Ponieważ regulatory odporne bazujące na metodzie H_∞ stosuje się najczęściej do obiektów liniowych oraz nie uwzględnia się w nich ograniczeń na sygnały sterujące zdecydowano się na zastosowanie regulatora ślizgowego. Dodatkową jego zaletą w stosunku do układów minimalizujących normę H jest możliwość kształtowania charakterystyki częstotliwościowej po przez dobór modelu referencyjnego.

Implementując regulator ślizgowy należy rozważyć w jaki sposób przyjąć hiperpłaszczyznę ślizgową. Dla układu redukcji drgań o dwóch stopniach swobody (model ćwiartkowy) dąży się do zapewnienia kompromisu pomiędzy efektywnością redukcji drgań w zakresie rezonansowym i poza rezonansowym. Przyjęto, że sygnał sterujący będzie wyznaczany w taki sposób aby zachowanie rzeczywistego obiektu odwzorowało ruch modelu idealnego. Najlepszym rozwiązaniem tego problemu jest zastosowanie układu redukcji drgań bazującego na idealnym algorytmie sky-hook, który będzie używany jako model referencyjny (*Model Reference Control* – MRC) w układzie regulacji. Zadaniem regulatora ślizgowego będzie minimalizacja błędu śledzenia pomiędzy obiektem a modelem referencyjnym.

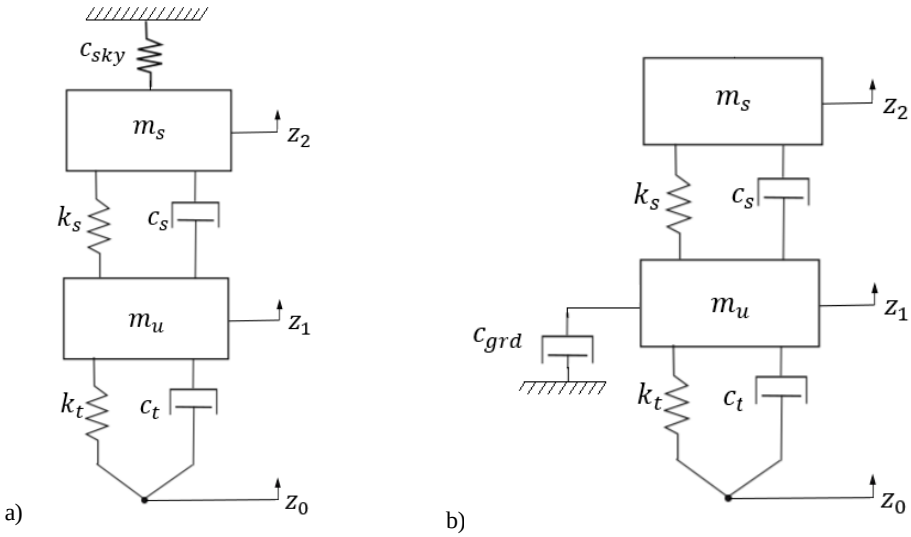
Trzecim rozważanym algorytmem jest sterowanie LQR. Efektywność działania takiego regulatora zależy od wybranej funkcji kosztu. W przypadku sterowania opartego na sprzężeniu od stanu obiektu, kluczowe dla jakości działania układu jest odpowiednie lokowanie biegunów układu zamkniętego. Wybór położenia biegunów układu zamkniętego w sposób arbitralny nie zapewnia optymalności otrzymanego rozwiązania. Wychodząc naprzeciw temu problemowi został opracowany regulator optymalny z kwadratowym wskaźnikiem jakości (LQR). Zastosowanie tego rozwiązania pozwala na określenie optymalnej macierzy sprzężenia od stanu umożliwiającej minimalizację funkcji kosztu. Dla postawionego problemu został wyznaczony regulator liniowo-kwadratowy dla modelu zlinearyzowanego a badania zostały przeprowadzone dla modelu nieliniowego.

2. Referencyjny model odniesienia

W literaturze z zakresu semi-aktywnych układów redukcji drgań można znaleźć rozwiązania, które mają na celu zapewnienie bezpieczeństwa, komfortu jazdy lub spełniają obydwa te kryteria [1], Najpopularniejszymi rozwiązaniami są układy sterowania typu: ground-hook, sky-hook lub rozwiązania godzące ze sobą te dwa podejścia.

2.1. Algorytm sky-hook

Algorytm sterowania typu sky-hook jest najczęściej rozważanym w literaturze algorytmem redukcji drgań w układach semi-aktywnych. Celem sterowania jest utrzymanie masy resorowanej (zahaczenie) w stałej odległości od wirtualnego punktu odniesienia (sky) niezwiązanego z nawierzchnią [3]. Działanie układu jest realizowane poprzez przymocowanie masy wibroizolowanej do pewnego punktu w niebie.



Rysunek 1. Algorytm sterowania a) typu sky-hook; b) typu ground-hook [opracowanie własne [7, 8, 10]]

W praktyce sprowadza się to do wprowadzenia siły proporcjonalnej do prędkości bezwzględnej masy wibroizolowanej m_s, a współczynnikiem proporcjonalności jest założony współczynnik c_{Sky}. Dzięki takiej budowie tłumienie pionowych oscylacji nadwozia może być realizowane w sposób niezależny od drgań koła. Schemat obliczeniowy przedstawiony na (Rysunek 1) opisano równaniami (1).

$$\begin{aligned} &m_s\ddot{z}_2 + k_s(z_2 - z_1) + c_s(\dot{z}_2 - \dot{z}_1) + c_{sky}\dot{z}_2 = 0 \\ &m_u\ddot{z}_1 + k_t(z_1 - z_0) + c_t(\dot{z}_1 - \dot{z}_0) - k_s(z_2 - z_1) \\ &\qquad - c_s(\dot{z}_2 - \dot{z}_1) = 0 \end{aligned} \tag{1}$$

Parametr c_{sky} reprezentuje współczynnik tłumienia elementu łączącego nadwozie z zewnętrznym punktem odniesienia (w niebie). Współczynnik tłumienia c_s reprezentuje tłumienie występujące w zawieszeniu. W rzeczywistych aplikacjach semi-aktywnych nie jest możliwe zbudowanie takiego układu redukcji drgań. Największym problemem jest zaczepienie tłumika we wspomnianym punkcie ze względu na to, że sterowane tłumiki generują siłę proporcjonalną do prędkości względnej (tłoka względem cylindra) i skierowaną przeciwnie do niej. Implementacja tego algorytmu w aplikacjach rzeczywistych sprowadza się do zastosowania tłumienia opisanego wzorem (2). W przypadku gdy $\dot{z}_2(\dot{z}_2 - \dot{z}_1) \leq 0$ należałoby przyjąć wartość współczynnika tłumienia równą 0. Nie jest to możliwe ze względów konstrukcyjnych, przyjmuje się wówczas wartość minimalną możliwą do realizacji c_{min}.

$$c_{vs} = \begin{cases} c_{sky} \dfrac{\dot{z}_2}{\dot{z}_2 - \dot{z}_1} \; dla \; \dot{z}_2(\dot{z}_2 - \dot{z}_1) > 0 \\ c_{min} \; dla \; \dot{z}_2(\dot{z}_2 - \dot{z}_1) \leq 0 \end{cases} \tag{2}$$

Z kolei w przypadku gdy $\dot{z}_2(\dot{z}_2 - \dot{z}_1) > 0$ - przyjmuje się wartość współczynnika tłumienia równą założonej wartości c_{sky}. Zwiększanie wartości tłumienia c_{sky} dla idealnego sterowania sky-hook poprawia efektywność układu redukcji drgań. W aplikacjach rzeczywistych wybór zbyt dużych wartości tłumienia c_{sky} może spowodować pogorszenie efektywności redukcji drgań. Wartość tłumienia związana jest z doborem elementu wykonawczego.

Podsumowując zastosowanie algorytmu sterowania typu sky-hook reprezentuje podejście zorientowane na komfort jazdy. Największą wadą wynikającą z zastosowania algorytmu typu sky-hook jest możliwość ograniczenia przyczepności zawieszenia. Wady tej pozbawiony jest algorytm ground-hook.

2.2. Algorytm ground-hook

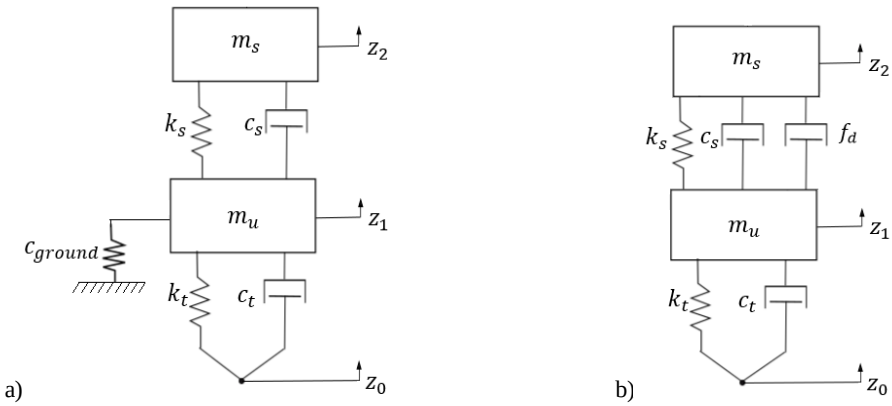
W celu poprawienia przyczepności pojazdu najczęściej stosowany jest algorytm ground-hook. Jego zadaniem jest zapewnienie ciągłego docisku masy niewibro-izolowanej do nawierzchni [10]. Jest to realizowane poprzez generowanie siły dociskającej tę masę do podłoża (Rysunek 2b). Analogicznie jak w przypadku sterowania typu sky-hook elementem wykonawczym jest tłumik przymocowany do wirtualnego nieruchomego punktu. Dynamikę rozważanego układu redukcji drgań opisuje układ równań (3) gdzie parametr c_{grd} reprezentuje współczynnik tłumienia elementu łączącego masę niewibroizlowaną z nieruchomym punktem.

$$\begin{aligned} &m_s\ddot{z}_2 + k_s(z_2 - z_1) + c_s(\dot{z}_2 - \dot{z}_1) = 0 \\ &m_u\ddot{z}_1 + k_t(z_1 - z_0) + c_t(\dot{z}_1 - \dot{z}_0) + c_{grd}\dot{z}_1 - k_s(z_2 - z_1) \\ &\qquad - c_s(\dot{z}_2 - \dot{z}_1) = 0 \end{aligned} \tag{3}$$

Tak jak w poprzednim przypadku nie jest możliwe zbudowanie takiego układu redukcji drgań. W rzeczywistych aplikacjach nie sposób przytwierdzić tłumik poruszającego się samochodu na stałe do umownego punktu. Podobnie jak w przypadku sterowania sky-hook implementacja algorytmu w aplikacjach rzeczywistych sprowadza się do zastosowania tłumienia opisanego wzorem (4).

$$c_{vg} = \begin{cases} c_{grd} \dfrac{\dot{z}_1}{\dot{z}_2 - \dot{z}_1} dla \; -\dot{z}_1(\dot{z}_2 - \dot{z}_1) > 0 \\ c_{min} \; dla \; -\dot{z}_1(\dot{z}_2 - \dot{z}_1) \leq 0 \end{cases} \tag{4}$$

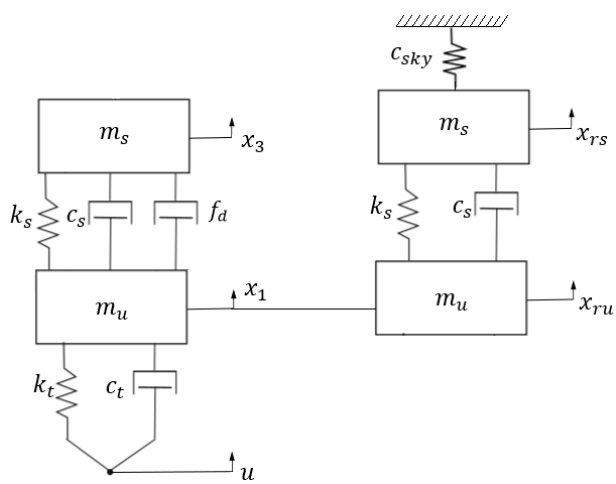
Zastosowanie algorytmu typu ground-hook umożliwia efektywną redukcję pionowych oscylacji osi koła, natomiast jedno z głównych zadań jakim jest skuteczna redukcja drgań masy wibroizolowanej nie zostaje zrealizowane. Zatem algorytm ten poprawia bezpieczeństwo jednak nie gwarantuje komfortu jazdy.



Rysunek 2. Model układu redukcji drgań ground-hook a) idealny, b) rzeczywisty [opracowanie własne [7, 8, 10]]

2.3. Model referencyjny sky-hook

W rozdziale tym skupiono się na przybliżeniu idei zastosowania modelu odniesienia w implementacji algorytmu sterowania ślizgowego z modelem referencyjnym typy sky-hook. Koncepcję sterowania ślizgowego z modelem referencyjnym dla regulatora typu sky-hook przedstawiono na (Rysunek 3). Jak wspomniano wcześniej celem sterowania ślizgowego w tym przypadku jest uzyskanie trajektorii współrzędnej opisującej ruch masy resorowanej m_S w obiekcie rzeczywistym odwzorowującej tę trajektorię w modelu referencyjnym [11].



Rysunek 3. Sterowanie ślizgowe z modelem odniesienia typu sky-hook [opracowanie własne [7], [8], [10]]

Model referencyjny został przyjęty jako układ o jednym stopniu swobody (1 DOF). Wymuszeniem dla tego układu jest przemieszczenie masy nieresorowanej obiektu x_1. Dla tak przyjętego modelu zapisano równanie równowagi dynamicznej (5).

$$m_s\ddot{x}_{rs} + k_s(x_{rs} - x_1) + c_s(\dot{x}_{rs} - \dot{x}_1) + c_{sky}\dot{x}_{rs} = 0 \tag{5}$$

Przyjęto model obiektu rzeczywistego o dwóch stopniach swobody (2DOF). Równania równowagi sił dla obiektu rzeczywistego opisano wzorami (6):

$$\begin{gathered} m_s\ddot{x}_3 + k_s(x_3 - x_1) + c_s(\dot{x}_3 - \dot{x}_1) - f_d = 0 \\ m_u\ddot{x}_1 + k_t(x_1 - u) + c_t(\dot{u} - \dot{x}_1) + k_s(x_1 - x_3) + c_s(\dot{x}_1 - \dot{x}_3) \\ + f_d = 0 \end{gathered} \tag{6}$$

Algorytm sterowania ślizgowego z modelem referencyjnym został sformułowany dla układu opisanego równaniami (5) i (6).

2.4. Model referencyjny ground-hook

W kolejnym kroku rozważono zastosowanie idealnego modelu ground-hook jako modelu referencyjnego. Dla tak przyjętego modelu odniesienia zadaniem regulatora ślizgowego jest uzyskanie trajektorii współrzędnej opisującej ruch masy nieresorowanej m_u w obiekcie rzeczywistym odwzorowującej tę trajektorię w modelu referencyjnym.

W przeciwieństwie do poprzedniej sytuacji model referencyjny został przyjęty jako układ o dwóch stopniach swobody (2 DOF). Wymuszeniem dla tego układu jest ten sam sygnał jak dla obiektu regulacji. Dla tak przyjętego modelu odniesienia zapisano równanie równowagi dynamicznej

(7).

$$\begin{gathered} m_s\ddot{x}_{r2} + k_s(x_{r2} - x_{r1}) + c_s(\dot{x}_{r2} - \dot{x}_{r1}) = 0 \\ m_u\ddot{x}_{r1} + k_t(x_{r1} - u) + c_t(\dot{x}_{r1} - \dot{u}) + c_{ground}\dot{x}_{r1} \\ - k_s(x_{r2} - x_{r1}) - c_s(\dot{x}_{r2} - \dot{x}_{r1}) = 0 \end{gathered} \tag{7}$$

Przyjęto model obiektu rzeczywistego o dwóch stopniach swobody (2DOF). Równania równowagi sił dla obiektu rzeczywistego opisano wzorami (8).

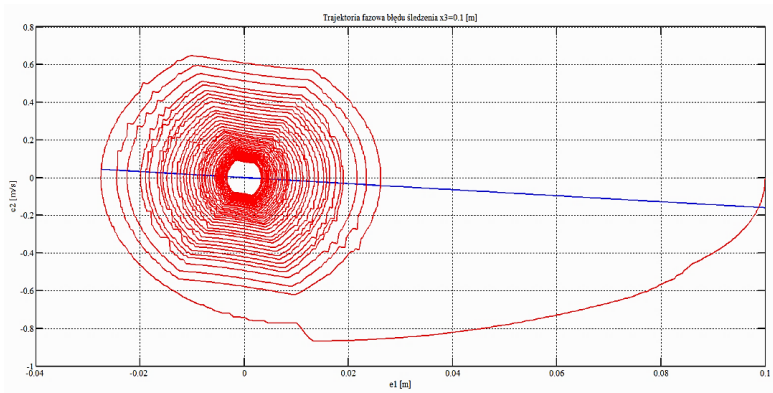
$$\begin{gathered} m_s\ddot{x}_3 + k_s(x_3 - x_1) + c_s(\dot{x}_3 - \dot{x}_1) - f_d = 0 \\ m_u\ddot{x}_1 + k_t(x_1 - u) + c_t(\dot{u} - \dot{x}_1) + k_s(x_1 - x_3) + c_s(\dot{x}_1 - \dot{x}_3) \\ + f_d = 0 \end{gathered} \tag{8}$$

Algorytm sterowania ślizgowego z modelem referencyjnym został sformułowany dla układu opisanego równaniami (7) i (8).

3. Regulator ślizgowy

Metodę sterowania ślizgowego wykorzystano do syntezy sterowania odpornego bazującego na model referencyjnym obiektu [2]. Celem sterowania jest osiągnięcie trajektorii obiektu, które będą odwzorowywały trajektorie modelu referencyjnego. W zaproponowanym algorytmie sterowania ślizgowego odwzorowanie trajektorii pomiędzy obiektem a modelem uzyskuje się generując sygnał sterujący minimalizujący błąd śledzenia trajektorii [6]. Jedną z głównych zalet takiego rozwiązania jest zredukowanie wpływu niestacjonarności parametrów obiektu [4]. W proponowanym układzie redukcji drgań zastosowano idealny model sky-hook oraz ground-hook jako model referencyjny. Ideą takiego rozwiązania jest stworzenie zamkniętego układu sterowania w taki sposób aby trajektorie układu rzeczywistego odpowiadały trajektoriom modelu referencyjnego (Rysunek 4).

Zadaniem regulatora jest zapewnienie aby błąd śledzenia pomiędzy obiektem a odpowiadającym mu modelem referencyjnym zmierzał asymptotycznie do zera.



Rysunek 4. Trajektoria fazowa błędu śledzenia [opracowanie własne]

Aby zobrazować działanie algorytmu sterowania ślizgowego dla rozpatrywanego przypadku przedstawiono wyniki symulacji działania aktywnego układu redukcji drgań. Wykres (Rysunek 4) przedstawia trajektorię fazową błędu śledzenia modelu referencyjnego przez obiekt sterowania. Z położenia równowagi została wychylona masa resorowana obiektu. Kolorem niebieskim zaznaczono powierzchnie ślizgową. Kolorem czerwonym trajektorie fazową obiektu zmierzającą od wartości wymuszenia do wartości 0.

3.1. Sterowanie ślizgowe z modelem referencyjnym sky-hook

Poniżej przedstawiono wyprowadzenie wzoru na siłę generowaną przez tłumik dla algorytmu sterowania ślizgowego z referencyjnym modelem typu sky-hook.

Wektor błędu śledzenia zdefiniowano wzorem (9):

$$e = [x_3 - x_{rs}\ \dot{x}_3 - \dot{x}_{rs}]^T = [e_1\ \dot{e}_1]^T = [e_1\ e_2]^T \tag{9}$$

Zmienna e_1 odpowiada różnicy pomiędzy przemieszczeniem masy resorowanej obiektu a przemieszczeniem masy resorowanej modelu referencyjnego. Natomiast zmienna e_2 odpowiada różnicy pomiędzy prędkościami tych mas. Po uwzględnieniu równań (5) i (6) oraz (9) sformułowano równania różniczkowe opisujące błąd śledzenia trajektorii (10).

$$e = A_e\dot{e} + H_r X_r + B_e f_d \tag{10}$$

Poszczególne macierze współczynników $A_e, H_r, oraz\ B_e$ przedstawiono wzorami (11).

$$A_e = \begin{bmatrix} 0 & 1 \\ -\frac{k_s}{m_s} & -\frac{c_s}{m_s} \end{bmatrix}, B_e = \begin{bmatrix} 0 \\ -\frac{1}{m_s} \end{bmatrix}, \quad H_e = \begin{bmatrix} 0 & 0 \\ 0 & \frac{c_{sky}}{m_s} \end{bmatrix}, \tag{11}$$

Zdefiniowano powierzchnię ślizgową dla wektora błędu śledzenia (12).

$$\begin{aligned} &s = ce, c = [c_1\ 1] \\ &s = c_1(x_3 - x_{rs}) + (\dot{x}_3 - \dot{x}_{rs}) \end{aligned} \tag{12}$$

Powierzchnia ślizgowa spełnia warunek (13).

$$s = \dot{s} = 0 \tag{13}$$

Po podstawieniu do równania (13) oraz zróżniczkowaniu otrzymujemy równanie (14).

$$\dot{s} = c1(\dot{x}_3 - \dot{x}_{rs}) + (\ddot{x}_3 - \ddot{x}_{rs}) \tag{14}$$

Siłę równoważną f_{deq} wyznacza się rozwiązując równanie (15). W przypadku gdy wektor stanu znajduje się na powierzchni ślizgowej oraz nie występują zakłócenia siła ta gwarantuje, że trajektoria systemu będzie znajdowała się na powierzchni ślizgowej.

Aby wyznaczyć siłę równoważną f_{deq} (15) należy do równania (14) podstawić równanie (6).

$$f_{deq} = [-k_s\ c_1 m_s - c_s]e + c_{sky}\dot{x}_3 \tag{15}$$

Ostatnim krokiem w syntezie sterowania ślizgowego jest określenie warunku zbieżności trajektorii do powierzchni ślizgowej. W tym celu wybrano funkcję w postaci (16).

$$V(s) = \frac{1}{2}s^2 \tag{16}$$

Warunkiem zbieżności trajektorii do powierzchni ślizgowej jest zagwarantowanie, że funkcja $V(s(x(t)))$ jest malejąca. Wynika stąd, że pochodna funkcji $V(s)$ po trajektoriach systemu musi być ujemna (17).

$$\frac{d}{dt}\frac{1}{2}s^2 = s\,\dot{s} = s(c_1\dot{e_1} + e_2) = s(c_1\dot{e_1} + \dot{e_2}) < 0 \tag{17}$$

Po przekształceniach uzyskano wzór na siłę generowaną przez tłumik (18).

$$f_d = \begin{cases} f_{deq} + \gamma sgn(s), \left(f_{deq} + \gamma sgn(s)(\dot{x}_3 - \dot{x}_1)\right) \geq 0 \\ 0, \left(f_{deq} + \gamma sgn(s)(\dot{x}_3 - \dot{x}_1)\right) < 0 \end{cases} \tag{18}$$

Podstawiając wzór (18) do (12) oraz uwzględniając wzór (11) otrzymano równanie (19). Pochodna funkcji opisującej wybraną powierzchnie ślizgową zależy od pewnej dodatniej stałej γ oraz znaku funkcji s.

$$\frac{ds}{dt} = -\gamma sgn(s), \gamma > 0 \tag{19}$$

Wynika stąd, że dla sterowania semi-aktywnego realizującego siłę zdefiniowaną wzorem (18) trajektorie systemu będą zmierzać asymptotycznie do zadanej powierzchni ślizgowej.

3.2. Sterowanie ślizgowe z modelem referencyjnym ground-hook

Poniżej w analogiczny sposób wyprowadzono wzór na siłę generowaną przez tłumik dla algorytmu sterowania ślizgowego z referencyjnym modelem typu ground-hook. W tym przypadku wektor śledzenie zostanie utworzony z różnic pomiędzy masą nieresorowaną obiektu i masą nieresorowaną modelu referencyjnego.

Wektor błędu śledzenia zdefiniowano wzorem (20):

$$e = [x_1 - x_{ru}\ \dot{x}_1 - \dot{x}_{ru}]^T = [e_1\ \dot{e}_1]^T = [e_1\ e_2]^T \tag{20}$$

Zmienna e_1 odpowiada różnicy pomiędzy przemieszczeniem masy nieresorowanej obiektu a przemieszczeniem masy nieresorowanej modelu referencyjnego. Natomiast zmienna e_2 odpowiada różnicy pomiędzy prędkościami tych mas.

Zdefiniowano powierzchnię ślizgową dla wektora błędu śledzenia (21).

$$\begin{aligned} &s = ce, c = [c_1\ 1] \\ &s = c_1(x_1 - x_{ru}) + (\dot{x}_1 - \dot{x}_{ru}) \end{aligned} \tag{21}$$

Powierzchnia ślizgowa spełnia warunek (22).

$$s = \dot{s} = 0 \tag{22}$$

Po podstawieniu do równania (21) oraz zróżniczkowaniu otrzymujemy równanie (23).

$$\dot{s} = c1(\dot{x}_1 - \dot{x}_{ru}) + (\ddot{x}_1 - \ddot{x}_{ru}) \tag{23}$$

Siła równoważna f_{deq} w przypadku gdy wektor stanu znajduje się na powierzchni ślizgowej oraz nie występują zakłócenia gwarantuje, że trajektoria systemu będzie znajdowała się na powierzchni ślizgowej.

Ostatnim krokiem w syntezie sterowania ślizgowego jest określenie warunku zbieżności trajektorii do powierzchni ślizgowej. W tym celu wybrano funkcję w postaci (24).

$$V(s) = \frac{1}{2}s^2 \tag{24}$$

Warunkiem zbieżności trajektorii do powierzchni ślizgowej jest zagwarantowanie, że funkcja $V(s(x(t)))$ jest malejąca. Wynika stąd, że pochodna funkcji $V(s)$ po trajektoriach systemu musi być ujemna (25).

$$\frac{d}{dt}\frac{1}{2}s^2 = s\,\dot{s} = s(c_1\dot{e_1} + e_2) = s(c_1\dot{e_1} + \dot{e_2}) < 0 \tag{25}$$

Po przekształceniach uzyskano wzór na siłę generowaną przez tłumik (26).

$$f_d = \begin{cases} f_{deq} + \gamma sgn(s), \left(f_{deq} + \gamma sgn(s)(\dot{x}_3 - \dot{x}_1)\right) \geq 0 \\ 0, \left(f_{deq} + \gamma sgn(s)(\dot{x}_3 - \dot{x}_1)\right) < 0 \end{cases} \tag{26}$$

Pochodna funkcji opisującej wybraną powierzchnie ślizgową (27) zależy od pewnej dodatniej stałej γ oraz znaku funkcji s.

$$\frac{ds}{dt} = -\gamma sgn(s), \gamma > 0 \tag{27}$$

Wynika stąd, że dla sterowania semi-aktywnego realizującego siłę zdefiniowaną wzorem (26) trajektorie systemu będą zmierzać asymptotycznie do zadanej powierzchni ślizgowej

4. Regulator liniowo-kwadratowy

Rozwój metod sterowania przy kwadratowym wskaźniku jakości (LQR) rozpoczął się w latach sześćdziesiątych. Zakłada się, że obiekt jest liniowy i znany jest jego opis w przestrzeni stanów. Zadanie polega na wyznaczeniu sterowania minimalizującego

kwadratowy wskaźnik jakości. Zakłada się ponadto, że system znajduje się w stanie równowagi. Celem sterowania jest utrzymanie systemu w stanie równowagi lub w zadanym punkcie pracy pomimo oddziałujących zakłóceń. Zadania regulacji typu nadążnego mogą być przekształcone do problemu regulacji stałowartościowej. Założenie stanu ustalonego umożliwia rozszerzenie horyzontu optymalizacji do nieskończoności. Przyjęcie horyzontu skończonego prowadzi do rozwiązań regulatorów typu niestacjonarnego, które są trudniejsze do analizy i implementacji.

W układach optymalnych struktura i parametry regulatora określone są na podstawie obliczonego ekstremum przyjętego wskaźnika jakości (tzw. kryterium kosztów) a sygnał sterujący optymalizuje ten wskaźnik. Sterowanie optymalne zasadniczo ma charakter dynamiczny (to znaczy dotyczy podejmowania decyzji w odniesieniu do pewnego przedziału czasu). Przykładem takiego układu może być układ sterowania ciągiem silników tak, aby samolot osiągnął określony pułap, przy minimalizacji wskaźnika jakości, którym jest zużycie paliwa.

System opisany równaniami (28) poddany zostaje zadaniu optymalizacji.

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u},\ \mathbf{x}(0) = \mathbf{x}_0, \tag{28}$$

$$\mathbf{y} = \mathbf{C}\mathbf{x}$$

$$J = \frac{1}{2}\int_0^{\tau} (\mathbf{x}^{\mathrm{T}}\mathbf{Q}\mathbf{x} + \mathbf{u}^{\mathrm{T}}\mathbf{R}\mathbf{u})dt \tag{29}$$

Q – macierz symetryczna dodatnio półokreślona,
R – macierz symetryczna dodatnio określona,
τ - horyzont optymalizacji.

Zadanie polega na wyznaczeniu minimalizującego wskaźnik jakości (29).
Dla systemu rzędu pierwszego wskaźnik jakości przyjmuje postać (30):

$$J = \frac{1}{2}\int_0^{\tau} (qx^2 + ru^2)dt \tag{30}$$

Wskaźnik reprezentuje ważoną sumę energii sterowania i stanu (dokładności sterowania). Jeżeli r jest bardzo duże, energia sterowania musi być niewielka, aby zapewnić niewielką wartość wskaźnika. Można wtedy stosować mniejsze silniki, elementy wykonawcze, wzmacniacze, itp. Jeżeli q jest większe od r, otrzymuje się układ zamknięty o dużym tłumieniu (bez wielkich zmian i przeregulowań).

Równanie (31) nazywane jest równaniem Riccatiego. Jest to nieliniowe macierzowe równanie różniczkowe, musi być rozwiązywane wstecznie w czasie ze względu na warunki końcowe. Rozwiązując równanie Riccatiego unika się rozwiązywania dwupunktowego problemu granicznego.

$$-\dot{\mathbf{P}} = \mathbf{A}^{\mathrm{T}}\mathbf{P} + \mathbf{P}\mathbf{A} + \mathbf{Q} - \mathbf{P}\mathbf{B}\mathbf{R}^{-1}\mathbf{B}^{\mathrm{T}}\mathbf{P},\ \mathbf{P}(\tau) = \mathbf{0} \tag{31}$$

Powyższe sformułowanie problemu LQR jest znane jako problem o horyzoncie skończonym. Jako wynik otrzymuje się regulator niestacjonarny.

$$\mathbf{u}(t) = -\mathbf{K}(t)\mathbf{x}(t) \tag{32}$$

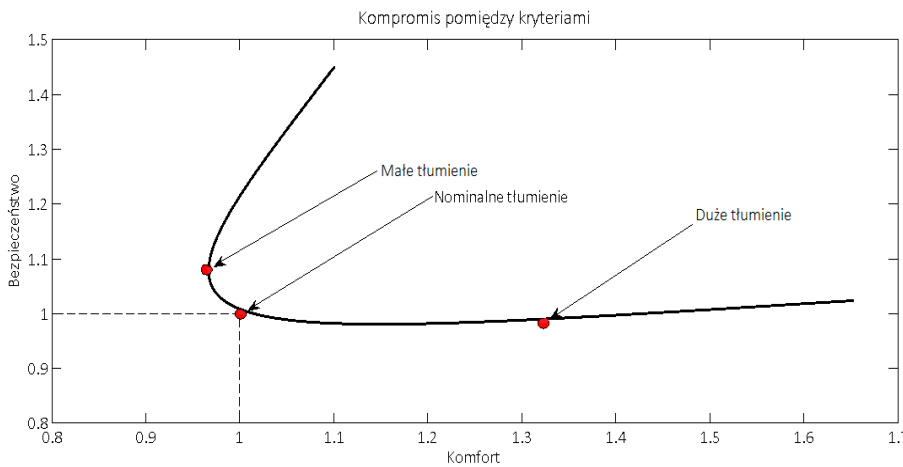
$$\mathbf{K}(t) = \mathbf{R}^{-1}\mathbf{B}^{\mathrm{T}}\mathbf{P}(t) \tag{33}$$

W problemie o horyzoncie nieskończonym $\tau \to \infty$ wskaźnik jakości musi być zbieżny, aby istniał regulator optymalny.

Jeżeli istnieje sterowanie optymalne, to układ zamknięty może być niestabilny.

5. Wybrane kryteria oceny zawieszeń

Na rysunku (Rysunek 5) przedstawiono realizację dwóch kryteriów będących miarą efektywności J zawieszenia. Można nadać kryteriom mierzalny charakter.



Rysunek 5. Kompromis pomiędzy bezpieczeństwem a komfortem jazdy [8]

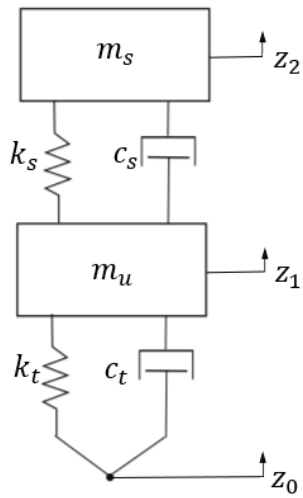
Pierwszy problem można rozwiązać poprzez zastosowanie sztywnego zawieszenia. Drugi poprzez zaprojektowanie komfortowego - miękkiego zawieszenia. Te dwa postulaty reprezentują klasyczny konflikt pomiędzy bezpieczeństwem a komfortem jazdy. Jak widać na rysunku (Rysunek 5) zastosowanie sztywnego zawieszenia nie idzie w parze z komfortem jazdy. Podobna sytuacja ma miejsce również w drugim przypadku.

Dla pasywnego układu zawieszenia przy ustalonych wartościach c *i* k możliwe jest uzyskanie kompromisowej konfiguracji tych parametrów. W odniesieniu do współczesnych wymagań stawianych zawieszeniom pojazdu kompromis ten nie zapewnia zarówno wymaganego komfortu jak i bezpieczeństwa. Semi-aktywne i aktywne metody redukcji drgań mają za zadanie pogodzić te dwa sprzeczne kryteria, a w szczególności:

- zapewnić pasażerowi odpowiedni poziom komfortu,
- utrzymać stabilność pojazdu przy różnych warunkach,

- kontrolować położenie masy wibroizolowanej względem koła,
- minimalizować przemieszczeniową funkcję przenoszenia drgań masy resorowanej.

Podstawowym wskaźnikiem stosowanym w ocenie układów wibroizolacji jest współczynnik przenoszenia drgań. Współczynnik ten często rozpatrywany jest w funkcji częstotliwości i w zawieszeniach definiowany jako przemieszczeniowa funkcja przenoszenia drgań [8]. Poniżej przedstawiono przyjęte kryteria oceny rozważanego układu wibroizolacji (Rysunek 6) w postaci transmitancji operatorowej.



Rysunek 6. Model pasywnego zawieszenia pojazdu (2 DOF) [opracowanie własne na podstawie [8]]

Przemieszczeniowa funkcja przenoszenia drgań związana z przemieszczeniem masy wibroizolowanej m_s (34) i wyznaczana jako moduł z transmitancji. Służy ona do oceny komfortu jazdy.

$$F_{z_2}(s) = \frac{z_2(s)}{z_0(s)} \tag{34}$$

Minimalizując ten wskaźnik zmniejsza się amplitudę drgań masy wibroizolowanej co znacząco poprawia wygodę.

Przemieszczeniowa funkcja przenoszenia drgań związana z przemieszczeniem masy niewibroizolowanej m_u (35) i wyznaczana jest jako moduł transmitancji. Odpowiada za ruch pionowy, w który wprawiane jest koło podczas pokonywania przeszkód.

$$F_{z_1}(s) = \frac{z_1(s)}{z_0(s)} \tag{35}$$

Maksymalizacja tego wskaźnika powoduje zwiększenie komfortu natomiast w przypadku układów aktywnych może ona zwiększyć pobór energii.

Kryterium ugięcia opony *„tire deflection”* odpowiada za oddziaływania dynamiczne między pojazdem a drogą i jest definiowane jako moduł z transmitancji opisanej wzorem (36).

$$F_{z_1-z_0}(s) = \frac{z_1(s) - z_0(s)}{z_0(s)} \tag{36}$$

Manewry takie jak: skręcanie, hamowanie wymagają stałego kontaktu z nawierzchnią (z ang. *road holding*). Poprawienie przyczepności jest możliwe jeżeli zmiany wartości normalnych obciążeń opony są minimalne. Dzieje się tak dlatego, że siły poprzeczne i wzdłużne generowane na oponie zależą właśnie od normalnego obciążenia opony [7]. Dzięki temu wskaźnikowi można także ocenić docisk koła do nawierzchni drogi. Jest to istotne dla współczesnych systemów bezpieczeństwa pojazdu takich jak ABS[2], ESP[3], EBD[4], EDS[5] itp. które wymagają stałego kontaktu koła z nawierzchnią.

Generalnie możemy dokonać podziału wskaźników ze względu na pełnione przez nie funkcje. Przemieszczeniowa funkcja przenoszenia drgań masy wibroizolowanej odpowiada za komfort jazdy pasażera. Natomiast docisk opony do podłoża odpowiada stricte za bezpieczeństwo jazdy samochodem. W przypadku pasywnego zawieszenia poprawa któregokolwiek z wymienionych kryteriów może skutkować pogorszeniem pozostałych. W przypadku zawieszenia sterowanego możliwe jest uzyskanie lepszych wyników dla poszczególnych kryteriów bez negatywnego wpływu na pozostałe.

5.1. Ograniczenia energetyczne

Dotychczas największym problemem związanym z aktywnym zawieszeniem było duże zapotrzebowanie na dostarczaną energię do układu w celu wygenerowania żądanej siły sterującej, mającej wpływ na kształtowanie charakterystyk tłumienia drgań. Wprowadzenie dodatkowej siły pozwala uzyskać wyższą wydajność w porównaniu do zawieszenia pasywnego. W tym momencie zazwyczaj pojawiał się drugi problem: w wielu zaprojektowanych aplikacjach utrzymanie tego efektu wiązało się z problemem zapewnienia niezawodności systemu. Ważnymi aspektami w projektowaniu zawieszeń aktywnych, na które w szczególności trzeba zwrócić uwagę są: możliwość awarii źródła zasilania oraz dostarczenie dużej ilości energii do układu, która może spowodować destabilizację systemu sterowania. Poniżej przedstawiono przyjęte dodatkowe kryteria oceny rozważanego układu wibroizolacji.

[2] ABS z ang. *Anti-Lock Braking System,*
[3] ESP z ang. *Electronic stability control,*
[4] EBD z ang. *Electronic Brakeforce Distribution,*
[5] EDS z niem. *Elektronische Differentialsperre*

Kryterium mocy chwilowej P_I odpowiada za moc chwilową pobieraną przez element aktywny (37) z układu zasilania, konwertowaną na pracę tego układu, zadaniem którego jest redukowanie drgań.

$$P_I = \frac{1}{T}\int_0^{\infty} P(t)dt, \qquad (37)$$
$$P(t) = |Fv|$$

Znajomość wartości mocy chwilowej pozwala na oszacowanie w jakich przedziałach mocy będzie pracował układ, jakie będzie jego zapotrzebowanie energetyczne.

Kryterium mocy maksymalnej P_{max}, reprezentuje moc maksymalną pobieraną przez element aktywny (38) z układu zasilania, konwertowaną na pracę tego układu, zadaniem którego jest redukowanie drgań.

$$P_{max} = \max(P_I) \qquad (38)$$

Wartość maksymalnej mocy przekłada się bezpośrednio na wielkość zastosowanego akumulatora.

6. Badania porównawcze

Przeprowadzono badania symulacyjne przedstawionych powyżej strategii sterowania. W praktyce implementacja przedstawionych modeli sprowadza się do zastosowania tłumika sterowanego. Do przeprowadzeniu symulacji numerycznych przyjęto model ćwiartkowy pojazdu o dwóch stopniach swobody (2DOF). Uwzględnia on oddziaływania pomiędzy masą resorowaną, nieresorowaną jak i podłożem. Jako element wykonawczy został wybrany tłumik o zmiennej (nastawialnej) sile tłumienia. Zapisano równania równowagi dynamicznej dla takiego układu (39).

$$m_s\ddot{x}_3 + k_s(x_3 - x_1) + c_s(\dot{x}_3 - \dot{x}_1) - f_d = 0 \ (3' \qquad (39)$$
$$m_u\ddot{x}_1 + k_t(x_1 - u) + c_t(\dot{u} - \dot{x}_1) + k_s(x_1 - x_3) + c_s(\dot{x}_1 -$$

W poniższym rozdziale przyjęto, że minimalna wartość tłumienia dla tego elementu wykonawczego wynosi $c_{min} = 0$. Oprócz sterowanego tłumika uwzględniono równolegle do niego położone tłumienie o stałej wartości wynikające z konstrukcji mechanicznej układu. Obydwa elementy odzwierciedlają rzeczywisty charakter działania modelu. Dla tak przyjętego układu zaimplementowano algorytmy sterowania ślizgowego typu ground-hook, sky-hook oraz algorytm LQR.

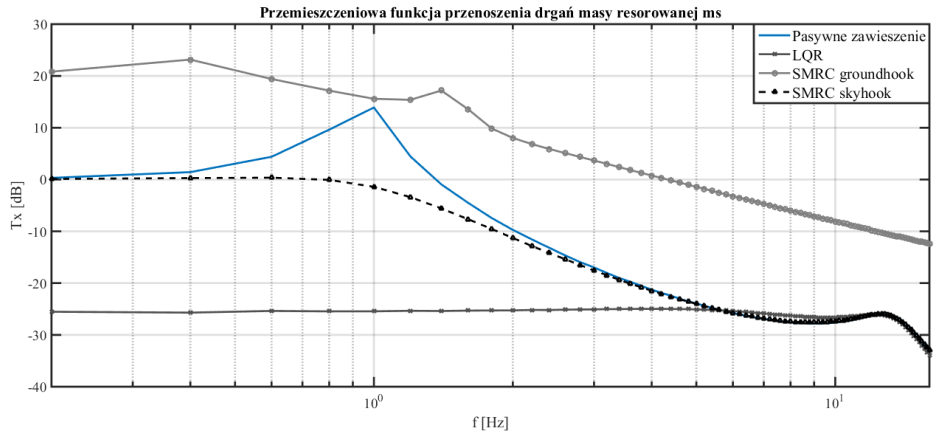
Do symulacji numerycznych przyjęto parametry jak w Tabeli 1. Jak wspomniano wcześniej wszystkie parametry modelu zawieszenia pojazdu odpowiadają parametrom samochodu typu SUV Isuzu D-MAX. W tabeli pojawia się również dodatkowy współczynnik α. Odpowiada on za przełożenie, które występuje w kolumnie zawieszenia tego samochodu.

Tabela 1. Parametry obiektu

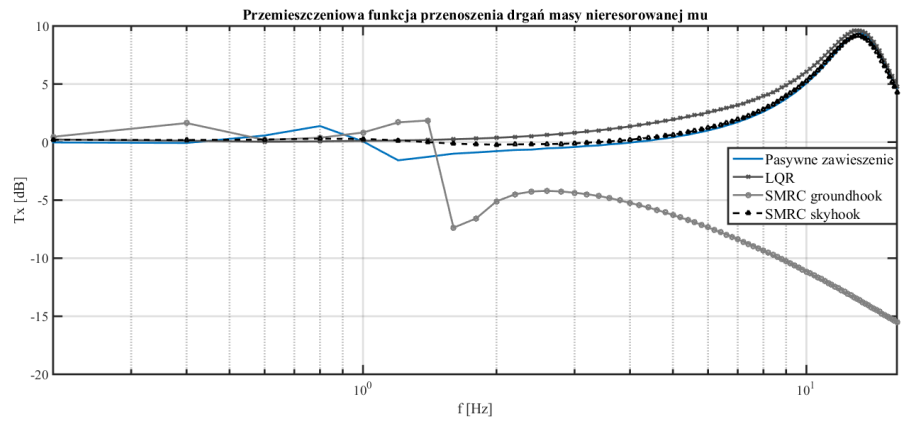
m_u	30 [kg]
m_s	470 [kg]
k_s	$\alpha * 100[\frac{kN}{m}]$
k_t	$200[\frac{kN}{m}]$
c_t	$202\ [\frac{Ns}{m}]$
c_s	$\alpha * 3250\ [\frac{Ns}{m}]$
c_{sky}	$3000\ \left[\frac{Ns}{m}\right]$
c_{ground}	$2000\ [\frac{Ns}{m}]$
$\alpha = 0.186$	

Źródło: opracowanie własne

Przeprowadzono badania symulacyjne działania poszczególnych algorytmów. Jako wymuszenie przyjęto sygnał sinusoidalny o amplitudzie 0.03 [m] oraz częstotliwości z zakresu $< 0.25, 16 >$ [Hz]. Na wykresie poniżej (Rysunek 7) możemy zaobserwować pierwsze przyjęte kryterium, odpowiadającego za komfort jazdy. Przemieszczeniowa funkcja przenoszenia drgań została wybrana jako sposób weryfikacji sytemu ze względu na jego nieliniowy charakter [9]. Na osi rzędnych wykresu (Rysunek 7) przedstawiono wzmocnienie w dB. Model pasywny w rozpatrywanym przedziale częstotliwości osiąga mniejsze wartości współczynnika przenoszenia drgań tylko od modelu SMRC ground-hook. W okolicach pierwszej częstotliwości rezonansowej osiąga on największe wzmocnienie $(15{,}59)$. Następnie plasuje się algorytm sterowania ślizgowego z modelem odniesienia sky-hook $(-1{,}41)$. Należy w tym miejscu podkreślić, że działanie dwóch wymienionych algorytmów w okolicy drugiej częstotliwości rezonansowej nie powoduje zwiększenia wartości współczynnika przenoszenia drgań. Ponadto dla układu sterowanego algorytmem LQR współczynnik przenoszenia drgań przyjmuje wartości najniższe w całym przedziale częstotliwości. Zastosowanie regulatora liniowo-kwadratowego zapewnia redukcję drgań w całym paśmie rozpatrywanych częstotliwości.

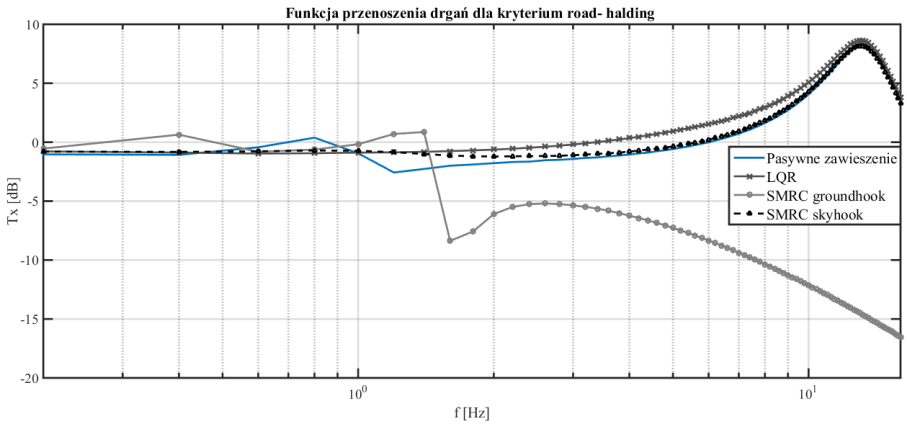


Rysunek 7. Przemieszczeniowa funkcja przenoszenia drgań m_s – porównanie realizacji poszczególnych algorytmów [opracowanie własne]



Rysunek 8. Przemieszczeniowa funkcja przenoszenia drgań dla masy nieresorowanej [opracowanie własne]

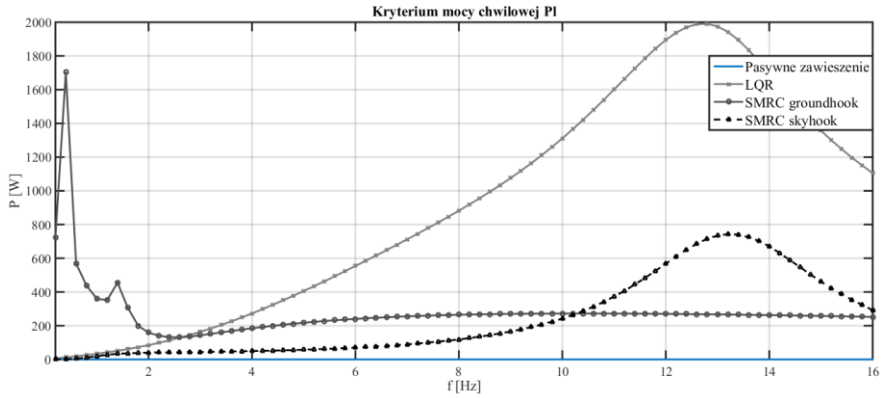
Przemieszczeniową funkcję przenoszenia drgań z wymuszenia z_0 na masę niewibroizolowaną m_u odpowiadającą za bezpieczeństwo w trakcie jazdy przedstawiono na (Rysunek 8). Funkcje te przedstawiono dla wszystkich wymienionych wcześniej algorytmów sterowania. W okolicy pierwszego rezonansu obserwujemy nieznaczne wzmocnienie dla układu redukcji drgań sterowanego algorytmem SMRC ground-hook (1,86) oraz dla pasywnego modelu zawieszenia (1,38). Układy sterowane algorytmami SMRC sky-hook oraz LQR w okolicy drugiego rezonansu osiągają największą wartość wzmocnienia (9,16). Największą redukcję drgań w okolicy drugiej częstotliwości drgań własnych uzyskano dla układu sterowanego algorytmem SMRC ground-hook ($-13{,}64$).



Rysunek 9. Realizacja kryterium road-holding [opracowanie własne]

Ze względu na fakt, że rozpatrywana jest składowa, która odpowiada za oddziaływania dynamiczne między pojazdem a drogą podobnie jak w poprzednich przypadkach należy minimalizować wskaźnik T_x. W przypadku kryterium road-holding (Rysunek 9) kształt charakterystyki jest bardzo podobny do kryterium przemieszczeniowej funkcji przenoszenia drgań masy nieresorowanej m_u. Wartości współczynnika przenoszenia drgań są przesunięte o ok -1 dB w całym rozważanym paśmie częstotliwości.

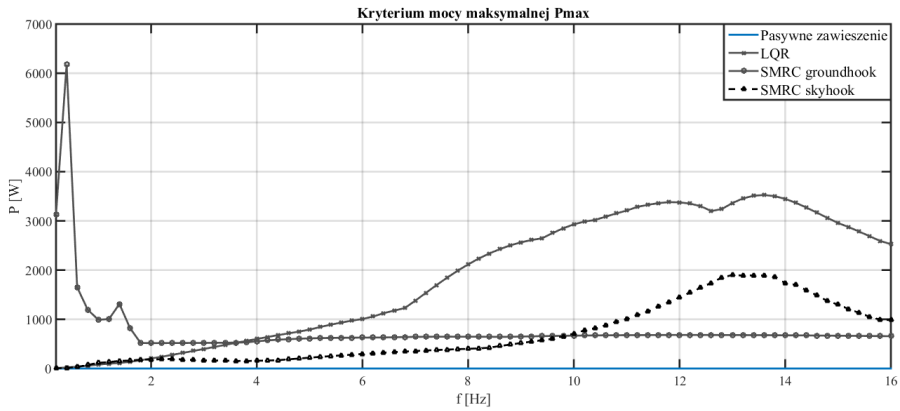
Poniżej przedstawiono wyniki drugiej części przeprowadzonych symulacji. Podobnie jak w pierwszym przypadku przeprowadzono badania symulacyjne działania poszczególnych algorytmów. Jako wymuszenie przyjęto sygnał sinusoidalny o amplitudzie 0.03 [m] oraz częstotliwości z zakresu $< 0.25, 16 >$ [Hz]. Na wykresie poniżej (Rysunek 10) możemy zaobserwować realizację kryterium mocy chwilowej, odpowiadającego za zużycie energii. Na osi rzędnych wykresu przedstawiono moc w jednostce Watt. Rozpatrywanie tego kryterium pozwala na oszacowanie jak dużo energii zużyje układ redukcji drgań dla wymuszenia sinusoidalnego o zadanej częstotliwości. Należy w tym miejscu podkreślić, że każdy punkt na dwóch zamieszczonych poniżej charakterystykach reprezentuje zużycie energii przez zawieszenie dla danej częstotliwości wymuszenia.



Rysunek 10. Realizacja kryterium mocy chwilowej P_I [opracowanie własne]

Algorytm SMRC ground-hook cechuje wysokie zapotrzebowanie na energię zewnętrzną (1704 W) dla niskich częstotliwości. Dla częstotliwości wyższych od 3 Hz zużycie energii stabilizuje się na poziomie ok 200 W. Pozostałe algorytmy dla niskich częstotliwości wykazują najniższe zapotrzebowanie na energię. Algorytm SMRC sky-hook w okolicy drugiej częstotliwości rezonansowej pobiera średnio 740 W. Natomiast regulator liniowo kwadratowy pobiera powyżej 1000W w szerokim paśmie częstotliwości pomiędzy 8,6 a 16 Hz.

Wykres (Rysunek 11) przedstawia kryterium mocy maksymalnej, odpowiadające za maksymalne zużycie energii. Rozpatrywanie tego kryterium pozwala na oszacowanie ile będzie maksymalnie potrzebne energii aby układ redukcji drgań działał zgodnie z założeniami. Innymi słowy jak duży będzie potrzebny akumulator. Kształt uzyskanej charakterystyki jest podobny do charakterystyki mocy chwilowej. Natomiast rząd wielkości jest znacznie większy. Algorytm SMRC ground-hook cechuje bardzo wysokie zapotrzebowanie na energię zewnętrzną (6178 W) dla niskich częstotliwości. Dla częstotliwości wyższych od 2 Hz zużycie energii stabilizuje się na poziomie ok 600 W. Pozostałe algorytmy dla niskich częstotliwości wykazują najniższe zapotrzebowanie na energię. Algorytm SMRC sky-hook w okolicy drugiej częstotliwości rezonansowej pobiera średnio 1888 W. Natomiast regulator liniowo kwadratowy pobiera powyżej 1000 W w szerokim paśmie częstotliwości pomiędzy 6 a 16 Hz, ale nie przekracza 3500 W.



Rysunek 11. Realizacja kryterium mocy maksymalnej P_{max} [opracowanie własne]

Niezależnie od wybranego kryterium można zaobserwować, że regulatory których głównym zadaniem jest uzyskanie największego komfortu uzyskują najniższe wartości funkcji przenoszenia drgań dla niskich częstotliwości. Natomiast dla wysokich częstotliwości uzyskują najwyższe wartości funkcji przenoszenia drgań z wszystkich badanych przypadków. Dla wyższych częstotliwości działanie układu wibroizolacji odbywa się kosztem zużycia większej ilości energii. Realizacja metod nastawionych na bezpieczeństwo jazdy wykazuje odwrotne właściwości: dla niskich częstotliwości wykazują najwyższe wartości funkcji przenoszenia drgań, a dla wysokich częstotliwości uzyskują najmniejsze wartości tej funkcji. Dla niższych częstotliwości działanie układu wibroizolacji odbywa się kosztem zużycia większej ilości energii.

7. Podsumowanie

Układy aktywne stosowane w zawieszeniach pojazdów zapewniają znacznie lepszą redukcję drgań niż tradycyjne zawieszenia bazujące na elementach pasywnych. W układach aktywnych odpowiednio sterowane zewnętrzne źródło mocy może dostarczać lub absorbować energię w określony sposób [5]. Najczęściej stosowane są układy ze sprzężeniem zwrotnym, w którym umieszczony jest odpowiedni przetwornik wielkości fizycznej (przemieszczenia, prędkości, przyśpieszenia, ciśnienia itp.) oraz regulator, wzmacniacz i element wykonawczy. Element wykonawczy wytwarza siłę, która równoważy siły wymuszające drgania. W praktycznych zastosowaniach aktywnych układów zawieszeń jako element wykonawczy stosuje się najczęściej siłowniki elektryczne, hydrauliczne lub pneumatyczne. Elementy aktywne stosowane w zawieszeniach pojazdów wymagają dostarczenia zewnętrznej energii niezbędnej do prawidłowej pracy układu. Jak już wspomniano wcześniej, duże zapotrzebowanie układu na energię jest jednym z głównych problemów dotyczących projektowania i eksploatacji aktywnych zawieszeń pojazdów.

Charakterystyczną cechą układów o dwóch stopniach swobody są dwie częstotliwości własne. Dla zawieszeń pojazdów kołowych zawierają się one w paśmie $< \mathbf{0.1, 16} > [\mathbf{Hz}]$. W celu poprawienia komfortu oraz bezpieczeństwa jazdy należy dążyć do jak największej redukcji drgań w tym paśmie częstotliwości przy

jednoczesnym zachowaniu jak największego ugięcia opony (docisku koła do nawierzchni). Jak zostało pokazane zawieszenie sterowane algorytmem SMRC sky-hook dobrze redukuje drgania w okolicach pierwszej częstotliwości własnej. W okolicach drugiej częstotliwości można zauważyć nieznaczne wzmocnienie. Tam zdecydowanie lepiej redukuje drgania algorytm sterowania SMRC ground-hook. Jeżeli chcemy uzyskać tłumienie drgań w całym paśmie przenoszenia najlepiej zastosować sterowanie LQR.

Najlepszą przyczepność (road holding) opony można uzyskać implementując algorytm SMRC ground-hook. Jeżeli chodzi o aspekt energetyczny rozwiązanie to cechuje konieczność zastosowania największej pojemności akumulatora spośród rozważanych algorytmów. Najlepszy kompromis pomiędzy komfortem a bezpieczeństwem zapewnia zastosowanie algorytmu LQR. Realizacja tego sterowania charakteryzuje się największym zapotrzebowaniem na moc chwilową.

W rozdziale skupiono się na przedstawieniu algorytmu sterowania ślizgowego (SMC) opartego o model referencyjny algorytmu typu sky-hook i ground-hook oraz regulator liniowo kwadratowy w porównaniu z modelem pasywnym. Dla sterowania aktywnego osiągane są lepsze wyniki w całym paśmie niż dla modelu zawieszenia pasywnego. Wynika to ze specyfiki sterowania aktywnego: siła generowana na elemencie aktywnym może przyjąć dowolny kierunek i dodatkowo nie narzucono żadnego ograniczenia na wartość siły generowanej przez element wykonawczy. W przypadku sterowania semi-aktywnego mamy do czynienia z ograniczeniem zwrotu siły względem prędkości.

Planuje się przeprowadzenie implementacji algorytmu sterowania ślizgowego mieszanego z modelem referencyjnym typu sky-ground-hook.

Podziękowania

Praca została sfinansowana ze środków z Grantu dziekańskiego (grant nr. 15.11130.183).

Literatura

1. Dupont P., Stokes A., Street C. *Semi-active Control*, Journal of Sound and Vibration, (1995), s. 203-218.
2. Edwards C., Spurgeon, S. K., *Sliding mode control : theory and applications*, Taylor & Francis, (1998).
3. Karnopp D., Crosby M. J., Harwood R. A. *Vibration Control Using Semi-Active Force Generators*, ASME. J. Eng. Ind., 96(May 1974), 619-626. http://doi.org/10.1115/1.3438373
4. Karnopp D. *Vehicle dynamics, stability, and control*, CRC Press, (2013).
5. Kowal, J. *Sterowanie drganiami*, Gutenberg, (1996).
6. Porumamilla H. *Modeling , analysis and non-linear control of a novel pneumatic semi-active vibration isolator : a concept validation study*, Retrospective Theses and Dissertations. Paper 15765, (2007).
7. Rajamani R. *Vehicle Dynamics and Control*, Springer, (2006), http://doi.org/10.1007/978-1-4614-1433-9_2
8. Savaresi S., Poussot-Vassal C., Spelta C. *Semi-Active Suspension Control Design for Vehicles*, Butterworth-Heinemann , (2010).

9. Sibielak M., Raczka W., Konieczny J., Kowal, J. *Optimal control based on a modified quadratic performance index for systems disturbed by sinusoidal signals,* Mechanical Systems and Signal Processing *64-65,* (2015), s.498-519.
10. Valášek M., Kortüm W., Šika Z., Magdolen L., Vaculín O. *Development of semi-active road-friendly truck suspensions,* Control Engineering Practice, 6(6), (1998), s.735-744.
11. Yokoyama M., Hedrick J. K., Toyama S. *A model following sliding mode controller for semi-active suspension systems with MR dampers,* Proceedings of the 2001 American Control Conference. (Cat. No.01CH37148), *4,* (2001), s.2652-2657.

Sterowanie aktywnym zawieszeniem pojazdu z uwzględnieniem regulatora liniowo-kwadratowego oraz algorytmu ślizgowego

Abstrakt
W niniejszym rozdziale wprowadzono trzy różne algorytmy sterowania: sterowanie ślizgowe z wykorzystaniem modelu referencyjnego (MRSMC) sky-hook, MRSMC ground-hook i regulator liniowo-kwadratowy (LQR) Zbadano wpływ zmiany rodzaju sterowania na efektywność układu wibroizolacji modelu ćwiartkowego zawieszenia. Każdy algorytm był testowany pod kątem trzech kryteriów: przemieszczeniowa funkcja przenoszenia drgań, kryterium ugięcia opony (‘tire deflection’) oraz zapotrzebowanie na energię zewnętrzną. Wygenerowano i przeanalizowano charakterystyki częstotliwościowe zawieszenia sterowanego dla wybranych wartości tłumienia c_{sky}, c_{ground}. Porównano wpływ zmian parametrów zawieszenia na przyjęte kryteria oceny wszystkich rozważanych regulatorów. Realizacja algorytmu sterowania ślizgowego (SMC) przy użyciu modelu odniesienia pozwala na uzyskanie porównywalnych charakterystyk częstotliwościowych do tych uzyskiwanych przez układy sterowane algorytmami sky-hook lub ground-hook. Co więcej użyty algorytm sterowania pozwala na obliczenie wartości generowanej siły przez element aktywny. Wygenerowana siła pozwala na osiągnięcie przez obiekt trajektorii fazowej wyznaczonej przez zastosowany model referencyjny. Trzecim rozważanym algorytmem jest sterowanie LQR. Regulator liniowo-kwadratowy został wyznaczony dla modelu zlinearyzowanego a badania zostały przeprowadzone dla modelu nieliniowego. Zaproponowany wskaźnik jakości uwzględnia zarówno poprawę bezpieczeństwa (‘road-holding’) oraz komfort (przemieszczeniowa funkcja przenoszenia drgań). Pozwala on również na ograniczenie mocy sygnału sterującego. Ponadto wszystkie parametry modelu zawieszenia pojazdu odpowiadają parametrom samochodu typu SUV Isuzu D-MAX.
Słowa kluczowe: sterowanie ślizgowe, aktywne zawieszenie pojazdu, model referencyjne, LQR

Control of the active suspension of the vehicle including the linear-quadratic regulator and the sliding mode algorithm

Abstract
Three different control algorithms have been introduced in this chapter: sky-hook, MRSMC ground-hook, and linear-quadratic (LQR) reference control (MRSMC). The influence of change of control type on the effectiveness of the quarter car suspension system. Each algorithm was tested for three criteria: displacement transmissibility function, tire deflection criteria, and external energy demand. Generated and analyzed the frequency characteristics of the controlled suspension for the selected attenuation values c_sky, c_ground. The impact of changes in suspension parameters on the assumed evaluation criteria for all considered regulators was compared. The implementation of the sliding mode control algorithm (SMC) using the reference model allows for comparable frequency characteristics to those obtained by sky-hook or ground-hook algorithms. What's more, the control algorithm used allows you to calculate the value of the force generated by the active element. The generated force allows the object to achieve the phase trajectory defined by used reference model. The third considered algorithm is LQR control. The linear-quadratic regulator was defined for a linearized model and the tests were performed for a nonlinear model. The proposed quality index takes into account both safety and comfort (displacement transmissibility function). It also allows you to limit the power of the control signal. In addition, all the parameters of the suspension model correspond to the parameters of the Isuzu D-MAX SUV.
Keywords: sliding mode control, active suspension, model reference control, LQR

Marta Paczkowska[1]

Zastosowanie technologii laserowych do zwiększania odporności na zużycie elementów maszyn

1. Wprowadzenie

Części maszyn i urządzeń podczas eksploatacji ulegają ciągłemu zużywaniu w wyniku tarcia (a niejednokrotnie również w wyniku oddziaływania środowiska korozyjnego). Intensywność tego procesu zależy od właściwości materiału z którego wykonana jest dana część. W szczególności dotyczy to właściwości jego warstwy wierzchniej.

W celu zapobiegania, bądź (w większości przypadków) opóźniania, procesu zużywania przeprowadza się różnego rodzaju modyfikacje warstwy wierzchniej. Do najbardziej popularnych należą obróbki cieplno-chemicznej, jak np. nawęglanie czy azotowane. Obróbki te, polegają na nasycaniu danym pierwiastkiem warstwy wierzchniej całej części maszyny. Pozwalają one efektywnie zwiększyć twardość warstwy oraz odporność na zużycie.

Niemniej jednak bardzo często zużycie tribologiczne dotyczy jedynie niewielkich fragmentów danej części. Oznacza to, że obróbka całej części nie jest konieczna.

Alternatywą do metod dyfuzyjnych mogą być obróbki powierzchniowe, które pozwalają na modyfikację jedynie fragmentów części maszyn – tych najbardziej narażonych na zużycie. Do jednej z takich obróbek należy laserowa obróbka cieplna (LOC).

Za pomocą selektywnego nagrzewania wiązką laserową możliwe jest uszlachetnienie tylko wybranych fragmentów warstwy wierzchniej. Dzięki takiej obróbce możliwe jest zwiększenie odporności na zużycie również elementów maszyn wykonanych ze stosunkowo tanich i popularnych materiałów do których należą żeliwa. Atutem LOC jest np. oszczędność zazwyczaj drogich pierwiastków stopowych, brak konieczności stosowania chłodziw, brak konieczności zagospodarowywania substancji szkodliwych, które czasami stasowane są w przypadku obróbek dyfuzyjnych. Po różnych próbach zużyciowych badania wykazały, że ubytek masy elementów zmodyfikowanych laserowo może być znacznie zredukowany. Ponadto, w niektórych przypadkach może być również zredukowany ubytek masy współpracujących z nimi elementów. Celem tego opracowania jest ocena możliwości jakie stwarza obróbka laserowa elementów maszyn przede wszystkim tych narażonych na zużycie tribologiczne.

[1] marta.paczkowska@put.poznan.pl, Instytut Maszyn Roboczych i Pojazdów Samochodowych, Wydział Maszyn Roboczych i Transportu, Politechnika Poznańska

2. Przykłady elementów maszyn narażonych na zużycie

Niektóre części elementów przeznaczonych dla przemysłu motoryzacyjnego narażone są na różnego rodzaju zużycia np. przez sczepianie I rodzaju (adhezyjne) czy utlenianie [1].

Wymienić tutaj można wały korbowe wykonywane ze stali lub żeliwa. Na zużycie przez tarcie szczególnie narażone są ich powierzchnie robocze, czyli współpracujące z panewkami czopy: główne i korbowodowe.

W przypadku wałków rozrządu, również wykonywanych ze stali bądź żeliwa, oprócz ich czopów na zużycie narażone również są krzywki (w szczególności ich noski), zazwyczaj stanowiące jedną całość z wałem rozrządu [2]. Warto dodać, że współpracujące z krzywką denko popychacza zaworu (wykonywane także ze stali lub żeliwa) powinno również mieć zapewnione odpowiednie właściwości warstwy wierzchniej. Zwiększoną odpornością na zużycie przez tarcie powinny cechować się również bieguny i powierzchnie czasz dźwigni zaworów (wykonywanie ze stali, staliwa lub żeliwa).

Z kolei, od takich powierzchni jak gładzie tulei cylindrowych wymaga się (poza odpornością na zużycie prze tarcie) dodatkowo odporności na oddziaływanie zwiększonych wartości temperatury. Podobnych właściwości wymaga się od pierścieni tłokowych oraz powierzchni roboczych tłoków (zwłaszcza rowków na pierścienie) [3].

Na zużycie przez tarcie oraz działanie agresywnych atmosfer w podwyższonych temperaturach narażone są gniazda zaworów w głowicach [3, 4] a także części wewnętrzne prowadnicy współpracującej z trzonkiem zaworu.

Natomiast, powierzchnie robocze takich elementów jak tarcze i bębny hamulcowe narażone są na zużycie ścierno-korozyjne oraz tzw. szoki temperaturowe.

Na zużycie tribologiczne narażone są również powierzchnie robocze takich części jak: koła zębate, wały wykorbione, tulejki, części hydrauliczne.

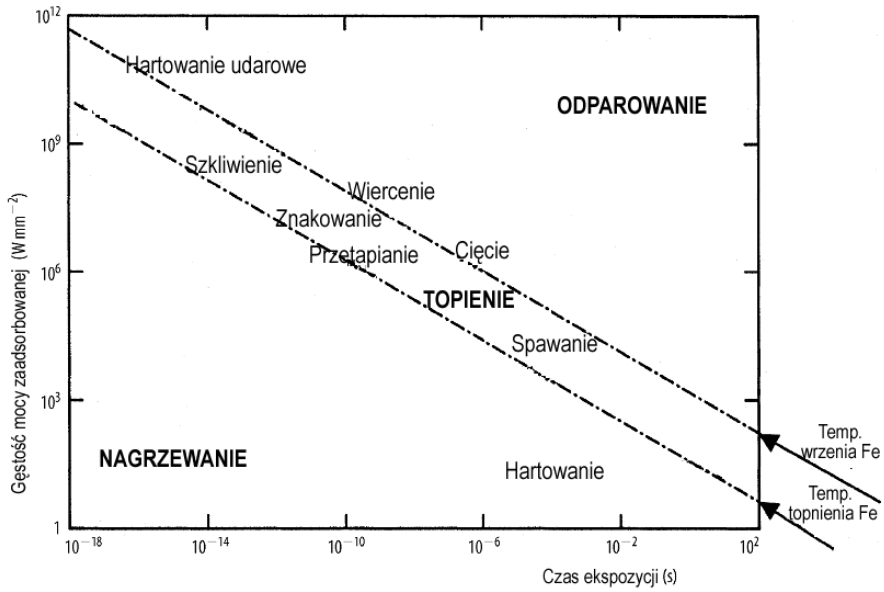
Wysoką odpornością na zużycie powinny charakteryzować się narzędzia do obróbki plastycznej np. walce do walcownia blach czy matryce do tłoczenia karoserii [5].

Wśród części maszyn narażonych na zużycie przez tarcie należy również wymienić części maszyn rolniczych jak wały maszyn żniwnych, zęby bron, brony talerzowe, stopki redlic.

We wszystkich wymienionych przypadkach fragmentów elementów maszyn wymaga się dużej twardości oraz odporności na zużycie ich warstw wierzchnich, a czasami nawet zwiększonej odporności na korozję czy też żaroodporności i żarowytrzymałości. Jedną z metod zapewniających odpowiednie właściwości może być obróbka laserowa.

3. Technologie laserowe i możliwości wykorzystania wiązki laserowej w inżynierii powierzchni

Technologie laserowe, odkąd zaczęły się rozwijać, znalazły szerokie zastosowanie w różnych dziedzinach życia, w tym również w inżynierii powierzchni. Na rysunku 1 przedstawiono różne grupy technologii laserowych stosowanych w inżynierii powierzchni.



Rysunek 1. Zależność procesów cieplnych zachodzących przy nagrzewaniu wiązką laserową od gęstości mocy wiązki i czasu ekspozycji. Linie ukośne reprezentują punkt wrzenia i topnienia żelaza wg [6]

Obróbka laserowa, a dokładniej laserowa obróbka cieplna należy do grupy nowoczesnych metod uszlachetniania powierzchni o wysokiej wydajności [7]. Jej specyfika związana jest z bardzo szybkimi przemianami zachodzących w materiale. Rezultat wywołany w nagrzewanym wiązką materiale zależy od szeregu różnych czynników, ale przede wszystkim od właściwości samego obrabianego materiału, od wytworzonej na jego powierzchni temperatury, czasu nagrzewania i chłodzenia. W związku z tym, kluczowe znaczenie mają takie czynniki jak: rodzaj lasera, parametry wiązki laserowej (w tym przede wszystkim moc i średnica wiązki oraz kształt rozkładu mocy na przekroju wiązki), czas działania wiązki na materiał (a więc np. prędkość przesuwu wiązki względem obrabianego materiału w przypadku lasera o pracy ciągłej lub czas trwania impulsów w przypadku lasera impulsowego) [8]. Temperatura warstw wierzchnich nagrzewanego materiału zależy przede wszystkim od ilości dostarczanej energii, która jest proporcjonalna do gęstości mocy wiązki laserowej oraz czasu jej działania. W konsekwencji, sterując tymi dwoma parametrami (rys. 1) można uzyskać różne wartości temperatury w warstwie wierzchniej, co umożliwia np.: powodowanie przemian fazowych w stanie stałym (np. hartowanie), przetapianie, bądź odparowanie [8].

Specyficzne warunki tworzenia mikrostruktury powodują, że za pomocą LOC pozwala otrzymywać takie właściwości, które są nieosiągalne dla tradycyjnych metod obróbki materiałów. Warto również zaznaczyć, że urządzenie laserowe umożliwia obróbkę powierzchni o dowolnie skomplikowanym kształcie zachowując pierwotne wymiary elementów obrabianych. Możliwe jest precyzyjne sterowanie dawką energii, a cały proces może zostać łatwo zautomatyzowany [9]. Obróbka laserowa zapewnia ponadto bezkontaktowość operacji gwarantując czystość miejsca obróbki co umożliwia

zdalne operowanie wiązką w warunkach próżni, atmosfer gazowych, czy nawet pod wodą [10]. Natomiast połączenie urządzenia laserowego np. z tokarką pozwala na zastosowanie obróbki hybrydowej podczas której można nagrzewać laserowo i toczyć twarde materiały ceramiczne [11].

Warto także zaznaczyć, że np. podczas tradycyjnej obróbki cieplno-chemicznej konieczne jest nieraz zabezpieczanie pozostałych części elementów (nie przeznczonych do obróbki) przed nasycaniem [3]. Z kolei za pomocą obróbki laserowej stopuje się tylko wybrane powierzchnie części maszyn i narzędzi narażone na szczególnie trudne warunki eksploatacyjne [12]. Możliwość obróbki jedynie fragmentów części maszyn za pomocą LOC (w porównaniu z obróbkami dyfuzyjnymi) pozwala na oszczędność stosowanych w procesie, niejednokrotnie drogich, pierwiastków stopowych. Nie ma konieczności stosowania chłodziw, ponieważ chłodzenie następuje automatycznie przez powierzchnię i rdzeń obrabianego elementu, czy też zagospodarowywania szkodliwych substancji, które czasami stasowane są w przypadku obróbek dyfuzyjnych.

Do tzw. technologii bezprzetopieniowych zaliczyć można wyżarzanie, hartowanie bezprzetopieniowe i odpuszczenie. Natomiast w zakres technologii tzw. przeto-pieniowych wchodzi nadtapianie do którego należy: hartowanie przetopieniowe, szkliwienie, zagęszczenie, wygładzanie; stopowanie do którego należy przetapianie, wtapianie; natapianie. Z kolei do technologii odparowaniowych zaliczyć można odparowanie czyste i utwardzanie detonacyjne. Wymienione technologie przede wszystkim stosowane są dla stali (węglowych, jak i stopowych) ale również żeliw oraz innych metali i ich stopów np.: miedzi, aluminium, tytanu [7].

4. Efekty modyfikacji warstw wierzchnich za pomocą laserowej obróbki cieplnej

Spośród wymienionych powyżej rodzajów laserowej obróbki cieplnej w celu zwiększania odporności na zużycie tribologiczne fragmentów części maszyn i urządzeń stosuje się zazwyczaj hartowanie (z przetopieniem lub bez przetopienia warstwy wierzchniej) bądź też stopowanie (przetopienie warstwy wierzchniej wraz z wprowadzeniem składnika stopowego), które zmieniając mikrostrukturę warstwy wierzchniej w większości przypadków istotnie (np. kilkukrotnie) zwiększają jej twardość w porównaniu do twardości rdzenia.

Znaczne rozdrobnienie mikrostruktury (i jej zahartowanie) w wyniku laserowego hartowania bez przetopienia pozwala zwiększyć twardość stali o ok. 30% w porównaniu do hartowania konwencjonalnego [7]. Odnotowano, że towarzyszy temu również zwiększenie umownej granicy plastyczności, doraźnej wytrzymałości na rozciąganie i odporności na zużycie. Na ogół zwiększa się również wytrzymałość zmęczeniowa [7]. Badania własne [13] wykazały, że hartowanie laserowe bez przetopienia warstwy wierzchniej żeliwa sferoidalnego powodujące uzyskanie mikrostruktury składającej się z martenzytu, ferrytu oraz nierozpuszczonych podczas obróbki kulek grafitu o twardości osnowy ok. 600HV0,1 (czyli ok. 2-krotnie większej od twardości osnowy rdzenia) pozwoliło zredukować blisko 6-krotnie ubytek masy tak obrobionego elementu po przeprowadzonej próbie zużycia. Natomiast głębokość w śladzie wytarcia była mniejsza blisko 4-krotnie w porównaniu do takiej głębokości uzyskanej w elemencie nie poddanym obróbce laserowej. Warto zaznaczyć, że

przeprowadzona próba tarciowa trwała jedynie 15 minut. Zaletą hartowania laserowego bez przetopienia jest brak zmian w topografii powierzchni po takiej modyfikacji. Oznacza to, że nie ma potrzeby przeprowadzania mechanicznej obróbki wykańczającej danego elementu.

Warto również zaznaczyć, że efekt zahartowania przy laserowej obróbce bezprzetopieniowej uzyskać w przypadku żeliw można już przy oddziaływaniu stosunkowo niewielkiej gęstości mocy wiązki laserowej np. 20 W/mm^2 (przy czasie jednostkowym jej oddziaływania na powierzchnię równym 1 s) [14] w porównaniu do obróbek przetopieniowych.

Z kolei, zastosowanie większych gęstości mocy wiązki laserowej wywołuje w warstwie wierzchniej przetopienie. Mikrostruktura powstała w wyniku szybkiego przetopienia, skrzepnięcia (i zazwyczaj zahartowania) jest bardziej jednorodna chemicznie i strukturalnie oraz drobnoziarnista w porównaniu do materiału rdzenia. Zazwyczaj pod obszarem przetopienia znajduje się również obszar zahartowany ze stanu stałego który dodatkowo wzmacnia tak zmodyfikowaną warstwę wierzchnią. Twardość warstwy wierzchniej stali może zwiększyć się 1,5-4-krotnie (w skrajnych sytuacjach nawet 10-krotnie) [7] (zależy to od rodzaju stali oraz warunków laserowej obróbki cieplnej). Niemniej jednak, należy zaznaczyć (jak odnotowano w [7]) twardość obszaru przetopionego może być czasami mniejsza niż występującego pod nim obszaru zahartowanego ze stanu stałego.

W wyniku przetapiania laserowego możliwe jest np. 1,5-krotne zwiększenie w przypadku stali szybkotnących odporności na zużycie. Odnotowano również zwiększenie o 250% trwałości niektórych stali narzędziowych do pracy na gorąco [7]. Efekt zwiększenia odporności na zużycie po przetapianiu laserowym pozwoliły stwierdzić również przeprowadzone badania własne na żeliwie [13]. Uzyskanie warstwy wierzchniej o drobnoziarnistej mikrostrukturze o charakterze zbliżonym do zahartowanego ledeburytu przemienionego (z częściowo rozpuszczonymi kulkami grafitu) poskutkowało twardością 1150HV0,1, czyli ok. blisko 4-krotnie większą od twardości osnowy materiału rdzenia. Modyfikacja ta spowodowała, że ubytek masy po badaniach tribologicznych tak obrobionych elementów był 8-krotnie mniejszy od ubytku elementów nieobrobionych, a odnotowana głębokość śladu wytarcia była ponad 4-krotnie mniejsza od głębokości takiego śladu w przypadku elementów nieobrobionych. Warto również dodać, że obróbka laserowa z przetopieniem żeliwa pozwoliła w tym przypadku nie tylko na wyraźne zmniejszenie ubytku masy samych obrobionych w ten sposób żeliw, ale i na 1,5-krotne zmniejszenie ubytku masy współpracujących z nimi pierścieni wykonanych z zahartowanej stali C45 o twardości ok. 550HV0,1 [15]. Jest to o tyle warte zaznaczenia, że zazwyczaj umocnione warstwy wierzchnie za pomocą różnych obróbek powierzchniowych charakteryzują się mniejszym zużywaniem przez tarcie, ale nie zawsze przekłada się to na zmniejszenie ubytku masy elementów z nimi współpracujących (a nawet powoduje to przyspieszone ich zużywanie). Efekt ten bardzo często stanowi ograniczenie stosowania metod obróbek powierzchni.

Zwiększenie odporności na zużycie ścierne w wyniku przetapiania laserowego pozwoliły stwardzić badania własne dotyczące obróbki elementów maszyn rolniczych, a mianowicie stopek redlic siewnika wykonywanych z szarego żeliwa płatkowego

[16]. Badania polegały na przeprowadzaniu obróbki laserowej z przetopieniem jedynie krawędzi stopki (uzyskując zmodyfikowaną warstwę wierzchnią o obszarze ok. 160 mm^2) tylko z jej jednej strony. Dla porównania przeprowadzono również inną obróbkę – natryskiwanie cieplne, które skutkowało wytworzeniem powłoki ze stopu Ni-Cr-B-Si-WC obejmującego obszar prawie 1/3 stopki redlicy z obu jej stron. Próba zużyciowa w warunkach tarcia ściernego w gruncie pozwoliła 6-krotnie zmniejszyć ubytek masy stopek po obróbce laserowej (w porównaniu do ubytku nieobrobionych). Podobny wynik uzyskano dla stopek po natryskiwaniu pomimo, że twardość ich powłoki wyniosła ok. 1500HV0,1 (i, jak wspomniano, obejmowała znacznie większą powierzchnię stopek niż po LOC) a po przetapianiu laserowym jedynie 900HV0,1. Do zwiększenia odporności na zużycie po LOC przyczyniła się jednorodna, drobnoziarnista mikrostruktura warstw wierzchniej (po natryskiwaniu uzyskana powłoka charakteryzowała się gruboziarnistą mikrostrukturą z wydzieleniami węglika WC o ostrych krawędziach). Oznacza to, że zapewniając drobnoziarnistą mikrostrukturę o dużym stopniu homogenizacji warstwy wierzchniej za pomocą przetapiania laserowego można efektywnie podnieść również odporność na zużycie ścierne.

Dodatkowo twardość, odporność na zużycie przez tarcie, korozję a także żaroodporność przetapianych laserowo warstw wierzchnich można zwiększyć przez wprowadzenie podczas tej obróbki substancji stosującej. Stop powstający w warstwie wierzchniej w wyniku jednoczesnego przetopienia dwóch materiałów, a mianowicie substancji stopującej i niewielkiej części materiału podłoża, po ich wymieszaniu i szybkim skrzepnięciu (co zapewnia laserowa obróbka cieplna) charakteryzuje się nowym składem chemicznym, mikrostrukturą i własnościami. Materiał ten jest inny niż materiał podłoża i materiał stopujący [17]. Substancję stopująca może stanowić jeden pierwiastek lub kilka, a także związki. Stosowane są również różne kombinacje pierwiastków lub związków.

Szybkie schłodzenia stopu, które towarzyszy laserowej obróbce cieplnej, umożliwia uzyskanie mikrostruktury drobnokrystalicznej, silnie przesyconych roztworów i faz metastabilnych, a także zupełnie nowych faz powstających w tych szczególnych warunkach nierównowagowych. Możliwe jest wytworzenie nanostruktury, a nawet formy amorficznej. Pozwala to na osiągnięcie zupełnie nowych efektów w mikrostrukturze, a co za tym idzie właściwościach (w porównaniu np. do tradycyjnych obróbek cieplno-chemicznych).

Uzyskane umocnienie warstwy przetopionej wynika z rozpuszczenia obcych atomów, efektów przemian fazowych oraz rozdrobnienia ziarn [18-20]). W zależności od zastosowanej substancji stopującej za pomocą takiej obróbki można uzyskiwać warstwy o mikrostrukturze i właściwościach stopów nierdzewiejących lub wysokostopowych [19]. W przypadku stopowania borem stali St2 twardość warstwy wierzchniej można zwiększyć prawie 3,5-krotnie (jak odnotowano w [7]), a w przypadku żeliwa sferoidalnego (jak wykazały badania własne [21]) nawet ponad 5-krotnie (w odniesieniu do twardości osnowy materiału rdzenia). Twardość taka wynika z wytworzenia mikrostruktury warstwy wierzchniej składającej się z borków żelaza w osnowie przesyconego roztworu stałego. W przypadku uzyskania twardości ok. 1300HV0,1 w wyniku stopowania laserowego borem odnotowano aż 17-krotne zmniejszenie ubytku masy tak obrobionych elementów w porównaniu do nierobionych

po przeprowadzonym teście tribologicznym w ramach badań własnych [13]. Elementy umocnione laserowo po próbie tarcia charakteryzowały się ponad 2-krotnie mniejszą szerokością i aż 8-krotnie mniejszą głębokością śladu wytarcia od elementów nieobrobionych. Badania tribologiczne przeprowadzone w ramach badań własnych [22] na stopowanych laserowo borem czopach wałów korbowych również pozwoliły stwierdzić zwiększenie ich odporności na zużycie w stosunku do czopów nieobrobionych [22]. Dodatkowo, analiza charakteru przebiegu procesu zużywania [23] pozwoliła zaobserwować, że przy ustabilizowanej pracy łożyska w przypadku zastosowania czopu obrobionego laserowego można się spodziewać większej grubości warstwy oleju na styku czop-panewka (niż w przypadku pracy czopu nieobrobionego), a także stwierdzono, że przy dłuższym czasie pracy takiej pary można spodziewać się mniejszej wartości temperatury oleju w tym miejscu co powinno sprzyjać zwiększeniu odporności na zużywanie przez tarcie badanego węzła kinematycznego. Ponadto, dzięki zastosowaniu stopowania laserowego jak wykazały kolejne badania własne [15] uzyskać również można zmniejszone zużycie elementów współpracujących (podobnie jak w przypadku wspomnianego samego przetapiania laserowego warstwy wierzchniej). Ubytek masy zahartowanych stalowych pierścieni współpracujących z elementami po stopowaniu laserowym po próbie zużyciowej był prawie 3-krotnie mniejszy niż w przypadku tych samych pierścieni współpracujących z elementami nieobrobionymi.

Należy zaznaczyć, że elementy po stopowaniu laserowym (bądź tylko po samym przetapianiu laserowym), które maja współpracować z innym elementem maszyny w węźle kinematycznym wymagają wykańczającej obróbki mechanicznej, ponieważ dość istotnie zmienia się struktura stereometryczna powierzchni. Badania własne wykazały, że np. parametr chropowatości Rz na skutek stopowania laserowego może zwiększyć się nawet blisko 9-krotnie (w zależności od parametrów obróbki laserowej) [21]. Zmniejszenie chropowatości powierzchni powstałej po obróbce laserowej z przetopieniem nie zawsze jest konieczne (np. elementów maszyn pracujących w gruncie). Badania własne [24] dotyczące wpływu stopowania laserowego (w tym przypadku azotkiem krzemu) fragmentu stopek redlic na ich zużycie przeprowadzono bez zastosowania przed próbą zużycia wykańczającej obróbki mechanicznej. Pomimo, tego odnotowano zwiększenie odporności na zużycie tak obrobionych elementów. Badania przeprowadzono jak wspomniane wcześniej, dotyczące samego przetapiania stopek redlic siewnika [16].

Zwiększeniu odporności na zużycie w wyniku tarcia sprzyja uzyskiwana, w większości przypadków przetapiania lub stopowania laserowego, drobnoziarnista i jednorodna mikrostruktura warstwy wierzchniej. Nie zawsze jednak wprowadzane pierwiastki stopowe mające na celu dodatkowo zwiększyć twardość warstwy (lub nadać inne właściwości) skutkują uzyskiwaniu zwiększonej odporności na zużycie. Odnotowano to w badaniach [25] polegających na stopowaniu laserowym chromem żeliwa sferoidalnego. Efektem tego było wytworzenie w warstwie węgliku chromu $(CrFe)_7C_3$, ale niestety po przeprowadzonej próbie zużycia odnotowano wzrost ubytku masy tak obrobionych elementów. Należy również zaznaczyć, że w przypadku stopowania laserowego w szczególności części żeliwnych (dodatkowo stopowania z dużymi prędkościami chłodzenia warstwy wierzchniej generującymi duże gradienty

naprężeń) należy liczyć z możliwością powstawania pęknięć w warstwie. Niemniej jednak, nie zawsze występujące pęknięcia w wyniku obróbki laserowej uniemożliwiają osiągnięcia zwiększenia odporności na zużycie przez tarcie. Na przykład, żeliwo sferoidalne po stopowaniu laserowym kombinacją pierwiastków C-B-W-Cr [26] skutkującym uzyskaniu bogatej w węgliki mikrostruktury warstwy wierzchniej o twardości 1200HV0,05 pomimo stwierdzonych pęknięć charakteryzowało się mniejszym ubytkiem masy po testach zużyciowych niż żeliwo nieobrobione.

5. Podsumowanie

Do zalet modyfikacji laserowej warstwy wierzchniej elementów maszyn należy niewątpliwie zaliczyć:

- możliwość efektywnego zwiększania odporności na zużycie tribologiczne za pomocą obróbki jedynie niewielkich fragmentów warstwy wierzchniej części maszyn przez jej hartowanie ze stałego, przetapianie, czy też stopowanie,
- możliwość uzyskiwania mikrostruktur, a co za tym idzie właściwości, które są nieosiągalne dla tradycyjnych metod obróbki materiałów (dotyczy to możliwości wytwarzania mikrostruktury bardzo drobnokrystalicznej, nawet nanostruktury, silnie przesyconych roztworów, faz metastabilnych, a także zupełnie nowych faz powstających w tych szczególnych warunkach nierównowagowych, a także witryfikacji),
- oszczędność pierwiastków stopowych wynikający z obróbki tylko fragmentów części maszyn,
- brak konieczności stosowania chłodziw, gdyż ciepło odprowadzane jest zaraz przez powierzchnię oraz przez rdzeń obrabianego elementu,
- brak konieczności zagospodarowywania substancji szkodliwych, które czasami stasowane są w przypadku obróbek dyfuzyjnych,
- brak konieczności zastosowania wykańczającej obróbki mechanicznej po modyfikacji laserowej przeprowadzonej w stanie stałym (np. hartowaniu ze stanu stałego), a w przypadku gdy element maszyny nie jest częścią węzła kinematycznego (typu pierścień-tłok, czop wału-panewka) i ma być eksploatowany np. w gruncie jak wiele części maszyn rolniczych (np. stopki redlicy siewnika) również nie ma takiej konieczności po modyfikacji laserowej z przetopieniem,
- możliwość obróbki powierzchni o dowolnie skomplikowanym kształcie zachowując pierwotne wymiary elementów obrabianych,
- możliwość precyzyjnego sterowania dawką energii,
- łatwą automatyzację oraz możliwość połączenia z innymi procesami w celu uzyskania obróbki hybrydowej,
- możliwość zdalnego operowania wiązką w warunkach próżni, atmosfer gazowych czy nawet pod wodą.

Niemniej jednak, należy mieć na uwadze ograniczenia tej metody polegające na:

- możliwości obróbki jedynie fragmentów elementów maszyn, co może stanowić przeszkodę gdy np. celem obróbki byłoby zwiększenie odporności na korozję całego elementu,

- zwiększeniu chropowatości powierzchni po modyfikacji laserowej z przetopieniem,
- konieczności zastosowania wykańczającej obróbki mechanicznej w celu zmniejszenia chropowatości po przetapianiu laserowym w przypadku gdy obrabiany element ma być częścią węzła kinematycznego i współpracować z innym elementem maszyny,
- możliwości występowania pęknięć w materiale warstwy wierzchniej w szczególności z zastosowaniem dużych prędkości chłodzenia podczas obróbki oraz w przypadku modyfikacji takich materiałów jak żeliwa.

Zastosowanie technologii laserowych do części maszyn i urządzeń daje możliwość efektywnego zwiększenia ich odporności na zużycie w wyniku tarcia nie tylko elementu obrobionego, ale w niektórych przypadkach również elementu z nim współpracującego w węźle kinematycznym. Zwiększyć również można odporność na zużycie elementów eksploatowanych w ośrodkach ściernym np. w glebie, jak ma to miejsce w wielu częściach maszyn rolniczych.

Literatura

1. Hebda M., Janecki J., *Tracie, smarowanie i zużycie części maszyn*, WNT, Warszawa (1972).
2. Sert H., Can A., Arican H., Selcuk B., Toprak H., *Wear behaviour of different cam spindles*, Wear 260, (2006), s. 1013-1019.
3. Wajand J.A., Wajand J.T., *Tłokowe silniki spalinowe średnio i szybkoobrotowe*, WNT, Warszawa (2000).
4. Maocai W., Weitao W., *Microstructure of laser-surface-alloyed cast iron with Cr-Al-Y alloy*, Surface and Coatings Technology 75 (1995), s. 181-188.
5. Piaskowski, J., Jankowski, A., *Żeliwo sferoidalne* WNT, Warszawa (1974).
6. Steen W.M., Mazumder J., *Laser material processing*, Fourth edition (2010), s. 1-55.
7. Burakowski T., Wierzchoń T., *Inżynieria powierzchni metali*, WNT, Warszawa (1995).
8. Waligóra W., *Odporność na zmęczenie powierzchniowe stali łożyskowej poddanej obróbce laserowej*, Rozprawy, nr 302, Politechnika Poznańska, (Poznań 1994).
9. Napadałek W., Przetakiewicz W., *Wpływ obróbki laserowej na właściwości wybranych elementów silnika spalinowego*, Inżynieria Materiałowa,5 (2002), s. 547-553.
10. Radziszewska A., Kusiński J., Kąc S., *Laserowe stopowanie stal niskowęglowej węglikiem krzemu*, Hutnik-Wiadomości Hutnicze 5 (2005), s. 232-237.
11. Kawalec M., Barbacki A., Pertek-Owsianna A., Jankowiak M., Nowak I., *Zastosowanie lasera technologicznego CO_2 do doskonalenia właściwości warstwy wierzchniej stali oraz wspomagania toczenia twardej ceramiki konstrukcyjnej Si_3N_4,* Archiwum Technologii Maszyn i Automatyzacji, KBM PAN, vol. 24, 2, (2004), s.139-157.
12. Młynarczak A., Jakubowski J., *Obróbka powierzchniowa i powłoki ochronne*, WPP (1998).
13. Paczkowska M., Wojciechowski Ł., *Modelowe badania zużycia adhezyjnego elementów z żeliwa sferoidalnego borowanego laserowo*, SIMP, Tribologia, 3 (2007), s. 71-81.
14. Paczkowska M., *The evaluation of the influence of laser treatment parameters on the type of thermal effects in the surface layer microstructure of gray irons*, Optics and Laser Technology, 76, (2016), s. 143-148.
15. Paczkowska M., Wojciechowski Ł., *Analiza zużycia stalowych pierścieni współpracujących z żeliwem sferoidalnym poddanym laserowej obróbce cieplnej*, Tribologia, 5 (2015), s. 99-107.

16. Paczkowska M., Selech J, Piasecki A., *Effect of surface treatment on abrasive wear resistance of seeder coulter flap*, Surface Review and Letters, 23, 3 (2015), s. 1650007-1÷1650007-11.
17. Burakowski T., *Lasery technologiczne i ich zastosowanie w inżynierii powierzchni metali*, PM, Zeszyt 11-12 (1993).
18. Kusiński J., *Lasery i ich zastosowanie w inżynierii materiałowej*. Wydawnictwo Naukowe Akapit, Kraków (2000).
19. Kusiński J., *Metalurgiczne podstawy kształtowania struktury i własności warstwy wierzchniej materiałów podczas obróbki laserowej*, V Sympozjum Techniki Laserowej, Szczecin – Świnoujście (1996), s. 275-287.
20. Kusiński J., *Zmiany struktury i własności mechanicznych stali pod wpływem obróbki laserowej*, Metalurgia i odlewnictwo, Zeszyty naukowe 132, Kraków (1989).
21. Paczkowska M., *The Possibility of Selected Surface Layer Modification of Nodular Iron Engine Parts by Laser Boronizing*, Vehicle Engineering (VE), Science and Engineering Publishing Company, 1 (3) September (2013), s. 64-70.
22. Paczkowska M., *Kształtowanie odporności na zużycie tribologiczne elementów maszyn z żeliwa przez laserową obróbkę cieplną (LOC)*, Wydawnictwo PP, Poznań (2016).
23. Paczkowska M, *The analysis of the influence of laser heat treatment of the crankshaft journal on wear resistance of the bearing*, Tribologia 2 (2016), s. 87-100.
24. Selech J., Paczkowska M., Kinal G, Baran B., Chruscinski W., *The comparison of effects of thermal spraying eutalloy 10112 and laser alloying with silicon nitride of cast iron outmost disk coulter*, Journal of Research and Applications in Agricultural Engineering, 60, 1, (2015), s. 88-92.
25. da Costa A.R., Vilar R., *Erosion by solid particle impingement: experimental results with cast-iron, laser-treated surface*, Tribology Letters, 3(4), (1997), s. 379-385.
26. Guifang S., Rui Z., Peng L., Aixin F., Yongkang Z., *Laser surface alloying of C-B-W-Cr powders on nodular cast iron Rolls*, Surface and Coatings Technology, Volume 205, Issues 8-9, 25 January (2011) s. 2747-2754.

Zastosowanie technologii laserowych do zwiększania odporności na zużycie elementów maszyn

Abstrakt

W pracy przedstawiono przykłady elementów maszyn narażonych na zużycie tribologiczne, scharakteryzowano możliwości wykorzystania wiązki laserowej w inżynierii powierzchni oraz przeanalizowano efekty jakie można osiągnąć dzięki zastosowaniu hartowania ze stanu stałego, przetapiania czy też stopowania laserowego w aspekcie zwiększania odporności na zużycie.

W związku z tym, że zazwyczaj zużycie w wyniku tarcia dotyczy jedynie niewielkich fragmentów części maszyny wystarczające jest stosowanie metod dających możliwość uszlachetnienia warstwy wierzchniej tylko wybranych obszarów. Taką możliwość daje obróbka laserowa. W różnych próbach zużyciowych badania wykazały, że tego rodzaju obróbka pozwala znacznie zredukować zużycie badanych obiektów. Badania własne wykazały, że elementy obrobione laserowo w których wytworzona została w wyniku stopowania odpowiednia warstwa wierzchnia (przede wszystkim o drobnoziarnistej, przesyconej, wzbogaconej o nowopowstałe fazy zawierające wprowadzony pierwiastek stopowy mikrostrukturze o twardości ponad 4-krotnie większej od twardości rdzenia materiału) charakteryzowały 17-krotnie mniejszym ubytkiem masy po przeprowadzonych testach zużyciowych niż elementy nieobrobione. Ponadto, badania własne wykazały, że w niektórych przypadkach może zostać również zredukowany ubytek masy współpracujących z nimi elementów. Zwiększenie odporności na zużycie można uzyskać nie tylko dla części maszyn pracujących w węźle kinematycznym z inną częścią maszyny (jak czop wału – panewka), ale również dla części maszyn rolniczych np. pracujących glebie (jak stopki redlic siewnika).

Laserowa obróbka cieplna jest bardzo obiecującą metodą, co powoduje, że właściwe jest podejmowanie dalszych badań w tym obszarze, w szczególności testów tribologicznych polegających na badaniu rezultatów obróbki w rzeczywistych węzłach kinematycznych.

Słowa kluczowe: laserowa obróbka cieplna, zużycie, elementy maszyn

Laser techniques application for increasing of the wear resistance of machine parts

Abstract

In the paper examples of machine elements exposed to the tribological wear have been presented, the possibility of the laser beam application in surface engineering has been characterized, the effects that can be achieved by laser hardening from the solid state, remelting or alloying in the aspect of increasing the wear resistance have been analyzed.

Due to the fact that usually the friction wear only occurs in relatively small areas of the machine part, it is enough to use methods that allow to modify the surface layer only in selected regions. Laser treatment gives such possibility. As a results of many tests, the research showed that this type of treatment significantly reduces the wear of the tested objects. Own research has shown that the laser treated (alloyed) objects with created appropriate surface layer (fine-grained, supersaturated, enriched with newly formed phases containing the alloying element microstructure with 4-times higher hardness than the hardness of the material) were characterized by 17-times reduction of the weight loss (in comparison to the untreated objects) after performed wear tests. In addition, own research has shown that in some cases the weight loss of the co-operating elements can also be reduced. Increase of the wear resistance can be achieved not only for parts of machines operating in the kinematic node with another machine part (like journal of the shaft and pan) but also for parts of agricultural machines for example working in the soil (like coulter flaps of the seeder).

Laser heat treatment is a very promising method, that causes that it is reasonable to continue the research in this area, in particular in the range of tribological tests consisting in studying the results of the treatment in real kinematic nodes.

Keywords: laser heat treatment, wear, machine parts

Krzysztof Foit[1]

Realizacja koncepcji Przemysłu 4.0 w ujęciu zagadnienia programowania robotów przemysłowych

1. Wprowadzenie

Pojęcie „Przemysł 4.0" (ang. *Industry 4.0*) związane jest z istotnymi przemianami, jakie dokonują się w sferze przemysłu. Nowoczesne techniki wytwarzania, w połączeniu ze współczesnymi trendami technologicznymi, kształtują nowe podejście do rynku i zarządzania przedsiębiorstwem, a w konsekwencji nowy model relacji pomiędzy przedsiębiorcami oraz pomiędzy wytwórcą a klientem.

Istota koncepcji „Przemysłu 4.0" leży w nowym sposobie pojmowania samego przedsiębiorstwa i procesu wytwarzania, który jest na tyle odmienny, że zasłużył sobie na określenie „Czwarta rewolucja przemysłowa". W istocie, pierwsze trzy rewolucje, kamienie milowe rozwoju przemysłowego – czyli napęd parowy, elektryfikacja i linia produkcyjna oraz automatyzacja – wprowadzały nową jakość, lecz kontynuowały pewną tradycję zawartą w określeniu „fabryka", które oznaczało pewną zwartą całość, związaną z konkretnym położeniem geograficznym, gdzie zarówno marszruta przedmiotu jak i jego ostateczna postać były sztywno określone. „Przemysł 4.0" zrywa całkowicie z klasycznym określeniem fabryki, oferując dynamicznie konfigurowalne środowisko, znacząco otwarte na wymagania klienta. Dzieje się tak ze względu na zatarcie granic pomiędzy przedsiębiorstwem, a światem zewnętrznym. Takie podejście jest przejawem zastosowania fundamentalnych zasad „Przemysłu 4.0", do których należą [1]:

- koncepcja Smart Factory, w której zastosowanie czujników wraz z odpowiednimi systemami komputerowego przetwarzania danych, pozwala na utworzenie cyfrowej kopii środowiska fabryki, co z kolei umożliwia zastosowanie różnych technik modelowania komputerowego do tworzenia autonomicznych i samoorganizujących się systemów cyfrowych, których działanie jest związane bezpośrednio z maszynami (aktorami) znajdującymi się w fabryce,
- systemy cyber-fizyczne, stanowiące łącze pomiędzy światem fizycznym (działaniem lub monitorowaniem pewnego procesu) a warstwą programową, umożliwiającą komunikację z innymi systemami, podejmowanie decyzji itd.,
- samoorganizacja, która jest ściśle związana z decentralizacją zarządzania procesem, polegającą na dobieraniu i uszeregowaniu konkretnych zasobów w zależności od potrzeb i celu, którym jest finalna postać produktu,

[1] krzysztof.foit@polsl.pl, Instytut Automatyzacji Procesów Technologicznych i Zintegrowanych Systemów Wytwarzania, Wydział Mechaniczny Technologiczny, Politechnika Śląska, ul. Konarskiego 18A, 44-100 Gliwice

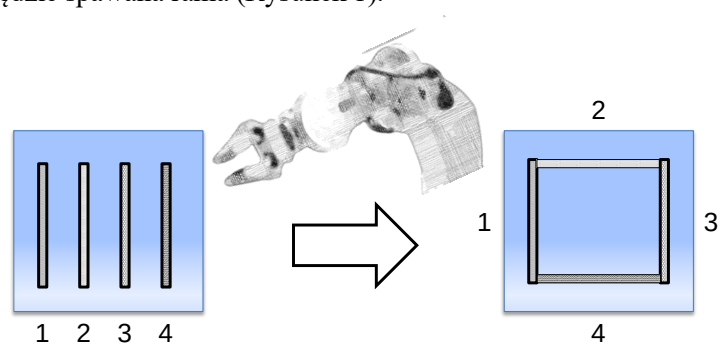
- elastyczność przejawiająca się m.in. w zarządzaniu dostawami i dystrybucją elementów, dostosowaniem do bieżących potrzeb klientów, personalizacji produktów, itp.,
- efektywne wykorzystanie zasobów, poszanowanie dla środowiska naturalnego, zrównoważony rozwój.

Odnosząc się do wymienionych fundamentów koncepcji „Przemysłu 4.0", należy zwrócić szczególną uwagę na wdrażanie „inteligentnych" systemów wytwarzania. Wiąże się to z odejściem od klasycznych układów sterowania, w których wykonywany proces znajduje swoje odzwierciedlenie w sztywnym, często sekwencyjnym przebiegu programu sterującego. Ten trend ma również zasadniczy wpływ na warstwę sprzętową układów automatyki przemysłowej, bowiem użycie bardziej skomplikowanych algorytmów wymaga zastosowania urządzeń o większych możliwościach szybkiego przetwarzania informacji, jak również odpowiedniej ilości pamięci operacyjnej, zdolnej pomieścić zarówno kod programu, jak i przetwarzane dane.

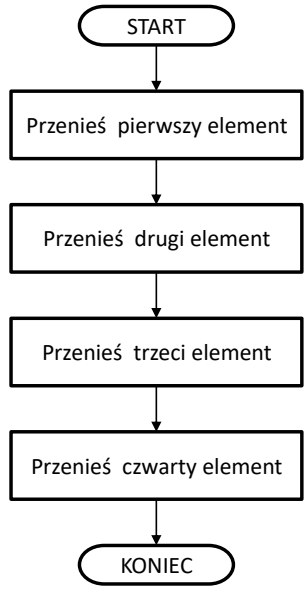
Odniesienie do koncepcji „Przemysłu 4.0" ma szczególne znaczenie w robotyce, bowiem to właśnie robot przemysłowy, w postaci stacjonarnej lub mobilnej (pojazdy AGV) ma przejąć działania manipulacyjne i transportowe, wykonywane dotychczas przez człowieka. Należy zatem zastanowić się, jak postępujące zmiany w postaci elastycznego podejścia do procesu wytwarzania, przekładają się na zmianę podejścia do zagadnień programowania układów sterowania robotów. Niniejsze opracowanie ma na celu przybliżenie wybranych zagadnień związanych ze współczesnymi metodami programowania robotów przemysłowych, które odnoszą się do realizacji koncepcji Przemysłu 4.0. Poruszone problemy będą poparte odpowiednim przeglądem dostępnej, aktualnej literatury naukowej i branżowej.

2. Definiowanie zadań robota – programowanie na poziomie robota

Klasyczne podejście do definiowania zadań realizowanych przez robota, sprowadza się przede wszystkim do określenia algorytmu postępowania, które ma doprowadzić do zrealizowania zamierzonego celu. W celu zilustrowania metody, rozpatrzmy prosty przykład. Zadaniem robota jest właściwe ułożenie czterech stalowych kształtowników, z których będzie spawana rama (Rysunek 1).



Rysunek 1. Schematyczne przedstawienie zadania wykonywanego przez robota [opracowanie własne]



Rysunek 2. Prosty algorytm opisujący sekwencję ruchów manipulatora robota [opracowanie własne]

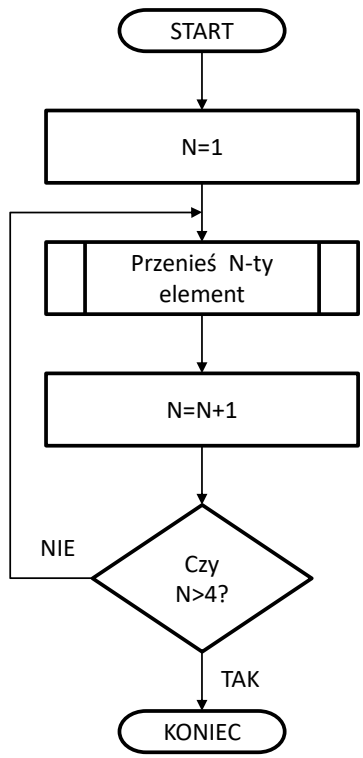
Wykonanie zadania wymaga zrealizowania konkretnych czynności manipulacyjnych, tj. pobrania odpowiedniego elementu z puli i ustawienie go we właściwej pozycji. W najprostszej postaci można wykonać sekwencję ruchów, tak jak przedstawiono to za pomocą schematu blokowego na Rysunku 2. W kolejnym kroku można algorytm nieco zmodyfikować, wprowadzając jeden podprogram (por. Rysunek 3) służący do przenoszenia elementów, który będzie wywoływany wielokrotnie, w pętli, w zależności od wartości stałej, określającej liczbę elementów (w omawianym przypadku podprogram będzie wywołany czterokrotnie).

Wprowadzona modyfikacja pozwoliła na uzyskanie pewnej elastyczności, polegającej na możliwości dostosowania programu do dowolnej liczby elementów poprzez zmianę jednej tylko wartości. Skutkuje to jednak większą złożonością kodu podprogramu przenoszącego elementy, gdyż w tym przypadku dane o położeniu osi manipulatora nie są „sztywno" zapisane w programie, lecz muszą być uszeregowane w pewnej indeksowanej strukturze, z której są kolejno odczytywane. Ceną za wspomnianą większą elastyczność kodu jest z pewnością większy nakład pracy podczas usuwania potencjalnych błędów, a w przypadku niewielkiej liczby iteracji, rozbudowany kod może zużyć więcej pamięci od prostej sekwencji poleceń.

Oprócz modelu programowania w formie graficznej lub tekstowej, prowadzonego przy użyciu komputera, umożliwiającego przygotowanie programu bez udziału robota (tzw. programowanie „off-line"), istnieje również sposób, który wymaga zaangażowania maszyny w jej środowisku działania. Jest to tzw. programowanie przez nauczanie, w którym programista-operator znajduje się bezpośrednio w zrobotyzowanej celi, scena jest skonfigurowana w identyczny sposób jak podczas normalnej pracy robota, natomiast gniazdo jest wyłączone z procesu produkcyjnego. Ten sposób

programowania określany jest jako „programowanie przez uczenie" lub „on-line". Może ono występować w dwóch odmianach:

- point-to-point (PTP): operator definiuje sekwencję ruchów robota, przemieszczając kiść z narzędziem w odpowiednie miejsca i ustawiając odpowiednio ramiona manipulatora, a każde położenie jest następne wprowadzane do pamięci i uzupełniane odpowiednią instrukcją odpowiadającą za realizację ruchu bądź operację za pomocą efektora,
- continuous path (CP): następuje wodzenie kiścią manipulatora po zadanej trajektorii, podczas którego system sterujący z określoną częstotliwością zapamiętuje kolejne punkty ścieżki i stany efektora.



Rysunek 3. Modyfikacja algorytmu przedstawionego na rysunku 2 – wprowadzenie podprogramu wywoływanego w pętli [opracowanie własne]

Ze względu na fakt, że metody „on-line" i „off-line" są względem siebie komplementarne, możliwe jest również mieszanie obydwu sposobów ze sobą tak, że część kodu powstaje na komputerze, natomiast inna przy bezpośredniej współpracy robota. Taki model nosi nazwę programowania hybrydowego. Przygotowany w ten sposób program jest zorientowany na robota, tj. opisuje jego środowisko pracy z perspektywy przestrzeni roboczej manipulatora i wykonywanych przez niego ruchów.

Odnosząc się do języków programowania robotów, Zieliński [2] zauważa, że większość z nich odnosi się do filozofii języków służących do programowania

komputerów, szczególnie BASIC-a i Pascala, zachowując podobną do nich składnię i wykorzystując podobnie brzmiące słowa kluczowe. Można więc wnioskować, że strukturalny model programowania jest nadal chętnie wykorzystywany do tworzenia oprogramowania robotów, nie tylko ze względu na swoją prostotę, ale również ze względu na ograniczenia sprzętowe starszych sterowników robotów, które wciąż są używane w procesie produkcyjnym. Sheu i Xue [3] wymieniają pewne wspólne cechy, którymi charakteryzują się języki programowania robotów. Są to m.in.:

- zapis rozmieszczenia i orientacji obiektów,
- zapis parametrów ruchu i sekwencji ruchów,
- przetwarzanie informacji z sensorów, jako źródła wiedzy o otoczeniu manipulatora robota; pozyskiwane dane służą do dostrojenia parametrów ruchu manipulatora czy np. siły uchwytu przedmiotu.

Na podstawie wymienionych własności można wywnioskować, że w klasycznym podejściu do definiowania zadań realizowanych przez robota, możliwe jest w miarę elastyczne podejście do wprowadzania ewentualnych zmian w kodzie programu, jednak reagowanie robota na zmiany zachodzące w jego środowisku jest mocno ograniczone, ze względu na zapisanie pewnych parametrów (np. rozmieszczenia obiektów) „na sztywno" w kodzie programu lub pamięci robota. Można sobie wyobrazić sytuację, w której manipulator ma możliwość transportu zamiennie dwóch różnych przedmiotów: ciężkiego i odpornego na zgniot oraz lekkiego i kruchego. Informacje uzyskane z czujników mogą być użyte do uruchomienia odpowiedniego podprogramu i wprowadzenia odpowiednich parametrów siły uchwytu czy zdefiniowania sposobu uchwycenia, tak by oba zadania mogły być wykonane za pomocą tego samego manipulatora.

Wspomniane podejście ujawnia swoje niedoskonałości w przypadku, gdy scena robota ulega dynamicznym i trudnym do przewidzenia zmianom. Przypadkowa orientacja przedmiotu, konieczność wyboru konkretnego typu przedmiotu spośród innych czy możliwość kolizji z obiektem przemieszczającym się w obrębie sceny robota, są zagadnieniami wykraczającymi poza możliwości programowania strukturalnego ze zdeterminowanym opisem konfiguracji przestrzeni roboczej i ścieżki manipulatora robota. Bardzo istotne są również zagadnienia, w których dochodzi do zamknięcia łańcucha kinematycznego manipulatorów, wykonujących działania na wspólnym przedmiocie. Jednoczesne oddziaływanie obu maszyn na ten sam przedmiot może występować zarówno wtedy, gdy jeden z manipulatorów ustawia odpowiednio przedmiot a drugi przeprowadza np. operację cięcia czy szlifowania jak i w przypadku, gdy oba manipulatory przenoszą równocześnie ten sam przedmiot. Zagadnienia takie – mimo że w teorii są dość łatwe do oprogramowania – w praktyce wymagają specjalnego podejścia, głównie od strony systemu sterowania, a to ze względu na różnice pomiędzy manipulatorami, powstające na etapie produkcji. W pracach Normana i in. [4] oraz Al-Yahmadiego i in. [5] zaprezentowano różne rozwiązania tego problemu. Charakterystyczne dla obu przedstawionych rozwiązań jest podejście skupiające się na manipulowanym obiekcie, nie zaś na manipulatorach, biorących udział w procesie oraz istnienie kanału komunikacyjnego pomiędzy robotami. Skupienie się na produkowanym przedmiocie cechuje również filozofię „Przemysłu 4.0" – to przedmiot zgłasza zapotrzebowanie na konkretne działania, natomiast

maszyny organizują się tak, by je zrealizować. Przykład ten pokazuje wyraźnie, że chcąc omawiać zagadnienia programowania robotów z punktu widzenia „Przemysłu 4.0" należy spojrzeć na realizowane przez nie zadania z innej perspektywy, skupiając się na manipulowanym przedmiocie.

3. Zaawansowane metody definiowania zadań robotów przemysłowych

Ograniczenia wynikające ze specyfiki klasycznych metod programowania robotów nie pozwalają na bezpośrednie wykorzystanie algorytmów związanych z inteligencją maszynową, takich jak logika rozmyta, sieci neuronowe czy programowanie agentowe. Wdrażanie tych rozwiązań jest możliwe dzięki szybkiemu rozwojowi rynku przemysłowych komputerów PC, w związku z czym osiągalne jest zapewnienie żądanej mocy obliczeniowej i odpowiedniego oprogramowania. Umożliwia to sprzężenie robotów i maszyn z jednostkami, które stanowią swego rodzaju uzupełnienie i rozszerzenie możliwości istniejącego parku maszynowego.

Wykorzystanie komputerów klasy PC jest istotnym aspektem rozszerzenia możliwości robotów przemysłowych z punktu widzenia wdrażania idei „Przemysłu 4.0". Robot, wspomagany przez komputer z odpowiednim oprogramowaniem, jest w stanie osiągnąć wysoki stopień autonomii w sferze podejmowania decyzji o przebiegu procesu manipulacji. Poprzez takie podejście zmienia się perspektywa, z której programista planuje przebieg procesu z udziałem robota, skupiając się na istotnych szczegółach, a pomijając w zasadzie proces planowania trajektorii. W dalszym ciągu będą omówione najważniejsze aspekty wybranych, zaawansowanych technik programowania robotów.

3.1. Programowanie przez demonstrację

Naturalnym procesem jest dążenie inżynierów do uproszczenia sposobu programowania robotów. Samo słowo „uproszczenie" nie pokrywa się jednak z rzeczywistością, jeśli wziąć pod uwagę całe zaplecze programowe i technologiczne, które umożliwia działanie bardziej zaawansowanych interfejsów programisty, wykorzystujących m. in. wirtualną lub rozszerzoną rzeczywistość. Opisany problem nie jest nowy, gdyż już w 1986 roku Norman i Draper [6] opisywali „przepaść" pomiędzy wyobrażeniem problemu, istniejącym w mózgu programisty a jego reprezentacją w postaci akceptowalnego przez komputer programu. Ich zdaniem istnieją dwa rozwiązania tego problemu: przybliżyć programistę do systemu lub przybliżyć system do programisty. O ile pierwsza koncepcja jest zwykle realizowana przez naukę odpowiednich języków programowania, o tyle druga wiąże się z „nauczeniem" maszyny, jak rozumieć intencje programisty.

Rysunek 4. Przykładowe rozwiązanie programowania robota przez demonstrację [źródło: Pires J.N. *Programming by demonstration*, online: YouTube, https://youtu.be/IXSho9IoGYU, dostęp 10.09.2017]

Początki rozwoju programowania przez demonstrację odnoszą się do metody programowania komputerów nazywanej „programowaniem poprzez przykład", chociaż równolegle był używany termin „programowanie przez demonstrację". Z czasem drugie z określeń trafiło na stałe do dziedziny robotyki, podczas gdy pierwsze pozostało związane z programowaniem komputerów. Opisując problem w dużym uogólnieniu można powiedzieć, że w świecie komputerów z programowaniem przez demonstrację spotykamy się głównie podczas pracy z programami biurowymi. Pierwszym przykładem może być korzystanie z systemu „nagrywania" makropoleceń. Rejestracja makropolecenia polega na zapamiętaniu przez aplikację sekwencji poleceń, która może zawierać kliknięcia myszą i naciśnięcia klawiszy w celu wywołania polecenia lub ustalenia pewnych parametrów. Po zakończeniu rejestracji, na podstawie działań użytkownika generowany jest kod makropolecenia, który następnie może być edytowany i wielokrotnie wykorzystywany. W przypadku pracy z robotem, programowanie przez demonstrację może kojarzyć się (częściowo słusznie) z metodą Continuous Path programowania przez nauczanie (por. Rysunek 4) – podobnie jak w CP, operator ręcznie ustawia ramiona robota, wykonuje ruchy, które są rejestrowane przez system. Zasadniczą różnicą jest późniejsze wykorzystanie tak utworzonego programu. W celu lepszego zilustrowania tego problemu, należy wrócić do przykładu pracy na komputerze z użyciem pakietu biurowego. Otóż makropolecenie, które zostało zarejestrowane przez użytkownika w jednym dokumencie, może być w większości przypadków wykorzystane globalnie, w każdym innym pliku. Niektóre edytory tekstu wspomagają użytkownika, proponując dokończenie wyrazów, automatycznie formatując tekst itp. Działania te nie muszą jednak być skojarzone z konkretną sekwencją naciśnięć klawiszy czy ruchów myszką, ale są nakierowane na styl działania użytkownika – tak więc funkcjonują na pewnym poziomie uogólnienia. Podobnie w przypadku programowania robotów, zdolność do uogólniania jest wyróżnikiem programowania poprzez demonstrację – na przykład napełnienie szklanki wodą z butelki wymaga zawsze wykonania podobnej do siebie sekwencji ruchów, jednak butelki mogą mieć różny kształt i ciężar, a szklanki mogą różnić się między

sobą kształtem i pojemnością. Zatem robot ma podany sposób wykonania zadania, jednak parametry ruchu musi doprecyzować samodzielnie, w oparciu o wiedzę pozyskaną z sensorów. Samo uogólnianie jest na tyle złożonym procesem, że konieczna jest przy tym asysta operatora [7-9]. Problemem jest przede wszystkim klasyfikacja zdarzeń zachodzących podczas demonstrowania. System nie jest w stanie samodzielnie zadecydować, czy zarejestrowana czynność jest istotną częścią procesu demonstracji, korektą innej czynności czy też niezamierzonym działaniem operatora. Należy również zaznaczyć, że nie zawsze zarejestrowana czynność jest zgodna z intencją operatora. Niezgodność ta wynika przede wszystkim ze sposobu prowadzenia efektora przez operatora – ruch ręki może być nieprecyzyjny, a zatem jego odwzorowanie może być obarczone błędem. Drugą przyczyną jest szum, który zakłóca wartość mierzoną przez czujniki, rejestrujące przebieg czynności. W fazie wczesnego rozwoju programowania przez demonstrację, jednym ze sposobów uniknięcia wspomnianych niezgodności była interpolacja ścieżki w odniesieniu do wybranych, uznanych za prawidłowe, punktów zarejestrowanych w czasie demonstracji. W ten sposób można było wyznaczyć podstawowe akcje, prowadzące do osiągnięcia założonych celów (przemieszczenie i zorientowanie efektora w punkcie) [9]. Na tej podstawie powstawał ciąg czynności, zapisanych w postaci „stan – czynność – stan". Taki ciąg mógł być z kolei przekonwertowany do postaci reguł warunkowych „jeśli – to", które opisywały podejmowane czynności w odniesieniu do relacji z elementami otoczenia robota lub innymi czynnościami. Graficzną reprezentację takiego zapisu stanowił graf, w którym wierzchołki odwzorowują poszczególne stany, natomiast łuki odpowiadają realizowanym czynnościom [9].

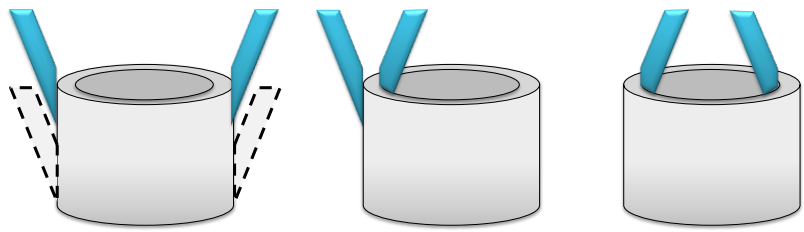
Pojęcie uogólniania zademonstrowanych zdarzeń czy czynności, może być rozumiane w różny sposób, jednak Kahn ujął to w bardzo ciekawy sposób: uogólnianie poprzez usunięcie szczegółów. W swoim opracowaniu odnoszącym się do programowania poprzez przykład [10], autor odnosi się do aplikacji „ToonTalk" – systemu opartego na animowanych aktorach, przeznaczonego dla dzieci, służącego do nauki podstaw programowania. Stwierdza on między innymi:

„Pierwsza z historii pokazuje uczącym się ToonTalk'a, że „dymek" [myślowy – przyp. wł.] nad głową robota wskazuje na rodzaj pudełka, z którym robot chce pracować. Jeśli dostanie pudełko różniące się od pożądanego, to przekaże je innemu członkowi zespołu. W sytuacji, gdy otrzyma pudełko opisane bardziej szczegółowo niż to, o którym „myśli", wszystko będzie w porządku i robot je zaakceptuje." [tłum. wł. na podst.10]

Opisane podejście sprawdzi się również w przypadku uogólniania podczas programowania rzeczywistego robota. Jest ono często wykorzystywane w fazie planowania trajektorii robota, gdy przedmiot manipulowany traktuje się jako objętość w przestrzeni, opisaną regularnymi kształtami (sześcian, prostopadłościan, kula) i nie wnika się w szczegóły związane z jego geometrią. Oczywiście jest to duże uproszczenie problemu uogólniania, jednak dobrze opisuje jego ogólną zasadę.

W większości przypadków, podczas programowania przez demonstrację, istotną rolę odgrywa proces uczenia maszynowego. Uogólnianie następuje na podstawie odpowiednio dużego zestawu zademonstrowanych czynności, jednak – podczas demonstrowania – nie zawsze można uwzględnić wszystkie możliwości, gdyż uczenie

(a przede wszystkim demonstrowanie) po prostu trwałoby zbyt długo. Przykładowo, na Rysunku 5 przedstawiono kilka możliwości uchwycenia przedmiotu typu tuleja, za pomocą chwytaka mechanicznego, równoległego. W zależności od rodzaju uchwycenia (zewnętrzne, mieszane, wewnętrzne) zmienia się nie tylko wymagana dokładność ustawienia chwytaka względem osi określającej średnicę wewnętrzną lub zewnętrzną tulei, ale również położenie względem osi wzdłużnej, czyli wysokość uchwycenia tulei, mierzona wzdłuż ścianki. Zbiór rozpatrywanych (a zarazem demonstrowanych) przypadków można zdecydowanie zmniejszyć poprzez narzucenie odpowiednich ograniczeń [11]. Na przykład w odniesieniu do wysokości uchwycenia tulei, można zdefiniować wartość minimalną i maksymalną, które to przypadki posłużą za wejście w procesie uczenia maszynowego.



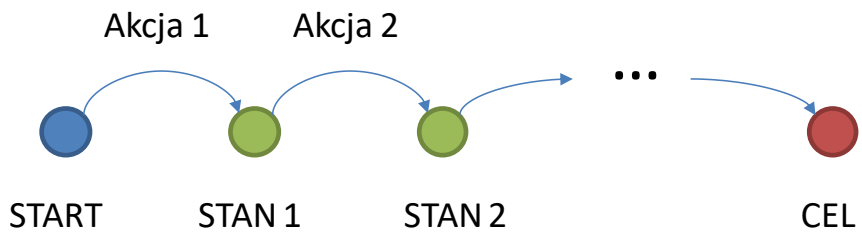
Rysunek 5. Różne możliwości uchwycenia przedmiotu typu tuleja za pomocą chwytaka równoległego [opracowanie własne]

Jako wartość dodaną procesu programowania przez demonstrację otrzymuje się bazę przypadków, która może być również wykorzystywana na innym poziomie ogólności, podczas programowania na poziomie zadania.

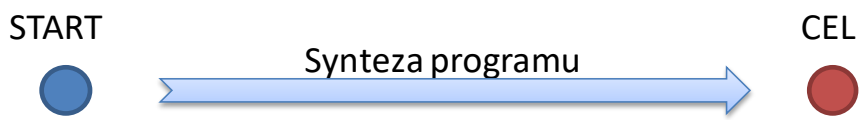
3.2. Programowanie na poziomie zadania

Pojęcie programowania na poziomie zadania, rozszerza omówioną wcześniej metodę programowania przez demonstrację. Opis na poziomie zadania czynności wykonywanych przez robota, stoi na wyższym poziomie abstrakcji, gdyż ten rodzaj programowania zmierza do wykorzystania języka naturalnego. Opisując to prostym przykładem, programowanie na poziomie zadania polega na wydaniu polecenia w stylu „pomaluj płot na zielono". Polecenie takie jest zrozumiałe dla człowieka, który przede wszystkim potrafi przypisać znaczenie każdego słowa danym czynnościom lub przedmiotom, a ponadto wie jak wykonać te czynności. Istnieje pewne prawdopodobieństwo, że kiedyś wykonywał czynność malowania jakiegoś przedmiotu, lub chociażby obserwował proces malowania (demonstracja) i na tej podstawie może wypracować własny plan działania w odniesieniu do przedmiotu, jakim jest płot. W bardzo podobny sposób proces ten przebiega w odniesieniu do programowania robota, gdzie definiuje się przede wszystkim cel działania, ale nie podaje się szczegółowego przepisu jak ten cel osiągnąć. Graficznie różnicę tę przedstawiono na Rysunku 6. W przypadku programowania przez demonstrację, podane są poszczególne stany i czynności, które pozwalają te stany osiągnąć. W przypadku programowania na poziomie zadania, podany jest stan początkowy i stan końcowy (cel), natomiast wszystkie czynności muszą być samodzielnie określone przez system sterowania robota.

a)



b)



Rysunek 6. Porównanie programowania przez demonstrację (a) z programowaniem na poziomie zadania (b) [opracowanie własne]

Lozano-Perez [12] zdefiniował trzy, podstawowe etapy programowania na poziomie zadania. Są to: modelowanie sceny, opis zadania i synteza programu robota. W części dotyczącej opisu zadania, autor wskazuje że jedną z możliwości jest przedstawienie sekwencji stanów, w jakich znajduje się scena robota, po wykonaniu każdej z czynności. Zatem, odmiennie do przypadku programowania przez demonstrację, tym razem wskazuje się poszczególne stany, nie zaś czynności, pozwalając by akcje podejmowane przez robota były wynikiem syntezy. Syntezę programu można z powodzeniem wspomagać komputerowo, używając języka programowania logicznego (np. Prolog lub podobne) [13, 14].

Poza syntezą kodu programu, konieczne jest również zaplanowanie trajektorii manipulatora robota. W tym przypadku konieczne jest również posiadanie informacji na temat rozmieszczenia obiektów na scenie robota, wliczając w to przedmioty manipulowane i ewentualne przeszkody. Z uwagi na zmieniającą się dynamicznie konfigurację sceny, roboty często wyposażone są w odpowiedni zestaw czujników, w tym czujniki wizyjne. Ostatnie z wymienionych sensorów umożliwiają nie tylko lokalizację obiektu, ale również jego rozpoznanie i określenie jego orientacji w przestrzeni.

Programowanie na poziomie zadania cechuje się dużą swobodą podejmowania decyzji przez system sterujący co do kształtu planowanej trajektorii i formy kodu programu. W tym sensie wpisuje się w założenia „Przemysłu 4.0”, tworząc maszynę o wysokim stopniu niezależności i zdolności do samoorganizacji.

3.3. Zaawansowane metody programowania robotów w kontekście filozofii „Przemysłu 4.0”

Przedstawione metody zaawansowanego programowania robotów w postaci programowania przez demonstrację i programowania na poziomie zadania wciąż są przedmiotem aktualnych badań naukowych. Jak dotąd większość badań ma charakter

teoretyczny, gdyż utylitarność wypracowanych metod jest bardzo ograniczona. Dzieje się tak głównie ze względu na brak standaryzacji w dziedzinie języków programowania robotów, stąd nawet wypracowanie jednolitych algorytmów działania wymaga za każdym razem poszukiwania sposobów przełożenia ich na konkretny język działający na konkretnym typie robota. Nieco mniejszym problemem jest zapewnienie odpowiedniej mocy obliczeniowej i zasobów pamięci operacyjnej, gdyż wielu czołowych producentów systemów sterowania przemysłowego postawiło na rozwój przemysłowych komputerów klasy PC. Wykorzystują one standardowe typy procesorów dostępnych na rynku i mogą działać pod kontrolą systemu Microsoft Windows lub innego, przewidzianego do eksploatacji z procesorami architektury x86/x64. Produkowane są również jednostki hybrydowe, które w jednej obudowie łączą sterownik programowalnego (PLC) i komputer przemysłowy PC. Umożliwia to sterowanie urządzeniami z zastosowaniem zaawansowanych metod numerycznych, takich jak teoria gier czy sieci neuronowe. Zatem nie ma przeszkód technicznych, by wprowadzić wspomniane metody do sterowania zaawansowanymi systemami zrobotyzowanymi i większości przypadków ich implementacja jest możliwa wporozumieniu z producentem (stanowiska realizujące konkretne zadania, jak np. malowanie, sortowanie czy pobieranie części z nieuporządkowanych zbiorów: pudła, kosze itp.) lub we współpracy z ośrodkami naukowymi (prace badawczo-rozwojowe, wdrażanie innowacji). Szerszy zakres działań umożliwiłoby wprowadzenie norm dotyczących języków programowania robotów, tak jak to ma miejsce w przypadku sterowników logicznych.

4. Przykłady zaawansowanych interfejsów programowania robotów przemysłowych

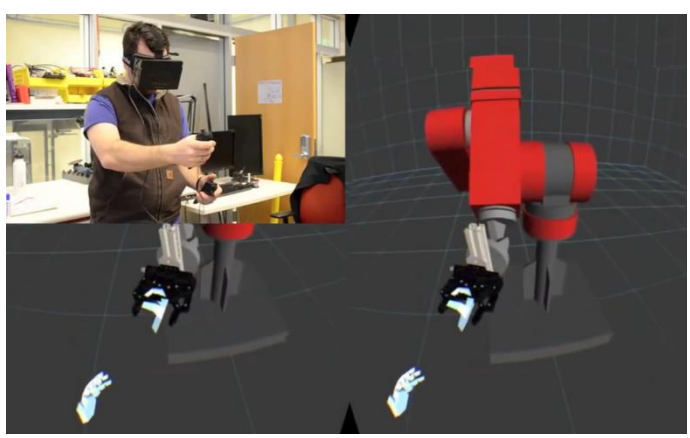
Założenia filozofii „Przemysłu 4.0" odnoszą się nie tylko do inteligentnych i samoorganizujących się fabryk, ale w dużej mierze przejawiają się w dążeniu do zatarcia różnic pomiędzy przemysłowymi interfejsami HMI a tymi używanymi codziennie w gospodarstwie domowym, przestrzeni publicznej czy Internecie. Uzyskanie odpowiedniej współpracy człowieka z robotem (podczas programowania) lub robota z innym robotem czy maszyną (podczas wykonywania zadań w ramach procesu produkcyjnego), wymaga użycia odpowiednio prostych interfejsów programistycznych jak również odpowiednio zaawansowanych sensorów, pozwalających na ciągłe śledzenie środowiska robota, upodabniając się w działaniu do funkcji realizowanych przez zmysły człowieka. W niniejszym rozdziale zostanie zaprezentowanych kilka przykładów tego typu zaawansowanych systemów.

Odnosząc się do programowania przez demonstrację, Folscher i Kruger [15] zaproponowali system programowania robota przemysłowego, oparty na śledzeniu znacznika w przestrzeni. System składał się z czterech kamer, monitorujących ruch znacznika w postaci pałeczki z zamocowaną na końcu diodą LED oraz komputera z oprogramowaniem MATLAB. Ruch znacznika w przestrzeni był śledzony przez kamery, a następnie składany w trójwymiarowy ślad. Kod dla robota KUKA w języku KRL (KUKA Robot Language) był generowany w postaci sekwencji poleceń ruchu do kolejnych punktów określonych na podstawie prześledzonej trajektorii znacznika. Autorzy dokonali porównania programowania ruchu robota za pomocą opracowanej przez nich metody z klasycznym programowaniem przy użyciu panelu operatorskiego.

Wykazano, że autorska metoda pozwoliła skrócić czas planowania trajektorii o 78,9% w stosunku do podejścia klasycznego, podczas realizacji zadania typu „pick&place" („chwyć i odłóż").

Odmienny, ale równie ciekawy sposób programowania zaprezentowano w [16]. Sterowanie manipulatorem odbywa się z zastosowaniem dwóch rodzajów interakcji: poprzez wydawanie komend głosowych i poprzez gesty. System oparty jest na oprogramowaniu ROS (*Robot Operating System*), które współpracuje z modułami rozpoznawania mowy i gestów. Gesty są rozpoznawane w oparciu o informacje z sensora Xbox Kinect. Proces programowania polega na wydaniu odpowiedniej komendy głosowej i wykonania gestu odpowiadającego za ruch manipulatora. Ruch jest realizowany tak długo, jak długo operator wykonuje określony gest. Sposób ten można porównać do „naprowadzania" kierowcy pojazdu czy operatora dźwigu, gdy obserwator słowami i gestami koryguje jego działania. W tym miejscu można pokusić się o dyskusję na temat przydatności takiego rozwiązania w procesie programowania robotów. Podejście to z pewnością wymaga nauczenia się odpowiednich komend i gestów – opanowanie sterowania manipulatorem w taki sposób też nie wydaje się łatwe, jednak po nabyciu pewnej wprawy może okazać się efektywniejsze niż korzystanie z panelu operatora. Przedmiotem dyskusji powinny też być kwestie bezpieczeństwa, gdyż operator powinien mieć bezpośredni, fizyczny dostęp do przycisku służącego do awaryjnego zatrzymania maszyny.

Rozwiązaniem o dużym potencjale jest wykorzystanie komputerowo generowanego świata do programowania robotów przemysłowych. Naukowcy prezentują różne typy metod oparte o wirtualną, poszerzoną lub mieszaną rzeczywistość. Przykład wykorzystania wirtualnej rzeczywistości przedstawiono na Rysunku 7. Operator wykorzystuje system Oculus Rift do poruszania manipulatorem robota w wirtualnym świecie, „chwytając" manipulator za zamocowany do niego chwytak i ustawiając go w pożądanej konfiguracji. Twórcy systemu IVRE [17] twierdzą, że ich celem jest umożliwienie zdalnej interakcji z robotem pracującym w niebezpiecznym środowisku lub posługującego się wyposażeniem, które może być niebezpieczne dla przebywającego w pobliżu człowieka.



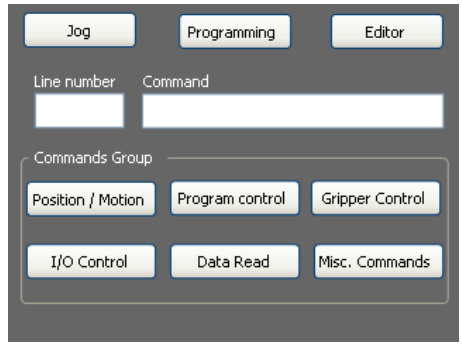
Rysunek 7. Interakcja z robotem w wirtualnej rzeczywistości [źródło: 17]

Odnosząc się do założeń Przemysłowego Internetu Rzeczy, Mateo i in. [18] zaprezentowali aplikację działającą na urządzeniach z systemem Android, służącą do programowania robota z wykorzystaniem reprezentacji wizualnej, bazującej na filozofii języka „Scratch" – programowanie odbywa się na zasadzie „dopasowywania" puzzli, reprezentujących odpowiednie instrukcje. Aplikacja pozwala zamienić tablet lub telefon z systemem Android w zaawansowany interfejs programisty. Szczególną funkcją programu jest możliwość użycia poszerzonej rzeczywistości, gdy na obraz z kamery, w czasie rzeczywistym nakładane są informacje dotyczące istotnych parametrów zadania, np. rozkład wektorów sił.

Na Rysunku 8 przedstawiono inne, przykładowe zastosowanie techniki poszerzonej rzeczywistości do planowania trajektorii robota. Ścieżkę wygenerowano w aplikacji uruchomionej na tablecie, stosując komendy w postaci gestów, które obserwowane przez kamerę urządzenia, pozwoliły na określenie kolejnych punktów trajektorii i ich wprowadzenie do pamięci robota. Rysunek 8 prezentuje moment odtwarzania zaplanowanej ścieżki, która jest prezentowana na ekranie urządzenia i złożona z obrazem pochodzącym z kamery.

Rysunek 8. Przykład użycia rozszerzonej rzeczywistości do intuicyjnego programowania i monitorowania ruchu manipulatora robota [źródło: 19]

Wśród aplikacji na smartfony i tablety, które są przeznaczone do programowania robotów, najczęściej spotyka się takie, które działają lokalnie, na konkretnym urządzeniu [18, 20, 21]. Odmienną zasadę działania zastosowano w przypadku interfejsu opisanego w [22], przeznaczonego do współpracy z robotem Mitsubishi Movemaster RV-M1. W celu zapewnienia maksymalnej przenośności, panel programowania wyświetlany jest na urządzeniu typu tablet lub smartfon w postaci strony WWW, pobranej z lokalnego serwera, który jednocześnie pełni rolę komputera służącego do programowania robota. Wynikające ze sposobu komunikacji z serwerem uproszczenia nadają rozwiązaniu jedynie eksperymentalny charakter – w zamierzeniu, interfejs ten miał uzupełniać działanie panelu operatora, który pozbawiony był funkcji umożliwiających programowanie i umożliwiał jedynie przemieszczanie ramienia oraz wprowadzanie do pamięci kolejnych położeń manipulatora. Przykładowy ekran utworzonego interfejsu przedstawiono na Rysunku 9.



Rysunek 9. Przykładowy ekran wizualnego interfejsu programowania robota RV-M1 [opracowanie własne]

5. Podsumowanie

Współczesna technika komputerowa umożliwia tworzenie coraz bardziej zaawansowanych technicznie i graficznie interfejsów programowania robotów. Wiele z nich wykracza poza granice tekstowego tworzenia programów, wykorzystując techniki sterowania za pomocą mowy czy gestów. Impulsem do wdrażania nowych, efektywniejszych i bardziej intuicyjnych sposobów programowania jest realizacja koncepcji „Przemysłu 4.0". Filozofia ta zaciera granice pomiędzy wnętrzem fabryki, specjalistycznym podejściem do procesu wytwarzania, a światem zewnętrznym, nastawionym na intuicyjny, pozbawiony aury tajemniczości, dialog z maszyną. W niniejszej pracy przedstawiono podstawowe założenia „Przemysłu 4.0", odnoszące się do ogółu tego pojęcia, a następnie skupiono się na zagadnieniach związanych z programowaniem robotów. Przedstawiono zarówno klasyczne podejście do programowania, jak i zaawansowane metody, wymagające użycia sztucznej inteligencji, technik uczenia maszynowego, ale oferujące w zamian możliwość uproszczenia procesu programowania (programowanie przez demonstrację) czy używania poleceń w języku zbliżonym do naturalnego (programowanie na poziomie zadania). W ostatniej części przedstawiono i omówiono kilka współczesnych interfejsów operatora, służących do programowania robotów przemysłowych i nawiązujących do filozofii „Przemysłu 4.0". Skupiono się między innymi na formie dialogu człowieka z maszyną, która w omawianych przypadkach często opiera się na metodach wykorzystujących wirtualną lub poszerzoną rzeczywistość.

Zaprezentowane w pracy przykłady podejścia do problemu programowania robotów obejmują ponad trzydziestoletni przedział czasu. Wskazują one między innymi na ciągłą potrzebę poszukiwania nowych rozwiązań, zmierzających do uproszczenia tego procesu. Na tym tle widać wyraźnie wpływ rozwoju technologicznego na sposób prowadzenia badań. Pierwsza, istotna zmiana dotyczy urządzeń docelowych, na które tworzone jest specjalistyczne oprogramowanie – współcześnie coraz częściej są to smartfony i tablety, w odróżnieniu od komputerów PC, które były szeroko wykorzystywane w latach 90 ubiegłego wieku i na przełomie wieków. Widoczne jest również odejście od tekstowej formy prezentacji i wprowadzania danych, na rzecz wykorzystywania wizualnych technik programowania i szeroko pojętych sensorów monitorujących środowisko. Można stwierdzić, że współczesny

trend w dziedzinie programowania robotów zmierza w kierunku bezpośredniego dialogu z maszyną, co jeszcze do niedawna wydawało się być wizją zaczerpniętą z literatury science fiction.

Literatura

1. Lasi H., Fettke P., Kemper H.-G., Feld T., Hoffmann M. *Industry 4.0*, Business & Information Systems Engineering, 6 (4) (2014), s. 239-242.
2. Zieliński C. *Języki programowania robotów przemysłowych,* Automatyka Online, 2012,online: http://automatykaonline.pl/Artykuly/Robotyka/jzyki-programowania-robotow-przemysowych, dostęp 16 sierpnia 2017.
3. Sheu P.C.-Y., Xue Q. *Intelligent Robotic Planning Systems*, WORLD SCIENTIFIC, 1993.1.
4. Norman A., Schönberg A., Gorlach I., Schmitt R. *Cooperation of Industrial Robots with Indoor-GPS*, Proceedings of the International Conference on Competitive Manufacturing (2010), s. 215-224.
5. Al-Yahmadi A.S., Abdo J., Hsia T.C. *Modeling and control of two manipulators handling a flexible object*, Journal of the Franklin Institute, 344 (5) (2007), s. 349-361.
6. Norman D.A., Draper S.W. *User centered system design: new perspectives on human-computer interaction*, L. Erlbaum Associates, 1986.
7. Muench S., Kreuziger J., Kaiser M., Dillman R. *Robot programming by demonstration (RPD)-Using machine learning and user interaction methods for the development of easy and comfortable robot programming systems*, Proceedings of the International Symposium on Industrial Robots, (1994), s. 685-685.
8. Billard A., Calinon S., Dillmann R., Schaal S. *Robot Programming by Demonstration*, [w:] *Springer Handbook of Robotics*. Springer Berlin Heidelberg, (2008), Berlin, Heidelberg.
9. Calinon S. *Robot programming by demonstration: a probabilistic approach*, EPFL Press, (2009).
10. Kahn K. *Generalizing by Removing Detail*, [w:] *Your Wish is My Command*. Elsevier, (2001), s. 21-43.
11. Somani N., Rickert M., Gaschler A., Cai C., Perzylo A., Knoll A. *Task level robot programming using prioritized non-linear inequality constraints*, 2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), IEEE, (2016), s. 430-437.
12. Lozano-Perez T. *Robot programming*, Proceedings of the IEEE, (1983), s. 821-841.
13. Foit K. *Introduction to solving task-level programming problems in logic programming language*, Journal of Achievements in Materials and Manufacturing Engineering, Vol. 64 (2) (2014), s. 78-84.
14. Foit K. *Task-level programming of a robotic system using high level logic language*, Selected Engineering Problems (6) (2012), s. 21-26.
15. Folscher D.J., Kruger K. *Saving time on robot programming: Programming by demonstration using stereoscopic motion capturing*, 2016 Pattern Recognition Association of South Africa and Robotics and Mechatronics International Conference (PRASA-RobMech), IEEE, (2016), s. 1-6.
16. Makris S., Tsarouchi P., Surdilovic D., Krüger J. *Intuitive dual arm robot programming for assembly operations*, CIRP Annals - Manufacturing Technology, 63 (1) (2014), s. 13-16.
17. *IVRE – An Immersive Virtual Robotics Environment*, online: https://cirl.lcsr.jhu.edu/research/human-machine-collaborative-systems/ivre/, dostęp 10.09.2017.
18. Mateo C., Brunete A., Gambao E., Hernando M. *Hammer: An Android based application for end-user industrial robot programming*, 2014 IEEE/ASME 10th International

Conference on Mechatronic and Embedded Systems and Applications (MESA), IEEE, (2014), s. 1-6.
19. *Intuitive Programming of Industrial Robots through Gestures and Augmented Reality*, Fachgebiet Industrielle Automatisierungstechnik, Technische Universität Berlin, online: YouTube, https://youtu.be/RPn4EnZ4V6M, dostęp 10.09.2017
20. Jan Y., Hassan S., Sanghun Pyo, Jungwon Yoon. *Smartphone Based Control Architecture of Teaching Pendant for Industrial Manipulators*, 2013 4th International Conference on Intelligent Systems, Modelling and Simulation, IEEE, (2013), s. 370-375.
21. *IPhone Industrial Robot Control – KUKA KR 6*, Fachgebiet Industrielle Automatisierungstechnik, Technische Universität Berlin, online: http://www.iat-tu.berlin.de, dostęp 10.09.2017
22. Foit K. *Visual Interface for Hybrid Programming of the Industrial Robot*, Advanced Materials Research, 1036 (2014), s. 743-748.
23. Norman A., Schönberg A., Gorlach I., Schmitt R. *Cooperation of Industrial Robots with Indoor-GPS*, Proceedings of the International Conference on Competitive Manufacturing (2010), s. 215-224.

Realizacja koncepcji Przemysłu 4.0 w ujęciu zagadnienia programowania robotów przemysłowych

Abstrakt
Filozofia „Przemysłu 4.0" znacząco zmienia podejście do procesu wytwarzania. Inteligentne fabryki, samoorganizujące się linie produkcyjne, przemysłowy Internet Rzeczy, czy w końcu zatarcie granic pomiędzy światem zewnętrznym a wnętrzem fabryki, stawia nowe wyzwania również przed operatorami robotów przemysłowych. Odejście od klasycznych form programowania robotów wymaga wypracowania nowych, efektywnych metod, które z jednej strony zapewnią prostotę programowania, z drugiej zaś pozostawią maszynie niezbędny margines swobodnego podejmowania decyzji, potrzebny do tworzenia dynamicznych struktur produkcyjnych oraz samoorganizacji. Niniejsza praca porusza aspekt rozwoju technik programowania robotów w kontekście filozofii „Przemysłu 4.0", przedstawiając zarówno wczesne osiągnięcia w tej dziedzinie, jak i współczesne trendy, wykorzystujące takie techniki jak wirtualna rzeczywistość czy Internet Rzeczy.
Słowa kluczowe: Przemysł 4.0, programowanie robotów, interfejsy programistyczne

The implementation of the „Industry 4.0" concept in terms of programming of industrial robots

Abstract
The philosophy of "Industry 4.0" significantly changes the approach to the manufacturing process. Smart factories, self-organizing production lines, Industrial Internet Things, blurring the boundaries between the outside world and the factory interior, present new challenges to the industrial robot operators. The abandon of the classic forms of robots programming, requires to develop new, effective methods that will – on the one hand – ensure the simplicity of programming, but also leave the necessary margin of machine independency that is needed to create dynamic production structures and to keep the ability of self-organization. This paper addresses the problem of the development of robot programming techniques in the context of the "Industry 4.0" philosophy, presenting both early achievements in this field and contemporary trends that are using such techniques as virtual reality or the Internet of Things.
Keywords: Industry 4.0, robots programming, programming interfaces

Piotr Löschner[1], Krzysztof Jarosz[2]

Przegląd wybranych programów MES do komputerowej symulacji procesów skrawania

1. Wstęp

Początkowo Metoda Elementów Skończonych (MES, *ang. Finite Element Method (FEM)*) wykorzystywana była do rozwiązywania prostych zagadnień matematycznych. Dzięki wzrostowi mocy obliczeniowej komputerów możliwe było poszerzenie obszaru zastosowania MES do rozwiązywania nieliniowych problemów z uwzględnieniem wielu czynników zmiennych w czasie, a także zagadnień o dużo większej skali. Dzięki zastosowaniu metod numerycznych w zagadnieniach związanych z mechaniką procesu skrawania możliwy był rozwój modeli symulacyjnych różnych operacji obróbkowych. Na podstawie metody elementów skończonych (MES), metody elementów brzegowych (MEB), czy metody różnic skończonych (MRS) tworzy się szereg dwu- oraz trójwymiarowych symulacji komputerowych różnych procesów statycznych, dynamicznych (i nie tylko) o coraz lepszej zgodności wyników z danymi pozyskanymi eksperymentalnie [1].

Metoda elementów skończonych jest obliczeniową metodą polegającą na oszacowaniu poszczególnych wartości charakterystycznych w lokalnych miejscach modelu (tzw. węzłach) według odpowiednich wzorów matematycznych. Bazuje ona na modelu geometrycznym, warunkach brzegowych i innych danych wejściowych potrzebnych do przeprowadzenia symulacji. W procesie dyskretyzacji model zamieniany na zbiór węzłów, które połączone ze sobą tworzą tak zwane elementy skończone. Elementy skończone to proste figury geometryczne płaskie lub przestrzenne, w których punktach końcowych znajdują się węzły. Węzły mogą znajdować się również na powierzchniach bocznych, a także wewnątrz elementu skończonego. Węzeł odpowiada punktowi, w którym przeprowadzone zostają obliczenia numeryczne według przyjętego modelu konstytutywnego (modelu materiałowego). Na podstawie modelu materiałowego, warunków brzegowych, itp. obliczone mogą być m.in. takie wartości jak: temperatura czy przemieszczenie w danym węźle oraz pochodne tych wartości jak np. składowy stan naprężenia i odkształcenia [2].

Metoda Elementów Skończonych jest szeroko wykorzystywana do komputerowej symulacji oraz analizy szeregu zjawisk i procesów związanych z obróbką skrawaniem. Szczególna popularność zdobyła w symulacjach procesu formowania wióra, pozwalając na lepsze zrozumienie skomplikowanych i wzajemnie powiązanych zjawisk zachodzących w strefie skrawania. Metoda Elementów Skończonych jest łatwiejsza i mniej czasochłonna do implementacji od skomplikowanych metod analitycznych, między innymi z powodu dostępności oprogramowań MES wyposażonych w graficzny interfejs użytkownika. Komputerowe symulacje procesów

[1] pleszner@op.pl, Katedra Technologii Maszyn i Automatyzacji Produkcji, Wydział Mechaniczny, Politechnika Opolska
[2] jarosz.krzysztof91@gmail.com, Katedra Technologii Maszyn i Automatyzacji Produkcji, Wydział Mechaniczny, Politechnika Opolska

skrawania mogą być wykonywane w jednym z wielu dostępnych na rynku programów MES. Dostępne są zarówno programy o szerokich możliwościach symulowania szeregu zjawisk, nie tylko tych związanych z obróbką skrawaniem, a także rozwiązania skupiające się na możliwości symulacji wyłącznie procesów związanych z obróbką wiórową. W zależności od wykorzystywanego programu spodziewać należy się różnic w interfejsie, intuicyjności i łatwości przygotowania symulacji, a przede wszystkim w jakości i zgodności otrzymanych wyników w porównaniu ze stanem rzeczywistym.

Głównymi zaletami skutkującymi coraz szerszym stosowaniem MES w obróbce skrawaniem są [1]:

- możliwość kompleksowego modelowania materiałów narzędziowych oraz przedmiotów obrabianych z uwzględnieniem wpływu szeregu czynników takich jak chociażby prędkość odkształcenia, właściwości sprężysto-plastyczne, czy temperatura w strefie skrawania,
- uwzględnianie zróżnicowanych modeli tarcia w celu dokładnego modelowania oddziaływań wiór-ostrze skrawające,
- możliwość badania zjawisk zachodzących podczas procesu formowania wióra,
- uzyskiwanie informacji o warunkach procesu takich jak przykładowo: składowe siły skrawania, postać wióra, odkształcenie i naprężenie w strefie skrawania, temperatura w strefie skrawania,
- uwzględnienie warunków brzegowych modelu i określenia ich wpływu na przebieg procesu skrawania,
- alternatywa dla kosztownych, czasochłonnych i trudnych w realizacji badań eksperymentalnych.

Do wykonania symulacji MES procesu skrawania niezbędne jest wprowadzenie szeregu danych wejściowych, takich jak:

- model geometryczny narzędzia i przedmiotu obrabianego przygotowany w zewnętrznym programie CAD, bądź zamodelowany bezpośrednio w programie MES,
- szereg danych materiałowych – należy przyjąć odpowiedni model konstytutywny uwzględniający możliwość dekohezji materiału,
- warunki brzegowe (utwierdzenia, przemieszczenia, prędkości),
- siatka elementów skończonych (ilość, rodzaj elementów skończonych i ich cechy).

2. Programy do symulacji obróbki skrawaniem Metodą Elementów Skończonych

W niniejszym rozdziale przedstawione zostały wybrane przez autorów na podstawie przeglądu cytowanej literatury programy MES umożliwiające przeprowadzenie symulacji procesów obróbki skrawaniem. Celem pracy było wskazanie odpowiedniego, w zależności od potrzeb użytkownika końcowego, programu MES do symulacji procesów obróbki skrawaniem.

2.1. DEFORM 2D/3D

Pierwszym z omawianych programów do komputerowej symulacji procesów skrawania jest oprogramowanie DEFORM 2D/3D firmy Scientific Forming Technologies Corporations. Jest to komercyjny pakiet oprogramowania MES zaprojektowany z myślą o symulacji procesów obróbkowych, takich jak procesy

obróbki bezubytkowej (przykładowo kucie swobodne i matrycowe, walcowanie, wyciskanie, ciągnienie) czy ubytkowej (między innymi toczenie, wiercenie, przeciąganie, rozwiercanie, dłutowanie). W związku z tym szczególny nacisk został położony na możliwość symulowania znacznych odkształceń, naprężeń oraz temperatur. DEFORM 2D/3D posiada możliwość współpracy ze skryptami przygotowanymi przez użytkowników końcowych, a także umożliwia definiowanie przez użytkownika własnych zmiennych i procedur [3].

Oprogramowanie Deform 3D zostało wyposażone w funkcję Automatic Mesh Generator przeznaczoną do lokalnego zagęszczania siatki elementów skończonych w miejscach, gdzie jest to wymagane ze względu na występowanie znacznych odkształceń, naprężeń itp., pozwalając na jednoczesne zachowanie relatywnie niskiej gęstości siatki w pozostałych obszarach. Pozwala to na zwiększenie dokładności obliczeń bez nadmiernego wydłużania ich czasu. Oprogramowanie posiada również preprocesor pomagający użytkownikowi końcowemu zdefiniować warunki brzegowe odpowiednie dla symulowanego rodzaju obróbki. W przypadku potrzeby symulacji bardziej złożonych procesów istnieje możliwość ręcznego wprowadzenia warunków brzegowych przez użytkownika programu. Możliwości systemu obejmują między innymi modelowanie składowych siły skrawania, postaci wióra, temperatury w strefie skrawania oraz naprężeń szczątkowych po obróbce oraz wynikających z nich odkształceń obrobionego przedmiotu [3].

2.2. AdvantEdge

Kolejnym z zaprezentowanych programów do komputerowej symulacji obróbki skrawaniem jest oprogramowanie firmy Third Wave Systems o nazwie AdvantEdge FEM. Interfejs programu przedstawiony został na rys. 1. Program ten służy do modelowania procesów obróbki skrawaniem Metodą Elementów Skończonych w płaskim bądź przestrzennym układzie współrzędnych. Składa się on z trzech podstawowych modułów [4]:

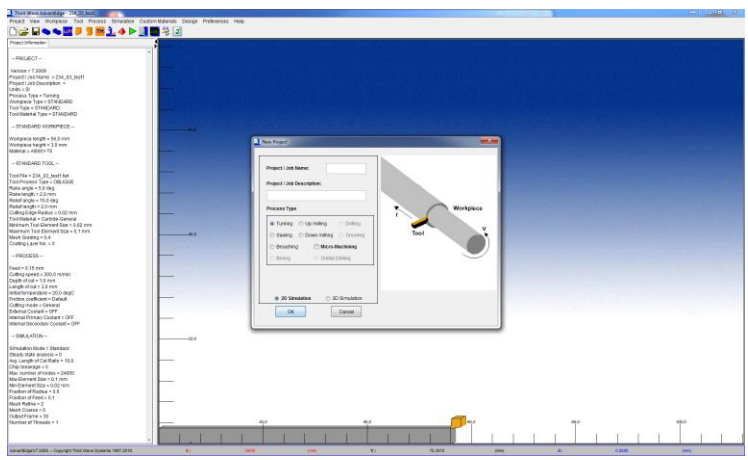
- moduł służący do zdefiniowania modelu geometrycznego, materiałowego przedmiotu oraz do wprowadzenia parametrów technologicznych – (dane wejściowe),
- moduł odpowiadający za proces obliczeniowy symulacji (obliczenia numeryczne),
- moduł, w którym przedstawione zostały wyniki obliczeń oraz służący do ich analizy.

Wszystkie trzy moduły są bardzo łatwe i intuicyjne w obsłudze. Podczas wprowadzania danych wejściowych wszystkie parametry, które należy wprowadzić są przedstawione na odpowiednich grafikach (rys. 1), co ułatwia użytkownikowi końcowemu ich właściwą identyfikację, a więc również poprawne wprowadzenie.

Jak zostało wspomniane w rozdziale 1, aby przeprowadzić symulację MES należy wprowadzić tzw. dane wejściowe. Ogromną zaletą oprogramowania AdvantEdge FEM jest fakt, iż do tego celu został stworzony moduł definiowania danych wejściowych i parametrów procesu. Dzięki zastosowaniu owego modułu w łatwy sposób można wprowadzić parametry dla procesów: toczenia, frezowania, przeciągania, wiercenia, gwintowania i wytaczania [5].

Definicja geometrii przedmiotu w programie AdvantEdge FEM odbywa się na zasadzie wyboru geometrii przedmiotu (przykładowo wałek o określonej średnicy)

bądź samodzielnej definicji poprzez zamodelowanie geometrii przygotówki według uznania użytkownika końcowego. Natomiast definiowanie geometrii narzędzia polega na wyborze gotowego modelu z obszernej biblioteki programu. Dodatkowo istnieje możliwość samodzielnego określenia kształtu ostrza. Po zdefiniowaniu kształtu narzędzia i przedmiotu obrabianego program prosi użytkownika o wpisanie w odpowiednie pola parametrów technologicznych modelowanego procesu. Definicja modelu materiałowego polega na wyborze materiału z bogatej biblioteki zintegrowanej z programem, natomiast wybór modelu konstytutywnego został ograniczony do z góry zdefiniowanych w programie modeli Power Law, Power Law – Temperature Dependent oraz modelu Drucker-Prager.

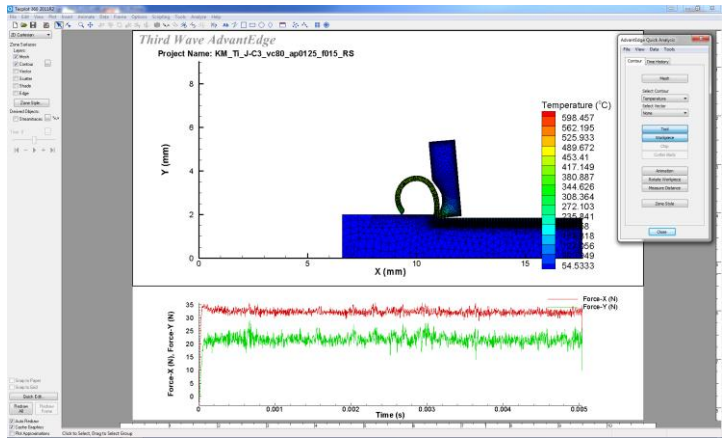


Rysunek 3. Interfejs program AdvantEdge wraz z widocznym oknem wyboru rodzaju obróbki [opracowanie własne]

W module obliczeń numerycznych definiowane są parametry siatki MES takie jak liczba węzłów, gęstość i kształt siatki, warunki brzegowe układu, własności związane z łamaniem wióra itp. Warto tu nadmienić, iż w programie AdvantEdge możliwe jest prowadzenie obliczeń numerycznych w dwóch trybach, mianowicie „Standard" oraz „Rapid". W trybie „Rapid" możliwe jest przeprowadzenie obliczeń numerycznym w dużo krótszym czasie niż w trybie „Standard". Czas obliczeń w trybie przyspieszonym może być nawet pięciokrotnie krótszy względem trybu standardowego. Tryb „Rapid" jest wykorzystywany głównie do weryfikacji poprawności budowy modelu materiałowego, obliczeniowego oraz do weryfikacji poprawności wprowadzenia danych materiałowych. Możliwość wyboru trybu obliczeń jest sporą zaletą AdvantEdge FEM, gdyż pozwala na oszczędność czasu oraz łatwą weryfikację poprawności wprowadzonych danych wejściowych. Tryb „Rapid" rzadko wykorzystywany jest do obliczeń końcowych ze względu na obarczenie obliczeń błędem sięgającym średnio ±15% względem wyników eksperymentalnych [4].

Do wizualizacji oraz analizy otrzymanych wyników służy zintegrowany z systemem AdvantEdge FEM moduł Tecplot przedstawiony na rys. 2. Jest to narzędzie służące do prezentacji otrzymanych wyników w formie graficznej na obrazach, wykresach 2D i 3D oraz animacjach. Główne zmienne możliwe do

przedstawienia w formie graficznej to temperatura, strumień ciepła generowany przy odkształceniu plastycznym, odkształcenie plastyczne oraz jego prędkość, naprężenia zredukowane, naprężenia główne i styczne oraz składowe X,Y, Z naprężeń [4].

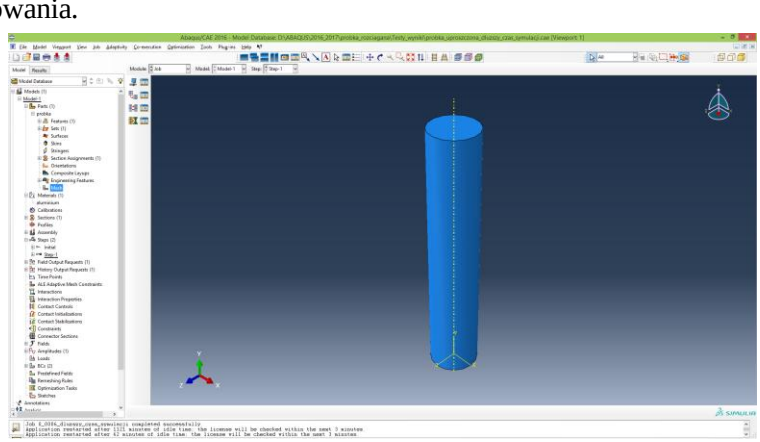


Rysunek 4. Widok modułu Tecplot programu AdvantEdge do wizualizacji i analizy wyników [opracowanie własne]

2.3. ABAQUS/CAE

Ostatnim z opisywanych programów do MES jest propozycja firmy Dassault Sytemes – oprogramowanie Abaqus/CAE. Okno programu przedstawione zostało na rys. 3. Jest to wysoce zaawansowany program komputerowy przeznaczony nie tylko do modelowania procesu skrawania. Znajduje on również zastosowanie przy rozwiązywaniu szeregu innych, zróżnicowanych zagadnień, przykładowo:

- w budownictwie do testowania wytrzymałości mostów, podpór, itp.,
- w lotnictwie do przepływów aerodynamicznych,
- w branży samochodowej do testów zderzeniowych,
- w mechanice płynów do symulacji przepływów w wymiennikach ciepła,
- w przemyśle wytwórczym, między innymi do modelowania procesów kucia, walcowania.



Rysunek 5. Okno programu Abaqus [opracowanie własne]

Oprogramowanie Abaqus oferuje możliwość kompleksowego i wydajnego rozwiązywania wielu zagadnień inżynierskich w jednym programie. Ważną cechą tego oprogramowania jest to, iż system ten nie posiada żadnych modułów/pakietów do symulacji. Pozwala to na modelowanie procesu na wysokim poziomie szczegółowości. Jednakże użytkownik musi sam jasno i precyzyjnie określić wszystkie cechy jakie powinien posiadać przedmiot oraz narzędzie (w tym warunki brzegowe, siatka elementów skończonych, własności materiału oraz parametry procesu), co zwiększa poziom trudności obsługi programu. Abaqus/CAE wyposażony jest w dwa solvery (Standard i Explicit), znajdujące zastosowanie w zależności od rodzaju przygotowywanej symulacji. Możliwe jest modelowanie zarówno nieskomplikowanych analiz liniowych i termicznych, jak również wysoce zaawansowane i rozwinięte nieliniowe analizy dynamiczne oraz różnego rodzaju sprzężone analizy termiczno-dynamiczne i obliczenia akustyczne. Znacząco rozwinięta funkcjonalność oprogramowania, a przede wszystkim implementacja postprocedur użytkownika pozwala realizować skomplikowane analizy i symulacje komputerowe zbliżone do realnych procesów. Znacząco rozwinięte narzędzia oprogramowania Abaqus/CAE pozwalają na definiowanie wielu zmiennych, a także umożliwiają implementację własnych praw, modeli, własności materiałowych itp. Oprogramowanie firmy Dassault Systemes ma prawie nieograniczone możliwości wspomagania różnego rodzaju prac badawczych. Narzędzia oprogramowania Abaqus pozwalają na wprowadzenie unikalnych lepko-plastycznych praw konstytutywnych, obciążeń, modeli tarcia, a także różnego rodzaju własności termicznych i elektromagnetycznych. Warto tu nadmienić, iż Abaqus/CAE nie posiada własnej bazy danych materiałowych. W związku z tym użytkownik musi posiadać wiedzę na temat modeli materiałowych, działania oprogramowania MES, a także wykazać się znajomością sposobu wprowadzania owych danych wejściowych do programu. Ważnym etapem podczas procesu modelowania jest tworzenie siatki elementów skończonych. W zależności od modelowanego zjawiska oraz pożądanej dokładności wyników przyjąć można różnorakie kryteria zagęszczenia siatki. Może to być lokalne zagęszczenie siatki (przykładowo w obszarze gdzie zachodzą najintensywniejsze odkształcenia), globalne zagęszczenie siatki MES (gdy wymagane jest wysoka dokładność obliczeń dla całego modelu) oraz adaptacyjne zagęszczenie siatki (realizowane automatycznie przez oprogramowanie według zadanych przez użytkownika kryteriów). Przed dyskretyzacją modelu geometrycznego należy podać podstawowe informacje dotyczące siatki elementów skończonych, tj.:

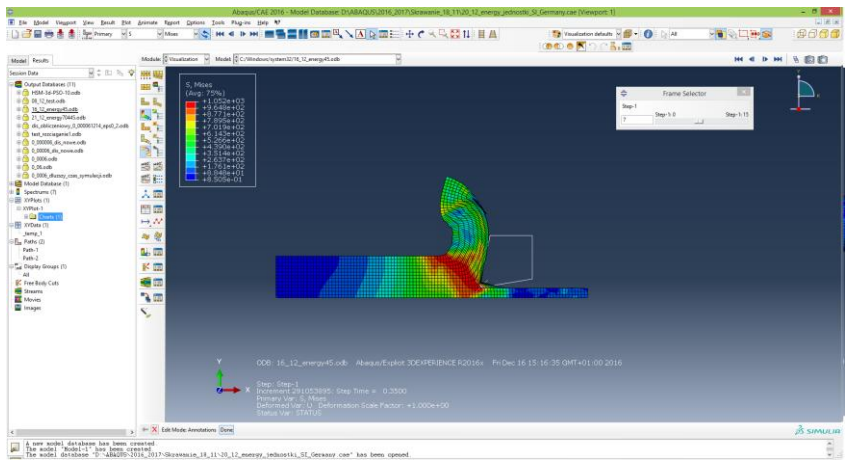
- wielkość elementów skończonych,
- typ elementu skończonego (trójkątne, czworokątne, itp.),
- właściwości siatki elementów skończonych.

W najmniej złożonych przypadkach wielkość każdego elementy skończonego jest taka sama. Istnieje również możliwość wprowadzenia siatki o skomplikowanej budowie poprzez zastosowanie dyskretyzacji z użyciem formuły Lagrange'a i Eulera. Takie rozwiązanie jest przydatne przy modelowaniu zjawisk związanych z obróbką ubytkową, podczas której zachodzą intensywne odkształcenia z wysokim prędkościami.

Modele materiałowe, które wykorzystywane są obecnie uwzględniają szereg danych, takich jak: wpływ wartości prędkości poślizgu, nacisków, temperatury i innych czynników na wartość współczynnika tarcia dla pary materiał obrabiany-narzędzie skrawające [6-8]. Pozwalają również na dokładniejsze przewidywanie wartości składowych całkowitej siły skrawania oraz sposób odkształcenia materiału

w trakcie obróbki [9, 10], a także na symulowanie zjawisk cieplnych występujących w trakcie procesu skrawania [11].

Po przeprowadzeniu symulacji procesu obróbki możliwa jest wizualizacja otrzymanych wyników. Na rys. 4 zaprezentowane zostało okno programu Abaqus służące do wizualizacji i analizy wyników. Wyniki mogą być zaprezentowane w formie wykresów bądź wizualizacji wybranych parametrów np. naprężenia, temperatury, odkształcenia, przemieszczenia, itp. Możliwa jest ich zaawansowana obróbka oraz zaprezentowanie przykładowo temperatury przedmiotu na przekroju geometrii oraz w dowolnym innym punkcie – w węźle wybranym przez użytkownika [12].



Rysunek 6. Widok okna programu Abaqus do wizualizacji i analizy wyników [opracowanie własne]

3. Podsumowanie i wnioski

W tabeli 1. zawarto podsumowanie przeglądu wybranych programów przeznaczonych do symulacji MES obróbki skrawaniem w formie ich wzajemnego porównania na podstawie wyszczególnionych w tabeli kryteriów.

Wybór odpowiedniego spośród zaprezentowanych programów jest przede wszystkim uzależniony od oczekiwań użytkownika końcowego. Jeżeli głównym kryterium jest łatwość obsługi oraz zadowalająca dokładność symulacji, zaś użytkownik planuje symulować jedynie procesy obróbki ubytkowej, powinien on wybrać oprogramowanie AdvantEdge FEM. W przypadku, gdy potrzebna jest dodatkowo funkcja modelowania procesów przeróbki plastycznej oraz wsparcie dla zewnętrznych skryptów i procedur, odpowiednim wyborem zdaje się DEFORM 2D/3D, będący systemem o charakterze półotwartym, jednocześnie nadal pozostając prostym w obsłudze. Najbardziej zaawansowanym, lecz również najtrudniejszym w obsłudze jest program Abaqus/CAE. Jego niezaprzeczalnymi zaletami są: wszechstronność zastosowania oraz możliwość pełnej ingerencji w warunki brzegowe, modele konstytutywne, a nawet samo algorytm MES. Konieczność ręcznego wprowadzania wszystkich parametrów modelu materiałowego, warunków brzegowych oraz pozostałych danych wejściowych czyni go jednak narzędziem przeznaczonym dla zaawansowanych użytkowników.

Tabela 2. Porównanie opisanych w pracy programów MES

Kryterium oceny	**DEFORM 2D/3D**	**AdvantEdge**	**Abaqus/CAE**
Sposób przygotowania symulacji	Preprocesor zaprojektowany w celu ułatwienia użytkownikowi wprowadzania warunków brzegowych	Przygotowanie symulacji w poszczególnych krokach przy pomocy wbudowanego kreatora symulacji	Ręczna definicja warunków brzegowych oraz geometrii przedmiotu i narzędzia. Program o wysokim stopniu złożoności
Implementacja modelu materiałowego	Biblioteka zdefiniowanych modeli materiałowych z możliwością edycji właściwości materiału	Wybór materiału z obszernej biblioteki programu wraz z edytorem właściwości materiału	Konieczność ręcznej definicji modelu materiałowego w oparciu o szeroki wachlarz parametrów. Wsparcie dla modeli materiałowych definiowanych przez użytkownika (procedury VUMAT)
Siatka MES	Siatka adaptacyjna z możliwością sterowania parametrami adaptacyjności w ograniczonym zakresie	Adaptacyjna gęstość siatki w obszarach identyfikowanych przez program jako krytyczne bez możliwości ingerencji użytkownika	Brak adaptacyjnej gęstości siatki dla symulacji typu Explicit (dynamicznych). Opcje kontroli jakości siatki (Distortion Control/ALE) bez remeshingu
Możliwość ingerencji w procedurę symulacji	Możliwość wprowadzania podstawowych zmian w procedurze symulacji. Niemodyfikowalny algorytm MES	Minimalne możliwości ingerencji – brak możliwości modyfikacji algorytmu MES	Pełna ingerencja we wszystkie aspekty symulacji wraz z modyfikacją bazowego algorytmu MES

Źródło: [13, 14]

Literatura

1. Niesłony, P., Grzesik, W. *Modelowanie procesu i operacji skrawania metodą elementów skończonych (MES) Cz. I*. Podstawy i programy symulacyjne., Mechanik., 10 (2013), s. 825-832.
2. Niesłony, P., Grzesik, W. *Modelowanie procesu i operacji skrawania z zastosowaniem metody elementów skończonych (MES). Cz. II.* Przykłady zastosowań praktycznych., Mechanik., 11 (2013), s. 909-917.
3. http://www.deform.com/.
4. Niesłony P. *Modelowanie skrawania z zastosowaniem MES w oparciu o pakiet AdvantEdge*, Stal Metale & Nowe Technologie., 5-6 (2012), s.80-85.
5. http://www.thirdwavesys.com/advantedge/.
6. Zemzemi F., Rech J., Ben Salem W., Dogui A., Kapsa P. *Identification of a friction model at tool/chip/workpiece interfaces in dry machining of AISI4142 treated steels*, Journal of Materials Processing Technology., 209 (2009) s. 3978-3990.
7. Rech J., Claudin C., D'Eramo E., *Identification of a friction model – Application to the context of dry cutting of an AISI 1045 annealed steel with a TiN-coated carbide tool*, Tribology International., 42 (2009) s. 738-744.
8. Courbon C., Mabrouki T., Rech J., Mazuyer D., D'Eramo E., Daguier P., *On a local contact definition in dry cutting simulations of AISI 1045 Steel*., Eighth International Conference on High Speed Machining, 8-10 Grudzień 2010, s. 38-43.

9. Mabrouki T., Rigal J.-F., *A contribution to a qualitative understanding of thermal-mechanical effects during chip formation in hard turning.*, Journal of Materials Processing Technology., 176 (2006) s. 214-221.
10. Kortabarria A., Rech J., Ruiz de Eguilaz E., Arrazola P. J., *A new Vuinter subroutine for Abaqus/Explicit™ to modeling rate dependent surface interactions laws in machining*., The 14th International ESAFORM Conference on material forming: ESAFORM 2011, 27–29 Kwieceń 2011 Belfast
11. Zhang Y.C., Mabrouki T., Nelias D., Gong Y.D., *Chip formation in orthogonal cutting considering interface limiting shear stress and damage evolution based on fracture energy approach*., Finite Elements in Analysis and Design., 7 (2011), s. 850-863.
12. https://www.3ds.com/products-services/simulia/products/abaqus/
13. Niesłony P. *Modelowanie przepływu ciepła i rozkładu temperatury w strefie skrawania dla ostrzy z twardymi powłokami ochronnymi*., Oficyna Wydawnicza Politechniki Opolskiej., 2008.
14. Gardner J. D., Vijayaraghavan A., Dornfeld D. A., *Comparative study of finite element simulation software., Laboratory for Manufacturing and Sustainability*., (2005).

Przegląd wybranych programów MES do komputerowej symulacji procesów skrawania

Abstrakt
W pracy przedstawione zostały wybrane programy MES (Metoda Elementów Skończonych), w których możliwe jest przeprowadzenie symulacji numerycznych procesów skrawania. W pierwszej części pracy zawarte zostały podstawowe informacje dotyczące problematyki symulacji komputerowych procesów skrawania. W dalszej części pracy opisane zostały wybrane przez autorów programy MES, wraz z ich wzajemnym porównaniem. Opisy poruszały tematy takie jak: sposób przygotowania numerycznej symulacji procesu toczenia wraz ze sposobem wprowadzania danych wejściowych oraz możliwe do uzyskania wyniki. W podsumowaniu zawarto wzajemnie porównanie opisanych wcześniej programów według określonych kryteriów, takich jak między innymi możliwości oprogramowania, łatwość i intuicyjność obsługi, sposób wprowadzania danych wejściowych. Podjęto próbę wyłonienia słabych i mocnych stron każdego z opisywanych programów, oraz ustalenia który z programów będzie najlepiej odpowiadał poszczególnym użytkownikom, w zależności od ich potrzeb.
Słowa kluczowe: MES, AdvantEdge, Deform, Abaqus, skrawanie

Overview of select software for FEM simulation of machining processes

Abstract
In this paper, select examples of software applicable for FEM simulation of machining processes were described and compared. In the first part of the work, basic information about FEM simulation of machining processes was provided. In the next part, three examples of commercial FEM software were described and compared. Descriptions of software contained topics such as: procedure of preparing the simulations, ways of entering input data and achievable results. The summary contained a comparison of previously described software, which were evaluated in accordance to a number of criteria. The authors tried to decide which software is most suitable for end users basing on their expectations and needs.
Keywords: FEM, AdvantEdge, Deform, Abaqus, machining

Małgorzata Franz[1]

Detekcja materiałów przy użyciu organicznych tranzystorów polowych

1. Wstęp

Ważny obszar aktualnych badań naukowych i technologicznych stanowią poszukiwania szybkich i efektywnych sposobów wykrywania substancji, zarówno organicznych jak i nieorganicznych [1-10]. Jednym z rodzajów czujników, który może być stosowany do wykrywania związków gazowych i ciekłych są organiczne tranzystory polowe (OFET – *Organic Field-Effect Transistor*). Czujniki zbudowane na bazie organicznych tranzystorów polowych działają w temperaturze pokojowej, co prowadzi do mniejszego zużycia energii i tym samym daje możliwość ich zastosowania w urządzeniach przenośnych, podobnie jak czujników bazujących na półprzewodnikach nieorganicznych. OFET mogą być wytwarzane na giętkich, lekkich podłożach przy wykorzystaniu technik drukarskich nanoszenia warstw z roztworów, co w efekcie umożliwia produkcję tanich, jednorazowych urządzeń o dowolnych kształtach i rozmiarach. Ważną cechą materiałów organicznych jest możliwość znacznej modyfikacji ich właściwości, co pozwala na wytworzenie warstwy aktywnej czujnika OFET, wrażliwej na dany analit. Ponadto biokompatybilność materiałów organicznych z żywymi organizmami umożliwia ich wykorzystanie w biosensorach. Czujniki typu OFET mogą również stanowić elementy składowe takich zaawansowanych urządzeń jak elektroniczny nos, czy elektroniczna skóra. Prace nad zastosowaniem OFET w czujnikach gazów, światła, ciśnienia, substancji chemicznych i biochemicznych stanowią pole badawcze nie tylko dla elektroniki organicznej, ale również inżynierii materiałowej biofizyki czy technologii. Tego typu detektory są niezwykle tanie oraz wszechstronne i należy oczekiwać, że, w przyszłości zastosowanie OFET do detekcji może stać się znaczące w różnych dziedzinach techniki, medycyny, analizy środowiska oraz w obronności. W rozdziale przedstawione zostaną zasady działania tego typu detektorów oraz ich potencjalne możliwości aplikacyjne.

2. Organiczny tranzystor polowy

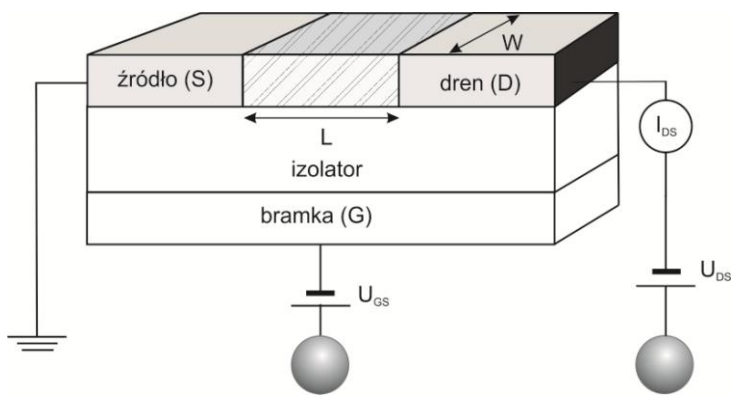
Tranzystor polowy (FET – *Field Effect Transistor*) jest unipolarnym przyrządem aktywnym, którego koncepcja została wprowadzona przez Lilienfeld'a w połowie lat dwudziestych ubiegłego stulecia [11]. Zasada działania tego urządzenia opiera się na kontroli przepływu prądu elektrycznego pomiędzy dwiema elektrodami poprzez przyłożenie pola elektrycznego do trzeciej elektrody zwanej bramką. Projekt Lilenfeld'a nie został zrealizowany w praktyce ze względu na brak odpowiedniej technologii, a pierwszy działający tranzystor polowy z izolowaną bramką (MOSFET – *Metal-Oxide-Semiconductor FET*) oparty na krzemie został wprowadzony w latach

[1] malobaro@pg.gda.pl, Katedra Fizyki Zjawisk Elektronowych, Wydział Fizyki Technicznej i Matematyki Stosowanej, Politechnika Gdańska, www.pg.edu.pl

60-tych ubiegłego stulecia przez Kahng'a i Atalla [12]. Wynalezienie MOSFET spowodowało przełom w elektronice, bowiem tranzystor zastąpił lampy elektronowe dając początek coraz większej miniaturyzacji urządzeń elektronicznych. Kolejnym ważnym wydarzeniem w rozwoju elektroniki było zastosowanie do budowy tranzystorów materiałów organicznych takich jak związki małocząsteczkowe [13] i polimery [14]. Początkowo tranzystory OFET posiadały dużo gorsze parametry od klasycznych tranzystorów nieorganicznych. Intensywny rozwój elektroniki molekularnej sprawił, że obecnie parametry OFET są zbliżone do tych wytwarzanych na amorficznym krzemie [15], a jedynym problemem, który spowalnia ich komercjalizację jest niska stabilność operacyjna [16] i wytrzymałość na czynniki zewnętrzne [17].

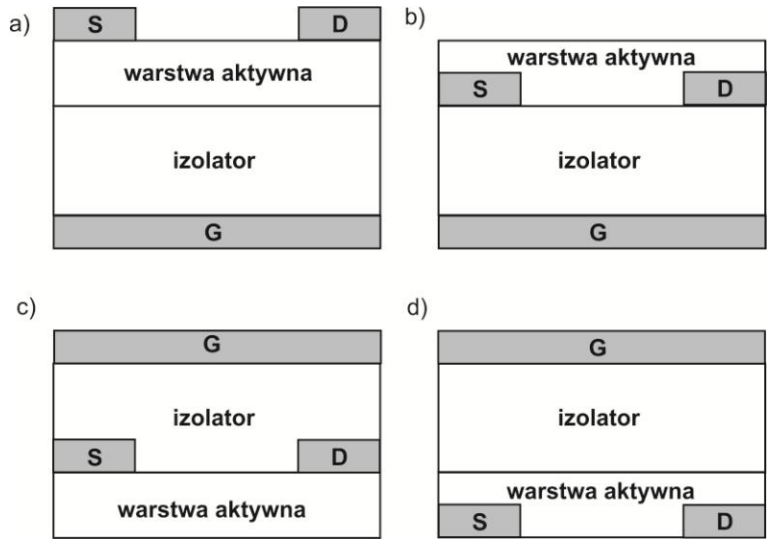
2.1. Budowa i zasada działania organicznego tranzystora polowego

Organiczny tranzystor polowy, którego schemat został przedstawiony na rys. 1 [18], zbudowany jest z warstwy materiału organicznego rozciągającej się od elektrody źródła (S – source) do elektrody zbiorczej/drenu (D – drain), oddzielonej od elektrody bramkującej (G – gate) warstwą izolatora. Szerokość elektrod źródła i drenu (W) oraz odległość między nimi (L) stanowią odpowiednio szerokość i długość kanału przewodzącego urządzenia.



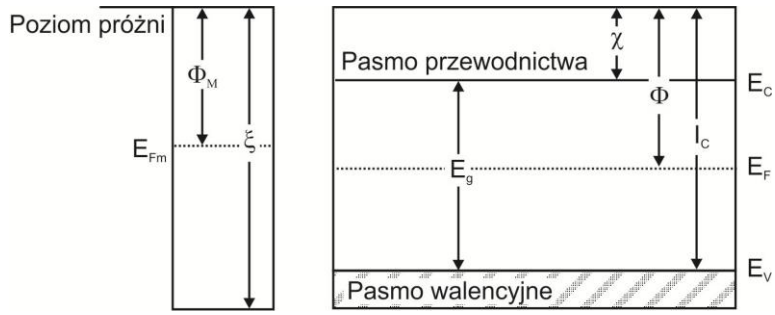
Rysunek 1. Schemat budowy organicznego tranzystora polowego [opracowanie własne]

Własności fizyczne zastosowanych w OFET materiałów, rodzaj podłoża, proces wytwarzania, a także przeznaczenie tranzystora mogą narzucać wybór odpowiedniej struktury warstwowej. Ze względu na sposób położenia warstwy półprzewodnika organicznego względem elektrod wyróżnia się cztery podstawowe konfiguracje budowy tranzystora polowego, które przedstawiono na rys. 2 [19].



Rysunek 2. Typowe struktury warstwowe stosowane przy wytwarzaniu tranzystorów polowych typu: a) TC BG (elektrody S i D na górze warstwy aktywnej, bramka ma dole); b) BC BG (elektrody S i D u dołu warstwy aktywnej, bramka na dole); c) TC TG (elektrody S i D na górze warstwy aktywnej, bramka na górze); d) BC TG (elektrody S i D u dołu warstwy aktywnej, bramka na górze) [opracowanie własne]

Zwykle zewnętrzny potencjał elektryczny przykładany jest do elektrody bramkującej (V_G) oraz elektrody zbiorczej (V_D), a elektroda źródła pozostaje uziemiona ($V_S = 0$). Różnica potencjałów pomiędzy elektrodą bramkującą, a źródłem jest określana, jako napięcie bramki (U_{GS}), natomiast różnica potencjałów pomiędzy elektrodą zbiorczą, a źródłem, jako napięcie drenu (U_{DS}). Napięcie przyłożone do bramki (U_{GS}) moduluje prąd płynący przez warstwę półprzewodnika organicznego (I_{SD}). Organiczne tranzystory polowe różnią się od ich nieorganicznych odpowiedników pochodzeniem ładunków biorących udział w transporcie przez kanał przewodzący. W przypadku FET zbudowanych z klasycznych półprzewodników nieorganicznych, prąd płynący od elektrody źródła do drenu stanowią nośniki ładunku generowane termicznie ze stanów domieszkowych w silnie domieszkowanych obszarach półprzewodnika stanowiących źródło i dren. Materiały organiczne, ze względu na dużą wartość przerwy energetycznej, mają niewielką koncentrację ładunków swobodnych, a głównym źródłem prądu w tych materiałach są ładunki wstrzykiwane z elektrod (głównie metalicznych) w wyniku tzw. iniekcji nośników ładunku [20]. Zatem prąd płynący w kanale przewodzącym tranzystora I_{SD} należy traktować, jako prąd iniekcyjny. Zrozumienie zjawiska iniekcji nośników ładunku z elektrody do półprzewodnika organicznego wymaga poznania układu ich charakterystycznych poziomów energetycznych. Rysunek 3 przedstawia diagram poziomów energetycznych półprzewodnika oraz elektrody metalicznej przed ich wzajemnym zetknięciem.



Rysunek 3. Schemat poziomów energetycznych półprzewodnika oraz elektrody metalicznej, tuż przed ich wzajemnym zetknięciem. Znaczenie symboli podano w tekście. Wartości energii mierzone są względem dna pasma przewodnictwa i poziomu próżni [opracowanie własne na podstawie [21]]

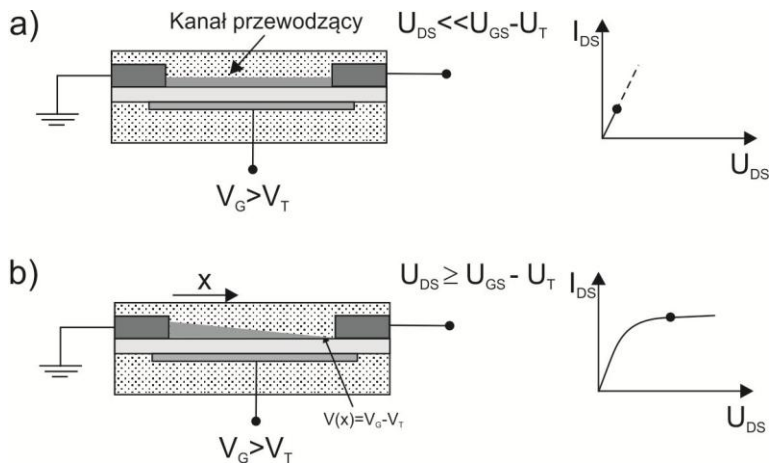
Na rysunku E_C oznacza energię pasma przewodnictwa, E_V energię pasma walencyjnego, E_g szerokość przerwy energetycznej, I_C energię jonizacji, χ powinowactwo elektronowe, natomiast ξ różnicę energii potencjalnych elektronu wewnątrz i na zewnątrz metalu. Linią przerywaną zaznaczono poziom energii Fermiego w metalu E_{Fm} i w półprzewodniku E_F. Symbolami Φ_M i Φ oznaczono pracę wyjścia odpowiednio dla elektrody metalicznej i półprzewodnika. Z rys. 3 wynika, że przeniesienie elektronu z elektrody do półprzewodnika wymaga dostarczenia mu energii o wartości: $\Phi_M - \chi$. Natomiast przeniesienie elektronu z półprzewodnika do metalu, co jest równoważne wstrzyknięciu dziury z elektrody do półprzewodnika, wymaga dostarczenia mu energii równej: $I_C - \Phi_M$. Rodzaj wstrzykiwanych z elektrody do materiału organicznego nośników ładunku zależy od położenia poziomu Fermiego w metalu względem pasm przewodnictwa i walencyjnego w materiale organicznym. Elektrony są, zatem wstrzykiwane z elektrody do półprzewodnika wtedy, gdy poziom Fermiego w metalu leży blisko dna pasma przewodnictwa. Natomiast iniekcja dziur z elektrody do materiału odbywa się wówczas, gdy poziom Fermiego znajduje się w pobliżu pasma walencyjnego. Modulacja wielkości prądu płynącego od źródła do elektrody zbiorczej wynika z faktu zmiany koncentracji nośników ładunku na złączu półprzewodnik/izolator, powodowanej zmianą napięcia bramki. Napięcie przyłożone do bramki wytwarza pole elektryczne w warstwie organicznej w kierunku prosto-padłym do kierunku przepływu prądu, wywołując rozdział nośników ładunku. Nośniki ładunku jednego znaku (dodatnie lub ujemne), zależnie od polaryzacji bramki, pod wpływem przyłożonego do bramki napięcia, przemieszczają się w kierunku górnej lub dolnej części warstwy, co utrudnia ich rekombinację powodując tym samym wzrost prądu w obwodzie: źródło – elektroda zbiorcza. Jest to zasadniczy mechanizm modulacji prądu w organicznym tranzystorze polowym. W przypadku, gdy potencjał elektrody bramki będzie dodatni na złączu izolatora z warstwą półprzewodnika organicznego będą gromadziły się elektrony. Natomiast, gdy potencjał bramki będzie ujemny na złączu izolator/półprzewodnik organiczny gromadzić się będą dodatnie nośniki ładunku (dziury). Ilość zgromadzonych na złączu nośników ładunku (N) zależy od wartości napięcia przyłożonego do elektrody bramkującej (U_{GS}) oraz pojemności układu (C) i może być obliczona na podstawie wzoru:

$$N = C \cdot U_{GS} / e \tag{1}$$

gdzie e oznacza ładunek elementarny.

Należy zaznaczyć, że nie wszystkie zgromadzone na złączu nośniki ładunku stanowią wkład do prądu płynącego przez warstwę przewodnika organicznego, gdyż część z nich zostaje spułapkowana. Stąd, aby w kanale przewodzącym nastąpił przepływ ładunku napięcie przyłożone do bramki musi być większe od tzw. napięcia progowego (U_T), a wartość napięcia efektywnego bramki jest, zatem równa U_{GS}-U_T.

Na rysunku 4 zilustrowano rozkład koncentracji nośników ładunku w kanale przewodzącym tranzystora oraz przedstawiono odpowiednie charakterystyki prądowo-napięciowe w zakresach pracy urządzenia: liniowym i nasycenia [19].

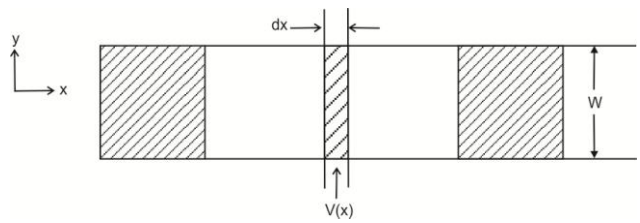


Rysunek 4. Schemat budowy organicznego tranzystora polowego wraz z rozkładem koncentracji nośników ładunku w kanale przewodzącym i odpowiednie charakterystyki prądowo-napięciowe w zakresach pracy urządzenia: a) liniowym i b) nasycenia [opracowanie własne na podstawie [19]]

W warunkach braku napięcia przyłożonego do elektrod źródło i dren (U_{DS} = 0), koncentracja nośników ładunku w kanale przewodzącym jest jednorodna, a natężenie prądu płynącego przez warstwę półprzewodnika organicznego równe zero. Niezbyt wysokie wartości napięć (U_{DS}<<U_G-U_T) zaburzają rozkład nośników ładunku w kanale sprawiając, że zależność prądu płynącego przez warstwę półprzewodnika od napięcia U_{DS} jest liniowa. Opisana sytuacja została przedstawiona na rysunku i wykresie 4a. W warunkach, gdy napięcie U_{DS} osiąga wartość U_{DS} = U_G-U_T, kanał tranzystora zostaje zamknięty, a w pobliżu elektrody zbiorczej tworzy się obszar zubożony w nośniki ładunku. Natężenie prądu nośników ładunku płynącego przez warstwę materiału organicznego osiąga wartość stałą, a dalszy wzrost napięcia między elektrodą źródło i dren nie powoduje jego wzrostu, a jedynie powiększa obszar zubożony w nośniki ładunku w pobliżu elektrody zbiorczej, skracając w ten sposób kanał przewodzący tranzystora. Rozkład koncentracji nośników ładunku w kanale przewodzącym tranzystora oraz charakterystyka prądowo napięciowa w zakresie nasycania zostały przedstawione na rys. 4b.

2.2. Charakterystyki prądowo-napięciowe oraz parametry pracy OFET

Natężenie prądu nośników ładunku płynącego przez warstwę materiału organicznego może zostać opisane analitycznie przy założeniu, że o wartościach prądu w układzie źródło-dren decyduje trzecia elektroda – bramka. Na rys. 5 przedstawiono schematycznie układ elektrod źródło i dren wraz z warstwą materiału organicznego, pozwalający określić prąd płynący w tranzystorze.



Rysunek 5. Schemat układu elektrod S i D wraz z warstwą materiału organicznego dla określenia prądu płynącego w kanale przewodzącym tranzystora (*opracowanie własne*)

Dla napięć przyłożonych do elektrody bramkującej, wyższych od napięcia progowego, ładunek zgromadzony w odległości x od elektrody źródła wyraża wzór:

$$Q = C_i \cdot \left(V_{GS} - V_T - V(x)\right) \tag{2}$$

gdzie V(x) oznacza potencjał pola elektrycznego w odległości x od elektrody źródła, natomiast C_i pojemność izolatora na jednostkę powierzchni.

Zaniedbując prąd dyfuzyjny nośników ładunku w kanale tranzystora, natężenie prądu dryfu może być zapisane, jako:

$$I_{DS} = W\mu QE(x), \tag{3}$$

gdzie E(x) oznacza natężenie pola elektrycznego w odległości x od elektrody źródła, natomiast μ ruchliwość nośników ładunku, zdefiniowaną, jako stosunek prędkości dryfu do natężenia pola elektrycznego $\left(\mu = \frac{\upsilon_d}{E}\right)$. Ruchliwość nośników ładunku jest ważnym parametrem określającym transport nośników ładunku. W szczególności jego zależność od temperatury bezwzględnej i natężenia pola elektrycznego jest ważnym kryterium weryfikacji mechanizmu ruchu nośników ładunku w molekularnym ciele stałym [22], który może być mechanizmem pasmowym bądź hoppingowym [23]. Wartości ruchliwości nośników ładunku w materiałach organicznych są zazwyczaj niższe od tych uzyskiwanych dla materiałów nieorganicznych, aczkolwiek dla nowych materiałów molekularnych mogą sięgać kilkadziesiąt $cm^2V^{-1}s^{-1}$ [24]. Należy jednak zaznaczyć, że na wartość tego parametru ma wpływ struktura chemiczna związku, proces przygotowania próbki, a przede wszystkim wybór doświadczalnej metody pomiaru.

Wykorzystując związek natężenia pola elektrycznego z potencjałem elektrycznym $(E = dV/dx)$ oraz wzór (2) prąd dryfu nośników ładunku może być zapisany w postaci:

$$I_{DS}dx = W\mu C_i\left(V_{GS} - V_T - V(x)\right)dV \tag{4}$$

Zakładając, że ruchliwość nośników ładunku jest stała i nie zmienia się w polu elektrycznym, po scałkowaniu powyższego równania w granicach od zera do wartości L dla wartości potencjału w zakresie od zera do V_{DS}, prąd płynący między źródłem i drenem możemy zapisać następującym wzorem:

$$I_{DS} = \frac{W\mu C_i}{L} V_{DS}\left(V_{GS} - V_T - \frac{V_{DS}}{2}\right) \tag{5}$$

Zróżniczkowanie powyższego równania pozwala znaleźć punkt nasycania prądu dryfu i po tej operacji otrzymujemy:

$$\frac{\partial I_{DS}}{\partial V_{DS}} = \frac{W\mu C_i}{L}\left(V_{GS} - V_T - V_{DS}\right) \tag{6}$$

Dla maksymalnej wartości natężenia prądu dryfu wartość pochodnej w równaniu (6) musi być równa zero, co prowadzi do warunku na wartość potencjału źródło dren, przy której prąd dąży do nasycenia:

$$V_{DS} = V_{GS} - V_T \tag{7}$$

Wartość prądu nasycenia przedstawia zależność:

$$I_{DS} = \frac{W\mu C_i}{L}\left(V_{GS} - V_T\right)^2 \tag{8}$$

W zakresie prądu nasycenia, ruchliwość nośników ładunku może zostać określona z nachylenia wykresu zależności $I_{DS}^{1/2}(V_{GS})$ przy założeniu, że ruchliwość ma wartość stałą niezależną od potencjału bramki. Jeżeli warunek ten nie jest spełniony, ruchliwość w zakresie nasycania może być wyznaczona ze wzoru:

$$\mu_{sat}\left(V_{GS}\right) = \frac{\partial I_{DS,sat}}{\partial V_{GS}} \cdot \frac{L}{WC_i} \cdot \frac{1}{\left(V_{GS} - V_T\right)} \tag{9}$$

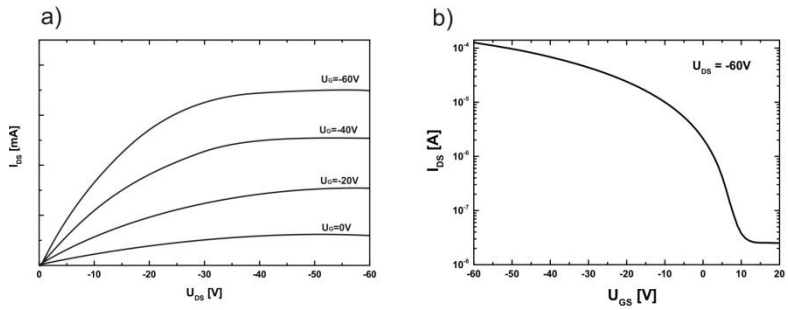
W liniowym zakresie pracy tranzystora $\left(V_{DS} << V_{GS}\right)$ wzór (8) może zostać uproszczony do postaci:

$$I_{DS} = \frac{W\mu C_i}{L} V_{DS}\left(V_{GS} - V_T\right) \tag{10}$$

Z powyższego wzoru wynika, że prąd dryfu nośników ładunku płynący pomiędzy źródłem i drenem jest wprost proporcjonalny do potencjału bramki, a ruchliwość nośników ładunku w tym przypadku może zostać wyznaczona z nachylenia prostej $I_{DS}(V_{GS})$:

$$\mu_{lin} = \frac{\partial I_{DS}}{\partial V_{GS}} \cdot \frac{L}{WC_i V_{DS}} \tag{11}$$

Typowe charakterystyki prądowo-napięciowe organicznego tranzystora polowego zostały przedstawione na rys. 6.



Rysunek 6. Charakterystyki prądowo-napięciowe OFET z warstwą aktywną TIPS pentacenu: a) wyjściowa ($I_{DS}(U_{DS})$) oraz b) charakterystyka przejściowa ($I_{DS}(U_{GS})$) [opracowanie własne [25]]

Odpowiednio rys. 6a przedstawia tzw. charakterystykę wyjściową określającą zależność prądu źródło-dren (I_{SD}) w funkcji napięcia dren-źródło (U_{DS}), przy ustalonym napięciu bramka-źródło (U_{GS}). Na podstawie charakterystyki wyjściowej można wyróżnić liniowy zakres pracy tranzystora, który odpowiada niskim wartościom napięć źródło-dren, oraz nasycenia, który odpowiada wysokim wartościom napięć U_{DS}. Rys. 6b przedstawia tzw. charakterystykę przejściową określającą zależność natężenia prądu I_{SD} w funkcji napięcia przyłożonego do bramki U_{GS}, przy stałym napięciu U_{DS}. W linowym zakresie pracy tranzystora, który odpowiada napięciom U_{DS} o wartościach $U_{DS}<<U_{GS}$, na podstawie wykresu, korzystając ze wzoru (11) można wyznaczyć ruchliwość nośników ładunku. Należy zaznaczyć, że większość półprzewodników organicznych wykazuje zmieniającą się wraz z napięciem bramki ruchliwość nośników ładunku, stąd zależność prądu źródło – dren (I_{DS}) od napięcia bramki (U_{GS}) może odbiegać od liniowej. W zakresie nasycania, aproksymacja linią prostą wykresu wyrażającego zależność $I_{DS}^{1/2}(U_{GS})$, pozwala wyznaczyć na podstawie wzoru (8) ruchliwość nośników ładunku oraz napięcie progowe (U_T).

Napięcie progowe jest jednym z najważniejszych parametrów charakteryzujących pracę organicznego tranzystora polowego. Przyczyną istnienia dla danego urządzenia napięcia progowego może być kilka efektów takich jak obecność pułapek nośników ładunku w warstwie półprzewodnika organicznego, stanów powierzchniowych na złączu półprzewodnik/izolator, a także ruchomych ładunków w warstwie izolatora itp., a na jego wartość ma wpływ rodzaj zastosowanego półprzewodnika organicznego, jako warstwy aktywnej oraz dielektryka izolującego elektrodę bramkującą [26]. Wartość napięcia progowego może zostać obniżona poprzez zastosowanie struktury bramka/dielektryk/półprzewodnik o większej pojemności elektrycznej, wykorzystując w tym celu dielektryk o dużej wartości stałej dielektrycznej bądź niewielkiej grubości, co w efekcie prowadzi do większej koncentracji nośników ładunku w kanale

przewodzącym urządzenia przy niższej wartości napięcia przykładanego do bramki. Częstym efektem obserwowanym dla organicznych tranzystorów polowych jest wzrost wartości napięcia progowego wraz z czasem pracy urządzenia [27]. Prowadzi to do powstania histerezy w przejściowych charakterystykach prądowo-napięciowych tranzystora, która przejawia się w wyższych wartościach natężeń prądów I_{DS} rejestrowanych przy rosnącym napięciu bramki od natężeń prądów I_{DS} mierzonych dla tych samych napięć U_{GS}, ale w odwrotnym porządku tzn. przy malejącym napięciu bramki. Ten efekt braku stabilności operacyjnej tranzystora jest istotny w zastosowaniach aplikacyjnych organicznych tranzystorów polowych, a także może zostać wykorzystany do zapisu informacji [28].

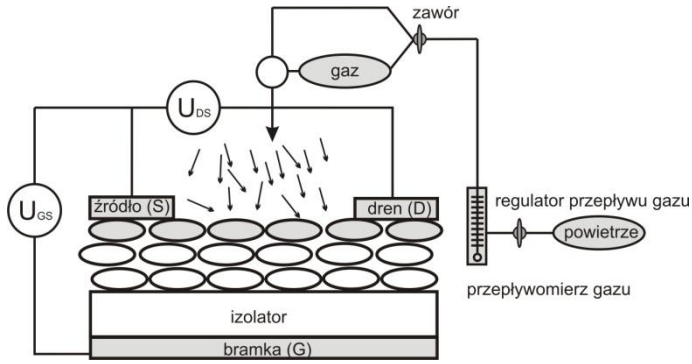
Kolejnym parametrem charakteryzującym OFET jest stosunek wartości natężenia prądu włączenia tranzystora (I_{ON}) do wartości natężenia prądu wyłączenia tranzystora (I_{OFF}), określana potocznie, jako ON/OFF. W przypadku, gdy organiczny tranzystor polowy ma spełniać rolę przełącznika, wartość tego parametru powinna być możliwie jak największa. Na wartość natężenia prądu włączenia, zaniedbując opór kontaktowy elektrod źródło i dren, decydujący wpływ ma ruchliwość nośników ładunku oraz pojemność zastosowanej w tranzystorze struktury bramka/dielektryk/półprzewodnik. Natomiast wartość natężenia prądu wyłączenia ogranicza prąd upływu bramki.

3. Wykorzystanie organicznych tranzystorów polowych do detekcji

Podstawową informacją, którą uzyskujemy wykorzystując tranzystor polowy jest jego charakterystyka prądowo – napięciowa, na której przebieg możemy wpłynąć zmieniając wartość napięcia progowego oraz ruchliwości nośników ładunku. Oddziaływanie badanej substancji (analitu) z materiałem tworzącym warstwę aktywną tranzystora może spowodować zarówno zmianę wartości ruchliwości nośników ładunku jak i napięcia progowego, a zmiany tych parametrów można skorelować z koncentracją oznaczanego analitu. Ruchliwość nośników ładunku w materiałach organicznych zależy od rodzaju materiału, jego struktury oraz obecności pułapek [29]. Podobne parametry mają wpływ na wartość napięcia progowego [30]. Aby badana substancja mogła swobodnie penetrować półprzewodnik organiczny, a w szczególności kanał przewodzący tranzystora, urządzenie musi posiadać odpowiednią konstrukcję. Z przedstawionych na rys. 2. podstawowych konfiguracji budowy tranzystora, konstrukcje przydatne do wykorzystania tranzystora jako czujnika to (a) i (b), w których zapewniony jest bezpośredni dostęp analitu do warstwy półprzewodnika.

3.1. OFET jako czujniki gazów

Układ do badania wpływu gazu na przewodnictwo warstwy aktywnej organicznego tranzystora polowego został przedstawiony na rys. 7 [5].



Rysunek 7. Schemat układu do badania wpływu gazu na parametry pracy OFET [opracowanie własne na podstawie [5]].

Odpowiedź czujnika OFET silnie zależy od własności substancji, która podlega oznaczeniu oraz zastosowanej w urządzeniu warstwy aktywnej. Przykład stanowić mogą badania zmiany parametrów pracy detektorów OFET z warstwą ftalocyjaniny miedzi (CuPc) oraz perfluorowanej ftalocyjaniny miedzi (F_{16}CuPc) na skutek oddziaływania z metylofosfonianem dimetylu (DMMP) [31], lotnym związkiem organicznym, który często używany jest, jako związek imitujący bojowy środek trujący sarin. W urządzeniach z warstwą aktywną CuPc, która stanowi półprzewodnik typu p, obserwowany jest spadek wartości prądu I_{DS} pod wpływem DMMP. Natomiast w OFET z warstwą aktywną F_{16}CuPc, stanowiącą półprzewodnik typu n, następuje wzrost prądu I_{DS} pod wpływem DMMP. Rezultaty badań pokazują, że cząsteczki gazu DMMP stanowią odpowiednio akceptory dziur i donory elektronów w OFET z warstwą aktywną typu p i n, zmniejszając w pierwszym przypadku przewodnictwo, a w drugim je zwiększając, co przekłada się bezpośrednio na wartość rejestrowanego prądu I_{DS} w tranzystorach. Słuszność tego mechanizmu detekcji potwierdzają wyniki badań wpływu DMMP na przewodnictwo detektora OFET z mieszaniną oligomerów fenylotiofenu (6PTTP6) funkcjonalizowanych grupą hydroksylową (OH), jako warstwą aktywną. Prąd nasycenia w tym tranzystorze zdecydowanie maleje po wprowadzeniu gazu do otoczenia półprzewodnika, który należy zaznaczyć jest półprzewodnikiem typu p [32].

Własności materiałów organicznych, a w szczególności obecność słabo związanych elektronów typu π, sprawia, że detektory OFET są czułe na związki utleniające, czyli przyjmujące elektrony. Oddziaływanie warstwy aktywnej OFET z substancją o własnościach utleniających może w konsekwencji powodować wzrost koncentracji dziur i przewodnictwa oraz zmianę wartości napięcia progowego. Dobrą detekcję nadtlenków wodoru (H_2O_2) i di-tert-butylowego (DTBP), związków o silnych własnościach utleniających, wykazują detektory OFET z warstwą aktywną typu p zbudowaną z kompleksu ftalocyjaniny z metalem [33].

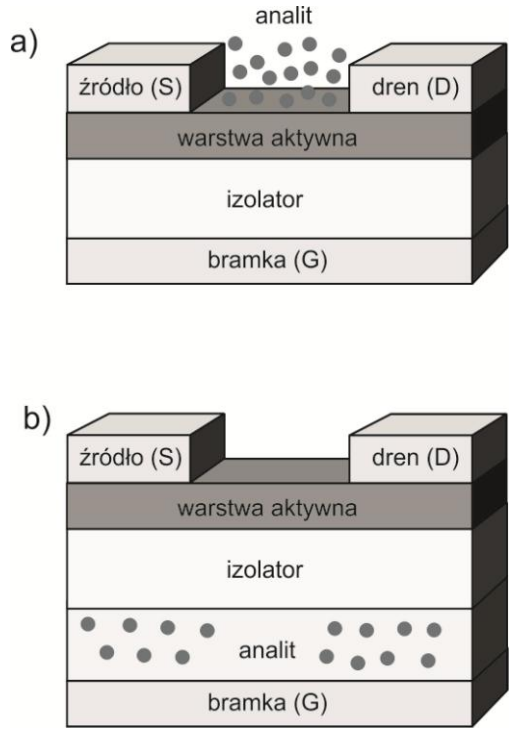
Zmianę prądu drenu w czujniku OFET może powodować zmiana struktury warstwy aktywnej tranzystora pod wpływem oddziaływania z gazem. Przykład stanowią badania detekcji przepływu gazu za pomocą czujnika OFET z warstwą polimeru poly(3-heksylotiofenu) (P3HT) pokrytą cząsteczkami jednego z najbardziej znanych związków ciekłokrystalicznych 5CB (pentylo-cyjanobifenylu) [34] o silnym momencie

dipolowym (około 4D). Pod wpływem padającego na urządzenie strumienia cząsteczek azotu następuje zmiana orientacji cząsteczek 5CB, które ustawiają się ujemnym biegunem dipola w kierunku warstwy P3HT, co w efekcie powoduje zwiększenie koncentracji dziur w kanale przewodzącym i tym samym wzrost wartości prądu drenu. Podsumowując mechanizm zmiany prądu w tranzystorze może być różny i powodowany przez szereg procesów, do których możemy zaliczyć min. ubytek bądź generację nowych pułapek nośników ładunku w warstwie aktywnej tranzystora, zmianę jej struktury pod wpływem reakcji chemicznych wywołanych oddziaływaniem z gazem czy też wprowadzeniem jonów lub domieszek. Reakcja chemiczna, jaka zachodzi pomiędzy gazem i materiałem stanowiącym warstwę aktywną tranzystora może być odwracalna lub nieodwracalna, w konsekwencji, czego sensor może być używany do detekcji wielokrotnie lub jednorazowo. Ponadto OFET może być czuły wybiórczo na jeden gaz lub na wiele różnych gazów jednocześnie. Generalnie do detekcji gazów przez tranzystory organiczne można wykorzystać zmianę prądu dryfu, ruchliwości nośników ładunku oraz wielkość napięcia progowego.

3.2. OFET jako czujniki chemiczne i biologiczne

Organiczne tranzystory polowe mogą być również wykorzystywane do detekcji substancji chemicznych i biologicznych w ciekłym stanie skupienia. W tym celu najczęściej stosowane są dwie konstrukcje detektorów. Pierwsza to czujniki OFET, w których analit pozostaje w bezpośrednim kontakcie z aktywną warstwą tranzystora zbudowaną z półprzewodnika organicznego. Natomiast druga to detektory OFET, w których działaniu badanego analitu poddawana jest elektroda bramkująca [4]. Na rys. 8 przedstawiono schematycznie budowę tych dwóch urządzeń.

Budowę detektora, w którym w bezpośrednim kontakcie z uwodnionym analitem pozostaje aktywna warstwa półprzewodnika organicznego przedstawiono na rys. 8a. Czujnik o takiej konstrukcji został zastosowany do detekcji substancji oznaczanej w środowisku wodnym przez Someya i współpracowników [35]. Aby uniknąć rozwarstwiania warstwy półprzewodnika organicznego od podłoża pod wpływem oddziaływania z ciekłym analitem zastosowano warstwę adhezyjną, natomiast elektrody źródło i dren zabezpieczono odpowiednią warstwą hydrofobową. Testom poddano tranzystory z warstwą aktywną zbudowaną z różnych półprzewodników organicznych, które poddawano działaniu rozpuszczonych w wodzie związków hydroksylowych i karboksylowych takich jak min. kwas mlekowy, glukoza, czy kwas pirogronowy. Badania wykazały, że pod wpływem analitu urządzenia wykazują spadek natężenia prądu źródło-dren płynącego w kanale przewodzącym oraz ruchliwości nośników ładunku. Efekt obserwowano zarówno w przypadku oddziaływania ze statycznym analitem jak i tym pozostającym w ruchu. W celu detekcji ruchomego analitu, warstwa aktywna urządzenia została pokryta systemem kanalików mikrofluidalnych, do wnętrza, których wstrzykiwana była badana substancja rozpuszczona w wodzie.



Rysunek 8. Schematy detektorów OFET w których: a) analit pozostaje w bezpośrednim kontakcie z aktywną warstwą tranzystora; b) działaniu badanego analitu poddawana jest elektroda bramkująca [opracowanie własne na podstawie [4]]

Na rys. 8b. przedstawiono schemat urządzenia wykrywającego obecność jonów w roztworze, które wykorzystuje w swojej budowie OFET. Jakkolwiek w odróżnieniu od struktury warstwowej organicznego tranzystora polowego typu TC BG przedstawionej na rys. 2a., elektroda bramkująca nie pozostaje w bezpośrednim kontakcie z warstwą izolującą, ale jest od niej oddzielona roztworem badanego analitu. Tego typu urządzenie jest nazywane w literaturze jonoselektywnym organicznym tranzystorem polowym (*Ion Sensitive Organic Field Effect Transistor*, ISOFET), a jego budowa i zasada działania jest zbliżona do klasycznego jonoselektywnego tranzystora polowego zbudowanego z materiałów nieorganicznych (*Ion Sensitive Field Effect Transistor*). Detektor ISFET został opracowany i opisany po raz pierwszy w 1970 roku przez Bergveld'a [36], a dalszy rozwój technologiczny pozwolił na wykorzystanie tego typu urządzeń do detekcji cząsteczek różnych substancji takich jak nieorganiczne sole [37], enzymy [38] czy glukoza [39]. W detektorze ISOFET oddziaływanie roztworu badanego analitu z warstwą izolującą powoduje pojawienie się potencjału na granicy faz: badany analit/warstwa dielektryka, którego wartość opisana równaniem Nernst'a zależy od koncentracji jonów w roztworze. Wzrost koncentracji jonów zmienia rozkład ładunku na złączu dielektryk/półprzewodnik organiczny, co w efekcie powoduje zmianę rozkładu potencjału elektrycznego w urządzeniu prowadząc do zmiany prądu płynącego pomiędzy źródłem i drenem. Wartość powstającego na granicy faz potencjału wynika z chemicznych właściwości warstwy izolacyjnej tranzystora. Rodzaj

dielektryka zastosowanego w urządzeniu determinuje również reakcje zachodzące w warunkach kontaktu z badanym analitem. W przypadku dielektryków nieorganicznych wytwarzanych z tlenków metali, ze względu na obecność na ich powierzchni grup wodorotlenowych, w kontakcie z roztworem badanego analitu mogą mieć miejsce reakcje protonowania lub deprotonowania, w zależności od jego kwasowości, wykazując w ten sposób czułość na pH [40]. W konstrukcji ISOFET wykorzystuje się również dielektryki organiczne, które wykazują brak bądź słabą czułość na obecność jonów w badanym roztworze, stąd konieczne jest w tym przypadku zastosowanie jonoselektywnych membran, które wzmacniają detekcję [41]. Jonoselektywne membrany nanoszone są na warstwę izolującą i pozostają w bezpośrednim kontakcie z roztworem badanego analitu. Stosowane w ISOFET membrany jonoselektywne można podzielić na dwie kategorie. Pierwsza to membrany z chemicznie wbudowanymi grupami funkcyjnymi takimi jak np. $–SO^-_3$, $–COO^-$ lub $–NR^+_3$ zdolnymi do oddziaływania z wybranym składnikiem roztworu, który zostaje związany wewnątrz membrany [42]. Natomiast druga kategoria to membrany zawierające jonofory, organiczne związki chemiczne zdolne do transferu jonów do wnętrza membrany [43]. W odróżnieniu od membran należących do pierwszej kategorii, których nanoszenie na powierzchnię warstwy izolującej tranzystora wymaga procesów wysoko temperaturowych takich jak naparowywanie próżniowe, związki należące do drugiej kategorii mogą być wytwarzane za pomocą procesów nisko temperaturowych takich jak np. technika LB. Technika Langmuira-Blodgett, w skrócie LB, jest jednym ze sposobów tworzenia uporządkowanych struktur do zastosowań w elektronice molekularnej, a polega na wynoszeniu warstw monomolekularnych na powierzchnię ciała stałego, co w efekcie prowadzi do powstania filmów wielowarstwowych [41]. Czujnik pH zbudowany na bazie jonoselektywnego organicznego tranzystora polowego z użyciem P3HT, jako warstwy aktywnej, azotku krzemu, jako warstwy izolującej oraz Ag/AgCl, jako elektrody referencyjnej został opracowany przez Bartic'a i współpracowników [44]. Zmiana wartości pH roztworu powodowała generację potencjału na złączu roztwór/dielektryk, w wyniku zachodzących na powierzchni dielektryka reakcji wiązania jonów wodorowych H^+ z roztworu z grupą silanolową SiOH oraz aminową $SiNH_2$ należących do azotku krzemu. W efekcie wartość mierzonego prądu była różna dla różnych wartości pH roztworu. Ponadto zastąpienie w detektorze azotku krzemu tlenkiem tantalu, pozwoliło na uzyskanie czujnika pH o większej czułości i niższych wartościach operacyjnych napięć [45]. W pełni organiczny ISOFET został zaproponowany przez Gao i współpracowników [1]. W odróżnieniu od opisanego powyżej czujnika pH, urządzenie wykorzystuje elastyczne polimerowe podłoże i posiada zintegrowaną z warstwą izolującą elektrodę bramkującą. Oprócz zastosowania OFET w detekcji pH roztworu, organiczne tranzystory zbudowane z warstwy aktywnej P3HT zostały wykorzystane do detekcji jonów K^+, H^+, Na^+ czy też Ca^{2+} [46].

Zastosowanie w konstrukcji OFET warstwy biologicznie czynnej pozwala na skonstruowanie bioczujnika. Podstawową konstrukcję BioOFET można uzyskać wykorzystując schemat budowy ISOFET z dodatkową warstwą substancji biologicznie czynnej na powierzchni warstwy dielektryka. W przypadku czujnika glukozy rolę substancji czynnej może pełnić tlenek glukozy (GOx). W detektorze glukozy

zaproponowanym przez Bartic`a wykorzystano konstrukcję ISOFET z warstwą aktywną P3HT oraz warstwą izolującą Ta_2O_5 pokrytą GOx [45]. Tlenek glukozy reaguje z glukozą, co w efekcie prowadzi do powstania w roztworze kwasu glukozowego, który zmienia pH, na który to czynnik jest czuła warstwa izolująca bramkę. W detektorze glukozy zaproponowanym przez Liu zastosowano konstrukcję OFET używając jako warstwy aktywnej półprzewodnika organicznego PEDOT:PSS, który pokryto warstwą GOx [47]. Działanie czujnika opiera się na zachodzącej reakcji redoks pomiędzy PEDOT:PSS i H_2O_2, który jest końcowym produktem reakcji glukozy z tlenkiem glukozy.

Ważnym zastosowaniem OFET na polu biosensorów są detektory cząsteczki kwasu deoksyrybonukleinowego (DNA). W tym celu stosuje się tranzystory polowe z warstwą aktywną zbudowaną z min. pentacenu [48], P3HT [49] czy DDFTTF [50]. W czujniku OFET z warstwą aktywną pentacenu [48], po umieszczeniu na jej powierzchni (w wyniku adsorpcji fizycznej) cząsteczki DNA następuje zmiana napięcia progowego, którego wartość ponadto jest różna dla formy jednoniciowej i dwuniciowej. Prace nad optymalizacją urządzenia, doprowadziły do wniosku, że czułość urządzenia jest większa w przypadku zastosowania w detektorze cienkich warstw pentacenu i obecności większych krystalitów na jej powierzchni. Mechanizm detekcji jest spowodowany wychwytem elektronów z warstwy pentacenu, a następnie ich pułapkowaniem w cząsteczce DNA.

Oprócz opisanych powyżej detektorów chemicznych i biologicznych prowadzone są badania nad wykorzystaniem OFET z warstwą aktywną pentacenu, α6T, DHα6T czy CuPc do detekcji takich substancji jak kwas mlekowy, kwas pirogronowy czy cysteina [3]. Mechanizm detekcji opiera się na zmianie transportu nośników ładunku w kanale przewodzącym urządzenia na skutek procesów pułapkowania i domieszkowania, które zachodzą w przypadku dyfuzji analitu na złączu półprzewodnik organiczny/warstwa izolująca bramkę. Na czułość detektorów ma wpływ struktura cząsteczkowa, morfologia i grubość warstwy aktywnej. Detektory można wytwarzać na giętkich podłożach, czego przykładem może być OFET wykonany na pod łożu PET do detekcji cysteiny [51]. Ponadto wykazano możliwość zintegrowania czujników OFET z żywymi komórkami w celach detekcyjnych. Scarpa zaprezentował OFET z warstwą aktywną P3HT, na powierzchni, której, po odpowiednim przygotowaniu zostały wyhodowane komórki fibroblastów [52]. Tak przygotowane urządzenie zostało poddane działaniu różnych elektrolitów i złożonych podłoży biologicznych. Detektor wykazywał czułość na obecność jonów i był stabilny, co pozwala stwierdzić możliwość wykorzystania tego bioczujnika w badaniach in vitro.

4. Wybrane zastosowania czujników OFET

Czujniki zbudowane na bazie organicznych tranzystorów polowych wykazują wiele zalet. Przede wszystkim są to urządzenia będące kombinacją dwóch elementów elektronicznych: sensora i wzmacniacza, stąd mała zmiana napięcia elektrody bramkującej, wywołana oddziaływaniem czujnika OFET z badanym analitem powoduje wyraźne zmiany prądu płynącego w kanale przewodzącym tranzystora, dając w efekcie detektor o dużej czułości. Inne ważne zalety czujników zbudowanych na bazie OFET to min. dobra selektywność, biokompatybilność, niewielkie rozmiary, potencjalnie niskie koszty produkcji oraz możliwość nanoszenia na elastyczne podłoża.

Dziedziną, w której potencjalnie mogą zostać zastosowane czujniki OFET to biomedycyna. Na szczególną uwagę zasługują tutaj takie urządzenia jak elektroniczna skóra (*electronic skin*, e-skin) czy też elektroniczny nos (*electronic nose*, e-nose). Idea stworzenia elektronicznej skóry nie jest nowa, ale przełomem w tej dziedzinie wydaje się zastosowanie elastycznych materiałów organicznych i elementów elektronicznych takich jak OFET [53]. Dotychczas, bowiem stosując technologie oparte na klasycznych półprzewodnikach nieorganicznych, przy konstruowaniu odpowiednika skóry naturalnej, napotykano na wiele przeszkód takich jak np. brak elastyczności, niewystarczająca czułość sensorów, czy też duże zużycie energii elektrycznej. Należy pamiętać, że skóra naturalna jest jednym z najbardziej skomplikowanych tworów ludzkiego organizmu, który zajmuje stosunkowo dużą powierzchnię, spełnia wiele czynności ochronnych, ale przede wszystkim odpowiada za percepcję dotyku, ciepła i bólu ze środowiska zewnętrznego. Upraszczając można napisać, że skóra naturalna działa jak siatka sensorów, odbierając i przekazując bodźce zewnętrzne do mózgu. Najbardziej obiecujące badania w tym zakresie są prowadzone przez zespół Takao Someya na Uniwersytecie Tokijskim. Naukowcom udało się tam stworzyć w pełni elastyczną wielkopowierzchniową strukturę sieci czujników temperatury i ciśnienia, które mogą działać równocześnie [54]. Czułość urządzeń nie zmienia się nawet wtedy, gdy podłoże, na którym się znajdują doznaje działania odkształceń mechanicznych. Ostatnie publikacje grupy donoszą o stworzeniu elastycznej skóry o grubości zaledwie 300 nm (co stanowi najmniejszą spośród uzyskanych do tej pory grubości) pokrytej pikselami, z których każdy składa się z tranzystora OFET i rezystancyjnego czujnika dotyku [55]. Urządzenie jest biokompatybilne i może być umieszczane na skórze ludzkiej w zastosowaniach diagnostyki medycznej. Zbliżonym do elektronicznej skóry, zastosowaniem czujników OFET są inteligentne materiały (*electronic textiles*, e-textiles) [56, 57], gdzie stosowane są głównie czujniki ciśnienia i bioczujniki. Zakłada się, że w przyszłości E-materiały znajdą zastosowanie w monitorowaniu funkcji ludzkiego organizmu takich jak min. prawidłowa postawa, czynności oddechowe, czy też skład płynów ustrojowych [58]. Podstawowe wymagania, jakie muszą spełniać czujniki, aby mogły zostać zastosowane w tym urządzeniu są zbliżone do tych dla elektronicznej skóry, ale przede wszystkim ważna jest ich biokompatybilność i elastyczność.

Ze względu na potrzeby współczesnej medycyny, która szuka tanich, ogólno dostępnych i bezinwazyjnych metod diagnostyki, ważnym potencjalnym wykorzystaniem czujników gazu jest ich zastosowanie w urządzeniu nazywanym elektronicznym nosem [59]. Badania naukowe wskazują jak istotne, w procesie diagnostyki i leczenia wielu chorób, jest wykrycie zmiany w żywych komórkach na poziomie molekularnym, które nie tylko umożliwia określenie procesu patofizjologicznego, jaki ma miejsce, ale pozwala określić sposób leczenia. Prace prowadzone są w kierunku stworzenia elektronicznego nosa, urządzenia, które będzie zdolne wykryć odpowiednie biomarkery chorobowe w wydzielanych przez organizm ludzki oparach lub płynach ustrojowych. Wiadomo, że wiele chorób można rozpoznać po unikalnej kombinacji produktów metabolizmu. Niektóre z nich to lotne związki organiczne, które mogą sygnalizować chorobę nowotworową, ale również wskazywać na zaburzenia psychiczne lub neurologiczne.

5. Podsumowanie

W rozdziale przedstawiono przegląd doniesień literaturowych dotyczących zastosowania organicznych tranzystorów polowych w detekcji różnych substancji ze szczególnym uwzględnieniem gazów i cieczy. Ogromny postęp w detekcji substancji za pomocą OFET pozwala oczekiwać nowych urządzeń, w których znajdą one zastosowanie już w najbliższym czasie. Niektóre z nich, takie jak elektroniczna skóra, inteligentne tekstylia czy też elektroniczny nos zostały po krótce opisane. Szczególnie biomedycyna i robotyka to dziedziny, które mogą niezmiernie zyskać na dalszym rozwoju detektorów OFET. Należy jednak podkreślić, że żadne z opisanych rozwiązań technologicznych nie znalazło do tej pory jeszcze praktycznego zastosowania, a wszystkie są w fazie testów. Ponadto ciągle jest wiele problemów, które wymagają znalezienia rozwiązań min. wysokie napięcie pracy, którego wymagają opisane w rozdziale niektóre z detektorów. Jakkolwiek takie cechy materiałów organicznych jak możliwość ich chemicznej modyfikacji i dostosowywania ich struktury do odpowiednich zadań oraz wysoka czułość na czynniki zewnętrzne pozwalają oczekiwać, że technologia detekcji z ich wykorzystaniem ma duży potencjał, który zostanie w przyszłości w pełni wykorzystany w takich zastosowaniach jak robotyka, biomedycyna czy biotechnologia.

Literatura

1. Bernards D.A., Owens R.M., Malliaras G.G. *Organic Semiconductors in Sensor Applications*, Springer Series in Material Science, (2007).
2. Torsi L., Magliulo M., Manoli K., Palazzo G. *Organic field-effect transistors: a tutorial review*, Chemical Society Reviews, 42 (2013), s. 8612-8628.
3. Lin P., Yan F. *Organic thin-film transistors for chemical and biological sensing*, Advanced Materials, 24 (2012), s. 34-51.
4. Mabeck J.T., Malliars G.G. *Chemical and biological sensors based on organic thin-film transistors*, Analytical and Bioanalytical Chemistry, 384 (2006), s. 343-353.
5. Someya T., Dodabalapur A., Huang J., See K.C., Katz H.E. *Chemical and physical sensing by organic field-effect transistors and related devices*, Advanced Materials, 22 (2010), s. 3799-3811.
6. Kergoat L., Piro B., Berggen M., Horowitz G., Pham M-Ch. *Advances in organic transistor-based biosensors: from organic electrochemical transistors to electrolyte-gated organic field-effect transistors*, Analytical and Bioanalytical Chemistry, 402 (2012), s. 1813-1826.
7. Shen H., Di Ch-A., Zhu D. *Organic transistor for bioelectronics applications*, Science China, 60 (2017), s. 437-449.
8. Zang Y., Huang D., Di Ch-A., Zhu D. *Device engineered organic transistors for flexible sensing applications*, Advanced Materials, 28 (2016), s. 4549-4555.
9. Lv A., Pan Y., Chi L. *Gas sensors based on polimer field-effect transistors*, Sensors, 17 (2017), s. 1-16.
10. Zhang C., Chen P., Hu W. *Organic field-effect transistor-based gas sensors*, Chemical Society Reviews, 44 (2015), s. 2087-2107.
11. Lilienfeld J.E. *Method and apparatus for controlling electric currents*, US Patent 1745175, (1930).
12. Kahng D., Atalla M.M. *Silicon – silicon dioxide field induced surface devices*, IRE-AIEE Solid-State Device Res. Conf., Carnegie Inst. of Technology, Pittsburgh, PA, 27th-29th October 1960.

13. Horowitz G., Fichou D., Peng X., Xu Z., Garnier F. *A field-effect transistor based on conjugated alpha-sexithienyl*, Solid State Communications, Vol. 72, 4 (1989), s. 381-384.
14. Tsumura A., Koezuka H., Ando T. *Macromolecular electronic device: field-effect transistor with a polythiophene thin film*, Applied Physics Letters, 49 (1986), s. 1210-1212.
15. Sirringhaus H. 5th Anniversary Article: *Organic Field-Effect Transistors: The Path Beyond Amorphous Silicon*, Advanced Materials, 26 (2014), s. 1319-1335.
16. Sirringhaus H. *Reliability of organic field-effect transistors*, Advanced Materials, 21 (2009), s. 3859-3873.
17. Subbarao N.V.V., Gedda M., Iyer P.K., Goswami D.K. *Enhanced environmental stability induced by effective polarization of a polar dielectric layer in a trilayer dielectric system of organic field-effect transistors: a quantitative study*, ASC Applied Materials&Interfaces, 7 (2015), s. 1915-1924.
18. Horowitz G. *Organic field-effect transistor*, Advanced Materials, 10 (1998), s. 365-377.
19. Zaumseil J., Sirringhaus H. *Electron and ambipolar transport in organic field-effect transistors*, Chemical Reviews, 107 (2007), s. 1296−1323.
20. Köhler A., Bässler H. *Electronic Processes in Organic Semiconductors*, Wiley-VCH Verlag GmbH & Co. KGaA, (2015).
21. Pope M., Swenberg Ch.E. *Electronic Processes in Organic Crystals*, Clarendon Press, New York (1999).
22. Karl N. *Charge-Carrier Mobility in Organic Crystals*, Organic Electronic Materials, Springer, (2001), s. 283-326.
23. Brütting W. Organic Semiconductors, Encyclopedia of Physics, Wiley-VCH Verlag GmbH & Co. KGaA, (2005), s. 1866-1876.
24. Minemawari H., Yamada T., Matsui H., Tsutsumi J., Haas S., Chiba R., Kumai R., Hasegawa T. *Inkjet printing of single-crystal films*, Nature, 475 (2011), s. 366-367.
25. Franz M. *Raport wewnętrzny projektu pt. Organic Stability Project*, KTP Program Merck Chemicals Ltd oraz University of Southampton (2010).
26. Veres J., Ogier S., Lloyd G. *Gate insulators in organic field-effect transistors*, Chemistry of Materials, 16 (2004), s. 4543-4555.
27. Hardigree J.F.M., Katz H.E. *Through thick and thin: tuning the threshold voltage in organic field-effect transistors*, Accounts of Chemical Research, 47 (2014), s. 1369-1377.
28. Chou Y-H., Chang H-Ch.,Liu Ch-L., Chen W-Ch. *Polymeric charge storage electrets for non-volatile organic field effect transistor memory devices*, Polymer Chemistry, 6 (2015), s. 341-352.
29. Coropceanu V., Cornil J.,Filho D.A.S., Olivier Y., Silbey R., Brédas J.L. *Charge transport in organic semiconductors*, Chemical Reviews, 107 (2007), s. 926-952.
30. Schroeder R., Majewski L. A., Grell M. *A study of the threshold voltage in pentacene organic field-effect transistors*, Applied Physics Letters, 83 (2003), s. 3201-3203.
31. Yang R.D., Park J., Colesniuc C.N., Schuller I.K., Royer J.E., Trogler W.C., Kummel A.C. *Analyte chemisorption and sensing on n- and-p-channel copper phthalocyanine thin-film transistors*, The Journal of Chemical Physics, 130 (2009), s. 164703-8.
32. Huang J., Miragliotta J., Becknell A., Katz H.E. *Hydroxy-terminated organic semiconductor-based field-effect transistors for phosphonate vapor detection*, Journal of the American Chemical Society, 129 (2007), s. 9366-9376.
33. Royer J.E., Kappe E.D., Zhang Ch., Martin D.T., Trogler W.C., Kummel A.C. *Organic thin-film transistors for selective hydrogen peroxide and organic peroxide vapor detection*, The Journal of Physical Chemistry, 116 (2012), s. 24566-24572.
34. Seo J., Park S., Nam S., Kim H., Kim Y. *Liquid crystal-on-organic field-effect transistor sensory devices for perceptive sensing of ultralow intensity gas flow touch*, Scientific Reports, 3 (2013), s. 1-6.

35. Someya T., Dodabalapur A., Gelperin A., Katz H.E., Bao Z. *Integration and response of organic electronics with aqueous microfluidics*, Langmuir, 18 (2002), s. 5299-5302.
36. Bergveld P. *Thirty years of ISFETOLOGY: what happened in the past 30 years and what may happen in the next 30 years*, Sensors and Actuators B: Chemicals, 88 (2003), s. 1-20.
37. Moss S.D., Johnson C.C., Janata J. *Hydrogen, calcium, and potassium ion-sensitive FET transducers: a preliminary report*, IEEE Transactions on Biomedical Engineering, 25 (1978), s. 49-54.
38. van der Schoot B.H., Bergveld P. *ISFET based enzyme sensors*, Biosensors, 3 (1987), s. 161-186.
39. Volotovskya V., Soldatkinb A.P., Shul'gac A.A., Rossokhatya V.K., Strikhaa V.I., El'skayab A.V. *Glucose-sensitive ion-sensitive field-effect transistorbased biosensor with additional positively charged membrane: dynamic range extension and reduction of buffer concentration influence on the sensor response*, Analytica Chimica Acta, 322 (1996), s. 77-81.
40. Yates D.E., Levine S., Healy T.W. *Site-binding model of the electrical double layer at the oxide/water interface*, Journal of the Chemical Society, Faraday Transactions 1: Physical Chemistry in Condensed Phases, 70 (1974), s. 1807-1818.
41. Osa T. *Potentiometric response of lipid modified ISFET*, Applied Biochemistry and Biotechnology, 41 (1993), s. 35-40.
42. Madou M., Morrison S. *Chemical sensing with solid state devices*, Academic Press, 1989.
43. Morf W.E. *The principles of ion-selective electrodes and of membrane transport*, Volume 2, Elsevier Science, 1981.
44. Bartic C., Palan B., Campitelli A., Borghs G. *Monitoring pH with organic-based field-effect transistors*, Sensors and Actuators B: Chemical Special Issues 83 (2002), s. 115-122.
45. Bartic C., Campitelli A., Borghs S. *Field-effect detection of chemical species with hybrid organic/inorganic transistors*, Applied Physics Letters, 82 (2003), s. 475-477.
46. Scarpa G., Idzko A.L., Yadav A., Thalhammer S. *Organic ISFET Based on Poly (3-hexylthiophene)*, Sensors, 10 (2010), s. 2262-2273.
47. Liu J., Agarwal M., Varahramyan K. *Glucose sensor based on organic thin film transistor using glucose oxidase and conducting polymer*, Sensors and Actuators B: Chemical, 135 (2008), s. 195-199.
48. Zhang Q., Subramanian V. *DNA hybridization detection with organic thin film transistors: Toward fast and disposable DNA microarray chips*, Biosensors and Bioelectronics, 22 (2007), s. 3182-3187.
49. Yan F., Moka S.M., Yub J., Chana H.L.W., Yangb M. *Label-free DNA sensor based on organic thin film transistors*, Biosensors and Bioelectronics, 24 (2009), s. 1241-1245.
50. Khan H.U., Roberts M.E., Johnson O., Förch R., Knoll W., Bao Z. *In situ, label-free DNA detection using organic transistor sensors*, Advanced Materials, 22 (2010), s. 4452-4456.
51. Roberts M.E., Mannsfeld S.C.B., Stoltenberg R.M., Bao Z. *Flexible, plastic transistor-based chemical sensors, Organic Electronics*, 10 (2009), s. 377-383.
52. Scarpa G., Idzko A-L., Yadav A., Martin E., Thalhammer S. *Toward cheap disposable sensing devices for biological assays*, IEEE Transactions On Nanotechnology, 9 (2010), s. 527-532.
53. Sokolov A.N., Tee B.C-K., Bettinger C.J., Tok J.B.-H., Bao Z. *Chemical and engineering approaches to enable organic field-effect transistors for electronic skin applications*, Accounts Of Chemical Research, 45 (2012), s. 361-371.
54. Lee W., Kim D., Rivnay J., Matsuhisa N., Lonjaret T., Yokota T., Yawo H., Sekino M., Malliaras G.G., Someya T. *Integration of organic electrochemical and field-effect transistors for ultraflexible, high temporal resolution electrophysiology arrays*, Advanced Materials, 28 (2016), 9722-9728.

55. Nawrocki R.A., Matsuhisa N., Yokota T., Someya T. *300-nm imperceptible, ultrafl exible, and biocompatible e-skin fit with tactile sensors and organic transistors*, Advanced Materials, 2 (2016), s. 1500452.
56. Carpi F., De Rossi D. *Electroactive polymer-based devices for e-textiles in biomedicine*, IEEE Transactions On Information Technology In Biomedicine, 9 (2005), s. 295-318.
57. Bonfiglio A., De Rossi D., Kirstein T., Locher I.R., Mameli F., Paradiso R., Vozzi G. *Organic field effect transistors for textile applications*, IEEE Transactions On Information Technology In Biomedicine, 9 (2005), s. 319-324.
58. Stoppa M., Chiolerio A. *Wearable electronics and smart textiles: a critical review*, Sensors, 14 (2014), s. 11957-11992.
59. Fitzgerald J.E., Bui E.T.H., Simon N.M., Fenniri H. *Artificial nose technology: status and prospects in diagnostics*, Trends in Biotechnology, 35 (2017), s. 33-42.

Detekcja materiałów przy użyciu organicznych tranzystorów polowych

Abstrakt
W ciągu ostatnich dwudziestu lat nastąpił rozwój elektroniki organicznej, która stała się ważnym obszarem badań naukowych i technologicznych. Organiczne ogniwa fotowoltaiczne i organiczne diody elektroluminescencyjne zostały wykorzystywane w urządzeniach komercyjnych, a jednym z najbardziej obiecujących zastosowań dla organicznych tranzystorów polowych są czujniki chemiczne i biologiczne. Postęp w tej dziedzinie jest tematem tego opracowania. W rozdziale przedstawiono zasadę działania, podstawowe parametry oraz charakterystyki prądowo-napięciowe organicznych tranzystorów polowych (OFET) oraz możliwości wykorzystania tych urządzeń jako czujników substancji gazowych i ciekłych. Przedstawiono zasadę działania czujników wykorzystujących OFET, a także potencjalne możliwości aplikacyjne tych urządzeń.
Słowa kluczowe: organiczny tranzystor polowy, czujniki chemiczne, czujniki biologiczne, czujniki gazów, materiały organiczne

Materials detection with organic field effect transistors

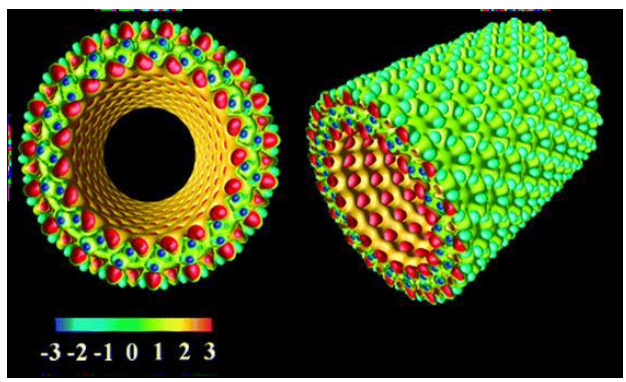
Abstract
Over the last two decades, organic electronics have developed into an important area of research and technology. Organic solar cells and organic light emitting diodes have been commercialized for a variety of applications and one of the most promising applications for organic field effect transistors seems to be their usage as chemical and biological sensors. The progress made in this direction is the topic of this review. The chapter presents the operating principle, the basic parameters and the current-voltage characteristics of organic field effect transistors (OFET) also the possibilities of using these devices as detectors of gas and liquid substances are described. The working principles of OFET-based sensors are discussed in detail, as well as the potential application of these devices.
Keywords: organic field-effect transistors, chemical sensors, biological sensors, gas sensors, organic materials

Małgorzata Skibińska[1]

Zastosowanie nanorurek haloizytowych

1. Wstęp

Haloizyt jest naturalnie występującym dwuwarstwowym glinokrzemianem, należącym do grupy kaolinitów o wzorze $Al_2Si_2O_5(OH)4*nH_2O$, gdzie wartości n= 0 przyjmuje dla formy bezwodnej (7Å) i n=2 dla formy uwodnionej (10Å) [1]. Może występować w formie płytek, sferycznej lub rurek (rysunek 1). Dzięki strukturze rurek znajduje szerokie zastosowanie w dynamicznie rozwijającej się nanotechnologii. Nanorurki te stanowią alternatywę dla drogich nanorurek węglowych, nie ustępują im jeśli chodzi o swe właściwości a przy tym są ekologiczne. Średnica nanorurek haloizytowych HNTs wynosi 10-150 nm a długość ok. 2 μm.



Rysunek 1. Struktura nanorurek haloizytowcyh [2].

2. Nanorurki haloizytowe

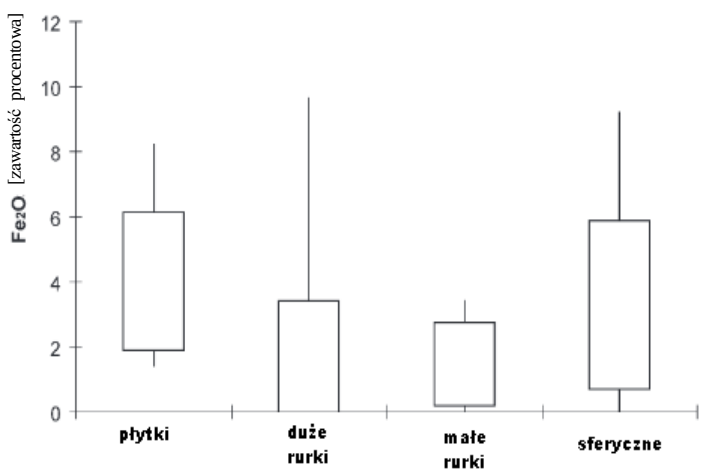
Haloizyt po raz pierwszy został opisany przez Berthiera w 1826 roku. Występuje on w dwóch odmianach polimorficznych [3]. Forma rurek powstała z kaolinitu na skutek działania warunków atmosferycznych oraz procesów hydrotermicznych. W odróżnieniu od innych materiałów występujących w formie rurek, tj. azotku boru, nanorurek węglowych, haloizyt jest powszechnie dostępny, co czyni go niezwykle atrakcyjnym materiałem. Naturalnie występuje on w Chinach, Francji, Belgii oraz Nowej Zelandii. Jednak jedne z największych jego złóż na świecie zlokalizowane są w Polsce – złoża Dunino, gdzie roczne wydobycie wynosi 12 mln ton. Nanokompozyty syntetyzowane na bazie haloizytu są stale badane ze względu na ich właściwości fizykochemiczne, strukturę, hydrofobowość oraz zdolność wymiany jonowej [4]. Szerokie spektrum zastosowań minerału wynika z jego struktury i właściwości:

[1] malgorzata_skibinska90@wp.pl, Zakład Krystalografii, Wydział Chemii, Uniwersytet Marii Curie-Skłodowskiej, www.umcs.pl

- Powierzchnia właściwa – 65 m^2/g
- Objętość porów – 1,25 cm^3/g
- Współczynnik załamania światła – 1,54
- Ciężar właściwy – 2,53 g/cm^3
- Zewnętrzna średnica rury – 50 nm
- Wewnętrzna średnica rury – 15 nm

Haloizyt ma podobny skład chemiczny do kaolinitu, zawiera jednak większe ilości wody [5]. Forma morfologiczna w jakiej występuje minerał uzależniona jest od warunków krystalizacji oraz procesów geologicznych. Dominująca formą morfologiczną są opisywane nanorurki haloizytowe-HNT. Mogą one występować w postaci długich, cienkich rur lub krótkich i grubych.

Haloizyt w zależności od struktury ma podobny skład chemiczny. Poza tlenkiem krzemu i glinu bogaty jest w żelazo. Istnieje pewna zależność pomiędzy zawartością żelaza w strukturze a morfologią minerału. Forma płytkowa jest stosunkowo bogata w Fe z kolei rurkowa ma najmniejszą zawartość tego pierwiastka.



Rysunek 2. Zależność zawartości jonów żelaza od struktury morfologicznej haloizytu [4].

Do zalet nanorurek należą:

- są to materiały naturalne, nietoksyczne, biokompatybilne,
- mają dużą powierzchnię właściwą,
- wysoką pojemność wymiany kationów,
- szybki proces uwalniania np. substancji biologicznie aktywnych, farmaceutyków,
- efektywny proces adsorpcji,
- wysoki współczynnik kształtu, duża porowatość,
- zdolność do regeneracji,
- mogą być stosowane pod wieloma postaciami np. w formie proszków, kremów, żeli, płynów, spreji,
- możliwość modyfikacji i przyłączania wielu grup funkcyjnych jednocześnie [6-9].

Przebadano stabilność nanorurek haloizytowych tak by przeanalizować ich możliwości aplikacyjne. Badano ją w środowisku kwaśnym oraz zasadowym w temperaturze pokojowej. Okazało się, że w 1M roztworze H_2SO_4 zostaje zainicjowane rozpuszczanie haloizytu na wewnętrznej powierzchni, co prowadzi do powstania bezpostaciowych nanocząstek SiO_2. Z kolei w 1M roztworze NaOH rozpuszczaniu wewnętrznej powierzchni towarzyszy tworzenie nanocząstek $Al(OH)_3$ [10].

3. Zastosowanie

Tradycyjnie haloizyt stosowany był do produkcji wysokiej jakości ceramiki, takiej jak cienkościenna porcelana czy do wyrobu tygli. Ostatnio odkryto jednak szereg ciekawych możliwości wykorzystania tanich i powszechnie dostępnych HNT.

Większość badań nanocząstek metalu osadzanych na HNTs koncentruje się głównie na wykorzystaniu metali przejściowych, takich jak Pt, Pd, Ru, Ag i Au. W ten sposób nadając materiałom właściwości katalityczne, magnetyczne i antybakteryjne. Haloizyt badano do zastosowania w nowych i zaawansowanych technologiach ze względu na jego dostępność, łatwość funkcjonalizacji oraz dobrze zdefiniowane struktury. Poprzez modyfikacje materiał ten miał uzyskać specyficzne właściwości chemiczne i fizyczne a tym samym szerokie zastosowanie [11].

4. Nanoreaktory

HNT stosowane są jako nanoreaktory do wytwarzania nanocząstek oraz nanodrutów. Interkalacja i polimeryzacja aniliny we wnętrzu haloizytu sugeruje, że może ona działać jako drut cząsteczkowy. Porowate węgle bogate w mezopory o dużych objętościach zostały przygotowane przez polimeryzację i karbonizację prekursora węgla, sacharozy w HNT. Proces ten pozwolił na otrzymanie węgli o dużej objętości porów oraz wysokiej powierzchni właściwej. Kapsułkowanie anionowych i kationowych metaloporfiryn we wnętrzu HNT również wykazało, że te nowe nieuruchomione katalizatory są obiecującym systemem selektywnych reakcji utleniania. Otrzymano także włókninową, porowatą tkaninę o dobrych właściwościach zwilżania przez bezpośrednie odlewanie dyspersji kompozytowej a następnie przez termiczne usieciowanie [12,13].

5. Procesy kontrolowanego uwalniania

Prowadzono również badania nad zastosowaniem haloizytu jako nośnika leków. Skupiono się tu przede wszystkim na cytokompatybilności oraz stabilności chemicznej HNT. Stwierdzono, że minerał ten może być stosowany do załadunku i przedłużonego uwalniania leków. Naukowcy poświęcili dużo uwagi na badanie termicznej transformacji haloizytu oraz oddziaływań fazowych pomiędzy minerałem a związkiem organicznym. Na ich podstawie wysnuto wnioski, że modyfikowany haloizyt może pełnić rolę gospodarza do załadunku różnych cząstek aktywnie biologicznych. Ponadto minerał ten może być wykorzystany jako metalowy środek kompleksujący [14,15].

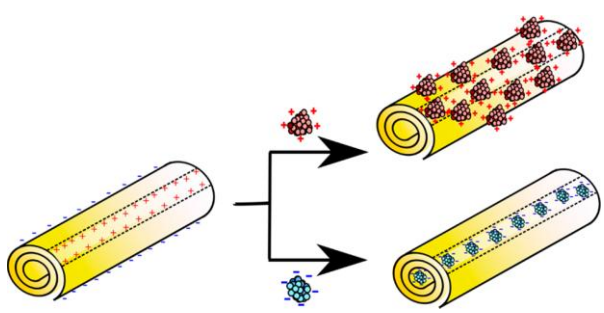
Mniejsze rozmiarowo nanorurki haloizytu są interesujące zwłaszcza dla procesów adsorpcji kompozytów zawierających cząstki o przedłużonym dostarczaniu, takich jak różnego rodzaju farmaceutyki. Haloizyt jest trwały do 460°C, w wyższych temperaturach dochodzi do dehydroksylacji. Najbardziej zewnętrzną warstwę rurek stanowi krzemionka o potencjale powierzchniowym -30 mV przy pH 4-8. Ładunek ten

pozwala na 2-3 godzinną stabilność w wodzie. Potencjał dzeta jest mniejszy niż typowy dla czystych cząstek krzemionki (-50 mV) co można wytłumaczyć różnicą w ładunku obu powierzchni nanorurki [16, 17].

Nanorurki haloizytowe mają gęstość 2,53 g *cm $^{-3}$, ich porowatość wynosi 50-60 cm^2/g, natomiast pojemność wymiany kationów jest równa 30-50* 10^{-2} moli*kg. Porowatość odzwierciedla tu dostępność wewnętrznych i zewnętrznych powierzchni rurek dla procesów adsorpcji, lecz nie obejmuje możliwości zastępowania molekułami gości cząsteczek wody. Przeniesienie interkalowanych cząsteczek wody jest możliwe dzięki jonom solnym lub małym cząsteczkom organicznym takim jak mocznik. Tak więc powierzchnia pomiędzy warstwami może być dostępna do adsorpcji gazowej.

Nanorurki haloizytowe są zdolne do uwięzienia czynników aktywnych wewnątrz nanorurki oraz w pustych przestrzeniach pomiędzy warstwami. Minerał ten może zatrzymać zarówno hydrofilowe, jak i hydrofobowe cząsteczki. W przypadku tych hydrofilowych (NAD, tetracyklina) typowy czas uwalniania wynosi 2-5 godzin, a dla cząsteczek o małej rozpuszczalności w wodzie (deksametazon, furosemid) 5-20 godzin. HNT w połączeniu z istniejącymi środkami farmaceutycznymi zapewniają długotrwałe dostarczanie leku. W porównaniu do nanorurek węglowych są tańsze i mają dużą powierzchnię, dzięki czemu zapewniają większą kontrolę profili obciążenia i elucji. Oprócz tego pozwalają na wyeliminowanie wysokiej początkowej szybkości podawania, uzyskanie niskiego stężenia początkowego. Gaber i Lwowa zasugerowali użycie haloizytu w procesach długotrwałego uwalniania farmaceutyków. Często stosowanymi rozpuszczalnikami są tu woda, alkohol lub aceton. Zazwyczaj skuteczne załadowanie leku osiąga się w przypadku 5-10% wagowych ciężaru nanorurki, biorąc pod uwagę gęstość leku. Jeżeli nanorurki są narażone na działanie wody, wówczas proces uwalniania trwa od 4-12 godzin. Korzystanie z haloizytu jako systemu dostarczania leków w przypadku oparzeń zwiększa czas skuteczności działania farmaceutyków oraz zmniejsza częstotliwość, z jaką musi być zmieniany bandaż. Minerał ten znajduje również zastosowanie w procesach kontrolowanego uwalniania glicerolu jako środka nawilżającego w kosmetykach. Stosowany także jako hipoalergiczny środek czyszczący przeznaczony do usuwania niepożądanych toksyn i olejów, ponieważ glin działa tu jako delikatny złuszczacz [18-26].

Haloizyt może być również wykorzystany w procesach immobilizacji enzymów (rysunek 3), zapewniający długoterminową stabilność enzymatyczną, zwłaszcza w kwaśnym pH.



Rysunek.3. Immoblizacja enzymów przy wykorzystaniu HNTs [18].

W przypadku np. oksydazy glukozowej stabilność temperatury wzrosła zasadniczo, bo aż do 70°C. Typowy czas uwolnienia rozpuszczalnych w wodzie środków chemicznych z nanorurek wynosi 5-10 godzin. Jednak dla wielu zastosowań potrzebne jest by proces ten trwał o wiele dłużej. Uwalnianie leków musi trwać przynajmniej 1-3 dni, preparatów przeciwbakteryjnych przez wiele tygodni, kompozytów przeciwutleniających przez wiele miesięcy a nawet lat, tak by ich działanie było skuteczne. Dopiero haloizyt modyfikowany materiałami polimerowymi pozwala na zwiększenie czas uwalnianie substancji chemicznych i warunkuje ich właściwości ochronne. W wielu przypadkach konieczne jest zaprojektowanie dodatkowej powłoki nanorurek zbudowanej z cienkich warstw polimerowych. Niedawno kilka doniesień wykazało skuteczność haloizytu w dystrybucji leków o działaniu przeciwrakowym. Długotrwałe uwalnianie leku trwało ponad dwa tygodnie [27] .

Mikrocząstki biokompozytowe stały się obiecującym biomateriałem. Komórki żywe mogą być powlekane nanocząstkami magnetycznymi, co pozwala na efektywną manipulację przestrzenną przy użyciu zewnętrznego pola magnetycznego. Kontakt tkanki z haloizytem nie jest szkodliwy. Minerał ten ma dobrą biokompatybilność, która została oceniona zarówno dla kultur komórkowych jak i modeli bezkręgowców. Ponadto w przypadku zastosowania go jako składnika kompozytu polimerowego wpływa on na poprawę biologiczną proliferacji komórek i powlekanie mikroorganizmów. Jednak nie ulega biodegradacji stąd nie można go wstrzykiwać do krwi dożylnie, ale raczej może być stosowany do zewnętrznego leczenia.

Dzięki niskiej toksyczności i biokompatybilności, magnetyczne nanocząstki o właściwościach super paramagnetycznych są intensywnie badane, ze względu na ich potencjalne zastosowania technologiczne, w tym w absorpcji mikrofalowej, bioseparacji, procesach dostarczania leków in vivo, czy też hipertermii leczenia raka. Stąd też zostało zgłoszonych wiele różnych syntez nanokompozytów opartych na haloizycie.

Opracowanie nowych nanokompozytów antybakteryjnych opartych na immobilizacji nanocząstek na powierzchni haloizytu powoduje wzrost zainteresowania tym minerałem. Haloizyt został użyty jako nośnik do ładowania środków przeciwbakteryjnych (Ag). Badania antybakteryjne wykazały, że nanocząstki HNT-Ag wykazują dobre właściwości antybakteryjne wobec bakterii Gram-ujemnych (*Escherichia coli*) i bakterii Gram-dodatnich (*Staphylococcus aureus*).

Wewnętrzne powierzchnie o średnicy 10-20 nm stanowią dobry pojemnik do ładowania białek. Możliwe jest tu wykorzystanie elektrostatycznych interakcji w celu zwiększenia wewnętrznego obciążenia rur. Nanorurki haloizytowe mogą adsorbować białka w ilości 4-8%. Adsorpcja zależy tu od rodzaju białka oraz warunków, w których procedura adsorpcji jest wykonywana. Charakterystyczne są tu interakcje elektrostatyczne pomiędzy ładunkiem białka a wewnętrznym ujemnym i zewnętrznym dodatnim ładunkiem nanorurek. Pozytywnie naładowane białka adsorbują w większych ilościach niż te naładowane ujemnie, ze względu na większą powierzchnię ujemnie naładowanej warstwy haloizytu. Zaadsorbowane białka wykazują zwiększoną stabilność termiczną oraz czasowe zdolności biokatalityczne, szczególnie przy pH powyżej lub poniżej optimum enzymu.

Porowatość materiału wzrasta 6-7 razy na skutek traktowania kwasem lub zasadą. Adsorpcja cząsteczek rozpuszczalnych w wodzie na haloizycie jest w przybliżeniu dwukrotnie wyższa niż w przypadku kaolinitu. Wynika to z podwójnego charakteru powierzchni haloizytu [28].

6. Nanokompozyty polimerowe

Haloizyt wykorzystywany jest jako dodatek do tworzyw sztucznych. Wpływa przede wszystkim na ich poprawę stabilności termicznej, nie zmieniając przy tym właściwości mechanicznych i użytkowych produktu. Jego wysoki współczynnik kształtu korzystnie wpływa na właściwości produktu, a ze względu na polarny charakter łatwo miesza się z niepolarnymi polimerami. Zewnętrzna powierzchnia HNT ma względnie małą gęstość grup hydroksylowych (Si-OH) i ładunek ujemny, podczas gdy ta wewnętrzna jest obfita w grupy hydroksylowe (Al-OH) i ładunek dodatni. Stąd też zewnętrzna powierzchnia jest doskonała do wytwarzania nanokompozytów polimerowych [29]. Różnice we właściwościach chemicznych i elektrycznych powierzchni minerału wynikają z występowania różnych struktur krystalicznych. Unikalna struktura sprawia, że nanokompozyty polimerowe mogą być wytwarzane różnymi metodami syntezy. Istnieje wiele publikacji na temat domieszkowania haloizytu polimerami. Rozpuszczalne w wodzie polimery można mieszać bezpośrednio z minerałem. Siłę rurek oszacowano na podstawie eksperymentów zginania bezpośredniego. Moduł Younga pojedynczych rur wynosi około 130 GPa i ma dużą elastyczność. Dodatek 3-5% wag. polimeru podwaja wytrzymałość na rozciąganie kompozytu i zasadniczo wzrasta jego przyczepność do podłoża stałego. Nanorurki haloizytowe są polarne a więc zapewniają dobrą hydrofilowość, dobrą dyspersję w polimerach polarnych, takich jak epoksyd, poliamidy, poliakrylany i biopolimery (pektyna, skrobia). Załadunek polimerów rozpuszczalnych w wodzie może być wysoki 50-60% wagowych i prowadzić do wytworzenia stałych kompozytów organiczno-nieorganicznych. Polimerowo-haloizytowe kompozyty modyfikowane przez półspalanie plastyfikatora (tributylocytratu) znajdują zastosowanie w przemyśle motoryzacyjnym. Polimery z domieszką 5-10% wagowych haloizytu mają dużo lepsze właściwości palne, zwykle prędkość spalania spada 2-3 razy w takich kompozytach. Ze względu na dobrą stabilność cieplną, wysoką przewodność cieplną, zdolność do zachowania kształtu podczas topienia np. wosku, oraz bogatej dostępności minerału kompozyty izolacyjne można wykorzystać do oszczędzania energii na dużą skalę. Jedyną wadą tych materiałów jest możliwość katalitycznego rozkładu polimerów z haloizytem [30-40].

7. Kataliza

Nanokompozyty oparte na HNT stosowane w katalizie mają wiele korzyści. Są nimi proste odzyskiwanie, stosowanie katalitycznych ilości, łatwość pracy, wysoka wydajność i selektywność. Cechy te czynią je niezwykle użytecznymi w zakresie tworzenia przyjaznych dla środowiska technologii. Nanorurki powleczone metalem lub tlenkami metali są skutecznymi mezoporowatymi mediami do zaawansowanej katalizy. W celu umieszczenia metalowych nanocząstek na powierzchni rurki przeprowadzane są standardowe reakcje tworzenia cząstek metalu. Może być to synteza nanocząstek w obecności HNT lub też załadowanie elementów reaktora do wnętrza nanorurki i przeprowadzenie syntezy cząstek wyłącznie wewnątrz rury. Wówczas tworzą się okrągłe lub podobne do prętów nanocząstki. Dzięki ograniczeniu objętości reakcji do nanoskali można wytwarzać metalowe nanocząstki, które znajdują się wewnątrz przewodu rurowego. Istnieje również możliwość zaprojektowania dwuskładnikowego systemu katalitycznego, gdy jeden typ nanocząstek jest umieszczony wewnątrz nanorurek gliniastych a drugi na zewnątrz rur.

Za przykład może posłużyć materiał, gdzie obecność HNT wpłynęła na poprawę dyspergowalności i fotokatalitycznej aktywności nanocząstek TiO_2. HNT działa jak stabilizator zapobiegając agregacji nanocząstek i zwiększając wydajność katalizatora. Związane jest to z jego strukturą, a mianowicie dużą powierzchniową właściwą, objętością i średnicą porów. Nanokompozyt HNT-TiO_2 otrzymany był metodą zol-żel z izopropoksylanem jako prekursorem. Ponadto połączenie to można stosować w celu zwiększenia usuwania organicznych zanieczyszczeń poprzez procesy absorpcji [41]. Nanocząstki rozproszone na zewnętrznej powierzchni HNT są bardziej aktywne niż te na powierzchni wewnętrznej. Zhang i Yang przedstawili skuteczną metodę syntezy, która umożliwiła selektywną lokalizację nanocząstek wewnątrz HNT w oparciu o różnicę w przepływie energii roztworu wodnego i organicznego. Brak aktywności fotokatalitycznej wykazywały nanocząstki naniesione na wewnętrzną powierzchnię. Co_3O_4 adsorbowany tylko na zewnętrznej powierzchni nanorurek wykazywał wyższą aktywność fotodegradacji do niebieskiego błękitu (MB) niż HNT, czyste nanocząstki Co_3O_4 i HNT-Co_3O_4 (otrzymany na skutek mieszania- tak, że Co_3O_4 adsorbowało się na zewnętrznej i wewnętrznej powierzchni nanorurek), ponieważ nanocząstki Co_3O_4 osadzone lub zapakowane we wnętrzu nie mogły zaadsorbować światła, a tym samym ograniczona była reakcja fotokatalityczna. Inne ciekawe materiały oparte na minerale haloizyt otrzymano stosując nanocząstki Fe_3O_4. HNT-Fe_3O_4 posiadało dobre własności magnetyczne, wyjątkowo małą pętlę histerezy i niską koercję, typową dla cząstek superparamagnetycznych. Magnetyczne HNTs-Fe_3O_4 wykazywały lepszą adsorpcję barwników i można je było łatwo wydzielić z roztworu wodnego w magnetycznym polu. Tak więc HNTs-Fe_3O_4 może być oszczędnym i wydajnym adsorbentem, gdzie mamy do czynienia z połączeniem technologii separacji magnetycznej i procesu adsorpcji.

Istnieje wiele metod opisanych w literaturze dotyczących nanoszenia nanocząstek metali na podłoża HNT, z których każda spodziewa się znacznej poprawy wydajności nowych opartych na haloizycie funkcjonalnych nanokompozytów. Nowy nanokatalizator HNTs-CdS zsyntetyzowano metodą hydrotermiczną z bezpośrednim wzrostem nanocząstki CdS na powierzchni HNT. Zaobserwowano, że nanocząstki CdS nieregularnie rozprowadzają się na powierzchni z HNT, a wielkość CdS na powierzchni HNT zmniejszyła się w porównaniu z czystym CdS. Papoulis stworzył nanokompozyty HNT-TiO_2 poprzez naniesienie anatazu na powierzchnię glinu metodą zol-żel. Metoda impregnacji jest stosowana w metalowych roztworach koloidalnych lub roztworach prekursora. Z kolei Tang zaproponował, aby wytworzyć katalizatory nanocząsteczkowe HNT-FeNiO i HNs-FeNi za pomocą metody koimprezenacji [42-45].

8. Adsorbent

HNT jest bardzo dobrą alternatywą dla adsorbentów do usuwania jonów metali ciężkich ze ścieków. Ma to ogromne znaczenie w ochronie środowiska. Naturalny haloizyt, bez modyfikacji najczęściej usuwa te jony poprzez mechanizmy fizycznej lub chemicznej sorpcji, specjacji. Jednak w celu zwiększenia powinowactwa do jonów metali ciężkich minerał często jest domieszkowany różnymi nanomateriałami lub grupami funkcyjnymi. HNT z octanem sodu wykazuje wysoką skuteczność jako adsorbent Cu(III), z dietanoaminą i trietanoaminą jako adsorbent Pb(III), Cd(III), Zn(II). W tym drugim przypadku mamy do czynienia z dwustopniowym procesem dyfuzji metali w przestrzeni międzywarstwowej a następnie wiązaniem azotu aminowego. Do usuwania Cr(VI) stosowano nanokompozyt wytworzony przez poli-

meryzację pirolu na HNT. By zwiększyć tu współczynnik adsorpcji zmodyfikowano go dodatkowo bromkiem heksadecylotrimetyloamoniowym. Z kolei haloizyt modyfikowany kwasem 2-hydroksybenzoesowym znalazł zastosowanie jako sorbent Fe(III). Wszystkie adsorbenty oparte na HNT charakteryzowały się niskimi kosztami produkcji, łatwością obsługi oraz możliwością wielokrotnego użytku [46-53]. Ze względu na różny ładunek powierzchni minerału może usuwać zarówno kationowe jak i anionowe barwniki. Magnetycznie modyfikowane HNT są łatwiejsze do wydzielenia z roztworu po procesie adsorpcji barwników. Fe_3O_4-HNT otrzymany metodą chemicznego strącania wykazał dużą zdolność adsorpcji fioletu metylowego. Funkcjonalizowanie powierzchni haloizytu przy pomocy TiO_2 czy też ZnO wykorzystano do usunięcia barwników z wody na bazie podwójnego mechanizmu adsorpcji i fotodegradacji. Zbadano również wpływ węgla grafitowego na HNT i wykazano, że znacznie poprawia on aktywność fotokatalityczną oraz przewodnictwo. Morfologiczne pory haloizytu mogą zawierać wodę, przy czym stosunek wilgotności jest proporcjonalny do jego porowatości, wyższa porowatość zapewnia więcej wody, co prowadzi do większej przewodności. Powłoka węglowa może tutaj zwiększyć zdolność adsorpcyjną wskutek synergii między cząstkami węgla a HNT. W opisanym poniżej przypadku pokryta węglem powierzchnia powoduje, że adsorbowane jony dodatnie mogą się swobodnie poruszać we wnętrzu nanorurki. Kompozyt ten ma dwie zauważalne zalety wynikające z unikalnych właściwości węgla, a mianowicie zwiększoną adsorpcję zanieczyszczeń i łatwiejszy transport oraz separację. Procesy adsorpcji są jedną ze skuteczniejszych metod usuwania barwników. Struktura adsorbentów jest ważna dla uzyskania jak najlepszych zdolności adsorpcji. Większa powierzchnia właściwa oraz szersze pory warunkują wyższą zdolność adsorpcji cząsteczek barwnika. W porównaniu z innymi minerałami z grupy kaolinitów, HNT wykazuje najwyższą pojemność sorpcji. Haloizyt jest biokompatybilnym i bezpiecznym nanomateriałem stąd może być stosowany w systemach filtracji wody. Skuteczność usuwania metali ciężkich jest wyższa niż w przypadku większości konwencjonalnych minerałów. Minerał ten może być dodatkowo regenerowany przez spalanie adsorbowanych substancji organicznych Jeszcze skuteczniejsze oczyszczanie wody występuje w przypadku modyfikacji nanorurki polikationami, takimi jak polietylenoimina lub chitosan. Minerał może być zastosowany również do wychwytywania CO_2 z otaczającego powietrza. Wysoka efektywność procesu oraz doskonała stabilność powierzchni haloizytu czyni procedurę przyjazną dla środowiska [54-56].

9. Inne zastosowania

Zewnętrzne powierzchnie HNT są zbudowane z tetraedrów SiO_4 i mają kilka grup hydroksylowych, co wskazuje, że HNT posiadają taką samą właściwość rozpraszania jak inne naturalne krzemiany (np. montmorylonit, kaolinit) i potencjalną zdolność do tworzenia

wiązania wodorowego. Przeprowadzono modyfikacje HNT Pd w celu zwiększenia pojemności wodoru HNT. Zdolność adsorpcji haloizytu modyfikowanego była o 1.14% wyższa niż HNT (0.43%).

Haloizyt stosowano również w leczeniu zwierząt, jako dodatek do pasz.

Przechwytywanie krążących komórek nowotworowych (CTC) we krwi może być skuteczniejsze dzięki wprowadzeniu HNT na powierzchnię urządzeń stosowanych w tym celu. Powłoki te promują zwiększone wychwytywanie komórek białaczkowych oraz okręcają kluczowe parametry kontrolujące przepływ, np. gęstość selekcji [10].

Haloizyt może być stosowany jako dodatek w farbach do wytwarzania funkcjonalnego kompozytowego materiału powłokowego. Możliwe jest otrzymanie aktywnej powłoki antykorozyjnej, złożonej z folii zol-żelu domieszkowanych haloizytem i uwalnianie zablokowanych inhibitorów korozji w sposób kontrolowany. Folia zol-żel wraz z dodatkiem minerału wykazuje ulepszoną i długotrwałą ochronę przed korozją. Wprowadzenie benzotriazolu do farb w nanohaloizytowym układzie kapsułkującym poprawia jego działanie antykorozyjne. Bezpośredni dodatek związku zmniejsza jego właściwości ochronne [1]. Opracowano również nowy nawóz mineralny superfosforan (SSP) w postaci sproszkowanej z dodatkiem haloizytu. Minerał wykazuje tu wiele korzyści:

- poprawia strukturę gleby i płodność,
- wzbogaca gleby o składniki odżywcze,
- poprawia retencję wody przez glebę, zwiększa wchłanianie składników odżywczych przez rośliny,
- zwiększa aktywność mikrobiologiczną,
- zmniejsza wchłanianie metali ciężkich przez rośliny (kadm, rtęć, ołów, nikiel) [57,58].

Opisane wcześniej różnice w budowie chemicznej wewnętrznej i zewnętrznej nanorurek sprawiają, iż mogą one być poddawane selektywnej modyfikacji. Nano i mikrocząstki mogą być stosowane do emulgowania oleju, gdzie składnik olejowy oddzielany jest od wody nanocząstką. Dodanie 0,5-1% wagowych haloizytu do 20% wag. dekanu pozwala na tworzenie koloidów o średnicy 40-50 μm, które są stabilne przez kilka tygodni. W literaturze znajdujemy informacje na temat zaproponowania takiego podejście w przypadku wycieku ropy naftowej. Adsorpcja środków powierzchniowo czynnych (DOSS, Tween, Span) na haloizycie jest stosowana w przemyśle naftowym do rozlewania. Powlekana powierzchnia haloizytu jest korzystna również do kotwiczenia i proliferacji komórek biologicznych [10].

10. Podsumowanie

Nanorurki haloizytowi są unikatowym nanomateriałem ze względu na swą podwójną strukturę. Dodatkowo nie wykazują właściwości toksycznych, mają wysoki stosunek L/D oraz niską gęstość grup hydroksylowych na powierzchni. Przede wszystkim jednak są materiałem tanim i powszechnie dostępnym, mogącym być w łatwy sposób modyfikowany, dlatego znajdują szereg zastosowań. HNT posiadają obiecujące perspektywy w przygotowaniu nowych materiałów konstrukcyjnych i funkcjonalnych.

Literatura

1. Du M., Guo B., Jia D., *Newly emerging applications of halloysite nanotubes: a review*, Polym Int, 59, (2010), s. 574-582.
2. Kamble R., Ghang M., Gaikawad S., Panda B. K., *Halloysite Nanotubes and Application: A Review*, Journal of Advanced Scientific Research, 3(2), (2012), s. 25-29.
3. Macewan MDC., *The Clay Minerals Halloysite and Meta-Halloysite*, Nature,157 (1946), s. 159.
4. Joussein E., Petit S., Churchman J., Theng B., Right D., Delvaux D., *Halloysite clay minerals-a review*, Clay Minerals, 40, (2005), s. 383.
5. Soheilmoghaddam M., Wahit M. U., Mohmoudian S., Hanid A. N., *Regenerated cellulose/halloysite nanotube nanocompositie films prepared with ioni liquid*, Materials Chemistry and Physics, 141, (2013), s. 936-943.

6. Gonchar K. A., Kondakova A. V., Jana S., Timoshenko V. Yu, Vasiliev A. N., *Investigation of Halloysite Nanotubes with Deposited Silver Nanoparticles by Methods of Optical Spectroscopy*, Physics of the Solid State, 58 (3), (2016), s. 601-605.
7. Lee S.Y., Kim S. J., *Adsorption of naphthalene by HDTMA modified kaolinite and halloysite* , Applied Clay Science, 22, (2002), s. 55.
8. Peixoto A. F. , Fernandes A. C. , Pereira C., Pires J. , Freire C., *Physicochemical characterization of organosilylated halloysite clay nanotubes,* Microporous and Mesoporous Materials, 219, (2016), s. 145.
9. Szczepanik B., Słomkiewicz P., Garnuszek M., Czech K., Banaś D., Kubala-Kukuś A., Stabrawa I., *The effect of chemical modification on the physico-chemical characteristics of halloysite: FTIR, XRF, and XRD studies,* J. Mol. Struct. 1084, (2015), s. 16-22.
10. Rawtani D., Agrawal Y.K., *Multifarious applications of haloysite nanotubes: a review,* Material Science, 30, (2012), s. 282-295.
11. Shemesh R., Krepker M., Natan M., Danin-POleg Y., Banin E., Kashi Y., Nitzan N., Vaxman A., Segal E., *Novel LDPE/halloysite nanotube films with sustained carvacrol release for broad-spectrum antimicrobial activity,* The Royal Society of Chemistry, 5, (2015), s. 87108.
12. Shchukin D. G., Sukhorukow G. B., Price R. R., Lvov Y. M., *Halloysite clay nanotubes for controlled release of protective agents,* Small, 1 (5), (2005), s. 510-513
13. Zhang Y., Yang H. M., Co3O4 nanoparticles on the surface of halloysite nanotubes, Phys.Chem.Miner.39, (2012), s. 789-795.
14. Yuan P., Tan D., Annabi-Bergaya F., *Properties and applications of halloysite nanotubes: recent research advances and future prospects,* Applied Clay Science, 75 (2015), s. 112-113.
15. Zhang Y., Tang A.,Yang H., Ouyang J., *Applications and interfaces of halloysite nanocompositions,* Applied Clay Science, 119, (2016), s. 8-17.
16. Sakiewicz P., Nowosielski R., Pilarczyk W., K. Gołombek, M. Lutyński, *Selected properties of the halloysite as a component of Geosynthetic Clay Liners*, Journal of Achievements in Materials and Manufacturing Engineering, 48, (2011), s. 117.
17. Matusik J., *Arsentae, orthophosphate, sulfate and nitrate sorption equilibris and kinetics for halloysite Chemicand kaolinities with an induced positive charg*e, Engineering Journal, 246, (2014), s. 244-253.
18. Tully J., Yendluri R., Lvov Y., *Halloysite Clay Nanotubes for Enzyme Immobilization,* Biomacromolecules, 17, (2016), s. 615-621.
19. Zhai R., Zhang B., Liu L., Xie Y., Zhang H., Liu J., *Immobilization of enzyme biocatalyst on natural halloysite nanotues,* Catalysis Communications, 12, (2010), s. 259-263.
20. Levis, S.R., Deasy, P.B., 2002. *Characterisation of halloysite for use as a microtubular drug delivery system*. Int. J. Pharm. 243, s. 125-134.
21. Tan, D., Yuan, P., Annabi-Bergaya, F., Yu, H., Liu, D., Liu, H., He, H., *Natural halloysite nanotubes asmesoporous carriers for the loading of ibuprofen*. MicroporousMesoporous Mater. 179, 2013, s. 89-98.
22. Tan, D.Y., Yuan, P., Annabi-Bergaya, F., Liu, D.,Wang, L.J., Liu, H.M., He, H.P., 2014. *Loading and in vitro release of ibuprofen in tubular halloysite*. Appl. Clay Sci. 96, s. 50-55.
23. Vergaro V., Lvov Y. M. , Leporatti S., *Cytocompability and uptake of halloysite clay nanotubes,* Colloids Surf. A Physicochem. Eng. Asp. 396, (2012), s. 182.
24. Shi Y. F. ,Tian Z., Zhang Y., Shen H. B., Jia N. Q, *Functonallized halloysite nanotube-based carrier for intracellular delivery of antisense,* Nanoscale Res. Lett. 6 1 (2011).
25. Konnova S. A., Sharipova I. R. , Demina T. A., Osin Y. N., D. R. Yarullina, Ilinskaya O. N. , Lvov Y. M., Fakhrullin R. F., *Biomimetic cell- mediated three-dimensional assembly of halloysite nanotubes,* Chem. Commun. 49 4208 (2013).
26. Rao K. M. , Nagappan S., Seo D. J. , Ha C.-S. , *pH sensitive halloyste- sodium hyaluronate/poly(hydroxyethyl methaacrylat) nanocomposites for cancer drug delivery,* Appl. Clay Sci., 33 (2014), s. 97-98.

27. Sakiewicz P. Nowosielski R., Pilarczyk W., Gołombek K., Lutyński M., *Selected properties of the halloysite as a component of Geosynthetic Clay Liners,* Journal of Achievements in Materials and Manufacturing Engineering, 48, (2011), s. 117.
28. Abdullayev E., Lvov Y., *Halloysite clay nanotubes as a ceramic "skeloton" for functional biopolymer composites with sustained drug release*, J. Mater. Chem. B, (2013).
29. Szpilska K., Kudła S., Czaja K., *Kompozyty poliolefinowe napełnione modyfikowanymi nanorurkami haloizytowymi*, Przemysł Chemiczny, 94/12, (2015), s. 2130-2132.
30. Marini J., Pollet E., Averous L., Bretas R. E. S., *Elaboration an properties of novel biobased nanocomposites with halloysit nanotubes and thermoplastic polyurethane from dimerize fatty acids*, Polymer, 55, (2014), s. 5226.
31. Carli L.N., Crespo J.S. , Mauler R.S, *PHBV nanocomposites based on organomodified montmorillonite and halloysite: the effect of clay type on the morphology and thermal and mechanical properties*, Compos. Part A, 42, (2011), s. 1601-1608.
32. Guo B., Lei Y., Chen F., Liu X., Du M., Jia D., *Styrene–butadiene rubber/halloysite nanotubes nanocomposites modified by methacrylic acid*, Applied Surface Science, 255, (2008), s. 2715-2722.
33. Ning N-Y., Yin Q-J., Luo F., Zhang Q., Du R., Fu Q., *Crystallization behavior and mechanical properties of polypropylene/halloysite composites*, Polymer, 48 (2007), s. 7374-7384.
34. Pasbakhsh P., Ismail H., Fauzi M. N. A., Bakar A. A., *EPDM/modified halloysite nanocomposites*, Applied Clay Science, 48, (2010), s. 405-413.
35. Prashantha K., Lacrampe M. F. , Krawczak P., *Processing and characterization of halloysite nanotubes filled polypropylene nanocomposites based on a masterbatch route: effect of halloysites treatment on structural and mechanical properties*, Express Polym. Lett. 5,(2011), s. 295-307.
36. Das A., Thakur V., Mahaling R. N., Bhowmick A. K., Heinrich G., *Preparation and properties of natural nanocomposites based on natural rubber and naturally occurring halloysite nanotubes*, Mater. Des., 31, (2010), s. 2151-2156.
37. Liu M., Jia Z., Jia D., Zhou C., *Recent advance in research on halloysite nanotubes-polymer nanocomposite*, Progress in Polymer Science, 39, (2014), s. 1498.
38. Carli L. N. , Daitx T. S. , Soares G. V., Crespo J. S. , Mauler R. S., *The effects of silane coupling agents on the properties of PHBV/halloysite nanocomposites*, Applied Clay Science, 87, (2014), s. 311.
39. Albdiry M. T., Yousif B. F., *Role of silanized halloysite nanotubes on structural, mechanical properties and facture toughness of thermoset nanocompsites,* Materials and Design, 57, (2014), s. 279-288.
40. Bischoff E., Daitx T., Simon D. A., Schrekker H. S., Liberman S. A., Mauler R. S., *Organosilane-functionalized halloysite for high performance halloysite/heterophasic ethylene–propylene copolymer nanocomposites*, Applied Clay Science, 112, (2015), s. 68-74.
41. Li C.P.,Wang J.Q., Feng S.Q., Yang Z.L., Ding S.J., Journal of Mater. Chem. A, 1, (2013), s. 804.
42. Machado G. S., Castro D. F., Wypych F., Nakagakia S. G. S., *Immobilization of metalloporphyrins into nanotubes of natural halloysite toward selective catalysts for oxidation reactions*, Journal of Molecular Catalysis: A Chemical, 283, (2008), s. 99-107.
43. Zatta L.,Gardolinski J.,Wypych F., *Raw halloysites asreusable heterogeneous catalyst for esterificationoflauric acid*., Applied Clay Science, 51, (2011), s. 165-169.
44. Barrientos-Ramírez S., Ramos-Fernández E. V., Silvestre-Albero J., Sepúlveda- Escribano A., Pastor-Blas M. M., González-Montiel A., *Use of nanotubes of natural halloysite as catalyst support in the atom transfer radical polymerization of methyl methacrylate*, Microporous and Mesoporous Materials, 120, (2009), s. 132-140.
45. Tang X., Li L., Shen B., Wang C., *Halloysite- nanotubes supported FeNi alloy nanoparticles for catalytic decomposition of toxic phosphine gas into yellow phosphorus and hydrogen*, Chemosphere, 91, (2013), s. 1368.

46. Yu L., Wang H., Zang Y., Zhang B., Liu J., *Recent advances in halloysite nanotube derived composities for water treatment*, Environmental Science, 3, (2016), s. 28.
47. Viseras M. T. , Aguzzi C., Cerezo P., Viseras C., Valenzuela C., *Equilibrium and kinetics of 5-aminosalicylic acid adsorption by halloysite*, Microporous and Mesoporous Materials, 108, (2008), s. 112-116.
48. Kiani G., *High removal capacity of silver ions from aqueous solution onto Halloysite nanotubes*, Applied Clay Science, 90, (2014), s. 159-164.
49. M. Lutyński, P. Sakiewicz, M. A. Gonzales, Inżynieria Mineralna, 111 (2014).
50. Zhang C-L., Cui S-J., Wang Y., *Adsorption removal of pefloxacin from water by halloysite nanotubes*, Journal of Industrial and Engineering Chemistry, 23, (2015), s. 12.
51. Ahammed Shabeer T. P. , Saha A., Gajbhiye V. T., Gupta S., Manjaiah K. M., Varghese E., *Exploitation of Nano-Bentonite, Nano-Halloysite and Organicalyy Modified Nano-Montrmorillonite as an Adsorbent and Coagulation Aid for the Removal of Multi-Pesticied from Water: A Sorption Modelling Approach*, Water Air Soil Pollut, 41, (2015), s. 226.
52. Ferrante F., Armata N., Cavallaro G., Lazzara G., *Adsorption Studies of Molecules on the Halloysite Surfaces: A Computational and Experimental Investigation*, The Journal of Physical Chemistry, 121, (2017), s. 2951.
53. Świercz A., Smorzewska E., Słomkiewicz P., Suchanek G., *Possible use of halloysite in phytoremediation of soils contaminated with heavy metals*, Journal of Elementology, 21(2), (2016), s. 559.
54. Liu L., Wan Y. Z., Xie Y. D., Zhai R., Zhang B., Liu J. D., *The removal of dye from aqueous solution using alginate-halloysite nanotube beads*, Chemical . Engeeniering Journal, 187, (2012), s. 210-216.
55. Kiani G., Dostali M., Rostami A., Khataee A. R., *Adsorption studies on the removal of Malachite Green from aqueous solutions onto halloysite nanotubes*, Applied Clay Science, 54, (2011), s. 34-39.
56. Zhao M. F. , Liu P., *Adsorption behaviour of methylene blue on halloysite nanotubes*, Microporous and Mesoporous Materials, 112, (2008), s. 419-424.
57. Ahammed Shabeer T. P., Saha A., Gajbhiye V. T. , Gupta S., K. M. Manjaiah, Varghese E., *Exploitation of Nano-Bentonite, Nano-Halloysite and Organically Modified Nano-Montmorillonite as an Adsorbent and Coagulation Aid for the Removal of Multi-Pesticides from Water: A Sorption Modelling Approach*, Water Air Soil Pollut, 226, (2015), s. 41.
58. Yah W. O., Takahara A., Lvov Y. M., *Selective Modification of Halloysite Lumen with Octadecylphosphonic Acid: New Inorganic Tubular Micelle*, J. Am. Chem. Soc., 134, (2012a), s. 1853-1859.

Zastosowanie nanorurek haloizytowych

Abstrakt
Haloizyt jest naturalnym dwuwarstwowym glinokrzemianem. Jest materiałem powszechnie dostępnym i tanim. Ze względu na strukturę znajduje szerokie zastosowanie w medycynie, katalizie, procesach sorpcji oraz kontrolowanego uwalniania.
Słowa kluczowe: haloizyt, naturalne nanorurki, glinokrzemian

The application of halloysite nanotubes

Abstract
Halloysite is a natural two-layer aluminosilicate. It is a widely available and cheap material. Thanks to its unique structure, it is widely used in medicine, catalysis, sorption and controlled release.
Keywords: halloysite, natural nanotubes, aluminosilicate

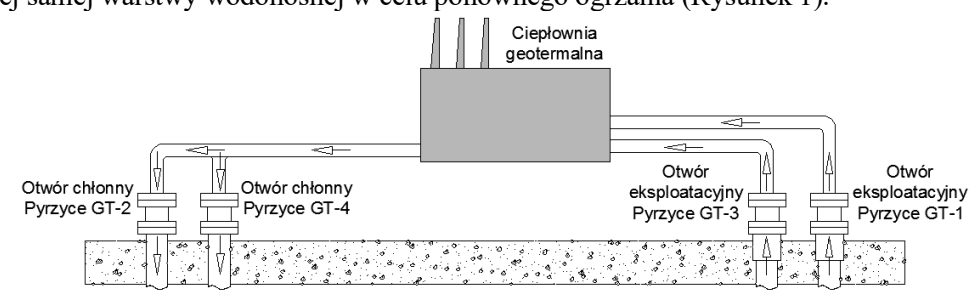
Bogdan Noga[1], Marcin Mazur[2], Agnieszka Wichowska[3]

Nowa technologia wspomagająca eksploatację wód termalnych w ciepłowniach geotermalnych

1. Wprowadzenie

Nowa technologia wspomagająca eksploatację wód termalnych w ciepłowniach geotermalnych została opracowana w ramach badań prowadzonych przez firmę HPC POLGEOL S.A (dawniej Przedsiębiorstwo Geologiczne POLGEOL S.A). Badania prowadzono również na instalacji geotermalnej udostępnionej przez Geotermię Pyrzyce. Badania były współfinansowane przez Narodowe Centrum Badań i Rozwoju oraz Narodowy Fundusz Ochrony Środowiska i Gospodarki Wodnej w ramach programu GEKON – Generator Koncepcji Ekologicznych.

Geotermia Pyrzyce pracuje w tzw. dublecie geologicznym, w którego skład wchodzą dwa otwory wydobywcze wody termalnej oraz dwa otwory chłonne (zatłaczające) (Rysunek 1). Woda termalna mająca temperaturę 58-61°C (na wypływie z otworu) oraz mineralizację około 121 g/dm^3 pobierana jest z piaskowców jury dolnej (lias, warstwy mechowskie). Temperatura wody termalnej w złożu wynosi 64°C. Następnie przetłaczana jest rurociągiem do hali ciepłowni geotermalnej gdzie następuje jej schłodzenie za pomocą wymienników ciepła. Po schłodzeniu przetłaczana jest ona w kierunku otworów chłonnych (na odległość około 1500 m) i zatłaczana do tej samej warstwy wodonośnej w celu ponownego ogrzania (Rysunek 1).



Rysunek 1. Uproszczony schemat obiegu wody termalnej w ciepłowni geotermalnej w Pyrzycach [opracowanie własne]

W procesie technologicznym woda termalna stosowana jest w obiegu zamkniętym przepływając przez wymiennik ciepła. Woda z otworów wydobywczych do otworów chłonnych przetłaczana jest bez kontaktu z otoczeniem. Przez większą część roku produkcja ciepła pochodzącego z układu geotermalnego pokrywa całe zapotrzebowanie

[1] bogdan.noga@polgeol.pl, Instytut Mechaniki Stosowanej i Energetyki, Uniwersytet Technologiczno-Humanistyczny im. K. Pułaskiego w Radomiu, www.imsie.pr.radom.pl
[2] marcin.mazur@polgeol.pl, HPC Polgeol S.A, www.polgeol.pl
[3] agnieszka.wichowska@polgeol.pl, HPC Polgeol S.A, www.polgeol.pl

jego odbiorców. W sezonie grzewczym, przy dużych mrozach, układ geotermalny wspomagany jest absorpcyjną pompą ciepła, która napędzana jest gorącą wodą o temperaturze 160°C. Aby uzyskać ciepło napędowe, pompa ciepła (moc pompy to 10 MW) współpracuje z wysokotemperaturowym kotłem gazowym o mocy 8 MW. Szczytowe źródło ciepła stanowią dwa niskotemperaturowe kotły o mocy 10 MW każdy.

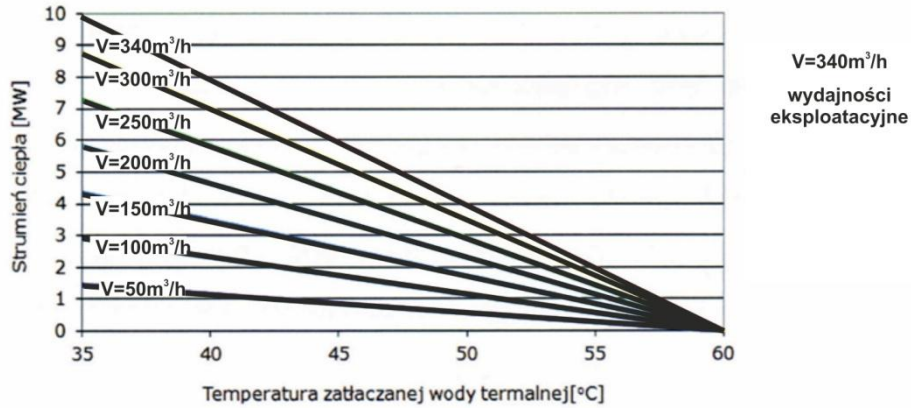
Kluczowym parametrem wpływającym na ilość wyprodukowanego ciepła geotermalnego jest wydajność zatłaczania wody termalnej oraz jej temperatura na wypływie z otworu i jej temperatura zatłaczania.

W Geotermii Pyrzyce otworami wydobywczymi są otwory Pyrzyce GT-1 oraz Pyrzyce GT-3, natomiast otworami chłonnymi są otwory Pyrzyce GT-2 oraz Pyrzyce GT-4. W zawiązku z funkcjami, jakie miały spełniać otwory geotermalne wykonano je parami – otwory chłonne około 1,5 km na północ od ciepłowni oraz wydobywcze w bezpośredniej jej bliskości. Takie rozmieszczenie było konieczne do wyeliminowania wpływu zatłaczanych, wychłodzonych wód, na otwory wydobywcze.

Plany wykorzystania ciepła ziemi w Geotermii Pyrzyce były bardzo ambitne. Eksploatacja wody termalnej miała wynosić 170 m^3/h [5] z każdego z otworów (Pyrzyce GT-1, Pyrzyce GT-3). Miało być to w sumie do 340 m^3/h wody termalnej o temperaturze 61°C (na wypływie z otworów). Zakładając schłodzenie wody do temperatury 35°C uzyskiwana moc cieplna członu geotermalnego miała wynosić około 10 MW.

Problemy związane z kolmatacją spowodowały, że wydajności zatłaczania wody termalnej do otworów chłonnych nie były w stanie umożliwić eksploatacji otworu Pyrzyce GT-1 z wydatkami większymi niż 130-140 m^3/h. Ilość ciepła uzyskana z członu geotermalnego jest związana w sposób liniowy z wydajnością eksploatacji oraz temperaturą schłodzenia (Rysunek 2). Zmniejszanie się chłonności otworów zatłaczających przekładało się na mniejszą ilość wyprodukowanego ciepła. Przy wydajnościach 130-140 m^3/h, na które składało się wydobycie wody termalnej wyłącznie otworem Pyrzyce GT-1 oraz zatłaczanie do dwóch otworów chłonnych, produkcja ciepła wynosiła zaledwie około 4-4, 5 MW (Rysunek 2), czyli stanowiła ok. 40% mocy planowanej.

Rozwój problemów z chłonnością otworów zatłaczających Pyrzyce GT-2 oraz Pyrzyce GT-4 doprowadził do zmiany statusu otworu Pyrzyce GT-3. W roku 2010 nastąpiła zmiana otworu wydobywczego Pyrzyce GT-3 na otwór chłonny. Od tego momentu ciepłownia geotermalna w Pyrzycach pracuje w oparciu o jeden otwór wydobywczy i 3 otwory chłonne. Woda termalna eksploatowana jest z wydajnością około 80-110 m^3/h. Otwór Pyrzyce GT-3 został włączony do obiegu geotermalnego jako kolejny otwór chłonny. Otwór Pyrzyce GT-3 nigdy nie pełnił roli otworu wydobywczego, gdyż wydajność zatłaczania otworów chłonnych była na tyle mała, że eksploatacja wody termalnej otworem Pyrzyce GT-1 była wystarczająca.



Rysunek 2. Możliwości pozyskiwania energii geotermalnej przy różnych wydajnościach eksploatacyjnych wody termalnej i temperatury jej zatłaczania [opracowanie własne]

2. Mechanizmy procesów kolmatacji

Procesy kolmatacji objawiające się stopniowym spadkiem chłonności otworu zatłaczającego są podstawowym problemem niemalże wszystkich ciepłowni geotermalnych na niżu Polskim [9]. Zatłaczanie schłodzonych wód termalnych do węglanowych skał zbiornikowych typu szczelinowego jest znacznie mniej kłopotliwe niż zatłaczanie do skał piaskowcowych typu porowego, które są ujmowane w ciepłowni geotermalnej w Pyrzycach.

Na procesy kolmatacji otworów chłonnych składa się kilka czynników. Jednym z nich jest korozja rur okładzinowych [1, 3, 4, 10]. W wyniku postępującej kolmatacji następuje zwężanie się średnicy wewnętrznej otworu chłonnego oraz zatykanie się stref czynnych filtrów, obsypek i skał zbiornikowych [2], doprowadzając do zatykania porów i szczelin poziomu wodonośnego cząstkami stałymi, a tym samym do zmniejszenia chłonności strefy przyodwiertowej i skał zbiornikowych. Promień strefy przyodwiertowej o przepuszczalności skał zmniejszonej w skutek postępującej kolmatacji może wynieść od kilku centymetrów do kilku metrów [10].

Podstawowym mechanizmem kolmatacji strefy przyodwiertowej (złożowej) jest wytrącanie się substancji osadotwórczych z wody termalnej. Podczas schładzania wody termalnej na wymiennikach ciepła zachodzą reakcje fizykochemiczne, w wyniku, których następuje wtórne wytrącanie się związków chemicznych m in. trudno rozpuszczalnych soli, a także w przypadku dostępu tlenu do rurociągów termalnych, również tlenków i hydroksytlenkow żelaza i manganu. W przypadku naruszenia równowagi oksydacyjno-redukcyjnej roztworu następuje wytrącanie siarczków żelaza, manganu i miedzi. Szczególnie intensywnie występuje kolmatacja w przypadku zatłaczania do złoża, schłodzonej w procesie odbioru ciepła, wody termalnej. To procesy fizykochemiczne zachodzące w wodzie termalnej, wykorzystanej w procesach energetycznych, są odpowiedzialne za pogarszającą się pracę całego obiegu geotermalnego. Odwrócenie kierunku zachodzenia tych procesów wymaga ingerencji w przemiany równowagowe, zachodzące na granicy faz: ciekłej (woda termalna), gazowej (uwalniane w wyniku rozprężania z fazy ciekłej gazy) i stałej

(wytrącające się osady). Niezbędne w tym celu jest poznanie skomplikowanego składu fizykochemicznego wody termalnej, a także równowag kwasowo-zasadowych związanych z krystalizacją węglanu wapnia, siarczanu baru, siarczanu strontu oraz równowag w procesach utleniająco-redukcyjnych, odpowiedzialnych za wytrącanie się tlenków i siarczków żelaza, miedzi i manganu, w filtrze oraz w strefie przyodwiertowej.

By zmniejszyć intensywność kolmatacji przed zatłoczeniem schłodzona woda termalna przepompowywana jest przez filtry świecowe gdzie następuje osadzanie się części produktów kolmatacji. Jak wynika z przeprowadzonych analiz dominującym składnikiem w osadzie z filtra jest żelazo. Występuje ono w kilku formach: bądź uwodnionego brunatnego tlenku żelaza (III) Fe_2O_3 – hematytu, bądź szaro-czarnego FeOOH – getytu. Mniejszy udział w osadzie ma mangan, występujący w formie uwodnionego tlenku MnO, który na powietrzu utlenia się do czarno-brunatnego uwodnionego tlenku MnO_2. W osadzie obecne są również siarczany. Stosunkowo niewiele jest natomiast osadów wapnia, którego udział w formie CaO wynosi zaledwie 1,2%.

W oparciu o wyniki analizy fizykochemicznej wody termalnej wyznaczono indeksy nasycenia trudno rozpuszczalnych związków nieorganicznych, dla których występuje potencjalne ryzyko wytrącenia z wody termalnej. Indeks nasycenia dla danego związku, to wskaźnik określający potencjał do wytrącania tego związku z badanego roztworu wodnego. Wyrażany jest, jako stosunek iloczynu aktywności kationu i anionu, składających się na dany związek chemiczny, do iloczynu rozpuszczalności tego związku. Wartości powyżej 1,0 wskazują na tendencję do wytrącania osadu, natomiast poniżej 1,0 informują, o stanie niewysycenia roztworu danym związkiem.

Obliczenia indeksów wykonano z użyciem specjalistycznych programów – AquaChemR oraz PHREEQC Interactive. Obliczenia odniesiono do parametrów ciśnienia i temperatury, odpowiadających tym panującym w ciepłowni geotermalnej w Pyrzycach. Interpretując indeksy nasycenia, należy zawsze pamiętać, że informują one jak silna jest tendencja do strącania osadów danych związków, nie określają natomiast, jaka ilość osadu strąci się z danej objętości roztworu. Rozważając wartości indeksów nasycenia należy brać również pod uwagę całkowite stężenie jonów składających się na dany związek. W omawianym przypadku, pomimo wysokiej wartości indeksu nasycenia, gibbsyt nie będzie stanowił istotnego zagrożenia, ze względu na bardzo niskie stężenie glinu w wodzie termalnej (0,012 mg/dm^3). Nawet zakładając strącenie całego glinu w postaci gibbsytu (co nie jest prawdą), maksymalna ilość wytworzonego osadu wyniesie 0,035 mg z 1 litra wody termalnej, co daje 0,035 g z 1 m^3. Kalkulując średnią wydajność przepływu wody na poziomie 100 m^3/h, osadu gibbsytu mogłoby powstać jedynie około 84 g/dobę, co jest wartością kompletnie nieznaczącą w porównaniu z rzeczywiście strącającymi się ilościami osadów. Nieznaczny wkład do masy osadów powinien wnieść także dolomit, a to ze względu na niski odczyn pH solanki. Pomimo dużych ilości wapnia i magnezu, przy odczynie pH na poziomie 6,0 ilość anionów węglanowych CO_3^{2-} będzie zbyt mała by mogły powstać duże ilości osadu dolomitu. W przypadku tlenków żelaza oraz siarczanu strontu występuje wystarczająco wysoki indeks nasycenia, a także stężenia jonów, sprzyjające wytrącaniu relatywnie dużych ilości osadów. W określonych warunkach ciśnienia

i temperatury należy więc oczekiwać wytrącania się istotnych ilości osadów trzech związków: celestynu, getytu i hematytu.

3. Wpływ procesów kolmatacji na ciśnienia zatłaczania i wydajność zatłaczania wody termalnej

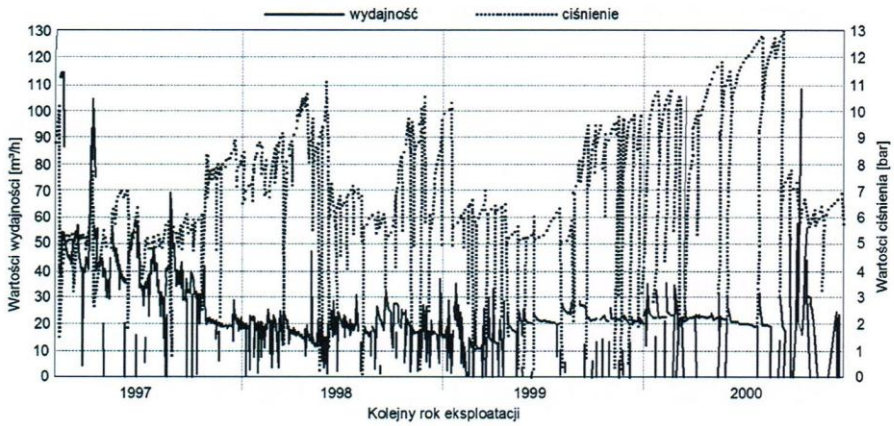
Na przestrzeni lat problemy związane z kolmatacją stanowiły istotny problem w funkcjonowaniu Geotermii Pyrzyce. Zmniejszanie wydajności zatłaczania wody termalnej połączone ze zwiększaniem się ciśnień zatłaczania na poszczególnych otworach doprowadzało do wyłączenia członu geotermalnego w produkcji ciepła lub zmniejszenie jego udziału [6, 7]. By utrzymać możliwość zatłaczania przeprowadzano zabiegi czyszczenia otworów chłonnych z użyciem urządzenia wiertniczego oraz metodą kwasowania.

Początki działania dubletu geotermalnego w Geotermii Pyrzyce to przełom roku 1995/1996. Pierwsze próby obiegu wody w skali przemysłowej podjęto w grudniu 1995 r. Przy wydajności 150 m^3/h ciśnienie w otworze zatłaczającym Pyrzyce GT-2 wynosiło 5,3 bara. Zatłaczanie przerwano ze względu na prace wykończeniowe w hali ciepłowni. Ponowne uruchomienie nastąpiło miesiąc później podczas oficjalnego startu całej ciepłowni geotermalnej. Uzyskano wtedy wydajność zatłaczania 148,6 m^3/h przy ciśnieniu ok. 10 barów. Podczas pierwszych miesięcy pracy ciepłowni wystąpiły 42 przestoje. Po każdym przestoju najczęściej wzrastało ciśnienie zatłaczania. W związku z tym już w październiku 1996 r. wykonano zabiegi czyszczenia mechanicznego i chemicznego otworu Pyrzyce GT-2. Po chwilowym uzyskaniu sprawności zbliżonej do początkowej (135 m^3/h przy ciśnieniu 5,7 bara) wydajność zatłaczania ponownie zaczęła maleć, a ciśnienie rosnąć.

Następnym zabiegiem mającym intensyfikować zatłaczanie do otworu Pyrzyce GT-2 było azotowanie wykonane w listopadzie 1998 i marcu 1999 r. Do czasu tego zabiegu wydajność zatłaczania spadła zaledwie do 10 m^3/h przy ciśnieniu ponad 10 barów. Reakcja otworu chłonnego Pyrzyce GT-2 na azotowanie była pozytywna, ale krótkotrwała – wydajność wzrastała do 23 m^3/h, a ciśnienie spadło do 6 barów.

Do roku 2000 pracę otworu Pyrzyce GT-2 można podzielić na 3 etapy (Rysunek 3):

- Rok 1997 – najczęściej spotykane wartości wydajności zatłaczania wynosiły 35-55 m^3/h przy ciśnieniu 4-6 barów,
- Lata 1998-1999 – obserwowany spadek wydajności zatłaczania do 15-30 m^3/h przy wzroście ciśnienia do około 10 barów,
- Rok 2000 – kolejny spadek wydajności zatłaczania do około 20 m^3/h przy ciśnieniu około 12 barów.



Rysunek 3. Wydajności i ciśnienia w otworze chłonnym Pyrzyce GT-2 w pierwszych czterech latach eksploatacji [opracowanie własne]

Również na przełomie roku 1995/1996 w otworze chłonnym Pyrzyce GT-4 przeprowadzono pompowania eksploatacyjno-zatłaczające. Uzyskano wydajność 143 m^3/h przy stabilizacji ciśnienia na poziomie 3,7 bara. Po krótkim okresie eksploatacji zaobserwowano spadek wydajności i znaczny wzrost ciśnienia zatłaczania. W związku z tym już w październiku 1996 roku zostały wykonane zabiegi czyszczenia mechanicznego oraz chemicznego, które pozwoliły uzyskać wydajność zatłaczania 170 m^3/h przy podciśnieniu. Kolejne lata pracy otworu Pyrzyce GT-4 odznaczały się stabilnymi parametrami. Było to około 100 m^3/h przy ciśnieniu 6 barów. Znaczne pogorszenie stanu technicznego otworu zaobserwowano w 2004 roku. W 2005 roku na otworze Pyrzyce GT-4 wykonano profilowanie średnicy filtra oraz stanu rur. Stwierdzono ubytki ścianek rur okładzinowych, a czasami ich zupełny brak. Uszkodzony był but rur oraz górna część filtra. Wymieniono filtr oraz wykonano chemiczne czyszczenie otworu. Nie uzyskano jednak wydajności początkowej. Każdy przestój powodował natychmiastowy wzrost ciśnienia zatłaczania oraz spadek chłonności.

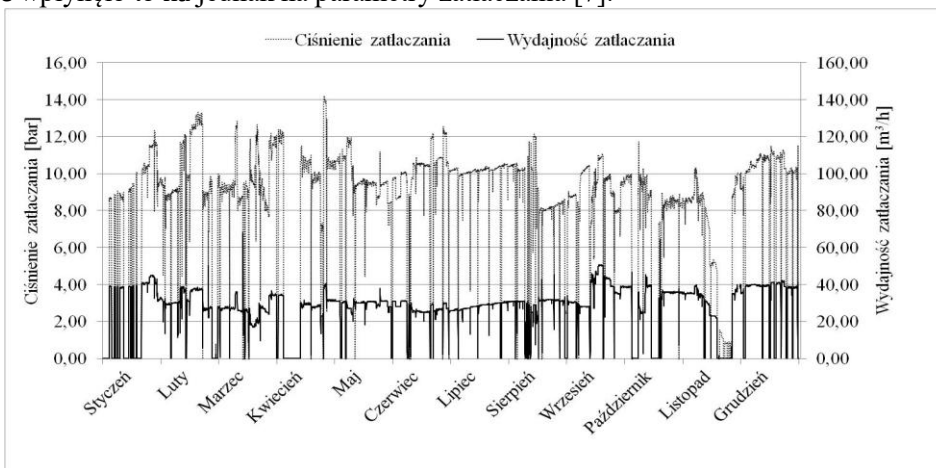
W związku z pogarszającymi się warunkami pracy otworu Pyrzyce GT-2 podjęto kolejne próby mające na celu zwiększenie intensywności zatłaczania. Na podstawie badań geofizycznych w otworze Pyrzyce GT-2 stwierdzono narosty osadu na filtrze w interwale 1428-1435 m p.p.t., które spowodowały zmniejszenie średnicy filtra ze 158 mm do 143 mm. Prace rekonstrukcyjne polegały na mechanicznym i chemicznym czyszczeniu filtra i przyniosły efekt w postaci wzrostu wydajności zatłaczania do 100 m^3/h przy zerowym ciśnieniu zatłaczania (podciśnienie). Podobnie jak w przypadku wcześniejszych zabiegów i te przynosiły efekt krótkotrwały – wydajności malały przy rosnących ciśnieniach zatłaczania.

Analogicznie jak w otworze Pyrzyce GT-4 należy zauważyć, że w przypadku otworu Pyrzyce GT-2 nagły wzrost ciśnienia zatłaczania pojawiał się po przestojach w działaniu obiegu geotermalnego.

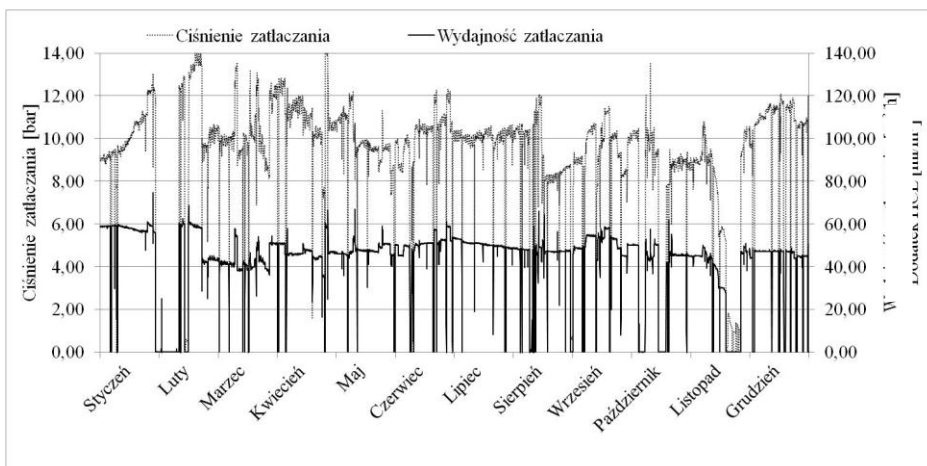
W 2005 roku wykonano kolejne badania geofizyczne w otworze Pyrzyce GT-2 oraz czyszczenie mechaniczne i chemiczne części roboczej filtra przy pomocy

urządzenia wiertniczego. Pozwoliło to na uzyskanie chłonności 170 m^3/h przy ciśnieniu zatłaczania 1,8 bara. Uzyskane parametry były lepsze niż te uzyskane podczas prac dokumentacyjnych w trakcie pompowań pomiarowych po wykonaniu otworu Pyrzyce GT-2. Jednak dłuższa eksploatacja powodowała ciągle spadek wydajności i wzrost ciśnienia zatłaczania.

W 2008 roku w otworze Pyrzyce GT-2 założono rury HDPE odporne na korozję. Nie wpłynęło to na jednak na parametry zatłaczania [7].



Rysunek 4. Kolejne miesiące eksploatacji otworu zatłaczającego Pyrzyce GT-2 w 2010 r. [opracowanie własne]



Rysunek 5. Kolejne miesiące eksploatacji otworu zatłaczającego Pyrzyce GT-4 w 2010 r. [opracowanie własne]

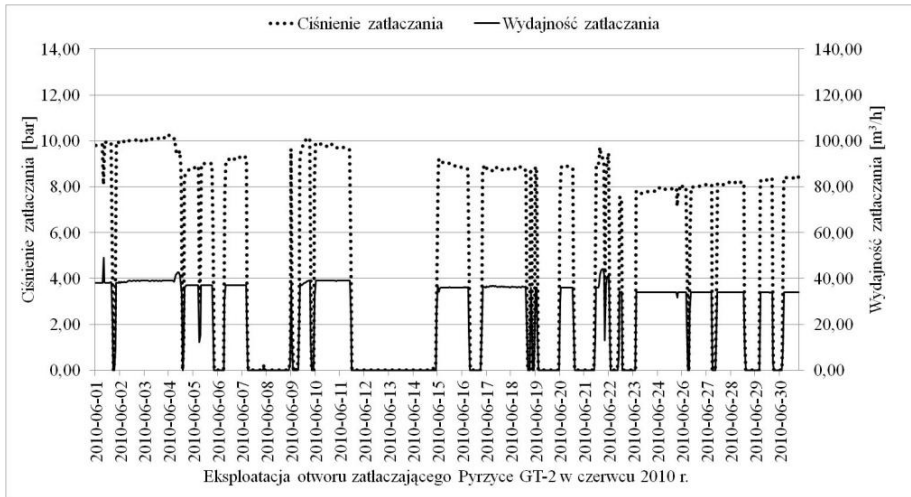
Procesy kolmatacji w otworach chłonnych nie zostały zahamowane. By umożliwić eksploatację otworu Pyrzyce GT-1 z wydajnościami około 130 m^3/h i utrzymać opłacalność produkcji ciepła z członu geotermalnego zdecydowano się w 2010 roku na włączenie do obiegu geotermalnego również otworu Pyrzyce GT-3, jako otworu

chłonnego. W 2010 roku wydajność zatłaczania do otworu Pyrzyce GT-2 wynosiła od 20 do 40 m^3/h przy ciśnieniach zatłaczania 8-12 barów (Rysunek 4), natomiast do otworu Pyrzyce GT-4 w ilości 40-50 m^3/h przy ciśnieniu ponad 9 barów (Rysunek 5). Pozostała woda w ilościach 40-60 m^3/h musiała zostać zatłaczana do otworu Pyrzyce GT-3.

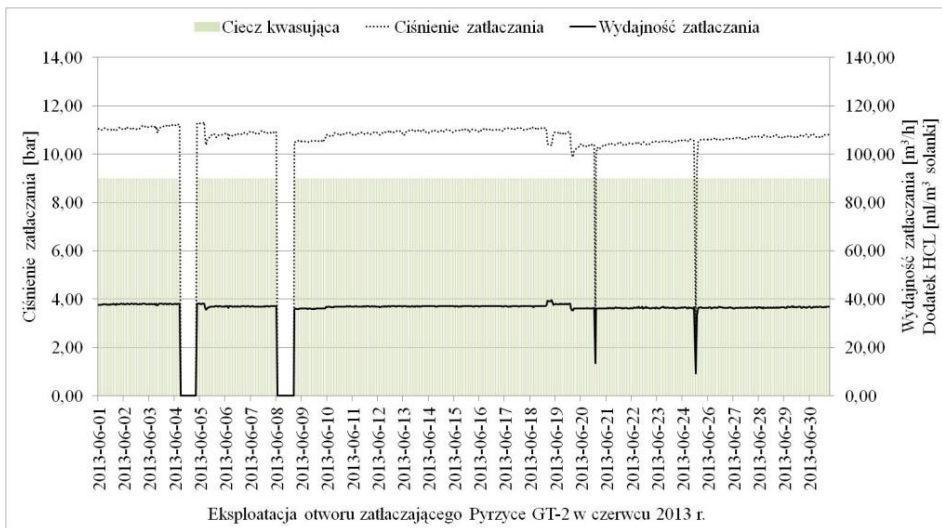
Zatłaczanie wód schłodzonych do otworu Pyrzyce GT-3 wpłynęło negatywnie na temperaturę wód ujmowanych pobliskim otworem Pyrzyce GT-1, która na przestrzeni lat spadła z około 61°C do około 58°C. Mimo, że od czerwca 2016 roku nie są zatłaczane wody do otworu Pyrzyce GT-3, temperatura eksploatowanej wody termalnej utrzymuje się na poziomie 58°C. Rok eksploatacji nie pozwolił na odbudowę pierwotnej temperatury wody termalnej uzyskiwanej otworem Pyrzyce GT-1. Problemy z kolmatacją w sposób pośredni wpłynęły na temperaturę ujmowanej wody termalnej. Włączenie otworu Pyrzyce GT-3 zostało wymuszone przez procesy kolmatacji otworów zaprojektowanych, jako chłonne. Zmniejszenie się temperatury wody termalnej o 3 stopnie spowodowało obniżenie produkowanej mocy cieplnej z członu geotermalnego aż o 10%. Podejrzewa się również, że warstwa wodonośna w otworze Pyrzyce GT-3 uległa wychłodzeniu, co powoduje, że nie będzie on mógł zostać wykorzystywany jako otwór eksploatacyjny.

Kolejnymi metodami mającymi na celu poprawę wydajności zatłaczania oraz zmniejszenie ciśnień były metody tzw. miękkiego kwasowania oraz super miękkiego kwasowania [8]. Różnica między metodą miękkiego kwasowania, a zwykłego kwasowania polega na tym, że kwas solny w metodzie miękkiego kwasowania jest zatłaczany w tej samej ilości lecz w dłuższym czasie. W super miękkim kwasowaniu kwas solny jest dozowany w mniejszej ilości, lecz przez cały czas działania obiegu geotermalnego. Miękkie kwasowanie było wprowadzane w różnych okresach do lutego 2013 roku, natomiast metoda super miękkiego kwasowania od tego czasu. Nowe metody zapobiegania powstawaniu kolmatacji pozwoliły na utrzymanie chłonności otworów zatłaczających na stałym poziomie. Przed ich wprowadzeniem instalacja pracowała z częstymi i dość długimi przestojami spowodowanymi trudnościami w zatłaczaniu.

W 2010 roku wydajność zatłaczania do otworu Pyrzyce GT-2 wynosiła około 20-40 m^3/h przy ciśnieniach 8-12 barów. W tym samym roku jednak kilkunastokrotnie dochodziło do przestojów w pracy ciepłowni np. w czerwcu 2010 roku (Rysunek 6). Metoda super miękkiego kwasowania, podczas której dozowano ciecz kondycjonującą w postaci kwasu solnego, pozwoliła w 2013 roku na osiągnięcie stabilizacji wydajności zatłaczania w otworze Pyrzyce GT-2 na poziomie ok. 40 m^3/h przy ciśnieniu około 8-9 barów (Rysunek 7). Pierwotnie, w trakcie testowania metody super miękkiego kwasowania, stosowany był kwas solny, do którego następnie dołączono odtleniacz i dyspergator. W 2013 r. dzięki tej metodzie wyeliminowano również liczne i długotrwałe wyłączenia instalacji geotermalnej spowodowane nagłymi wzrostami ciśnienia zatłaczania [8]. W wyniku ww. metody zwiększyła się ilość ciepła z członu geotermalnego, szczególnie w okresie sezonu grzewczego, co przełożyło się na zmniejszenie zużycia gazu ziemnego i na redukcję emisji CO_2 przez Geotermię Pyrzyce.



Rysunek 6. Wykres pracy otworu Pyrzyce GT-2 w czerwcu 2010 r. [opracowanie własne]



Rysunek 7. Wykres pracy otworu Pyrzyce GT-2 w czerwcu 2013 r. [opracowanie własne]

4. Badania nad technologią zapobiegania kolmatacji chłonnych otworów geotermalnych

W 2015 roku, w ramach programu GEKON prowadzonego przez HPC POLGEOL S.A (dawniej Przedsiębiorstwo Geologiczne POLGEOL S.A) pod nazwą „*Nowa technologia zapobiegania kolmatacji chłonnych otworów geotermlanych"*, rozpoczęto prace badawcze nad technologią zapobiegania kolmatacji chłonnych otworów geotermalnych. Zakładała ona udostępnienie nowych horyzontów do zatłaczania wody termalnej w istniejących otworach chłonnych oraz kondycjonowanie wody termalnej. Opracowywana technologia miała na celu przywrócenie początkowych parametrów

zatłaczania (ok. 120-130 m^3/h przy możliwie jak najniższych ciśnieniach) oraz zapewnić trwałość uzyskanych parametrów.

Badania związane z udostępnieniem nowej warstwy wodonośnej polegały na wykonaniu perforacji rur okładzinowych oraz jej oczyszczeniu, które zostało wykonane w ramach opracowanej technologii.

Wcześniejsze próby (2015 rok) przeprowadzone w otworze Pyrzyce GT-4 zakończyły się niepowodzeniem. Prace badawcze nad opracowaniem nowej technologii czyszczenia chłonnych otworów geotermalnych w 2016 roku prowadzone były na otworze Pyrzyce GT-2. Badania te zakończyły się sukcesem.

Otwór Pyrzyce GT-2 jest zafiltrowany filtrem typu Johnsona o konstrukcji prętowo-rurowej ze szczeliną 0,5 mm, o powierzchni czynnej 15% w interwale 1375,56-1523,00 m p.p.t. Wykonana jest obsypka żwirowa (ziarna 0,8-1,4 mm) filtra. W związku z powyższym zaplanowano zapuszczenie przewodu wiertniczego do około 1375 m p.p.t. do strefy filtra. Użyto urządzenia wiertniczego Wirth. Jest to urządzenie samojezdne na samochodzie Mercedes do wiercenia z lewym obiegiem płuczki. Z racji, że otwór Pyrzyce GT-2 był wykorzystywany do zatłaczania wody termalnej przed początkiem prac rozmontowano wiatę ochronną, rurociągi oraz głowicę otworu. Zejście przewodu wiertniczego zakończyło się sukcesem. Udało się zejść przewodem do 1376 m p.p.t. Nie doszło do uszkodzeń rur osłonowych i stwierdzono drożność otworu, stąd przystąpiono do następnego etapu prac badawczych.

Szablonowanie otworu, wykonane w trakcie badań geofizycznych, było konieczne do określenia miejsc perforacji rur okładzinowych. Perforacja musiała być wykonana w miejscach gdzie występują piaskowcowe warstwy litologiczne, o podobnej charakterystyce jak te ujęte filtrem. Docelowo poszukiwane były piaskowce drobnoziarniste i średnioziarniste tzw. warstwy komorowskie jury dolnej zalegające nad zafiltrowanymi warstwami mechowskimi.

Badania geofizyczne zostały wykonane przez Geofizykę Toruń S.A. Polegały na zapuszczeniu na linie karotażowej próbnika promieniowania gamma wraz z mufolokatorem (Gamma Ray, mufolokator – GR-CCL), w rurach Ø$9^5/_8$" do głębokości około 1376 m p.p.t. Podstawą do korelacji głębokościowej i wytypowania interwałów do perforacji był pomiar GR-CCL precyzyjnie porównany z archiwalnymi badaniami geofizycznymi oraz profilem litologicznym.

Do perforacji wytypowano warstwy piaskowcowe w interwałach:

- 1284-1294 m p.p.t,
- 1308-1318 m p.p.t,
- 1351-1371 m p.p.t.

Perforacja rur okładzinowych miała za zadanie udostępnić warstwę wodonośną, wytypowaną w badaniach geofizycznych, nad strefą już zafiltrowaną. By to zrealizować konieczne było przebicie za pomocą ładunku wybuchowego ścianki przewodu, na który składała się rura HDPE – 10 mm, rura okładzinowa – 10 mm oraz warstwa cementu.

Perforacja rur okładzinowych została wykonana perforatorem rurowym o średnicy 114 mm firmy DYNAenergetics zapuszczonym na kablu karotażowym inicjowanym zapalnikiem elektrycznym Z480 i 0015 FDE. Ładunkami wybuchowymi były ładunki kumulacyjne DYNAenergetics 39g HMXDP2. Gęstość perforacji wynosiła 16,5

przestrzeliny na 1 metr, fazowanie 60°. Zasięg perforacji to 65,4", a średnica przestrzeliny to 0,47''.

Korelacja głębokości odpalenia perforatora rurowego zapuszczonego na kablu karotażowym wykonywana była na podstawie pomiaru mufolokatorem, który był zapuszczany wraz z zestawem perforacyjnym. Pomiar wykonany mufolokatorem (CCL) podczas zapuszczania zestawu perforacyjnego korelowany był do wcześniej wykonywanego (na etapie szablonowania otworu) pomiaru Gamma Ray – mufolokator (GR-CLL). Głębokość odpalenia perforatora dokumentowana była na pomiarze korelacyjnym (mufolokatora), na którym zaznaczono pozycje perforatora.

Odpłukanie strefy perforowanej miało na celu oczyszczenie jej z pozostałości rur okładzinowych, cementu oraz skał. Wykonano je za pomocą rurek syfonowych Ø$3^1/_2$" zapuszczonych do głębokości wykonania perforacji. Prowadzono strefowe płukania poprzez stopniowe dokładania kolejnych syfonówek. Na każdym poszczególnym etapie płukania (dla danej strefy perforowanej np. interwał 1284-1294 m p.p.t) rury były podciągane i opuszczane z jednoczesnym odsysaniem płynu złożowego aż do uzyskania klarownej cieczy. Po odpłukaniu danej strefy perforowania dokładano kolejne syfonówki, tym samym schodząc do kolejnej warstwy. Płukanie rozpoczęto od strefy w interwale 1284-1294 m p.p.t. Płukanie wykonano za pomocą air liftu.

Celem zabiegu kwasowania było wykonanie wanny kwasowej w celu oczyszczenia strefy sperforowanej oraz zafiltrowanej.

Kwasowanie wykonano za pomocą rurek syfonowych Ø$3^1/_2$". W I etapie, którym było kwasowanie strefy zafiltrowanej (poniżej 1375 m p.p.t.) zapuszczono zestaw do głębokości 1375 m p.p.t. i zatłoczono 6 m^3 cieczy kwasującej, w której skład wchodził 15% roztwór kwasu solnego oraz inhibitor korozji. Następnie wtłoczono 7,3 m^3 solanki na przybitkę. Po godzinie usunięto ciesz poreakcyjną przy pomocy airliftu z wydajnością 22 m^3/h. Czas pompowania wyniósł 2 godziny. Ciecz poreakcyjną odprowadzano do zbiornika i neutralizowano węglanem wapnia. Wykonywano monitorowanie pH pehametrem i po stwierdzeniu osiągnięcia odczynu obojętnego ciecz przerzucono na dół zrzutowy Geotermii Pyrzyce.

W etapie II podciągnięto zestaw do głębokości 1362 m p.p.t. Był to etap kwasowania strefy perforowanej. Wtłoczono 9 m^3 cieczy kwasującej (15% kwas solny + inhibitor korozji). Następnie wtłaczano przybitkę 1,3 m^3, podciągano zestaw do 1312 m p.p.t., wtłaczano przybitkę 3 m^3, podciągano zestaw do 1288 m p.p.t. i wtłaczano ostatnią część przybitki – 3 m^3. Po godzinie od kwasowania rozpoczęto ponownie usuwanie cieczy poreakcyjnej przy pomocy airliftu z wydajnością około 22 m^3/h. Wykonano to strefowo w głębokościach 1288, 1296, 1312, 1320, 1360, 1368 oraz 1376 m p.p.t. Ciecz poreakcyjną odprowadzano do zbiornika, neutralizowano i przerzucano na dół zrzutowy.

Do zabiegów kwasowania wykorzystano agregat CA-320 oraz sprężarkę Atlas Copco zapewniają ciśnienie robocze 2 MPa i maksymalną wydajność 25 m^3/min.

Przeprowadzone badania były częścią planu, który zakładał utrzymanie dobrych parametrów zatłaczania w otworze Pyrzyce GT-2. Kolejnym krokiem było wprowadzenie nowej metody kondycjonowania wody termalnej za pomocą środków chemicznych.

W trakcie testów laboratoryjnych wytypowano substancje, które mogłyby kondycjonować wodę termalną i pozwolić na uzyskanie stabilnych wysokich parametrów pracy dubletu geotermalnego otworów Pyrzyce GT-1 oraz Pyrzyce GT-2.

Testy polegały na dozowaniu wytypowanych środków chemicznych do pobranej z otworu Pyrzyce GT-1 wody termalnej oraz obserwowanie stopnia zahamowania procesów wytrącania się związków osadotwórczych. Dla jakościowej analizy wykonywano testy mętnościomierzem. Pozwoliło to na wytypowanie następujących środków chemicznych takich jak: kwas mrówkowy, chlorowodorek mocznika, kwas cytrynowy i kwas askorbinowy.

W następstwie kolejnych testów na wodzie termalnej, w której zaszły już procesy wytrącania osadów, stwierdzono, że wymienione wyżej substancje mają również potencjał do rozpuszczania związków osadotwórczych.

Po wykonaniu pierwszego etapu badań, który polegał na udostępnieniu nowego horyzontu wodonośnego w trakcie zatłaczania wody termalnej dozowano środki wytypowane w badaniach laboratoryjnych.

5. Analiza wyników

Podczas przeprowadzonych badań nad nową technologią zapobiegania kolmatacji chłonnych otworów geotermalnych, w czerwcu 2016 roku, udostępniono nową warstwę wodonośną do zatłaczania wód termalnych w otworze Pyrzyce GT-2 oraz wykonano kwasowanie i płukanie strefy zafiltrowanej.

Prace badawcze związane z opracowaniem nowej technologii wspomagania eksploatacji wody termalnej w ciepłowni geotermalnej zakończyły się sukcesem.

Perforowano rury okładzinowe w interwałach 1284-1294 m p.p.t, 1308-1318 m p.p.t, 1351-1371 m p.p.t. udostępniając piaskowce warstw komorowskich do zatłaczania wody termalnej.

Przeprowadzone prace na żadnym z etapów nie napotkały problemów technicznych. Otwór był drożny, nie było problemów z zejściem przewodu wiertniczego oraz próbników geofizycznych. Perforacje za pomocą metod strzałowych wykonano zgodnie z planem.

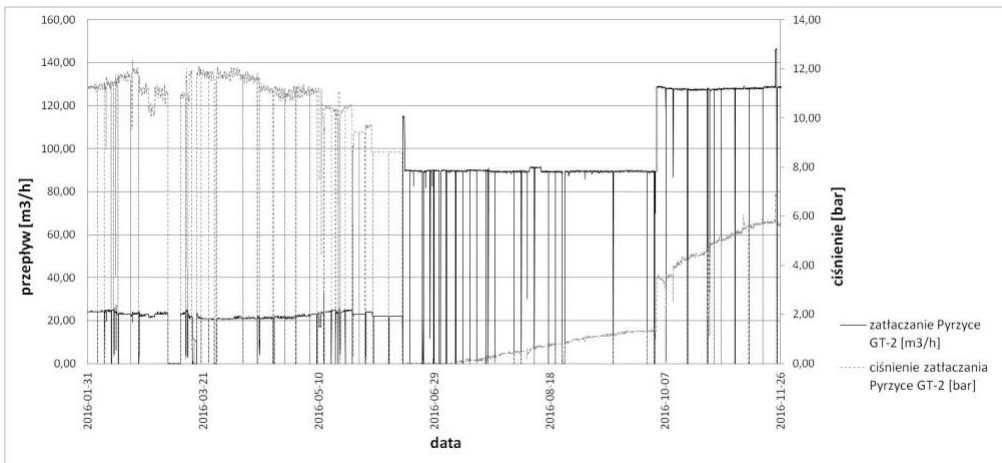
Parametry uzyskane w wyniku przeprowadzonych badań były lepsze od tych, które zakładano. Otwór Pyrzyce GT-2 pracował na bardzo słabych parametrach. Wydajność zatłaczania, do czasu wykonania badań (czerwiec 2016 r.), wynosiła około 20 m^3/h przy ciśnieniach zatłaczania dochodzącymi do 12 barów (Rysunek 8). W tym czasie otwór Pyrzyce GT-4 pracował z minimalną wydajnością (2-5 m^3/h) co powodowało, że pozostała woda termalna musiała trafiać do otworu Pyrzyce GT-3.

W czasie pierwszych 3 miesięcy po wykonaniu perforacji rur okładzinowych i czyszczeniu strefy zafiltrowanej zatłaczano do otworu Pyrzyce GT-2 wodę termalną z wydajnością około 90 m^3/h. Następnie od października z wydajnością około 130 m^3/h (Rysunek 8). Ciśnienie zatłaczania w pierwszych tygodniach było zerowe (podciśnienie). Zaczęło rosnąć do 2 barów, a następnie po zwiększeniu wydajności zatłaczania do maksymalnie 7 barów (Rysunek 8).

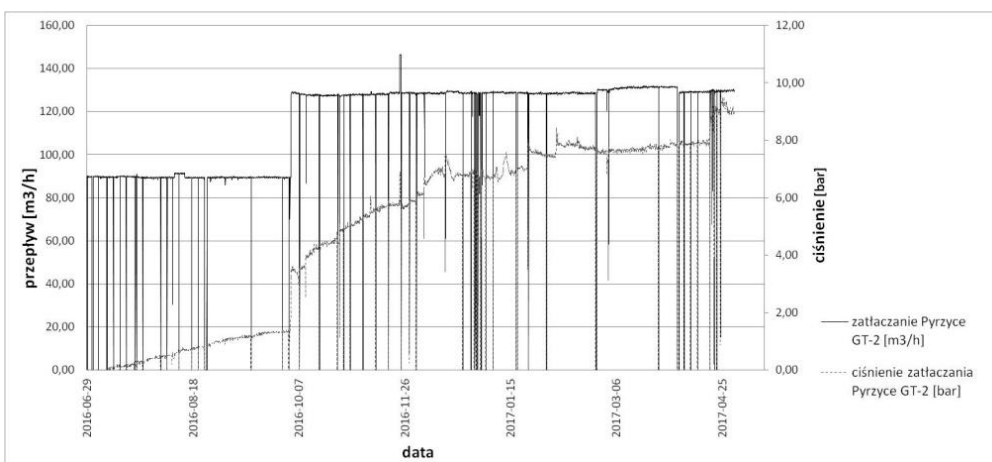
Dozowanie środków do kondycjonowania wody termalnej pozwoliło na utrzymanie wysokiej sprawności zatłaczania otworu Pyrzyce GT-2. W 2017 roku wydajności zatłaczania są na poziomie 130 m^3/h przy ciśnieniu zatłaczania dochodzącym do 10 barów (Rysunek 9). Nie występują skoki ciśnienia, co zapewnia ciągłą pracę członu

geotermalnego. W trakcie ostatnich miesięcy wiosenno-letnich energia w Geotermii Pyrzyce była pozyskiwana wyłącznie z członu geotermalnego.

W związku z wykonanymi pracami od czerwca 2016 roku praca Ciepłowni Pyrzyce opiera się wyłącznie na otworze eksploatacyjnym Pyrzyce GT-1 oraz chłonnym Pyrzyce GT-2.



Rysunek 8. Wydajność zatłaczania i ciśnienie zatłaczania na otworze chłonnym Pyrzyce GT-2 od stycznia 2016 r. do grudnia 2016 r. [opracowanie własne]



Rysunek 9. Wydajność zatłaczania i ciśnienie zatłaczania na otworze chłonnym Pyrzyce GT-2 od czerwca 2016 r. do maja 2017 r. [opracowanie własne]

6. Wnioski

Na podstawie wyników badań nad nową technologią wspomagającą eksploatację wód termalnych w ciepłowniach geotermalnych należy stwierdzić, że spełnia ona założone cele.

Otwór Pyrzyce GT-2 obecnie pracuje z wydajnością zatłaczania do 130 m^3/h przy ciśnieniach dochodzących do 10 barów, a kondycjonowanie zatłaczanej wody

termalnej zapewnia trwałość uzyskanych parametrów 15 miesięcy po przeprowadzeniu badań.

Uzyskane w wyniku badań nad nową technologią wydajności zatłaczania są ponad 6-krotnie wyższe niż przed jej zastosowaniem – wzrost z około 20 m^3/h do 130 m^3/h.

Technologia polegająca na udostępnieniu nowego horyzontu do zatłaczania w istniejących otworach chłonnych, połączona z kondycjonowaniem zatłaczanej do nich wody termalnej, znacząco usprawnia i wspomaga pracę ciepłowni geotermalnej.

Prezentowane w pracy wyniki badań uzyskano w ramach realizacji projektu pn. „Nowa technologia zapobiegania kolmatacji chłonnych otworów geotermalnych" dofinansowanego ze środków Narodowego Centrum Badań i Rozwoju oraz Narodowego Funduszu Ochrony Środowiska i Gospodarki Wodnej w ramach programu GEKON – Generator Koncepcji Ekologicznych.

Umowa nr GEKON1/O4/214087/42/2015

Literatura

1. Banaś J., Mazurkiewicz B., Solarski W., *Korozja metali w wodach geotermalnych*, Technika Poszukiwań Geologicznych. Geotermia, Zrównoważony Rozwój, Nr 2 (2007), s. 5-12.
2. Barkman J.H., Davidson D.H., *Measuring water quality and predicting well important, Journal of Petroleum Technology*, July 1972.
3. Biernat H., Kulik S., Noga B., Kosma Z., 2010b. *Problemy korozji przy zatłaczaniu wykorzystanych wód termalnych*. Modelowanie Inżynierskie, Tom 8, Nr 39, s. 13-18.
4. Biernat H., Kulik S., Noga B., Kosma Z. 2010c. *Problemy inkrustacji przy zatłaczaniu wykorzystanych wód termalnych*. Modelowanie Inżynierskie, Tom 8, Nr 39, s. 7-12.
5. Bujakowska K., Biernat H., Bentkowski A., Kapuściński J., 1995. *Dokumentacja hydrogeologiczna zasobów eksploatacyjnych ujęcia wód termalnych w utworach jury dolnej dla potrzeb m. Pyrzyce województwo Szczecińskie*. Arch. PG POLGEOL, Warszawa.
6. Noga B., Biernat H., *Sprawozdanie końcowe z prac związanych z realizacją kontroli w zakresie poprawności i efektywności dozowania czystego kwasu solnego, odtleniacza i dyspergatora do otworu geotermalnego i analizę ich wpływu na poprawę chłonności warstwy złożowej w otworach chłonnych poprzez pomiary pH i potencjału redox wody termalnej*, Arch. PG POLGEOL, Warszawa, 2014.
7. Noga B., Biernat H., Martyka P., Saletowicz G., *Program prac dla poprawy chłonności i zapobiegania kolmatacji warstwy złożowej w otworach chłonnych poprzez wykonanie zabiegów intensyfikujących z dozowaniem preparatów kondycjonujących w Ciepłowni Geotermalnej Pyrzyce*. Praca niepublikowana, archiwum HPC Polgeol Spółka Akcyjna, 2012.
8. Noga B., Biernat H., Zwierzyński M., Martyka P., *Ocena wyników z demonstracji nowatorskiej technologii poprawy chłonności warstwy złożowej wód geotermalnych*. Praca niepublikowana, archiwum HPC Polgeol Spółka Akcyjna, 2014.
9. Parecki A., Biernat H., *Próba rozwiązania problemów towarzyszących eksploatacji ciepłowni geotermalnych wykorzystujących wody termalne z kolektorów porowych*. Technika Poszukiwań Geologicznych. Geotermia, Zrównoważony Rozwój, Nr 2, 2007, s. 107-110.
10. Wirght C.C., Chilingarian G.V., *Water quality for subsurface injection* [w:] Surface operations for petroleum, II. Eds. Chilingarian, Robertson, Kummar: Developments in Petroleum Science. Elsevier, 1989.

Nowa technologia wspomagająca eksploatację wód termalnych w ciepłowniach geotermalnych

Abstrakt
Opracowanie ma za zadanie przedstawiać zasadę działania geotermii na przykładzie Geotermii Pyrzyce oraz pokazać problemy związane z kolmatacją chłonnych otworów geotermalnych. Opisana technologia zapobiegania kolmatacji w otworze chłonnym polega na udostępnieniu nowego horyzontu do zatłaczania wody, czyszczeniu chemicznym i mechanicznym filtra oraz na kondycjonowaniu wody termalnej na dalszym etapie. Z przeprowadzonych badań na otworze chłonnym Pyrzyce GT-2 wynika, że nowa technologia pozwoliła uzyskać wydajności zatłaczania na poziomie 130 m^3/h i zapewnić trwałość wysokich parametrów zatłaczania.
Słowa kluczowe: geotermia, woda termalna, kolmatacja, zatłaczanie, otwór chłonny

A new technology supporting the exploitation of thermal water in geothermal heating

Abstract
This article has the task of presenting the geotermy on the example of Pyrzyce Geotermy and show problems related to clogging of absorptive wellbore. This technology of preventing clogging in absorptive wellbore involves sharing a new aquifer for water injection, chemical and mechanical cleaning of the filter and conditioning of the thermal water. The research on the absorption wellbore Pyrzyce GT-2 shows that the new technology led to an injection capacity of 130 m^3/h and ensured high performance of injection.
Keywords: geotermy, thermal water, clogging, injection, absorptive wellbore

Krystyna Niesiobędzka[1], Ewelina Skiba[2]

Nowoczesne technologie oczyszczania gleb zanieczyszczonych metalami ciężkimi

1. Wstęp

Dokonujący się na przełomie ostatnich dziesięcioleci rozwój cywilizacyjny, a także powszechna chemizacja życia ludzkiego, umożliwiły nie tylko istotne i w przeważającej mierze pozytywne zmiany i ułatwienia dotyczące większości dziedzin życia ludzkiego, ale także spowodowały zakłócenie równowagi ekologicznej środowiska przyrodniczego [1]. Gwałtowny rozwój większości gałęzi przemysłu spowodował uwolnienie się do środowiska glebowego różnych ksenobiotyków – substancji chemicznych, obcych dla organizmów żywych. Nazwą **ksenobiotyki** określa się substancje, które nie są naturalnymi metabolitami organizmu, nie posiadają wartości odżywczych, a także nie są syntetyzowane przez organizm. Ważną grupę ksenobiotyków stanowią związki chemiczne otrzymane przez człowieka, o strukturze chemicznej niewystępującej w przyrodzie, do których organizmy nie przystosowały się na drodze wcześniejszej ewolucji. Za jedne z najbardziej uciążliwych ksenobiotyków z ekotoksykologicznego punktu widzenia uważa się metale ciężkie [2]. Są one substancjami o tyle problematycznymi, że w odróżnieniu od zanieczyszczeń organicznych praktycznie nie ulegają biodegradacji, a jedynie procesom biotransformacji [3, 4]. Z tego powodu niezwykle istotne staje się poszukiwanie skutecznych, a przy tym bezpiecznych dla środowiska metod remediacji gleb zanieczyszczonych metalami ciężkimi. Jedną z prężnie rozwijających się obecnie technologii jest fitoremediacja, opierająca się na zastosowaniu roślin w celu usunięcia zanieczyszczeń z gleby lub przekształcenia ich w formy charakteryzujące się jak najniższą toksycznością w stosunku do organizmów żywych. Znajduje ona szerokie zastosowanie zwłaszcza na terenach poprzemysłowych, gdzie umożliwia przywrócenie ich do takiego stanu, aby mogły zostać oddane do użytku jako obszary rekreacyjne czy przeznaczone pod zabudowę [5-7]. Celem niniejszej pracy jest przegląd wybranych technologii fitoremediacyjnych stosowanych w oczyszczaniu gleb zanieczyszczonych metalami ciężkimi oraz ich ocena pod względem mechanizmu działania oraz możliwości ich stosowania.

2. Metale ciężkie w środowisku glebowym

2.1. Definicja metali ciężkich

Za metale ciężkie przyjęto uważać metale o gęstości powyżej 4,5 $g \cdot cm^{-3}$. Obecnie pojęcie „metale ciężkie" jest na ogół znacznie szerzej rozumiane. Często określa się je jako metale, które w nadmiarze stanowią potencjalne zagrożenie dla całego środowiska, zwłaszcza jeśli będą występowały także w łatwo przyswajalnej postaci. Możemy do nich

[1] krystyna.niesiobedzka@pw.edu.pl, Wydział Instalacji Budowlanych, Hydrotechniki i Inżynierii Środowiska, Politechnika Warszawska, http://is.pw.edu.pl/index.php/pl/
[2] ewelina.skiba@pwszciechanow.edu.pl, Państwowa Wyższa Szkoła Zawodowa w Ciechanowie, http://pwszciechanow.edu.pl/

zaliczyć przede wszystkim: miedź, kobalt, chrom, kadm, żelazo, cynk, ołów, krzem, rtęć, mangan, nikiel oraz molibden. Niektóre z nich (z wyłączeniem m. in. Cd, Hg, Pb) są niezbędne do prawidłowego rozwoju i funkcjonowania organizmów żywych, czyli pełnią funkcję mikroelementów. Należy jednak pamiętać, że zarówno niedobór jak i nadmiar któregokolwiek z nich jest szkodliwy. Źródła metali ciężkich w glebach rozpatruje się w dwóch kategoriach: a) jako źródła naturalne, b) jako źródła antropogeniczne. Źródła naturalne są związane z różnymi procesami, które naturalnie zachodzą w przyrodzie, np.: wietrzenie skał, procesy glebotwórcze, pożary lasów czy erupcja wulkanów [4]. Źródła antropogeniczne to między innymi opad pyłów oraz odpady, ścieki, nawozy organiczne i mineralne, a także pestycydy, które bardzo często są produkowane na bazie metali ciężkich [8-11].

2.2. Akumulacja metali w poszczególnych organach roślin

Akumulacja metali w roślinach może przebiegać w zróżnicowany sposób w zależności od przebiegu jednego z czterech rodzajów mechanizmu: 1) wykluczanie z pędów (ang. shoot exclusion), 2) indykacja, 3) akumulacja w pędach, 4) hiperakumulacja [6-7, 12-14].

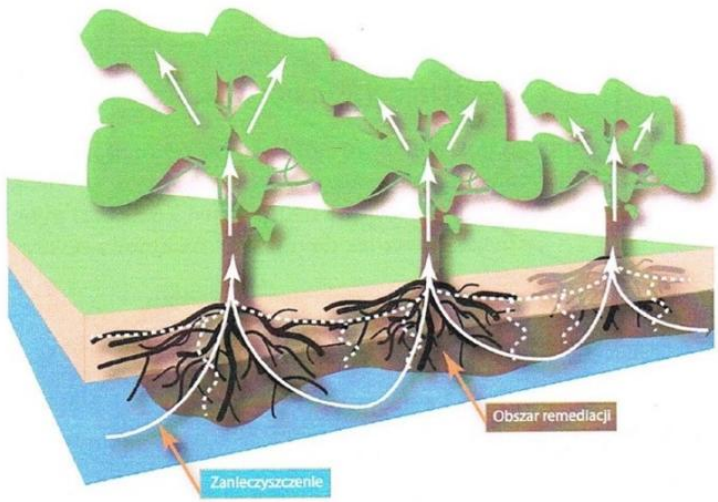
1) Rośliny wykazujące strategię wykluczania metali z pędów nazywane są roślinami wykluczającymi (ang.*excluders*). Pobierają i akumulują metale ciężkie w korzeniach roślin oraz ograniczają ich transport do pędów. Iloraz zawartości metali w liściach w stosunku do korzeni wynosi < 1,0.
2) Mechanizm indykacji stosują rośliny wskaźnikowe, czyli bioindykatory. Wykazują one pośrednią odpowiedź na wysokie stężenia metali ciężkich w glebie. Iloraz zawartości metali w roślinie do ich zawartości w glebie jest równy około 1.
3) Rośliny wykorzystujące mechanizm akumulacji w pędach przemieszczają i gromadzą duże ilości metali ciężkich w częściach nadziemnych roślin bez objawów toksyczności zarówno z gleb o niskiej jak i wysokiej zawartości metali. Iloraz stężenia metali w liściach w stosunku do korzeni wynosi > 1,0.
4) Rośliny, które wykształciły mechanizm hiperakumulacji, charakteryzujący się ekstremalnym pobieraniem metali transportowanych następnie do ich pędów. Stanowią one obiekt wielu badań naukowych z uwagi na dużą efektywność działania w fitormediacji.

3. Fitoremediacja jako alternatywny sposób oczyszczania gleb

Termin fitoremediacja (ang. *phytoremediation*) stanowi połączenie słów: greckiego phyton, oznaczającego „roślinę" oraz łacińskiego remediare, czyli „naprawiać" albo „zwalczać zło". Nazwę tę zaproponował i wprowadził do literatury w 1991 roku amerykański naukowiec z Rutgers University [3]. Fitoremediacja jest obiecującą, szybko rozwijającą się techniką remediacyjną. Pierwsze udane próby zastosowania roślin w rolniczym zagospodarowaniu ścieków komunalnych podjęto w Niemczech na początku XIX wieku. Jednak największe zainteresowanie fitooczyszczaniem przypada na wczesne lata osiemdziesiąte XX wieku. Zostały dostrzeżone wówczas szersze możliwości wykorzystania roślin, które nie były już związane tylko z oczyszczaniem ścieków bytowo-gospodarczych, ale także z usuwaniem z gleby metali ciężkich oraz ksenobiotyków organicznych, które powstały na skutek rozwoju przemysłu [12].

3.1. Fitoekstrakcja

Fitoekstrakcja (ang. phytoextraction) nazywana jest również fitoakumulacją [13]. Jest zdecydowanie najpopularniejszą i wykazującą największą efektywność działania metodą fitoremediacji gleb zanieczyszczonych metalami ciężkimi. Polega na zastosowaniu roślin w celu „ekstrakcji" czyli usunięcia metali ciężkich z gleby [Rys. 1]. Mechanizm fitoekstrakcji opiera się na trzech elementarnych procesach, czyli: mobilizacji metali w glebie, pobraniu uruchomionych metali przez korzenie roślin, a następnie ich translokacji do części nadziemnych [14]. Odkładanie zanieczyszczeń w pędach jest o tyle istotne, że części nadziemne, w odróżnieniu od korzeni, łatwo można usunąć w całości z pola [13]. Konieczne jest także, aby usuwane tą metodą metale występowały w formie rozpuszczonej w roztworze glebowym albo w innej formie przyswajalnej dla roślin. W kontakt z korzeniami wchodzą one przeważnie na drodze procesu transpiracji albo gazowej adsorpcji zachodzącej na powierzchni korzeni w strefie aeracji [6].

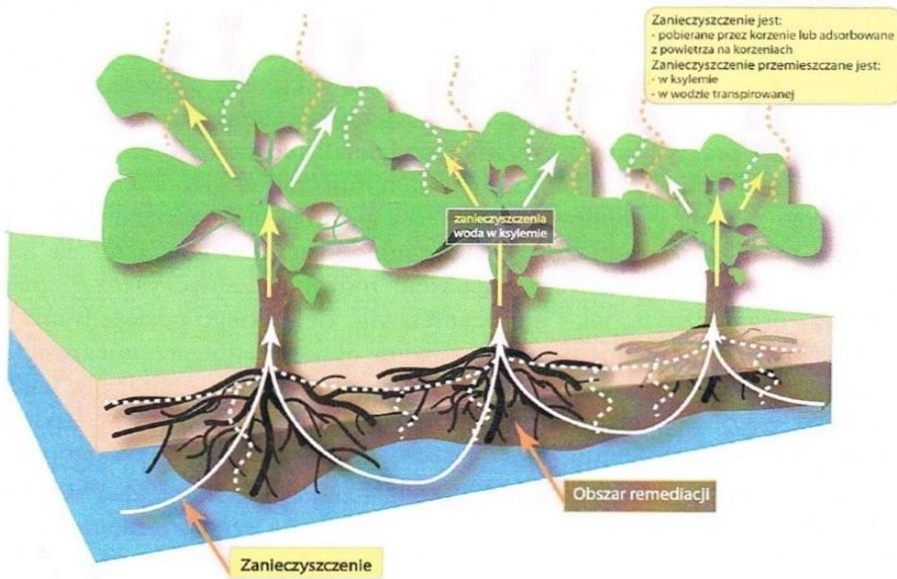


Rysunek 1. Mechanizm fitoekstrakcji
(Kacprzak M.– *Fitoremediacja gleb zanieczyszczonych metalami ciężkimi* [6])

Pojęcie fitoekstrakcji często stosowane jest błędnie jako synonim fitoremediacji, szczególnie jeśli chodzi o metale ciężkie. Należy pamiętać, ze fitoremediacja to koncepcja ogólna, fitoekstrakcja natomiast mówi o konkretnych mechanizmach. Dotychczas ze względu na odmienne mechanizmy procesów, a także stosowane gatunki roślin, zostały opracowane dwie główne strategie fitoekstrakcji: 1) fitoekstracja ciągła, 2) fitoekstrakcja indukowana, czyli wspomagana związkami chelatującymi [6,13].

3.2. Fitoulatnianie

Fitoulatnianie (ang. phytovolatilization) to proces, w którym metale ciężkie w rozpuszczalnej w wodzie formie jonowej pobrane poprzez rośliny z gleby ulegają wolatylizacji czyli przekształceniu w formy lotne, a następnie ulatniają się w procesie transpiracji wraz z parą wodną [Rys. 2.]. Metoda ta dotyczy tylko tych metali, które są w stanie wytworzyć formy lotne, czyli: rtęci, selenu i arsenu [7].



Rysunek 2. Mechanizm fitoulatniania
(Kacprzak M.– *Fitoremediacja gleb zanieczyszczonych metalami ciężkimi* [6])

W nazewnictwie polskim pojawiło się kilka terminów na określenie tej metody fitoremediacji, między innymi fitowolatylizacja, fitoewaporacja, fitoodparowanie. Spośród nich jedynie termin fitoodparowanie oddaje prawidłowo sens metody i powinien być używany jako synonim fitoulatniania [13]. Cały proces jest w dużym stopniu zależny od właściwości substancji zanieczyszczającej: jej rozpuszczalności w wodzie, polarności, hydrofobowości, współczynnika podziału: oktanol-woda (K_{ow}).

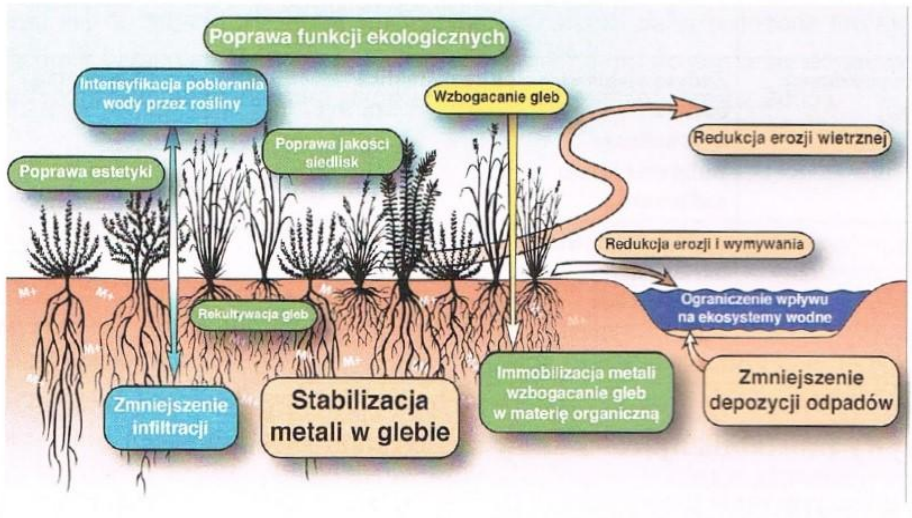
Po przeprowadzeniu licznych badań stwierdzono [15, 16], że współczynnik podziału oktanol-woda jest najlepszym czynnikiem określającym potencjał danego polutanta do pobierania go przez korzenie roślin. Zanieczyszczenia, dla których $\log K_{ow}$ przyjmuje wartości od 1 do 3,5 mogą być pobierane przez korzenie. Przy wartości $\log K_{ow}>3,5$ polutanty są zwykle zbyt hydrofobowe, słabo rozpuszczają się w roztworze glebowym lub są bardzo mocno związane z powierzchnią korzeni, uniemożliwiając w ten sposób translokację do tkanek roślinnych. Z drugiej strony przy wartości $\log K_{ow}<1$ zanieczyszczenia są zbyt hydrofilowe, nie są pobierane przez korzenie w wystarczającym stopniu ani aktywnie transportowane z powodu zbyt wysokiej polarności [4, 17].

Mechanizm fitoodparowania został opisany już w latach 40. XX wieku. Zaobserwowano wówczas, że niektóre z roślin mają zdolność nie tylko do hiperakumulacji selenu, stanowiącego ogromny problem środowiskowy, ale również wydzielają go w formie lotnej do atmosfery. Jest to najlepiej poznany przykład fitoodparowania, stąd najbardziej zaawansowane badania dotyczą właśnie fitoulatniania selenu. *Astragalus racemosus,* czyli traganek groniasty wydziela selen w postaci selenku dimetylu, o charakterystycznym czosnkowym zapachu. Generalnie selenki są transformowane dzięki metylotransferazie SeCys do metyloorganicznego dimetylu selenianiu, który jest

od 500 do 700 razy mniej toksyczny w porównaniu z formą organiczną. Inne rośliny posiadające zdolność do wolatylizacji selenu to *Stanleya piniata* oraz *Astragalus bisulcatus* [4]. Nieco inaczej przebiega proces fitoulatniania rtęci. Wynika to z faktu, że nie są znane obecnie rośliny zdolne do pobierania z gleby, a następnie odparowywania jej w taki sposób. Proces ten w warunkach naturalnych prowadzą liczne bakterie. Stąd też w ostatnich latach prowadzi się intensywne badania dotyczące modyfikacji genetycznych, mających na celu przeszczepianie roślinom zidentyfikowanych genów bakteryjnych, które są odpowiedzialne za metylację i redukcję rtęci [6-7]. Metoda fitoutleniania jest najbardziej kontrowersyjna ze wszystkich metod fitoremediacji gleb. Jest ona niewątpliwie bardzo atrakcyjna głównie ze względu na brak szkodliwych odpadów, jednak duże zastrzeżenia budzi uwalnianie do atmosfery form lotnych metali. Nie następuje więc usunięcie ich ze środowiska, lecz jedynie rozproszenie na dalej położone tereny [18]. Istnieją jednak badania, które udowadniają, że zanieczyszczenia po przedostaniu się do atmosfery i rozproszeniu nie stanowią już zagrożenia dla środowiska. Dodatkowo efektywne fitoulatnianie często wiąże się ze stosowaniem roślin genetycznie zmodyfikowanych, co także budzi kontrowersje [6, 13].

3.3. Fitochemostabilizacja

Fitochemostabilizacja (ang. phytochemostabilization) to metoda, która ma na celu immobilizację metali ciężkich w glebie poprzez użycie odpowiednich gatunków roślin oraz wprowadzanie dodatków do gleby [Rys. 3]. Unieruchomienie metali ciężkich w glebie przez rośliny oparte jest na procesach: absorpcji i kumulacji w korzeniach, adsorpcji na powierzchni korzeni oraz przekształcania w obrębie korzeni ryzosfery w trudno rozpuszczalne związki [19]. Metoda ta nie powoduje oczyszczania gleby z metali, a jedynie minimalizuje ich toksyczny wpływ na składniki ekosystemu poprzez unieruchomienie [3]. Proces ten nazywany jest często terminem fitostabilizacji (ang. *phytostabilization*). Sugeruje on, że jedyny czynnik, który decyduje o immobilizacji metali w glebie to rośliny. Jednak w celu uzyskania pełnej i trwałej stabilizacji metali w glebie konieczne jest zastosowanie nie tylko roślin (fito), ale także odpowiednich dodatków wprowadzanych do gleby (chemo). W związku z tym wydaje się, że pojęcie „chemophytostabilization" lepiej oddaje sens procesu. Z drugiej strony w metodzie tej wytwarza się trwała i zwarta okrywa roślinna, a jej ostatecznym celem jest wzrost bioróżnorodności, stąd zaproponowano stosowanie terminu „fitochemio-stabilizacja" [13]. Metoda fitochemostabilizacji polega na wprowadzeniu do gleby dodatków oraz ich wymieszania z górną warstwą gleby, czyli tą najbardziej zanieczyszczoną metalami. Zadaniem roślin wprowadzonych na tereny skażone jest zapobieganie migracji zanieczyszczeń w głąb gleby lub po jej powierzchni wraz ze spływem wód opadowych oraz ich przenikaniu do wód gruntowych.



Rysunek 3. Zalety zastosowania technologii fitostabilizacji na terenach zanieczyszczonych metalami ciężkimi (Kacprzak M.– *Fitoremediacja gleb zanieczyszczonych metalami ciężkimi* [6])

4. Ocena metod fitoremediacyjnych pod względem możliwości stosowania i ograniczeń

- Metody fitoremediacyjne mogą być stosowne jako odrębne metody oczyszczania terenów zdewastowanych lub też jako uzupełnienie metod tradycyjnych [3]. Używa się ich zarówno w przypadku remediacji małych jak i dużych obszarów. Często znajdują zastosowanie tam, gdzie nie opłaca się stosować metod tradycyjnych lub na terenach, które nie posiadają odpowiednio przystosowanej infrastruktury do prowadzenia remediacji innymi metodami [20]. W zdecydowanej większości metody fitoremediacji stosowane są jako technologie *in-situ*, czyli w miejscu skażenia, dzięki czemu nie generują dodatkowych szkód dla środowiska oraz kosztów związanych z transportem gleby [7].
- Stosowanie fitoremediacji generalnie nie wymaga specjalistycznego sprzętu ani także wyspecjalizowanego personelu. Rośliny wykorzystywane do oczyszczenia gleby tworzą swoistą „okrywę", która przyczynia się do ograniczenia procesów erozji wietrznej i zmniejszenia natężenia hałasu pochodzącego z działalności przemysłowej czy motoryzacji. Ponadto obsadzenie terenu zdegradowanego zdecydowanie poprawia jego walory estetyczne [3, 21].
- Niewątpliwą zaletę stanowi również opłacalność ekonomiczna samego procesu fitoremediacji. Czas oczekiwania na efekty jest dość długi, ale dzięki temu całkowity koszt prowadzonego procesu remediacji zostaje w pewnym sensie „rozłożony" w czasie [3, 20].
- Każdej z przedstawionych metod fitoremediacji towarzyszy powstawanie zanieczyszczonego „materiału poprocesowego". Skażona biomasa roślinna jest o tyle „użyteczna", że można ją wykorzystać w celu odzysku metali lub na cele energetyczne.

- Problematyczne może okazać się nieumiejętne stosowanie dodatków do gleby, skutkujące zanieczyszczeniem obszarów sąsiadujących z remediowaną glebą. W tym przypadku uzasadniona wydaje się być konieczność stałego monitoringu obszarów poddawanych remediacji (techniką fitochemostabilizacji) [7].
- Istotne ograniczenie każdej z technik fitoremediacyjnych dotyczy właściwego doboru rośliny (istotna jest głębokość penetracji korzeni i zdolności akumulacyjne). Problem ten jednak można zredukować poprzez stosowanie roślin z rozbudowanym systemem korzeniowym albo drzew, których korzenie sięgają długości 5-10 m [3].
- Słabą stroną technik fitoremediacyjnych jest niewątpliwie zależność samych roślin od wielu czynników. Często ich sezonowość lub zależność od klimatu stanowi poważne ograniczenie w możliwości stosowania fitoremediacji gleb zanieczyszczonych metalami ciężkimi [20].
- Efektywność procesu fitoremediacji może zostać ograniczona w wyniku uszkodzenia roślin, na przykład wskutek wystąpienia skrajnych warunków pogodowych albo pojawienia się chorób lub szkodników.
- Ograniczenie dla stosowania technik fitoremediacji może również wynikać ze zbyt wysokiego stężenia metali ciężkich w glebie, które stają się toksyczne nawet dla wysoko tolerancyjnych gatunków roślin [21].

Literatura

1. Smolińska B. *Metody oczyszczania gleb zanieczyszczonych rtęcią*, Zeszyty Naukowe Chemia Spożywcza i Biotechnologia, Politechnika Łódzka, 74 (2010), s. 121-136.
2. Przedpełska-Wąsowicz E. M., Wierzbicka M. *Wprowadzenie do ekotoksykologii w:* M. Wierzbicka (red.), Ekotoksykologia Wydawnictwa Uniwersytetu Warszawskiego, Warszawa 2015, s. 17-21.
3. Astel A., Czyżyk A., Parzych A., *Fitoremediacja metodą obniżania toksyczności gleb zanieczyszczonych metalami ciężkimi,* Czasopismo LAB Laboratoria, Aparatura, Badania, 19(4) (2014), s. 6-12.
4. Szczygłowska M., *Zbadanie możliwości kapusty białej (Brassica oleracea L. var. capitata f. alba) w procesie fitoekstrakcyjnym wybranych jonów metali ciężkich (Zn,Cd, Cu, Pb oraz Hg) z gleby*, rozprawa doktorska, Politechnika Gdańska, Gdańsk 2015.
5. Piotrowska-Niczyporuk A., Bajguz A., *Fitoremediacja – alternatywa na czyste środowisko* w: I. Ciereszko, A. Bajguz (red.) - Różnorodność biologiczna – od komórki do ekosystemu. Rośliny i grzyby w zmieniających się warunkach środowiska, Polskie Towarzystwo Botaniczne, Oddział w Białymstoku, Białystok 2013, s. 97-110.
6. Kacprzak M., *Fitoremediacja gleb zanieczyszczonych metalami ciężkimi*, Wydawnictwo Politechniki Częstochowskiej, Częstochowa 2013.
7. Karczewska A., Spiak Z., Kabała C., Gałka B., Szopka K., Jezierski P., Kocan K., *Ocena możliwości zastosowania metody wspomaganej fitoekstrakcji do rekultywacji gleb zanieczyszczonych emisjami hutnictwa miedzi*, Wydawnictwo Zante, Wrocław 2008.
8. Kabata-Pendias A., Pendias H., *Trace elements in soil and plants*, 3rd edn, CRC Press, Boca Raton, Fl, 2001.
9. Niesiobędzka, K., *Mobile forms and migration ability of Cu, Pb and Zn in forestry system in Poland*, Environmental Earth Sciences 75(2) (2016), s.1-8.
10. Krajewska E., Niesiobędzka K., *Wpływ zasolenia spływów powierzchniowych na wymywanie metali z pyłów ulicznych do fazy wodnej*, Ochrona Środowiska i Zasobów Naturalnych, (40) (2009), s.137-144.

11. Niesiobędzka K., Krajewska, E., *Badanie wpływu składowiska odpadów na wybrane elementy środowiska przyrodniczego*, Chemistry-Didactics-Ecology-Metrology, 15(2) (2010), s.179-184.
12. Marecik R., Króliczak P., Cyplik P. *Fitoremediacja – alternatywa dla tradycyjnych metod oczyszczania środowiska*, Biotechnologia, 3 (74) (2006), s. 88-97.
13. Małkowski E., *Fitoremediacja metali ciężkich w: M. Wierzbicka (red.)*, Ekotoksykologia, Wydawnictwa Uniwersytetu Warszawskiego, Warszawa 2015, s. 469-506.
14. Antonkiewicz A., *Ocena przydatności kukurydzy (Zea mays L.) do fitoekstracji kadmu, ołowiu, niklu, miedzi i cynku z gleby zanieczyszczonej tymi pierwiastkami*, Acta Scientiarum Polonorum. Formatio Circumiectus 2(1)(2003), s. 61-69.
15. Kłoczko E., *Metody eksperymentalne w chemii*, PWN, Warszawa 1978.
16. Szyszko E., *Biblioteka farmaceuty*, Zarys Biofarmacji, tom 5, PZWL, Warszawa 1984.
17. Świtnicki M. P., *Fitoremediacja – technologie usuwania zanieczyszczeń ze środowiska przy wykorzystaniu roślin,* praca licencjacka, Katedra Genetyki Ogólnej, Biologii Molekularnej i Biotechnologii Roślin Uniwersytetu Łódzkiego, Łódź 2009.
18. Cyplik P., Króliczak P., Marecik R., *Fitoremediacja – alternatywa dla tradycyjnych metod oczyszczania środowiska*, Biotechnologia, 3 (74) (2006), s. 88-97.
19. Krzyżak J., *Wspomagana fitostabilizacja metali ciężkich w glebach*, praca doktorska, Politechnika Wrocławska, Wrocław 2013.
20. Siwek M., *Biologiczne sposoby oczyszczania środowiska – fitoremediacja*, Wiadomości Botaniczne, 52(1/2) (2008), s. 23-28.
21. http://e-czytelnia.abrys.pl/zielen-miejska/2009-10-434/rosliny-4380/fitoremediacja-a-tereny-zieleni-10333 (dostęp: 01.2017).

Nowoczesne technologie oczyszczania gleb zanieczyszczonych metalami ciężkimi

Abstrakt
W artykule dokonano przeglądu technologii fitoremediacyjnych stosowanych do oczyszczania gleb zanieczyszczonych metalami ciężkimi. Zdefiniowano pojęcie metali ciężkich, podano ich źródła występowania w glebie oraz przybliżono przeciwstawne strategie, jakie wykształciły rośliny w celu przystosowania się do wysokich stężeń metali w glebach. Przedstawiono nowoczesne metody fitoremediacji takie, jak: fitoekstrakcja, fitochemostabilizacja oraz fitoulatnianie służące usuwaniu metali ciężkich ze środowiska glebowego. Dokonano analizy tych technik pod względem możliwości ich stosowania oraz ograniczeń, jakie pociągają za sobą rozważane technologie.
Słowa kluczowe: metale ciężkie, fitoekstrakcja, fitoulatnianie, fitochemostabilizacja,

Modern technologies of soils purification contaminated with heavy metals

Abstrakt
A review of modern technology for the removal of heavy metals from contaminated soils has been made.The concept of heavy metals is defined, their sources of occurrence in the soil and the opposing strategies developed by the plants to adapt to the high concentrations of metals in the soil were given. Modern phytoremediation methods such as phyto-extraction, phyto-volatilization phyto-chemostabilization used for the removal of heavy metals from the soil environment were presented. These techniques have been analyzed in terms of their applicability and their limitations.
Keywords: heavy metals, phyto-extraction, phyto-volatilization phyto-chemostabilization

Kinga Górniak[1], Monika Tuleja[2]

Indukcja wydajnej ryzogenezy z wybranych eksplantatów *Euphorbia milii* Des Moul. w warunkach kultur in vitro jako alternatywne źródła lateksu

1. Wstęp

Euphorbia milii Des Moul. (wilczomlecz okazały) (Fot.1) to gatunek wywodzący się z wyżynnej części Madagaskaru, gdzie w postaci krzewów tworzy gęste, cierniste zarośla [1]. Należy do rodziny Euphorbiaceae, rodzaju *Euphorbia* w obrębie którego poznanych jest około 2008 gatunków, z czego ponad 400 zaliczanych jest do sukulentów, powszechnie występujących w południowej i wschodniej Afryce. Dodatkowo liczne, dziko rosnące gatunki pojawiają się w Środkowej oraz Południowej Ameryce, a także w cieplejszych strefach Azj [2]. Jest to jeden z najbardziej licznych oraz różnorodnych rodzajów w królestwie roślin zaraz obok *Rumex* i *Senecio* [3].

Fot.1 *Euphorbia milii* Des Moul. [4]

Wilczomlecz w języku angielskim często określany jest jako „spurge", nazwa ta wywodzi się z łacińskiego słowa – „expurgare", które znaczy „oczyszczać". Związane jest to z przeczyszczającymi właściwościami wielu zielnych gatunków *Euphorbia,* stosowanych w tradycyjnej medycynie ludowej. W Chinach do dzisiejszego dnia używa się tej rośliny w ziołolecznictwie, w takich chorobach jak np. zapalenie wątroby oraz obrzęk brzucha a nawet w leczeniu nowotworów [5, 6]. Niektórzy Brazylijczycy wykorzystują ją w zwalczaniu brodawek [7, 8]. W ostatnich latach badania nad

[1] kinga.gorniak@student.uj.edu.pl, Zakład Cytologii i Embriologii Roślin, Instytut Botaniki, Wydział Biologii i Nauk o Ziemi, Uniwersytet Jagielloński w Krakowie, www.ib.uj.edu.pl
[2] monika.tuleja@uj.edu.pl, Zakład Cytologii i Embriologii Roślin, Instytut Botaniki, Wydział Biologii i Nauk o Ziemi, Uniwersytet Jagielloński w Krakowie, www.ib.uj.edu.pl

wieloma gatunkami roślin z rodzaju *Euphorbia* dowiodły ich skuteczności w zwalczaniu mikrobów i pasożytów wewnętrznych [9, 10]. Powszechnie liczne gatunki stosowane są jako rośliny ozdobne ogrodowe oraz doniczkowe, uprawiane w domach w wielu częściach świata [11].

Interesującą właściwością *E. milii* jest wydzielanie dużych ilości białego soku – lateksu. Lateks to złożona emulsja, zawierająca proteiny, alkaloidy, cukry, oleje, garbniki, żywice oraz gumy roślinne. Jest produkowany oraz przechowywany w specjalnych komórkach, występujących pojedynczo lub tworzących charakterystyczne rurki mleczne, określane jako „laticifers". Przewody mleczne u tych roślin tworzą system rozgałęzień, dochodzących do korzeni, pędów, liści, czasami nawet do owoców [12]. Lateks u roślin spełnia wiele funkcji, może być formą zmagazynowanych substancji odżywczych lub metabolitów wtórnych. Może chronić roślinę przed zjedzeniem przez zwierzęta roślinożerne dzięki obecności toksycznych substancji. Co więcej jest on ważnym czynnikiem obronnym rośliny przeciw patogenom, w przypadku zranienia zabezpiecza uszkodzone miejsce przed wniknięciem bakterii, wirusów oraz grzybów do jej wnętrza [13].

Dobrze zbadane zostały substancje wchodzące w skład lateksu, takie jak: euphol oraz euphorbol mające silne działanie przeciwzapalne, przeciwwirusowe a także wykazujące skuteczne zastosowanie w leczeniu bólów neuropatycznych [11]. Inne wyizolowane substancje to: euphorbol hexacosanoate, β-sitosterol oraz makrolit – lasiodiplodin, który wykazuje skuteczne działanie przeciw rozwojowi białaczki [11].

Niedawno odkryte w lateksie *E. caducifolia* związki lipofilne mogą posłużyć jako surowiec do produkcji biopaliwa, co stanowiłoby alternatywę dla paliw kopalnianych, których zasoby kończą się w bardzo szybkim tempie. Okazuję się, że rośliny z rodzaju *Euphorbia*, które mają bardzo oszczędną gospodarkę wodą, są odporne na stres związany z suszą mogłyby być hodowane w celach pozyskiwania surowców do biopaliw na terenach nieużytków rolnych [12].

Lateks pomocny jest również w zwalczaniu schistosomatozy, choroby, która obok malarii, stanowi najpoważniejsze zagrożenie parazytologiczne w skali światowej i należy do tzw. zaniedbanych chorób tropikalnych (NTD). Występuje na kontynencie afrykańskim, wschodnich wybrzeżach Ameryki Południowej, w południowo-wschodniej Azji oraz Bliskim Wschodzie. Szacowanych jest ponad 200 milionów zachorowań w 76 krajach (85% w Afryce) [14], a w grupie ryzyka zachorowań znajduje się ponad 600 milionów ludzi [15].

Schistosomatoza jest wywoływana przez larwy przywr z rodzaju *Schistosoma,* które infekują słodkowodne ślimaki, stające się wówczas ich żywicielami pośrednimi. Mogą one infekować ślimaki z wielu rodzajów jak: *Biomphalaria, Bulinus, Oncomelania*. Po osiągnięciu dojrzałości larwy *Schistosoma* wydostają się do wód słodkich w celu znalezienia żywiciela ostatecznego, którym jest człowiek. Zakażenie ma wiele objawów, które powodują liczne dolegliwości, często przechodzące w ciężkie powikłania lub nawet doprowadzają do śmierci gospodarza [15].

Światowa Organizacja Zdrowia wyznaczyła wiele sposobów zwalczania schistosomatozy. Okazuje się, że najskuteczniejszą metodą, pozwalającą wyeliminować tak dużą liczbę zachorowań, zaraz obok masowego podawania doustnego leku przeciwpasożytniczemu – parazykwantelu, może być ingerencja w cykl życiowy pasożytniczych przywr poprzez wyeliminowanie ze środowiska jej niezbędnego żywiciela pośredniego [16]. W wielu państwach, w których zastosowano tę metodę doszło do znacznego obniżenia liczby zachorowań w bardzo krótkim czasie [17, 18].

Najczęściej stosowanym syntetycznym środkiem przeciwko ślimakom jest niklozamid. Badania wykazują jednak, że po zastosowaniu tego środka następuje ponowna rekolonizacja ślimaków [19, 20]. Dodatkowo niklozamid działa toksycznie na wiele innych organizmów jak np. żaby i ryby a także na liczne rośliny wodne [21]. Niemniej jednak program kontroli ślimaków jest niezbędny, aby zahamować rozprzestrzenianie się schistosomatozy. Badania na terenach narażonych na występowanie schistosomatozy wykazały, że lateks pozyskiwany z *E. milii* jest idealną alternatywą dla niklozamidu, ponieważ jest to substancja naturalna, biodegradowalna, produkowana w dużych ilościach przez rośliny, nisko toksyczna oraz skuteczna w niewielkiej koncentracji [22, 23]. Jak wykazują badania, lateks w konkretnych dawkach, w niskich stężeniach jest zabójczy dla poszczególnych gatunków ślimaków z rodzaju *Biomphalaria,* np. badania przeprowadzone w warunkach polowych na terenie Minas Gerais w Brazylii wskazują, że mleczny sok *E. milii* zabija 100% *B. glorbata* oraz *B. tenagophila* w stężeniu 5 i 12 ppm odpowiednio [27-30].

W związku z tym, że lateks wykazuje pewną rozbieżność w swojej potencji w zależności od terenu występowania rośliny oraz pory roku w jakiej został pozyskany [25], uzasadnionym wydaję się opracowanie metody pozyskiwania stabilnej formy lateksu w warunkach laboratoryjnych np. wykorzystując metodykę hodowli roślinnych kultur in vitro. Jedną z możliwości jest indukcja ryzogenezy w warunkach kultur in vitro tak jak w przypadku *Heliotropium indicum* [31] lub innych roślin o znaczeniu farmaceutycznym [32, 33]. Ryzogeneza to proces, powstawania korzeni przybyszowych z komórek wyizolowanego z rośliny macierzystej eksplantatu. Zawiązki organu mogą powstawać z komórki lub grup komórek kompetentnych do odbierania i reagowania na sygnały. Komórki te mogą być kompetentne w momencie izolowania z rośliny macierzystej lub nabywać tę cechę pod wpływem różnorodnych czynników wykorzystywanych w kulturach in vitro [34, 35].

U około 60% roślin wykorzystywanych w celach medycznych, korzenie stanowią główny rezerwuar substancji leczniczych [36] Zdecydowana większość tych roślin (90%) pozyskiwana jest ze stanowisk naturalnych co w znaczniej mierze powoduje wyniszczanie ich naturalnych siedlisk [37]. Tylko nieznaczny procent roślin wykorzystywanych do produkcji leków, jest pozyskiwany z hodowli tradycyjnych. Kultura korzeni przybyszowych stanowi więc alternatywne i wydajne źródło sposobu pozyskiwania substancji leczniczych w warunkach laboratoryjnych [31].

Celem pracy było opracowanie wydajnego systemu regeneracyjnego drogą ryzogenezy dla wybranego typu eksplantatu *Euphorbia milii* w warunkach kultury *in vitro,* w kontekście efektywnego uzyskiwania lateksu oraz zobrazowanie zmian morfo-histologicznych zachodzących podczas hodowli.

2. Materiały i metody

2.1. Założenie hodowli

W przeprowadzonych eksperymentach wykorzystano eksplantaty pochodzące z roślin hodowanych w Zakładzie Cytologii i Embriologii Roślin UJ. Były to fragmenty blaszek liściowych oraz fragmenty pędów *Euphorbia milii*. W badaniach zastosowano 7 układów doświadczalnych, które oparte były na standardowej pożywce MS [38] z dodatkiem 30g/l sacharozy oraz 8g/l agaru. Skład regulatorów wzrostu ustalono, tak aby w zastosowanych warunkach można było wykorzystać pożywki z różną kombinacją roślinnych regulatorów wzrostu (Tabela 1). W ten sposób

wykorzystano pożywki: (1) z przewagą auksyny, (2) z przewagą cytokininy (3) z równoważną ilością auksyny i cytokininy. Wartość pH każdej pożywki została doprowadzona do 5,7-5,8. Wysterylizowane liście cięto podłużnie wzdłuż nerwu głównego oraz w poprzek blaszki liściowej na 4 części. W przypadku mniejszych, młodych liści krojono je jedynie poprzecznie na 2 części lub pozostawiano w całości. Powstałe w ten sposób eksplantaty posiadały średnie wymiary 1 cm x 1 cm. Fragmenty łodyg przecięto w poprzek na krążki o grubości 2 mm. Tak przygotowane eksplantaty zostały wyłożone na szalki Petriego z wszystkimi rodzajami pożywek. Blaszki liściowe oraz fragmenty pędów układano w liczbie 6 do 7 na jednej szalce. Pożywki 1-4 (Tabela 1) zawierały po co najmniej 12 eksplantatów blaszek liściowych oraz fragmentów pędów. Natomiast pożywki 5-7 (Tabela 1) zawierały po co najmniej 21 eksplantatów obu rodzajów. Kontrole stanowiły sterylne eksplantaty umieszczone na standardowej pożywce MS [38].

Tabela 1. Skład pożywek wykorzystywanych w doświadczeniu

Numer Pożywki	Auksyna	Cytokinina
0 (kontrola)	Brak	Brak
1	0,80 mg/L IAA	1,00 mg/L BAP
2	0,35 mg/L IAA	1,00 mg/L BAP
3	0,80 mg/L IBA	0,25 mg/L BAP
4	0,40 mg/L IBA	1,00 mg/L BAP
5	0,80 mg/L NAA	2,00 mg/L 2iP
6	5,00 mg/L NAA	2,00 mg/L 2iP
7	2,00 mg/L NAA	2,00 mg/L 2iP

Źródło: opracowanie własne

2.2. Inkubacja

Kultury inkubowano w pomieszczeniu hodowlanym w temperaturze 23-25^{o}C, w 16-godzinnym fotoperiodzie (chłodne białe światło fluorescencyjne o natężeniu 39 $\mu mol \cdot m^{-2} \cdot s^{-1}$).

2.3. Monitorowanie przebiegu hodowli

Obserwacji dokonywano przez okres kilkunastu tygodniu i dokumentowano w postaci zdjęć z wykorzystaniem lupy binokularnej typu Opta-Tech połączonej z kamerą cyfrową. Zdjęcia preparatów histologicznych wykonywane były za pomocą mikroskopu świetlnego typu Opta-Tech połączonego z aparatem cyfrowym firmy Nikon.

2.4. Analiza histologiczna

W celu analizy histologicznej wykonano preparaty histologiczne z wyizolowanych części eksplantatów z regenerantami pochodzącymi z kultur ryzogennych w 43 dniu hodowli. Materiał roślinny został utrwalony za pomocą roztworu 5% glutaraldehydu w 0,1 M buforze fosforanowym o pH 7.2. Tak przygotowane tkanki roślinne przesycono żywicą epoksydową, następnie pokrojono na cienkie skrawki za pomocą mikrotomu rotacyjnego (Microm; Thermo Fisher Scientific, USA).

Preparaty zostały wybarwione z wykorzystaniem wodnego roztworu 0,1% błękitu toluidyny. Szkiełka ze skrawkami umieszczono na płycie rozgrzanej do temperatury 70°C. Kilka kropel barwnika naniesiono na ciepłe preparaty i po upływie 40s inkubacji spłukano delikatnym strumieniem bieżącej wody i następnie pozostawiono do wysuszenia w temperaturze pokojowej. Barwienie to wykorzystano w celu wybarwienia struktur komórkowych: cytoplazmy, jąder komórkowych, jąderek i wyróżnienia ścian komórkowych. Preparaty również poddano reakcji PAS w celu uwidocznienia skrobi i nierozpuszczalnych węglowodanów. Zastosowano 0,1% kwas nadjodowy przez40 min. Następnie preparaty przepłukiwano wodą destylowaną przez 10 min. Po tym czasie szkiełka z preparatami przeniesiono do odczynnika Schiff'a na 40 minut (Przygotowanie odczynnika Schiff'a: 0,5 g fuksyny zasadowej rozpuszczono w 100 ml wrzącej wody destylowanej. Roztwór przesączono po oziębieniu do temperatury 55°C. Otrzymany przesącz oziębiono do 26°C. Następnie dodano 10 ml 1NHCl oraz 0,5 g pirosiarczanu potasu. Po upływie doby do roztworu dodano 0,5 g węgla aktywnego. Tak przygotowaną mieszaninę wytrząsano przez 1 godzinę. Po upływie tego czasu ponownie przesączono). Po wyjęciu z odczynnika przepłukano uprzednio przygotowaną wodą siarkową (10 ml $K_2S_2O_5$ + 200 ml wody destylowanej + 10 ml HCl) dwukrotnie po 10 minut. Ostatnim etapem było płukanie preparatów w wodzie wodociągowej. Po 45 minutach szkiełka wyjęto i pozostawiono w temperaturze pokojowej do wyschnięcia.

2.5. Analiza statystyczna

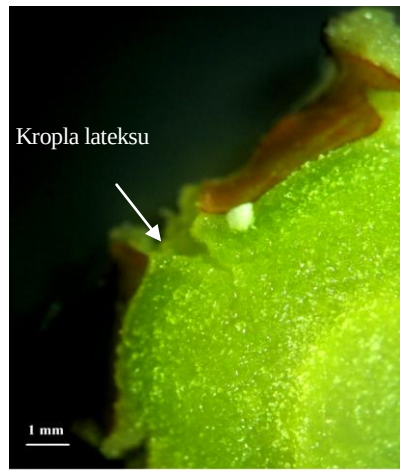
Po 7 tygodniach zliczono ilość utworzonych zawiązków korzeni na blaszkach liściowych (liczba ekspantantów tworzących korzenie x 100/ liczba wszystkich eksplantatów na danej pożywce oraz ilość wszystkich powstałych korzeni/ ilość ryzogenicznych eksplantatów). Wykonano obliczenia średnich i odchyleń standardowych dla 21 eksplantatów w trzech odrębnych doświadczeniach z użyciem programu Microsoft Excel. Odchylenia standardowe przedstawiono w postaci słupków błędu na wykresie obrazującym średnie wartości liczby korzeni na eksplantat dla każdego układu doświadczalnego.

3. Wyniki

3.1. Eksplantaty pochodzące z fragmentów pędów

3.1.1. Indukcja Kalusa

Pierwsze symptomy odpowiedzi morfogenetycznej zaobserwowano w 7 dniu hodowli. Kalus pojawił się na wszystkich eksplantatach (100%) utrzymywanych na pożywkach zawierających NAA oraz 2iP w każdej z zastosowanych kombinacji (Tabela 2.), poczynając od ich powierzchni ciętych a następnie rozchodząc się na dalsze partie, w konsekwencji całkowicie je pokrywając. W 14 dniu hodowli zanotowano różnice w wyglądzie kalusa we wszystkich zastosowanych układach doświadczalnych. Obserwowano kalus żółto-zielony, biało-różowy powstający prawdopodobnie w obrębie miękiszu kory pierwotnej a także intensywnie zielony, tworzący się w obrębie steli. W trakcie hodowli obserwowano pojawianie się na poszczególnych eksplantatach niewielkich ilości lateksu z komórek kalusa (Fot 2.), pojedynczych cierni oraz włosków przyrannych. Na eksplantatach utrzymywanych na pożywce kontrolnej oraz zawierających auksynę IAA lub IBA w różnych kombinacjach z cytokininą BAP nie obserwowano indukcji kalusa.



Fot 2. Kalusujący eksplantat z białą kroplą lateksu na powierzchni komórek kalusa uzyskany na pożywce nr 6 (Tabela 1) [opracowanie własne]

3.1.2. Indukcja ryzogenezy

Na eksplantatach pochodzących z fragmentów pędów w żadnym z zastosowanych układów doświadczalnych nie doszło do formowania korzeni.

3.2. Eksplantaty pochodzące z fragmentów blaszek liściowych

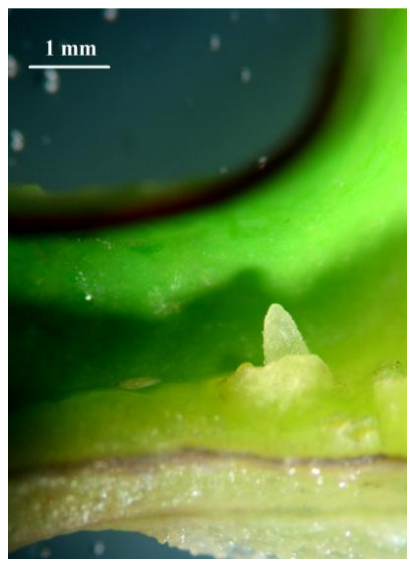
3.2.1. Indukcja kalusa

Po 7 dniach od założenia hodowli nieduży procent eksplantatów (13,2%) znajdujących się na pożywce MS bez dodatku roślinnych regulatorów wzrostu wytworzyło tkankę przyranną (Tabela 2.) Również po tygodniu trwania hodowli na pożywkach z auksyną NAA oraz cytokininą 2iP doszło do indukcji kalusa w tkankach eksplantatów. Kalus zwykle tworzył się na obrzeżach blaszki liściowej a także często od spodniej strony eksplantatu lub przy ogonku liściowym. Nigdy nie pokrywał całości eksplantatu. Przyjmował odcienie od żółtego przez kremowy do białego tworząc dość zwartą formę komórek. W największej ilości kalus obserwowany był na pożywce nr 6 z przewagą auksyny NAA (5 mg/l) w stosunku do cytokininy 2iP (2 mg/l). W niewiele mniejszym stopniu doszło do wytworzenia tkanki przyrannej na pożywce 5 i 7 (Tabela 2). Eksplantaty pochodzące z blaszki liściowej utrzymywane na pożywce 1, 2, 3 i 4 po 3 tygodniach hodowli całkowicie obumarły nie dając żadnej odpowiedzi morfogenetycznej (Tabela 2.).

3.2.2. Indukcja ryzogenezy

W 14 dniu, hodowli na eksplantatach z blaszek liściowych zaobserwowano pierwsze symptomy organogenezy – ryzogenezy (Fot 3). Zawiązki korzeni formowały się głównie bezpośrednio z tkanki eksplantatu bez uprzedniego wytworzenia kalusa i były widoczne zarówno na obrzeżach blaszki liściowej jak i na górniej powierzchni liścia, a także powstawały z okolic nerwu głównego. W zdecydowanie mniejszym stopniu obserwowano ryzogenezę pośrednią. Jedynie w sporadycznych przypadkach dochodziło do formowania się korzeni z komórek kalusa.

W trakcie trwania hodowli wyraźnie widoczny był intensywny wzrost elongacyjny regenerantów (Fot 4). Najdłuższy korzeń dorastał do 2 cm długości. Regeneranty nie tworzyły korzeni bocznych, jak również nie przypominały systemu korzeniowego wiązkowego. Oprócz obecności starszych korzeni, na tych samych eksplantatach zauważono w tym samym czasie hodowli, formowanie się nowych, młodych korzeni, obserwowana więc ryzogeneza miała charakter niezsynchronizowany. Znaczna większość reagujących eksplantatów wytwarzała bardzo liczne korzenie (Tabela 2). Na pożywce z przewagą cytokinin (pożywka nr 5) 90% ekspalntatów wytworzyło największą liczbę korzeni na eksplantat. Na tej pożywce korzenie były najdłuższe. Na podłożu nr 7 otrzymano najniższy procent eksplantatów dających ryzogenezę – 75% jak również uzyskano najmniejszą liczbę korzeni na eksplantat (Wykres 1).



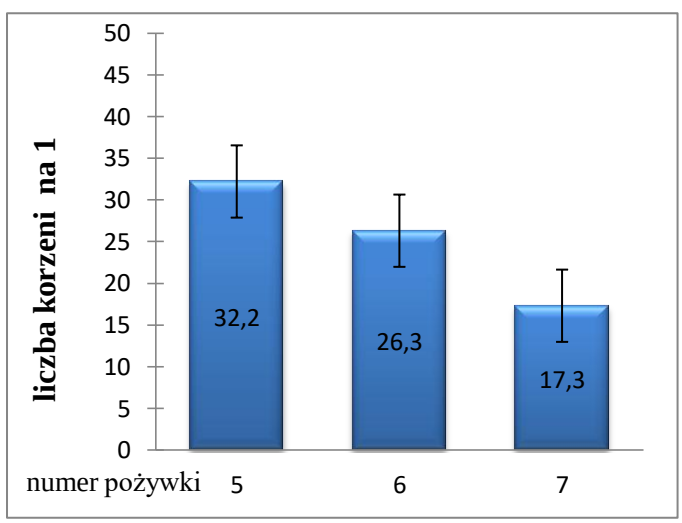
Fot 3. Zawiązek korzenia przybyszowego wyrastający na eksplantatach fragmentów blaszek liściowych w 2 tygodniu hodowli na pożywce nr 6 (Tabela 1.) [opracowanie własne]

Fot 4. Korzenie przybyszowe z blaszki liściowej na pożywce nr 5. (Tabela 1.) w 7 tygodniu hodowli [opracowanie własne]

Tabela 2. Rodzaj odpowiedzi morfogenetycznej i procent reagujących eksplantatów pochodzących z fragmentów pędów oraz blaszek liściowych *Euphorbia milii* na pożywkach MS w 7 różnych kombinacjach auksyn i cytokinin

	Eksplantaty pochodzące z fragmentów pędów		Eksplantaty pochodzące z fragmentów blaszek liściowych	
Numer pożywki	% ekslpantatów indukujących zawiązki korzeni	% eksplantatów indukujących kalus	% ekslpantatów indukujących zawiązki korzeni	% eksplantatów indukujących kalus
0 (kontrola)	Brak	Brak	5,6	13,9
1 [0,80 mg/L IAA + 1,00 mg/L BAP]	Brak	Brak	Brak	Brak
2 [0,35 mg/L IAA + 1,00 mg/L BAP]	Brak	Brak	Brak	Brak
3 [0,80 mg/L IBA + 0,25 mg/L BAP]	Brak	Brak	Brak	Brak
4 [0,40 mg/L IBA + 1,00 mg/L BAP]	Brak	Brak	Brak	Brak
5 [0,80 mg/L NAA + 2,00 mg/L 2iP]	Brak	100,0	90,0	71,4
6 [5,00 mg/L NAA + 2,00 mg/L 2iP]	Brak	100,0	84,2	78,9
7 [2,00 mg/L NAA + 2,00 mg/L 2iP]	Brak	100,0	75,0	65,0

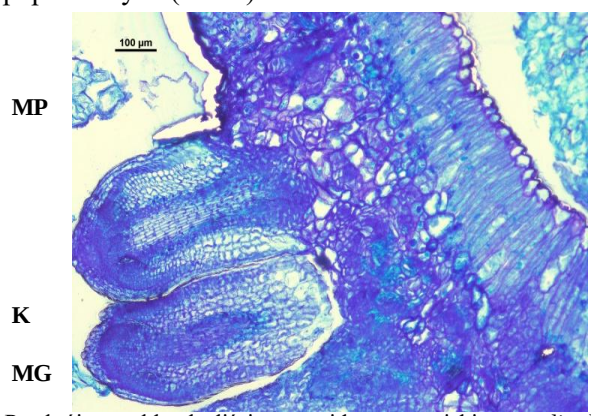
Źródło: opracowanie własne



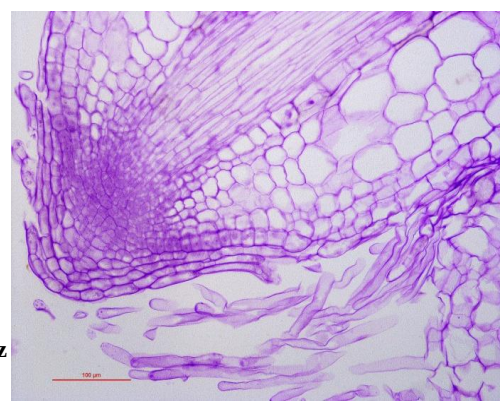
Wykres 1. Średnia liczba korzeni powstałych na eksplantatach z blaszki liściowej na trzech rodzajach pożywek [opracowanie własne]

3.3. Analiza histologiczna

Analiza histologiczna potwierdziła przeważający, bezpośredni charakter otrzymanej morfogenezy na zdecydowanej większości badanych eksplnatatów, w zastosowanych układach doświadczalnych. Korzenie powstawały z warstw komórek mezofilu znajdujących się tuż pod epidermą, najprawdopodobniej komórek miękiszu gąbczastego (Fot 5.). Przekrój korzeni zarówno poprzeczny jak i podłużny wykazał, że anatomia powstałych regenerantów nie odbiega od schematu budowy korzeni rośliny okrytonasiennej. Korzenie formowały liczne włośniki, widoczne na przekrojach podłużnych oraz poprzecznych (Fot 6.).



Fot 5. Przekrój przez blaszkę liściową z widocznym miękiszem palisadowym **(MP)** i miękiszem gąbczastym **(MG),** z którego powstają dwa korzenie **(K)** na pożywce nr 7 (powiększenie 10x) [opracowanie własne]



Fot 6. Przekrój podłużny przez korzeń z widocznymi włośnikami (**W)**, merystemem wierzchołkowym korzenia (**M**) oraz czapeczką korzenia (**Cz**) na pożywce nr 5 (powiększenie 20x) [opracowanie własne]

4. Analiza wyników – dyskusja

Z niniejszych badań przeprowadzonych nad eksplantatami pochodzącychz fragmentów blaszek liściowych oraz fragmentów pędów *Euphorbia milii* wynika, że zastosowanie pożywki w oparciu o pożywkę MS [41] z przeważającą zawartością cytokininy 2iP (2 mg/l) względem ilości auksyny NAA (0,8 mg/l) jest optymalne w inicjacji i wydajności procesu ryzogenezy na eksplatatach pochodzących z blaszki liściowej, dając średnio 32,2 ± 20,7 SD korzeni powstałych na jednym eksplantacie

(Wykres 1.). Wykorzystanie tej kombinacji fitohormonów daje wysoki procent ryzogenezy w każdym z przypadków. Takie rezultaty nie zostały dotychczas opisane w żadnej pracy badawczej. Jest to więc pierwszy otrzymany proces ryzogenezy uzyskany na eksplantatach pochodzących z blaszek liściowych *E. milii*, mogący posłużyć do prowadzenia kultur tkankowych, między innymi w celach aplikacyjnych. W niniejszej pracy obecność cytokininy 2iP nie wpłynęła hamująco na proces formowania się korzeni. Wyniki uzyskane w przeprowadzonym doświadczeniu pokazują, że przewaga cytokininy nad zastosowaną auksyną wpływa korzystnie na indukcję i intensywność ryzogenezy, co stoi w sprzeczności z ogólnie przyjętą hipotezą [39], że zwiększenie ilości cytokinin w pożywce powoduje obniżenie odpowiedzi ryzogenicznej. Zjawisko to jest związane z wywoływanymi przez cytokininy zaburzeniami polarnego transportu auksyn w eksplantacie, co z kolei doprowadza do ich niedoboru i skutkuje brakiem indukcji ryzogenezy [40]. Tak odbiegające wyniki mogą sugerować, że u tego rodzaju wilczomlecza proces ten może odbiegać od przyjętych norm. Jest to jednak jedynie przypuszczenie, które należałoby dokładnie zweryfikować, poprzez dalsze badania nad *E. milii* w celu poznania dokładnego wpływu poszczególnych roślinnych regulatorów wzrostu na wybrane tkanki tej rośliny. Przeważający bezpośredni przebieg ryzogenezy gwarantuje stabilność genetyczną, co może stanowić zaletę tych kultur i mogłoby być wykorzystane w hodowli tkankowej korzeni w bioreaktorach w celu pozyskiwania substancji zwalczających gospodarzy pośrednich chorobotwórczych przywr z rodzaju *Schistosoma*. Kultury korzeni w bioreaktorach charakteryzuje szybkie tempo proliferacji oraz wysoki współczynnik namnażania, dlatego najczęściej wykorzystywane są do pozyskiwania naturalnych związków roślinnych na skalę produkcyjną.

W niniejszej pracy przedstawiono innowacyjną procedurę indukcji wydajnej ryzogenezy z blaszek liściowych *E. milii*, w obecności auksyny i cytokininy (0,80 mg/L NAA oraz 2,00 mg/L 2iP), która może stanowić alternatywny sposób pozyskiwania lateksu zawartego w korzeniach. Co więcej, otrzymana tkanka kalusowa mogłaby posłużyć do prowadzenia kultur agregatów komórek lub pojedynczych komórek w płynnych pożywkach również w celu pozyskiwania lateksu. Uprzednio należałoby przeanalizować obecność oraz skład chemiczny lateksu zawartego w korzeniach wyprodukowanych w hodowli *in vitro* oraz w tkance kalusowej za pomocą odpowiedniej metody, co jest przedmiotem kontynuowanych badań.

Literatura

1. Masternak A. Wielka Encyklopedia Przyrody. *Rośliny kwiatowe 1*.Muza: (1998), s. 394-415.
2. Carter S., „Euhorbia", In: U.Eggli, Ed., *Dicotyledons:Illustrated Handbook of Succulent Plants,* Springer,Heidelberg (2002), s. 102.
3. Dhayalan M., Jegadeeshwaril A., Gandhin N., *Biological activity sources from traditionally used tribe and herbal plants material*, Asian Journal of Pharmacutical and clinical Research Vol 8 (6) (2015).
4. http://plantemania.fr/caudex-pachycaules/793-euphorbia-milii.html (31.03.2017).
5. Yasukawa K., Akihisa T., Yoshida Z.Y., *Inhibitory effect of euphol, a triterpene alcohol from the roots of Euphorbia kansui, on tumor promotion by 12-o- tetradecanoylphorbol-13-acetate in two-stage carcinogenesis in mouse skin.* J. Pharm. Pharmacol. 52(2000) s. 119-124.

6. Kedei N., Lundberg D.J., Toth A., Welburn P., Garfield S.H., Blumberg P.M., (2004) *Characterization of the Interaction of Ingenol 3-Angelate with Protein Kinase C. Cancer.*, (2004), (64), s. 3243-3255.
7. Fatope M.O., Zeng L., Ohayagha J.E., McLaughlin J.L., *New 19-acetoxyingol diterpenes from the latex of Euphorbia poisonii.* Journal of Medicinal Chemistry, (1996a) 39(4), s.1679-1683.
8. Fatope M.O., Zeng L., Ohayagha J.E., Shi G. McLaughlin J.L., *Selectively cytotoxic diterpenes from Euphorbia poisonii,* Journal of Medicinal Chemistry, (1996b), 39(4), s.1005-1008.
9. Mazoir N., Banharref A., Bailen M., Reina M., Gonzalez-Coloma A., Martinez-Diaz R.A., *Antileishmanial and Antitrypanosomal Activity of Triterpene Derivatives from Latex of Two Euphorbia Species,* Z. Naturforsch., (2011), (66c), s. 360-366.
10. Duarte N., Kayser O., Abreu P., and Ferreira M. J., *Antileishmanial activity of piceatannol isolated from Euphorbia lagascae seeds. Phytother*. (2008), (22), s. 455-457.
11. Yadav S.C., Pande M., Jagannadham M.V., *Highly stable glycosylated serine protease from the medicinal plant Euphorbia milii. Phytochemistry*, (2006), (67), s. 1414-1426.
12. Rajeswari B., Kumar, S.P., Rao A.P., Khan P.S.S.V., *A distribution and ultrastructure oflaticifers in the Pholloclade od Euphorbiacaducifolia Haines a Potential Hydrocarbon Yielding CAM Plant*, American Journal of Plant Sciences, (2014), (5), s. 70-79.
13. Delgado I.F. and Paumgartten F.J.R., *Effects of Euphorbia milii on mitogen-induced lymphocyte proliferation*, Rev.Bras.PI.Med.,Campinas, 16 (1) (2014), s.107-111.
14. Delgado I.F., De-Carvalho R.R., De-Oliveira A.C.A., Kuriyama S.N., Oliveira-FilhoE.C., Souza C.A.M., Paumgartten F.J.R., *Absence of tumor promoting activity of Euphorbia milii latex on the mouse back skin.* Toxicology letters, 145 (2) (2003), s. 175-180.
15. Oliveira-Filho E. and Francisco Paumgartten J.R., *Toxicity of Euphorbia milii Latex and Niclosamide to Snails and Nontarget Aquatic Species.* Ecotoxicology and Environmental Safety 46 (1999), s. 342-350.
16. Sun Q., Peng Z.,Zhang J., Jiang J., *Response of Oncomelania snail distribution on land use in Sichuan, China,* African Jurnal of Biotechnology. 10(63)(2011), s.13835-13840.
17. King C. H.,Sutherland L.J., Bertsch D., *Systematic Review and Meta-analysis of the impact of Chemical-Based Mollusciciding for Control of Schistosoma mansoni and S. haematobium Transmission,* PLOS Neglected Tropical Diseases. 10 (2015).
18. King C.H.,Bertsch D., Historical Perspective Snail Control to Prevent Schistosomiasis, PLOS Neglected Tropical Diseases. 10 (2015).
19. Khallaayoune K., Madsen H., Laamrani H., *Evaluation of three methods to control Bulinus truncatus, the intermediate host of Schistosoma haematobium in an irrigation scheme, Tessaout-Amont, Morocco.* Acta Tropica 69 (1998), s.51-63.
20. Pieri O.S, Gonçalves J.F, Sarquis O., *Repeated focal molluscicide for snail control in a sugar-cane area of northeast Brazi*l. Mem Inst Oswaldo Cruz 90,(1995), s. 535-536.
21. Andrews P., Thyssen J., Lorke D, *The biology and toxicology of molluscicides, Bayluscide*. Pfarmocol. Ther., 19 (1983), s. 245-249.
22. De Vasconcellos M.C., De Amorim A., *Activity of Euphorbia splenders var. hislopii N.E.B. (Euphorbiaceae) Latex against Lymnaea comulla (Say,1817) (Pulmonata: Lymnaeidae), Intermediate Host of Fasciola hepatica, Linnaeus, 1758 (Trematoda:Fasciolidae). 2: Limited Filed-testing,* Mem inst. Oswaldo Cruz, Rio de Janeiro. 98 (7) (2003).
23. Schall V.T., Vasconcellos C., Rocha R.S., Souza C.P., Mendes N.M., *The control of the schistosome-transmitting snail Biomphalaria glorbata by the plant Molluscicide Euphorbia splenders var.hislopii (syn milii Des. Moul): a longitudinal field study in an endemic area in Brazil*, Acta Tropica, 79 (2001), s. 165-170.
24. De Vasconcellos MC, Schall VT. (1986) *Latex of "coroa de cristo" (Euphorbia splendens): an effective molluscicide*. Mem Inst Oswaldo Cruz 81 (1986), s.475-476.
25. Schall V.T., de Vasconcellos M.C., de Souza C.P., Baptista D.F. *The molluscicidal activity of Crown of Christ (Euphorbia splendens var. hislopii) latex on snails acting as*

intermediate hosts of Schistosoma mansoni and Schistosoma haematobium. Am.J.Trop.Med.Hyg. 58(1), (1998), s.7-10.
26. Zamith, H. P. S.,Paumgartten,F.J.R., and Speit, G. *Evaluation of the mutagenicity of the molluscicidal latex of ' Christ's Crown' (Euphorbia milii var. hislopii) in mammalian cell in vitro and in vivo.* Mutat. Res., 368 (1996), s. 15-20.
27. Baptista D.F., Vasconcellos M.C., Lopes F.E.F., Silva I.P, Schall V.T *Evaluation of the molluscicidal properties of Euphorbia splendens var. hislopii (N.E.B.) (Euphrbiaceae). Investigation in lotic habitat.* Me. Inst Oswaldo Cruz., 87 (1992), s.549-553.
28. Baptista D.F., Vasconcellos M.C., Lopes F.E., Silva I.P., Schall V.T. *Perspectives of using Euphorbia splendens as a molluscicide in schistosomiasis control programs*. The Southeast Asian Journal of Tropical Medicine and Public Health 25(3) (1994), s. 419-424.
29. Mendes N.M., Baptista D.F, De Vasconcellos M.C., Schall V.T. *Evaluation of the molluscicidal properties of Euphorbia Splendens var. Hislopii (N. E. B.) (Euphorbiaceae) – 1. Experimental test in a lentic habitat.* Mem. Inst. Oswaldo Cruz, Vol. 87 (1) (1992), s.21-23.
30. Mendes N.M., De Vasconcellos M.C., Baptista D.F, Rocha R.S., Schall V.T., *Evaluation of the molluscicidal properties of Euphorbia Splendens var. Hislopii (N. E. B.). Latex Experimental test in an endemic area in the state of Minas Gerais, Brazil.* Mem. Inst. Oswaldo Cruz 92 (5) (1997), s. 719-724.
31. Bagadekar A.N., Jayaraj M., *In vitro rhizogenesis from leaf and stem callus of Heliotropium Indicum, L.- Medicinal Herb., International Jurnal of Plant, Animal and environmental sciences*, (2011) Vol.1, Issue-2.
32. Sakhare S. P., Mendhulkar V. D., *Embryo excised callus induction and rhizogenesisin Linum Usitatissimum L.*, International Journal of Pharma and Bio Sciences,(2016), 7(3), s. 507-511.
33. Dunal P.A. Wadegaonkar K.A., Bhagwat M.K., *Direct rhizogenesis and establishment of fast growing normal root organ culture of Withania somnifera*, Rai Plant Cell, Tissue and Organ Culture (2006), (82), s. 223-225.
34. Woźny A., Przybył K., *Komórki roślinne w warunkach stresu. Tom II. Komórki in vitro.* Wydawnictwo Naukowe UAM, (2007).
35. Naija S., Nadhra Elloumi N., Jbir N., Saida Ammar S., Claire Kevers C., *Anatomical and biochemical changes during adventitious rooting of apple rootstocks MM 106 cultured in vitro*, C. R. Biologies 331, Science Direct (2008), s. 518-525.
36. Agarwal V. S., Ghosh B., *Drug plants of India (Root drugs)*, Kalyani Publishers, Ludhiana, India, (1995).
37. Ved D. K., Mudappa A., Shankar D., *Regulating export of endangered medicinal plant species-need for scientific rigor.* Curr Sci, (1995), (75), s. 341-341.
38. Murashige T., Skoog F. *A Revised Medium for Rapid Growth and Bio Assays with Tobacco Tissue Culture*. Physiologia Plantarum vol. 15 (3) (1962), s. 473-497.
39. Choffe K. L., Victor J. M. R., Murch S. J., Saxena P. K. *In vitro regeneration of Echinacea purpurea L.: direct somatic embryogenesis and indirect shoot organogenesis in petiole culture.* In Vitro cellular & Development Biology – Plant Vol., 36 (2000), s. 30-36.
40. Laplaze L., Benkova E., Casimiro I., Maes L., Vanneste S., Swarup R., Weijers D., Calvo V., Parizot B., Herrera-Rodriques M., B., Offringa R., Graham N., Doumas P., Friml J., Bogusz D., Beeckman T., Brnnrtt M. *Cytokinins act directly on lateral root founder cells to inhibit root initiation.* The Palnt Cell Vol. 19 (2007), s. 3889-3900.

Indukcja wydajnej ryzogenezy z wybranych eksplantatów *Euphorbia milii* Des Moul. w warunkach kultur in vitro jako alternatywne źródła lateksu

Abstrakt
Schistosomatoza (bilharzia) obok malarii stanowi najpoważniejsze zagrożeniem parazytologiczne w skali światowej. Należy do tzw. zaniedbanych chorób tropikalnych (NTD), gdzie zamieszczona została na pierwszym miejscu. Szacuję się 258 mln zachorowań z czego 91% występuje w Afryce (WHO, 2014). Schistosomatoze wywołują pasożytnicze przywry z rodzaju *schistosoma*, których żywicielami pośrednimi są ślimaki słodkowodne z rodzajów *Bulinus*, *Biomphalaria*, *Oncomelania* natomiast żywicielami ostatecznym są kręgowce w tym człowiek. Światowa Organizacja Zdrowia wyznaczyła wiele sposobów zwalczania schistosomatozy. Jednym z nich jest Program Kontroli Ślimaków mający na celu wyeliminować gospodarzy pośrednich chorobotwórczych przywr, obejmujący tereny endemiczne dla tej choroby. Do tej pory stosuje się głównie syntetyczny środek, który wykazał toksyczne działanie na niektóre nie docelowe organizmy wodne, dlatego ważne stało się znalezienie alternatywnej substancji, jaką okazał się lateks pochodzący z Wilczomlecza okazałego *(Euphorbia milii* Des Moul.). Ten mleczny sok jest produkowany oraz przechowywany w specjalnych komórkach, występujących pojedynczo lub tworzących charakterystyczne rurki mleczne, określane jako „laticifers". Przewody mleczne u tych roślin tworzą system rozgałęzień, dochodzących do korzeni, pędów oraz liści. Aktualnie przeprowadzanych jest wiele badań nad wpływem tej naturalnej, biodegradowalnej substancji w porównaniu do syntetycznych środków. Aby wyeliminować rozbieżność w składzie chemicznym oraz skuteczności pozyskiwanego lateksu, a także w celu ograniczenia kosztów podjęto próby testowania zdolności morfogenetycznych eksplantatów pochodzących z E.milii. Wybrane eksplantaty (blaszki liściowe oraz fragmenty pędów) były umieszczane na pożywkach indukujących MS (Murashige i Skoog, 1962) w różnych kombinacjach auksyn i cytokinin. Zaobserwowaną odpowiedzią morfogenetyczną była obfita indukcja kalusa na przeważającej ilości eksplantatów. Innym rodzajem odpowiedzi morfogenetycznej była ryzogeneza indukowana wyłącznie z blaszek liściowych. Proces ten został otrzymany na 90% eksplatatach formujących średnio 32,2 korzenie na eksplantat. Otrzymana tkanka kalusowa, jak i korzenie włośnikowe mogłyby posłużyć do prowadzenia roślinnych kultur in vitro w celu pozyskiwania dużych ilości lateksu na szeroką skalę. Uprzednio należałoby przeanalizować obecność oraz skład chemiczny lateksu zawartego w korzeniach i tkance kalusowej wyprodukowanych w hodowli in vitro za pomocą odpowiednich metod, co stanowi przedmiot kontynuacji badań.
Słowa kluczowe: Euphorbia milii, Euphorbiaceae, Lateks, Ryzogeneza, Schistosomatoza

Induction of high rhizogenesis of selected explants *Euphorbia milii* Des Moul in *in vitro* cultures as alternative sources of latex

Abstract
Schistosomosis (bilharzia) next to malaria is the most serious parasitic disease in the world. It belongs to the so-called. Neglected tropical diseases (NTD), where it was ranked first. It is estimate 258 million cases, 91% of which occur in Africa (WHO, 2014). Schistosomatosis induces parasitic flukes of the genus Schistosoma, whose intermediate hosts are freshwater snails of the genera Bulinus, Biomphalaria, Oncomelania and the ultimate hosts are vertebrates in this man. The World Health Organization has identified many ways to combat schistosomiasis. One is the Snail Control Program designed to eliminate hosts of intermediate pathogenic flukes, including endemic areas for this disease. So far, it is mainly used a synthetic agent that has shown toxic effects on some non-target aquatic organisms, so it is important to find an alternative substancje, which turned out to be latex derived from (*Euphorbia milii* Des Moul.). This milk juice is produced and stored in special cells, either individually or forming characteristic milk tubes, termed "laticifers". The milk ducts in these plants form a branching system that reaches the roots, shoots and leaves. Many studies are currently underway on the impact of this natural, biodegraded substance compared to synthetic drugs. In order to eliminate the discrepancies in the chemical composition and effectiveness of the latex obtained and to reduce costs, attempts were made to test the morphogenetic capacity of explants from E.milii. Selected explants (leaf blades and shoot fragments) were placed on MS induction media (Murashige and Skoog, 1962) in various combinations of auxin and cytokines. The observed morphogenetic response was the generous induction of the callus on the predominant number of explants. Another type of response was was rhizogenesis induced exclusively from leaf blades. This process was obtained on 90% of explorations forming an average of 32.2 roots per explant. Obtained callus tissue and root canal roots could be used to drive in vitro plant cultures to obtain large amounts of latex on a wide rock. Previously, it is necessary to analyze the presence and chemical composition of latex contained in roots and callus tissue produced in in vitro culture by appropriate methods, which is the subject of follow-up studies.
Keywords: Euphorbia Milii, Euphorbiaceae, Latex, Rhisogenesis, Schistosomiasis,

Indeks Autorów

Brachaczek W. 18

Ferenc D. 49

Foit K. 113

Franz M. 138

Górniak K. 192

Jarosz K. 129

Krawiec A. 30

Löschner P. 129

Mazur M. 169

Mazurkiewicz T. 49

Niesiobędzka K. 184

Noga B. 169

Okulewicz J. 65

Paczkowska M. 102

Skiba E. 184

Skibińska M. 157

Stasiak P. 30

Tuleja M. 192

Wichowska A. 169

Zawartka M. 80

Zdeb K. 7

Żukowski P. 65