Bydgoszcz, 04.05.2022

Prof. dr hab. inż. Michał Choraś

Wydział Telekomunikacji, Informatyki i Elektrotechniki
Politechnika Bydgoska im. J.J. Śniadeckich, Bydgoszcz

Recenzja rozprawy doktorskiej

Rozproszony algorytm analizy cyberbezpieczeństwa w sieci korporacyjnej,

której Autorem jest Pan

mgr inż. Michał Walkowski

realizowanej na Wydziale Informatyki i Telekomunikacji PWr

1. Wprowadzenie.

Niniejsza recenzja rozprawy doktorskiej, której Autorem jest Pan mgr inż. Michał Walkowski, została wykonana na zlecenie Rady Dyscypliny Informatyka Techniczna i Telekomunikacja Politechniki Wrocławskiej (Uchwała nr RDN/ITT/38/2022 z dnia 16 lutego 2022 r. oraz na podstawie zawiadomienia o wyznaczeniu na Recenzenta w postępowaniu o nadanie stopnia doktora podpisanego przez Przewodniczącego RDN ITT Politechniki Wrocławskiej Pana Profesora Michała Woźniaka, z dnia 18 lutego 2022 r.)

Rozprawę odebrałem w marcu 2022 r., a recenzję wysłałem w wyznaczonym terminie w maju 2022 r.

Promotorem niniejszej rozprawy jest Pan Profesor dr hab. inż. Sławomir Sujecki. Promotorem pomocniczym w niniejszym przewodzie jest dr inż. Jacek Oko.

Praca doktorska składa się z sześciu rozdziałów, bibliografii, a także pięciu załączników. Załączniki stanowią obszerną część rozprawy przedstawiając dokumentację projektową oprogramowania wytworzonego w tamach rozprawy oraz wdrożonego w ramach projektu RegSoc. Niektóre załączniki zawierają także oddzielne bibliografie oraz własne oddzielnie numerowane rysunki.

Niniejsza recenzja (poza wprowadzeniem i wnioskiem) zawiera odpowiedzi na siedem pytań dotyczących rozprawy doktorskiej.

2. **Jaki jest problem naukowy (teza) rozprawy? Czy został on trafnie i jasno sformułowany? Jaki charakter ma rozprawa?**

Rozprawa, której Autorem jest Pan mgr inż. Michał Walkowski, dotyczy cyberbezpieczeństwa. W szczególności, Autor zajął się problemem zarządzania podatnościami oraz ich priorytetyzacją w celu poprawy bezpieczeństwa w sieciach i systemach teleinformatycznych. Autor dostarczył także mechanizmy wizualizacji dla ekspertów ds. bezpieczeństwa pracujących np. w lokalnych SOC (*j.ang. Security Operations Centre*).

Autor zaproponował własną metodykę i narzędzia do analizy, zarządzania i priorytetyzacji podatności wykrytych w środowiskach teleinformatycznych. Autor zaprojektował i zaimplementował także platformę nazwaną VMC (*j.ang. Vulnerability Management Center*), która jest informatycznym narzędziem będącym rezultatem niniejszej rozprawy.

Niniejsza praca naukowa ma charakter koncepcyjno-eksperymentalny i wdrożeniowy. Warto podkreślić, iż rezultaty rozprawy oraz wytworzone oprogramowanie zostało wdrożone i przetestowane przez Wrocławskie Centrum Sieciowo-Superkomputerowe Politechniki Wrocławskiej w ramach projektu RegSoc.

Problemy naukowe rozprawy zostały jasno i trafnie sformułowane, a także rozwiązane przez Autora. Dość długa, ale czytelna i dobrze sformułowana teza znajduje się pod koniec Wstępu (Rozdział 1) pracy na stronie 28. Teza została potwierdzona przez Autora pracy w dalszych częściach rozprawy.

3. **Czy w rozprawie przeprowadzono w sposób właściwy analizę źródeł, w tym literatury światowej, stanu wiedzy i zastosowań w przemyśle? Czy wnioski z przeglądu źródeł sformułowano w sposób jasny i przekonywujący?**

W pracy brakuje typowego oddzielnego rozdziału lub wyraźnej części poświęconej innym naukowym pracom w podobnej tematyce. Autor przedstawił zarys problemu oraz wykorzystywane standardy (np. CVSS 2.0, CVSS 3.x), a także modele w Rozdziale 1, czyli we wstępie rozprawy.

Wyraźnie widać, że Autor jest praktykiem i w pracy skupił się na wykorzystywanych rozwiązaniach i standardach, bez szerokiej analizy innych naukowych prac w tym zakresie.

Autor w sposób ograniczony zdefiniował i przedstawił samo pojęcie podatności (na str. 3 w rozdziale 1.1) jako błędów w oprogramowaniu. Taka definicja jest oczywiście słuszna, ale w rozprawie doktorskiej Autor mógł ukazać podatności w szerszym kontekście. Podatności mogą być także (poza błędami w oprogramowaniu) m.in. słabościami użytkownika systemu, błędną konfiguracją, słabościami technologii, słabościami protokołów lub kanałów telekomunikacyjnych, lub błędami w projekcie.

Autor mógłby w sposób szerszy opisać stan wiedzy dot. bezpieczeństwa sieci komputerowych, zagrożeń i podatności, sposobów ich wykrywania, zapobiegania, a także możliwych reakcji i remediacji.

Zabrakło także szerszego spojrzenia na problem cyberbezpieczeństwa, a także zarysowania pojęć i umiejscowienia pojęcia podatności w szerszej perspektywie. Mam na myśli ukazanie podatności w relacji do pojęć takich jak zagrożenie (*j.ang. threat*), atak (*j.ang. cyber attack*), włamanie (*j.ang. intrusion*), itp. Innymi słowy, zabrakło omówienia znanych ontologii lub standardów dotyczących tych pojęć. Autor pominął także inne standardy (np. CVE, CWE) itp., a także istotne zagadnienie dotyczące ujawniania podatności (*j.ang. vulnerability disclosure*).

Tym niemniej, Autor jest bez wątpienia ekspertem w dziedzinie bezpieczeństwa, a w szczególności w analizie i zarządzaniu podatnościami. Autor celowo skoncentrował się na omówieniu praktycznie stosowanych przez niego standardów i węższemu spojrzeniu na podatności. Sama bibliografia zawiera odpowiednią liczbę źródeł (101), w tym większość z ostatnich kilku lat, co oceniam pozytywnie.

4. **Czy autor rozwiązał postawione zagadnienia? Czy użył do tego właściwych metod dowodząc, że posiadł umiejętności związane z metodyką i metodologią prowadzenia badań naukowych? Czy przyjęte założenia są uzasadnione?**

Generalnie, Autor w sposób odpowiedni rozwiązał problemy, których dotyczy rozprawa. Nie mam wątpliwości, iż Autor posiada dużą wiedzę dot. zagadnień związanych z zarządzaniem podatnościami, standardami priorytetyzacji podatności oraz z cyberbezpieczeństwem, w jego praktycznym ujęciu.

Przyjęte założenia są uzasadnione i merytorycznie poprawne. Teza rozprawy została dowiedziona.

Autor posiada duże umiejętności w projektowaniu i implementacji platformy dla wizualizacji i zarządzania podatnościami. Autor wykorzystał w sposób właściwy wiele znanych i typowych obecnie technologii informatycznych, m.in. dokeryzację, Elasticsearch, Kibana, Nessus itp.

Autorskim i głównym elementem rozprawy jest projekt oraz implementacja platformy VMC, a także: wykorzystania metody optymalizacji obliczeń dla dużej ilości danych, metoda określania potencjalnych szkód dodatkowych CDP (*j.ang. Collateral Damage Potential*), metodę określania dystrybucji podatności, sposób konwersji ze standardu CVSS 2.0 do standardu CVSS 3.x z wykorzystaniem mechanizmów uczenia maszynowego.

Ponadto, Autor dokonał głębokiej analizy i eksperymentów dla trzech teleinformatycznych środowisk testowych nazwanych A,B,C opisanych w Rozdziale 3. Wpływ parametrów oraz głęboka analiza znajdują się w Rozdziałach 4 i 5.

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

Autor wykorzystał i opisał w Rozprawie szereg nowych rozwiązań, w tym projekt oraz implementację platformy VMC, metodę optymalizacji obliczeń dla dużej ilości danych, metodę określania potencjalnych szkód dodatkowych CDP (*j.ang. Collateral Damage Potential*), metodę określania dystrybucji podatności, sposób konwersji ze standardu CVSS 2.0 do standardu CVSS 3.x z wykorzystaniem mechanizmów uczenia maszynowego.

Bardzo ważnym aspektem rozprawy jest jej wartość praktyczna: rezultaty pracy zostały wdrożone i przetestowane w projekcie RegSoc we Wrocławskim Centrum Sieciowo-Superkomputerowym Politechniki Wrocławskiej.

Ponadto, co oceniam bardzo pozytywnie, Autor umieścił własne kody źródłowe w otwartych repozytoriach (github.com).

Ponadto, rezultaty Autora rozprawy zostały zweryfikowane i zaakceptowane do fazy inkubator przez organizację OWASP, co świadczy o ich jakości oraz szerokich możliwościach potencjalnego wykorzystania.

6. **Czy autor wykazał umiejętność poprawnego i przekonywującego przedstawienia uzyskanych przez siebie wyników? Jaka jest poprawność redakcyjna rozprawy?**

Niniejsza rozprawa stanowi przykład profesjonalnie przygotowanej pracy doktorskiej. Praca napisana jest na wysokim poziomie edycyjnym oraz graficznym.

W pracy występują oczywiście drobne usterki, literówki, ale jest ich niewiele i nie są znaczące, m.in.:

- Literówki i błędy na Rysunkach, np. „modułu pobierające dane" na Rys. 2.6
- Niespójny język (polski czy angielski) na Rys. 2.3
- „mulit-klientów' na str. 40
- W słowie angielskim na str. 47 powinno być 'collateral', a nie 'colleteral'
- Nie jest jasne w jakim standardzie/notacji stworzono diagramy, czy to na pewno standardowy BPMN?
- Na Rysunkach 1.2-1.5 brak konsekwencji dot. dużej/małej litery w słowie „zasobu"
- W chwili pisania recenzji, na widoku projektu na stronie OWASP widoczna była literówka w słowie 'daily'

Te drobne usterki nie zmieniają ogólnej opinii o dobrym poziomie językowym i edycyjnym rozprawy.

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

Rolą recenzenta jest zauważenie ewentualnych niedociągnięć i mankamentów przedstawianej pracy, oraz zgłoszenie uwag, które mogą być pomocne i przydatne w dalszych pracach.

Uwagi krytyczne to między innymi:

- Brak typowego oddzielnego rozdziału lub wyraźnej części poświęconej innym naukowym pracom w podobnej tematyce, o czym wspomniałem w punkcie 3 niniejszej recenzji.
- Brak krytycznej dyskusji na temat zalet i wad wybranych technologii oraz uzasadnienia tych wyborów, w tym np. dlaczego Nessus (czy nie ma innych skanerów), dlaczego Kibana (czy nie ma innych sposobów wizualizacji, nawet wykorzystujących Elasticsearch), itp. itd. Tego typu pytania można zadać dla każdej z wybranych technologii.
- Autor bez uzasadnienia wybrał do szerszego opisu część technologii, a część zupełnie pominął. Nie jest dla mnie jasne dlaczego opisano typowe, znane i mało naukowe rozwiązania jak dokeryzacja czy technologię Kubernetes. Nie opisano szerzej działania skanera Nessus, ani technologii Elasticsearch oraz kilku innych.
- Uboga analiza zagadnienia cyberbezpieczeństwa, a także wykorzystywanych pojęć. Autor w sposób ograniczony zdefiniował i przedstawił samo pojęcie podatności (na str. 3 w rozdziale 1.1) jako błędów w oprogramowaniu. Taka definicja jest oczywiście słuszna, ale w rozprawie doktorskiej Autor mógł ukazać podatności w szerszym kontekście. Podatności mogą być także (poza błędami w oprogramowaniu) m.in. słabościami użytkownika systemu, błędną konfiguracją, słabościami technologii, słabościami protokołów lub kanałów telekomunikacyjnych, lub błędami w projekcie.
- Brak odniesienia podatności do konkretnych zagrożeń, ataków i skutecznych włamań, także w odniesieniu do wybranych testowych środowisk teleinformatycznych A, B, C.
- Brak porównania z innymi rozwiązaniami, lub porównania z sytuacją, gdy platforma VMC nie jest wykorzystywana. Innymi słowy zabrakło ew. konkretnych zwymiarowanych korzyści wynikających z zastosowania VMC. Zabrakło także opinii użytkowników końcowych.
- Brak dyskusji na temat ewentualnych błędów dot. skanowania zasobów sieci. Autor pomija fakt, iż pewne zasoby (topologia, użytkownicy, usługi itp.) zmieniają się bardzo często, a skanery nie są przecież skuteczne w 100%.
- Pewien niedosyt pozostawia część dot. uczenia maszynowego – Autor bardzo skąpo potraktował tę część rozwiązania.

- Brak dyskusji dot. ujawniania podatności, czyli m.in. CVD (j. ang. *Coordinated Vulnerability Disclosure*), zarówno na poziomie małych organizacji, a także na poziomie krajowym w EU.
- Priorytetyzacja podatności jest rzeczywiście bardzo ważnym zagadnieniem. Tym niemniej, Autor mógł szerzej przedyskutować aspekty związane z analizą ich konsekwencji (w tym biznesowymi). Czy podatność klasyfikowana jako „niska" może mieć większe negatywne konsekwencje, niż podatność klasyfikowana jako „wysoka"?
- Zabrakło także odniesienia/dyskusji dotyczącej interoperacyjności przedstawionego zagadnienia z innymi mechanizmami i narzędziami cyberbezpieczeństwa, w tym np. informacjami o reakcjach, rozwiązaniach klasy MISP, itp.
- Zabrakło także pełnej informacji na temat licencji wykorzystywanych technologii oraz ich implikacji dla rozwoju platformy VMC.

Tym niemniej, praca jest bardzo wartościowa i ma **szereg mocnych stron**, a wymienione uwagi krytyczne nie wpływają na pozytywną ocenę niniejszej rozprawy.

Autor zajął się bardzo ważnym i aktualnym zagadnieniem, a przedstawione i zaimplementowane przez Autora propozycje rozwiązań dotyczących zarządzania i priorytetyzacji podatności, a także platforma VMC, zostały odpowiednio przetestowane oraz wdrożone przez Wrocławskie Centrum Sieciowo-Superkomputerowe Politechniki Wrocławskiej w ramach projektu RegSoc.

Warto nadmienić, iż poza udostępnieniem kodów na platformie GitHub, Doktorant jest także współautorem kilku artykułów naukowych wymienionych na str. iv oraz v niniejszej rozprawy.

8. Jaka jest przydatność rozprawy dla nauk technicznych?

Praca dotyczy bardzo aktualnych i potrzebnych zagadnień nowoczesnej informatyki technicznej i telekomunikacji, a w szczególności istotnego problemu dla społeczeństw, czyli cyberbezpieczeństwa.

Opracowane przez Autora mechanizmy mogą znaleźć zastosowanie i być rozwijane w tzw. SOC i innych jednostkach dbających o analizę bezpieczeństwa sieci.

Zaproponowane i wdrożone przez Autora mechanizmy zarządzania podatnościami mogą pozwolić na szybszą analizę, a tym samym na konkretne oszczędności w tzw. roboczogodzinach pracy ekspertów od cyberbezpieczeństwa.

Przedstawione w rozprawie rozwiązania, a w szczególności platforma VMC, mają duży potencjał wykorzystania oraz wpływu na otoczenie i społeczeństwo w zakresie bezpieczeństwa.

9. Wniosek

Biorąc pod uwagę przedstawioną przez Doktoranta rozprawę stwierdzam, że recenzowana praca **spełnia wymagania stawiane rozprawom doktorskim** przez obowiązujące przepisy. Dlatego wnoszę o przyjęcie niniejszej rozprawy i **dopuszczenie** mgr inż. Michała Walkowskiego do publicznej obrony.