RADA NAUKOWA DYSCYPLINY INFORMATYKA TECHNICZNA I TELEKOMUNIKACJA POLITECHNIKI WARSZAWSKIEJ

zaprasza na
PUBLICZNĄ OBRONĘ ROZPRAWY DOKTORSKIEJ

mgr. inż. Mariusza Gajewskiego

która odbędzie się w dniu 14 stycznia 2021 roku o godzinie $10^{\underline{00}}$ w trybie zdalnym na platformie MS Teams*.

Tytuł rozprawy doktorskiej:
„Systemy zarządzania infrastrukturą sieciową inteligentnych budynków”

Promotor – dr hab. inż. Jordi Mongay Batalla Wydział Elektroniki i Technik Informacyjnych Politechnika Warszawska

Recenzenci: dr hab. inż. Piotr Chołda, prof. uczelni Wydział Informatyki, Elektroniki i Telekomunikacji Akademia Górniczo-Hutnicza

dr hab. inż. Adrian Kliks, prof. uczelni Wydział Informatyki i Telekomunikacji Politechnika Poznańska

Na stronie internetowej wydziału **www.elka.pw.edu.pl/Wydzial/Rada-Wydzialu/Harmonogram-obron-doktorskich-streszczenia-i-recenzje** znajdują się streszczenie rozprawy oraz recenzje, jak również dostęp do tekstu rozprawy umieszczonej w Bazie Wiedzy Politechniki Warszawskiej.

Sposób uczestniczenia w publicznej obronie:

https://teams.microsoft.com/l/meetup-join/19%3ameeting_YjI5MjY5OTgtMmY5OC00MDZlLTgxYjctMGZkOGRjOTQ1OTU3%40thread.v2/0?context=%7b%22Tid%22%3a%223b50229c-cd78-4588-9bcf-97b7629e2f0f%22%2c%22Oid%22%3a%22393ff735-7b26-4f34-bdcd-f7ccab968b42%22%7d

W imieniu Rady Naukowej Dyscypliny

Informatyka Techniczna i Telekomunikacja

Politechniki Warszawskiej

Prof. dr hab. inż. Zbigniew Kotulski

Przewodniczący Komisji Doktorskiej

Rozprawa doktorska

mgr inż. Mariusz Gajewski

Systemy zarządzania infrastrukturą sieciową inteligentnych budynków

Streszczenie

W rozprawie przedstawiono nowe rozwiązania dla sieci teleinformatycznych stosowanych do zarządzania infrastrukturą inteligentnego domu (np. czujnikami, przełącznikami, zasilaniem itp.). Zaproponowano nowe mechanizmy do wykorzystania w zarządzaniu ww. sieciami. W szczególności zaproponowane mechanizmy znajdują zastosowanie w następujących obszarach:

- zastosowania nowego systemu adresacji opartej na identyfikatorach do określania nazwy i lokalizacji obiektów dołączonych do sieci inteligentnego budynku
- rozdzielenia wybranych funkcji zarządzania (w szczególności detekcja intruzów) obiektami dołączonymi do sieci inteligentnego budynku pomiędzy użytkownika oraz dostawcę usługi.

W pierwszym przypadku przedstawiono stan sztuki dotyczący stosowanych schematów nazewnictwa do adresowania urządzeń oraz usług Internetu Rzeczy. Na podstawie wniosków z przeglądu rozwiązań sformułowano przesłanki do opracowania schematu adresacji opartego na identyfikatorach (*ID Layer*), obejmującego zarówno adresowanie urządzeń jak i realizowanych z ich pomocą usług. Zaproponowany schemat jest ściśle powiązany z topologią połączeń sieciowych, co czyni go naturalnym kandydatem do wykorzystania w sieciach przewodowych do zarządzania infrastrukturą inteligentnych budynków czy w sieciach przemysłowych. W oparciu o założenia zaproponowano specyfikację protokołu komunikacyjnego obejmującego warstwy L3-L7 modelu OSI. To rozwiązanie może mieć

praktyczne zastosowane tam, gdy lokalizacja węzłów jest ściśle powiązana ze strukturą otoczenia, np. jak ma to miejsce w inteligentnych domach lub budynkach.

Przedstawiono również implementację protokołu oraz mechanizmów sieciowych a następnie wykonano testy porównujące funkcjonalność uzyskanego rozwiązania *proof of concept* z funkcjonalnością modułu przekazywania pakietów (*forwardingu*), zaimplementowaną w systemie Linux. Wykonano również testy mające na celu porównanie czasów przekazywania pakietów *od końca do końca* przez węzły oparte na systemie operacyjnym Linux. Wykonano również testy porównujące opóźnienie wprowadzane przez mechanizm buforowania pakietów zaproponowanego rozwiązania z opóźnieniami obserwowanymi przy dostępie do danych zapisanych w bazie danych.

W drugiej części rozprawy omówiono problem wykrywania anomalii w sieciach Internetu Rzeczy ze szczególnym uwzględnieniem lokalnych sieci teleinformatycznych służących do komunikacji w ramach inteligentnego domu/budynku. Zaproponowano i opisano rozwiązanie łączące w sobie wykrywanie anomalii lokalne oraz centralne. Warunkiem zastosowania takiego podejścia jest jednak to, że urządzenie odpowiedzialne za wykrywanie anomalii – brama domowa (*ang.* Home Gateway) – jest zarządzane przez dostawcę usługi, który m.in. gromadzi i przetwarza informację o anomaliach wykrytych lokalnie. Anomalie wykrywane są lokalnie dla każdej podsieci działającej w ramach sieci domowej i w tym celu wykorzystywane jest uczenie maszynowe. Z kolei w procesie centralnego korelowania anomalii zaproponowano klasteryzację.

W rozprawie zamieszczono opis analizy matematycznej zagadnienia oraz przeprowadzono symulacje dla tak zaproponowanego procesu wykrywania anomalii. Opisano skuteczność takiego podejścia, zastawiając wyniki z przykładami literaturowymi. W pracy opisano szczegółowo warunki przeprowadzenia symulacji oraz przesłanki dla wyboru wektora cech wykorzystanych w procesie uczenia maszynowego.

W podsumowaniu zamieszczono najważniejsze argumenty wspierające obie przedstawione tezy. Przywołano również ograniczenia zaproponowanych rozwiązań.

Słowa kluczowe: Internet Rzeczy, sieć domowa, sieć sensorowa, model zachowania, wykrywanie anomalii, rezerwacja zasobów, informacja o bezpieczeństwie i zarządzanie zdarzeniami , brama domowa.

Abstract

This dissertation describes methods for the improvement of local networks for Smart Home applications (i.e., sensors, switches, controllers, etc.). It includes new mechanisms directed to the management of Smart building networks; concretely, the novel mechanisms are:

- a new ID based addressing system for packet networks that makes use of an unique identifier and location pointer of intelligent nodes connected to Home Area Network,
- a framework for splitting selected management functions between the user and the service provider.

In the first part of the Thesis, a discussion about modern naming schemes and protocols supporting message forwarding is presented. Based on that, it is proposed a new forwarding framework at the ID layer and the corresponding modules have been implemented. The ID layer node prototype joints addresses of objects and services in an easy-to-manage and flexible way. Moreover, a specification of the communication protocol was proposed covering the L3-L7 layers of the OSI model. This solution is appropriate when the location of the nodes is closely related to the structure of the environment as it occurs in intelligent buildings/enterprises for the wired infrastructure.

Implementation of the protocol and network mechanisms was also presented. Then, tests were performed comparing the functionality of the proof of concept solution with the functionality of the packet forwarding module implemented in the Linux system. Tests were performed to compare end-to-end packet forward delays on Linux-based nodes. Tests were also carried out to check the delay introduced by the packet buffering mechanism of the proposed solution. The obtained test results were also compared it with the delays observed when accessing data saved in the databases.

In the second part, the dissertation demonstrates that functions responsible for Smart Home anomalous traffic detection may be split between client and service provider. It encompasses a novel strategy for anomaly detection that consists of shared responsibilities between user and network provider. The proposed two-tier Intrusion Detection System uses a machine learning method for classifying the monitoring records and searching suspicious anomalies across the network at the service provider's data center. Result show that local anomaly detection combined with anomaly correlation at the service providers level can provide reliable information on the most frequent IoT devices misbehavior which may be caused by infection.

The dissertation describes also the mathematical analysis of the proposed solution as well as the results of simulations performed for the proposed anomaly detection process. This includes tables of measurement results, configuration details of experimental simulation environment and the premises for choosing the features used in the machine learning process. The effectiveness of my approach was compared with other solutions proposed in the literature.

The summary contains the most important arguments supporting both presented theses. Limitations of the proposed solutions are also highlighted.

Keywords: Internet of Things, Home Area Network, Sensor Network, Behavior model, Anomaly detection, provisioning, Security Information and Event Management (SIEM), Home Gateway.

AKADEMIA GÓRNICZO-HUTNICZA
IM. STANISŁAWA STASZICA W KRAKOWIE

Wydział Informatyki, Elektroniki i Telekomunikacji

KATEDRA TELEKOMUNIKACJI

Dr hab. inż. Piotr CHOŁDA, prof. AGH

Kraków, dn. 18 września 2020 r.

RECENZJA ROZPRAWY DOKTORSKIEJ

Tytuł rozprawy: Systemy zarządzania infrastrukturą sieciową inteligentnych budynków
Autor rozprawy: mgr inż. Mariusz GAJEWSKI

1. WSTĘP

Niniejsza recenzja została przygotowana na potrzeby postępowania ws. nadania stopnia doktora nauk inżynieryjno-technicznych. Postępowanie prowadzi Politechnika Warszawska (PW). Pełną dokumentację dot. rozprawy doktorskiej, wraz z informacją o powołaniu przez Radę Dyscypliny Informatyka Techniczna i Telekomunikacja PW na recenzenta rozprawy, otrzymałem od p. dr. hab. inż. Jarosława ARABASA, prof. PW, w dn. 21 lipca 2020 r.

Recenzowana rozprawa została złożona przez p. mgr. inż. Mariusza GAJEWSKIEGO. Nosi tytuł „Systemy zarządzania infrastrukturą sieciową inteligentnych budynków" i została napisana w całości po polsku (poza anglojęzycznym streszczeniem, *Abstract*, zamieszczonym na str. 7). Promotorem doktoratu jest p. dr hab. inż. Jordi MONGAY BATALLA.

Dostarczona mi rozprawa doktorska liczy 145 stron. Składa się z pięciu numerowanych rozdziałów: 1. Wstęp i tezy pracy (str. 11-13), 2. Wprowadzenie (str. 15-29), 3. Adresowanie na podstawie identyfikatorów (str. 31-77), 4. Współdzielenie funkcji wykrywania anomalii w sieci domowej (str. 79-133), 5. Podsumowanie (str. 135-136). Oryginalne wyniki zasadniczo ujęto w rozdziałach 3 i 4. Uzupełnieniem zawartości pracy są: zamieszczone na początku streszczenia (w językach polskim i angielskim); spis treści; jak również zawarte na końcu: ułożona w kolejności cytowania bibliografia licząca 102 pozycje oraz wykaz anglojęzycznych skrótów. Praca zawiera szereg kolorowych ilustracji, które ułatwiają odbiór tekstu. W pracy zamieszczono również pewną liczbę czytelnych tabel, które służą podsumowaniu prezentowanej wiedzy zastanej oraz wizualizacji parametrów prowadzonych doświadczeń.

Poniżej odnoszę się do poszczególnych punktów składowych, których skomentowania oczekuje się ode mnie według informacji dostarczonej w zleceniu przysłanym przez Radę Dyscypliny.

Akademia Górniczo-Hutnicza | Wydział Informatyki, Elektroniki i Telekomunikacji
Katedra Telekomunikacji
al. A. Mickiewicza 30, 30-059 Kraków,
tel. +48 12 617 48 08 fax +48 12 634 23 72
e-mail: piotr.cholda@agh.edu.pl

2. CEL BADAŃ (W ODNIESIENIU DO TEZY ROZPRAWY). Jakie zagadnienie naukowe jest rozpatrzone w pracy (teza rozprawy) i czy zostało ono dostatecznie jasno sformułowane przez Autora?

Tematyka rozprawy obejmuje zagadnienia teleinformatyki. Doktorant skupia się na obszarze tworzenia sieci Internetu Rzeczy (*Internet of Things*, IoT) oraz zarządzania nimi. Interesuje go specyfika zastosowania tych sieci w tzw. inteligentnych budynkach. Wpływa to na pewne założenia techniczne związane z samym celem badań oraz prowadzeniem prac. Kwestie szczegółowe dotyczą projektowania pewnych elementów protokołów komunikacyjnych, implementacji programistycznej systemów, bezpieczeństwa działania sieci (a właściwie wykrywania nietypowych zachowań, które mogą być skutkami ataków) oraz zastosowań uczenia maszynowego w telekomunikacji. W ogólności cała problematyka pracy jest aktualna i bardzo atrakcyjna od strony teoretycznej i praktycznej.

W rozprawie wyróżniono dwie tezy dotyczące w zasadzie odmiennych aspektów:

- Teza 1 brzmi: „Jest możliwe zastosowanie nowego systemu adresacji opartej na identyfikatorach do określania nazwy i lokalizacji obiektów dołączonych do sieci inteligentnego budynku." Użyte w tezie pojęcia są generalnie jasne, zresztą w tekście rozprawy Doktorant doprecyzowuje, co przez nie rozumie. Ponadto Autor przy wprowadzaniu tezy na str. 12 mówi, że zamierza sprawdzać efektywność przyjętych rozwiązań. Być może elementy związane z owym badaniem efektywności warto byłoby wprost zamieścić w tezie.
- Teza 2 została sformułowana następująco: „Jest możliwe rozdzielenie wykrywania anomalii w sieci teleinformatycznej inteligentnego domu pomiędzy użytkownika oraz dostawcę usługi." W tym przypadku elementy tezy również są dosyć jasne, chociaż można byłoby od razu bliżej podać, jakiego typu anomalie są interesujące dla Doktoranta (oraz w ogóle doprecyzować warunki adekwatne dla prowadzonych badań, np. typ stosowanych mechanizmów sterowania). Ponadto — biorąc pod uwagę w jaki sposób teza jest dowodzona (o czym niżej) — byłoby zasadne uwzględnienie w niej, jakie są kryteria stwierdzania, że przedmiotowe rozdzielenie faktycznie jest możliwe. W tym przypadku Doktorant pokazał, że zyski wynikające ze specyficznych metod detekcji anomalii opartych na klasyfikacji zero-jedynkowej oraz analizy skupień z użyciem metod uczenia maszynowego mogą być wykonywane w innych obszarach sieci i na podstawie odmiennych danych. Taka klaryfikacja na pewno wzmocniłaby brzmienie tezy.

Odrębną kwestią jest użycie w przypadku obu tez modalności związanej z potencjałem uzyskania założonych celów. Jest to wprawdzie popularne podejście, ale wprowadza jednak pewną niejasność dotyczącą oczekiwań. Jak wynika co najmniej z podsumowania, w przypadku tezy 2 trafne byłoby mówienie nie tyle o możliwości, co raczej o zasadności wprowadzenia wskazanego rozdzielenia funkcji detekcji w celu uzyskania wyższej jakości działania systemu.

Niemniej jednak, obie tezy są dosyć **przejrzyste, wartościowe technicznie, dobrze zrozumiałe i stanowią cel istotnych badań**. Doktorant dowodzi ich, z jednej strony konstruując po prostu odpowiedni system adresacji poszerzony też o inne elementy o charakterze protokołowym (to w odniesieniu do tezy 1), zaś z drugiej strony proponując mechanizmy wykrywania anomalii na obu poziomach:

użytkownika (w oparciu o klasyfikator binarny) oraz dostawcy usługi (który może zaaplikować analizę skupień ze względu na potencjalnie dużą liczbę danych rozpoznanych przez różnych użytkowników). Trzeba tutaj zwrócić uwagę, że obie tezy dotyczą jednak bardzo zróżnicowanych obszarów zarządzania tytułowymi sieciami, są dowodzone z użyciem odmiennych metod i w zasadzie nie mają na siebie większego wzajemnego wpływu. Można na to patrzeć z punktu widzenia pozytywnego (szerokość zainteresowań i umiejętności Doktoranta), ale też bardziej krytycznego (pewna niespójność koncepcyjna rozprawy).

3. **CHARAKTER ROZPRAWY. Jaki charakter ma rozprawa (teoretyczny, doświadczalny, inny)?**

Praca ma **charakter konstrukcyjno-doświadczalny**, ponieważ Doktorant dowodzi obu tez, proponując pewne rozwiązania (element konstrukcyjny), które następnie testuje z użyciem eksperymentu (w przypadku tezy 1, tj. w odniesieniu do zaproponowanego protokołu) oraz symulacji (w przypadku tezy 2, tj. zdefiniowania systemu wykrywania anomalii). System adresacji i elementy protokołu mają charakter autorski, natomiast w przypadku opracowania systemu wykrywania anomalii Dyplomant zaproponował w oryginalny sposób rozdzielenie go na dwa podsystemy (tj. oparty na obszarach funkcjonalnych użytkownika oraz operatora), w ramach których zastosował (ale z niewielkimi elementami wkładu własnego) znane metody uczenia maszynowego.

Podejście zaproponowane przez Doktoranta, tj. polegające na sformułowaniu tezy dotyczącej możliwości wykonania czegoś, a następnie udowodnieniu jej w oparciu o konstrukcję odpowiedniej metody, która jest potem zilustrowana (ewentualnie przetestowana), jest zasadniczo typowe i powszechnie akceptowane. Element doświadczalny ma na celu wzmocnienie koncepcji konstrukcyjnej, co wpisuje się w standardowe postępowanie w informatyce technicznej i telekomunikacji. Biorąc to wszystko pod uwagę stwierdzam, że praca ma przede wszystkim **walor empiryczny**.

Aspekty skupione na modelowaniu teoretycznym w pracy praktycznie nie występują (poza niewielkim wtrąceniem sformalizowanego opisu używanych metod uczenia maszynowego). W świetle przyjętej koncepcji pracy modelowanie teoretyczne nie było jednak konieczne.

4. **SPOSÓB PRZEPROWADZENIA ANALIZY ŹRÓDEŁ. SPOSÓB SFORMUŁOWANIA WNIOSKÓW WYNIKAJĄCYCH Z ANALIZY ŹRÓDEŁ. Czy w rozprawie przeprowadzono w sposób właściwy analizę źródeł (w tym literatury światowej, stanu wiedzy i zastosowań w przemyśle) świadczący o dostatecznej wiedzy Autora. Czy wnioski z przeglądu źródeł sformułowano w sposób jasny i przekonywujący?**

Tekst rozprawy jest napisany w sposób wskazujący na dobrą znajomość tematyki. Lista źródeł literaturowych obejmuje 102 pozycje, chociaż warto zwrócić

uwagę, że Doktorant zdecydował się nie włączać do niej bardzo licznych przypisów-odnośników do stron internetowych, które w sumie też tworzą podstawy wiedzy dla obszaru stanowiącego przedmiot badań. Przeważająca większość źródeł to pozycje anglojęzyczne (właściwie tylko z wyjątkiem prac, w których zawarto niektóre aspekty koncepcyjne prezentowane w rozprawie doktorskiej, a które powstały w zespole, którego Doktorant jest częścią). Wykorzystana bibliografia składa się z tekstów funkcjonujących w obiegu międzynarodowym (w większości przypadków były publikowane w uznanych periodykach branżowych lub na cieszących się renomą konferencjach).

Analiza literaturowa jest rozproszona i Doktorant w zasadzie nie poświęca jej odrębnego rozdziału, co częstokroć bywa praktykowane w rozprawach doktorskich oraz artykułach naukowych (i jest wygodne dla czytelnika, dla którego jest wtedy oczywiste, które koncepcje uważa się za istotne z punktu widzenia inspiracji do przedstawionych przez autora badań, ewentualnie co stanowi element polemiczny). W przypadku recenzowanej rozprawy analiza literatury jest zamieszczona przede wszystkim w rozdziale 2 (wprowadzenie do technik sieciowych wspólnych dla obszarów związanych z obiema tezami), w podrozdziale 3.1 (zagadnienia adresacji odnoszące się do tezy 1) i w podrozdziale 4.2, a także niektórych sekcjach podrozdziału 4.3 (problematyka monitorowania systemów z punktu widzenia wykrywania ataków z niewielkimi i raczej niekonsekwentnie opisanymi elementami wprowadzenia do mechanizmów uczenia maszynowego).

Dobór literatury na pewno wskazuje **na dobre rozeznanie** Doktoranta w zakresie dziedziny objętej wnioskiem, przede wszystkim systemów IoT konstruowanych na potrzeby infrastruktur sieciowych inteligentnych budynków. Bardzo dobrze reprezentowane są elementy związane z właściwościami działania, architekturami systemów i protokołami. W zakresie związanym z konstrukcją systemu adresacji Doktorant przedstawił zastaną wiedzę, ale bardziej skupiał się na opisie rozwiązań blisko powiązanych ze standardami i sposobem działania sieci, a mniej na pracach typowo badawczych (najlepiej reprezentowane są wyniki wybranych projektów). Doktorant przedstawia wprawdzie wyniki uzyskane przez innych specjalistów, ale w stosunkowo niewielkim stopniu podkreśla aspekty inspirujące dla siebie i nie do końca jasno odróżnia swój oryginalny wkład od wcześniejszych osiągnieć w branży. Ponadto trzeba zauważyć, że w odniesieniu do wspomnianego obszaru badawczego przegląd literatury nie poświęca zbyt wiele miejsca pracom nowszym, nacisk kładąc raczej na przegląd prac pochodzących głównie sprzed 2016 roku. W odniesieniu do zagadnień związanych z wykrywaniem anomalii Doktorant analizuje także prace nowsze i trafnie przedstawia wiedzę zastaną (oraz bardziej bogato odwołuje się do propozycji zawartych w oryginalnych pracach badawczych). Zwraca również uwagę na pewne inspirujące go elementy, ale nie pokazuje zbyt mocno, na ile źródła dostarczają np. oczekiwanych poziomów wykorzystywanych wskaźników, które mogłyby świadczyć o jakości wyników raportowanych w części doświadczalnej.

5. ROZWIĄZANIE PRZEDSTAWIONEGO ZADANIA, WŁAŚCIWOŚCI PRZYJĘTYCH METOD I ZAŁOŻEŃ. Czy Autor rozwiązał postawione zagadnienia, czy użył właściwej do tego metody i czy przyjęte założenia są uzasadnione?

Oba aspekty wyróżnione w tezach rozprawy dotyczą zarządzania infrastrukturami tzw. inteligentnego domu, przy czym inteligencja jest tutaj związana np. z obsługą przez różnego rodzaju czujniki, które są połączone w sieć. W ten sposób tworzą elementy Internetu Rzeczy, tj. w większości są to urządzenia o ograniczonych możliwościach przetwarzania danych (co niesie konsekwencje np. dotyczące potencjału do wykrywania anomalii itp.). W odniesieniu do tezy 1 Doktorant z jednej strony wprowadza autorski system adresacji w oparciu o zaproponowane identyfikatory definiujące nazwę i lokalizację połączonych w sieć elementów. W odniesieniu do tezy 2 Doktorant proponuje w dziedzinie detekcji anomalii (tu: ataków) oparty na uczeniu maszynowym system rozdzielający wybrane funkcje zarządzania.

Patrząc dokładniej, można powiedzieć że w celu udowodnienia obu założonych tez Doktorant przeprowadził następujące prace, przedstawione w rozprawie:

1. Opracował schemat adresacji urządzeń i realizowanych z ich pomocą usług, przy czym w odniesieniu do tego schematu zapronował protokół komunikacyjny obejmujący warstwy od sieciowej w górę stosu (aspekt konstrukcyjny dot. tezy 1).
2. Zaimplementował schemat adresacji z protokołem i przeprowadził eksperymenty porównujące ich wydajność w stosunku do mechanizmów przekazywania pakietów traktowanych jako standardowa funkcję systemu operacyjnego Linux; zaobserwował również inne właściwości działania sieci w oparciu o zaproponowane rozwiązania protokołowe, w tym wartości opóźnień wprowadzanych przez mechanizm buforowania pakietów (aspekt doświadczalny dot. tezy 1).
3. Zaproponował system wykrywania anomalii w lokalnych sieciach teleinformatycznych służących do obsługi inteligentnego domu, przy czym system rozdziela funkcje detekcyjne dokonywane lokalnie (obszar użytkownika) i centralnie (obszar dostawcy); system korzysta z mechanizmów uczenia maszynowego (naiwny klasyfikator Bayesa i analiza skupień) w celu realizacji procesu przetwarzania danych na potrzeby uzyskania informacji o wystąpieniu anomalii oraz jej typie (aspekt konstrukcyjny dot. tezy 2).
4. Zaimplementował opisany wyżej system i przetestował go symulacyjnie, ilustrując jego działanie (aspekt doświadczalny dot. tezy 2).

Poniżej podaję nieco więcej informacji oraz komentarzy na temat realizacji poszczególnych punktów:

Ad. 1. Obie tezy są przedstawione w jednym kontekście, tj. zarządzania inteligentnymi sieciami domowymi, przy czym podstawy techniczne (ogólne koncepcje architektury, protokoły itp.) Doktorant wprowadza wspólnie w rozdz. 2, który zawiera w sobie pewne elementy analizy literaturowej. Bliższe dane bibliograficzne dotyczące zagadnień identyfikacji i adresowania w kontekście sieci IoT oraz infrastruktur sieciowych budowanych na potrzeby inteligentnych budynków Doktorant zamieścił w podrozdziale 3.1. (faktem jednak jest, że niektóre z wprowadzonych koncepcji są również przydatne przy zrozumieniu treści rozdz. 4 odnoszącego się głównie do tezy 2).

Zaproponowany przez Doktoranta system adresacji opiera się na zastosowaniu jednolitej warstwy identyfikatorów, które jednak są bardzo bogate semantycznie i mogą zawierać różnorodne informacje (np. lokalizacja, obsługa konkretnych serwisów itp.). Zdefiniowane identyfikatory (ciągi symboli ASCII o strukturze hierarchicznej) są oczywiście używane do trasowania danych (pakietów), ale ich nowość polega na tym, że można istotnie używać ich do wykrywania węzłów, jak również stosować w transmisji jeden-do-wielu (w formie *multicast* lub *anycast*). Z punktu widzenia konstrukcji protokołu ważne jest także założenie nt. możliwości buforowania danych w poszczególnych węzłach sieci. Dzięki temu da się oszczędzać zasoby energetyczne, bo węzły mogą być usypiane. Całość protokołu korzysta z techniki Ethernet w warstwie łącza danych.

Ad. 2. Implementacja środowiska eksperymentalnego została dokonana w ramach projektu IDSECOM, w którym uczestniczył Doktorant. Zaproponowane rozwiązanie o charakterze protokołowym zostało przetestowane głównie w opozycji do sposobu przekazywania pakietów opartego na implementacji w systemie operacyjnym Linux (m.in. uzyskano mniejsze wartości strat pakietów). Nie neguję, że jest to pouczające porównanie, które zresztą wypada na korzyść propozycji opracowanej przez Doktoranta, trzeba jednak powiedzieć że można byłoby oczekiwać, iż część doświadczalna posłuży do pokazania w sposób ilościowy zdefiniowanych na str. 59 cech o charakterze jakościowym. Są one faktycznie istotne i na pewno część z nich mogła zostać również przełożona na oczekiwane wartości parametrów mierzalnych (np. oszczędność energii, elastyczność transferu plików itp.). W ogóle byłoby zasadne, gdyby Doktorant najpierw zdefiniował satysfakcjonujące poziomy wartości wskaźników wydajności, gdyż — przy rezygnacji z porównania ilościowego z innymi rozwiązaniami proponowanymi w literaturze przedmiotu — wyniki wydajnościowe mają charakter raczej opisu fenomenów niż zobiektywizowanej wiedzy. Skądinąd Doktorant ma jakąś wizję, jakie wartości byłyby pożądane, poniważ mówi nawet o „wystarczająco wydajnej" implementacji — niemniej jednak nie podaje uprzednio żadnych oczekiwań, które mogłyby zostać zweryfikowane lub sfalsyfikowane. Zaproponowane testy walidacyjne nie zawsze jasno odnoszą się do wskazanych cech, ale na pewno mają walor ilustracyjny. W przypadku propozycji związanych z nowymi mechanizmami o charakterze protokołowym w gruncie rzeczy jest to także podejście atrakcyjne, tym bardziej że sformułowane scenariusze badawcze istotnie odpowiadają warunkom spotykanym w inteligentnych budynkach (np. topologia sieci testowej czy struktura interakcji z bazą danych), a walidacja odbywa się z użyciem eksperymentu w warunkach rzeczywistych. Ta część zatem dobrze służy jako *proof-of-concept* pomysłu zaproponowanego przez Doktoranta.

Ad. 3. Pomysł związany z tezą 2 obejmuje co najmniej dwie koncepcje. Pierwsza z nich wprost dotyczy zawartych w tezie kwestii rozdzielenia specyficznych funkcji zarządzania związanych z bezpieczeństwem (wykrywanie anomalii) na dwa obszary odpowiedzialności: jeden jest związany z samym użytkownikiem urządzenia (urządzeń), zaś drugi z dostawcą usług inteligentnego budynku. Druga koncepcja dotyczy sposobu realizacji funkcji we wskazanych obszarach (chociaż tutaj Doktorant raczej ilustruje koncepcję i korzysta z osiągnięć innych badaczy) — w przypadku obszaru użytkownika Autor proponuje zastosowanie klasyfikatora zero-jedynkowego (opartego na jednym z naiwnych klasyfikatorów Bayesa), a w przypadku obszaru dostawcy proponuje zastosowanie analizy skupień z użyciem cieszącego się dobrą opinią algorytmu nieparametrycznego DBSCAN, który faktycznie musi korzystać z bogatszych danych niż dostępne w obszarze

użytkownika. Samo zastosowanie wspomnianych mechanizmów uczenia maszynowego nie jest jednak sproblematyzowane przez Doktoranta i są one aplikowane czysto narzędziowo, nie stanowiąc przedmiotu badań jako takich.

Koncepcja wykrywania anomalii w obszarze użytkownika opiera się na obserwacji tzw. normalnego (typowego) profilu aktywności ruchowej. Jest to jedno z dopuszczalnych i akceptowalnych podejść, chociaż oczywiście możliwe byłoby również zastosowanie innych (np. zawartość przesyłanych komunikatów itp.). Ten problem też nie jest sproblematyzowany przez Doktoranta, chociaż wybór podejścia jest przez niego umotywowany przez analizę literaturową i jasno wyrażony (co widać np. w pustych polach tab. 4.5). Z kolei aspekt korelacyjny, za który odpowiada obszar dostawcy, jest oparty na bogatszym zestawie danych pobieranych z obszaru użytkownika i w tym przypadku Doktorant wyraźnie wprowadza koncepcję korzystania ze wskazanego w tezie 2 rozdzielenia funkcji wykrywania anomalii. W związku z tym, że urządzenia obecne w obszarze dostawcy mają istotnie większy obszar obliczeniowy jest to rozwiązanie realistycznie zakładające specyfikę tytułowych sieci.

Z punktu widzenia problematyki obliczeniowej, tj. uczenia maszynowego, warto również podkreślić, że Doktorant wykorzystał metodę analizy głównych składowych PCA do wyboru odpowiednich cech służących jako dane wejściowe do narzędzia umożliwiającego wykrywanie anomalii. Faktem jednak jest, że jak na użycie metod uczenia maszynowego stosunkowo mało miejsca zostało poświęcone świadomej obróbce danych, która nie jest zbyt jasno opisana (np. nie dostrzegłem zwrócenia uwagi na problem normalizacji danych, na które prawie wszystkie metody uczenia maszynowego — może poza niestosowanymi przez Doktoranta drzewami klasyfikacyjnymi — są bardzo czułe). Nie uważam ponadto, żeby zastosowanie konkretnego klasyfikatora było w pełni uzasadnione. Nie neguję wartości zastosowanego klasyfikatora Bayesa, ale mógłby on być zastąpiony także innym (np. sztuczną siecią neuronową). Wybór dokonany przez Doktoranta nie jest wystarczająco umotywowany, nawet z punktu widzenia jego własnej analizy literaturowej. W każdym razie decyzja nt. zastosowania tylko jednego klasyfikatora jest dosyć radykalna. Doktorant wprawdzie uzasadnia swój wybór analizą złożoności obliczeniowej, która być może wskazuje na przewagę klasyfikatora Bayesa, ale trzeba powiedzieć, że jest to raczej podejście nietypowe w środowisku badaczy zajmujących się uczeniem maszynowym. Kładą oni większy nacisk w zakresie wyboru metody kładzie się na wskaźniki związane z samymi wynikami klasyfikacji, przynajmniej na razie godząc się na to, że metody uczenia maszynowego działają niezbyt efektywnie obliczeniowo. Na usprawiedliwienie wyboru dokonanego przez Doktoranta może powiedzieć, że urządzenia IoT obecne w obszarze użytkownika mają istotnie bardzo ograniczone możliwości obliczeniowe i w ich przypadku mniejsza złożoność faktycznie jest dużą zaletą. Poza wskazanymi dyskusyjnymi aspektami założenia dotyczące samej propozycji są przekonująco uzasadnione i dobrze oddają warunki praktyczne: obserwacja aktywności na interfejsach, skupienie na specjalizowanych urządzeniach Internetu Rzeczy, przyjęcie że w normalnej sytuacji charakterystyka opisująca zachowanie jest stabilna itd.

Ad. 4. Implementacja środowiska doświadczalnego została dokonana w ramach projektu FUSE, w którym uczestniczył Doktorant. Zasadność zaproponowanego podejścia związanego z wykrywaniem anomalii jest przez Doktoranta dowodzona z użyciem symulacji. W tym przypadku brak eksperymentów nie dziwi i jest często praktykowany, gdyż konstrukcja środowiska, w którym mogłyby istotnie występować ataki, jest niezwykle trudna i łączyłaby się zapewne z bardzo długimi

badaniami, dużymi kosztami itp. Oprogramowanie użyte do symulacji to powszechnie akceptowany w badaniach sieciowych symulator Omnet++. Jest to otwarte oprogramowanie, co podnosi jego wiarygodność. Gdyby Doktorant konstruował samodzielnie środowisko symulacyjne, wymagałoby to od niego także dużo większego trudu, przy niepewnej wiarygodność wyników. Z tego powodu wybór Autora rozprawy jest dobry. Symulacje przeprowadzono w odniesieniu do dwóch różnych sposobów komunikacji (korzystającego ze stosu IP oraz bez niego), którym odpowiadają różne topologie. Doktorant przekonująco uzasadnia wybór cech, na których opiera się rozpoznawanie anomalii w odniesieniu do profilu ruchowego. Doktorant samodzielnie symuluje zachowania anomalne, np. zwiększając intensywność generowania pakietów przez wybrane urządzenie. Takie podejście jest dobrze umotywowane (np. z pomocą tab. 4.2) i trafnie oddaje fenomeny związane z szeroką klasą ataków. W ogólności przyjęta metodologia doświadczeń w odniesieniu do tezy 2 nie budzi zastrzeżeń.

Jakość klasyfikacji jest oceniana z użyciem krzywych ROC, opisujących zachowanie klasyfikatorów rozróżnianych z użyciem różnych wartości punktów odcięcia służących do odróżnienia zachowania typowego od anormalnego. Jest to prawidłowe podejście. Zestawione wyniki pokazują, że proponowany klasyfikator Bayesa dobrze się sprawdza. Jest to wprawdzie ocena „na oko", gdyż Doktorant nie podaje wartości odpowiednich wskaźników AUC, mimo że o nich wspomina. Bez wątpienia są one jednak znacznie większe niż 0,5. Aspekt dotyczący drugiego obszaru detekcji, tj. opartego na analizie skupień został głównie zilustrowany w oparciu o wybrane przykłady działania.

W podsumowaniu przedstawionego podejścia do problemów rozwiązywanych w pracy stwierdzam, że **Doktorant ze znawstwem rozwiązuje postawione problemy (które są istotne), wpisując się w najnowsze trendy techniczne oraz istniejące standardy**. W ogólności Doktorantowi udało się dowieść obu zdefiniowanych przez niego tez. Od strony metodologicznej sposób prowadzenia przez niego prac doświadczalnych (eksperymentów i symulacji) jest trafny. Największy zarzut jaki można postawić od strony metodologii badawczej dotyczy tego, że nie zawsze jasno uzasadniono podejmowane decyzje (chociaż nie można im zarzucić nietrafności). Niemniej jednak przyjęte założenia są przeważnie klarownie przedstawione, co stanowi rzetelne podejście, gdyż jest jasne, co Doktorant robi. Krytycznie oceniam jednak też fakt, że Doktorant w zasadzie rezygnuje z porównania proponowanych rozwiązań z innymi propozycjami obecnymi w literaturze przedmiotu.

W tej części mojej recenzji przy okazji chcę zwrócić uwagę na kilka mniejszych nieścisłości merytorycznych, które znalazłem w rozprawie:

- Str. 26: OFDM to raczej technika zwielokrotnienia niż modulacji, dlatego pełne rozwinięcie nazwy to „Orthogonal Frequency Division Multiplexing", a nie „Modulation".
- Str. 36: Doktorant pisze „Adres MAC ma postać 64-bitowego ciągu, podobnie jak adresy MAC w Etherniecie", ale te ostatnie korzystają z sześciu bajtów, a więc 48 bitów.
- Str. 48 i 51: Doktorant doprecyzowuje pojęcie drzewa utożsamiając je z grafem acyklicznym. Trzeba pamiętać, że graf acykliczne to tzw. las (który można traktować jako uogólnienie drzewa), natomiast istotna jest (chyba nie tylko z punktu widzenia teorii grafów, ale także właściwości topologicznych) także cecha spójności.

- W zakresie opisu szczegółów eksperymentalnych Doktorant odsyła czytelnika do charakterystyki testów podanej w publikacji [52] (str. 68). W zasadzie rozprawa doktorska powinna funkcjonować jako odrębna całość i o ile możliwe jest odesłanie do pracy innych autorów, to akurat w przypadku wyników własnych Doktoranta zasadne byłoby oczekiwanie, że wszystkie niezbędne aspekty zostaną wyjaśnione.
- Wyniki zamieszczone w sekcji 3.4.2 wiążą się z rozwiązywaniem nazw z pomocą zapytań do bazy danych. W tym przypadku Doktorant wykonał pięć serii testów, które — jak domniemywam na podstawie rys. 3.16 — są rozmieszczone niezbyt gęsto (czy nie za rzadko?). Ponadto podane wartości wykazują bardzo duży rozrzut wokół średnich, które dodatkowo mają charakter niemonotoniczny. Brakuje mi interpretacji uzyskanego wyniku, który robi wrażenie bardzo losowego i nie wiadomo nawet, czy jest poprawny.
- Rozważania związane z anomaliami Doktorant ogranicza do problematyki ataków (wspomina jeszcze o niezdatnościach wynikających z niepoprawnego zaprogramowania urządzeń). Jest to jasno wskazane i takie zawężenie nie stanowi w zasadzie problemu, tym bardziej że w odniesieniu do rozważanych typów ataków Doktorant odnosi się do powszechnie poważanych źródeł (raporty OWASP). Warto jednak zwrócić uwagę, że anomalie w zachowaniu mogą też wynikać z uszkodzeń, na które sieci Internetu Rzeczy są przecież narażone (niekiedy ze względu na ograniczoną żywotność i niewielkie koszty urządzeń nawet bardziej niż klasyczny sprzęt spotykany w sieciach lokalnych). W tym przypadku zachowanie ruchowe może być zupełnie odmienne. Nie jest jasne, skąd wynika fakt pominięcia tego rodzaju anomalii w badaniach. Nie jestem pewien, czy takie zaniedbanie da się to uzasadnić statystykami.
- W odniesieniu do wyników zaprezentowanych w rozdz. 4 (np. na rys. 4.7-4.8) brak mi bliższych danych opisujących rozproszenie poszczególnych wartości wokół podanych średnich.
- Analiza skupień opiera się na reprezentacji danych w określonej przestrzeni wielowymiarowej (czego Doktorant jest świadomy). W celu stwierdzenia podobieństwa między różnymi obiektami wprowadza się metrykę odległości oraz odpowiednie progi decyzyjne (co Doktorant także uczynił). Brakuje mi jednak przy okazji definiowaniu tych progów — odnoszących się do promienia sąsiedztwa (na str. 114 i 132) — podania informacji, jakiego rodzaju metryka jest używana.

6. ORYGINALNOŚĆ ROZPRAWY, SAMODZIELNY DOROBEK AUTORA, POZYCJA ROZPRAWY W STOSUNKU DO STANU WIEDZY (POZIOM TECHNIKI) PREZENTOWANEGO W LITERATURZE ŚWIATOWEJ. Na czym polega oryginalność rozprawy, co stanowi samodzielny i oryginalny dorobek Autora, jaka jest pozycja rozprawy w stosunku do stanu wiedzy czy poziomu techniki reprezentowanych przez literaturę światową?

Nowatorski charakter rozprawy opiera się na dwóch opisywanych wcześniej propozycjach: adresacyjnej/protokołowej oraz dotyczącej systemu wykrywania

anomalii w specyficznych infrastrukturach sieci budynkowych. **Zaproponowane rozwiązania są wartościowe i stanowią postęp badawczy, przy czym lepiej umotywowany i ciekawszy wydaje mi się system nastawiony na wsparcie bezpieczeństwa** (np. lepsze osadzenie na tle stanu techniki prezentowanego w literaturze).

O oryginalności dokonań Doktoranta świadczą także pośrednio (ale za to bardzo dobitnie) jego dokonania publikacyjne, na których oceniana rozprawa jest w dużym stopniu oparta. Jest to istotny dorobek wspierający pracę doktorską (zapoznałem się również z publikacjami Autora przywoływanymi w spisie literatury) i — według mojej wiedzy — znacznie wykracza poza typowe dorobki publikacyjne doktorantów zajmujących się w Polsce telekomunikacją:

- Doktorant jest już współautorem co najmniej dziewięciu publikacji (pozycje [17], [21], [40], [47], [49], [50], [52], [80-82] w spisie literatury); część z nich ukazała się w periodykach o zasięgu międzynarodowym. Niektóre to uznane czasopisma z dziedziny telekomunikacji (ACM Computing Surveys czy Computing Networks). Jedna z prac ma już 52 cytowania w bazie Google Scholar. To ogromny sukces na tym etapie kariery naukowej.
- Doktorant nie tylko aktywnie publikuje. Jego dorobek obejmuje także udział w międzynarodowych projektach badawczych, przy okazji których powstała rozprawa doktorska. Skutkuje to także publikacjami w gronie autorów zagranicznych.

Wszystkie wspomniane tu aspekty wskazują na istotny dorobek Autora rozprawy, który jest zauważalny w środowisku międzynarodowym.

7. Poprawność przedstawienia uzyskanych wyników. Czy Autor wykazał umiejętność poprawnego i przekonywującego przedstawienia uzyskanych przez siebie wyników (zwięzłość, jasność, poprawność redakcyjna rozprawy)?

Doktorant posiada **umiejętność niezwykle jasnego przekazywania myśli**. Nie zauważyłem żadnego fragmentu, który byłby trudno zrozumiały z punktu widzenia merytorycznego. Opis jest sporządzony w taki sposób, że w zasadzie nawet osoba, która nie jest specjalistą w przedmiocie rozprawy, może ją swobodnie zrozumieć, o ile tylko dysponuje wiedzą nt. podstaw działania protokołów teleinformatycznych. Doktorant nadbudowuje następujące po sobie opisy techniczne w sposób zrozumiały, wychodząc od podstaw. Doceniam również dbałość o objaśnienia terminologiczne, także w odniesieniu do słów, które w teleinformatyce miewają nieco inne znaczenie niż w obszarze interesującym dla Doktoranta (dobry przykład to zwrócenie uwagi na specyficzne rozumienie terminu „kontroler" w przypisie 2 na str. 16). Odbiór pracy polepszają również liczne kolorowe, czytelne, eleganckie i dobrze pomyślane merytorycznie ilustracje (np. szczegóły proponowanych elementów protokołów wprowadzono z użyciem bardzo jasnych w interpretacji diagramów sekwencji). Przydatne, szczególnie od strony jasnego zestawiania różnych treści dotyczących wiedzy zastanej, są również sporządzone przez Doktoranta tabele.

Faktem jednak jest, że **praca zawiera niemało niedoróbek edycyjnych**:

- Struktura pracy jest wprawdzie czytelna, ale ziarnistość zdefiniowania rozdziałów niezbyt służy dobremu podziałowi treści. Należy tutaj przede wszystkim zwrócić uwagę na niezbalansowane długości rozdziałów. Przede wszystkim występują nieproporcjonalnie długie w stosunku do reszty rozprawy kluczowe z punktu widzenia oryginalności rozdziały 3 i 4. Ponadto konstrukcja odpowiadających sobie elementów nie jest przeprowadzona konsekwentnie, np. w przypadku podrozdziału 3.1, który służy do wprowadzenia wiedzy zastanej związanej z tezą 1, Doktorant wyraźnie podzielił tematykę na różne aspekty i nawet dla każdego z nich zdefiniował pięć różnych sekcji. Inaczej jest w przypadku przeglądu literaturowego związanego z tezą 2, którego dokonano w podrozdziale 4.2 — można byłoby co najmniej wyraźnie wydzielić różne podtematy (np. ataki na infrastruktury będące przedmiotem zainteresowań Autora, sposoby ich monitorowania, opis metod uczenia maszynowego stosowanego w wykrywaniu anomalii itp.).
- Uwagi terminologiczne:
    - W podrozdziale 2.2. Doktorant mówi o „przepływności", „przepustowości" oraz „szybkości transmisji". Są to w zasadzie różne terminy (ich odpowiedniki angielskie to zazwyczaj „capacity", „throughput" oraz „bitrate"). W tym przypadku mam jednak wrażenie, że Doktorant rozumie przez nie to samo. Sprawa ta wymagałaby objaśnienia poszczególnych terminów (jeśli jednak chodzi o różne pojęcia) lub ujednolicenia terminologii. W przypadku tekstu technicznego nagromadzenie synonimów nieco utrudnia odbiór treści, chociaż w przypadku tej pracy opisywana tu kwestia nie ma akurat znaczenia krytycznego.
    - Autor konsekwentnie nadużywa terminu „technologia", co jest kalką z języka angielskiego. Tymczasem odpowiednikiem polskim słowa *technology* jest w wielu przypadkach raczej „technika", gdyż aspekt wytwórczy w odniesieniu to tematyki tej pracy doktorskiej czy w ogóle informatyki technicznej i telekomunikacji bardzo rzadko jest przedmiotem zainteresowań. Inne często występujące kalki językowe to zastosowanie zwrotu „bazować na" zamiast „opierać się na" czy „mapować" zamiast „odwzorować".
- Rozprawa doktorska ma charakter monograficzny, zatem zgodnie ze zwyczajem akademickim należałoby uporządkować bibliografię nie według kolejności cytowania, a raczej z uwzględnieniem porządku opartego na nazwiskach pierwszych autorów prac, do których odwołuje się Doktorant.
- Doceniam, że Autor zdecydował się zamieścić wykaz anglojęzycznych skrótów. Zwracam jednak uwagę, że dla czytelnika byłoby wygodne otrzymać jednocześnie odpowiedniki polskie wylistowanych terminów, tym bardziej że jest ich dosyć dużo i dotyczą różnorodnych obszarów. Ponadto w tekście rozprawy pojawiają się skróty, których nie ma we wspomnianym wykazie (np. BLE, DSSS) — nie jest jasne, czym właściwie Doktorant kierował się wybierając skróty do pominięcia. W każdym razie dla mnie nie jest to oczywiste (wydaje mi się, że niekoniecznie jest to związane z istotnością odpowiedniego terminu w kontekście zagadnień opracowanych w rozprawie).

- W pracy umieszczono szereg przydatnych rysunków i tabel, ale niestety brakuje spisu tych rysunków i tabel, a to nieco utrudnia czytanie, jeśli chciałoby się sięgnąć do wcześniej widzianego materiału.
- W tekście rzuca się w oczy duża liczba potknięć interpunkcyjnych. Z jednej strony Doktorant w wielu przypadkach przenosi zasady interpunkcyjne z języka angielskiego (które niekiedy są odmienne od przyjętych w języku polskim), a czasem po prostu gubi znaki przestankowe (np. „...oraz mechanizmów sieciowych a następnie" na str. 5; „Użycia tak zdefiniowanego identyfikatora HI w każdym pakiecie jest nieefektywne ponieważ..." na str. 39).
- Niedopatrzenia redakcyjne, typograficzne itd.:
    - tabele bywają sformatowane niewygodnie dla czytelnika: na przykład „przeskakują" między stronami, mimo że ich długość tego nie uzasadnia, tj. nie przekracza strony (jak w przypadku tab. 4.1.), używają kolumn o długości prowadzącej do niepoprawnego dzielenia wyrazów (np. tab. 4.2) albo zawierają w różnych kolumnach te same informacje, chociaż nieco inaczej wyglądające, co powoduje, że niepotrzebnie należy dokładnie czytać redundantne teksty — warto byłoby połączyć kolumny, co też pozwoliłoby łatwiej uchwycić podobieństwa (np. tab. 4.8);
    - pozostawianie dużych wolnych przestrzeni na końcu strony ze względu na zamieszczenie w dalszym ciągu tabeli lub rysunku (np. str. 17, 55, 72 czy 120) — jest to zapewne pochodna edycji rozprawy w sposób domyślny w programie Word (który jednak jest w stanie poradzić sobie poprawnie typograficznie z takim problemem);
    - niekonsekwencje lub niewielkie niepoprawności opisu, niezręczności językowe itp.:
        - niekiedy bardzo irytujące (przez nagromadzenie potknięć i chyba świadczące o niedokładnym przejrzeniu tekstu przed drukiem) fragmenty zawierające ewidentnie błędne gramatyczne lub nielogiczne konstrukcje (np. „opisano analizy matematycznej zagadnienia oraz zrobiono symulacji [...] zastawiając wyniki..." (str. 6) , „zaproponowano nowego rozwiązania IoT wykorzystujących" czy „opisano współdzielenia funkcji" na str. 13, „2.1.1. Zarządzanie zużyciem energii elektrycznej i wykorzystania..." na str. 19, „przedstawione szerzej w m.in. w..." na str. 22, „wykorzystywany stosach protokołów" na str. 25, „zostanie wyemitowane zostaje" na str. 36, w ogóle urwane zdanie „Przedstawienie propozycji zostało poprzedzone przeglądem literatury dotyczącym wykrywania anomalii w" na str,. 79, „pomiary pochodzące z pomiarów RSS" na str. 86, „urządzenia końcowe komunikują się ze sobą za pośrednictwem węzła centralnego, np. urządzenia wykorzystujące." na str. 97);
        - wielokrotnie występujące nagromadzenie w jednym zdaniu takich samych wyrazów, co często nie razi tak bardzo w języku angielskim, ale w polskim uchodzi za błąd stylistyczny (np. „W szerszym kontekście zarządzania

zużyciem energii obejmuje również zarządzanie wykorzystaniem energii..." na str. 19, „...obniżenia kosztów dostaw energii a także zwiększenia efektywności wykorzystania energii oraz integracji rozproszonych źródeł energii w tym energii źródeł energii..." czy „urządzenia mierzące pobór mocy urządzeń..." na str. 20, „Minimalizowanie [...] funkcji zarządzania ogranicza możliwości zarządzania..." na str. 79);

- gubienie numerowanych odnośników do rysunków lub tabel (np. „została przedstawiona na rysunku" na str. 41, „jak pokazano w tabeli." na str. 63);
- zapis formalny: opuszczone formatowanie zmiennych, które powinno odbywać się z użyciem kursywy matematycznej (np. powinno być raczej „*n*-tego" niż „n-tego" albo „segmentu n+1" zamiast „segmentu n+1" na str. 63, „A_j" na str. 117), niezbyt konsekwentny zapis formuł w sekcji 4.3.3, użycie kropki dziesiętnej zamiast przecinka (np. „2.4 GHz" na str. 25);
- wadliwe odniesienia do obiektów strukturalnych, których nie ma (np. „przedstawione w rozdziale 0" na str. 68, „W kolejnych rozdziałach niniejszej rozprawy" użyte na str. 81 w rozdz. 4, podczas gdy jedyny kolejny rozdział to już tylko rozdz. 5, który stanowi po prostu podsumowanie pracy; „W załączniku wyjaśniono..." na str. 100 — praca nie zawiera załączników);
- w spisie literatury:
  - niejasności odnośnie do typu dokumentu ze względu na niepełny opis (np. [30], [33], [39], [91], [102]),
  - czasem brak wielkich liter w tytułach czasopism (np. [7], [9], [10], [95]) lub tytułach przywoływanych tekstów (np. [24], [38], [100]),
  - braki w opisach bibliograficznych: miejsce wydania książek czy odbywania się konferencji naukowych,
  - dziwaczne symbole „${$" i „$}$" w opisie pozycji [84];
- brak odmiany nazwisk (np. „został opisany m.in. przez Nikos Komninos" na str. 20, „pod kierownictwem Tiago Mendes" na str. 21, „opracowanych przez Nikhil Naikal..." na str. 22);
- pomijanie polskich znaków diakrytycznych (np. „zarzadzania" na str. 23, „powiazanie" na str. 36, „z kontrola siły" na str. 79);
- używanie czasownika „posiadać" zamiast „mieć" w stosunku do rzeczowników nieosobowych (np. „urządzenia końcowe powinny posiadać" na str. 11);
- rusycyzmy/anglicyzmy związane z niepoprawną kolejnością złożeń rzeczowników z przydawkami (np. „bezprzewodową komunikację" na str. 26, „4.4.4 Lokalna klasyfikacja" na str. 121);

- wprowadzanie niepotrzebnego odstępu między słowem a następującym po nim znakiem przestankowym (np. „zdarzeniami ," na str. 6), brak kropki na koniec zdania (np. „...system nazewnictwa EPC Stał się jednym..." na str.42, podobny problem na str. 68), brak przerwy między wartością a związaną z nią jednostką (np. „50m zamiast 100m" na str. 25) albo wyjustowania fragmentu tekstu (np. wprowadzającego do podrozdziału 2.2 na str. 22).

Gdyby Doktorant ponownie wydawał rozprawę na pewno warto z tego punktu widzenia przejrzeć jej nową wersję, stąd podaję liczne przykłady, ale nie wyczerpują one listy wszystkich niedoróbek edycyjnych.

8. SŁABE STRONY ROZPRAWY, JEJ GŁÓWNE WADY. Jakie są słabe strony rozprawy i jej główne wady?

W przypadku rozprawy można wskazać trzy słabsze punkty, które podsumowują wcześniej omówione szerzej kwestie:

- **Umiarkowana zwartość koncepcyjna pracy:** wprawdzie całość pracy dotyczy jednego obszaru, tj. zarządzania infrastrukturą sieciową urządzeń Internetu Rzeczy w inteligentnych budynkach, ale Doktorant zajmuje się dwoma w zasadzie nie połączonymi zagadnieniami. W efekcie przedstawia szczegółowe rozwiązania dwóch odrębnych problemów. Pracę oczywiście spina wprowadzenie (także opisujące aspekty techniczne) i podsumowanie, niemniej jednak wkład rozprawy w dyscyplinę nie stanowi jednej spójnej całości.
- **Zbyt uboga analiza literaturowa:** niewyodrębnienie w jasny sposób części dotyczącej analizy bibliografii wraz z jasnymi wnioskami i źródłami inspiracji oraz skromniejsza reprezentacja prac opublikowanych po 2016 r. Rozumiem, że wynika to z faktu, iż Doktorant od wielu lat prowadził prace badawcze, które były indukowane przez różne projekty, a dużą część pracy oparł na tekście opublikowanych wcześniej przez siebie artykułów naukowych — niemniej jednak praca doktorska stanowi w niezależną całość, która jako taka została złożona i jest oceniana z punktu widzenia roku 2020.
- **Nie w pełni staranna edycja pracy**: wprawdzie nie jest to aspekt istotnie zaburzający odbiór pracy od strony merytorycznej, ale na pewno wskazane byłoby lepsze zadbanie o stronę formalną.

9. PRZYDATNOŚĆ ROZPRAWY DLA NAUKI, PRZEMYSŁU, OBRONNOŚCI KRAJU ITP. Jaka jest przydatność rozprawy dla nauk technicznych?

Opiniowana praca doktorska bez wątpienia jest **przydatna z punktu widzenia nauk inżynieryjno-technicznych oraz zastosowań badań w przemyśle.**

Po pierwsze, dotyczy niezwykle interesujących dla środowiska naukowego, ale też biznesowego, tematów IoT oraz *Smart Home* (zagadnienia o dużym potencjale ze względu na poprawę komfortu życia przeciętnego konsumenta). Po drugie, Doktorant zajmuje się ważnymi aspektami tych tematów, tj. w ogólności zarządzaniem (ale też sterowaniem) takimi sieciami, w tym mechanizmami organizacji komunikacji oraz problematyką zapewniania bezpieczeństwa. Po trzecie, uzyskane przez Doktoranta **wyniki są pomysłowe, oryginalne, przydatne, jak również przedstawione na tle uznanej literatury i zilustrowane doświadczalnie w sposób pokazujący ich potencjał aplikacyjny**.

10. Podsumowanie (czy rozprawa spełnia wymagania przez obowiązujące przepisy)

Przedstawiona rozprawa doktorska proponuje oryginalne rozwiązania dwóch różnych problemów w zakresie nauk inżynieryjno-technicznych. Zagadnienia zostały poprawnie postawione przez Doktoranta. Całość pokazuje, że Autor ma odpowiednią wiedzę oraz doświadczenie w zakresie informatyki technicznej i telekomunikacji, przede wszystkim ze wskazaniem na obszary: Internetu Rzeczy, projektowania protokołów telekomunikacyjnych i użycia uczenia maszynowego. Nie mam wątpliwości, że Doktorant potrafi formułować ważkie problemy techniczne i prowadzić związane z nimi badania naukowe, jak również trafnie je opisywać (czego dowodzi imponująca liczb publikacji). Niezaniedbywalne są również jego umiejętności praktyczne.

Stwierdzam zatem, że **recenzowana rozprawa doktorska spełnia wymagania** stawiane przez obowiązujące przepisy. Z tego powodu wnoszę o dopuszczenie jej do publicznej obrony.

11. Ocena rozprawy. Do której z następujących kategorii Recenzent zalicza rozprawę (niepotrzebne skreślić)?

~~a. Nie spełniająca wymagań stawianych rozprawom doktorskim przez obowiązujące przepisy.~~
~~b. Wymagająca wprowadzenia poprawek i ponownego recenzowania.~~
~~c. Spełniająca wymagania.~~
d. Spełniająca wymagania z wyraźnym nadmiarem.
~~e. Wybitnie dobra, zasługująca na wyróżnienie.~~

Dr hab. inż. Adrian Kliks, prof. uczelni
Instytut Radiokomunikacji
Wydział Informatyki i Telekomunikacji
Politechnika Poznańska

Poznań, 15 września 2020 r.

KWESTIONARIUSZ - RECENZJA ROZPRAWY DOKTORSKIEJ DLA RADY WYDZIAŁU ELEKTRONIKI I TECHNIK INFOMRACYJNYCH POLITECHNIKI WARSZAWSKIEJ

Tytuł rozprawy: „Systemy zarządzania infrastrukturą sieciową inteligentnych budynków"

Autor rozprawy: mgr. inż. Mariusza GAJEWSKIEGO

Informacje wstępne

Niniejsza recenzja rozprawy doktorskiej została wykonana na podstawie uchwały Rady Naukowej Dyscypliny Informatyka Techniczna i Telekomunikacja Politechniki Warszawskiej z dnia 30 czerwca 2020 r. oraz pisma przewodniego z dnia 21 lipca 2020 r. Przedmiotem recenzji jest rozprawa doktorska pt. „Systemy zarządzania infrastrukturą sieciową inteligentnych budynków" autorstwa p. mgr. inż. Mariusza GAJEWSKIEGO, realizowana w Politechnice Warszawskiej, której kierownikiem naukowym i promotorem jest p. dr. hab. inż. Jordi Mongay Batalla. Rozprawa została zgłoszona jako zrealizowana w dyscyplinie „informatyka techniczna i telekomunikacja".

1. **Jakie zagadnienie naukowe jest rozpatrzone w pracy / teza rozprawy / i czy zostało ono dostatecznie jasno sformułowane przez autora? Jaki charakter ma rozprawa (teoretyczny, doświadczalny, inny)?**

Pomijając streszczenie przedłożona praca doktorska składa się z pięciu rozdziałów, w których Autor przedstawia kolejno:

- w rozdziale pierwszym – wstęp i tezę pracy,
- w rozdziale drugim – wprowadzenie teoretyczne dotyczące obecnych obszarów zastosowań domowych sieci telekomunikacyjnych,
- w rozdziale trzecim – zagadnienia związane z adresowaniem węzłów w sieci na podstawie identyfikatorów,
- w rozdziale czwartym – zagadnienia dotyczące funkcji wykrywania anomalii w sieciach domowych,
- w ostatnim rozdziale – podsumowanie osiągniętych rezultatów pracy badawczej.

Tematyka rozprawy doktorskiej dotyka dwóch, powiązanych tematycznie jednak dość istotnie niezależnych od siebie zagadnień. W pierwszej kolejności Autor przedstawia ocenę możliwości zastosowania nowego sposobu adresacji węzłów w sieci domowej Internetu Rzeczy (ang. *Internet of Things*, IoT) opartego na identyfikatorach, które mogą być zrozumiałe przez człowieka. Drugim podejmowanym przez Doktoranta zagadnieniem jest możliwość współdzielenia wybranych zasad zarządzania pomiędzy różne podmioty odpowiedzialne za korzystanie i dostarczanie usług IoT dla zastosowań domowych (w szczególności pomiędzy użytkownika domu i dostawcę usług internetowych).

Wspólnym mianownikiem obu zagadnień jest wykorzystanie proponowanych rozwiązań do sieci IoT w budynkach mieszkalnych, ale także w firmach czy budynkach użyteczności publicznej. Sieci te, co również słusznie podkreśla Autor, charakteryzują się istotnymi wymaganiami na poziom zużycia energii i bezpieczeństwa. Mianowicie poszczególne węzły sieci IoT mogą być zasilane bateryjnie, w konsekwencji łączne zużycie energii w sieci i ogólnie poziom skomplikowania stosowanych rozwiązań powinny być pod tym kątem zoptymalizowane. Dodatkowo istotnym aspektem jest również bezpieczeństwo takich sieci – wraz ze wzrostem ich popularności należy spodziewać się coraz silniejszych i częstszych ataków cyber-przestępców. Kluczowe jest więc opracowanie takich rozwiązań, które będą mogły być stosowane powszechnie i jednocześnie które będą uwzględniać specyfikę sieci IoT (wspomniane założone niskie zużycie energii przez węzły). Oba zagadnienia są w pracy jednak traktowane dość rozłącznie, a Autor stawia w pracy dwie odrębne tezy badawcze:

- Teza 1 (powiązana z pierwszym ze wspomnianych zagadnień naukowych podejmowanych w pracy): *Jest możliwe zastosowanie nowego systemu adresacji opartej na identyfikatorach do określenia nazwy i lokalizacji obiektów dołączonych do sieci inteligentnego budynku*
- Teza 2 (powiązana z drugim zagadnieniem – cyberbezpieczeństwa sieci): *Jest możliwe rozdzielenie wykrywania anomalii w sieci teleinformatycznej inteligentnego domu pomiędzy użytkownika oraz dostawcę usług.*

Tezy te i proponowane przez Autora rozwiązania są omawiane odpowiednio w dwóch głównych rozdziałach tematycznych pracy – rozdziale trzecim (Teza 1) oraz czwartym (Teza 2). W tym kontekście uważam, że postawione tezy badawcze zostały przedstawione i wyjaśnione przez Autora w sposób klarowny, choć w mojej ocenie nieco zbyt ogólny. Niemniej jednak cele pracy zostały nakreślone w sposób prawidłowy i nie budzą wątpliwości.

Praca ma charakter koncepcyjno-eksperymentalny, co na pewno stanowi jest istotny walor. Względem pierwszej tezy Autor zaproponował nowe podejście do sposobu adresacji węzłów w sieci IoT, przedstawiając efektywność rozwiązania za pomocą przeprowadzonych eksperymentów wykonanych na przygotowanej platformie komputerowej. W odniesieniu do drugiej tezy Autor zaproponował sposób współdzielenia funkcji detekcji anomalii pomiędzy użytkownika systemu i dostawcę usług z wykorzystaniem algorytmów sztucznej inteligencji, a także przedstawił rezultaty przeprowadzonych eksperymentów symulacyjnych z użyciem m.in. środowiska Omnet++.

2. Czy w rozprawie przeprowadzono w sposób właściwy analizę źródeł / w tym literatury światowej, stanu wiedzy i zastosowań w przemyśle / świadczący o dostatecznej wiedzy autora? Czy wnioski z przeglądu źródeł sformułowano w sposób jasny i przekonywujący?

Przedstawione przez Autora rozwiązania dla obu zagadnień badawczych zostały poprzedzone analizą obecnego stanu wiedzy oraz techniki. Autor przedstawił zarówno istniejące rozwiązania z zakresu sposobów adresacji stosowanych w różnego rodzaju sieciach, kładąc nacisk oczywiście na sieci IoT, jak również opisał techniki używane w celu detekcji anomalii w funkcjonowaniu sieci, które mogą świadczyć o przeprowadzanych lub planowanych atakach. Analiza została wykonana w sposób prawidłowy i przejrzysty, pozwalając czytelnikowi na właściwe umiejscowienie proponowanych rozwiązań w nurcie prowadzonych prac na arenie międzynarodowej.

Przedstawiony przez mgr. inż. Mariusza Gajewskiego spis wykorzystywanych źródeł i publikacji naukowych jest wystarczający (102 pozycje). Warto podkreślić różnorodność cytowanych prac (np. publikacje standaryzacyjne, publikacje naukowe, książkowe), a także aktualność i kompletność dobranej literatury. Dobór ten świadczy o dużej wiedzy Doktoranta i ogólnym rozeznaniu w prowadzonych na świecie działaniach.

Jedynie w przypadku drugiego zagadnienia powstaje wątpliwość, czy przegląd obecnego stanu techniki został wykonany w sposób wyczerpujący – zabrakło mi odniesień do stosowanych obecnie rozwiązań znanych z praktyki np. z ogólnych programów antywirusowych oferowanych przez dużych (ale i mniejszych) graczy oferujących usługi związane z cyber-bezpieczeństwem. Nie jest bowiem niczym nowym obserwacja, że wielostopniowe podejście w zakresie analizy bezpieczeństwa (nawet antywirusowego) jest stosowane powszechnie (część zadań wykonywanych lokalnie, część u dostawcy usług, a część także u dostawców samego oprogramowania na poziomie krajowym czy globalnym). Znani dostawcy usług z zakresu bezpieczeństwa (np. Kaspersky, Norton i inne) mają w swojej ofercie systemy bezpieczeństwa dla IoT. Także w odniesieniu do systemów zarządzania inteligentnymi budynkami (ang. *Building Management System,* BMS) można odnaleźć bogatą ofertę zaawansowanych narzędzi do oceny bezpieczeństwa systemów automatyki budynkowej (w tym więc także IoT), np. rozwiązania stosowane przez firmę Honeywell. Ostatecznie również dostawcy systemów dla inteligentnych domów (np. Nice i inni) także oferują różnego typu rozwiązania mające na celu zapewnienie bezpieczeństwa swoich klientów. W pracy zabrakło mi analizy tego typu istniejących rozwiązań, które mogłyby być albo zastosowane bezpośrednio albo z drobną modyfikacją dla systemów IoT. Tego typu analiza wydaje się szczególnie istotna w sytuacji, gdy praca doktorska ma dominujący charakter eksperymentalny i praktyczny.

3. Czy autor rozwiązał postawione zagadnienia, czy użył właściwej do tego metody i czy przyjęte założenia są uzasadnione?

Pomimo wspomnianego mankamentu przedstawiona analiza stanu wiedzy jest wystarczająca i pozwoliła Autorowi na prawidłowe rozwiązanie analizowanych problemów badawczych. W odniesieniu do postawionych tez i rozpatrywanych zagadnień uważam, że przyjęte założenia są słuszne i uzasadnione. Ponadto Doktorant zaproponował przetestowanie poprawności swoich rozwiązań koncepcyjnych na gruncie eksperymentów komputerowych – założenie to jest również poprawne z perspektywy metodologii pracy.

Jedyną wątpliwość budzi przyjęcie założenia o relatywnej statyczności ruchu generowanego przez poszczególne węzły w sieci IoT inteligentnego budynku (braku gwałtownych zmian). Kwestię tę wyjaśniam bardziej szczegółowo w dalszej części recenzji.

4. Na czym polega oryginalność rozprawy, co stanowi samodzielny i oryginalny dorobek autora, jaka jest pozycja rozprawy w stosunku do stanu wiedzy czy poziomu techniki reprezentowanego przez literaturę światową?

Analizując wkład własny Autora przy powstaniu zaproponowanych rozwiązań należy zauważyć, że dominująca większość publikacji przypisanych p. mgr inż. Mariuszowi Gajewskiemu jest wieloautorska. Świadczy to oczywiście z jednej strony o rozpoznawalności kandydata w środowisku naukowym i o umiejętności prowadzenia współpracy na gruncie międzynarodowym, co jest pożądaną cechą, z drugiej strony utrudnia precyzyjne wskazanie, które fragmenty przedstawionej pracy są rzeczywistym pomysłem Doktoranta. Bardzo pomocne byłoby jednoznaczne wskazanie przez Doktoranta, które dokładnie fragmenty opisywanych w dysertacji rozwiązań są wyłącznie jego pomysłu. Niemniej jednak opierając się na przedłożonej rozprawie doktorskiej można wskazać na następujące, kluczowe osiągnięcia Kandydata:

- w odniesieniu do Tezy 1, został zaproponowany nowy schemat adresowania obiektów Internetu Rzeczy, który może być z powodzeniem zastosowany do budynków mieszkalnych i przemysłowych. Podstawową cechą tego podejścia jest możliwość użycia identyfikatorów urządzeń (węzłów) do bezpośredniego ich adresowania, uwzględniając także oferowane przez nie usługi. W szczególności wykorzystano specjalną warstwę w ramach stosu protokołów nazwaną jako *ID Layer*, która

pozwala przeprowadzić routing wiadomości opierając się na wspomnianych identyfikatorach. Ponadto w rozwiązaniu założono, że węzły przechowują i aktualizują informację o swoich sąsiadach, co umożliwia uproszczenie procedur kwerendowych kierowanych do sieci ze strony klienta. Następnie zaproponowana została metoda adresowania węzłów i usług (do 64 poziomów w hierarchii), w której wykorzystano adresowanie alfanumeryczne zrozumiałe łatwo także przez człowieka. W pracy Autor opisał także różne przypadki aplikacyjne, przedstawiając szczegółowo proponowane formaty wiadomości wymieniane przez węzły, jak również sekwencje ich przekazywania. Doktorant przedstawił także wyniki przeprowadzonych przez niego eksperymentów symulacyjnych. Ocena powyższego rozwiązania wskazuje na jej innowacyjny charakter w porównaniu z dostępnymi na rynku produktami, jest ono kompletne i rzetelnie przetestowane w różnych wariantach aplikacyjnych. W dalszej części recenzji chciałbym jednak podjąć polemikę z Autorem na temat sugerowanych rozwiązań.

- w odniesieniu do Tezy 2 Autor wskazuje na model architektoniczny systemu, w którym decyzje o potencjalnych anomaliach w ruchu sieci są podejmowane na dwóch poziomach: lokalnym realizowanym w bramie domowej (nazywanej przez Autora *Home Gateway*, HG) oraz globalnym w centrum przetwarzania danych dostawcy usług. Podkreślona została zaleta tego podejścia w postaci dużych możliwości obliczeniowych oraz większej świadomości potencjalnych ataków, które posiada dostawca usług. Doktorant przeprowadził analizę ruchu i wytypował kilka klas potencjalnych zagrożeń w sieciach IoT, a następnie przedstawił propozycję algorytmów klasyfikacji i grupowania anomalii. Całość rozwiązania została przetestowana na drodze eksperymentów komputerowych. Oceniając tę część pracy należy podkreślić spory wkład Doktoranta w rzetelne przeprowadzenie badań eksperymentalnych. Ich nowatorskość została podkreślona kilkoma wieloautorskimi publikacjami naukowymi. Jak jednak wspomniano wcześniej, analiza rzeczywistej innowacyjności zaproponowanych rozwiązań w porównaniu ze stanem techniki w zakresie cyberbezpieczeństwa jest utrudniona przez brak – według recenzenta – dogłębnego porównania praktycznych rozwiązań w tym zakresie. Pomimo to uważam wkład Doktoranta w rozwój dziedziny bezpieczeństwa sieci domowych IoT za istotny.

5. Czy autor wykazał umiejętność poprawnego i przekonywującego przedstawienia uzyskanych przez siebie wyników / zwięzłość, jasność, poprawność redakcyjna rozprawy/?

Analizując sposób przedstawienia swoich osiągnięć należy podkreślić bardzo dobrą strukturę rozprawy. Założenia prowadzonych prac badawczych zostały przedstawione w sposób czytelny, wprowadzenie do problematyki naukowej także nie budzi wątpliwości. Ze względu na postawienie dwóch tez badawczych przyjęta przez Doktoranta struktura pracy została przyjęta właściwie, tzn. przygotowane zostały dwa dominujące rozdziały pokrywające tematycznie dwa obszary badacze. Myśli formułowane przez Doktoranta są w sposób czytelny i zwięzły. Prezentowane wyniki w postaci wykresów i rysunków są w ogólności czytelne, a ich opis jest przeważnie wystarczający do zrozumienia ich znaczenia. Niestety dość sporym mankamentem pracy jest duża liczba uchybień zarówno edytorskich, jak i językowych, świadcząca prawdopodobnie o nie dość starannym przejrzeniu wersji końcowej pracy. O tych błędach piszę w dalszej części recenzji. Ogólnie przygotowanie całej pracy oceniam jako dobre.

6. Jakie są słabe strony rozprawy i jej główne wady?

Jak wspomniano we wcześniejszych fragmentach recenzji, pewnych wyjaśnień lub komentarzy wymaga kilka kwestii, które wymieniam poniżej:

a. Właściwe byłoby precyzyjne wskazanie (poprzez zebranie w jednym miejscu), które propozycje i osiągnięcia opisane w pracy doktorskiej są samodzielnym dorobkiem Doktoranta. Obecnie można

to wywnioskować z samego opisu, jednak takie zwięzłe (aczkolwiek precyzyjne) podsumowanie byłoby dobrym uzupełnieniem rozprawy.

b. Pewnym mankamentem pracy są dość liczne błędy edycyjne, interpunkcyjne i językowe, świadczące jednoznacznie o nie dość rzetelnej korekcie ostatecznej wersji pracy (są bowiem spore fragmenty rozprawy napisane bardzo ładnie). Poniżej podaję kilka przykładów:
   - *(s. 5) ... znajdują zastosowanie w następujących obszarach: zastosowania ...*
   - *(s. 5) To rozwiązanie jest może mieć praktyczne zastosowanie...*
   - *(s. 6) W rozprawie opisano analizy matematycznej zagadnienia oraz zrobiono symulacji wykonanych dla tak zaproponowanego procesu...*
   - *(s. 6) Opisano skuteczność takiego podejścia, zastawiając wyniki...*
   - *(s. 10) W rozdziale 3 zaproponowano nowego rozwiązania IoT wykorzystujących warstwę...*
   - *(s. 16) Charakteryzują się tym, że, są zasilane..*
   - *...*
   - *(s. 36) zostanie wyemitowane zostaje ogłoszenie*
   - *(s 37) Numery te są przydzielane przez w trakcie procesu... Urządzenia mogą również w udostępniać informacje...*
   - *(s. 42) W ten sposób dla kod EPC może być reprezentowany...*
   - *(s. 48) w celu poprawę wydajność...*
   - *(s. 50)* brak czasownika w zdaniu *„Natomiast w zaproponowanym rozwiązaniu..."*
   - *...*
   - *(s. 70) tym przypadku prze przeprowadzono*
   - *Itd. Itd.*

   W świetle oceny treści pracy należy jeszcze zwrócić uwagę na pewien brak precyzji w stosowaniu niektórych zwrotów – co prawda typowych w mowie potocznej, jednak nie powinny one się pojawiać w rozprawie. Mam tu na myśli użycie np. słowa dedykować (rozwiązanie dedykowane do czegoś), ekosystem (*eko* jednak oznacza związek z naukami przyrodniczymi) czy funkcjonalność (która nie ma liczby mnogiej i jest rzeczownikiem niepoliczalnym). Oczywiście, na styku języka specjalistycznego i informatycznego jest to zawsze pewien problem, jednak w przypadku treści rozprawy powinna być zachowana szczególna dbałość o precyzyjność wypowiedzi. Także użycie zwrotu „węzły potomne" zamiast „podrzędne" wydaje się niewłaściwe, wskazuje bowiem na relację czasową (następstwa czasu), która przecież w rozważanym kontekście nie występuje.

c. W analizie literatury oraz dostępnych rozwiązań technologicznych, zwłaszcza w odniesieniu do kwestii cyberbezpieczeństwa, zabrakło dogłębnej analizy systemów oferowanych przed dostawców systemów dla inteligentnych budynków i automatyki przemysłowej (np. rozwiązania światowych liderów w cyberbezpieczeństwie jak Kaspersky, F-Secure, Norton, rozwiązania typu AWS Cloud lub dostawców typu Honeywell, Nice itd.)

d. Obszary zastosowań podane w rozdziale 2.1 są niepełne lub nieprecyzyjnie określone – jak na przykład zakwalifikować wszelkie działania określane z angielskiego jako *„well-being"* np. zdalne włączanie zaparzania kawy z poziomu telefonu przed przyjazdem do posesji albo automatyczne przyciemnianie światła łącznie z opuszczeniem żaluzji na potrzeby oglądania filmu po wybraniu w systemie inteligentnego budynku trybu kinowego? To oczywiście tylko przykłady.

e. Następnie w rozdziale 2.1.4 Autor pisze o systemach IoT dla bezpieczeństwa i nadzoru powołując się na rozpoznawanie twarzy itd. Jak w tym kontekście należy wg. Autora umiejscowić wszelkie

Kl-Ls

systemy monitoringu, biometrii i kontroli dostępu? Są to podklasy systemów IoT czy odrębne systemy?

f. Odnośnie tezy 1:
   - Autor w sposób zwięzły i poprawny prezentuje różnego rodzaju podejścia w adresowaniu węzłów. W pracy zabrakło jednak jasnego porównania (np., tabelarycznego), jak proponowane innowacyjne podejście przedstawione w pracy wypada na tle innych rozwiązań; jakie są jego wady i zalety w porównaniu do rozwiązań już istniejących lub proponowanych w wymienionych w rozprawie projektach?
   - W jaki sposób można efektywnie zastosować zaproponowaną koncepcję świadomości węzłów w systemach IoT, w których z założenia węzły będą miały ograniczoną pamięć lub nie będą miały jej wcale? Przecież węzły będące wyżej w hierarchii będą miały lawinowo przyrastającą ilość informacji do przechowania (sterowaną oczywiście czasem ważności wiadomości). Jak obecność takich węzłów wpływa na funkcjonowanie całego systemu? Czy konieczność przechowywania danych o węzłach sąsiadujących (podrzędnych) nie kłóci się trochę z założeniem systemu IoT w ogólności? W jaki sposób mają być uwzględnione w systemie węzły będące np. kodami QR czy znacznikami RFID?
   - Jaki jest rzeczywisty zysk z zastosowania alfanumerycznych identyfikatorów? Przecież stosowanie znaków alfanumerycznych wprowadza ogromną nadmiarowość (administratorzy systemów jako ludzie będą stosowali pewne szablony nazewnictwa – np. jak zaproponowano w pracy floor001; wiele sekwencji będzie niewykorzystanych np. XH3Gdd2i). Znaki alfanumeryczne mają wprowadzać ułatwienie m.in. w zarządzaniu. Czy jednak bardziej efektywnym energetycznie (co jest ważne dla sieci IoT) rozwiązaniem nie byłoby istotne zredukowanie rozmiaru adresu z 64 bajtów (2^{64} lub nawet 2^{512} kombinacji, niemożliwe do wykorzystania nawet w bardzo dużym budynku czy fabryce) na rzecz adresowania binarnego i automatycznego *parsera* z reprezentacji binarnej na alfanumeryczną? Dla budynku np. o 20 piętrach, po 100 pokoi na piętrze i po 100 czujników w pokoju wystarczy 5 bitów na pierwszy poziom hierarchii, 7 na drugi i 7 na trzeci (długości tych bloków binarnych mogą być ustalane w sposób dynamiczny przez administratora), zamiast 512. Przy takim podejściu sposoby np. przekazywania pakietów i ich usuwania mogłyby być niezmienione, bowiem z perspektywy analizy programistycznej porównuje się i tak ostatecznie sekwencje binarne.
   - S. 54. – w procesie rejestracji obiekty przekazują informacje do najbliższego węzła *ID layer* – w jaki sposób i kto ustala, który węzeł jest najbliższy? Czy proponowane rozwiązanie jest stosowane tylko do sieci przewodowych (jak sugeruje wpis na stronie 76)?
   - Zastosowana struktura drzewa ma pewną wadę – jeżeli mamy jeden węzeł opisany jako build001.floor1, to on obsługuje bardzo dużo węzłów podrzędnych; jak wygląda wydajność komunikacji, przechowywania danych w takim węźle w przypadku dużej liczby węzłów podrzędnych np. liczonych w setkach?
   - Skąd w równaniu (3.1) jest liczba 37?
   - Jaki wpływ na przedstawione wyniki ma długość kolejki/pamięci stosowanej na poszczególnych poziomach węzłów?
   - Czy porównanie na rys 3.16 jest sprawiedliwe – zapytanie do bazy danych jest oczywiście koncepcyjnie bardzo podobne, jednak przecież struktura zapytań, sposób ich obsługi itd. Itd. mocno wpływa na wydajność, a jest inne niż w zastosowanej metodzie dla adresacji IoT. Ponadto, dlaczego wykres pomarańczowy nie rośnie wraz ze wzrostem liczby poziomów (wydaje się, że wzrost powinien być większy)?

- Na rys. 3.17 dla 8 węzłów bez buforowania mamy opóźnienie rzędu 4 ms, zaś takie liczby nie występują na rysunku 3.16? Czym się różni pusta tablica od braku buforowania?
- Istotnym brakiem jest nieuwzględnienie wykresów pokazujących zyski energetyczne zaproponowanego rozwiązania, a przecież buforowanie miało pozwalać wyłączać węzły. Widoczne są koszty (opóźnienie, konieczna pamięć) a nie widać precyzyjnie zysków.

g. Odnośnie tezy 2:
   - Ciekawym (choć pewnie pobocznym) wątkiem proponowanego rozwiązania dla cyberbezpieczeństwa byłaby krótka analiza w perspektywie przekazywania danych osobowych o sposobie używania wybranych modułów sieci IoT dla podmiotu zewnętrznego (problem GIODO). Czy samo zanonimizowanie danych będzie wystarczające wg. Autora do praktycznego zastosowania tego rozwiązania?
   - Autor pisze (s. 13) o „modelowym zachowaniu obserwowanym w normalnych warunkach". Są to oczywiście warunki typowe dla danego budynku. Jednak czujniki (ogólnie węzły IoT) mogą posiadać przecież całą gamę parametrów konfiguracyjnych, określających sposób funkcjonowania węzła. Dla każdego domu te ustawienia mogą być różne i to istotnie – w jednym miejscu temperatura może być zbierana na żądanie z mała dokładnością, w drugim okresowo z dużą intensywnością i dużą dokładnością. Oba zachowania są typowe (normalne). Jak w tym kontekście zachowuje się zaproponowany system modelowania i analizy w centrum dostawcy usług?
   - Kontynuując powyższą myśl, potrzebne byłoby uzasadnienie założenia nr 3 (4.3 kropka 3) – w rozdziale 2 założono, że w HAN może występować ruch multimedialny, który w zależności od ustawień może się dość istotnie różnić. Dodatkowo, bardziej rozbudowane czujniki mogą reagować adaptacyjnie do zaistniałej sytuacji – np. wykrycie jakiegoś zjawiska (podwyższenie nagłe temperatury) powoduje nagłe zwiększenie częstotliwości i dokładności raportów. Jak to odnieść do przyjętego założenia? Czy możliwość adaptacji (także skokowej) różnych modułów sieci IoT nie jest właśnie cechą typową czujników sieci inteligentnej?
   - Dlaczego wskazano jako odrębne cechy w tabeli 4.3 (rząd pierwszy, czwarty i piąty), które należą do obu rozważanych grup? Jak je wyróżnić i sklasyfikować w rozważanym podejściu?
   - W treści pracy nie podano, jak w zbiorze WSN-DS definiowany jest przepływ i czy ta definicja jest zbieżna z przyjętą przez Autora pracy.
   - Jaka jest różnica pomiędzy rysunkami 4.4a i 4.4b?
   - Pod równaniem (4.3) – czy istotnie miał być wspomniany tutaj licznik ułamka czy raczej mianownik?
   - Równanie (4.10) – tutaj prawdopodobnie po prawej stronie nierówności należałoby użyć zmiennej $n_i^{(1)}$ zamiast $n_i^{(0)}$, także mianownik pewnie powinien mieć N_1. Ponadto, obie strony nierówności można uprościć poprzez przemnożenie przez $(N_0 + N_1)\Delta$. Wówczas po obu stronach nierówności mamy sumę po *i*, która równa jest odpowiednio N_0 po lewej stroni i N_1 po prawej stronie. Cały estymator sprowadzałby się więc do porównania wielkości N_0 i N_1.
   - Skoro przyjęto średnią liczbę pakietów na poziomie 3 p/min, dlaczego na wykresie 4.7 u góry żaden słupek nie osiąga tej wartości? Czy oznacza to jakieś straty?

h. Pewnym niedociągnięciem jest brak umieszczenia w pracy choćby próby teoretycznej połączenia obu naukowych części pracy i ich wzajemnego wpływu na siebie.

7. Jaka jest przydatność rozprawy dla nauk technicznych?

Pomimo przedstawionych mankamentów wysoko oceniam przydatność opisywanych w pracy rozwiązań dla nauk technicznych, głównie ze względu na aplikacyjny charakter pracy. Przedstawione koncepcje adresowania umożliwiające jednocześnie lokalizację węzłów i usług, jak i ogólnie metody zastosowania algorytmów sztucznej inteligencji do wykrywania anomalii ruchu zostały przebadane eksperymentalnie. Stanowi to więc istotny wkład w tych obszarach nauki.

8. **Do której z następujących kategorii Recenzent zalicza rozprawę:**
   a. **Nie spełniająca wymagań stawianych rozprawom doktorskim przez obowiązujące przepisy**
   b. **Wymagająca wprowadzenia poprawek i ponownego recenzowania**
   c. **Spełniająca wymagania**
   d. **Spełniająca wymagania z wyraźnym nadmiarem**
   e. **Wybitnie dobra, zasługująca na wyróżnienie**

Podsumowując swoją recenzję uważam, że praca – choć nie pozbawiona pewnych niedociągnięć przedstawionych wcześniej - spełnia z wyraźnym nadmiarem wymagania stawiane rozprawom doktorskim. Wnoszę także o dopuszczenie do publicznej obrony przedstawionej rozprawy.

Adrian Kliks

15.09.2020 r.