Szczecin, 23.01.2025 r.

prof. dr hab. inż. Krzysztof Okarma
Wydział Elektryczny
Zachodniopomorski Uniwersytet Technologiczny w Szczecinie

RECENZJA ROZPRAWY DOKTORSKIEJ
dla Senatu Politechniki Opolskiej

Tytuł rozprawy:	**Threat Modeling methods and detecting vulnerabilities in 5G networks**
Autor rozprawy:	**mgr inż. Wiktor Sędkowski**
Dyscyplina naukowa:	**automatyka, elektronika, elektrotechnika i technologie kosmiczne**
Promotor:	**dr hab. inż. Rafał Stanisławski, prof. PO**

I. TEMATYKA, TEZA NAUKOWA I CEL ROZPRAWY

Rozprawa doktorska pt. *„Threat Modeling methods and detecting vulnerabilities in 5G networks"* przedstawiona przez mgr. inż. Wiktora Sędkowskiego została napisana w całości w języku angielskim. Dotyczy ona istotnych oraz aktualnych zagadnień związanych z modelowaniem zagrożeń, jak również wykrywaniem podatności w sieciach 5G, co z jednej strony mocno wiąże się z dyscypliną naukową *informatyka techniczna i telekomunikacja*, ale tematyka ta ma równie bliski związek z nowoczesnymi rozwiązaniami z zakresu automatyki, zwłaszcza dotyczącymi automatyzacji procesu modelowania zagrożeń oraz wykrywania luk w systemach dynamicznych – w tym przypadku sieci 5G. Ze względu na użycie metod sztucznej inteligencji (SI), wywodzących się z automatyki, jak również skupienie się na modelowaniu zagrożeń w systemach dynamicznych o zmiennej strukturze, rozprawa ta dobrze wpisuje się w dyscyplinę *automatyka, elektronika, elektrotechnika i technologie kosmiczne,* choć ma silne zabarwienie interdyscyplinarne związane z *informatyką techniczną i telekomunikacją.* Należy to jednak uznać za zaletę pracy a nie wadę, gdyż w niektórych obszarach obie wymienione dyscypliny naukowe często się wzajemnie przenikają, co umożliwia efektywne użycie rozwiązań lokujących się na ich „styku", jak również ich dostosowywanie do specyficznych potrzeb. Takim przykładem są m.in. technologie sieci 5G, szczególnie istotne dla rozwoju Przemysłu 4.0 oraz rozwiązań Przemysłowego Internetu Rzeczy, czy też nowoczesnej energetyki, co słusznie zauważył Doktorant we wprowadzeniu do rozprawy, wskazując na związek tematyki rozprawy z dyscypliną naukowej, w której ubiega się On o nadanie stopnia doktora.

Ze względu na zmienną i trudną do przewidzenia naturę zagrożeń bezpieczeństwa w sieciach 5G, Doktorant postawił tezę, że narzędzia heurystyczne oparte na sztucznej inteligencji mogą okazać się szczególnie skuteczne w modelowaniu tej klasy problemów. Wskazał dwie szczegółowe hipotezy badawcze dotyczące:

- możliwości efektywnego modelowania i ustalania priorytetów zagrożeń występujących w sieciach 5G, biorąc pod uwagę zmienność struktury sieci,
- możliwości użycia zautomatyzowanych, adaptacyjnych algorytmów obliczeniowych działających w czasie rzeczywistym w celu modelowania i ustalania priorytetów zagrożeń w sieciach 5G.

Za cel rozprawy Autor postawił sobie opracowanie metodologii i algorytmów modelowania oraz oceny podatności na zagrożenia w sieciach 5G, motywując to potrzebami zapewnienia bezpieczeństwa interfejsów radiowych, jak też zachowania integralności płaszczyzny użytkownika. Wziął On pod uwagę podatność na ataki typu Denial of Service (DoS), skierowane zarówno na infrastrukturę, jak i urządzenia użytkowników końcowych. Wskazał, iż architektura sieci 5G wiąże się z nowymi, dynamicznie zmieniającymi się zagrożeniami bezpieczeństwa, co wymaga opracowania nowych sposobów modelowania tych zagrożeń, tak aby operator sieci mógł dostosować się do dynamicznych zmian infrastruktury i zapewnić bezpieczeństwo nie tylko swoich zasobów, ale także użytkowników sieci. Biorąc pod uwagę dynamiczny rozwój technologii związanych z tematyką rozprawy, jak również uzasadnienie wskazania celu oraz tez badawczych, uznać należy, iż zarówno cel rozprawy, jak też tezy zostały wskazane trafnie.

II. ZAWARTOŚĆ MERYTORYCZNA ROZPRAWY

Motywując wybór tematyki badawczej Kandydat wskazał, iż technologia 5G oferuje znaczące zalety w zastosowaniach przemysłowych, które wymagają zwiększonej szybkości transmisji danych oraz zredukowanego opóźnienia. Dotyczy to m.in. systemów opartych na rzeczywistości rozszerzonej (*Augmented Reality* - AR) i sztucznej inteligencji (*Artificial Intelligence* - AI), czy też streamingu wideo.

Jak wskazał Kandydat, badania przedstawione w rozprawie podzielone zostały na dwie zasadnicze części. Pierwsza z nich ma charakter teoretyczny i dotyczy wyboru metodologii modelowania zagrożeń z zastosowaniem iteracyjnego odkrywania wiedzy, jak również metod porównawczych, czy też podejść mieszanych bazujących zarówno na podejściach jakościowych, jak też ilościowych, m.in. w celu określenia kryteriów jakości oraz metryk dotyczących zagrożeń. Druga część pracy, wynikająca wprost z przeprowadzonej analizy, dotyczy wdrożenia zaproponowanych rozwiązań z użyciem podejścia mieszanego przy założeniu integracji metod sztucznej inteligencji z przyjętymi praktykami zapewniania cyberbezpieczeństwa.

W pierwszym rozdziale rozprawy Autor skupił się na przedstawieniu zakresu oraz celu pracy, metod stosowanych podczas badań, a także głównych elementów wkładu badawczego. Rozdział ten ma charakter wprowadzający w tematykę badawczą i systematyzuje w pewnym zakresie zawartość całej dysertacji. W rozdziale drugim przedstawiony został szczegółowy przegląd sieci 5G wraz z omówieniem ich właściwości. Przedstawione zostały zagadnienia transmisji bezprzewodowej

związane z kształtowaniem wiązki (ang. *beamforming*), technologią massive MIMO bazującą na zestawach wielu anten, dynamicznym współdzieleniem pasma, a także wirtualizacją radiowej sieci dostępowej (RAN – *Radio Access Network*). Przedstawiono także główne komponenty sieciowe, w tym m.in. wielodostępowe przetwarzanie brzegowe (ang. *multi-access edge computing*), jak też architekturę sieci 5G bazującą na usługach. Szczególnie istotnymi elementami poruszonymi w tym rozdziale są zagrożenia wymierzone w sieci 5G oraz krytyczne elementy sieci szkieletowej 5G.

W kolejnym rozdziale przedstawiono przegląd metodologii modelowania zagrożeń, w tym opartych na danych, atakach, zasobach, oprogramowaniu oraz systemie. Ponadto, przestawiono możliwości zastosowania sztucznej inteligencji do automatyzacji modelowania zagrożeń. Rozdział 4. dotyczy z kolei źródeł danych do modelowania zagrożeń, jak również metody wspomagające automatyzację przygotowywania diagramów zagrożeń. Przeanalizowano również możliwości zastosowania narzędzia ChatGPT jako bazy wiedzy do opisu zagrożeń.

Ostatni rozdział dotyczy automatycznego wykrywania luk bazującego na wykrywaniu zasobów oraz przewidywaniu zagrożeń wraz z opisem procesu pozyskiwania danych treningowych, opracowywania modeli AI do przewidywania podatnych zasobów i wzorców wykorzystania luk. Zawiera on również prezentację wyników przeprowadzonych badań.

III. OGÓLNA OCENA ROZPRAWY i UWAGI DYSKUSYJNE

Rozprawa koncentruje się na automatyzacji procesów wspierających wykrywanie zagrożeń z użyciem metod sztucznej inteligencji, co stanowi aktualny temat badawczy o znacznym potencjale wdrożeniowym. Jest to zresztą dość charakterystyczne dla tej rozprawy, choćby ze względu na fakt, iż stanowi ona doktorat wdrożeniowy, gdzie „punkt ciężkości" przesuwa się nieco bardziej w kierunku zagadnień związanych z implementacją i wdrożeniem proponowanych rozwiązań aniżeli w stronę aspektów typowo naukowych. Istotnym elementem rozprawy jest przeprowadzone badanie potencjału metod sztucznej inteligencji w odniesieniu do usprawnienia procesów modelowania zagrożeń w sieciach 5G. Do osiągnięć Doktoranta zaliczyć można również dokonanie kompleksowego usystematyzowania metodologii modelowania zagrożeń w takich sieciach wraz z rozpoznawaniem unikalnych podatności oraz wektorów potencjalnych ataków.

Istotnym osiągnięciem o praktycznym charakterze, co jest niewątpliwie cechą specyficzną dla doktoratów wdrożeniowych, jest wdrożenie i ocena prototypowego narzędzia do modelowania zagrożeń opartego na metodach sztucznej inteligencji. Integruje ono istniejące bazy wiedzy dotyczące modelowania zagrożeń właśnie z algorytmami sztucznej inteligencji oraz automatyzacją procesów identyfikacji zasobów, analizy wektorów ataku oraz określania priorytetów podatności. Warto zauważyć także identyfikację kierunków dalszych badań związanych z zastosowaniem sztucznej inteligencji w domenie cyberbezpieczeństwa, w tym integrację metod SI z innymi metodami, które zostały trafnie określone przez Kandydata.

Wśród wymienionych w recenzji osiągnięć nie wszystkie mają jednakową wagę, także pod względem naukowym, jednak za najistotniejszy element wkładu pod tym względem uznać można opracowanie nowego rozwiązania dla modelowania zagrożeń wspomaganego metodami sztucznej

inteligencji, a także jego praktyczną implementację. W pewnym aspekcie rozprawa stanowi swoisty „pomost" pomiędzy modelami teoretycznymi a scenariuszami rzeczywistych zagrożeń w sieci 5G, gdyż przedstawione spostrzeżenia zostały zweryfikowane na podstawie zebranych danych empirycznych dotyczących ruchu sieciowego. Kandydat użył do tego celu dostępnych narzędzi o otwartym kodzie źródłowym takich jak OpenVAS (*Open Vulnerability Assessment System*) czy Nmap (*Network mapper*), dzięki czemu możliwe było dokonanie oceny skuteczności różnych środków bezpieczeństwa, w tym analiza zaproponowanych modeli zagrożeń, a także symulacja różnych scenariuszy ataków w sieciach 5G. Dobrane narzędzia, jak również wybór języka Python do implementacji części praktycznej pracy, uznać należy za właściwe.

Tematyka rozprawy bardzo dobrze wpisuje się w koncepcję Przemysłu 4.0, jak również rozwoju technologii Internetu Rzeczy, systemów wbudowanych oraz technik transmisji danych docelowo stanowiących fundament koncepcji tzw. „Internetu wszystkiego" (*Internet of Everything – IoE*). Chociaż sporą część rozprawy zajmują opisy sieci 5G, nie stanowi to istotnej wady rozprawy, gdyż przykładowo podrozdział 2.3 zawiera ważne informacje dotyczące architektury sieci 5G, wskazujące także na jej złożoność, co ma niebagatelny wpływ na stopień komplikacji budowy kompletnego modelu zagrożeń. Ułatwiają one lekturę dalszych części pracy, pozytywnie wpływając na kompletność rozprawy, także z punktu widzenia czytelników mniej zorientowanych w szczegółach rozwiązań stosowanych w sieciach 5G.

Jak zauważa sam Doktorant (str. 74), idea zastosowania metod sztucznej inteligencji w obszarze cyberbezpieczeństwa, nie jest nowa, jednak pojawiające się coraz nowsze i wciąż udoskonalane rozwiązania z zakresu SI, ale także nowe rodzaje zagrożeń, powodują, iż tematyka ta niezmiennie cieszy się zainteresowaniem nie tylko naukowców, ale także przemysłu. W rozprawie przedstawiona została usystematyzowana analiza narzędzi SI stosowanych w celu modelowania zagrożeń, co jest autorskim osiągnięciem Kandydata, choć nie ma ono typowo naukowego charakteru.

Zasadnicze osiągnięcia Kandydata zawarte są w rozdziale 4. (w szczególności w części 4.3) oraz 5. rozprawy. W podrozdziale 4.3 opartym na współautorskiej publikacji Doktoranta zaprezentowano wyniki przeprowadzonej ankiety dotyczącej różnych opisów zagrożenia dotyczącego wykorzystania taj samej luki. Ankieta ta została przeprowadzana przez Internet wśród 20 specjalistów z zakresu cyberbezpieczeństwa. Wyniki zostały jednak przedstawione wyłącznie w postaci wartości średnich – szkoda, że zabrakło informacji o rozrzucie wartości ocen (w postaci wariancji lub przynajmniej wartości minimalnych oraz maksymalnych ocen dla poszczególnych opisów). Kandydat zaproponował użycie połączenia skanera sieciowego z bazą wiedzy dostarczoną przez ChatGPT, weryfikując przydatność tego narzędzia, także w kontekście tzw. halucynacji związanych z udzielaniem nieprawidłowych odpowiedzi zawierających nieścisłości a nawet przekłamania. Wyjaśnienia wymagałoby jednak użycie tzw. pre-promptów, ponieważ w pracy zabrakło szczegółowych informacji, w jaki sposób były używane i jak wpłynęły na poprawę jakości wyników uzyskiwanych z systemu opartego na generatywnej sztucznej inteligencji. Nie jest również jasne, z jakiego względu spośród 860 użytecznych plików uzyskanych za pomocą narzędzia Nmap, w dalszych analizach użyto zaledwie 200 losowo wybranych plików, jak podano na stronie 100. W szczególności brak informacji, czy przy każdym uruchomieniu skryptu użyto tych samych 200 plików, czy były one wybierane losowo za każdym razem.

Wyjaśnienia wymaga również wykres przedstawiony na rysunku 17, ponieważ nie jest jasne, w jaki sposób określany był stopień podobieństwa uzyskiwanych wyników – czy była to ocena ekspercka, czy może został użyty jakiś wskaźnik podobieństwa ciągów tekstowych. Nie jest również jasne, dla jakiego zbioru danych uzyskano wyniki przedstawione pod tabelą na str. 107 (podane wartości z dokładnością do setnych części procenta sugerują, iż był on dość duży).

Wyniki przedstawione w rozdziale 5. dotyczące porównania czasu skanowania przez dwa różne narzędzia (GSA oraz Nmap) pokazują znaczne różnice pomiędzy nimi, jednak porównanie takie może być dość mylące ze względu na istotne różnice dotyczące ich działania. Tym niemniej jest to wątek poboczny rozprawy a przedstawione wnioski dotyczące celowości stosowania narzędzia Nmap nie budzą wątpliwości. Ciekawą oryginalną koncepcją zaproponowaną przez Doktoranta jest użycie informacji uzyskanych na podstawie raportów skanowania do zautomatyzowanej wstępnej selekcji potencjalnie podatnych elementów sieciowych.

Do oryginalnych osiągnięć przedstawionych w rozprawie można także zaliczyć utworzenie bazy danych raportów zawierających oznaczenia podatności dokonane przez doświadczonego eksperta ds. cyberbezpieczeństwa. Dane te zostały następnie użycie do trenowania modelu SI przy dość typowym podziale na 80% próbek dla zbioru uczącego oraz 20% dla zbioru testowego. W rozprawie zabrakło jednak informacji dotyczących walidacji krzyżowej modelu – nie jest jasne, czy eksperymenty były powtarzane i ewentualnie ile razy, a także czy dokonywane było ponowne losowanie próbek stanowiących zbiór uczący. Ciekawe byłoby również porównanie wyników uzyskiwanych dla innych podziałów np. 70%/30%. Tym niemniej testy uzyskanego modelu dla innego zbioru danych pozwoliły osiągnąć wartość wskaźnika F1-score na poziomie 0,872, co uznać należy za dobry wynik, potwierdzający słuszność użycia zaproponowanego podejścia opartego na połączeniu narzędzia Nmap oraz metod sztucznej inteligencji do identyfikacji potencjalnie podatnych zasobów sieciowych. Innym istotnym osiągnięciem jest zaproponowanie podobnego modelu służącego do analizy wzorców wykorzystania luk w zabezpieczeniach VEPA (*Vulnerability Exploit Pattern Analyzer*), przedstawionego w jednej ze współautorskich publikacji konferencyjnych. Doktorant wykazał, iż generowanie pakietów o różnych wzorcach dla danego exploita, a następnie trenowanie modelu SI przy ich użyciu może tworzyć rozwiązania bazujące na sztucznej inteligencji do wykrywania exploitów lub ich wariantów, ukierunkowanych na luki w zabezpieczeniach rzeczywistych sieci. Uzyskana dokładność wykrywania pakietów zawierających warianty kodu exploita na poziomie powyżej 99% stanowi przekonujące potwierdzenie słuszności przyjętych założeń, choć wyjaśnienia wymaga sposób zapobiegania zjawisku przeuczenia sieci (*overfitting*).

IV. UWAGI SZCZEGÓŁOWE

Praca napisana jest na ogół poprawnym i zrozumiałym językiem, choć zdarzają się pomyłki językowe, czy też typograficzne. Przykładowo w tabelach na str. 55–58 zamiast „*effected area*" powinno być sformułowanie „*affected area*", na str. 98 zamiast „*To simply this step*" powinno być „*To simpify this step*"), w sformułowaniu „*It is possible to use if to fully automate the process*" powinno wystąpić słowo „*it*" zamiast „*if*", podobnie jak we frazie „*Research proved that is is possible*" (str. 92) zamiast „*is*". Niektóre frazy, jak np. „*Presented in Figure 8 is the architectural difference*" na str. 41, mogłyby być przeredagowane dla poprawy czytelności tekstu. Innego rodzaju zauważonymi

usterkami jest używanie wielkich liter przy kontynuacji zdań po wzorach („*Where*" str. 64). Sposób numeracji samych wzorów z zastosowaniem podpisów, zbliżony do numeracji tabel czy rysunków, również nie jest typowy dla rozpraw doktorskich w naukach inżynieryjno-technicznych związanych z automatyką, czy też informatyką. Zdarzają się także równoważniki zdań, np. „*P(B) probability of event B.*" (str. 64), czy też rozpoczęcie zdania od słowa „*whereas*" (str. 104). Sformułowanie „*It also significantly smaller amount of traffic in the network*" (str. 110) również nie jest poprawne gramatycznie i powinno zostać poprawione. Tym niemniej, tego rodzaju raczej drobne potknięcia, nieuniknione w większości rozpraw doktorskich, zwłaszcza nie pisanych w języku ojczystym Doktoranta, nie utrudniają w żaden sposób lektury pracy i nie wpływają na ogólną pozytywną opinię o rozprawie.

Pozytywnym elementem pracy jest umieszczenie w niej wykazu skrótów i oznaczeń, co jest szczególnie przydatne ze względu na tematykę rozprawy, dla której cechą charakterystyczną jest dość częste ich stosowanie.

V. WNIOSKI KOŃCOWE

W rozprawie przedstawione zostało oryginalne rozwiązanie problemu o charakterze naukowym, choć przedstawione koncepcje mają w znacznym stopniu charakter wdrożeniowy. Jest to jednak cecha typowa dla doktoratów wdrożeniowych. Rozprawa ma charakter interdyscyplinarny – chociaż równie dobrze mogłaby być złożona w aktualnie obowiązującej dyscyplinie naukowej *informatyka techniczna i telekomunikacja,* to ze względu na obszar zastosowań związany z rozwojem technologii Internetu Rzeczy oraz Przemysłu 4.0, jej tematyka mieści się w dyscyplinie *automatyka, elektronika, elektrotechnika i technologie kosmiczne.*

Pomimo uwag krytycznych oraz dyskusyjnych zawartych w recenzji, stwierdzam, iż przedstawiona do recenzji rozprawa doktorska **mgr. inż. Wiktora Sędkowskiego** pt. ***Threat Modeling methods and detecting vulnerabilities in 5G networks*** (tytuł w języku polskim: *Metody modelowania zagrożeń i wykrywania podatności w sieciach 5G*), której promotorem jest dr hab. inż. Rafał Stanisławski, prof. Politechniki Opolskiej, spełnia wymagania stawiane rozprawom doktorskim przez *Ustawę Prawo o szkolnictwie wyższym i nauce* z dnia 20 lipca 2018 roku (tekst jednolity Dz. U. 2023 poz. 742, z późn. zm.). **Wnioskuję o jej przyjęcie oraz dopuszczenie do publicznej obrony.**