dr hab. inż. Janusz Stokłosa
prof. Wyższej Szkoły Bankowej w Poznaniu

Poznań, 1 grudnia 2020 r.

Recenzja rozprawy doktorskiej mgr inż. Natalii Krzyworzeki

pt. „**Percepcyjne techniki zabezpieczania danych i weryfikacji użytkowników**"

wykonana dla

Rady Dyscypliny Informatyki Technicznej i Telekomunikacji
Akademii Górniczo-Hutniczej im. Stanisława Staszica w Krakowie

0. Recenzję rozprawy wykonałem na prośbę Przewodniczącego Rady, skierowaną do mnie w piśmie nr RD.ITiT.WEAIiIB.510-1/18/517/2020 z dn. 19.10.2020 r.

1. Przedłożona mi do recenzji rozprawa doktorska mgr inż. Natalii Krzyworzeki dotyczy technik percepcyjnych w postaci modeli CAPTCHA służących do weryfikacji użytkowników systemów informatycznych. CAPTCHA (***C**ompletely **A**utomated **P**ublic Turing test to tell **C**omputers and **H**umans **A**part*) jest prostym protokołem typu wyzwanie-odpowiedź, służącym do stwierdzenia czy użytkownik jest człowiekiem, a nie programem komputerowym (botem) imitującym aktywność człowieka, próbującym włamać się na konto w sieci WWW chronione hasłem. Wyzwaniem jest prośba o rozwiązanie relatywnie prostego dla człowieka (a niekoniecznie dla maszyny) testu, a odpowiedzią podanie rozwiązania.

Autorka sformułowała tezę badawczą w następującej postaci: „Możliwe jest opracowanie nowych algorytmów wykorzystujących wizualne kody CAPTCHA, a także technik percepcyjnych do zabezpieczania danych i weryfikacji użytkowników". Jako cel badawczy przyjęła opracowanie nowych algorytmów weryfikacji użytkowników i zabezpieczania danych.

2. Rozprawa zawiera streszczenia w języku polskim i angielskim, wstęp, pięć głównych rozdziałów, wykaz bibliografii (63 pozycje) i spis rysunków – razem 102 strony. Rozdziałem 1 jest wstęp do rozprawy, w którym zarysowano problematykę badawczą, sformułowano tezę rozprawy, określono cel i zakres pracy oraz przedstawiono jej układ.

Rozdział 2 poświęcono omówieniu podstaw technik kognitywnych oraz przeglądowi literatury przedmiotu. Punktem wyjścia do dalszych rozważań jest stwierdzenie, że kognitywne techniki zabezpieczeń „wykorzystują przewagę człowieka nad maszyną w sposobie przetwarzania określonych typów danych", polegającą „nie tylko na niedoskonałości algorytmów detekcyjnych, bądź braku odpowiedniej bazy danych do ich analizy, ale przede wszystkim na

złożoności interpretacji i procesu myślowego ludzkiego umysłu". Zwrócono uwagę na fakt, że CAPTCHA oraz inne metody percepcyjne powinny być łatwe do zrozumienia i zapamiętania przez człowieka oraz szybko wykonywalne.

W dalszym ciągu tego rozdziału Autorka omawia modele CAPTCHA:

- lingwistyczne – oparte o rozumienia znaczenia słów i kontekstu zdania;
- tekstowe – wykorzystujące zdolność człowieka do identyfikacji zniekształconych liter bądź cyfr;
- obrazowe – polegające na rozpoznaniu i wskazaniu obiektu, odgadnięciu kontekstu obrazu, umieszczenie obiektu w odpowiednim położeniu, zidentyfikowaniu niepasującego elementu, bądź zaznaczeniu na obrazku wymienionych w teście obiektów;
- wideo – wymagające udzielenia poprawnej odpowiedzi po obejrzeniu krótkiego filmu lub animacji;
- 3D, u podstaw której leży brak zdolności programów komputerowych do rozpoznawania trzeciego wymiaru na podstawie obrazu 2D, np. przepisanie znaków widocznych w postaci trójwymiarowej;
- spersonalizowana – służąca do weryfikacji określonej grupy ludzi lub konkretnych osób;
- wymagająca wiedzy eksperckiej – do użytkownika kierowane jest zapytanie z określonej dziedziny wiedzy, np. geologii.

Autorka zwraca także uwagę na inne techniki percepcyjne – w postaci kryptografii wizualnej i metod biometrycznych.

Ochrona danych za pomocą technik percepcyjnych jest przedmiotem rozważań zawartych w rozdziale 3 rozprawy. Autorka omawia wymagania dotyczące bezpieczeństwa i użyteczności technik percepcyjnych, podaje przykłady stosowania technik percepcyjnych do ochrony systemów informatycznych, zwraca uwagę na kwestie bezpieczeństwa technik percepcyjnych. Bezpieczeństwo protokołu CAPTCHA zależy do algorytmów detekcyjnych, na działanie których jest odporny. W procesie łamania protokołu przez algorytmy komputerowe wyróżnia się kilka kroków: przetwarzanie wstępne, segmentacja, post-segmentacja, rozpoznawanie i klasyfikacja oraz przetwarzanie końcowe i analiza – wszystkie omówione są w rozprawie wystarczająco szczegółowo. Kilka końcowych stronach tego rozdziału Autorka poświęca kierunkom rozwoju i przyszłości CAPTCHA.

W rozdziale 4 mgr inż. Natalia Krzyworzeka przedstawiła własne propozycje wizualnych technik uwierzytelniania użytkowników – modeli CAPTCHA korzystających z technik percepcyjnych. Mają one – w zamierzeniu Autorki – rozwiązać problem pamiętania przez użytkowników własnego hasła dostępu do systemu. Pierwsza propozycja to CAPTCHA Password Reminder, oparta na dwóch przesłankach: (i) mózg ludzki radzi sobie lepiej z zapamiętywaniem obrazów i skojarzeń niż słów oraz (ii) istnieją silne bezpośrednie relacje pomiędzy ośrodkiem mowy a korą wzrokową człowieka. Zaproponowane rozwiązanie ma pomóc w utworzeniu w pamięci użytkownika asocjacji obraz-hasło. Propozycja ta wymaga na etapie rejestracji konta zapisania dowolnego obrazu i udzieleniu odpowiedzi na jedno z pytań służących do odzyskiwania konta. W oparciu o obraz i pytanie użytkownik przypomina sobie hasło.

Drugą autorską propozycją jest CAPTCHA personalna wykorzystująca dane lokalizacyjne. W tym przypadku włączono dodatkowy element wiedzy podnoszący poziom trudności, a jednocześnie gwarantujący sukces wyłącznie określonemu użytkownikowi. Zaproponowane przykładowe rozwiązanie korzysta z aplikacji Google Street View. Na etapie rejestracji konta użytkownik typuje odpowiednie miejsca na mapie (wybiera ulice). Algorytm generujący CAPTCHA personalną usuwa nazwę ulicy automatycznie umieszczanej na obrazie w Street View. W tym celu (i) wyznacza obszar, na którym znajduje się nazwa, (ii) odfiltrowuje najjaśniejsze piksele z obszaru zawierającego nazwę, (iii) dla tak przygotowanego wstępnie obrazu tworzy negatyw poddawany następnie dylatacji za pomocą autorskiego algorytmu, (iv) wstawia otrzymaną strukturę do oryginalnego zdjęcia ulicy. W ekranie logowania z użyciem CAPTCHA użytkownik jest proszony o padanie nazwy ulicy. Zaproponowane rozwiązanie CAPTCHA personalnej nie musi ograniczać się do korzystania z aplikacji Google Street View i podawania nazwy ulicy; może odnosić się do innych miejsc prosząc użytkownika o padanie ich nazwy bądź numeru.

Trzecią propozycją autorską przedstawioną w rozprawie jest CAPTCHA tekstowa, tworzona w oparciu o równoczesne łączenie kilku różnych metod przetwarzania obrazów. Algorytm działa następująco: (i) losowane są litery i cyfry oraz kolejność, w jakiej zostaną przedstawione na obrazie, (ii) każdej literze i cyfrze przydzielany jest pseudolosowo algorytm modyfikujący (zniekształcający) ją, (iii) zmienione symbole wstawiane są do CAPTCHA w odległości pseudolosowej względem poprzedniego symbolu, (iv) pseudolosowo generowane jest tło, (v) sprawdzane jest podobieństwo pikseli znaków do graniczących z nimi pikseli tła i wykonywana jest ewentualna korekta. W algorytmie użyto kilkanaście różnych stylów czcionek. Użyte litery i cyfry zostają poddane przekształceniom: ustawiane są pod odpowiednim kątem, zmianie ulegają proporcje znaku, tekstura zostaje zaszumiona, kontur pogrubiony, stosowane są różne skale szarości, wprowadzany efekt cienia, ekstrakcji konturu, erozji szkieletu.

Opisane algorytmy były tworzone z myślą o spełnieniu następujących wymagań: (i) pełnej automatyzacji, (ii) szybkiej weryfikacji, (iii) łatwej obsługi, (iv) skuteczności identyfikacji, (v) bezpieczeństwa. Co do ostatniej kwestii, to bezpieczeństwo algorytmu nie wynika ze stopnia skomplikowania zapytań CAPTCHA, lecz z przewagi człowieka nad komputerem w rozwiązywaniu wybranych problemów kognitywnych. Bezpieczeństwo CAPTCHA byłoby zagrożone wówczas, gdyby udowodniono, że odpowiedni problem z zakresu sztucznej inteligencji daje się rozwiązać efektywnie (w sensie złożoności obliczeniowej algorytmów).

Wszystkie autorskie algorytmy prezentowane w rozprawie oraz opracowany przez Autorkę protokół uwierzytelniania użytkowników zostały uruchomione w środowisku Matlab v. 2010a.

Opracowana przez Autorkę koncepcja percepcyjnego protokołu uwierzytelniania użytkowników (przedstawionego w rozdz. 4.4 jako „percepcyjny protokół autentykacji") została opracowana w oparciu o CAPTCHA personalną i została pomyślana jako rozwiązanie do stosowania przez użytkownika wyposażonego w urządzenie mobilne z aparatem fotograficz-

nym. Działanie protokołu polega na podaniu przez użytkownika hasła w postaci obrazu (sfotografowanego na bieżąco obiektu bądź sceny) i porównaniu go z hasłem kontrolnym (kluczem), także w postaci obrazu, przechowywanym w bazie danych, która powinna zawierać odpowiednio duży zbiór lokalnych charakterystyk fotografowanego obiektu. Wprowadzony obraz nie musi być całkowicie zgodny z hasłem kontrolnym, jednak aby logowanie zakończyło się sukcesem, to wyznaczony dla niego stopień podobieństwa musi być odpowiednio wysoki. We wstępnym kroku działania protokołu użytkownik jest proszony o umieszczenie w bazie danych kilku zdjęć dowolnego rekwizytu (bądź rekwizytów), który będzie później używany jako hasło. W procesie logowania używany jest nowy obraz, który zostaje poddany normalizacji, tworzonych jest kilka jego kopii w różnych rozmiarach, a następnie każda z nich zostaje poddana analizie pod kątem podobieństwa (dopasowania wartości pikseli) z obrazami z bazy danych. Jeśli założony procent pikseli jest niezgodnych z kluczem, to uwierzytelnianie kończy się niepowodzeniem.

Jak wspomniano powyżej protokół został zakodowany w środowisku Matlab v. 2010a. Autorka dokonała analizy porównawczej działania protokołu dla kilku przypadków: dwóch identycznych obrazów (klucz-hasło), dwóch podobnych obrazów z tym samym rekwizytem, dwóch różnych obrazów przedstawiających ten sam rekwizyt, dwóch różnych obrazów zawierających różne rekwizyty. Przedstawione rozwiązanie jest prototypem, który umożliwił analizę zaproponowanego rozwiązania protokołu uwierzytelniania użytkowników systemu informatycznego.

W rozdziale 5 podsumowano uzyskane w rozprawie wyniki oraz wskazano kierunki dalszych badań. Autorka podkreśla, że zaproponowane przez nią rozwiązania stanowią alternatywę dla metod istniejących. Zwraca uwagę na cechy charakterystyczne związane z łatwością przypominania sobie hasła na podstawie doświadczeń wynikających z posiadanej wiedzy, własnych skojarzeń, wspomnień itp.

3. Rozprawa ma charakter analityczno-eksperymentalny. Obszar badań został wybrany trafnie. Głównym osiągnięciem Autorki rozprawy jest opracowanie trzech algorytmów CAPTCHA oraz propozycji percepcyjnego protokołu uwierzytelniania użytkowników systemów informatycznych, utworzenie prototypu tego protokołu oraz poddanie go analizie krytycznej. Zaproponowane rozwiązania korzystające z technik CAPTCHA mają na celu ominięcie trudności polegającej na zapamiętywaniu przez użytkowników dużej liczby haseł.

Praca napisana jest językiem zrozumiałym i oszczędnym. Pewnym jej mankamentem jest brak definicji, a w konsekwencji nie zawsze poprawne użycie terminów podstawowych: uwierzytelnianie, weryfikacja, identyfikacja, autoryzacja (*nb.* terminu *autentykacja* nie odnotowuje żaden ze znanych mi słowników języka polskiego). W wielu miejscach Autorka używa niepoprawnie rzeczownika *ilość* (zamiast *liczba*). Drobnych usterek redakcyjnych zauważyłem kilka.

4. Reasumując stwierdzam, że:
   - teza rozprawy została wykazana,

- rozprawa stanowi oryginalne rozwiązanie problemu naukowego,
- tematyka rozprawa jest aktualna i ważna,
- Autorka rozwiązała zdefiniowany przez siebie problem naukowy i użyła do tego celu odpowiednich metod, tak więc wykazała się umiejętnością samodzielnego prowadzenia badań naukowych,
- rozprawa świadczy o dużej wiedzy teoretycznej mgr inż. Natalii Krzyworzeki w zakresie informatyki technicznej, w szczególności w obszarze technik ochrony danych.

Przedstawiona mi do recenzji dysertacja doktorska, mieszcząca się w dyscyplinie naukowej informatyka techniczna i telekomunikacja, spełnia wymagania stawiane rozprawom doktorskim w obowiązujących przepisach prawa o szkolnictwie wyższym i nauce.

Wnoszę o dopuszczenie mgr inż. Natalii Krzyworzeki do publicznej obrony rozprawy.