



ZESZYTY NAUKOWE WYDZIAŁU

ELEKTRONIKI I INFORMATYKI

POLITECHNIKI KOSZALIŃSKIEJ NR **15**



000101101010

POLITECHNIKA KOSZALIŃSKA

**Zeszyty Naukowe
Wydziału Elektroniki i Informatyki**

Nr 15

KOSZALIN 2019

Zeszyty Naukowe Wydziału Elektroniki i Informatyki Nr 15

ISSN 1897-7421

Przewodniczący Uczelnianej Rady Wydawniczej
Zbigniew Danielewicz

Przewodniczący Komitetu Redakcyjnego
Aleksy Patryn

Komitet Redakcyjny
Krzysztof Bzdrya
Walery Susłow
Wiesław Madej
Józef Drabarek
Adam Słowik

Strona internetowa
<https://weii.tu.koszalin.pl/nauka/zeszyty-naukowe>

Projekt okładki
Tadeusz Walczak

Skład, łamanie
Maciej Bączek

© Copyright by Wydawnictwo Uczelniane Politechniki Koszalińskiej
Koszalin 2019

Wydawnictwo Uczelniane Politechniki Koszalińskiej
75-620 Koszalin, ul. Raławicka 15-17

Koszalin 2019, wyd. I, ark. wyd. 4,74, format B-5, nakład 100 egz.
Druk: INTRO-DRUK, Koszalin

Spis treści

<i>Grzegorz Radzki, Grzegorz Bocewicz, Henryk Budzisz</i>	5
Problem marszrutyzacji floty dronów	
<i>Daniel Czyczyn-Egird</i>	23
Prognozowanie defektów w oprogramowaniu z wykorzystaniem modeli predykcyjnych opartych na danych historycznych	
<i>B. Andriyevsky, A. Kashuba, H. Ilchuk</i>	35
Effect of crystal defects on the electronic structure and dielectric functions of $\text{In}_{0.5}\text{Tl}_{0.5}\text{I}$ solid state solutions	
<i>Marek Gąsiorowski, Aleksy Patryn, Leszek Bychto</i>	57
Możliwości oraz obszar zastosowań małowabarytowego spektrometru DLP NIRScan Nano do błyskawicznych pomiarów spektralnych	
<i>Mirosław Maliński</i>	71
Metody fototermiczne i ich zastosowania - Wykład inauguracyjny roku akademickiego 2019/2020 na Wydziale Elektroniki i Informatyki Politechniki Koszalińskiej	
<i>Patryk Widuliński</i>	83
Badanie wykrywalności anomalii w pliku monitorowanym przez system wykrywania intruzów w zależności od parametrów generacji receptorów	

Grzegorz Radzki
Grzegorz Bocewicz
Henryk Budzisz
Wydział Elektroniki i Informatyki
Politechnika Koszalińska

Problem marszrutyzacji floty dronów

Słowa kluczowe: optymalizacja, drony, marszrutyzacja i harmonogramowanie

1. Wstęp

Od momentu rozpoczęcia świadczenia usług transportowych, przedsiębiorstwa codziennie stają przed wyzwaniem w jaki sposób dostarczyć towary do klientów, aby wkład finansowy był jak najmniejszy, a klienci zadowoleni. Wsparcie decyzji w tym zakresie wielokrotnie sprowadza się do rozwiązania odpowiedniego problemu marszrutyzacji [1].

Zauważalny w ostatnich latach dynamiczny rozwój technologii bezzałogowych statków powietrznych (w skrócie dronów) pozwolił na wykorzystywanie tego typu maszyn do transportu towarów. Bardzo szybko zorientowano się, że dostarczanie towarów do klientów za pośrednictwem dronów niesie ze sobą korzyści w postaci: niewielkiego czasu realizacji zamówienia, łatwości dostarczania produktów (np. brak utrudnień związanych z ruchem ulicznym) oraz niższych kosztów energii. Zauważono również, że takie rozwiązania wpływają korzystnie na środowisko [5]. Zastosowanie dronów do zadań transportowych generuje jednak nowe wyzwania związane z organizacją transportu. Typowymi ograniczeniami wynikającymi z wdrożenia systemów transportu powietrznego bazujących na bezzałogowych statkach powietrznych są: konieczność okresowej (częstej) wymiany baterii, niewielki zasięg limitowany pojemnością baterii, niewielka ładowność, możliwości występowania kolizji między dronami realizującymi swoje misje itp. [6].

W tym ujęciu rozważany problem jest rozszerzeniem klasycznego problemu marszrutyzacji pojazdów [1] o możliwość uwzględniania nierównomiernego zużycia energii wynikającego z panujących w danym momencie warunków pogodowych [1]. W ogólności tego typu problem należy do klasy problemów

NP – trudnych [6]. Opracowany model problemu marszrutyzacji, implementowany w środowiskach programowania z ograniczeniami (IBM ILOG), zakłada, że zużycie energii jest funkcją nieliniową, która zależy od warunków pogodowych, „geometrii” drona [13] oraz przyjętej trajektorii (marszrutu) lotu [10]. Poszukiwane są takie plany lotów (marszrutu i harmonogramy), które gwarantują brak kolizji oraz nie prowadzą do wyczerpania baterii przy jednoczesnej maksymalizacji ilości towarów dostarczanych do klientów.

Struktura artykułu jest następująca. W rozdziale drugim przedstawiono przegląd literatury przedmiotu. Rozdział trzeci zawiera przykład ilustrujący problem marszrutyzacji bezzałogowych statków powietrznych. Rozdziały czwarty i piąty zawierają opracowany model oraz sformułowanie problemu. W szóstym rozdziale opisane zostały eksperymenty weryfikujące poprawność opracowanego modelu, wykonane w środowisku programowania deklaratywnego IBM ILOG. Eksperymenty ilustrują możliwości praktycznego wykorzystania proponowanego rozwiązania. Ważniejsze wnioski i kierunki przyszłych badań sformułowano w rozdziale siódmym.

2. Przegląd literaturowy

Problem transportu towarów za pośrednictwem dronów sprowadza się do znanego w literaturze problemu marszrutowania pojazdów – VRP (*ang. Vehicle Routing Problem*) [1]. W najprostszej wersji VRP zakłada, że dostarczany jest jeden rodzaj towaru z wykorzystaniem jednorodnych pojazdów. Poszukiwane są rozwiązania minimalizujące całkowity koszt podróży [2]. Rozszerzenie problemu o założenie, że towary powinny być dostarczone do klientów w określonych przedziałach czasu prowadzi do tzw. Problemu marszrutyzacji z oknami czasowymi – VRPTW (*ang. Vehicle Routing Problem with Time Window*) [1]. Tego typu problemy są najczęściej spotykane w praktyce. Problemy marszrutyzacji bezzałogowych statków powietrznych klasy VRPTW znajdują swoje zastosowania w zarówno rozwiązaniach militarnych jak i cywilnych [1, 3, 4].

Według literatury [6, 14, 15] problem wyznaczania tras dla bezzałogowych statków powietrznych modelowany jest zwykle jako standardowy problem VRP z dodatkowymi ograniczeniami odzwierciedlającymi specyfikę zastosowania. Pierwsze prace w tym zakresie dotyczyły opracowania metod pozwalających na uzyskanie akceptowalnych rozwiązań w zadanym przedziale czasu. Ze względu na specyfikę zagadnienia wykorzystania dronów w problemie VRP oprócz minimalizacji kosztu podróży można wyodrębnić inne kryteria związane z: zmniejszeniem indywidualnych kosztów pracy drona (zużycie baterii), skrócenie czasu operacji, zwiększenie bezpieczeństwa operacji [7]. Kolejnym aspektem odróżniającym rozważany problem od standardowej wersji VRP jest środowisko, w którym planowane są marszrutu. Przy transporcie z wykorzystaniem pojazdów

lądowych lub statków morskich trasy wyznaczone są w przestrzeni 2D [8]. Środowisko pracy bezałogowego statku powietrznego należy do przestrzeni trójwymiarowej, stąd też, wyznaczanie tras dla tego typu obiektów należy rozpatrywać w przestrzeni 3D [9]. W pracy [10] przedstawiony został model matematyczny opisujący wyznaczanie marszrut bezałogowych statków powietrznych w przestrzeni 3D. Należy zwrócić uwagę, że spotykane w literaturze podejścia koncentrują się głównie na wyznaczaniu marszrut dronów bez uwzględniania zmienności warunków pogodowych oraz wynikającego z nich nieliniowego zużycia energii [11]. Stosowanie w praktyce takich rozwiązań wymusza, w trakcie lotu, ciągłe dostosowywanie wcześniej wyznaczonych marszrut. Proces ten jest bardzo złożony i nie zabezpiecza planu misji przed jej ewentualnym niepowodzeniem np. spowodowanym wyczerpaniem baterii [6].

Alternatywą dla takiego postępowania może być podejście umożliwiające uwzględnianie niepewności/zmienności warunków pogodowych i wynikającej z niej poziomu zużycia energii, już na etapie planowania misji lotów. Tego typu problem stanowi rozszerzenie VRPTW o elementy związane z zużyciem energii zależnej od warunków pogodowych oraz przyjętej marszruty lotu [10]. W pracach [12, 13] zostały zaproponowane wstępne heurystyki umożliwiające wyznaczenie tego typu rozwiązań. Przedstawiona praca stanowi rozszerzenie prowadzonych badań o możliwość wykorzystania środowisk programowania deklaratywnego (IBM ILOG).

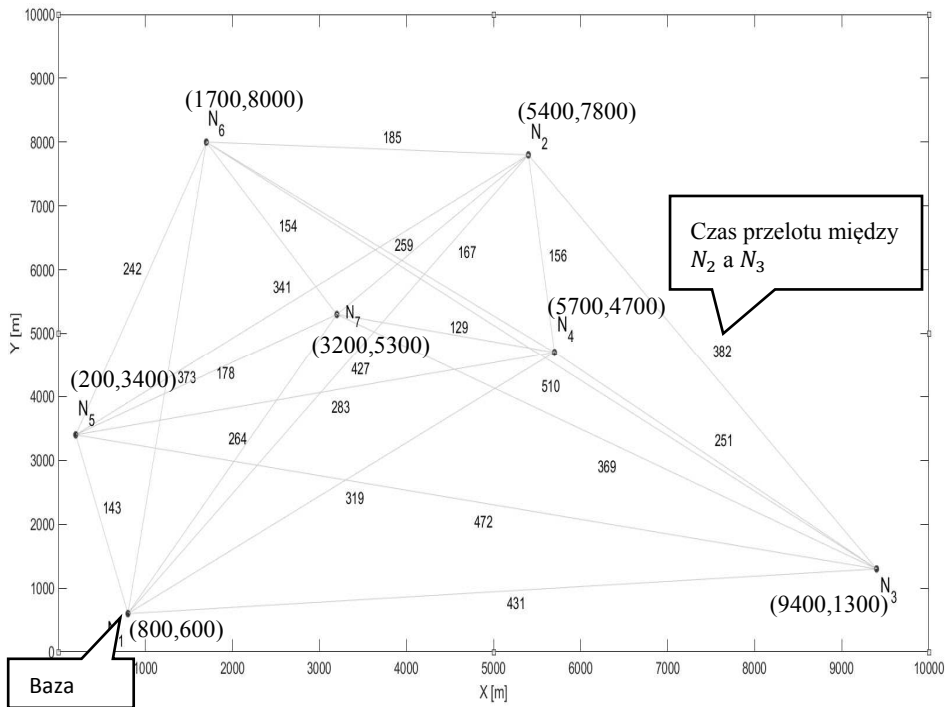
3. Przykład motywacyjny

Rozważany problem można scharakteryzować następująco. Dane jest przedsiębiorstwo, które prowadzi usługi transportu powietrznego za pośrednictwem floty dronów. W rozważanym przypadku w skład floty wchodzi trzy drony o jednakowych parametrach technicznych – tab.1. Drony dostarczają towar do 6 klientów rozlokowanych na powierzchni 10 km^2 – sieć połączeń została zilustrowana na rys. 1. Wierzchołek N_1 określa pozycję przedsiębiorstwa (bazy, z której startują/lądują drony), wierzchołki $N_2 - N_7$ określają pozycję poszczególnych klientów. Przyjęto, że drony przemieszczają się ze stałą prędkością $v = 20 \frac{\text{m}}{\text{s}}$. Wynikające z tego założenia czasy przelotów wzdłuż krawędzi $\{N_i, N_j\}$ grafu zostały przedstawione na rys. 1. Przykładowo czas przelotu pomiędzy węzłem N_6 , a węzłem N_2 wynosi 185 sekund.

Znane są zapotrzebowania poszczególnych klientów na transportowany towar – tab. 2. Zakłada się, drony muszą dostarczyć dokładnie taką ilość towaru jaka jest wymagana przez danego klienta.

Tab. 1. Parametry techniczne drona

Parametry techniczne drona	Wartość	Jednostka
Udźwign Q	100	kg
Pojemność baterii CAP	6000	kJ
Prędkość przelotu v	20	m/s
Aerodynamiczny współczynnik oporu C_D	0,54	-
Powierzchnia czołowa drona A	1,2	m
Szerokość drona b	8,7	m



Rys. 1. Mapa pozycji odbiorców

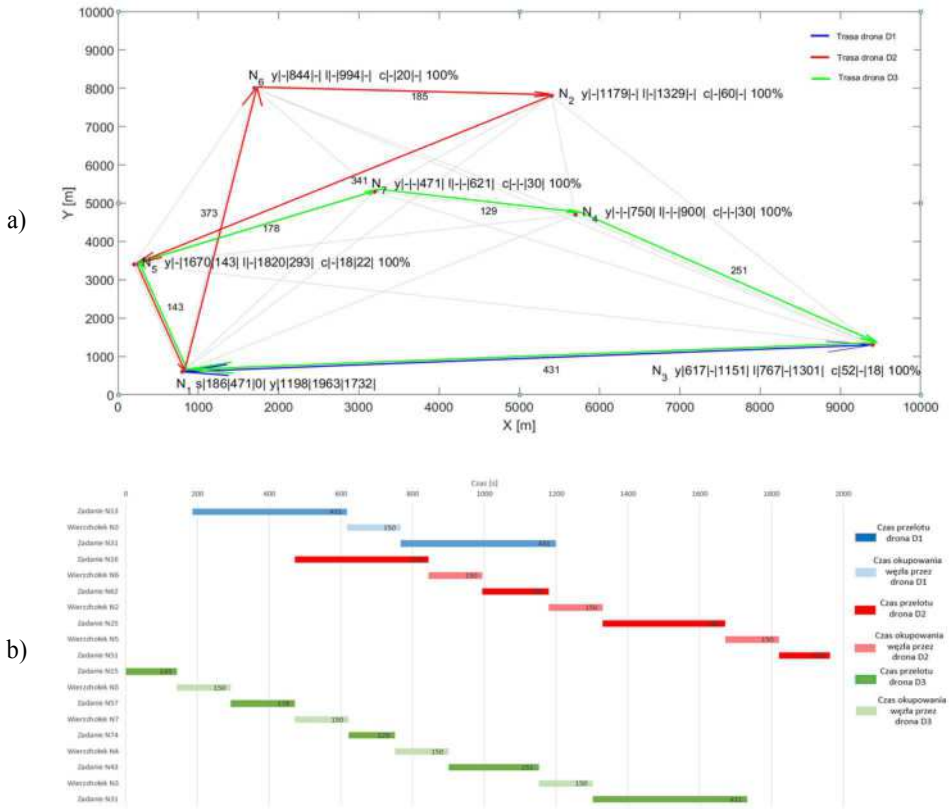
Tab. 2. Zapotrzebowanie klientów na transportowany towar

Węzeł	Zapotrzebowanie na towar [kg]
N1	0
N2	60
N3	70
N4	30
N5	40
N6	20
N7	30

Tab. 3. Warunki pogodowe

Warunki pogodowe	Wartość	Jednostka
Prędkość wiatru v_w	10	m/s
Kierunek wiatru θ	30	°
Gęstość powietrza D	1,225	kg/m ³

Dodatkowo zakłada się, że transport towarów odbywa się w różnych warunkach pogodowych, które wpływają na tempo rozładowania baterii (uwzględnia się prędkość i kierunek wiatru – tab. 3). Zakłada się, że każdy z dronów musi utrzymać zadaną stałą prędkość lotu ($v = 20 \frac{m}{s}$) bez względu na panujące warunki pogodowe. Oznacza to, że przyjęcie stałych, niezmiennych w czasie, planów dostaw (obejmujących marszruty dronów i harmonogram lotów) może skutkować różnym stopniem wykorzystania baterii, a w szczególnych przypadkach ich całkowitym wyczerpaniem. Na rys. 2 przedstawiono przykład marszrut i harmonogramu lotów dronów gwarantujących dostarczenie do klientówżądanego towaru (100% satysfakcji klientów) w zadanych warunkach pogodowych (tab. 3). Jak łatwo zauważyć w proponowanym rozwiązaniu towar zostanie dostarczony w czasie nie przekraczającym 2000 s.

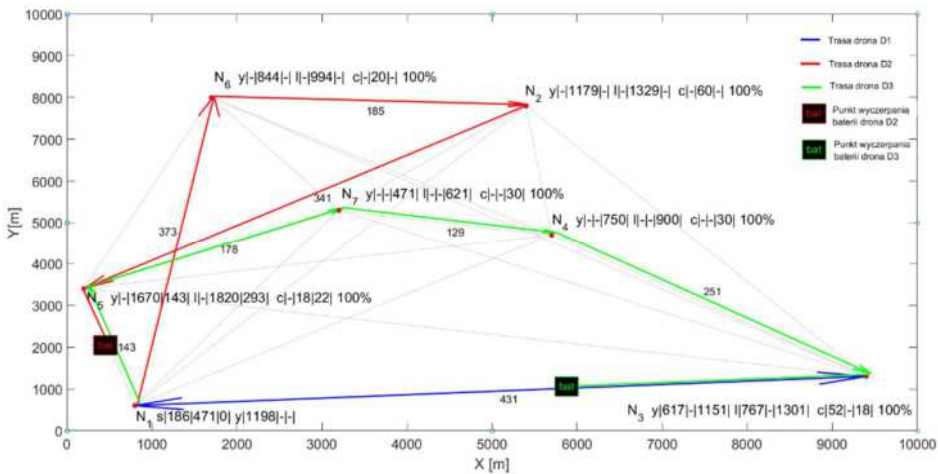


Rys. 2. Marszruty a) i harmonogram lotu b) floty dronów które gwarantują dostarczenie do klientów żądanego towaru (warunki pogodowe: prędkość wiatru = 10 m/s, kierunek wiatru = 30°)

W przyjętym rozwiązaniu drony przemieszczają się odpowiednio wzdłuż marszrut $\Pi_1 = (N_1, N_3, N_1)$, $\Pi_2 = (N_1, N_6, N_2, N_5, N_1)$, $\Pi_3 = (N_1, N_5, N_7, N_4, N_3, N_1)$. Przykładowo dron D_1 (kolor niebieski) przemieszcza się z bazy N_1 do węzła N_3 , gdzie dostarcza 52 kg towaru, a następnie wraca do bazy. Dron numer D_2 (kolor czerwony) przemieszcza się z bazy N_1 przez wierzchołki N_6 , N_2 , N_5 po czym wraca do bazy N_1 , zostawiając kolejno 20,60 i 18 kg towaru. Dron D_3 (kolor zielony) dostarcza towar do klientów w węzłach: N_5 - 22 kg, N_7 - 30 kg, N_4 - 30 kg, N_3 - 18 kg, a następnie wraca do bazy.

Taka organizacja planu lotów nie prowadzi do kolizji (drony nie wykorzystują współdzielonych krawędzi jednocześnie -np. krawędź łącząca wierzchołki N_3 i N_1) oraz gwarantuje dodatni poziom baterii na koniec misji. Stopień wykorzystania baterii dla każdego z dronów wynosi odpowiednio 77,58%, 88,93% oraz 97,53%.

Przedstawione rozwiązanie nie jest jednak dopuszczalne w każdych warunkach pogodowych. Przykładowo gdy kierunek wiatru ulegnie zmianie z 30° na 60° , a prędkość wiatru zmieni się z $10 \frac{m}{s}$ na $12 \frac{m}{s}$ zastosowanie planu lotów z rys. 2 doprowadzi do całkowitego wyczerpania baterii drona D_2 jak i D_3 zanim jeszcze powrócą one do bazy. Sytuację tę ilustruje rys. 3. Jak łatwo zauważyć D_2 zakończy swój lot w trakcie przelotu lotu z wierzchołka N_5 do N_1 , zaś dron D_3 w trakcie przelotu z wierzchołka N_3 do N_1 . W takiej sytuacji należy poszukiwać innych, dedykowanych dla zadanych warunków pogodowych, planów lotów.



Rys. 3. Wyniki symulacji, gdy: prędkość wiatru = 12 m/s, kierunek wiatru = 60°

Biorąc pod uwagę, że przyjęte warunki pogodowe mogą wpływać na możliwość realizacji danego planu lotów rozważany problem sprowadza się do poszukiwania odpowiedzi na następujące pytanie:

Czy zadana flota dronów umożliwia zaspokojenie potrzeb klientów (dostarczenie żądanych ilości towarów) w zadanej sieci transportowej w określonych warunkach pogodowych ?

Inaczej mówiąc rozważany problem sprowadza się do problemu marszrutyzacji i harmonogramowania floty dronów dla zadanych zmiennych warunków pogodowych (prędkość i kierunek wiatru). Poszukiwane są rozwiązania maksymalizujące satysfakcje klientów (funkcja określająca stopień zaspokojenia potrzeb klientów) w zadanych warunkach pogodowych i ograniczonych poziomach baterii.

4. Model deklaracyjny

W rozważanym problemie zakłada się, że znana jest struktura sieci dystrybucji dostaw (liczba, lokalizacja i zapotrzebowanie klientów). Transport towaru odbywa się za pomocą bezałogowych statków powietrznych. Znany jest również przedział czasu tzw. horyzont, w którym powinny być zakończone wszystkie loty. W tym kontekście uwzględnia się następujące założenia:

- Warunki pogodowe są znane (prędkość v_w i kierunek wiatru θ).
- Warunki pogodowe są niezmiennie w całym horyzoncie czasowym.
- Wszystkie drony znajdują się w bazie przed rozpoczęciem misji transportu towaru do klientów.
- Do klientów dostarczany jest ten sam rodzaj ładunku.
- Podczas lotu masa całkowita drona jest stała (tzn. nie uwzględnia się zmniejszenia masy na wskutek pozostawienia części ładunku).
- Drony przemieszczają się ze stałą prędkością $v = 20 \frac{m}{s}$.
- Poszukiwane są plany lotów, które nie powodują kolizji.
- Poszukiwane są marszruty gwarantujące najwyższy poziom satysfakcji klienta.

Opracowany model definiowany jest następująco:

Parametry

Sieć

I – liczba węzłów

$t_{i,j}$ – czas przelotu pomiędzy węzłami N_i , a N_j

m_i – zapotrzebowanie na towar w i -tym węźle $i = 1..I$, $m_1 = 0$

w_i – priorytet i -tego węzła $i = 1..I$, $w_1 = 0$

TN – czas okupowania węzła przez drona (czas rozładowania towaru)

TS – odstęp czasu z jakim drony mogą startować z bazy

$block_{\{i,j\};\{a,b\}}$ – zmienna binarna odpowiadająca przecinającym się krawędziom

$$block_{\{i,j\};\{a,b\}} = \begin{cases} 1 & \text{gdy krawędzie } \{i,j\} \text{ i } \{a,b\} \text{ się przecinają} \\ 0 & \text{w innych przypadkach} \end{cases}$$

Parametry techniczne floty dronów

K – liczebność floty dronów

Q – maksymalna ładowność drona

CAP – pojemność baterii drona

$e_{i,j}$ – zużycie energii przez drona podczas lotu z węzła N_i do N_j

C_D – aerodynamiczny współczynnik oporu

A – powierzchnia czołowa drona

b – szerokość drona

W – masa całkowita drona

$va_{i,j}$ – prędkość generowanego ciągu silnika

Parametry środowiska

H – horyzont planowania $H = [0, t_{max}]$

D – gęstość powietrza

v_w – prędkość wiatru

θ – kierunek wiatru

Zmienne decyzyjne

$x_{i,j}^k$ - zmienna binarna opisująca, czy k -ty dron przemieszcza się z węzła N_i do węzła N_j

$$x_{i,j}^k \begin{cases} 1 & \text{gdy } k - \text{ty dron przemieszcza się z węzła } N_i \text{ do węzła } N_j \\ 0 & \text{w innych przypadkach} \end{cases}$$

s^k – moment startu k -tego drona

y_i^k – moment, w którym k -ty dron przybywa do węzła N_i

c_i^k – masa ładunku dostarczonego do węzła N_i przez k -tego drona

cp_i – łączna masa ładunku dostarczona do węzła N_i

bat^k – poziom baterii k -tego drona

Zbiory

Y^k – zbiór momentów y_i^k - harmonogram k -tego drona

Y – rodzina Y^k – harmonogram floty dronów

C^k – zbiór c_i^k – masa ładunku dostarczona przez k -tego drona

C – rodzina C^k

Π – zbiór marszrut floty dronów

Ograniczenia

Wyznaczenie marszrut. Relacje między zmiennymi opisującymi czasy rozpoczęcia oraz kolejność zadań.

$$s^k \geq 0, k = 1 \dots K \tag{1}$$

$$(k \neq q) \Rightarrow (|s^k - s^q| \geq TS), k, q = 1 \dots K \tag{2}$$

$$\sum_{j=1}^I x_{1,j}^k = 1, k = 1 \dots K \quad (3)$$

$$(x_{1,j}^k = 1) \Rightarrow y_j^k = s^k + t_{1,j}, j = 1 \dots I, k = 1 \dots K \quad (4)$$

$$(k \neq q \wedge y_i^k \neq 0 \wedge y_i^q \neq 0) \Rightarrow (|y_i^k - y_i^q| \geq TN), i = 1 \dots I, k, q = 1 \dots K \quad (5)$$

$$(x_{i,j}^k = 1) \Rightarrow y_j^k = y_i^k + t_{i,j} + TN, j = 1 \dots I, i = 2 \dots I, k = 1 \dots K \quad (6)$$

$$y_i^k \geq 0, i = 1 \dots I, k = 1 \dots K \quad (7)$$

$$\sum_{j=1}^I x_{i,j}^k = \sum_{j=1}^I x_{j,i}^k, i = 1 \dots I, k = 1 \dots K \quad (8)$$

$$y_i^k \leq H * \sum_{j=1}^I x_{i,j}^k \quad \forall i \in I, \forall k \in K \quad (9)$$

Unikanie kolizji. Przecinające się krawędzie ($b_{\{i,j\};\{a,b\}} = 1$) nie mogą być zajęte w tym samym czasie przez drony ($x_{i,j}^k = 1, x_{i,j}^q = 1$)

$$\begin{aligned} & (\text{block}_{\{i,j\}\{a,b\}} \wedge x_{i,j}^k = 1 \wedge x_{a,b}^q = 1) \\ & \Rightarrow (y_b^q \leq y_j^k - t_{i,j}) \vee (y_j^k \leq y_b^q - t_{a,b}) \\ & i, j = 1 \dots I, k, q = 1 \dots K, k \neq q \end{aligned} \quad (10)$$

Dostarczanie ładunków. Relacje między zmiennymi opisującymi ilość dostarczonego ładunku do węzłów przez drony, a zapotrzebowaniem na towar w danym węźle.

$$c_i^k \geq 0, i = 1 \dots I, k = 1 \dots K \quad (11)$$

$$c_i^k \leq Q * \sum_{j=1}^I x_{i,j}^k, i = 1 \dots I, k = 1 \dots K \quad (12)$$

$$\sum_{i=1}^I c_i^k \leq Q, k = 1 \dots K \quad (13)$$

$$(x_{i,j}^k = 1) \Rightarrow c_j^k \geq 1, k = 1 \dots K, i = 1 \dots I, j = 2 \dots I \quad (14)$$

$$x_{i,i}^k = 0, i = 1 \dots I, k = 1 \dots K \quad (15)$$

$$cp_i \leq m_i, i = 1 \dots I \quad (16)$$

$$\sum_{k=1}^K c_i^k = cp_i, i = 1 \dots I \quad (17)$$

$$\sum_{j=1}^I x_{1,j}^k = 1, k = 1 \dots K \quad (18)$$

Zużycie baterii. Ilość potrzebnej energii do wykonania zadania nie może przekraczać maksymalnej pojemności baterii drona.

$$bat^k \leq CAP, k = 1 \dots K \quad (19)$$

$$\sum_{i=0}^I \sum_{j=0}^I x_{i,j}^k * t_{i,j} * e_{i,j}, k = 1 \dots K \quad (20)$$

$$e_{i,j} = \frac{1}{2} C_D A D (v a_{i,j})^3 + \frac{W^2}{D b^2 v a_{i,j}}, i, j = 1 \dots I \quad (21)$$

$$v a_{i,j} = \sqrt{(v \cos \theta - v_w \cos \theta)^2 + (v \sin \theta - v_w \sin \theta)^2}, \quad (22)$$

$$i, j = 1 \dots I$$

Funkcja celu. Maksymalizacja zadowolenia klienta. Zadowolenie klienta wyrażone jest w postaci sumy iloczynu zmiennej w_i oraz cp_i . Satisfakcję klienta należy rozumieć jako stosunek ilości dostarczonego towaru do jego zapotrzebowania w danym węźle wyrażoną w procentach.

$$CS = \max \sum_{i=0}^I w_i * cp_i \quad (23)$$

5. Sformułowanie problemu

W kontekście wprowadzonego modelu rozważany problem wiąże się z odpowiedzią na następujące pytanie:

Czy dla floty K-dronów, przemieszczających się w sieci transportowej składającej się z I wierzchołków, istnieją takie marszruty Π , które gwarantują maksymalny poziom satysfakcji klienta CS przy spełnieniu ograniczeń związanych z unikaniem kolizji (10), dostarczeniem wymaganej ilości ładunku (11) - (18) oraz zużyciem energii (19)-(22)?

Badany problem można rozpatrywać jako optymalizacyjny problem z ograniczeniami (COP, ang. *Constraint Optimization Problem*) [12]:

$$COP = (V, D, C, CS) \quad (24)$$

gdzie: $\mathcal{V} = \{\Pi, Y, C\}$ – zbiór zmiennych decyzyjnych: Π – zmienne opisujące marszruty dronów, Y – zmienne opisujące harmonogram zadań floty dronów, C – zmienne opisujące ilości towarów dostarczanych przez drony.

\mathcal{D} – skończony zbiór opisów zmiennych decyzyjnych,

\mathcal{C} – zestaw ograniczeń opisujących relację między trasami, harmonogramem dronów oraz transportowanymi ładunkami (1)-(22),

CS – funkcja celu reprezentująca poziom satysfakcji klienta.

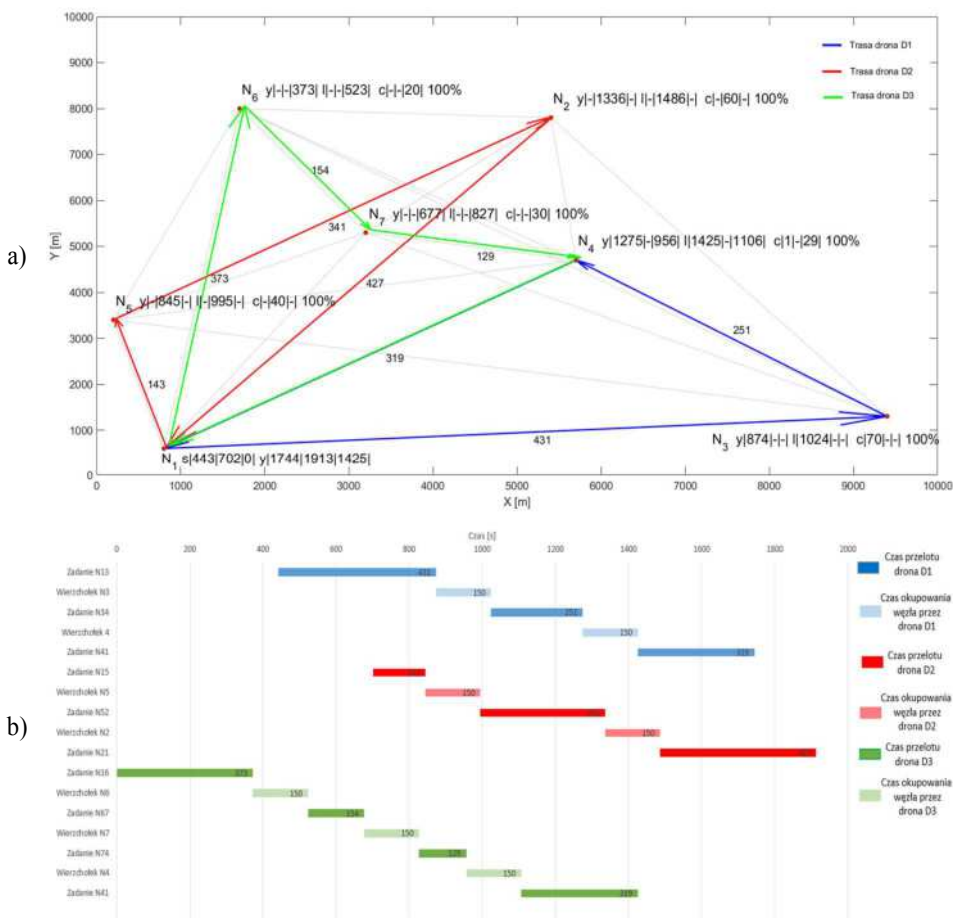
Rozwiązywanie problemu optymalizacji z ograniczeniami polega na określeniu takich wartości zmiennych decyzyjnych, dla których spełnione są dane ograniczenia, a funkcja celu osiąga maksimum. Implementacja COP w środowisku programowania z ograniczeniami np. IBM ILOG pozwala na zbudowanie silnika obliczeniowego, który może być wykorzystany w systemach wspomagania decyzji (DSS).

6. Eksperymenty obliczeniowe

W rozdziale 3. zauważono, że dla rozważanej sieci dostaw marszruta z rys. 2, w przypadku warunków pogodowych: $v_w = 12 \frac{m}{s}$; $\theta = 60^\circ$, nie sprawdza się - skutkuje wyczerpaniem baterii dronów D_2 oraz D_3 . Opracowany model może zostać wykorzystany w próbie wyznaczenia misji lotów gwarantującej powrót wszystkich dronów do bazy przy zadanych warunkach pogodowych, przy jednoczesnej maksymalizacji poziomu zadowolenia klienta. Inaczej mówiąc, poszukiwana jest odpowiedź na następujące pytanie:

Czy dla floty 3-dronów, przemieszczających się w sieci transportowej z rys. 1, istnieją takie marszruty Π , które gwarantują najwyższy poziom satysfakcji klienta w określonych warunkach pogodowych ($v_w = 12 \frac{m}{s}$; $\theta = 60^\circ$)?

Odpowiedź na tak postawione pytanie wiąże się z rozwiązaniem problemu (23). Problem ten został zaimplementowany i rozwiązany w środowisku programowania deklaratywnego IBM ILOG (Intel Core i7-M4800MQ 2.7 GHz, 32 GB RAM). Rozwiązanie otrzymane zostało w czasie 5,14 s. Na rys. 4. przedstawiono otrzymane marszruty i harmonogram lotów dronów. Przyjęcie takiego rozwiązania gwarantuje dostarczenie do klientów żadanego towaru w zadanych warunkach pogodowych. Satysfakcja klientów we wszystkich węzłach wyniosła 100%. Zużycie baterii dla każdego z dronów przy danej marszrucie oraz warunkach pogodowych wynosi odpowiednio: 94,86%, 98,12% oraz 99,5%.



Rys. 4. Marszruty a) i harmonogram lotu b) floty dronów, które gwarantują dostarczenie do klientów żądanego towaru (warunki pogodowe: prędkość wiatru = 12 m/s, kierunek wiatru = 60°)

Wykorzystując opracowany model przeprowadzono serie eksperymentów mających na celu ocenę wpływu pogody na otrzymywane rozwiązania oraz skali problemów, które można efektywnie rozwiązywać (w trybie online: czas obliczeń < 300 s).

W pierwszym przypadku dla zadanej sieci (rys. 1) i stałej floty dronów (3 drony – tab. 1) wyznaczone zostały misje lotów przy zmiennych warunkach pogodowych tj. $v_w = 5 - 15 \frac{m}{s}$; $\theta = 30^\circ - 360^\circ$. Wyniki badań zostały przedstawione w tab. 4 oraz tab. 5.

Tab. 4. Wyniki symulacji przy stałej prędkości wiatru =10 m/s oraz zmiennym kierunku wiatru

Kierunek wiatru [°]	Satysfakcja klienta [%]	Czas symulacji [s]
30	100	6,03
60	100	5,33
90	100	6,39
120	100	6,27
150	100	6,52
180	100	6,82
210	100	5,47
240	100	5,31
270	100	6,32
300	100	5,29
330	100	5,3
360	100	6,33

Tab. 5. Wyniki symulacji przy zmiennej prędkości wiatru oraz stałym kierunku wiatru = 30°

Prędkość wiatru [m/s]	Satysfakcja klienta [%]	Czas symulacji [s]
5	100	5,91
6	100	6,25
7	100	26,83
8	100	6,42
9	100	6,11
11	100	7,62
12	100	5,02
13	83,3	21,73
14	66,6	23,59
15	66,6	12,03

Na podstawie tab. 4 łatwo zauważyć, że przy prędkości wiatru $v_w = 10 \frac{m}{s}$ oraz kierunku zakresie 30°-360° możliwe jest wyznaczenie marszrut gwarantujących satysfakcję klientów na poziomie 100%. W przypadku, gdy kierunek wiatru jest stały, a zmienia się wartość prędkości wiatru zgodnie z tab. 5, to przy prędkościach wiatru większej niż $13 \frac{m}{s}$ znalezienie rozwiązania gwarantującego 100% zadowo-

lenia klientów nie jest już niemożliwe. Innymi słowy zadana flota dronów pozwala na zaspokojenie wszystkich potrzeb klientów tylko w sytuacji gdy prędkość wiatru nie przekracza $12 \frac{m}{s}$.

Kolejnym etapem badań była ocena efektywności opracowanego podejścia. Podczas eksperymentów założono, że maksymalny udźwig drona Q jest zbilansowany względem łącznej masy zapotrzebowania na towar we wszystkich węzłach. Wyniki zostały przedstawione w tab. 6. Jak łatwo zauważyć uzyskanie odpowiedzi na postawione pytanie w czasie < 300 s jest możliwe dla sieci, w których liczba dronów ≤ 4 liczby klientów ≤ 8 .

Tab. 6. Wyniki poszukiwań maksymalnej złożoności sieci, przy prędkości wiatru = 10 m/s oraz kierunku wiatru = 30°

Lp.	Liczba klientów	Liczba dronów	Satysfakcja klientów = 100%	Czas obliczeń [s]
1	6	2	tak	7,46
2	6	3	tak	6,03
3	6	4	tak	10,18
4	6	5	tak	113,93
5	6	6	brak rozwiązania	300
6	7	2	tak	5,6
7	7	3	tak	11,05
8	7	4	tak	187,57
9	7	5	nie	300
10	7	6	nie	300
11	8	2	tak	15,82
12	8	3	nie	28,81
13	8	4	tak	253,22
14	8	5	brak rozwiązania	300
15	8	6	nie	300
16	9	2	nie	300
17	9	3	brak rozwiązania	300
18	9	4	nie	300
19	9	5	brak rozwiązania	300
20	9	6	brak rozwiązania	300

7. Wnioski

Zaproponowany model deklaracyjny (zaimplementowany w środowisku IBM ILOG) pozwala na wyznaczenie marszrut floty dronów gwarantujących maksymalny poziom satysfakcji klienta dla różnych warunków pogodowych. Jak pokazały wyniki eksperymentów, dopuszczalny rozmiar sieci, dla których możliwe jest poszukiwanie tego typu rozwiązań to 8 węzłów przy flocie zawierającej 4 drony. Oznacza to, że zaproponowane podejście może być wykorzystane w metodach zakładających dekompozycje sieci na wiele niewielkich klastrów. W tym ujęciu opracowane podejście może stanowić element heurystyk prezentowanych w pracach [12], [13].

Przyszłe prace będą się koncentrować na rozszerzeniu opracowanego modelu o możliwość uwzględnienia zmiennej masy drona oraz na poszukiwaniu tzw. marszrut „odpornych” na zadane warunki pogodowe.

Bibliografia

1. Abdelhafiz, M., Mostafa, A., Girard, A.: Vehicle Routing Problem Instances: Application to Multi-UAV Mission Planning. In: AIAA Guidance, Navigation, and Control Conference. American Institute of Aeronautics and Astronautics (2010)
2. Golden, B.L., Raghavan, S., Wasil, E.A.: The Vehicle Routing Problem: Latest Advances and New Challenges. Springer Science + Business Media, New York (2010)
3. Yakici, E.: Solving location and routing problem for UAVs. In: Computers & Industrial Engineering. Vol. 102. 294-301 (2016) doi.org/10.1016/j.cie.2016.10.029
4. Ullah, S., Kim, K.I., Kim, K.H., Imran, M., Khan, P., Tovar, E., Ali, F.: UAV enabled healthcare architecture: Issues and challenges. In: Future Generation Computer Systems. Vol. 97. 425-432 (2019). doi.org/10.1016/j.future.2019.01.028
5. Wen-Chyuan, C., Yuyu, L., Shang, Urban, T.L.: Impact of drone delivery on sustainability and cost: Realizing the UAV potential through vehicle routing optimization. In: Applied Energy. Vol. 242. 1164-1175 (2019). doi.org/10.1016/j.apenergy.2019.03.117
6. Bocewicz, G., Nielsen, P., Banaszak, Z., Thibbotuwawa, A.: Routing and scheduling of unmanned aerial vehicles subject to cyclic production flow constraints. In: Proceedings of 15th International Conference on Distributed Computing and Artificial Intelligence. Advances in Intelligent Systems and Computing. Vol. 801. 75-86 (2019), doi.org/10.1007/978-3-319-99608-0_9

7. Enright, J.J., Frazzoli, E., Pavone, M., Savla, K.: Handbook of Unmanned Aerial Vehicles (2015). doi.org/10.1007/978-90-481-9707-1
8. Karpenko, S., Konovalenko, I., Miller., et al.: UAV control on the basis of 3D landmark bearing-only observations. *Sensors (Switzerland)* 15, 29802-29820 (2015), doi.org/10.3390/s151229768
9. Goerzen, C., Kong, Z., Mettler, B.: A survey of motion planning algorithms from the perspective of autonomous UAV guidance. In: *Journall of Intelligent and Robotic Systems: Theory and Application*. Vol. 57. 65-100 (2010). doi.org/10.1007/s10846-009-9383-1
10. Guerriero, F., Surace, R., Loscrí, V., Natalizio, E.: A multi-objective approach for unmanned aerial vehicle routing problrmeith soft time windows constraints. In: *Applied Mathematical Modeling*. Vol 38. 839-852 (2014), doi.org/10.1016/j.apm.2013.07.002
11. Wang, X., Poikonen, S., Golden, B.: The Vehicle Routing Problem with Drones : A Worst-Case Analysis Outline Introduction to VRP Introduction to VRPD, pp. 1–22 (2016)
12. Thibbotuwawa, A., Nielsen, P., Banaszak Z., Bocewicz, G.: Energy Consumption in Un-manned Aerial Vehicles: A Review of Energy Consumption Models and Their Relation to the UAV Routing. In: *Advances in Intelligent Systems and Computing*. Vol. 853. 173–184 (2019). doi.org/10.1007/978-3-319-99996-8_16
13. Thibbotuwawa, A., Nielsen, P., Zbigniew, B., Bocewicz, G.: Factors affecting Energy consumption of unmanned aerial vehicles: an analysis of how energy consumption changes in relation to UAV routing. In: *Advances in Intelligent Systems and Computing*, pp. 228–238. Springer International Publishing (2018)
14. Dai, R., Fotebar, S., Radmanesh, M., Kumar, K.: Quality-aware UAV coverage and path planning in geometrically complex environments. In: *Ad Hoc Network*. Vol.73.95-105 (2018), doi.org/10.1016/j.adhoc.2018.02.008
15. Chauchan, D., Unnikrishnan, A., Figliozzi, M.: Maximum coverage capacitated facility location problem with range constrained drones. In: *TransportationResearch Part C: Emerging Technologies*. Vol. 99. 1-18 (2019). doi.org/10.1016/j.trc.2018.12.001

Streszczenie

W artykule rozważany jest problem dostarczania towarów w sieci dystrybucyjnej, w której operacje transportu realizowane są przez flotę bezałogowych statków powietrznych (BSP). Szczególny nacisk położony jest na możliwość uwzględniania warunków pogodowych, w których odbywa się transport.

Rozważany problem sprowadza się do wyznaczenia marszruty gwarantującej bezkolizyjne dostarczenie towaru do węzła, przy uwzględnieniu warunków pogodowych oraz poziomu baterii dronów. Zaproponowany został deklaracyjny model, którego implementacja w środowisku programowania IBM ILOG pozwala na rozwiązywanie tego typu problemów w trybie on-line. Przedstawione eksperymenty obliczeniowe, potwierdzają poprawność opracowanego modelu.

Abstract

A problem of delivering goods in a distribution network is considered in which, transport operations are carried out by a fleet of unmanned aerial vehicles (UAV). The weather conditions in which a transport operations take place and a UAV battery level are taken into account. The above-mentioned weather conditions and the battery level affect the determination of the route guaranteeing collision-free delivery. The goods must be delivered to a specific node in a given time window. Specifying the route is the focus of this study. Solutions maximizing the level of customer satisfaction are sought. Computational experiments which show impact of the weather conditions to on route determination are presented.

Keywords: optimization, drones, routing and scheduling

Daniel Czyczyn-Egird
Wydział Elektroniki i Informatyki
Politechnika Koszalińska
daniel.czyczyn-egird@dce-systems.pl

Prognozowanie defektów w oprogramowaniu z wykorzystaniem modeli predykcyjnych opartych na danych historycznych

Słowa kluczowe: eksploracja danych w oprogramowaniu, modele predykcji defektów, metryki oprogramowania

1. Wprowadzenie

W dobie nieustannego rozwoju komputerów oraz oprogramowania, istnieje duże zapotrzebowanie na nowe i coraz bardziej zaawansowane systemy informatyczne, od których wymaga się oprócz określonych funkcjonalności, także jak najwyższej niezawodności. Niestety zdecydowana większość oprogramowania obciążona jest defektami, które powodują niestabilne działanie określonych funkcjonalności lub też potrafią być przyczyną wadliwego działania całego systemu.

Defekt pojawia się w oprogramowaniu, kiedy osoba tworząca dany system popełnia błąd, który może się pojawić na różnych etapach tworzenia oprogramowania takich jak analiza wymagań, projektowanie dokumentacji systemowej (projekt ogólny/szczegółowy), plan testów, nieodpowiednie skrypty testowe, kod źródłowy etc.

Zatem ważnym aspektem jest proces testowania, podczas którego tester wykonuje określone przypadki testowe i może obserwować czy wyniki tych testów pokrywają się z oczekiwaniami. Wszelkie odstępstwa od oczekiwań są traktowane jako incydenty, które trzeba zbadać i wyjaśnić. Wszystkie wykryte usterki i problemy powinny być zapisywane w systemach do śledzenia błędów (ang. ITS) i / lub w systemach kontroli wersji (ang. VCS) w celu dalszej analizy i próby znalezienia rozwiązania problemu.

Repozytoria danych, w których przechowywane są powyższe informacje, mogą stanowić ciekawe źródło wiedzy dla badaczy i naukowców, którzy zajmują się problematyką procesów ulepszania oprogramowania (ang. SPI) [1] lub zapewnienia jakości (ang. QA).

Artykuł ma na celu przedstawienie ogólnego podejścia do problemu predykcji defektów w oprogramowaniu, w oparciu o modele operujące na informacjach historycznych z repozytoriów danych oraz dokonując ich analizy.

2. Prace powiązane

Istnieje wiele narzędzi wspierających pracę programistów i testerów w ich codziennej pracy nad systemami informatycznymi. Wśród tych narzędzi są takie, które pozwalają na wygodne i szybkie testowanie tworzonych rozwiązań zarówno na poziomie kodu źródłowego jak i na poziomie postkompilacyjnym. Wyłapywanie błędów w oprogramowaniu to rzecz niezmiernie ważna, aby finalnie dostarczać produkty pozbawione defektów. Jednakże ciągle testowanie i debugowanie systemów wiąże się z poniesieniem wydatków związanych z użyciem zasobów ludzkich (programiści, testerzy) jak i zasobów finansowych [2]. Dlatego też predykcja defektów w oprogramowaniu jest ważna, gdyż ma realne zastosowanie w projektach komercyjnych, powinna pozwalać na wymierne oszczędności. Badania nad predykcją defektów mają zatem coraz to większe zainteresowanie zarówno po stronie praktyków jak i badaczy.

Ramler i Himmelbauer [3] proponują predykcję defektów przy użyciu modeli predykcyjnych powiązanych z systemami oprogramowania na poziomie ich modułów.

Moduły mogą być plikami, klasami, komponentami a także podsystemami danego systemu. Moduły te są opisywane poprzez zestawy atrybutów (np. przez metryki kodu czy liczbę zmian w danej iteracji), które to są dostępne poprzez ekstrakcję ich z różnych źródeł danych takich jak bazy metryk czy repozytoria kodów źródłowych.

Istnieje także wiele prac, w których autorzy oprócz skupiania się wyłącznie na modelach, sporo czasu poświęcają zagadnieniom akwizycji danych z repozytoriów.

Powstało kilka narzędzi przydatnych do zadań związanych z pozyskiwaniem danych oraz wspierających modelowanie predykcyjne (np. predykcję defektów).

Jureczko i Magot [4] przygotowali Quality Spyframework [5] o otwartym kodzie (licencja Apache 2.0 [6]), którego zadaniem było odczytywać i zbierać surowe dane z kodu źródłowego i repozytoriów zdarzeń oraz metryki zdefiniowane przez użytkowników. Projekt skupiał się na dwóch modułach dotyczących pozyskiwania danych oraz raportowania. W ostatniej wydanej wersji framework

pozwalal na odczytywanie metryk z klas dla technologii Java, odczytywanie zdarzeń systemu JIRA [7], a także wpisów z systemu kontroli wersji Subversion (SVN) [8].

D'Ambros i Lanza [9] zaproponowali narzędzie wspierające analizę ewolucji oprogramowania poprzez interfejs sieciowy – Churrasco. Jest to narzędzie o otwartym kodzie źródłowym pobierające i przetwarzające dane z systemu Bugzilla oraz SVN, opierające się na meta-modelu FAMIX, który jest niezależny od zastosowanej technologii programowania. Oprócz tego został wykorzystany obiektowy moduł mapowania relacyjnego (GLORP), moduł ekstrakcji faktów (MOOSE) oraz moduł SVG do wizualizacji.

Madeyski oraz Majchrzak [10] opracowali DefectPrediction for software systems (DePress), specjalny framework, którego celem jest rozszerzalny pomiar oprogramowania oraz integracja danych w celach predykcyjnych (predykcja defektów, predykcja kosztów/nakładów). Framework DePress bazuje na projekcie KNIME [11] i jest jego rozszerzeniem (zestawem wtyczek), pozwalającym budować modele graficzne przepływu danych w graficzny, prosty i przejrzysty dla użytkownika sposób.

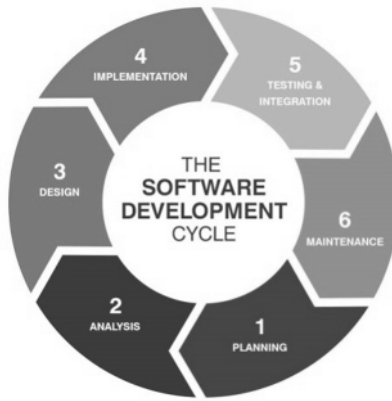
Głównym założeniem projektu była predykcja defektów oprogramowania w oparciu o dane historyczne. W tym celu przygotowano zestaw wtyczek odpowiedzialnych kolejno za kolekcjonowanie, transformację oraz analizę danych. Autorzy skupili się na operacjach akwizycji i transformacji danych oraz raportowaniu, jednocześnie zostawiając operacje statystyczne oraz eksplorację danych [12] sprawdzonemu środowisku KNIME, które posiada odpowiednie wbudowane mechanizmy.

Ważnym aspektem, na który autorzy postawili jest także nacisk na archiwizację oraz udostępnianie zbiorów danych na zewnątrz (np. dla innych badaczy, którzy chcieli by testować własne rozwiązania). W tego typu badaniach ważne jest, aby repozytoria danych historycznych oraz przygotowane modele predykcyjne, były w miarę możliwości publiczne (po wcześniejszym procesie anonimizacji treści chronionych komercyjnym prawem autorskim).

3. Prognozowanie defektów w oparciu o dane historyczne

3.1. Predykcja defektów a cykl życia oprogramowania

Defekt może zostać wprowadzony na dowolnym etapie procesu zwanego SLDC (ang. Software Development Life Cycle) [13] dlatego bardzo ważne jest, aby testerzy byli zaangażowani od początku cyklu życia oprogramowania, po to aby wykrywać i usuwać wady.



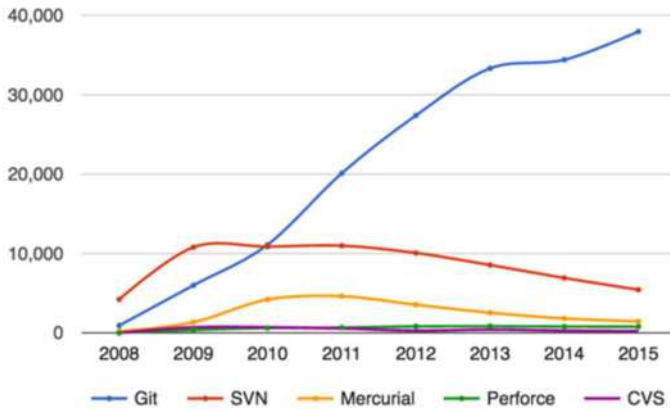
Rys. 1. Diagram cyklu życia procesów wytwarzania oprogramowania

Im szybciej dany defekt zostanie zlokalizowany i naprawiony, tym samym koszt utrzymania jakości będzie mniejszy. Przykładowo jeżeli defekt jest zidentyfikowany w fazie analizy wymagań, wtedy koszt naprawy sprowadza się do zmodyfikowania wymagań na odpowiednim dokumencie. Jednakże jeżeli, wymagania zostaną źle opisane i zaimplementowane, a defekt zostanie wykryty dopiero podczas fazy testowania, wtedy koszt naprawy będzie bardzo wysoki i będzie wiązał się z poprawą wymagań i specyfikacji oraz zmianą w implementacji. Będzie wymagał także dalszego procesu testowania.

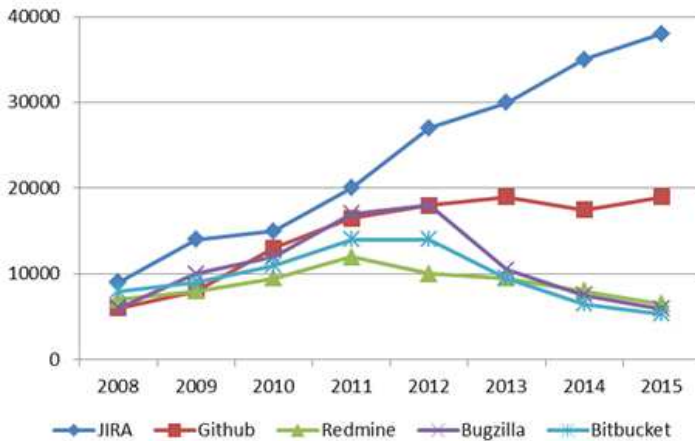
W niniejszym artykule autor skupia się na fazie implementacji i testowania, gdyż z poziomu tych faz możliwe jest uzyskanie odpowiednich danych historycznych z systemów zarządzania konfiguracją oprogramowania oraz systemów śledzenia błędów (oczywiście o ile te są przechowywane i pielęgnowane).

3.2. Główne założenia

Jednym z głównych założeń operacji predykcji defektów jest określenie źródeł danych historycznych, na podstawie których będzie odbywał się cały proces predykcyjny. Interesujące nas dane można uzyskać z systemów zarządzania konfiguracją oprogramowania oraz systemów śledzenia błędów. Obecnie na rynku istnieje wiele typów owych systemów, gdzie jednymi z najbardziej popularnych oraz używanych w profesjonalnych zespołach programistycznych są odpowiednio dla systemów kontroli wersji (pierwsza piątka): Git, Subversion, Mercurial, Perforce, CVS, a dla systemów śledzenia błędów (pierwsza piątka): JIRA, Github, Redmine, Bugzilla, BitBucket. Poniższe diagramy przedstawiają procentowy udział wspomnianych narzędzi (badanie przeprowadzi na podstawie zbadania popularności zapytań według słów kluczowych na portalu dla programistów – stackoverflow.com) [14]:



Rys. 2. Popularność systemów kontroli wersji (VCS)



Rys. 3. Popularność systemów do śledzenia błędów (ITS)

Akwizycja danych z wyżej wymienionych systemów może odbywać się na dwa sposoby: bezpośredni lub pośredni. Pierwszy z nich pozwala na dostęp do repozytoriów systemowych najczęściej poprzez podłączenie się do ich bazy danych lub też za pomocą odpowiednich mechanizmów, które pozwalają na odczyt informacji z tych baz.

W przypadku systemu JIRA, oprócz sparametryzowanych kwerend SQL bezpośrednio wykonywanych na odpowiednich tabelach w bazie danych (silnikiem

bazodanowym obsługującym system JIRA jest MSSQL), dostępne jest jeszcze API, które za pomocą odpowiednich żądań, zwraca oczekiwane wyniki. JIRA oferuje także swój własny mikro-język JIRA Query Language (JQL). Jest to najbardziej elastyczny sposób wyszukiwania danych w JIRA i jest dla wszystkich: programistów, testerów, kierowników projektów, a nawet nietechnicznych użytkowników biznesowych. Ta metoda może być dedykowana dla tych, którzy nie mają doświadczenia z zapytaniami do baz danych, a także dla tych, którzy chcą szybszego dostępu do informacji w JIRA.

Drugi sposób dostępu do danych opiera się na plikach wymiany, które są eksportowane ręcznie z systemów przez odpowiednie interfejsy np. do formatu CSV lub XML. Analiza wyeksportowanych plików może być najprostszym sposobem dostępu do danych oraz często jedynym, jeżeli np. w powodu braku uprawnień nie mamy dostępu bezpośredniego do systemu lub/i jego bazy danych.

W przypadku systemów kontroli wersja, tak samo do wyboru są dwie takie same opcje. Dla przykładu dla Gita lub Subversion, także możemy spróbować podłączyć się do ich baz (plikowych) oraz wyszukiwać interesujące nas artefakty lub też za pomocą odpowiednich narzędzi wyeksportować dane do plików wymiany.

Pobrane dane historyczne powinny być w pewien sposób ujednoczone oraz oczyszczone z niepotrzebnych pól (np. z informacji o autorach danego wpisu), następnie przetransformowane do wspólnego formatu np. formy tabelarycznej, która zostanie w następnych etapach poddana procesom eksploracji.

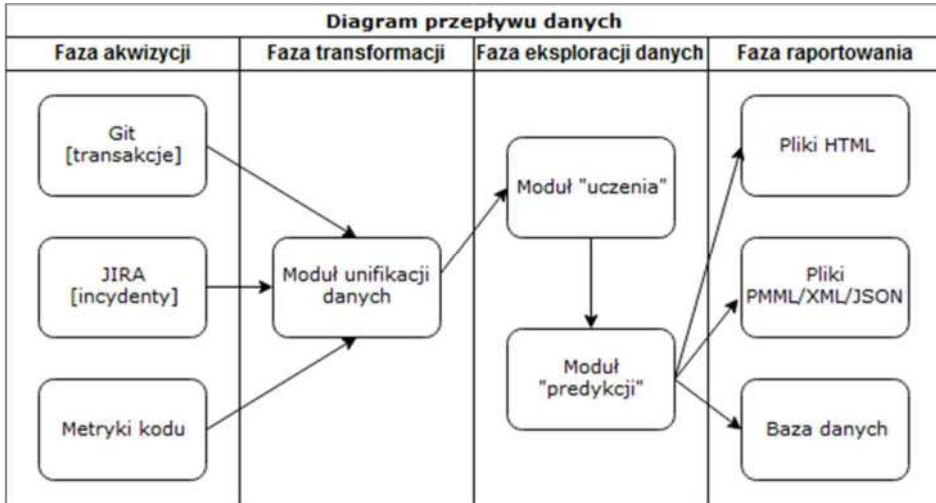
Do pobranych danych historycznych należy dołączyć także metryki kodu źródłowego. Metryki te powinny zostać odpowiednio powiązane z danymi historycznymi, w celu umożliwienia budowy klasyfikatorów predykcyjnych. Bez dołączenia stosownych metryk (np. liczba linii kodu w pliku, liczba klas w pliku, stopień zagnieżdżenia klas), nie było by możliwe zastosowanie modelu predykcyjnego w nowotworzonych systemach, gdzie dane historyczne nie występują.

3.3. Przykładowe podejście do tworzenia modelu predykcji

Jednym ze sposobów budowy modelu predykcyjnego, jest podejście dwu-etapowe, gdzie w pierwszym etapie budujemy oraz trenujemy nasz model, a w drugim etapie jest on używany do predykcji na nowych danych wejściowych. W każdym z dwóch etapów można wyróżnić kilka następujących faz:

1. Faza akwizycji danych – procesy pobierania danych.
2. Faza transformacji danych – procesy ujednoczania i dopasowania danych wejściowych.
3. Faza eksploracji danych – procesy związane z odkrywaniem nowej wiedzy.
4. Faza raportowania – procesy przedstawiania wyników oraz ich archiwizacja.

Z punktu widzenia badawczego, istotnym elementem do dalszych badań, ale także najtrudniejszym, jest dobór odpowiednich klasyfikatorów w fazie trzeciej – w tym celu stosowane są różne techniki z obszarów związanych np. z algorytmami ewolucyjnym, inteligencji roju [15] czy też logiką rozmytą.



Rys. 4. Diagram przepływu danych w procesach uczenia i predykcji

4. Studium przypadku

4.1. Wybór projektu do analizy

Dla poniższego opracowania, jednym z głównych kryteriów wyboru oprogramowania do analizy, był swobodny dostęp do danych archiwalnych zawartych w systemach kontroli wersji oraz w systemach śledzenia błędów. Takie kryterium spełniają projekty o otwartym kodzie źródłowym – wybór padł na system Apache Netbeans [16]. Jest to zintegrowane środowisko programistyczne (IDE) dla języka Java, którego głównym celem jest przyspieszenie budowy aplikacji Java, w tym również usług sieciowych oraz aplikacji mobilnych. Oprócz standardowej funkcjonalności, możliwe jest również rozszerzenie narzędzia o między innymi wsparcie dla języków programowania C i C++, wsparcie dla tworzenia aplikacji w architekturze SOA, użycia XML i schematów XML, BPEL i Java Web Services czy modelowania UML. Tak jak pozostałe produkty fundacji Apache –NetBeans jest rozprowadzany na licencji Apache License. Narzędzie to zostało wydane po raz pierwszy wiosną 1999 roku i jest cały czas nieustannie rozwijane, a jego ostatnia wydana wersja pochodzi z lipca 2019 roku (wersja o numerze 11.1 dla Javy 11). NetBeans jest w całej swojej historii rozwijany przez dużą grupę programistów – zwolenników otwartego oprogramowania (ponad 100 osób zaangażowanych

w projekt, na przestrzeni kilkunastu lat, ostatnie wydania wspierane przez 50 stałych programistów). Fakt tak aktywnego rozwoju wybranego narzędzia niewątpliwie jest spowodowany brakiem ograniczeń w rozwoju tego oprogramowania wraz z nielimitowanym dostępem do repozytoriów kodów źródłowych, zgodnie z polityką licencyjną fundacji Apache.

Repozytoria kodów źródłowych są przechowywane w systemie Git [17], natomiast wykryte defekty są raportowane w systemie JIRA.

4.2. Akwizycja danych

Kody źródłowe narzędzia NetBeans są magazynowane w systemie Git, ostatnie czasy pokazują, że twórcy oprogramowania powiązani z fundacją Apache, migrują swoje projekty do systemu Git (np. z SVN) jako zyskującego coraz większą popularność. Zatem wybór padł na pobranie danych z systemu Git – w tym celu zostały użyte odpowiednie narzędzia między innymi GraphQL API v4 [18] dla Githuba oraz zestaw skryptów typu bash operujących na systemie Git. Zakres dat dla opisywanych działań to: styczeń 2018 – grudzień 2018.

Tabela 1. Główne informacje z systemu Git

Zakres czasu	01.01.2018 – 31.12.2018
Wydanie	9.0
Transakcje (commity)	560
Zmiany w plikach	2554
Liczba akt. programistów	118

Następnym krokiem była akwizycja danych z systemu JIRA dla zadanego okresu czasu. Z repozytorium JIRA, brane były pod uwagę wszystkie te zdarzenia, które były opisane jako: Bug o statusie Closed/Resolved, wyniku końcowym Fixed/Duplicated. Pominięto parametr priorytetu potrzeby naprawy defektu, w celu uzyskania szerszego zakresu informacji o zdarzeniach.

Tabela 2. Główne informacje z systemu JIRA

Zakres czasu	01.01.2018 – 31.12.2018
Release	9.0
No. of issues	312
No. of issues (Bug)	128
Improvementtasks	60

W następnym etapie podjęta jest próba sparowania artefaktów pobranych z systemu Git z artefaktami pobranymi z systemu JIRA. Takie parowanie jest możliwe, jeżeli przykładowo w opisach transakcji (Git) występują opisy rozwiązywanych problemów zgłoszonych w systemie JIRA. Przykładowo należy poszukać (np. z użyciem wyrażeń regularnych) ciągu z numerem danego zdarzenia np. „ResolvedissueNETBEANS-953”, gdzie po stronie JIRA zdarzenie te opisane jest: „(NETBEANS-953) Profile jdk10 remote platform NullPointerException”.

Ostatnim etapem akwizycji danych jest pobranie metryk kodu źródłowego. Można tego dokonać przy pomocy narzędzi dedykowanych pod daną technologię programowania, w przypadku NetBeansa dla Javy, można użyć biblioteki JavaNCSS.

Tabela 3. Wybrane metryki z JavaNCSS

Moduł	Pakiety	Klasy	Metody
org-apache-tools-ant-module	33	45	56
org-netbeans-api-annotations-common	5	17	32
org-netbeans-api-debugger	12	77	152
org-netbeans-api-debugger-jpda	11	34	100
org-netbeans-api-intent	19	39	89
org-netbeans-api-io	20	62	122
org-netbeans-api-progress	5	17	77
org-netbeans-api-search	8	19	39
org-openide-loaders	13	37	99
org-openide-modules	14	29	101
org-netbeans-core-multiview	8	15	45
org-netbeans-core-network	14	29	120
org-netbeans-lib-v8debug	10	19	22
org-netbeans-modules-db	15	49	107

4.3. Transformacja danych

Podczas procesu transformacji i normalizacji, dane (kolumny) zbędne z punktu widzenia predykcji, są usuwane lub wartości im przypisane ustawiane są na zero. Każde artefakty, które zostały pobrane w poprzednim etapie i ze sobą powiązane, otrzymują opis właściwości: HasDefects oraz NoOfDefects z odpowiednimi wartościami opisującymi czy dla danej klasy/metody z kodu źródłowego, występują jakieś defekty, które zostały zgłoszone w systemie JIRA i naprawione w kolejnej transakcji do systemu kontroli wersji (Git).

4.4. Proces predykcji oraz wyniki

Metodologia predykcji powinna opierać się na minimum trzech etapach tj. wyborze atrybutów do predykcji, budowie modelu oraz jego walidacji. W pierwszym etapie wybieramy atrybuty, na których będzie opierał się budowany model. Stosowanie wszystkich atrybutów (danych historycznych oraz metryk) może prowadzić do przeładowania modelu, dlatego zaleca się wybrać tylko te wyróżniające się (można skorzystać np. z eliminacji wstecznej). Kolejny etap to budowa modelu, który może opierać się na jednym klasyfikatorze lub na kilku powiązanych ze sobą, tworząc tzw. klasyfikator hybrydowy. Popularnym klasyfikatorem stosowanym do wszelakich predykcji są drzewa decyzyjne. Dla zadanego przypadku atrybutem predykcyjnym jest HasDefects, który mówi czy w danym zestawie wejściowym znaleziony jest defekt. Zmiennymi niezależnymi są w tym przypadku atrybuty wejściowe wybrane w pierwszym etapie. Walidacja modelu jako ostatni etap, może opierać się przykładowo na walidacji krzyżowej (np. K-fold cross validation), gdzie dane wejściowe dzielone są na dwa segmenty: uczący i walidujący, a cały proces jest powtarzany kilkakrotnie z losowym podziałem (innym za każdym razem). O ile wyniki pracy modelu będą zadowalające, można go próbować zastosować do kolejnych wydań projektowych, zwłaszcza w fazie tworzenia kodu oraz testowania.

Wszystkie operacje związane z modelowaniem oraz testowaniem można przeprowadzić np. w środowisku badawczym KNIME, które ma pewien zestaw modułów wspierających techniki tworzenia modeli predykcyjnych. Pozwala także na zapisanie wyników do bazy danych, czy też eksport do formatu XML jako model PMML (przydatne do dalszych badań).

5. Podsumowanie i przyszłe prace

Artykuł ma na celu przedstawienie ogólnego podejścia do problemu predykcji defektów w oprogramowaniu, w oparciu o modele predykcyjne oparte na atrybutach historycznych uzyskiwanych z systemów kontroli wersji, systemów śledzenia błędów oraz metryk, uzyskiwanych z poziomu kodu źródłowego.

Został przedstawiony ogólny mechanizm procesu predykcji defektów opierający się na analizie wybranych repozytoriów danych projektu Apache NetBeans. Narzędzie to charakteryzuje się długoletnią obecnością na rynku informatycznym (od około 1999 roku), zatem systemy kontroli wersji jak i systemy śledzenia błędów zawierają sporą ilość interesujących danych do eksploracji. Na ich podstawie można zbudować model predykcyjny, który po odpowiedniej walidacji będzie w stanie przewidywać wystąpienia defektów w przyszłości w zadanym projekcie. Co przeloży się na wymierne korzyści w postaci oszczędności czasu (testowanie) i środków finansowych (poprawki).

Temat dotyczący skutecznej predykcji defektów jest tematem bardzo rozwojowym i porusza realne, współczesne problemy rynku informatycznego, dlatego też przewidziane są dalsze prace nad predykcją defektów. Przyszłe plany prac powinny skupiać się na możliwościach własnej, konkurencyjnej implementacji narzędzi wspierających akwizycję, normalizację oraz modelowanie predykcyjne. Oprócz tego istnieje potrzeba budowania bardziej skutecznych modeli predykcyjnych, w tym celu dalsze zainteresowania mogą skupiać się przykładowo na hybrydowych klasyfikatorach predykcyjnych.

Bibliografia

1. Petersen K., Wohlin C.: Software process improvement through the Lean Measurement (LEAM) method, *Journal of Systems and Software*, vol. 83, no. 7, pp. 1275-1287, 2010.
2. Wojszczyk R.: Quality Assessment of Implementation of Strategy Design Pattern. *Advances in Intelligent Systems and Computing*, Springer , Vol. 620, pp. 37 - 44, 2018.
3. Ramler R., Himmelbauer J.: Building Defect Prediction Models in Practice, *Handbook of Research on Emerging Advancements in Software Engineering*, pp. 540-565, 2014.
4. Jureczko M., Magott J.: QualitySpy: a framework for monitoring software development processes, *Journal of Theoretical and Applied Computer Science*, vol. 6, no. 1, pp. 35-45, 2012.
5. Jureczko M. and Contributors: "Quality Spy.", <http://java.net/projects/qualityspy>.
6. The Apache Software Foundation, "Apache License, Version 2.0." <http://www.apache.org/licenses/LICENSE-2.0.html>.
7. JIRA, Atlassian, <https://www.atlassian.com/software/jira>.
8. SVN, "Enterprise-class centralized version control ", <https://subversion.apache.org/>
9. D'Ambros M., Lanza M., Distributed and Collaborative Software Evolution Analysis with Churrasco, *Sci. Comput. Program.*, vol. 75, pp. 276-287, Apr. 2010.
10. Madeyski L., Majchrzak M.: Software Measurement and Defect Prediction with Depress Extensible Framework, *Foundations of Computing and Decision Sciences*, vol. 39, no. 4, 2014.
11. Berthold M. R., Cebon N., Dill F., Gabriel T. R., Meinel T., Ohl P., Sieb C., Thiel K., and Wiswedel B.: KNIME: The Konstanz Information Miner, *Studies in Classification, Data Analysis, and Knowledge Organization (GfKL 2007)*, Springer, 2007.

12. Czyczyn-Egird D., Wojszczyk R.: The effectiveness of data mining techniques in the detection of DDoS attacks. *Distributed Computing and Artificial Intelligence*, 14th International Conference, Springer, Vol. 620, pp. 53-60, 2018.
13. Kazim A., A Study of Software Development Life Cycle Process Models, *International Journal of Advanced Research in Computer Science*, Volume 8, No. 1, 2017.
14. Version Control Systems Popularity in 2016, <https://rhodecode.com/insights/version-control-systems-2016> (access on 08.2019).
15. Slowik A., Kwasnicka H., Nature Inspired Methods and Their Industry Applications – Swarm Intelligence Algorithms, *IEEE Transactions on Industrial Informatics*, Volume 14, Issue 3, pp. 1004-1015, March 2018.
16. Apache NetBeans, <https://netbeans.apache.org/>
17. GitHub Inc., <http://www.github.com>.
18. GraphQL API v4, <https://developer.github.com/v4/>

Streszczenie

W dzisiejszych czasach istnieje wiele metod i dobrych praktyk w inżynierii oprogramowania, które mają na celu zapewnienie wysokiej jakości tworzonego oprogramowania. Jednakże pomimo starań twórców oprogramowania, często w projektach występują defekty, których usuwanie wiąże się często z dużym nakładem finansowym oraz nakładem czasu. Artykuł prezentuje przykładowe podejście do predykcji defektów w projektach informatycznych opierając się na modelach predykcyjnych zbudowanych w oparciu o informacje historyczne oraz metryki produktu, zebrane z różnych repozytoriów danych.

Abstract

Nowadays, there are many methods and good practices in software engineering that are aimed at providing high quality of created software. However, despite the efforts of software developers, there are often defects in projects, the removal of which is often associated with a large financial and time expenditure. The article presents an example approach to defect prediction in IT projects based on predictive models based on historical information and product metrics, collected from various data repositories.

Keywords: data mining, defect prediction models, software metrics

B. Andriyevsky ¹

A. Kashuba ²

H. Ilchuk ²

¹Faculty of Electronics and Computer Sciences
Koszalin University of Technology,
Śniadeckich str. 2, PL-75-453, Koszalin, Poland

²Lviv Polytechnic National University,
Bandera str. 12, 79646 Lviv, Ukraine
bohdan.andriyevskyy@tu.koszalin.pl

Effect of crystal defects on the electronic structure and dielectric functions of $\text{In}_{0.5}\text{Tl}_{0.5}\text{I}$ solid state solutions

Keywords: semiconductors, point defects, electronic band structure, electrical transport, optical properties, computer simulations

1. Introduction

Searching the new functional materials with tunable physical properties is of key interest for current micro and optoelectronics. Over the past few decades the attention of scientists is being focused on the searching new materials with possibilities the controlled tuning of the physical properties. In particular, modifications of initial material for achieve the expected properties suitable for potential applications can be performed in many ways. One of the popular and effective methods is a gradual substitution one of the components by another belonging to the same group of the periodic system of elements. Such a procedure introduces internal stress into the already known crystals (mostly caused by mismatch of ionic radiuses) and creates solid solutions based on a pristine crystal [1, 2]. So far, we have successfully tuned properties of several crystals by changing composition and utilizing structure-properties relations for such modified compounds, using both experimental methods and computer simulations [3 - 5].

As well known, during crystal growth processes, all kinds of defects always appear: twinning's, interstitials, vacancies or atomic positions exchange [6].

Structural defects substantially affect the electronic structure of a material, cause changes of the optical properties, transport phenomena and finally influence its structural stability. In many cases, the controlled level of certain defects may positively influence the expected properties of materials, improving desired features such as non-linear optical, thermoelectric, etc. Often, significant differences between the experimentally measured and computationally simulated data are caused by not only the limitations of the applied theoretical models or computational techniques (e.g. the well-known underestimation of the band gap obtained with using the density functional theory) but because of the disregard the ubiquitous defects existing in real crystals. In the present study, the influence of crystal structure deviations of the ternary crystals $\text{In}_x\text{Tl}_{1-x}\text{I}$ and its solid solution from the perfect one on the corresponding electron related properties was computationally investigated by using the density functional theory (DFT) approach. Comparison of the computational results obtained in the present investigation with our previous experimental and theoretical studies of the same materials confirms the effectiveness of the structure defects models adopted. The method proposed in this work can be extended to modeling and prediction the properties of new crystalline materials or to verify compliance the previous calculations made with the assumption of a perfect structure and corresponding experimental data.

2. Materials: $\text{In}_x\text{Tl}_{1-x}\text{I}$

Thallium, embedded in a stable crystal matrix, having saturated bonds, is not as dangerous as a free atom or ion, while the chemical compounds that contain it possess extremely interesting optical properties [3,4] and are excellent thermoelectric materials [7]. Among the mixed semiconductors of the $\text{A}^{\text{III}}\text{V}^{\text{VII}}$ group the three-component crystals $\text{In}_x\text{Tl}_{1-x}\text{I}$, representing a continuous series of solid solutions of substitution (SSS), may be promising for applications in optoelectronics in spite of investigations these SSSs many years ago [8 - 10]. Their band gap E_g vary within the range 2.01 - 2.84 eV [8 - 11]. The crystals possess a layered structure. However, in contrast to typical layered crystals of the $\text{A}^{\text{II}}\text{B}^{\text{VI}}$ and $\text{A}^{\text{III}}\text{B}^{\text{VI}}$ (CdI_2 , HgI_2 , GaSe , InSe) groups [12, 13], in which the van der Waals interlayer gaps are formed mainly by the corresponding anions, in $\text{In}_x\text{Tl}_{1-x}\text{I}$ SSS, this weakest chemical bonding is probably created by indium and thallium cations and iodine anions. Investigations of the crystals $\text{In}_x\text{Tl}_{1-x}\text{I}$ is of interest due to the possibility of samples growth combining the properties of both InI and TlI . The application prospect of these compounds is associated with their use as materials in the ionizing radiation detectors and optical modulators of CO_2 laser [14]. The crystal structure of TlI was firstly investigated at ambient temperature [15]: $Z = 4$, orthorhombic, space group $Cmcm$, $a = 4.57$, $b = 12.92$, $c = 5.24$ Å. In turn, the crystal structure of InI has been determined in works [16, 17]: InI , structure type TlI , $Z = 4$, orthorhombic,

space group No. 63 (*Cmcm*), $a = 4.763$, $b = 12.781$, $c = 4.909$ Å. The space group of symmetry for InI crystal is the same as for TlI one.

Thallium iodide crystallizes in a rhombic structure at the temperatures up to 178 °C, but at higher temperatures, it transforms into the CsCl structure of the cubic symmetry. Information on the corresponding structural phase transition is presented in work [18]. In halides of indium, no phase transitions occur except in InCl. The crystals InI and TlI cleave perpendicularly to the crystallographic b -axis [8 - 13]. Smooth controlled changes of the energy band gap E_g , mechanical and photovoltaic characteristics and spectral range of the recombination radiation may be realized in them. Results of investigation of the band structure and photoconductivity with change of the indium vs thallium content of the crystals $\text{In}_x\text{Tl}_{1-x}\text{I}$ are presented in the recent study [11].

So far, studies of the phonon spectra, mechanical and optical characteristics are known for $\text{In}_x\text{Tl}_{1-x}\text{I}$ SSS [19 - 22]. In Ref. [19], a possibility of forming the InI structure as a matrix for the TlI quantum dots inclusion is discussed. The reduction of the crystal unit cell dimensions b and c and the unit cell volume V and the increase of dimension a together with the increase of content index x are the main features indicating the continuous structure changes of $\text{In}_x\text{Tl}_{1-x}\text{I}$ [11] (in the previous Refs. [8 - 13] and [17 - 22], the convention $a \approx c < b$ for the unit cell dimension was adopted).

The optical anisotropy, which is expected due to the layer structure, and the continuous structure transformations mentioned above make $\text{In}_x\text{Tl}_{1-x}\text{I}$ SSS perspective for the nonlinear optics and electrical engineering applications. For this purpose, the details of the phenomena associated with such applications of $\text{In}_x\text{Tl}_{1-x}\text{I}$ SSS should be studied. In particular, it was found [11] that the photoconductivity spectra in the ac layer of $\text{In}_x\text{Tl}_{1-x}\text{I}$ SSS correspond to the n - and p -types carriers. It was also suggested that the appearance of the n -type conductivity, occurred primarily in the a - crystallographic direction of $\text{In}_x\text{Tl}_{1-x}\text{I}$, is caused by the anionic vacancy or cationic interstitials. If an alternation of the conductive and insulating sub-layers along the b unit cell dimension of the crystal is possible, one can assume that $\text{In}_x\text{Tl}_{1-x}\text{I}$ SSS may be used as a material for the low-dimensional capacitors. From this point of view, one of the main goals of this work is to study the origin of electric conductivity and the quasi-metallization of $\text{In}_x\text{Tl}_{1-x}\text{I}$ SSS.

If the suggested anionic vacancies or cationic interstitials in $\text{In}_x\text{Tl}_{1-x}\text{I}$ SSS lead to the quasi-metallization, this should be revealed in the band structure and dielectric function of the material. The present study is based on the first principles calculations in the framework of the density functional theory (DFT) of the band structure (BS) and dielectric functions of the nominally pure crystal $\text{In}_{0.5}\text{Tl}_{0.5}\text{I}$ and on the changes of these characteristics caused by the anionic vacancies and cationic interstitials in the compound.

3. Calculations

Calculations of BS and related properties of the materials were carried out using the academic CASTEP code [23] in the framework of DFT approach. The calculations were performed within the generalized gradient approximation (GGA) with the Perdew-Burke-Ernzerhof (PBESOL) exchange-and-correlation functional [24]. The interaction of electrons with atomic cores was described with the norm-conserving pseudopotentials supplied with the academic CASTEP [23].

The electronic wave functions are expanded in a plane wave basis set with the energy cut-off of 650 eV. The atomic levels $4d^{10}5s^25p^1$ for indium, $5d^{10}6s^26p^1$ for thallium and $5s^25p^5$ for iodine atoms are treated as valence electron states. For the Brillouin zone (BZ) sampling, we have used the Monkhorst-Pack K -points meshes [25] of 0.0625 \AA^{-1} . The self-consistent convergence of the total energy was chosen to be $1.0 \times 10^{-6} \text{ eV/atom}$. For DFT calculations of the non-defective and defective $\text{In}_x\text{Tl}_{1-x}\text{I}$ SSS the supercells, representing $2 \times 2 \times 2$ unit cells containing 64 atoms, were created. The crystal symmetry of the optimized non-defective $\text{In}_x\text{Tl}_{1-x}\text{I}$ SSS at $x = 0.5$ is lower (No. 38, *Amm2*) than the space group of symmetry for InI or TlI crystals (No. 63, *Cmcm*). The geometry optimization of lattice parameters and internal atomic coordinates were determined using the Broyden-Fletcher-Goldfarb-Shanno minimization technique (BFGS) with the maximum ionic Hellmann-Feynman forces within 0.02 eV/\AA , the maximum ionic displacement within $1.0 \times 10^{-3} \text{ \AA}$, and the maximum stress within 0.05 GPa.

4. Results and discussion

The crystal structure of $\text{In}_{0.5}\text{Tl}_{0.5}\text{I}$ SSS has been obtained by the replacement of two indium atoms by the thallium ones in the conventional unit cell of InI (In_4I_4). For the purpose of the proper BZ sampling of the crystals studied using the 'SeeK-path' tool [26] the sequence of the Cartesian axes and corresponding unit cell dimensions and atomic coordinates has been changed to satisfy the relation $a(X) < b(Y) < c(Z)$, where $a = 4.763$, $b = 4.909$, $c = 12.781 \text{ \AA}$.

Symmetry requirements of $\text{In}_{0.5}\text{Tl}_{0.5}\text{I}$ crystal (space group No. 38) cause that x - and y -conventional coordinates of each atom remain constant (0.0, 0.25, 0.5 or 0.75) and only the z -coordinates of atoms may change during structure optimization together with the unit cell dimensions a , b and c . We have found, however, that the parameter of the unit cell elongation $2c/(a+b)$ of the optimized $\text{In}_2\text{Tl}_2\text{I}_4$ crystal (2.638) is close to the similar values in In_4I_4 (2.645) and Tl_4I_4 (2.638) ones.

We have studied also the defective $\text{In}_{0.5}\text{Tl}_{0.5}\text{I}$ SSSs: with iodine vacancy (1) and with thallium interstitial atom (2). With this aim, the primitive or conventional unit cells, possessing too small crystal cell dimensions, are not convenient. To prevent the direct interaction between neighboring iodine vacancies or thallium interstitials

the supercell $2 \times 2 \times 2$ of the conventional unit cell of $\text{In}_2\text{Tl}_2\text{J}_4$ with the dimensions $a = 9.010 \text{ \AA}$, $b = 9.284 \text{ \AA}$ and $c = 26.216 \text{ \AA}$, containing 64 atoms ($\text{In}_{16}\text{Tl}_{16}\text{I}_{32}$), has been used for calculations (Figs. 1, 2). Two different localizations of the iodine vacancy I1 and I2 are possible relating to the initial structure $\text{In}_{16}\text{Tl}_{16}\text{I}_{32}$ (Fig. 1). In the first localization I1, the smallest distance I1 - Tl (3.302 \AA) is smaller than the smallest distance I1 - In (3.409 \AA). In turn, in the second localization I2, the smallest distance I2 - In (3.157 \AA) is smaller than the smallest distance I2 - Tl (3.432 \AA). To study the defective $\text{In}_{0.5}\text{Tl}_{0.5}\text{J}$ SSS with thallium interstitial atom, we have used the initial structure $\text{In}_{16}\text{Tl}_{16}\text{I}_{32}$, to which one thallium atom has been added into the supercell center corresponding to the most probable localization of large interstitial atom (Fig. 2). Thus, three crystal structures of the defective $\text{In}_{0.5}\text{Tl}_{0.5}\text{J}$ SSS have been studied: two supercells $2 \times 2 \times 2$ with one I1 or I2 vacancy and one supercell $2 \times 2 \times 2$ with one Tl interstitial. Every of these three defective crystal structures possess the space group of symmetry No. 25 ($Pmm2$).

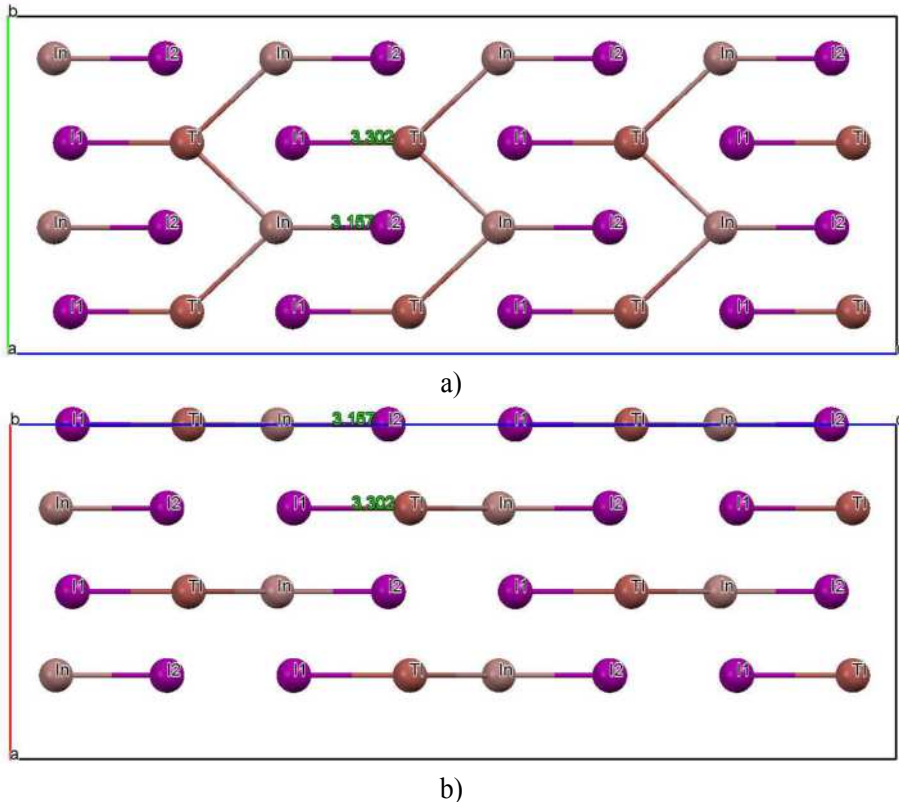


Fig. 1. Views along a - (a) and b -axis (b) of the optimized structure of $\text{In}_{16}\text{Tl}_{16}\text{I}_{32}$ crystal ($a = 9.527 \text{ \AA}$, $b = 9.575 \text{ \AA}$, $c = 25.193 \text{ \AA}$). Interatomic distances Tl - I1 and In - I2 are presented in Angstrom

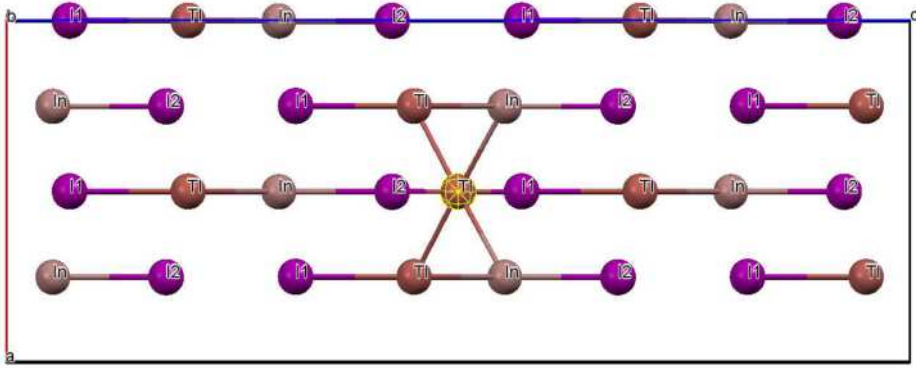


Fig. 2. View along a -axis of the initial structure of $\text{In}_{16}\text{Tl}_{17}\text{I}_{32}$ crystal ($a = 9.527 \text{ \AA}$, $b = 9.575 \text{ \AA}$, $c = 25.193 \text{ \AA}$) with one interstitial thallium atom indicated by yellow net

The CASTEP calculations of BS were carried out at the points of Brillouin zone (BZ) determined using the 'SeeK-path' tool [26]. Because of the different space groups of symmetry in the cases of crystals InI (TII), $\text{In}_{0.5}\text{Tl}_{0.5}\text{I}$ and defective crystals $\text{In}_{0.5}\text{Tl}_{0.5}\text{I}$, the corresponding sets of K -points for the BS calculations are different.

For the cases of the primitive unit cells used, the calculated band gap E_g of InI has been found to be indirect. In turn, the direct band gaps E_{gd} have been detected for $\text{In}_{0.5}\text{Tl}_{0.5}\text{I}$ and TII crystals, which are not located at Γ -point of BZ (Fig. 3). For $\text{In}_{0.5}\text{Tl}_{0.5}\text{I}$ the direct energy band gap E_{gd} takes place at the K -point $-1/2, 1/2, 1/3$ at the boundary of the first BZ, not at the Γ -point $0, 0, 0$ (Fig. 3). The value of band gap found, $E_g = 1.315 \text{ eV}$, may correspond to either direct or indirect electron transitions (Fig. 3). Almost inverse proportional dependence of the band gap $E_g(x)$ on the content parameter x ($\text{In}_x\text{Tl}_{1-x}\text{I}$) can be noticed: $E_g(x=0) = 2.2 \text{ eV}$, $E_g(x=0.5) = 1.3 \text{ eV}$, $E_g(x=1.0) = 0.61 \text{ eV}$ (see Table 1), what is in agreement with previous results [11]. The calculated values of E_g are smaller than the experimentally observed ones, that occurs often for DFT calculations of crystals. A clear dependence between Hirshfeld charges of ions in InI, $\text{In}_{0.5}\text{Tl}_{0.5}\text{I}$ and TII crystals and their band gap values is observed (Table 1). In $\text{In}_{0.5}\text{Tl}_{0.5}\text{I}$ SSS, the smallest interatomic distance $d_{\text{In-I}} = 3.13 \text{ \AA}$ is smaller than the corresponding value in InI (3.22 \AA), and the smallest distance $d_{\text{Tl-I}} = 3.33 \text{ \AA}$ is larger than the similar one in TII (3.23 \AA). This may be explained using the electronegativity values for indium (1.7), thallium (1.8) and iodine (2.5) atoms [27]. The larger the difference of the electronegativity values of two atoms the larger the ionicity of the corresponding chemical bond. In view of the prevailing ionic bonding in the crystals studied, one may expect larger ion-ion interactions In - I in comparison to Tl - I ones, that may lead to the mentioned changes of equilibrium interatomic distances in $\text{In}_{0.5}\text{Tl}_{0.5}\text{I}$ in comparison to InI and TII.

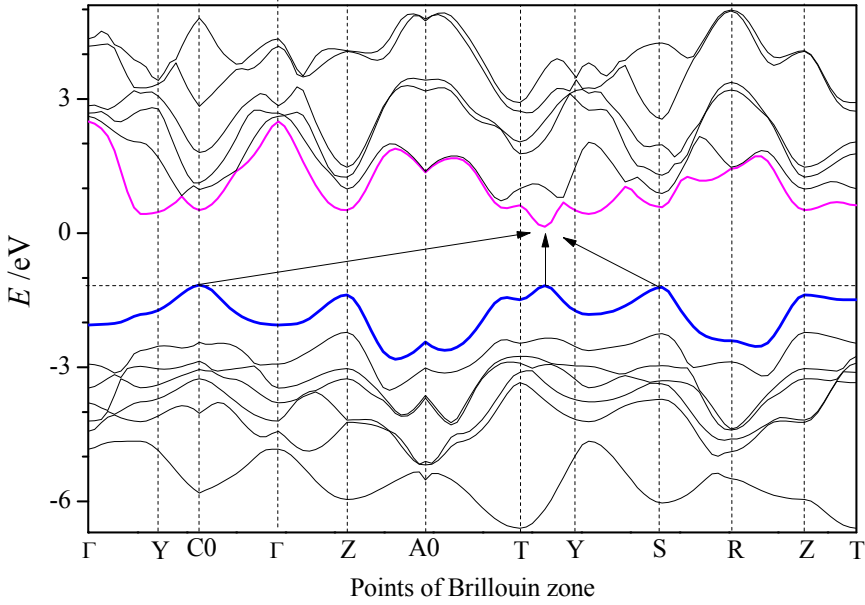


Fig. 3. Band structure of InTl_2 SSS. Top VB and bottom CB are indicated by color lines. Probable transitions corresponding to band gap E_g are shown by arrows. The K -points list used, $\Gamma - 0,0,0$; $Y - -1/2,1/2,0$; $C0 - -0.286,0.714,0$; $\Gamma - 0,0,0$; $Z - 0,0,1/2$; $A0 - 0.286,0.286,1/2$; $T - -1/2,1/2,1/2$; $Y - -1/2,1/2,0$; $S - 0,1/2,0$; $R - 0,1/2,1/2$; $Z - 0,0,1/2$; $T - -1/2,1/2,1/2$ corresponds to the space group of symmetry No. 38 ($Amm2$)

Table 1. The Hirshfeld charges q and smallest interatomic distances d for InI, $\text{In}_{0.5}\text{Tl}_{0.5}\text{I}$ and TlI crystals calculated at primitive unit cell

Crystal	InI	$\text{In}_{0.5}\text{Tl}_{0.5}\text{I}$	TlI
Value			
q_{In1} or q_{Tl1} /e	0.11	0.12 (In)	0.14
q_{In2} or q_{Tl2} /e	0.11	0.13 (Tl)	0.14
q_{I1} /e	-0.11	-0.13	-0.14
q_{I2} /e	-0.11	-0.12	-0.14
$d_{\text{In-I}}$ / \AA	3.22	3.13	-
$d_{\text{Tl-I}}$ / \AA	-	3.33	3.23
E_g /eV	0.61	1.3	2.2

Another characteristic property of the band structure of semiconductors is a tensor of effective electron mass m_{ij}^* [28],

$$\left(m^{*-1}\right)_{ij} = \frac{1}{\hbar^2} \frac{\partial^2 E}{\partial K_i \partial K_j} \quad (1)$$

which is usually presented in terms of the free electron mass m_e . Information on the effective electron mass m^* is important because this parameter determines the dynamics of electron conductivity in material and therefore is significant for the practical applications.

In order to see difference between the effective electron masses in $\text{In}_x\text{Tl}_{1-x}\text{I}$ for $x = 0, 0.5$ and 1.0 at the same K -point set corresponding to the space group of symmetry No. 38, the band structure calculations have been performed at the corresponding primitive unit cells. The effective mass tensor m_{ij}^* was calculated using the Effective Mass Calculator code [29] at the bands extrema using finite difference method with the step value of 0.02 (Bohr^{-1}). The calculated diagonal components m_{ii}^* ($i = 1, 2, 3$) of the effective mass tensor m_{ij}^* ($i, j = 1, 2, 3$) are presented in Table 2 and Table 3. There are no clear dependences of effective mass m^* on the content parameter x . However, by more detailed analysis of the Table 2 and Table 3 one may reveal that for the selected components of the effective mass tensor m_{ij}^* a tendency to decrease the absolute values $|m_{ii}^*|$ with increase of the content parameter x can be noticed. More reliable conclusion relating the dependence of the electron effective mass in $\text{In}_x\text{Tl}_{1-x}\text{I}$ SSS may be derived from the study of the problem for more values of the content parameter x . This, however, was not among our first aims in this work.

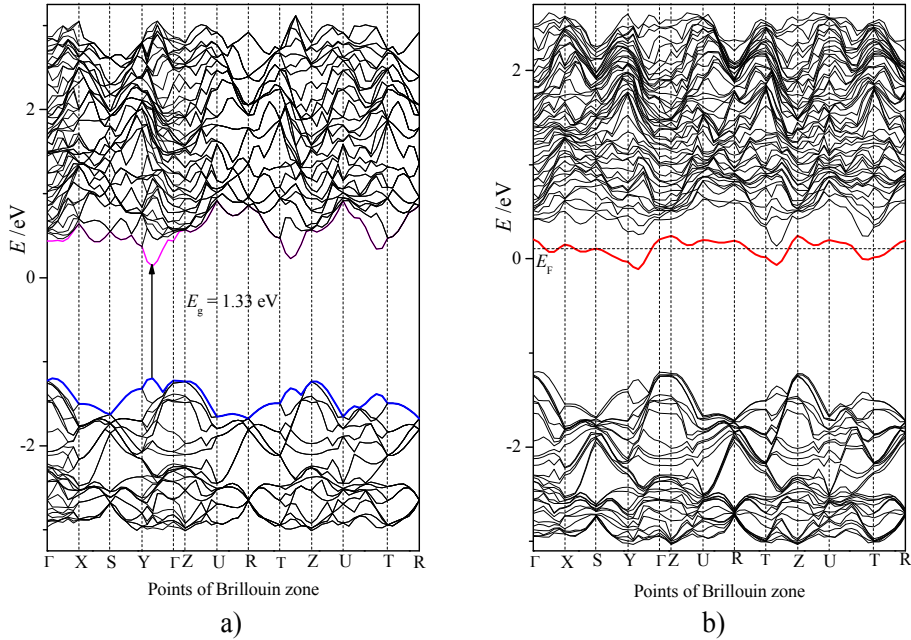
Table 2. Diagonal components m_{ii}^* of the effective mass tensor m_{ij}^* for top VB of $\text{In}_x\text{Tl}_{1-x}\text{I}$ crystals, calculated at selected K -points of BZ of primitive unit cells: C0 – $-0.285, 0.714, 0$; Z – $0, 0, 0.5$; A0 – $-0.285, 0.285, 0.5$; TY – $-0.5, 0.5, 0.2727$; S – $0, 0.5, 0$ (see Fig. 3)

VB	$x = 1.0$			$x = 0.50$			$x = 0$		
	m_{11}^*	m_{22}^*	m_{33}^*	m_{11}^*	m_{22}^*	m_{33}^*	m_{11}^*	m_{22}^*	m_{33}^*
C0	-0.070	-0.294	-0.456	-0.043	-0.177	-0.256	0.296	0.277	0.215
Z	-0.053	-0.453	0.050	-0.050	-0.165	-0.569	-0.065	-0.182	-0.926
A0	0.443	0.159	0.096	0.381	0.205	0.203	-1.850	0.143	0.084
TY	-0.039	-0.079	-0.088	-0.086	-0.090	-0.186	-0.131	-0.224	-13.414
S	-0.025	-0.074	0.259	-0.044	-0.134	0.731	-0.174	-0.328	0.273

Table 3. Diagonal components m_{ii}^* of the effective mass tensor m_{ij}^* for bottom CB of $\text{In}_x\text{Tl}_{1-x}\text{I}$ crystals, calculated at selected K -points of BZ of primitive unit cells: C0 – $-0.285, 0.714, 0$; Z – $0, 0, 0.5$; A0 – $-0.285, 0.285, 0.5$; TY – $-0.5, 0.5, 0.2727$; S – $0, 0.5, 0$ (see Fig. 3)

CB	$x = 1.0$			$x = 0.50$			$x = 0$		
	m_{11}^*	m_{22}^*	m_{33}^*	m_{11}^*	m_{22}^*	m_{33}^*	m_{11}^*	m_{22}^*	m_{33}^*
C0	1.372	0.431	0.067	0.378	0.112	0.032	0.357	0.101	0.037
Z	-0.042	0.176	0.075	0.291	0.264	0.066	-0.131	0.576	0.070
A0	-0.023	-0.909	-1.135	-0.023	-0.673	-2.505	-0.020	-0.695	0.707
TY	0.193	0.063	0.034	0.176	0.073	0.057	1.394	0.645	0.091
S	-0.410	0.121	0.028	-0.418	0.112	0.039	-0.407	0.102	0.052

When the supercell $2 \times 2 \times 2$ of the non-defective $\text{In}_{0.5}\text{Tl}_{0.5}\text{I}$ SSS crystal was used, the direct energy band gap $E_g = 1.332$ eV has been found to be at the BZ point $0, 1/3, 0$ (Fig. 4a). This value of E_g is a bit larger in comparison to that ($E_g = 1.315$ eV) found for the primitive unit cell (Fig. 3). These differences may be due to the inclusion of interactions between more distant atoms in the supercell $2 \times 2 \times 2$ in comparison to the primitive one.



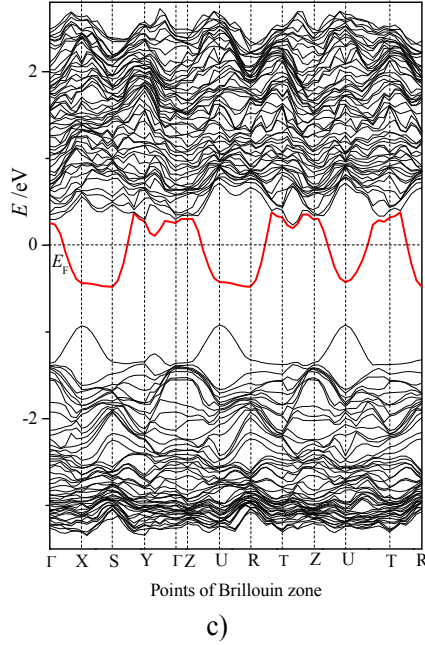


Fig. 4. Band structures of (a) $\text{In}_{16}\text{Tl}_{16}\text{I}_{32}$ crystal (the space group of symmetry No. 38, $Amm2$), (b) $\text{In}_{16}\text{Tl}_{16}\text{I}_{31}$ (the space group of symmetry No. 25, $Pmm2$) with one iodine vacancy of the type II (see Fig. 1) and (c) $\text{In}_{16}\text{Tl}_{17}\text{I}_{32}$ crystal with one interstitial thallium atom (the space group of symmetry No. 25, $Pmm2$). Top VB and bottom CB are indicated at (a) in blue and magenta. The top half-occupied VBs are indicated in (b) and (c) by thick red lines and the Fermi level E_F by horizontal dashed lines. The set of BZ points $\Gamma - 0,0,0$; $X - 1/2,0,0$; $S - 1/2,1/2,0$; $Y - 0,1/2,0$; $Z - 0,0,1/2$; $U - 1/2,0,1/2$; $R - 1/2,1/2,1/2$; $T - 0,1/2,1/2$ corresponds to the space group of symmetry No. 25 ($Pmm2$)

On the basis of the crystal $\text{In}_{16}\text{Tl}_{16}\text{I}_{32}$ described above three defective crystals, the first, $\text{In}_{16}\text{Tl}_{16}\text{I}_{31}$, with iodine vacancy close to thallium atom, the second, $\text{In}_{16}\text{Tl}_{16}\text{I}_{31}$, with iodine vacancy close to indium atom, and the third, $\text{In}_{16}\text{Tl}_{17}\text{I}_{32}$, with interstitial thallium atom have been studied. Band structure of the crystal $\text{In}_{16}\text{Tl}_{16}\text{I}_{31}$ with vacancy defect (Fig. 4b) is similar to that of the non-defected one $\text{In}_{16}\text{Tl}_{16}\text{I}_{32}$ (Fig. 4a). One of the main differences is the larger number of energy bands in the case of $\text{In}_{16}\text{Tl}_{16}\text{I}_{31}$ (Fig. 4b) comparing to $\text{In}_{16}\text{Tl}_{16}\text{I}_{32}$ (Fig. 4a). The above feature may be caused by cancelation of the energy levels degeneracy due to the lower crystal symmetry of $\text{In}_{16}\text{Tl}_{16}\text{I}_{31}$ (space group No. 25) when compare to $\text{In}_{16}\text{Tl}_{16}\text{I}_{32}$ (space group No. 38). The other main difference is a position of the highest occupied energy band, which, in the defective crystal $\text{In}_{16}\text{Tl}_{16}\text{I}_{31}$, is located close to the unoccupied levels of the conduction band (Fig. 4b). The direct energy gap between

half-occupied and unoccupied bands, $\Delta E_{vc} \approx 0.2$ eV, may be, however, not sufficiently small to ensure the electron conductivity of the reasonable value at ambient temperature. An analysis of the partial density of states (PDOS) near the Fermi level E_F has revealed a high degree of the indium and thallium p -states hybridization at this half-occupied energy band. These p -states of indium and thallium form in predominant degree the top half-occupied band near Fermi level E_F . The contribution of the iodine p -states here is relatively smaller (Fig. 5a).

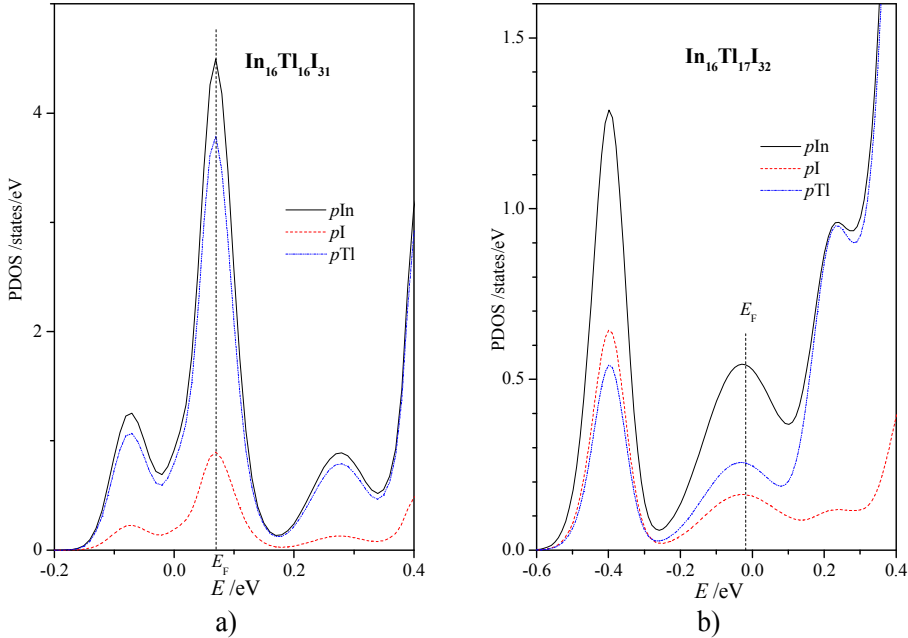


Fig. 5. Partial densities of (a) p -states of $\text{In}_{16}\text{Tl}_{16}\text{I}_{31}$ crystal with one iodine vacancy of the type I1 and (b) p -states of $\text{In}_{16}\text{Tl}_{17}\text{I}_{32}$ crystal with one interstitial thallium. The Fermi energy E_F is indicated by vertical dashed lines

Because of the large dimensions of thallium atom and the interstitial character of the corresponding defect the interatomic distances of Tl17 atom to the surrounding iodine and indium ones are smaller than the analogous values in the non-defective crystal $\text{In}_{16}\text{Tl}_{16}\text{I}_{32}$. In the present study, the interstitial Tl17 atom has been placed at the most symmetrical position with the nearest neighboring three iodine and two indium atoms. In this case, the symmetry of the optimized $\text{In}_{16}\text{Tl}_{17}\text{I}_{32}$ belongs to the space group No. 25 and is lower than that for the non-defective $\text{In}_{16}\text{Tl}_{16}\text{I}_{32}$ crystal (space group No. 38). In the relaxed crystal structure of $\text{In}_{16}\text{Tl}_{17}\text{I}_{32}$, the interatomic distances of interstitial thallium atom (Tl17) to the five nearest neighbors are the following: 3.089, 3.089 and 3.117 Å for Tl17 - I and 3.152 Å for

two Tl17 – In bonds. The absolute values of Hirshfeld charges for Tl17 and these neighboring five ions are much smaller ($q_{\text{Tl17}} = 0.02$, $q_{\text{I}} = -0.09$, $q_{\text{In}} = 0.05$) than the similar charges of such ions located far from the interstitial defect ($q_{\text{Tl}} = 0.14$, $q_{\text{I}} = -0.13$, $q_{\text{In}} = 0.12$). This leads to the conclusion that the ionic type bonding of thallium atom with the above mentioned three iodine ones is much weaker in comparison to that between thallium and iodine atoms located far from the interstitial defect associated with Tl17 atom.

The Fermi energy E_F of $\text{In}_{16}\text{Tl}_{17}\text{I}_{32}$ corresponds to the averaged energy of the top half-occupied valence band, which is shown in Fig. 4c by thick red line. The dispersion of this energy band, $\Delta E \approx 0.9$ eV, is larger than the similar value for the case of $\text{In}_{16}\text{Tl}_{16}\text{I}_{31}$ crystal with one iodine vacancy (Fig. 4b). Upper parts of this half-occupied band are located very close ($\Delta E_{\text{vc}} \approx 0.02$ eV) to the more higher lying conduction bands (Fig. 4c). Due to this feature of the top half-occupied valence band one may expect increased electron conductivity associated with the unoccupied conduction bands of $\text{In}_{16}\text{Tl}_{17}\text{I}_{32}$ crystal in comparison to $\text{In}_{16}\text{Tl}_{16}\text{I}_{31}$ one with one iodine vacancy (Fig. 4b), where the similar binding energy is much larger ($\Delta E_{\text{vc}} \approx 0.2$ eV).

An analysis of PDOS near the Fermi level E_F (Fig. 5b) indicates lower degree of the indium and thallium p -states hybridization in comparison to the case of iodine vacancy (Fig. 5a). This is in agreement with much smaller Hirshfeld charge of the interstitial thallium ion, $q_{\text{Tl17}} = 0.02$, in comparison to that of the main matrix of $\text{In}_{16}\text{Tl}_{17}\text{I}_{32}$ crystal ($q_{\text{Tl}} = 0.14$). Thus, the interstitial thallium ion of $\text{In}_{16}\text{Tl}_{17}\text{I}_{32}$ participates not so much in the formation of the half-occupied top VB (Fig. 5b) in comparison to the formation of similar top VB in $\text{In}_{16}\text{Tl}_{16}\text{I}_{31}$ crystal with iodine vacancy (Fig. 5a). These results suggest the conclusion that the main reason for the half-occupied band formation near Fermi level may be the local distortion of crystal structure but not the deviations from the stoichiometric content of the chemical elements itself forming the nominal structure of a crystal.

For the defective crystals $\text{In}_{16}\text{Tl}_{16}\text{I}_{31}$ and $\text{In}_{16}\text{Tl}_{17}\text{I}_{32}$ the Fermi level E_F is shifted to the bottom conduction bands in comparison to the non-defective one (Fig. 4), that means an appearance of the donor states with the odd number of electrons in the top VB. As a result, the increased electron conductivity is expected in the defective crystals $\text{In}_{16}\text{Tl}_{16}\text{I}_{31}$ and $\text{In}_{16}\text{Tl}_{17}\text{I}_{32}$, caused by the intra- and inter-band electron transitions. In viewpoint of the relation between conductivity σ and dielectric permittivity ε ,

$$\sigma = \sigma_1 + i\sigma_2 = -i\frac{\nu}{2}(\varepsilon_1 + i\varepsilon_2 - 1) \quad (2)$$

the spectral features of the imaginary part of dielectric function $\varepsilon_2(h\nu)$ at small photon energies are expected to be more pronounced than those for the real part of conductivity $\sigma_1(h\nu)$. In the case of defective crystals $\text{In}_{16}\text{Tl}_{16}\text{I}_{31}$ and $\text{In}_{16}\text{Tl}_{17}\text{I}_{32}$, the

total dielectric function $\varepsilon(h\nu)$ is a sum of two components corresponding to the inter- and intra-band electron transitions,

$$\varepsilon(h\nu) = \varepsilon_{\text{inter}}(h\nu) + \varepsilon_{\text{intra}}(h\nu) \quad (3)$$

Here, the intra-band component $\varepsilon_{\text{intra}}(h\nu)$ is usually associated with the low frequency and stationary electron conductivity, which is observed in metals and heavily doped semiconductors and is described satisfactorily by the monotonous Drude frequency dependence [30].

In view of the further discussion of the defective $\text{In}_{0.5}\text{Tl}_{0.5}\text{I}_{0.5}$ SSS we have approximated the CASTEP-calculated dielectric function $\varepsilon(h\nu)$ of the non-defective and defective SSS, taking into consideration the inter- and intra-band electron transitions using the OptaDOS code [31]. The intra-band part of dielectric function $\varepsilon_{\text{intra}}(h\nu)$ was approximated by the following relation according to the Drude model [28],

$$\varepsilon_{\text{intra}}(\nu) = 1 - \frac{\nu_p^2}{\nu^2 + i\gamma\nu} \quad (4)$$

where the constant value of broadening $\gamma = 10^{14} \text{ s}^{-1}$ was used. The larger the plasmon frequency ν_p , the larger the imaginary part of dielectric function $\varepsilon_{2\text{intra}}$ at the frequency $\nu = 0$, and the larger the expected direct (stationary) electron current conductivity $\sigma_{1\text{intra}}(0)$.

The polarized imaginary parts of dielectric functions $\varepsilon_2(h\nu)$ for the SSS studied are shown in Figs. 6 - 9. The main characteristic parameters of the crystals studied with $2 \times 2 \times 2$ supercells, obtained using the OptaDOS code, are presented in Table 4. For the inter-band electron transitions, the dielectric function $\varepsilon_2(h\nu)$ of the non-defective $\text{In}_{16}\text{Tl}_{16}\text{I}_{32}$ reveals large anisotropy in the range of band gap, $h\nu \geq E_g = 1.33 \text{ eV}$, which appears in the low energy shift of the dependence $\varepsilon_{2y}(h\nu)$ relating $\varepsilon_{2x}(h\nu)$ and $\varepsilon_{2z}(h\nu)$ ones (Fig. 6). This anisotropy of dielectric function is caused by the features of the top valence and bottom conduction bands of the crystal.

Analysis of dielectric functions of $\text{In}_{16}\text{Tl}_{16}\text{I}_{31}$ and $\text{In}_{16}\text{Tl}_{17}\text{I}_{32}$ crystals in the photon energy range about $0.4 - 1.8 \text{ eV}$ (Figs. 7 - 9) with the corresponding band structures (Figs. 4b, 4c) permits to claim that the fine structure of $\varepsilon_2(h\nu)$ in this narrow range is formed by the electron transitions between the half-occupied top valence band and the lowest conduction bands. In turn the low-energy unstructured tail of the dependence $\varepsilon_2(h\nu)$ in the range $0 - 0.4 \text{ eV}$ is formed due to the intra-band electron transitions within this half-occupied top valence band.

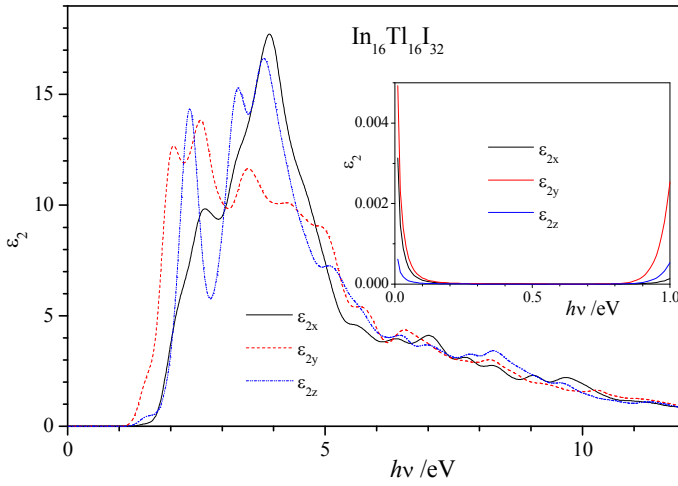


Fig. 6. Dielectric function $\varepsilon_2(h\nu)$ of $\text{In}_{16}\text{Tl}_{16}\text{I}_{32}$ crystal. ε_{2x} and ε_{2z} components correspond to the shortest (a) and longest (b) supercell dimensions ($a = 9.527 \text{ \AA}$, $b = 9.575 \text{ \AA}$, $c = 25.193 \text{ \AA}$). In the inset, the dependence $\varepsilon_2(h\nu)$ in the range $0 - 1.0 \text{ eV}$

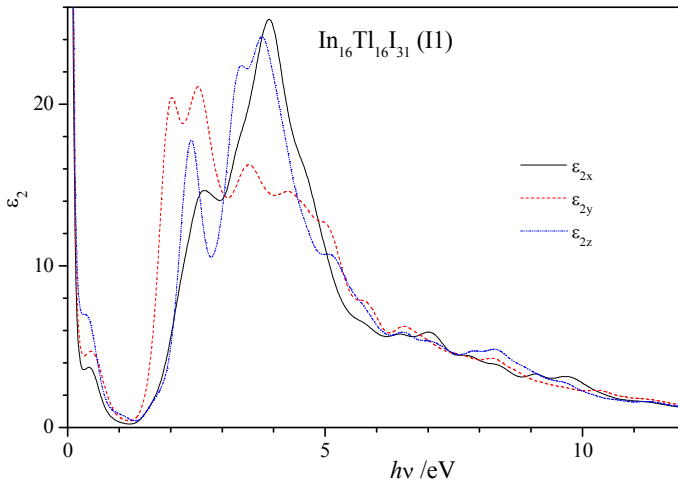


Fig. 7. Dielectric function $\varepsilon_2(h\nu)$ of $\text{In}_{16}\text{Tl}_{16}\text{I}_{31}$ with one iodine vacancy of the type II (see Fig. 1). ε_{2x} and ε_{2z} components correspond to the shortest (a) and longest (c) supercell dimensions ($a = 9.517 \text{ \AA}$, $b = 9.518 \text{ \AA}$, $c = 25.185 \text{ \AA}$)

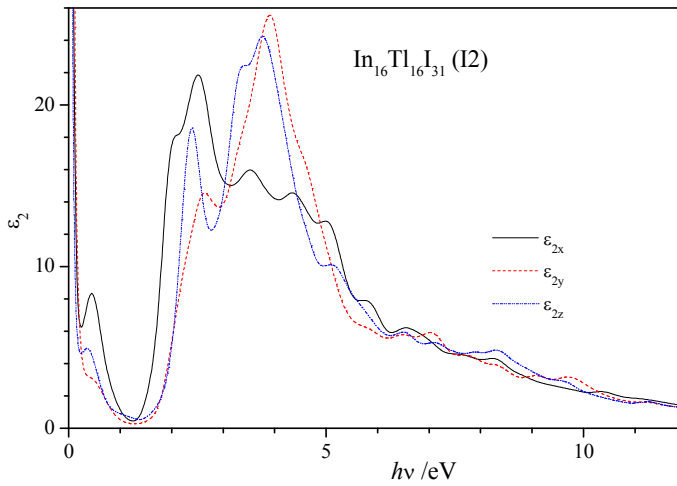


Fig. 8. Dielectric function $\varepsilon_2(h\nu)$ of $\text{In}_{16}\text{Tl}_{16}\text{I}_{31}$ with one iodine vacancy of the type I2 (see Fig. 1). ε_{2x} and ε_{2z} components correspond to the shortest (a) and longest (c) supercell dimensions ($a = 9.466 \text{ \AA}$, $b = 9.540 \text{ \AA}$, $c = 25.266 \text{ \AA}$)

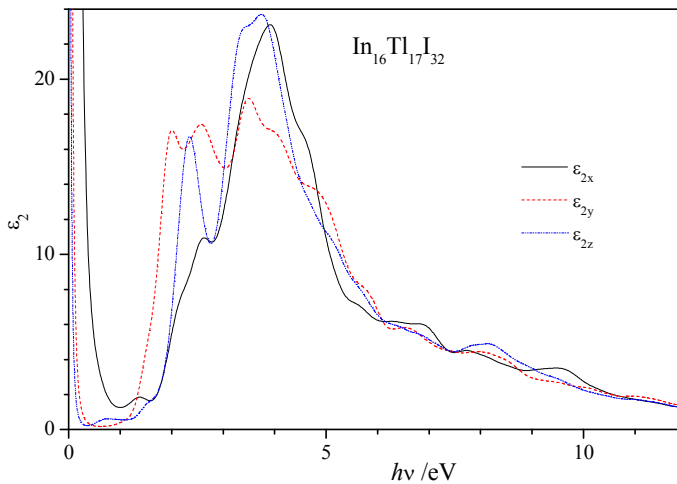


Fig. 9. Dielectric function $\varepsilon_2(h\nu)$ of $\text{In}_{16}\text{Tl}_{17}\text{I}_{32}$ crystal with one interstitial thallium atom. ε_{2x} and ε_{2z} components correspond to the shortest (a) and longest (c) supercell dimensions ($a = 9.397 \text{ \AA}$, $b = 9.848 \text{ \AA}$, $c = 25.358 \text{ \AA}$)

The iodine vacancies and thallium interstitials lead to the clear increase of the imaginary part of dielectric function $\varepsilon_2(h\nu)$ in the close vicinity of $h\nu = 0$ (Figs. 7 - 9). The largest increase of dielectric function $\varepsilon_2(h\nu)$ is observed for the case of thallium interstitial atom for the light polarization along X -axis (Fig. 9). For Y - and Z -polarizations the similar increase is much smaller (Fig. 9). Due to the features of the relation (3) the imaginary part of dielectric function $\varepsilon_2(h\nu)$ at the photon energy $h\nu$ close to zero is proportional to the plasmon frequency ν_p .

When comparing the plasmon energies ν_p for $\text{In}_{16}\text{Tl}_{17}\text{I}_{32}$ and $\text{In}_{16}\text{Tl}_{16}\text{I}_{31}$ (Table 4), the large spatial anisotropy in ν_p is observed for the crystal $\text{In}_{16}\text{Tl}_{17}\text{I}_{32}$ with thallium interstitial atom when compare to the relatively small one for $\text{In}_{16}\text{Tl}_{16}\text{I}_{31}$ crystals with iodine vacancy. For $\text{In}_{16}\text{Tl}_{17}\text{I}_{32}$ crystal this anisotropy correlates well with the corresponding energy vs. electron wave vector dispersion $E(\mathbf{K})$ of the top half-occupied band (Fig. 4c). For this top half-occupied band we have calculated the inverse electron mass tensor m^{*-1} at the K -points X , S , U , and R (Table 5) corresponding to the lowest band energies of the dispersion curve $E(K)$ (Fig. 4c). At these K -points, the mentioned band will rather be filled by the sole electron. The relation between three values $m^*_{\text{averaged}}^{-1}$ polarized along three Cartesian directions, $m^*_{\text{averaged}}^{-1}(100) > m^*_{\text{averaged}}^{-1}(010) > m^*_{\text{averaged}}^{-1}(001)$ (Table 5), is similar to the relation between three corresponding plasmon frequencies $\nu_p(100) > \nu_p(010) > \nu_p(001)$ (Table 4). This demonstrates that anisotropy of the dielectric function $\varepsilon_2(h\nu \rightarrow 0)$, the plasmon frequency ν_p and conductivity $\sigma_1(h\nu \rightarrow 0)$ for the intra-band electron transitions is closely related with anisotropy of the inverse electron mass $m^*_{\text{averaged}}^{-1}$ for the corresponding half-occupied band and BZ points at the lowest band energies. This conclusion is expected if one takes into account the known direct proportionality of the electron mobility μ and the inverse effective electron mass m^{*-1} [32, 33].

Table 4. Parameters related to the inter- and intra-band electron transitions in non-defective ($\text{In}_{16}\text{Tl}_{16}\text{I}_{32}$) and defective ($\text{In}_{16}\text{Tl}_{17}\text{I}_{32}$, $\text{In}_{16}\text{Tl}_{16}\text{I}_{31}$ (I1), $\text{In}_{16}\text{Tl}_{16}\text{I}_{31}$ (I2)) SSS calculated using OptaDOS code [31]

Parameter	Light polarization vector	Plasmon energy $h\nu_p$ /eV	Indirect band gap E_{gi} /eV	Direct band gap E_{gd} /eV	DOS at E_F (spin up \uparrow)	DOS at E_F (spin down \downarrow)
Crystal	(1 0 0)	$1.45 \cdot 10^{-3}$	1.531	1.54	$1.3 \cdot 10^{-5}$	
	(0 1 0)	$1.82 \cdot 10^{-3}$				
	(0 0 1)	$6.43 \cdot 10^{-4}$				

In ₁₆ Tl ₁₇ I ₃₂	(1 0 0)	1.66	0.054	0.53 (↑)	1.08 (↑)
	(0 1 0)	0.308		0.52 (↓)	1.08 (↓)
	(0 0 1)	0.173			
In ₁₆ Tl ₁₆ I ₃₁ (I1 vacancy)	(1 0 0)	0.444	0.011	0.32 (↑)	2.79 (↑)
	(0 1 0)	0.486		0.28 (↓)	2.61 (↓)
	(0 0 1)	0.497			
In ₁₆ Tl ₁₆ I ₃₁ (I2 vacancy)	(1 0 0)	0.420	0.009	0.31 (↑)	2.65 (↑)
	(0 1 0)	0.526		0.32 (↓)	2.61 (↓)
	(0 0 1)	0.311			

Table 5. Principal components of the inverse electron masses m^{*-1} for three Cartesian directions of the top half-occupied band of In₁₆Tl₁₇I₃₂ crystal calculated at four BZ points and the averaged one $m^*_{\text{averaged}}^{-1}$

Cartesian direction	m^{*-1}/m_e		
	100	010	001
BZ point			
$X(0.5,0,0)$	3.74	-0.241	0.473
$S(0.5,0.5,0)$	6.236	0.245	0.044
$U(0.5,0,0.5)$	4.122	-0.280	-0.446
$R(0.5,0.5,0.5)$	6.248	0.35	-0.039
$m^*_{\text{averaged}}^{-1}/m_e$	5.086	0.0185	0.008

Conclusion

DFT-based calculations of the band electronic structure and dielectric function have been performed for both the non-defective and defective (with iodine anion vacancy and thallium cation interstitial) solid state solutions of In_{0.5}Tl_{0.5}I. On the basis of the band structures obtained the tensors of effective electron mass have been calculated for the non-defective InI, In_{0.5}Tl_{0.5}I and TlI crystals. Absolute values of the effective electron masses for the non-defective crystals studied have been found to be of the same order of magnitude. Minor decrease of the averaged effective electron masses takes place in the direction TlI → In_{0.5}Tl_{0.5}I → InI.

Imaginary part of the dielectric functions calculated for the defective solid state solutions of In_{0.5}Tl_{0.5}I with taking into consideration the inter- and intra-band

electron transitions possesses the considerable, in comparison to the non-defective $\text{In}_{0.5}\text{Tl}_{0.5}\text{I}$ SSS, low-frequency component ($h\nu < 0.4$ eV). That points to the large low-frequency and stationary electron conductivity in the defective $\text{In}_{0.5}\text{Tl}_{0.5}\text{I}$ crystals. This finding is related to the crucial changes of band structure (formation of the donor half-occupied band close to the unoccupied conduction bands) due to the corresponding defects of SSS crystal structure (iodine vacancy or thallium interstitial atom). By taking into account defects in the calculation the results obtained are in qualitative agreement with experimental observations related to the band gap as well as n -type conductivity in $\text{In}_x\text{Tl}_{1-x}\text{I}$ SSS [11]. Finally, it should be emphasized that in the case of real crystals, in particular metal-halides, in order to better match the calculated and experimental data, the existing structural imperfections should be taken into account in a suitable manner at the corresponding *ab initio* quantum chemical calculations.

Acknowledgments

Computer calculations have been performed using the Academic release of the CASTEP code (ver. 18.1) at WCSS of Wrocław University of Technology, Poland (project No. 053). The work was supported by the Project of Young Scientists (Co-66Hp) of Ukraine.

References

1. M.G. Kanatzidis, Discovery-synthesis, design, and prediction of chalcogenide phases. *Inorg. Chem.* 56 (2017) 3158 - 3173.
2. I. Chung, G. Mercouri, M.G. Kanatzidis, Metal chalcogenides: a rich source of nonlinear optical materials, *Chem. Mater.* 26 (2014) 849 - 869.
3. M. Piasecki, M.G. Brik, I.E. Barchiy, K. Ozga, I.V. Kityk, A.M. Al-Naggar, A.A. Albassam, T.A. Malakhovskaya, G. Lakshminarayana, Band structure, electronic and optical features of Tl_4SnX_3 ($X=\text{S}, \text{Te}$) ternary compounds for optoelectronic applications, *J. Alloy. Compd.* 710 (2017) 600-607. <https://doi:10.1016/j.jallcom.2017.03.280>.
4. Piasecki, M., Myronchuk, G.L., Zamurueva, O.V., Khyzhun, O.Y., Parasyuk, O.V., Fedorchuk, A.O., Albassam, A., El-Naggar, A.M., Kityk, I.V., 02/2016. Huge operation by energy gap of novel narrow band gap $\text{Tl}_{1-x}\text{In}_{1-x}\text{B}_x\text{Se}_2$ ($B = \text{Si}, \text{Ge}$): DFT, x-ray emission and photoconductivity studies. *Mater. Res. Express*, 3(2):025902. <https://doi:10.1088/2053-1591/3/2/025902>.
5. Brik, M.G., Piasecki, M., Kityk, I.V., Structural, Electronic, and Optical Features of $\text{CuAl}(\text{S}_{1-x}\text{Se}_x)_2$ Solar Cell Materials, 02/2014. *Inorg. Chem.*, 53:2645. <https://doi:10.1021/ic403030w>.

6. Atuchin V.V., et al., Structure, defects, mechanical and optical properties of hexagonal semiconductor $\text{GaSe}_{1-x}\text{S}_x$ single crystals, *Siberian Conference on Control and Communications*, Tomsk, IEEE Xplore 2007, pp. 179-184. [https://doi: 10.1109/SIBCON.2007.371321](https://doi.org/10.1109/SIBCON.2007.371321).
7. Q. Guo, A. Assoud, H. Kleinke, Thallium-based chalcogenides as thermoelectrics, chapter 12, ed. C. Uher, *Materials Aspect of Thermoelectricity*, Boca Raton, CRC Press, 2016.
8. I.V. Blonsky, M.I. Kolinko, Yu.O. Lun, A.V. Franiv, Influence of static lattice disordering on optical properties of $\text{In}_x\text{Tl}_{1-x}\text{I}$ crystals, *Proc. SPIE* 2647 (1995) 452-454. [https://doi: 10.1117/12.226732](https://doi.org/10.1117/12.226732).
9. A. Franiv, R. Peleshchynshyn, Y. Kolosivski, Optical properties of quantum size nanocrystals $\text{In}_x\text{Tl}_{1-x}\text{I}$ embedded in solid matrices, *Ukr. J. Phys. Opt.* 1 (2000) 24-27. [https://doi: 10.3116/16091833/1/1/24/2000](https://doi.org/10.3116/16091833/1/1/24/2000).
10. Y.O. Dovhyi, A.V. Franiv, S.V. Ternavska, Isostructural phase transition in $\text{In}_x\text{Tl}_{1-x}\text{I}$ system, *Ukr. J. Phys. Opt.* 2 (2001) 141-147. [https://doi: 10.3116/16091833/2/3/141/2001](https://doi.org/10.3116/16091833/2/3/141/2001).
11. A.I. Kashuba, M. Piasecki, O.V. Bovgyra, V.Yo. Stadnyk, P. Demchenko, A. Fedorchuk, A.V. Franiv, B. Andriyevsky, Specific features of content dependences for energy gap in $\text{In}_x\text{Tl}_{1-x}\text{I}$ solid state crystalline alloys, *Acta Phys. Pol. A* 133 (2018) 68-75. [https://doi: 10.12693/APhysPolA.133.68](https://doi.org/10.12693/APhysPolA.133.68).
12. M.S. Brodin, I.V. Blonskii, B.M. Nitsovich, A.S. Krochuk, A.V. Franiv, Dynamical properties of excitons in layer crystals of PbI_2 , *Phys. Status Solidi B* 111 (1982) 625-630. [https://doi: 10.1002/pssb.2221110226](https://doi.org/10.1002/pssb.2221110226).
13. R. Zallen, M.L. Slade, Inter-polytype conversion and layer-layer coupling in PbI_2 , *Solid State Commun.* 17 (1975) 1561-1566. [https://doi: 10.1016/0038-1098\(75\)90996-5](https://doi.org/10.1016/0038-1098(75)90996-5).
14. A.V. Franiv, A.I. Kashuba, O.V. Bovgyra, Non-linear transducer of medium-infrared radiation, *Patent of Ukraine*, Bulletin No. 16, published on 2017-08-28 (<http://base.uipv.org/searchINV/>, application number u201702658).
15. L. Helmholz, *Z. Kristallogr.* 95 (1936) 129-137.
16. R.E. Jones, D.H. Templeton, *The crystal structure of indium (I) iodide*, *Acta Crystallogr.* 8 (1955) 847. <https://doi.org/10.1107/S0365110X55002594>.
17. G. Meyer, T. Staffel, Note on the red monohalides of indium, InCl , InBr , InI , *Z. Anorg. Allg. Chem.* 574 (1989) 114-118. [https://doi: 10.1002/zaac.655740112](https://doi.org/10.1002/zaac.655740112).
18. R.P. Lowndes, C.H. Perry, Molecular structure and anharmonicity in thallium iodide, *J. Chem. Phys.* 58 (1973) 271-278. [https://doi: 10.1063/1.1678917](https://doi.org/10.1063/1.1678917).
19. A.V. Franiv, V.Y. Stadnyk, A.I. Kashuba, R.S. Brezvin, O.V. Bovgira, A.V. Futei, Temperature behavior of thermal expansion and birefringence of

- $\text{In}_x\text{Tl}_{1-x}\text{I}$ - substitution solid solutions, *Opt. Spectrosc.* 123 (2017) 177-180. [https://doi: 10.1134/S0030400X17070074](https://doi.org/10.1134/S0030400X17070074).
20. A.V. Franiv, A.I. Kashuba, O.V. Bovgyra, O.V. Futey, Elastic properties of substitutional solid solutions $\text{In}_x\text{Tl}_{1-x}\text{I}$ and sounds wave velocities in them, *Ukr. J. Phys.* 62 (2017) 679-684. [https://doi: 10.15407/ujpe62.08.0679](https://doi.org/10.15407/ujpe62.08.0679).
21. A.I. Kashuba, A.V. Franiv, R.S. Brezvin, O.V. Bovgyra, Birefringence of $\text{In}_x\text{Tl}_{1-x}\text{I}$ solid state solution, *Functional materials* 23 (2017) 26–30. [https://doi: 10.15407/fm24.01.026](https://doi.org/10.15407/fm24.01.026).
22. A.I. Kashuba, Ya.A. Zhydachevskyy, I.V. Semkiv, A.V. Franiv, O.S. Kushnir, Photoluminescence in the solid solution $\text{In}_{0.5}\text{Tl}_{0.5}\text{I}$, *Ukr. J. Phys. Opt.* 19 (2018) 1-8.
23. S.J. Clark, M.D. Segall, C.J. Pickard, P.J. Hasnip, M.J. Probert, K. Refson, M.C. Payne, First principles methods using CASTEP, *Zeitschrift für Kristallographie* 220 (2005) 567-570. [https://doi: 10.1524/zkri.220.5.567.65075](https://doi.org/10.1524/zkri.220.5.567.65075).
24. Perdew, J.P., Ruzsinszky, A., Csonka, G.I., Vydrov, O.A., Scuseria, G.E., Constantin, L.A., Zhou, X., Burke, K., Restoring the Density-Gradient Expansion for Exchange in Solids and Surfaces, 2008. *Phys. Rev. Lett.* 100, 136406. [https://doi: 10.1103/PhysRevLett.101.239702](https://doi.org/10.1103/PhysRevLett.101.239702).
25. H.J. Monkhorst, J.D. Pack, Special points for Brillouin-zone integrations, *Phys. Rev. B* 13 (1976) 5188-5192. <https://doi.org/10.1103/PhysRevB.13.5188>.
26. Y. Hinuma, G. Pizzi, Y. Kumagai, F. Oba, I. Tanaka, Band structure diagram paths based on crystallography, *Comp. Mater. Sci.* 128 (2017) 140-184. [https://doi: 10.1016/j.commatsci.2016.10.015](https://doi.org/10.1016/j.commatsci.2016.10.015).
27. L. Pauling, *The Nature of the Chemical Bond*, 3rd edition. Copyright 1939 and 1940, 3rd edition copyright © by Cornell University, 1960.
28. M. Grundmann. *The Physics of Semiconductors*, Springer-Verlag Berlin Heidelberg, 2006.
29. Fonari, A., Sutton, C., Effective mass calculator, 2012.
30. H. Fujiwara, *Spectroscopic ellipsometry: principles and applications*, John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, 2007.
31. A.J. Morris, R. Nicholls, C.J. Pickard, J. Yates, OptaDOS: A tool for obtaining density of states, core-level and optical spectra from electronic structure codes, *Comp. Phys. Comm.* 185 (2014) 1477-1485. [https://doi: 10.1016/j.cpc.2014.02.013](https://doi.org/10.1016/j.cpc.2014.02.013).
32. P.Y. Yu, M. Cardona, *Fundamentals of Semiconductors: Physics and Materials Properties*. Springer, 2010.

33. Hou, Q.Y., Li, W.C., Xu, Z.C., Zhao, C.W., Study on the effect of high V doping on the conductivity of anatase TiO₂, 2016. *Int. J. Modern Phys. B* 30, 165001. [https://doi: 10.1142/S0217979216500016](https://doi.org/10.1142/S0217979216500016).

Abstract

We investigate an influence of the various crystal structure imperfections on the electronic properties and dielectric functions for In_{0.5}Tl_{0.5}I semiconductor in the frame of the density functional theory calculations. The tensor of electron effective mass m^*_{ij} of InI, In_{0.5}Tl_{0.5}I and TlI crystals has been calculated for the valence and conduction bands and different *K*-points of Brillouin zone. Dielectric functions $\epsilon(h\nu)$ of the defective crystals based on In_{0.5}Tl_{0.5}I solid state solution with iodine vacancy and thallium interstitial atom were calculated taking into consideration the inter-band and intra-band electron transitions. The studies of the defective crystals reveal increased low-frequency and stationary electron conductivity with anisotropy resulted from the anisotropy of the electron effective mass tensor. Our findings explain the origin of crucial changes in the band structure by formation the donor half-occupied levels close to the unoccupied conduction bands due to the crystal structure defects, i.e. iodine vacancy or thallium interstitial atom. It has been shown that in the case of real crystals, in particular metal-halides, the proper consideration of defects in quantum-chemical calculations results in a better matching of the theoretical and experimental results in comparison to the case when the perfect crystal structure had been used for calculations.

Streszczenie

Zbadano wpływ różnych niedoskonałości struktury krystalicznej na właściwości elektronowe i funkcje dielektryczne półprzewodnika In_{0.5}Tl_{0.5}I w ramach teorii funkcjonału gęstości. Został obliczony tensor efektywnej masy elektronów m^* kryształów InI, In_{0.5}Tl_{0.5}I i TlI dla pasm walencyjnych i przewodnictwa oraz różnych *K*-punktów strefy Brillouina. Funkcje dielektryczne $\epsilon(h\nu)$ domieszkowanych kryształów roztworów stałych In_{0.5}Tl_{0.5}I z wakansami jodu i atomami międzywęzłowymi talu zostały obliczone z uwzględnieniem międzypasmowych i wewnątrz-pasmowych przejść elektronowych. Badania domieszkowanych kryształów ujawniły zwiększoną przewodność elektronową niskoczęstotliwościową i stacjonarną o anizotropii wynikającej z anizotropii tensora efektywnej masy elektronów. Przeprowadzone badania wyjaśniają obserwowane duże zmiany struktury pasmowej pochodzące z utworzenia pól wypełnionych poziomów donorowych w pobliżu niezajętych pasm przewodnictwa wynikających z defektów struktury krystalicznej, tj. wakansów jodu czy atomów międzywęzłową talu. Wykazano, że w przypadku kryształów rzeczywistych, w szczególności halogenków

metali, właściwe uwzględnienie defektów w obliczeniach kwantowo-chemicznych daje możliwość lepszego dopasowania obliczeń teoretycznych do wyników doświadczalnych w porównaniu do obliczeń bazujących na strukturze krystalicznej doskonałej.

Marek Gąsiorowski *

Aleksy Patryn

Leszek Bychto

Wydział Elektroniki i Informatyki

Politechnika Koszalińska

*gasiorowski.marek@o2.pl

**aleksy.patryn@tu.koszalin.pl

***leszek.bychto@tu.koszalin.pl

Możliwości oraz obszar zastosowań małogabarytowego spektrometru DLP NIRScan Nano do błyskawicznych pomiarów spektralnych

Słowa kluczowe: DLP, NIRScan, Nano EVM, mobilność pomiarów, spektroskopia

1. Wprowadzenie

Spektroskopia to potężne narzędzie do analizy różnych materiałów poprzez oświetlanie próbek za pomocą światła, które w zależności od właściwości fizycznych materiałów badanych jest pochłonięte, odbite bądź powoduje emisję innego światła [1-3]. Wszystkie te zmiany w funkcji długości fali świetlnej są rejestrowane przez aparaturę pomiarową i charakteryzują dany materiał. W zależności od urządzenia przedział pomiarowy długości fal świetlnych próbki jest zróżnicowany. Ze względu na typ wykorzystywanych detektorów spektrometry dzielą się na działające w zakresie ultrafioletu (UV), światła widzialnego (VIS), bliskiej, średniej i dalekiej podczerwieni (NIR, MIR i IR). Bardziej rozbudowane urządzenia umożliwiają pomiar w kilku zakresach np. UV-VIS, UV-VIS-NIR. Tradycyjnie badania różnorodnych obiektów organicznych lub nieorganicznych znajdujących się w różnym stanie fazowym odbywa się w następujący sposób – zbieranie materiału, przygotowanie próbek, dostarczenie ich do laboratorium gdzie znajduje się specjalistyczna aparatura pomiarowa [1-3]. Tradycyjne laboratoryjne spektrometry są zwykle duże i drogie.

W ostatnich latach progres w rozwoju spektroskopii doprowadził do powstania na rynku nowej klasy zminiaturyzowanych spektrometrów, w tym na zakresy widzialny i bliskiej podczerwieni (VS, NIR) [4]. Miniaturyzacja urządzeń i systemów jest powiązana z tendencją do dokonywania pomiaru w fizycznej lokalizacji „próbki”, niezależnie od tego, czy jest to gotowy produkt, proces produkcyjny, czy środowisko.

Ważnym atutem stosowania mniejszych przyrządów jest potencjalna możliwość realizacji rozproszonych schematów pomiarowych (wiele lokalizacji w procesach) i podejściu do „zdalnego” monitorowania pomiarów środowiskowych i „polowych”.

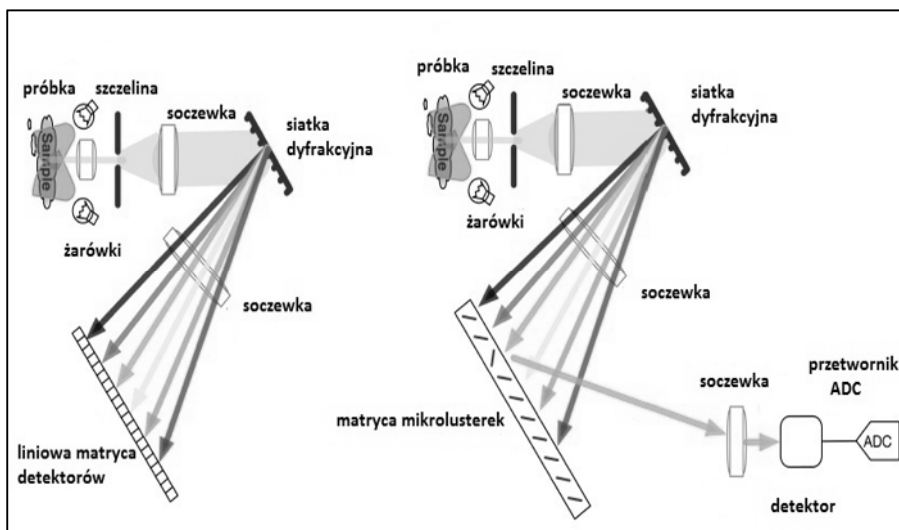
Mikrominiaturowe spektrometry na zakres NIR spowodowały aktywny wzrost badań w zakresie ich zastosowań do analizy różnorodnych obiektów z punktu widzenia kontroli parametrów i procesów dot. różnych obiektów, np. nanokompozytów [5-6], półprzewodników, w tym cienkowarstwowych [7-8] oraz obiektów organicznych (rośliny) [9-13].

W niniejszej pracy została przeprowadzona analiza działania miniaturowego „*fullintegrated*” spektrometru z takiej grupy urządzeń, mianowicie marki DLP NIRscan Nano EVM [14], zaprezentowano wyniki pewnej modernizacji jego konstrukcji w celu usprawnienia pomiarów i zwiększenia wydajności oraz zaprezentowano wyniki pomiarów testowych widma refleksyjności i absorpcyjności serii wybranych obiektów testowych przeprowadzonych w celu oszacowania możliwości i skuteczności dalszych badań.

2. Opis urządzenia, dane techniczne, zasada pracy, oprogramowanie

Koncepcja działania urządzenia i zastosowania matrycy mikro-lusterek pozwalająca na wykorzystanie tylko pojedynczego detektora a nie całej linii (matrycy) detektorów została przedstawiona na rys. 1.

Pomiary w tym, jednopikselowym spektrometrze są realizowane w inny sposób niż w „klasycznym” siatkowym, w tym matrycowym spektrometrze. Zasada pracy polega na mikromechanicznym sterowaniu zestawem mikrolusterek i kierowaniu kolejno wybranej wiązki światła o odpowiedniej długości do okna detektora. Wykorzystywany jest tylko pojedynczy detektor, a nie cała linia detektorów. Przez takie rozwiązanie koszt urządzenia jest znacznie niższy.

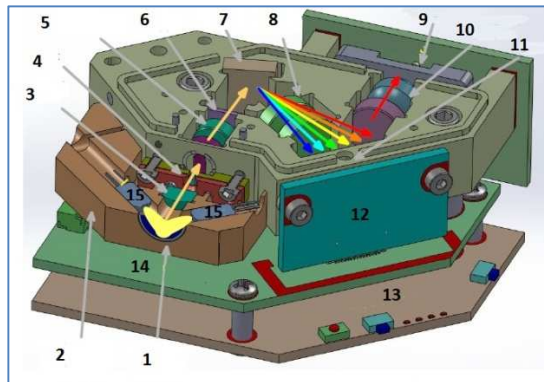


Rys. 1. Zasada pracy „klasycznego” siatkowego spektrometru z detektorem matrycowym (po lewej) oraz spektrometru wykorzystujące matryce ruchomych lusterek (po prawej), wg [14].

Ważnym atutem spektrometru DLP NIRScan Nano EVM jest prędkość realizacji pomiarów oraz ich dokładność. Zasadniczy tryb konfiguracji pracy polega na połączeniu przewodem USB urządzenia wraz z komputerem. Za pomocą oprogramowania przystosowanego do obsługi, można przeprowadzać pomiary, konfigurować ilość pomiarów, tryb oraz implementować własne algorytmy oświetlania próbek. Urządzenie może również współpracować z smartfonami za pomocą specjalnej aplikacji ogólnodostępnej. Komunikacja pomiędzy urządzeniami realizowana jest za pomocą modułów Bluetooth. Po krótkim opisie łatwo można dostrzec największe zalety – prosta obsługa, mobilność dzięki niewielkim rozmiarom (wymiary: 58x62x36 mm – rys. 3) oraz prostota wykonywania pomiarów co umożliwia wachlarz potencjalnych zastosowań.

Oświetlanie badanych obiektów odbywa się za pomocą dwóch żarówek typu ILT 1088-1 o mocy 1,4 W zasilanych napięciem 5 V. Przesłona toru optycznego ma wymiary 1,8 x 25 μm , liczba falowa „f” przesłony wynosi 2,5. Urządzenie uzyskuje widmo odbiciowe fali świetlnej próbek w zakresie od 900 do 1700 nm z rozdzielczością optyczną co 10 nm. Zastosowano detektor półprzewodnikowy oparty na arsenie indu galu (InGaAs), niechłodzony, o średnicy okna 1 mm [14].

Urządzenie składa się z następujących podstawowych elementów, które wyróżniono na rys. 2 [14]:



Rys. 2. Widok podstawowych elementów składowych urządzenia (1 – okno spektrometru; 2 – oświetleniowo-odbiciowy moduł; 3 – soczewka skupiająca; 4 – szczelina; 5 – soczewki kolimacyjne; 6 – filtr szerokopasmowy; 7 – siatka dyfrakcyjna; 8 – soczewki skupiające; 9 – detektor, 10 - soczewki kondensujące; 11 – DMD (ang. Digital Micromirror Device); 12 – płytka DMD; 13 – płytka mikrokontrolera; 14 – płytka sterownika spektrometru DLP; 15 – żarówki oświetlające)

Jak można zauważyć z poglądowego rysunku, mikro-lusterka pojedynczo nakierowują odpowiednie wiązki światła o danej długości fali do detektora. Transmisja światła odbywa się szeregowo do okna detektora co wydłuża czas analizy badanego widma światła. Jest to pewnego rodzaju ograniczenie. Nie możliwe jest jednoczesna analiza całego przedziału pomiarowego ze względu na pojedynczy detektor.

Zasadniczo urządzenie udostępnia przeprowadzenie badań w trzech podstawowych trybach pomiarowych: absorbancja, reflektancja oraz intensywność światła, które to zmieniają się w funkcji długość fali świetlnej z krokiem co 10 nm w zakresie pomiarowym urządzenia.

Po uruchomieniu pomiaru za pomocą aplikacji służącej do obsługi urządzenia następuje oświetlenie próbki. Fotorezystor umieszczony przed szczeliną wejściową dokonuje automatycznej kalibracji i dostosowania zakresu intensywności światła. Następnie odbite światło przenika przez soczewkę skupiającą, przechodząc przez szczelinę trafiając na kolejne soczewki skupiające oraz filtr szerokopasmowy. Światło dociera do siatki dyfrakcyjnej i ulega dyspersji na składowe o różnej długości fali świetlnej. Za pomocą soczewek skupiających trafia na matrycę cyfrowo sterowanych mikro-lusterek (ang. Digital Mirror Device - DMD). Jest to innowacyjne rozwiązanie, które ma pewne zalety oraz kilka wad. Zastosowanie takiego rozwiązania pozwala na obniżenie kosztów produkcji urządzenia oraz znaczne zmniejszenie jego rozmiarów w porównaniu do klasycznych spektrometrów, w których wykorzystuje się matrycę detektorów, jednak powoduje to zwiększenie czasu realizacji pomiaru gdyż zakres pomiarowy jest analizowany

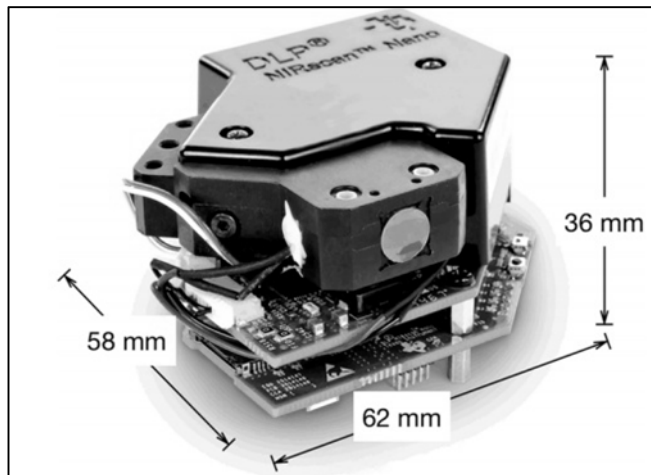
szeregowo podpasdami. Na rysunku numer 1 przedstawiono różnicę w działaniu tych dwóch trybów.

Producent zapewnia dostęp do oprogramowania działającego na komputerze z systemem operacyjnym Windows. Do uruchomienia spektrometru potrzebny jest kabel USB z wtyczką mini USB oraz odpowiednie sterowniki udostępniane przez producenta. Możliwe jest również pobranie aplikacji na smartfon lub tablet, która łączy się z urządzeniem DLP za pomocą wbudowanego modułu Bluetooth w konfiguracji pracy „slave”. Tak więc sterowanie urządzeniem możliwe jest na różny sposób.

Po podłączeniu spektrometru do komputera za pomocą kabla USB oraz uruchomieniu oprogramowania „DLP NIRscan Nano GUI v2.1.0 urządzenie jest gotowe do pracy – należy tylko wybrać źródło kalibracji urządzenia oraz pozostawić domyślne ustawienia realizacji pomiarów bądź zmienić je. Można wybrać zapisany wzorcowy pomiar kalibracyjny lub przeprowadzić własny pomiar kalibracyjny (odniesienia).

3. Rozszerzenia funkcjonalności urządzenia

Spektrometr został opracowany przez producenta w wersji „otwartej”, bez żadnej obudowy i elementów mocujących. Jego widok zewnętrzny zaprezentowano na rys. 3.



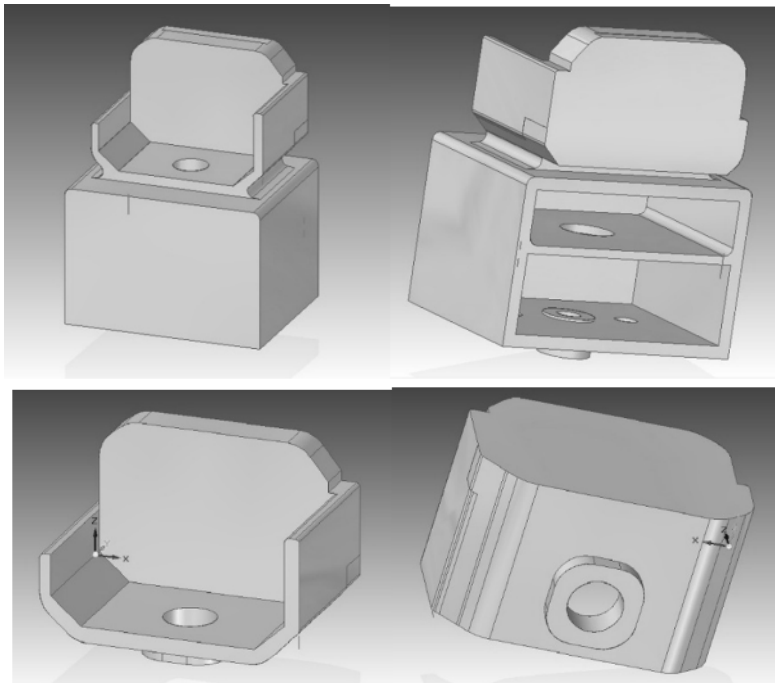
Rys. 3. Widok zewnętrzny DLP NIRScan Nano EVM, wg. [15]

Po wstępnej analizie konstrukcji była opracowana własna konstrukcja dodatkowych elementów umożliwiająca rozszerzenie jego funkcjonalności. Opracowano dwa typy nasadek (rys. 4), pierwszy typ służy do wykorzystania

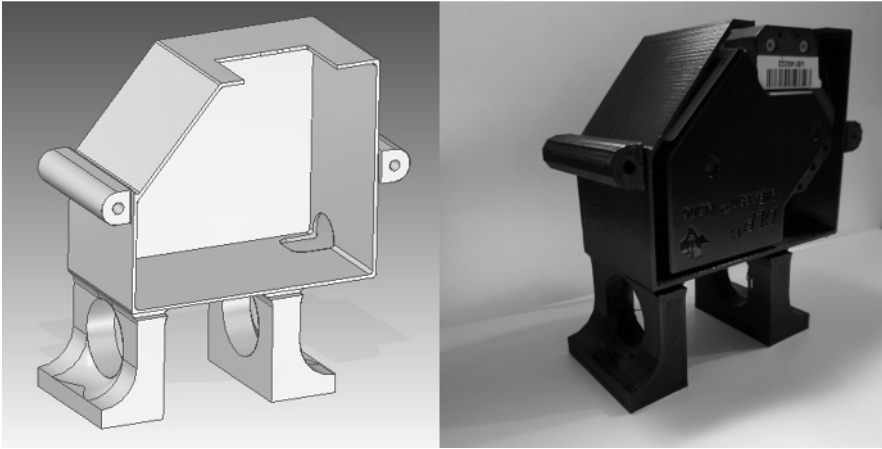
światłowodu w torze pomiarowym, który może być wykorzystany do realizacji badań w trudno dostępnych miejscach jak przedłużenie szczeliny. Drugi typ opracowanej nasadki umożliwia badanie różnego rodzaju materiałów pomiaru tłumienia (ilościowe tłumienie obiektów przezroczystych). Nasadka mocowana jest bezpośrednio na urządzeniu, wykonane jest w postaci małego pudełka z przysłoną. W górnej części umieszczono źródło światła – jest to żarówka ILT 1088-1 o mocy 0,7 W zasilana napięciem stałym wynoszącym 5 V i kompatybilna z układem zasilania i sterowania DLP NIRScan Nano [14, 15].

W opracowanym układzie z dodatkowymi elementami badanie polega na oświetlaniu przedmiotu umieszczonego w komorze za pomocą żarówki, w zależności od badanego materiału część światła ulega odbiciu, część jest absorbowana a w niektórych przypadkach światło penetruje próbkę i dociera do szczeliny spektrometru. Dodatkowo zaprojektowano i wykonano specjalną obudowę (rys. 5), która zabezpiecza układy elektroniczne urządzenia (standardowo płytkę elektroniczną sterownika urządzenia jest odłożona, istnieje możliwość jej dotknięcia a co za tym idzie uszkodzenia obwodów sterowania) oraz gwarantuje ergonomiczną pozycję spektrometru podczas pracy z nasadką z rysunku nr 4 – badanie przezroczca.

Wszystkie elementy te zostały wykonane w technice druku 3d.



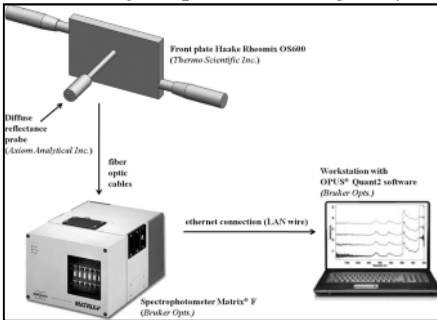
Rys. 4. Projekt nasadek umożliwiających podłączenie światłowodu lub badania przezroczca



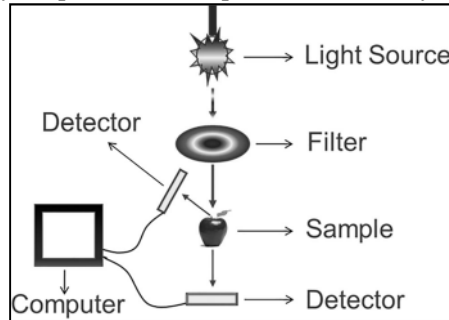
Rys. 5. Projekt (a) i realizacja fizyczna obudowy spektrometru DLP, spektrometr w opracowanej obudowie (b).

4. Testowy próby zastosowań

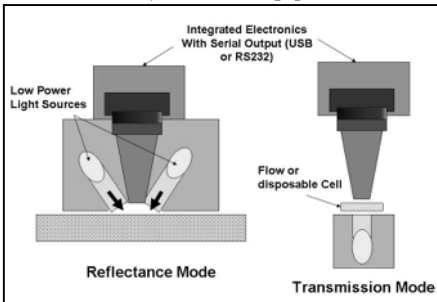
Zestawienie porównawcze konfiguracji do pomiaru widm refleksyjności i absorpcyjności za pomocą małowagarytowych spektrometrów przestawiono na rys. 6.



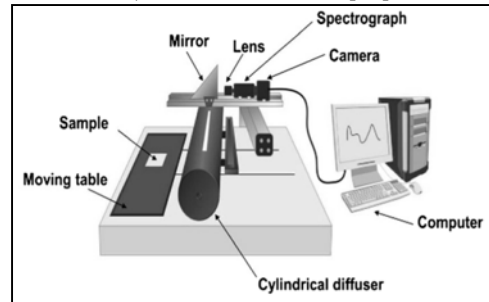
a) Barbas [6]



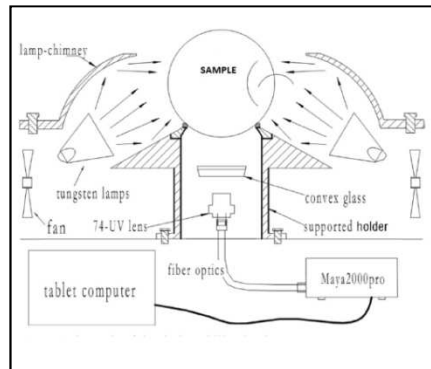
b) Chandrasekaran [13]



c) Coates [4]



d) Yuan-Yuan Pu [9]



e) Lei-Ming Yuan [10]

Rys. 6. Zestawienie porównawcze. Konfiguracja widma reflektancji i absorbcyjności za pomocą małogabarytowych spektrometrów

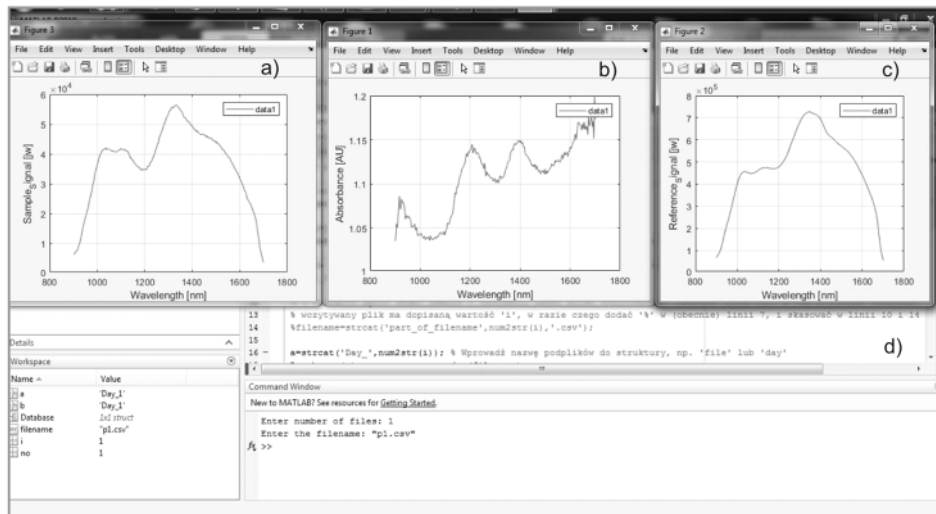
Do realizacji pomiarów została wybrana metoda zbliżona do wariantu rys. 6(c) z ustawieniem okna spektrometru praktycznie w bezpośrednim kontakcie z badaną powierzchnią. Badania zapisywane są w plikach o formacie .dat lub .xlms, który jest łatwy do przetwarzania za pomocą programu Matlab bądź Orgin.

Ważnym aspektem pozwalającym kontrolować powtarzalność warunków powtarzania i wiarygodność wyników jest możliwość stałego kontrolowania temperatury nie tylko systemu w całości, ale też temperaturę detektora oraz wilgotności w obszarze pomiarowym. W trakcie pomiarów testowych prezentowanych w tej serii pomiarowej nie było konieczności dążyć do unifikacji warunków pomiarowych, jednak na dalszych etapach badawczych jak i przy opracowaniu zaawansowanych algorytmach obróbki danych to może być konieczne.

Wszystkie pomiary były wykonane w następującym trybie: ekspozycja do pomiaru w pojedynczym punkcie – 0,635 ms, ilość powtórzeń pomiaru w każdym punkcie widma – 6, ilość punktów pomiarowych na widmie – 227, rozdzielczość na punkt pomiarowy – 7,07 nm, ogólny czas pomiaru jednego widma w obszarze 900-1700 nm – 2,70 s. Do analizy uzyskanych pomiarów wykorzystano oprogramowanie Matlab ze względu na duży wachlarz możliwości implementacji różnego rodzaju algorytmów przetwarzających dane. Dzięki oprogramowaniu istnieje możliwość porównywania serii pomiarowych między poszczególnymi pomiarami. Program umożliwi również wizualizację danych, określanie współczynników korelacji, wskazywanie charakterystycznych punktów w zależności od zastosowania urządzenia.

Na rys. 7 przedstawiono wizualizację wyników pomiaru i obróbki widma dla przykładowej próbki (a - sygnał bezpośrednio mierzony, b - widmo absorbcyjności)

otrzymane po obróbce mierzonego sygnału i uwzględnienia widma referencyjnego (c) oraz fragment programu do wizualizacji wyników za pomocą programu Matlab.



Rys. 7. Wizualizacja wyników pomiaru i obróbki widma dla przykładowej próbki (a - sygnał bezpośrednio mierzony, b - widmo absorpcji otrzymane po obróbce mierzonego sygnału i uwzględnienia widma referencyjnego (c) oraz fragment programu (d) do wizualizacji wyników za pomocą programu Matlab.

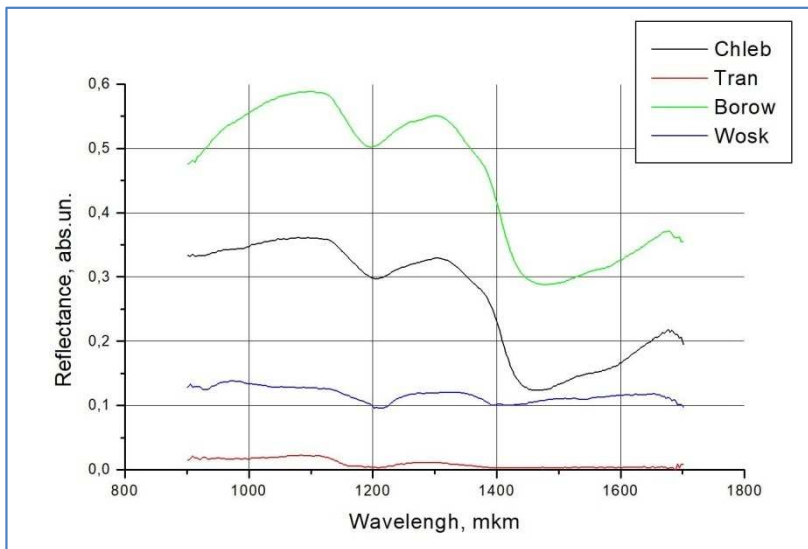
Do serii pomiarowej testowej zostały wybrane obiekty o różnym charakterze powierzchni i składu wewnętrznego. Typ obiektów i wybrane parametry i warunki w trakcie pomiarów przedstawiono w Tab. 1.

Tab. 1. Typ obiektów i wybrane parametry i warunki w trakcie pomiarów.

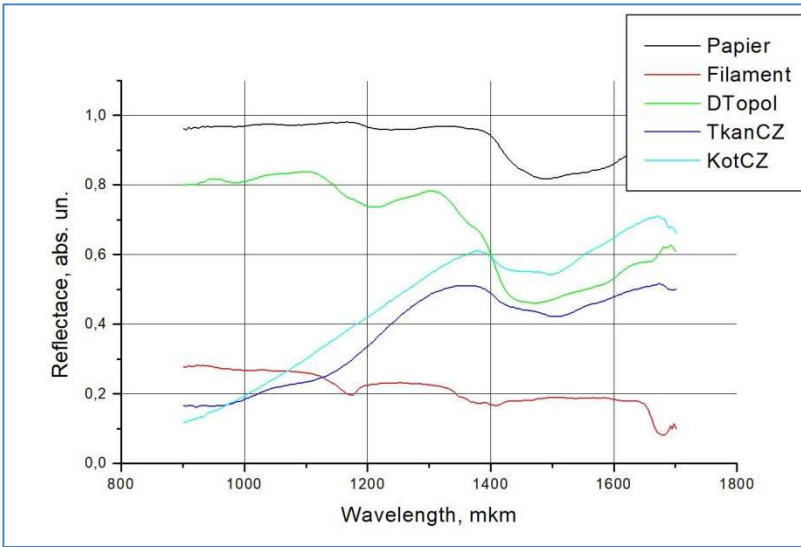
Nr	Opis obiektu	Temperatura systemu, °C	Temperatura detektora widma, °C	Parametr warunkowy lampy, „Lamp PD”	Wilgotność otoczenia, %	Uwagi
1	Sierść kota	31,65	27,53	3813	43,77	Struktura 3D
2	Tkanina	30,27	27,93	3814	36,4	Makroskopowo płaska
3	Papryka niedojrzała	29,61	25,55	3817	42,23	Makroskopowo nie płaska
4	Papryka dojrzała	28,39	24,84	3820	45,35	Makroskopowo nie płaska
5	Chleb świeży	31,64	27,64	3812	36,65	Struktura chropowata

6	Deska topola	31,36	28,14	3812	31,36	Makroskopowo płaska (gładka)
7	Kartka papieru	30,71	26,61	3811	30,19	Makroskopowo płaska
8	Kapsułka tranu	31,37	27,73	3809	31,37	Makroskopowo nie płaska
9	Filament 1 mm czerwony	30,35	26,16	3811	37,63	Makroskopowo płaska (gładka)
10	Wosk pszczeli	38,64	24,68	3812	36,64	Struktura z powietrzną makroskopowo nierówna
11	Borowik suszony	31,59	27,75	3812	31,59	Struktura z powietrzną nierówna

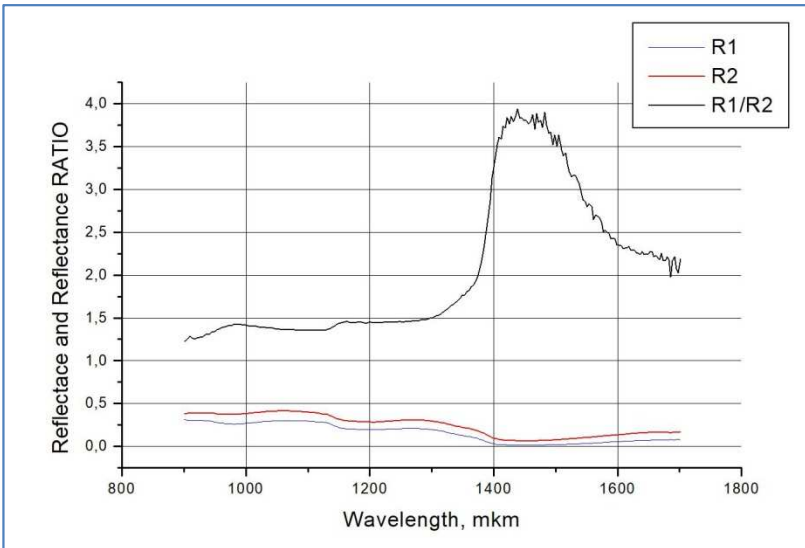
Na rys. 8-9 przedstawiono wyniki pomiarów widma reflektancji i absorpcji testowej serii obiektów badań.



Rys. 8. Widmo reflektancji próbek serii 1: Chleb – chleb jasny, Tran – kapsułka tranu, Borow – borowik suszony, Wosk – wosk pszczeli.



Rys. 9. Widmo reflektancji próbek serii 2: Papier – papier biały, Filament – filament o grubości 1 mm czerwony gładki, DTopol – deska topola sucha, TkanCZ – tkanina czarna, KotCZ – sierść kota czarna



Rys. 10. Widmo reflektancji papryki niedojrzałej (R1), papryki dojrzałej (R2) oraz widmo parametru $R2(\lambda)/R1(\lambda)$.

Przedstawione wyniki ilustrują różnorodność możliwości zastosowania ekspres-metody pomiarowej do różnorodnych obiektów. Ilustrację zastosowania metody do analizy zmian widma „w dziedzinie” czasu wybranego obiektu przedstawia rys. 10, na którym przedstawiono widma refleksyjności papryki niedojrzałej $R1(\lambda)$, dojrzałej $R2(\lambda)$ oraz wynik obliczenia widma stosunku parametrów $R2(\lambda)/R1(\lambda)$.

5. Podsumowanie

Jak wynika z przeprowadzonej analizy, urządzenie typu DPL NIR Scan może być zastosowane w laboratoriach oraz w warunkach mobilnych do analizy różnych materiałów.

Jako testowa została wybrana seria obiektów badawczych w postaci głównie obiektów organicznych, głównie z powodu dostępu do dużej bazy znanych porównawczych widm różnych obiektów tego rodzaju. Metoda pomiarowa i zastosowane urządzenie z powodzeniem może być zastosowane również do badania różnorodnych cienkowarstwowych materiałów półprzewodnikowych, ze względu na zakres pomiarowy. Badania wstępne już wykonane na innych obiektach, w tym z grupy materiałów stosowanych w elektronice, m.in. filmów półprzewodnikowych i nanokompozytowych, rokuje perspektywistyczność poszerzenia grupy obiektów pomiarowych też i na materiały z tej grupy.

Za pomocą urządzenia można tworzyć bazy danych pomiarowych, które później mogą być wykorzystywane do diagnostyki badanych obiektów, w tym zmian w funkcji czasu lub na etapach obróbki technologicznej.

Bibliografia

1. Handbook of Spectroscopy, Volume 1. Edited by Günter Gauglitz and Tuan Vo-Dinh. Copyright © 2003 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim ISBN 3-527-29782-0
2. Near-Infrared Spectroscopy. Principles, Instruments, Applications, ed. Siesler H. W., Ozaki Y., Kawata S., Heise H. M., Wiley-VCH, Weinheim 2001.
3. Burns D. A., Ciurczak E. W., Handbook of Near-Infrared Analysis, Marcel Dekker, New York 1992.
4. John Coates, A Review of New Small-Scale Technologies for Near Infrared Measurements, American Pharmaceutical Review 17(4) · 2014
5. Stefan Laske et. Al., In-line quality control of polymer nanocomposites using near-infrared spectroscopy, nanoconConf, Roznov pod Radhostem, Czech Republic, EU, 20. - 22. 10. 2009

6. J.M. Barbas, A.V. Machado, J.A. Covas, In-line near-infrared spectrometry for the characterization of dispersion in polimer-clay nanocomposites, *Polimer Tisting* v.31, 527-536, 2012
7. Grolik Benno, Kopp Joachim, Optical Properties of Thin Semiconductor Films, Raport Fermi Lab, 2003, https://home.fnal.gov/~jkopp/tum/pdf/F/hl_spekt.pdf
8. Z. Sultan and N. Sultana, Analysis of Reflectance and Transmittance Characteristics of Optical Thin Film for Various Film Materials, Thicknesses and Substrates, *J Electr Electron Syst.* 2015, V.4 (3), 2015, DOI: 10.4172/2332-0796.1000160
9. Yuan-Yuan Pu, Calibration Transfer from Micro NIR Spectrometer to Hyperspectral Imaging: a Case Study on Predicting Soluble Solids Content of Bananito Fruit (*Musa acuminata*), *Food Analitical Methods* (2018) 11:1021–1033, <https://doi.org/10.1007/s12161-017-1055-3>
10. Lei-Ming Yuan, et al., Nondestructive Measurement of Soluble Solids Content in Apples by a Portable Fruit Analyzer, *Food Analitical Methods*, Published online, 2015, DOI 10.1007/s12161-015-0251-2
11. Claudia Santos et al., A review on Application of Portable Near-Infrared Spectrometers in Agro-Food Industry, *Applied Spectroscopy*, 2013, 1215-1233, DOI: 10.1366/13-07228
12. M. Smiga-Kout, L. Bychto et al., Use of buckwheat straw to produce ethyl alcohol using ionic liquids, *Energy*, Volume 12, Issue 10, 2019, Article number 2014
13. I. Chandrasekaran, S.S. Panigrahi, L. Ravikanth, C.B. Singh, Potential of Near-Infrared (NIR) Spectroscopy and Hyperspectral Imaging for Quality and Safety Assessment of Fruits: an Overview, *Food Analytical Methods*, (2019) 12, 2438–2458, doi.org/10.1007/s12161-019-01609-1
14. P. Gelabert, E. Pruet, et al., DLP NIRscan Nano: An ultra-mobile DLP-based near-infrared Bluetooth spectrometer, *Proceedings of SPIE - The International Society for Optical Engineering* Volume 9761, 2016, Article number 97610B
15. P. Gelabert. DLP NIRscan Nano EVM User's Guide User's Guide; Texas Instrument, TI, 2015

Streszczenie

Zmodyfikowana wersja minispektrometru typu DPL NIR Scan została testowo zastosowana w laboratoriach oraz w warunkach mobilnych do analizy różnych materiałów.

Jako testowa została wybrana seria obiektów badawczych w postaci głównie obiektów organicznych. Metoda pomiarowa i zastosowane urządzenie z powodzeniem może być zastosowane również do badania różnorodnych

cienkowarstwowych materiałów półprzewodnikowych, ze względu na zakres pomiarowy.

Badania wstępne już wykonane na innych obiektach, w tym z grupy materiałów stosowanych w elektronice, m.in. filmów półprzewodnikowych i nanokompozytowych, rokuje perspektywistyczność poszerzenia grupy obiektów pomiarowych też i na materiały z tej grupy.

Za pomocą urządzenia można tworzyć bazy danych pomiarowych, które później mogą być wykorzystywane do diagnostyki badanych obiektów, w tym zmian w dziedzinie czasu lub na etapach obróbki technologicznej.

Abstract

A modified version of the DPL NIR Scan type mini-spectrometer was tested in laboratories and in mobile conditions for the analysis of various materials. A series of research objects in the form of mainly organic objects was chosen as the testing probes. The measuring method and the device used can also be successfully used for testing various thin-film semiconductor materials, due to the measuring range. Preliminary tests already performed on other objects, including from the group of materials used in electronics, among others semiconductor and nanocomposite films, promising the prospect of expanding the group of measurement objects also for materials from this group. Using the device, it is possible to create measurement databases that can later be used to diagnose the examined objects, including changes in the time domain or at the stages of technological processing.

Keywords: DLP, NIRScan, Nano EVM, measurement mobility, spectroscopy

Mirosław Maliński
Katedra Elektroniki
Wydział Elektroniki i Informatyki
Politechnika Koszalińska
miroslaw.malinski@tu.koszalin.pl

Metody fototermiczne i ich zastosowania - Wykład inauguracyjny roku akademickiego 2019/2020 na Wydziale Elektroniki i Informatyki Politechniki Koszalińskiej

Słowa kluczowe: fotoakustyka, metoda fototermiczna, modulowana absorpcja na swobodnych nośnikach, metoda piezoelektryczna, radiometria fototermiczna

1. Wprowadzenie

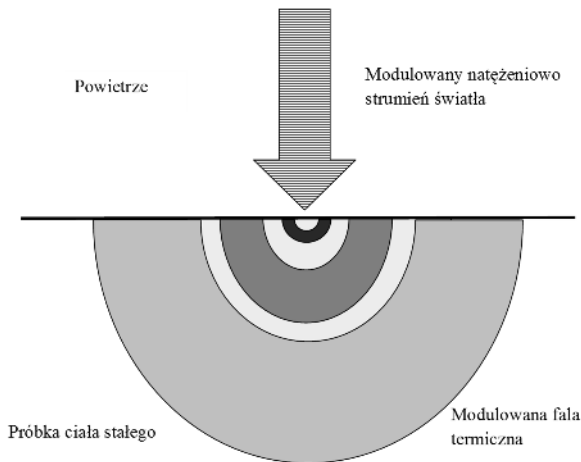
W pracy opisano szereg zagadnień związanych z zastosowaniem metod fototermicznych w metrologii materiałów półprzewodnikowych. Po krótkim wprowadzeniu w podstawy fizyczne zjawiska fototermicznego opisano wybrane rodzaje detekcji sygnałów fototermicznych oraz przedstawiono metody pomiarowe stosowane w metodzie fototermicznej. Praca przedstawia także opis różnych zastosowań fototermiki oraz stanowiska pomiarowe metod fototermicznych wykorzystywane w Katedrze Elektroniki Wydziału. Dla ilustracji zagadnienia przedstawiono przykładowe charakterystyki doświadczalne uzyskane opisanymi metodami fototermicznymi. Na zakończenie przedstawiono związek fototermiki z automatyką, elektroniką i elektrotechniką, a więc dyscypliną naukową uprawianą na Wydziale w Katedrze Elektroniki, a związaną ściśle z kierunkami kształcenia prowadzonymi na Wydziale takimi jak elektronika i telekomunikacja oraz informatyka.

2. Zjawisko fototermiczne

Fototermika jest dziedziną nauki, która wykorzystuje zjawisko absorpcji światła w oświetlanych materiałach i zmianę pochłoniętej w nich energii świetlnej na energię cieplną. Energia cieplna rejestrowana jest jako zmiana temperatury badanego ciała. Temperatura ta niesie informację o wielu procesach i parametrach badanych ciał takich jak: parametry termiczne, optyczne, rekombinacyjne, wydajnościowe.

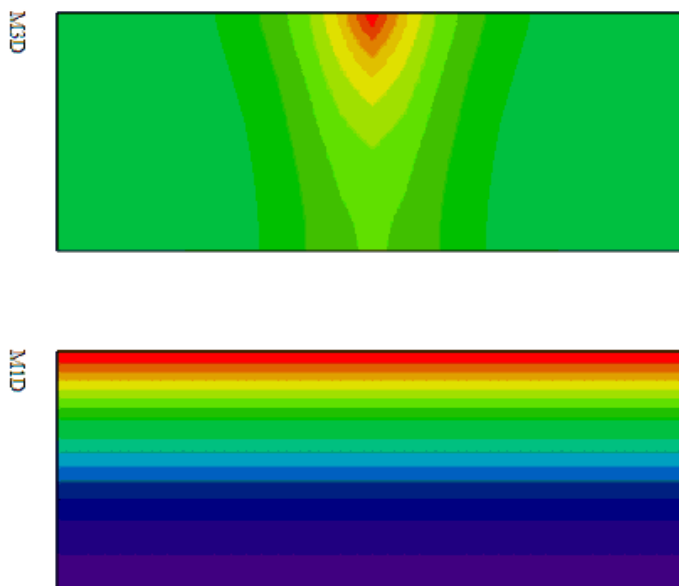
Fototermika jest więc działem metrologii. Pod pojęciem fototermiki, jako dziedziny techniki, rozumie się również zmianę energii słonecznej na energię cieplną jaka ma miejsce w kolektorach słonecznych w systemach grzewczych. Zagadnienie to nie jest jednak opisywane w niniejszej pracy.

Idea zjawiska fototermicznego jest przedstawiona schematycznie na Rys. 1. Modulowany natężeniowo strumień światła oświetla próbkę ciała stałego. W wyniku absorpcji promieniowania i zmianę energii świetlnej na energię termiczną, w wyniku rekombinacji bezpromienistej nośników, w próbce rozchodzi się modulowana fala termiczna.



Rys. 1. Idea zjawiska fototermicznego

Rozkład przestrzenny amplitudy fali termicznej może przyjmować różne formy zależnie od wielu parametrów. Przykładowe rozkłady amplitudy fali termicznej przedstawiono na Rys. 2.



Rys. 2. Rozkład przestrzenny amplitudy fali termicznej rozchodzącej się w próbce w przypadku wąskiej i szerokiej wiązki światła oświetlającej próbkę. Wg [1].

3. Rodzaje detekcji i metody pomiarowe

W metodzie fototermicznej stosuje się szereg rodzajów detekcji zależnie od wielkości fizycznej, która ma być rejestrowana, zależnej od rozchodzącej się fali termicznej w badanym materiale.

Jeżeli jest rejestrowana zmiana ciśnienia w komorze fotoakustycznej, będąca wynikiem zmiany temperatury powierzchni próbki, stosuje się metodę fotoakustyczną z detekcją mikrofonową.

Jeżeli rejestruje się zmiany kształtu próbki, będące wynikiem jej rozszerzania termicznego, stosuje się metodę piezoelektryczną z detekcją przetwornikiem piezoelektrycznym.

Jeżeli rejestruje się promieniowanie podczerwone emitowane przez próbkę, będące efektem okresowych zmian temperatury w całej objętości próbki, stosuje się detekcję czujnikiem na podczerwień.

Jeżeli rejestruje się okresową składową transmisji światła podczerwonego, będącą wynikiem okresowej zmiany koncentracji nośników w objętości próbki, to stosuje się metodę modulacji absorpcji na swobodnych nośnikach.

Do innych rodzajów detekcji i metod pomiarowych stosowanych w metodzie fototermicznej zalicza się: metodę piroelektryczną, metodę mirażu oraz metodę termoodbiciową.

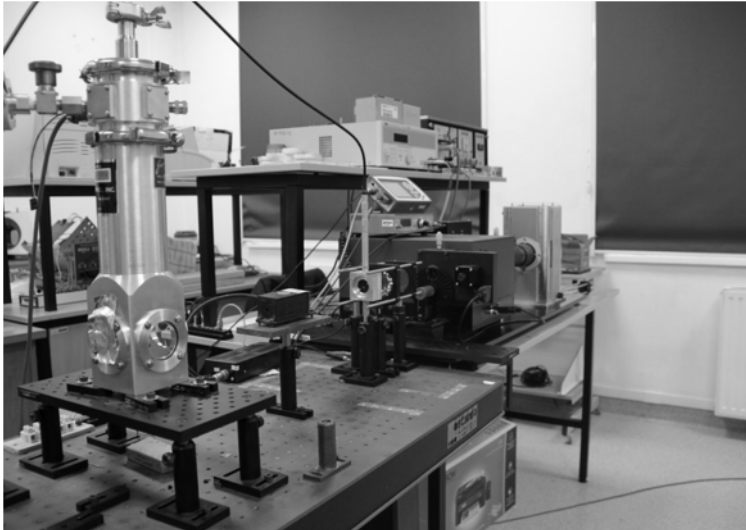
4. Zastosowania fototermiki

Metody fototermiczne podane powyżej stosuje się między innymi do:

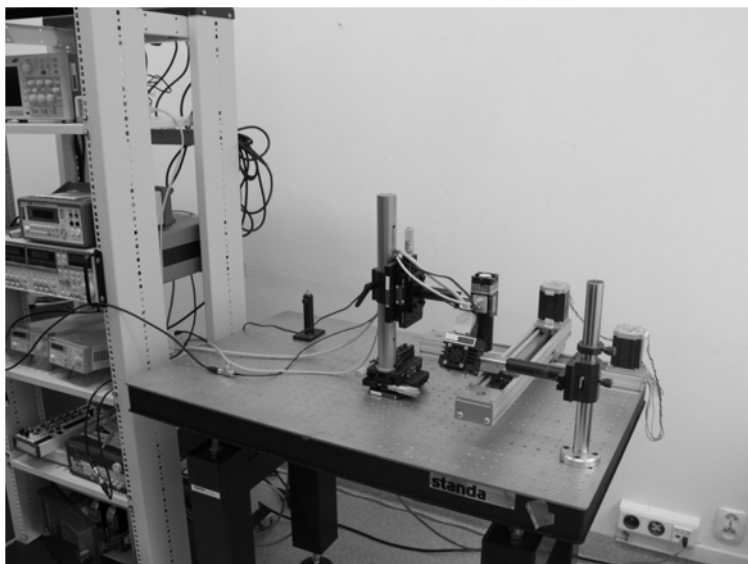
- Wyznaczania parametrów fizycznych materiałów i urządzeń takich jak: parametry cieplne, parametry optyczne i widma optyczne oraz parametry rekombinacyjne.
- Wyznaczania rozkładów przestrzennych parametrów materiałowych takich przykładowo jak czas życia nośników.
- Wyznaczania wydajności kwantowej i energetycznej fotoluminescencji oraz elektroluminescencji.
- Wyznaczanie wydajności energetycznej konwersji światła na prąd.

5. Stanowiska pomiarowe fototermiczne

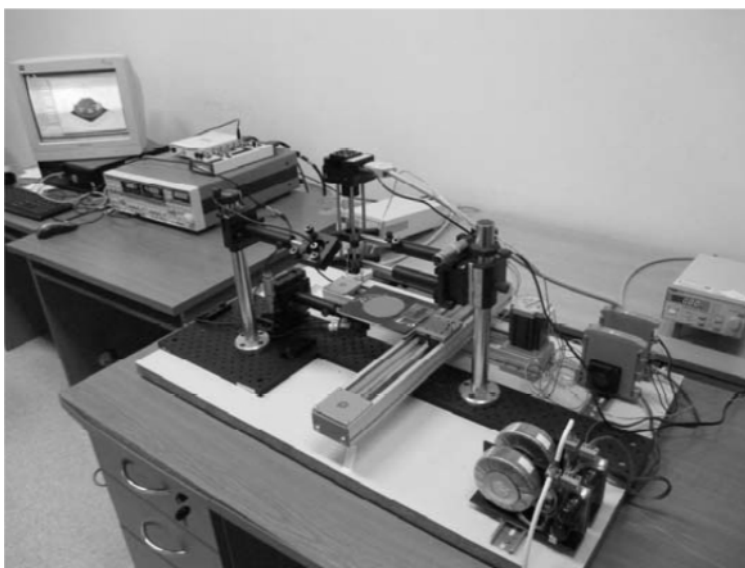
Na zdjęciach poniżej przedstawiono stanowiska pomiarowe Katedry Elektroniki stosowane do pomiarów fotoakustycznych (Rys. 3), do pomiarów radiometrycznych w podczerwieni (Rys. 4), do pomiarów rozkładu przestrzennego modulowanej absorpcji na swobodnych nośnikach oraz prądów indukowanych wiązką światła (Rys. 5).



Rys. 3. Stanowisko do badań widm fotoakustycznych



Rys. 4. Stanowisko do badań charakterystyk radiometrii fototermicznej PTR

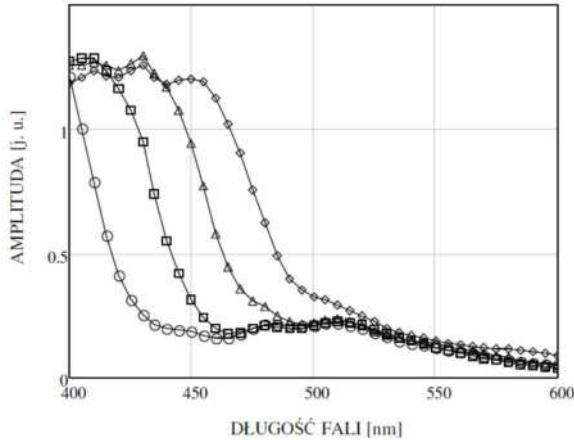


Rys. 5. Stanowisko do badań rozkładu modulowanej absorpcji na nośnikach swobodnych MFCA oraz LBIC

6. Przykładowe charakterystyki uzyskiwane metodami fototermicznymi

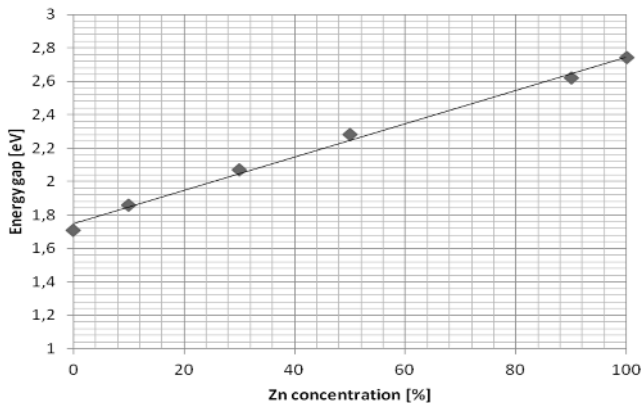
Poniżej przedstawiono przykładowe charakterystyki doświadczalne uzyskiwane metodami fototermicznymi.

Przykładowe widma doświadczalne fotoakustyczne 'grubych' kryształów mieszanych przedstawiono na Rys. 6.



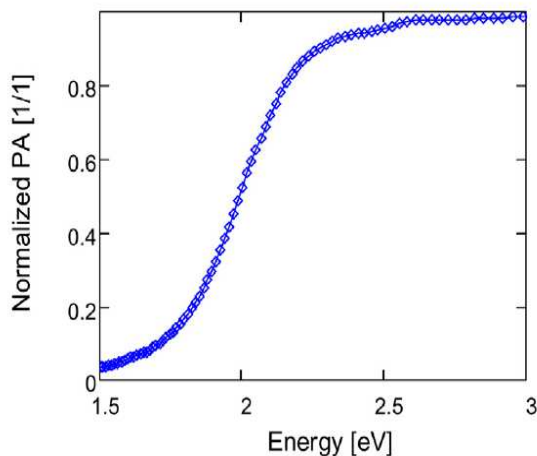
Rys. 6. Widma fotoakustyczne kryształów mieszanych $Zn_{1-x-y}Be_xMn_ySe$ dla różnej zawartości jonów Be. Wg [2].

Zależność wielkości przerwy energetycznej 'grubych' kryształów mieszanych $CdZn_xSe_{1-x}$ od koncentracji Zn przedstawiono na Rys. 7.

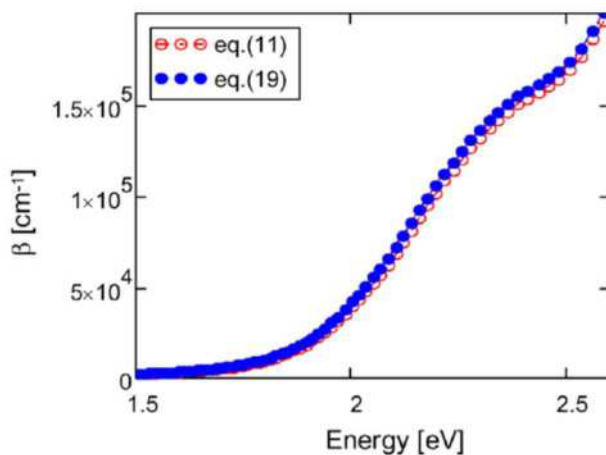


Rys. 7. Zależność wielkości przerwy energetycznej od składu kryształu mieszanego $CdZn_xSe_{1-x}$ Wg [3].

Widma fotoakustyczne oraz widma współczynnika absorpcji optycznej 'cienkich' warstw półprzewodnikowych In_2S_3 przedstawiono odpowiednio na Rys. 8 i 9.

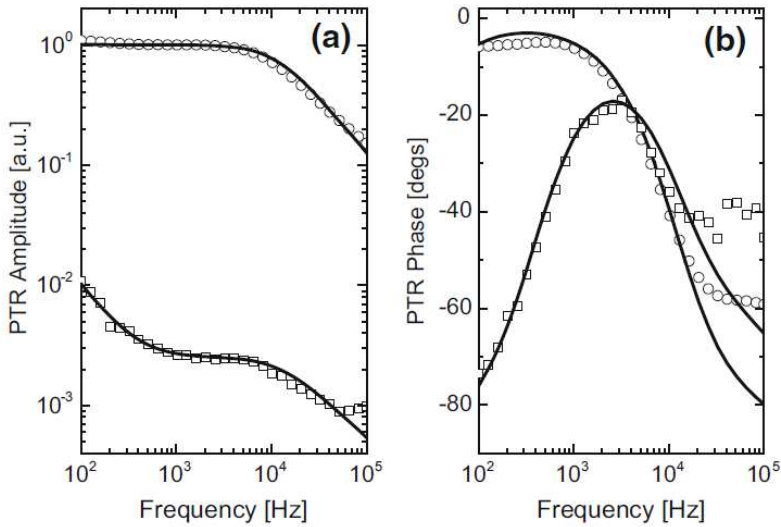


Rys. 8. Widmo fotoakustyczne cienkiej warstwy $d=0.17 \mu\text{m}$ In_2S_3 . Wg [4].



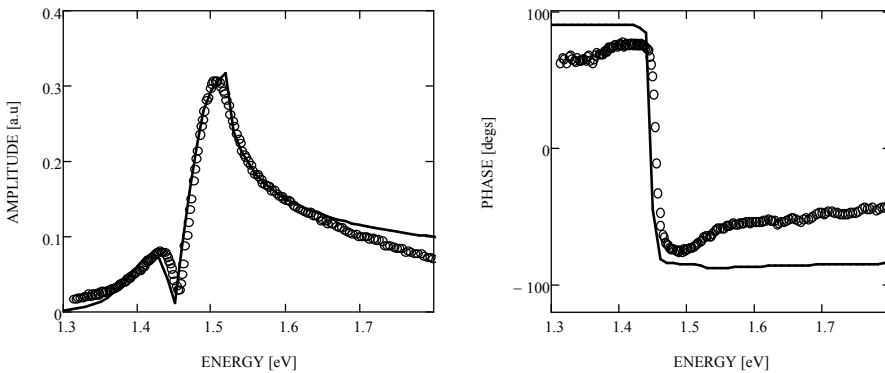
Rys. 9. Widmo współczynnika absorpcji optycznej cienkiej warstwy, $d=0.17 \mu\text{m}$, In_2S_3 . Wg [4].

Przykładowe charakterystyki doświadczalne uzyskane w metodzie radiometrii fototermicznej przedstawiono na Rys. 10.



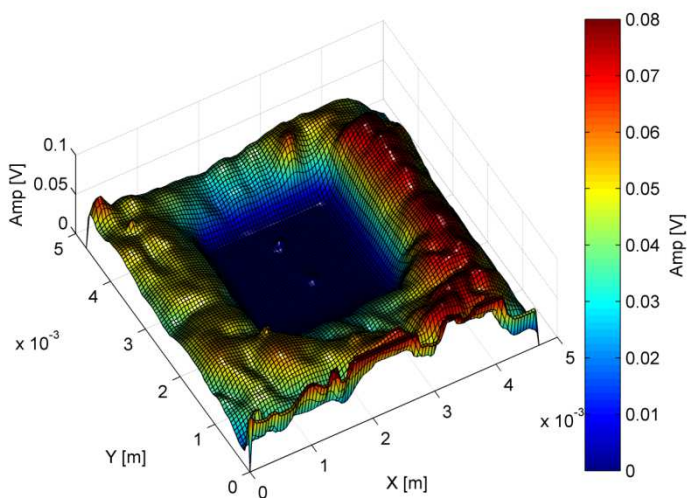
Rys. 10. Charakterystyki sygnału PTR uzyskane dla próbek Si przed i po implantacji wysoko energetycznymi jonami O^{+6} . Wg [5].

Przykładowe charakterystyki doświadczalne uzyskane w detekcji piezoelektrycznej przedstawiono na Rys. 11.



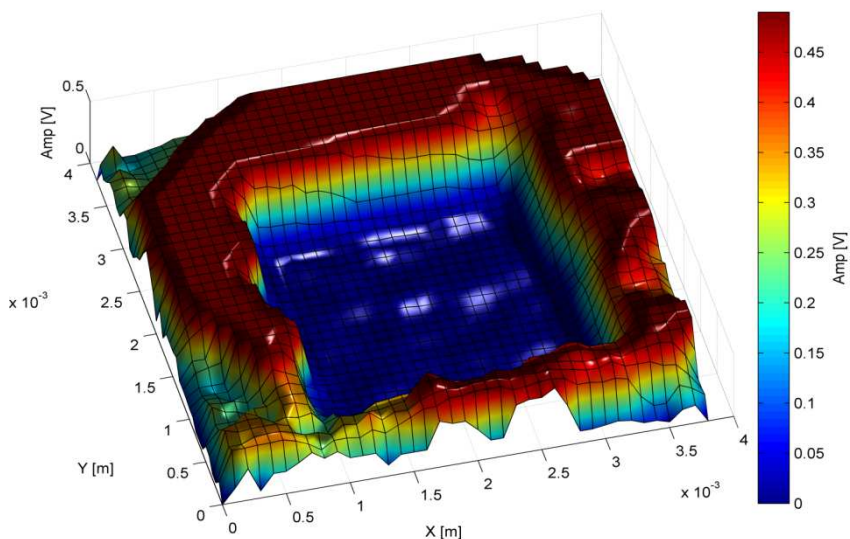
Rys. 11. Widma piezoelektryczne amplitudowe i fazowe kryształu CdTe – model dwuwarstwowy. Grubość warstwy zdefektowanej $d=0.5 \mu\text{m}$. Wg [6].

Przykładowe charakterystyki doświadczalne rozkładu przestrzennego sygnału PTR uzyskane doświadczalnie przedstawiono na Rys. 12.



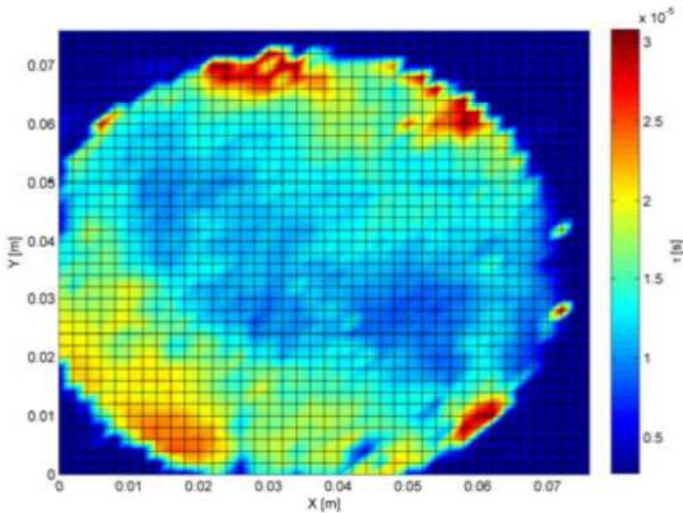
Rys. 12. Rozkład przestrzenny sygnału PTR uzyskany na próbce Si implantowanego Au^{2+} $\lambda_{\text{ex}} = 480 \text{ nm}$, $f=1\text{kHz}$. Wg [7].

Przykładowe rozkłady przestrzenne sygnału MFCA przedstawiono na Rys. 13.



Rys. 13. Rozkład przestrzenny sygnału MFCA uzyskany na próbce Si implantowanej jonami Mn^+ . Wg [8].

Przykładowy rozkład przestrzenny czasu życia nośników uzyskany na płycie krzemowej metodą MFCA przedstawiono na Rys. 14.



Rys. 14. Rozkład przestrzenny czasu życia nośników uzyskany na płycie krzemowej metodą MFCA. Wg [9].

7. Związek fototermiki z elektroniką, automatyką i informatyką

Metody fototermiczne mają silny związek z elektroniką, automatyką i informatyką co przedstawiono poniżej.

- Czujniki stosowane w badaniach są elementami elektronicznymi: mikrofony, przetworniki piezoelektryczne, detektory podczerwieni
- Mierzone sygnały są bardzo małe wymagające przedwzmacniaczy i technik fazoczułych
- Źródła światła to najczęściej lasery sterowane sygnałami elektrycznymi
- Pomiary są sterowane komputerowo z wykorzystaniem oprogramowania sterującego: stoliki X-Y, monochromatory, wzmacniacze fazoczułe, modulatory
- Wyniki pomiarów zapisywane są w komputerach
- Wyniki pomiarów są następnie obrabiane komputerowo celem wyciągnięcia z nich użytecznych informacji.

Wymaga to przeprowadzenia szeregu działań:

- stworzenia modeli badanych zjawisk w oparciu o ich analizę fizyczną
- napisania dedykowanych arkuszy obliczeniowych w Matlabie, Mathcadzie czy też aplikacji w C++ czy Javie
- przeprowadzenia dopasowań krzywych teoretycznych do krzywych doświadczalnych, w których parametrem dopasowania jest szukany parametrem materiałowym.

8. Podsumowanie

Proces badawczy w fototermice jest więc mariażem szeregu dyscyplin naukowych:

- fizyki
- elektroniki
- automatyki
- informatyki
- matematyki

Wszystkie te dyscypliny poznają studenci w trakcie studiów na Wydziale Elektroniki i Informatyki Politechniki Koszalińskiej.

Bibliografia

1. L. Bychto, M. Maliński, „*ID or 3D spatial distribution – comparison*”, Journal de Physique IV 117(2004) 17-23
2. M. Maliński, Ł. Chrobak, J. Zakrzewski, K. Strzałkowski, „*Determination of the Quantum Efficiency of Luminescence in Mn²⁺ Ions in Zn_{0.75} Be_{0.20} Mn_{0.05} Se Crystals by the Nondestructive Photoacoustic Method*”, Optical Materials 33 (2010) 75-78.
3. J. Zakrzewski, K. Strzałkowski, M. Maliński, „*Piezoelectric Photothermal Spectroscopy of CdZnSe Mixed Crystals*”, Proceedings 20th ICPPP Moscow (2019).
4. L. Bychto, M. Maliński, „*Photoacoustic spectroscopy analysis of thin semiconductor samples*”, Opto-Electronics Review 26 (2018) 217-222
5. M. Maliński, M. Pawlak, Ł. Chrobak, S. Pal, A. Ludwig, „*Monitoring of amorfization of the oxygen implanted layers in silicon wafers using photothermal radiometry and modulated free carrier absorption methods*”, Applied Physics A 118 (2015) 1009-1014.

6. J. Zakrzewski, K. Strzałkowski, M. Maliński, Ł. Chrobak, „*Two-layer model in piezoelectric photothermal spectra of CdTe crystals*”, International Journal of Thermophysics 40 (2019) 56-66.
7. Ł. Chrobak, M. Maliński, „*On investigations of the optical absorption coefficient of gold and germanium implanted silicon with the use of the nondestructive contactless photothermal infrared radiometry*”, Journal of Electronic Materials 48(8) (2019) 5273-5278
8. Ł. Chrobak, M. Maliński, „*Properties of Silicon Implanted with Fe⁺, Ge⁺, Mn⁺ Ions Investigated Using a Frequency Contactless Modulated Free-Carrier Absorption Technique*”, Optical Materials 86 (2018) 484-491.
9. M. Maliński, L. Bychto, Ł. Chrobak, W. Madej, „*Mapowanie czasu życia nośników w materiałach półprzewodnikowych z wykorzystaniem techniki modulacji absorpcji na swobodnych nośnikach*”, Przegląd Elektrotechniczny 91(9) (2015) 113-116.

Streszczenie

Praca przedstawia metody fototermiczne jakie są stosowane w badaniach materiałów półprzewodnikowych w Katedrze Elektroniki Wydziału Elektroniki i Informatyki Politechniki Koszalińskiej. W pracy przedstawiono również badawcze stanowiska fototermiczne w Katedrze, a także uzyskane wybrane charakterystyki doświadczalne.

Abstract

This paper presents photothermal methods which are used in the research of semiconductor materials in the Department of Electronics of the Faculty of Electronics and Computer Science of the Koszalin University of Technology. In this paper the experimental photothermal research set ups of the Department as also the chosen obtained experimental characteristics are presented.

Keywords: photoacoustics, photothermal method, modulated free carriers absorption modulowana, piezoelectric method, photothermal radiometry.

Patryk Widuliński
Wydział Elektroniki i Informatyki
Politechnika Koszalińska
patryk.widulinski@tu.koszalin.pl

Badanie wykrywalności anomalii w pliku monitorowanym przez system wykrywania intruzów w zależności od parametrów generacji receptorów

Słowa kluczowe: sztuczny system immunologiczny, receptor, anomalia, szkodliwe oprogramowanie

1. Wstęp

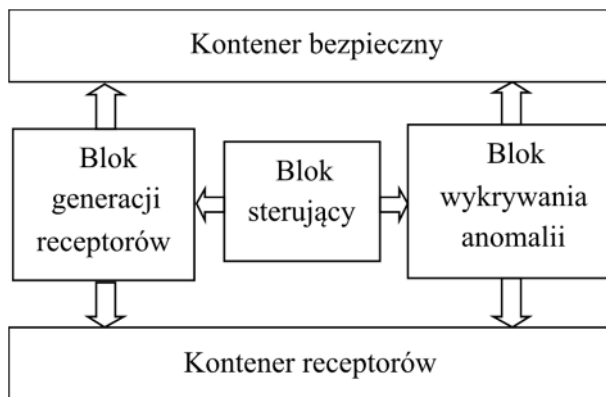
Wykrywanie zagrożeń w systemie operacyjnym jest popularnym wśród naukowców zagadnieniem związanym z bezpieczeństwem. Konstruktorzy oprogramowania wykrywającego włamania stale poszukują nowych rozwiązań, które pozwoliłyby na coraz skuteczniejszą detekcję zagrożeń. Jednym z nowatorskich sposobów detekcji zagrożeń jest zastosowanie sztucznego systemu immunologicznego w systemie operacyjnym [1-3]. Systemy takie inspirowane są ludzkim układem odpornościowym. Przykłady takich systemów opisane są w [4-13]. Jednym z algorytmów implementowanych w sztucznych systemach immunologicznych jest algorytm negatywnej selekcji opisany w [11], który ogólnie opiera się na generacji ciągów zwanych receptorami. Receptory są ciągami binarnymi posiadającymi zdolność wykrywania struktur obcych.

W pracy [15] zaproponowano system wykrywania intruzów oparty na algorytmie negatywnej selekcji z wykorzystaniem wzorców [14]. Do poprawnego działania algorytm generacji receptorów z [15] wymaga podania parametrów generacji, takich jak liczba bitów receptora i próg aktywacji. W niniejszym artykule zaprezentowano, przeanalizowano i podsumowano wyniki badań wykrywalności anomalii dla metody wzorców w zależności od liczby bitów receptora, progu aktywacji i rozmiaru anomalii.

Artykuł zorganizowany jest w następujący sposób. Badany system omówiono w rozdziale 2. System przebadano eksperymentalnie i przedstawiono wyniki w rozdziale 3. Analiza wyników znajduje się w rozdziale 4, a w rozdziale 5 przedstawiono wnioski.

2. System wykrywania intruzów

Badany system wykrywania intruzów (SWI), zaprezentowany w [15], pozwala na wykrywanie nieregularności (anomalii, modyfikacji, intruzów) w programach napisanych dla systemu operacyjnego Windows. SWI monitoruje określoną lokalizację w systemie operacyjnym zawierającą pliki do ochrony. Na początku działania system generuje ciągi zwane receptorami na podstawie niezmodyfikowanych, poprawnych wersji plików. Receptory są ciągami binarnymi o długości l bitów posiadającymi zdolność wykrywania struktur obcych. Następnie system skanuje ciągle wszystkie monitorowane pliki w poszukiwaniu anomalii.



Rys. 1. Schemat systemu wykrywania intruzów

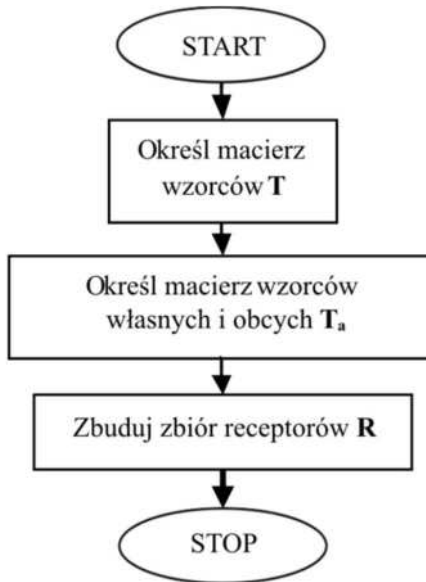
SWI przedstawiono na Rys. 1. System składa się z trzech głównych bloków operacyjnych: bloku sterującego BS, bloku generacji receptorów BGR i bloku wykrywania anomalii BWA. W systemie obecne są także dwa kontenery zawierające odpowiednio pliki monitorowane (kontener bezpieczny) i receptory wykorzystywane do detekcji anomalii. Blok sterujący nadzoruje pracę bloków BGR i BWA i ich dostęp do kontenerów.

2.1. Blok BGR

Blok BGR posiada dwa wbudowane algorytmy generacji: losową i wzorców (szablonów). W niniejszym artykule wykorzystano do badań metodę wzorców opisaną w [14]. Wzorce są ciągami binarnymi o długości l zawierającymi bity

istotne i nieistotne. Liczbę bitów istotnych określa parametr zwany progiem aktywacji m . Liczba bitów nieistotnych wynosi więc $l - m$.

Schemat blokowy generacji receptorów przy użyciu metody wzorców zaprezentowano na Rys. 2. W pierwszym kroku określa się macierz wzorców T . Macierz T zawiera wszystkie możliwe kombinacje bitów istotnych, a ich liczba wynosi 2^m . Dla każdej kombinacji bitów istotnych (dla każdego wiersza) rozpisywane są w poszczególnych kolumnach wszystkie możliwe kombinacje pozycji bitów nieistotnych w ramach długości wzorca l . Bity nieistotne w macierzy T oznacza się gwiazdką ("*").



Rys. 2. Schemat blokowy działania algorytmu generacji receptorów na podstawie wzorców

Na podstawie T powstaje następnie macierz pomocnicza T_a zawierająca informację o przynależności ciągów binarnych wzorców do zbiorów wzorców własnych (*self templates*) i wzorców obcych (*nonself templates*). Przynależność ta określana jest na podstawie macierzy T i treści monitorowanego, niezmodyfikowanego pliku.

Na podstawie macierzy T i T_a powstaje ostatecznie zbiór receptorów R , zdolnych do wykrywania struktur obcych. Zbiór R umieszczany jest w kontenerze receptorów.

2.2. Blok BWA

Blok BWA zawiera implementację algorytmu wykrywania anomalii. Schemat blokowy działania BWA przedstawiono na Rys. 3. Algorytm odczytuje fragment o długości l bitów z aktualnej wersji monitorowanego pliku, a następnie porównuje ze wszystkimi możliwymi receptorami z kontenera. W przypadku dopasowania m kolejnych bitów między dowolnym receptorem a fragmentem programu algorytm raportuje wykrycie anomalii.



Rys. 3. Schemat blokowy działania algorytmu generacji receptorów na podstawie wzorców

3. Badania eksperymentalne

Do badań wykorzystano pojedynczy monitorowany plik wykonywalny „sample2.exe” o rozmiarze 9216 bajtów. Badaną wielkością jest procentowa wykrywalność anomalii w pliku. Rozmiar anomalii a mierzony jest w bajtach. Metodą generacji receptorów wykorzystaną do badań jest metoda wzorców, przyjmująca jako parametry liczbę bitów receptora l i próg aktywacji m . Badania przeprowadzono dla $l \in \{16, 32\}$, $m \in \{7, 8, 9, 10\}$ i $a \in \{1, 2, 3, 4, 5, 6, 7, 8\}$. Dla każdej kombinacji l i m wygenerowany został nowy zbiór receptorów \mathbf{R} . Liczba receptorów w zbiorze oznaczona jest przez R_n . Pojedynczy test polegał na wstrzyknięciu anomalii o rozmiarze a w losowym miejscu w programie, a następnie przeskanowaniu go przy pomocy algorytmu wykrywania anomalii. Dla każdego rozmiaru anomalii a przeprowadzono 1000 testów.

Wyniki badań przedstawiono w Tabelach 1 – 7. Oznaczenia w tabelach są następujące: l – liczba bitów receptora, m – próg aktywacji receptora w bitach, TGT – czas generacji macierzy \mathbf{T} i \mathbf{T}_a wzorców w milisekundach, RGT – czas generacji receptorów, R_n – liczba wygenerowanych prawidłowych receptorów, M_{rec} – zajętość pamięci przez receptory w bajtach wyliczona ze wzoru:

$$M_{rec} = R_n \cdot l \div 8, \quad (1)$$

a – rozmiar anomalii w bajtach, R_a – średnia liczba aktywowanych receptorów, DT_{avg} – średni czas wykrycia anomalii w milisekundach, D – wykrywalność w procentach.

Dla przypadku $l = 16$, $m = 7$ nie został wygenerowany ani jeden receptor, więc żadna z anomalii nie została wykryta. Z tego powodu nie zamieszczono osobnej tabeli wyników dla tej kombinacji (czas generacji wzorców wyniósł 76 ms, a czas generacji receptorów 2 ms).

Tabela 1. Wykrywalność dla $l = 16$ i $m = 8$

l	16	m	8
TGT [ms]	431	RGT [ms]	58
R_n	38	M_{rec} [B]	76
a	R_a	DT_{avg} [ms]	D
1	7	6	16,4%
2	12	12	30,0%
3	12	16	38,8%
4	11	21	50,6%
5	17	24	57,5%
6	14	26	63,2%
7	15	29	69,2%
8	15	32	76,1%
\acute{S} rednia	12,88	20,75	50,2%

Tabela 2. Wykrywalność dla $l = 16$ i $m = 9$

l	16	m	9
TGT [ms]	873	RGT [ms]	3249
R_n	269	M_{rec} [B]	538
a	R_a	DT_{avg} [ms]	D
1	9	127	43,1%
2	13	207	70,3%
3	16	238	81,0%
4	17	260	88,4%
5	20	278	94,5%
6	23	284	96,6%
7	27	288	98,1%
8	27	290	98,6%
\acute{S}rednia	19	246,5	83,8%

Tabela 3. Wykrywalność dla $l = 16$ i $m = 10$

l	16	m	10
TGT [ms]	1624	RGT [ms]	72942
R_n	1006	M_{rec} [B]	2012
a	R_a	DT_{avg} [ms]	D
1	9	762	60,0%
2	16	1031	82,2%
3	18	1147	92,8%
4	24	1204	96,9%
5	24	1232	98,7%
6	28	1256	99,8%
7	30	1256	99,4%
8	34	1269	100,0%
\acute{S}rednia	22,88	1144,63	91,2%

Tabela 4. Wykrywalność dla $l = 32$ i $m = 7$

l	32	m	7
TGT [ms]	384	RGT [ms]	10
R_n	9	M_{rec} [B]	36
a	R_a	DT_{avg} [ms]	D
1	9	0	5,7%
2	9	1	10,6%
3	9	1	15,4%
4	13	2	22,5%
5	18	2	22,9%
6	18	3	28,7%
7	18	3	31,6%
8	27	3	38,2%
\acute{S}rednia	15,13	1,88	22,0%

Tabela 5. Wykrywalność dla $l = 32$ i $m = 8$

l	32	m	8
TGT [ms]	724	RGT [ms]	676
R_n	103	M_{rec} [B]	412
a	R_a	DT_{avg} [ms]	D
1	23	43	36,5%
2	27	69	58,4%
3	32	84	71,7%
4	35	95	81,0%
5	37	104	88,2%
6	42	104	88,4%
7	45	113	94,1%
8	51	115	96,9%
\acute{S}rednia	36,5	90,88	76,9%

Tabela 6. Wykrywalność dla $l = 32$ i $m = 9$

l	32	m	9
TGT [ms]	986	RGT [ms]	30308
R_n	451	M_{rec} [B]	1804
a	R_a	DT_{avg} [ms]	D
1	22	342	66,9%
2	31	454	88,5%
3	30	496	96,8%
4	38	507	99,0%
5	44	511	99,6%
6	45	514	100,0%
7	47	514	100,0%
8	51	514	100,0%
\acute{S}rednia	38,5	481,5	93,9%

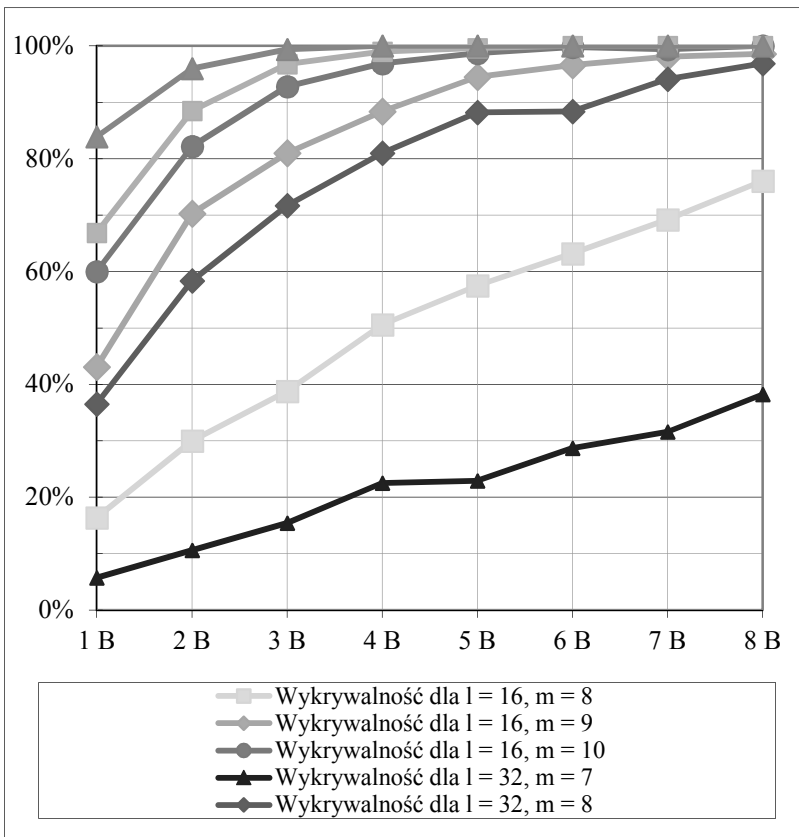
Tabela 7. Wykrywalność dla $l = 32$ i $m = 10$

l	32	m	10
TGT [ms]	2026	RGT [ms]	457063
R_n	1336	M_{rec} [B]	5344
a	R_a	DT_{avg} [ms]	D
1	25	1302	83,9%
2	25	1483	96,0%
3	32	1542	99,4%
4	34	1662	100,0%
5	41	1662	100,0%
6	47	1662	100,0%
7	49	1662	100,0%
8	52	1662	100,0%
\acute{S}rednia	38,13	1579,63	97,4%

4. Analiza wyników badań

Dla każdej kombinacji $l \in \{16, 32\}$, $m \in \{7, 8, 9, 10\}$ i $a \in \{1, 2, \dots, 8\}$ przeprowadzono 1000 testów, co daje łączną liczbę $1000 \cdot 2 \cdot 4 \cdot 8 = 64000$ testów wykrycia anomalii. Za „dobry” poziom wykrywalności anomalii przyjęto wykrywalność na poziomie 75% lub wyższym, a za „bardzo dobry” poziom wykrywalności przyjęto wykrywalność na poziomie co najmniej 90%.

Wykres wykrywalności w funkcji rozmiaru anomalii przedstawiono na Rys. 4. Dla $m = 7$ nie udało się wygenerować receptorów 16-bitowych (każdy kandydat na receptor wykrywał przynajmniej jeden fragment niezmodyfikowanego kodu monitorowanego), a wykrywalność dla $l = 32$, $m = 7$ wyniosła średnio 22% przy zajętości pamięci 36 bajtów.



Rys. 4. Wykres wykrywalności dla zmiennych parametrów l i m w funkcji rozmiaru anomalii z zakresu [1 B; 8 B]

Zwiększając próg aktywacji m do 8 bitów wygenerowano 38 receptorów 16-bitowych zajmujących łącznie 76 bajtów pamięci, czyli prawie dwukrotnie więcej niż dla $l = 32$ i $m = 7$, osiągając wykrywalność na poziomie 50,2%. W przypadku receptorów 32-bitowych i progu aktywacji $m = 8$, zajętość pamięci wzrasta do 412 B, a wykrywalność do 76,9%, osiągając dobry wynik. Generacja receptorów 32-bitowych dla progu $m = 8$ zajęła jednak znacznie więcej czasu niż dla receptorów 16-bitowych: dla $l = 16$ zajęła 58 ms, a dla $l = 32$ zajęła 676 ms. Można więc zaobserwować, że przy przyroście liczby m czas generacji receptorów 32-bitowych wzrasta znacznie bardziej niż czas generacji receptorów 16-bitowych.

Dla parametru $m = 9$ zajętość pamięci w przypadku $l = 16$ wyniosła 538 B, a w przypadku $l = 32$ wyniosła 1804 B. Receptory 16-bitowe osiągnęły wtedy dobrą wykrywalność na poziomie 83,8%, a receptory 32-bitowe osiągnęły po raz pierwszy bardzo dobrą średnią wykrywalność 93,9%. Dla $m = 9$ i wielkości anomalii ≥ 3 B receptory 32-bitowe osiągały bardzo dobrą wykrywalność (większą niż 90%), a dla wielkości anomalii 6 B, 7 B i 8 B osiągnęły one pełną, 100% wykrywalność. Dla $l = 16$, $m = 9$ bardzo dobra wykrywalność pojawiła się po raz pierwszy dopiero przy $a = 5$ B (wyniosła ona 94,5%), a średnia uzyskana wykrywalność osiągnęła wtedy dobry wynik 83,8%. Można zauważyć, że zajętość pamięci receptorów 32-bitowych jest znacznie większa niż receptorów 16-bitowych dla tej samej wartości m (przykładowo dla $m = 9$: 538 B dla $l = 16$, 1804 B dla $l = 32$). Czas generacji receptorów 32-bitowych dla $m = 9$ wyniósł aż 30,3 sekund, w porównaniu z czasem 3,2 sekundy dla $l = 16$. Ostatnim badanym parametrem m był 10-bitowy próg aktywacji. Średnia wykrywalność wzrosła do bardzo dobrego poziomu 91,2% w przypadku $l = 16$ i 97,4% dla $l = 32$. Bardzo dobre wykrywalności osiągnięte zostały kosztem znacznie dłuższego czasu generacji receptorów, który dla $l = 16$ wyniósł 72,9 sekund, a dla $l = 32$ aż 457 sekund.

5. Wnioski

Sztuczne systemy immunologiczne są nowatorskim rozwiązaniem umożliwiającym wykrywanie nieprawidłowości w programach komputerowych. Do ich budowy wykorzystuje się między innymi algorytm negatywnej selekcji. W artykule omówiono system wykrywania intruzów oparty o sztuczny system immunologiczny pracujący w systemie operacyjnym. W rozwiązaniu tym zaimplementowano algorytm generacji receptorów oparty o wzorce, który do działania wymaga podania parametrów takich, jak liczba bitów receptora i próg aktywacji. W artykule przedstawiono badania wykrywalności anomalii w zależności od rozmiaru anomalii, liczby bitów receptora i liczby bitów progu aktywacji.

Analizując wyniki badań łatwo zauważyć, że im większy próg aktywacji, tym lepsze osiągi systemu immunologicznego. Zwiększanie wartości progu aktywacji skutkuje jednak nie tylko lepszą wykrywalnością anomalii, lecz również większą

zajętością pamięci i znacznie dłuższym czasem generacji receptorów, szczególnie w przypadku receptorów 32-bitowych, gdzie czas generacji wyniósł dla zadanych parametrów 457 sekund. Dalsze prace nad badaniami mogą uwzględniać liczby bitów receptora niepodzielne przez 8 i anomalie mniejsze niż 1 pełny bajt.

Literatura

1. Somayaji A., Forrest S., Hofmeyr S., Longstaff T., *A sense of self for unix processes*, w: *IEEE Symposium on Security and Privacy*, 1996, s. 120-128.
2. Somayaji A., Hofmeyr S., Forrest S., *Principles of a computer immune system*, w: *New Security Workshop*, Langdale, Cumbria 1997, s. 75-82.
3. Forrest S., Perelson A.S., Allen L., Cherukuri R., *Self-nonsel self discrimination in a computer*, w: *IEEE Symposium on Security and Privacy*, IEEE Computer Society, 1994, No 202.
4. Kephart J., *A biologically inspired immune system for computers*, w: *Fourth International Workshop on Synthesis and Simulation of Living Systems, Artificial Life IV*, 1994, s. 130-139.
5. Dasgupta D., *Immunity-based intrusion detection systems: a general framework*, w: *22nd National Information Systems Security Conference (NISSC)*, 1999.
6. Andrews P.S., Timmis J., *Tunable detectors for artificial immune systems: from model to algorithm*, w: *Bioinformatics for Immunomics*, Springer, New York, NY, USA 2010, vol. 3, s. 103-127.
7. Sobh T.S., Mostafa W.M., *A cooperative immunological approach for detecting network anomaly*, w: *J Applied Soft Computing*, 2011, vol. 11(1), s. 1275-1283.
8. Wang D., Zhang F., Xi L., *Evolving boundary detector for anomaly detection*, w: *Expert Systems with Applications*, 2011, vol. 38(3), s. 2412-2420.
9. Powers S.T., He J., *A hybrid artificial immune system and self organizing map for network intrusion detection*, w: *Information Sciences*, 2008, vol. 78(15), s. 3024-3042.
10. Li G.Y., Guo T., *Receptor editing-inspired real negative selection algorithm*, w: *Computer Science*, 2012, vol. 39, s. 246-251.
11. Laurentys C.A., Ronacher G., Palhares R.M., Caminhas W.M., *Design of an artificial immune system for fault detection: a negative selection approach*, w: *Expert Systems with Applications*, 2010, vol. 37(7), s. 5507-5513.
12. Fanelli R., *A hybrid model for immune inspired network intrusion detection*, Springer-Verlag, Phuket, Thailand 2008.

13. Mostardinha P., Faria B.F., Zúquete A., Vistulo de Abreu F., *A negative selection approach to intrusion detection*, w: Coello, C.A.C., Greensmith, J., Krasnogor, N., Liò, P., Nicosia, G., Pavone, M., *Artificial Immune Systems, Lecture Notes in Computer Science*, 2012, vol. 7597, s. 178-190.
14. Wierzchoń S.T., *Generating optimal repertoire of antibody strings in an artificial immune system*, *Intelligent Information Systems*, 2000, s. 119-133.
15. Widuliński P., Wawryn K., *Detekcja anomalii w plikach za pomocą wybranych algorytmów inspirowanych mechanizmami immunologicznymi*, *Zeszyty Naukowe Wydziału Elektroniki i Informatyki Politechniki Koszalińskiej*, 2019.

Streszczenie

Zabezpieczenie systemu operacyjnego przed zagrożeniami jest od lat obszarem badań wielu naukowców i konstruktorów oprogramowania antywirusowego. Nowe, skomplikowane zagrożenia pojawiają się w szybkim tempie, co skłoniło badaczy do poszukiwania nowoczesnych metod ich wykrywania. W artykule zaprezentowano badania wykrywalności zagrożeń w systemie operacyjnym przy pomocy sztucznego systemu immunologicznego w zależności od podanych parametrów wejściowych algorytmów. System składa się z bloków: sterującego, generacji receptorów i wykrywania anomalii. Blok generacji receptorów konstruuje ciągi binarne zwane receptorami do wykrywania przy ich pomocy anomalii w programach. Blok wykrywania anomalii korzysta z wygenerowanego zestawu receptorów do detekcji zagrożeń. W pracy przedstawiono wyniki badań wykrywalności anomalii w zależności od podanej liczby bitów receptora i proggu aktywacji, a następnie przeanalizowano je i podsumowano.

Abstract

The protection of operating systems against malware has been a field of research for many scientists and antivirus software designers for years now. New, complicated dangerous software appears rapidly, which inspired the researchers to look for unconventional, novel solutions for malware detection. In the paper, an original research of malware detection rates achieved by an artificial immune system using specific input parameters is presented. The system consists of control, receptor generation and anomaly detection units. The receptor generation unit constructs binary strings called receptors used to recognize foreign program structures. The anomaly detection unit uses generated receptors to detect malware in the monitored program. In the work, presented are the results of research of malware detection rates with regard to receptor bit count and activation threshold. The results are analyzed and concluded.

Keywords: artificial immune system, receptor, anomaly, malware