Vilniaus universitetas

LIETUVOS TEISMO EKSPERTIZĖS CENTRAS

Ataskaitinė informatikos krypties doktorantų konferencija 2022-03-24

Andrius Chaževskas (VU DMSTI doktorantas, Išmaniųjų technologijų tyrimų grupė)

Darbo tema.

Teksto semantinės analizės ir mašininio mokymosi algoritmų taikymo slaptažodžių parinkimui tyrimas.

Application of text semantic analysis and machine learning algorithms for passwords guessing.

Darbo vadovas.

Prof. dr. Igoris Belovas.

Doktorantūros studijų laikotarpis.

2020 m. spalio mėn. 1 d. – 2024 m. rugsėjo mėn. 30 d..

Ataskaitinis laikotarpis.

2021 m. rugsėjo mėn. 30 d. – 2022 m. kovo mėn. 24 d..

Visų studijų planas ir jo vykdymo suvestinė

Studijų metai	Egzaminai		Dalyvavimas konferencijose		Publikacijos		
	Planas	Įvykdyta	Planas	Įvykdyta	Planas	Įvykdyta	Būklė
I (2020/2021) Pirmas pusmetis	1	1		1 (L[1])			
I (2020/2021) Antras pusmetis	1	1	1 (L)	1 (T[2])		1 (KD/R[3])	Publikuota
II (2021/2022) Pirmas pusmetis	1	1		1 (L)	1 (KD[4])		
II (2021/2022) Antras pusmetis	1		1 (L)		1 (KD / R)		
III (2022/2023) Pirmas pusmetis							
III (2022/2023) Antras pusmetis			1 (T)		1 (CA WoS[5])		
IV (2023/2024) Pirmas pusmetis							
IV (2023/2024) Antras pusmetis			1 (T)		1 (CA WoS)		

[1] Tarpinių rezultatų pristatymas konferencijoje Lietuvoje.

[2] Tyrimo rezultatų pristatymas tarptautinėje mokslinėje konferencijoje.

[3] Tarpinių rezultatų publikavimas (recenzuojamoje konferencijos darbų medžiagoje).

[4] Mokslinių tyrimų disertacijos tema apžvalga (konferencijos darbų medžiagoje).

[5] Rezultatų publikavimas (recenzuojamame periodiniame leidinyje CA WoS su Impact Factor).

Vilniaus universitetas

Ataskaitinių metų darbo planas ir jo vykdymo suvestinė

Egzaminai	**Dalyvavimas konferencijose**	**Publikacijos**
Planas	**Įvykdyta**	**Būklė**
„Fundamentalieji informatikos ir informatikos inžinerijos mokslų metodai“.	„Fundamentalieji informatikos ir informatikos inžinerijos mokslų metodai“, 2022-01-28.	Išlaikytas.
Dalyvavimas konferencijose		
Planas	**Įvykdyta**	**Konferencijos tipas**
Tarpinių rezultatų pristatymas konferencijoje Lietuvoje.	Dalyvauta dvyliktojoje „Data Analysis Methods for Software Systems Conference” DAMSS konferencijoje, gruodžio 2-4 d., 2021, Druskininkai, Lietuva.	Nacionalinė konferencija.
Publikacijos		
Planas	**Įvykdyta**	**Būklė**	**Publikacijos tipas**
Mokslinių tyrimų disertacijos tema apžvalga (konferencijos darbų medžiagoje).	Chaževskas, Andrius; Belovas, Igoris; Marcinkevičius, Virginijus. Forensic password examination in leaked user databases // Zborník príspevkov 17. medzinárodný kongres kriminalistika a forenzné vedy: veda, vzdelávanie, prax, 16. - 17. September 2021 Bratislava, Slovenská Republika / Štefan Zachar, Jozef Meteńko, Miriam Meteńková (eds.). Bratislava : Akadémia Policajného zboru v Bratislave, 2021. ISBN 9788080549053. p. 241-257.).	Publikuota	Recenzuojamoje konferencijos darbų medžiagoje (be cituojamumo rodiklio).

Vilniaus universitetas

Kvalifikacijos kėlimas

- Darbas su VU Duomenų mokslo ir skaitmeninių technologijų instituto Informacinių sistemų inžinerijos bakalauro programos IV k. studentais (jaunesniojo asistento pareigose). Informacinės saugos kurso laboratorinių darbų kuravimas (2021 ruduo).
- Užduočių tarptautinėms kibernetinio saugumo varžyboms VU Cyberthon 2022 parengimas.
- Darbas su VU Kauno fakulteto III k. studentais, dėstomas kursas Skaitmeninio turinio teisminė analizė (2022 pavasaris).

Vilniaus universitetas

Visų mokslinių tyrimų ir disertacijos rengimo etapai

	Darbo pavadinimas	Atlikimo terminai	Pastabos
1.	Mokslinių tyrimų disertacijos tema apžvalga ir analizė (Lietuvoje ir užsienyje): 1.1. Analitinės apžvalgos atlikimas. 1.2. Disertacijos tyrimo objekto detalizavimas. 1.3. Mokslinių problemų susietų su tyrimo objektu identifikavimas ir tyrimo tikslo suformavimas.	2020 m. spalio mėn. – 2021 m. rugsėjo mėn.	Atlikta, apibendrinti rezultatai mokslinėje ataskaitoje.
	Mokslinio tyrimo vykdymas:		
2.	**2.1.Tyrimo metodikos sudarymas:** 2.1.1. Uždavinių, skirtų tyrimo tikslui pasiekti, suformulavimas. 2.1.2. Tyrimo metodikos išsikeltiems uždaviniams spręsti parinkimas. 2.1.3. Teorinio ir empirinio tyrimų suplanavimas pagal pasirinktą metodiką.	2021 m. spalio mėn. – 2022 m. sausio mėn.	Atlikta, apibendrinti rezultatai mokslinėje ataskaitoje.

Visų mokslinių tyrimų ir disertacijos rengimo etapai

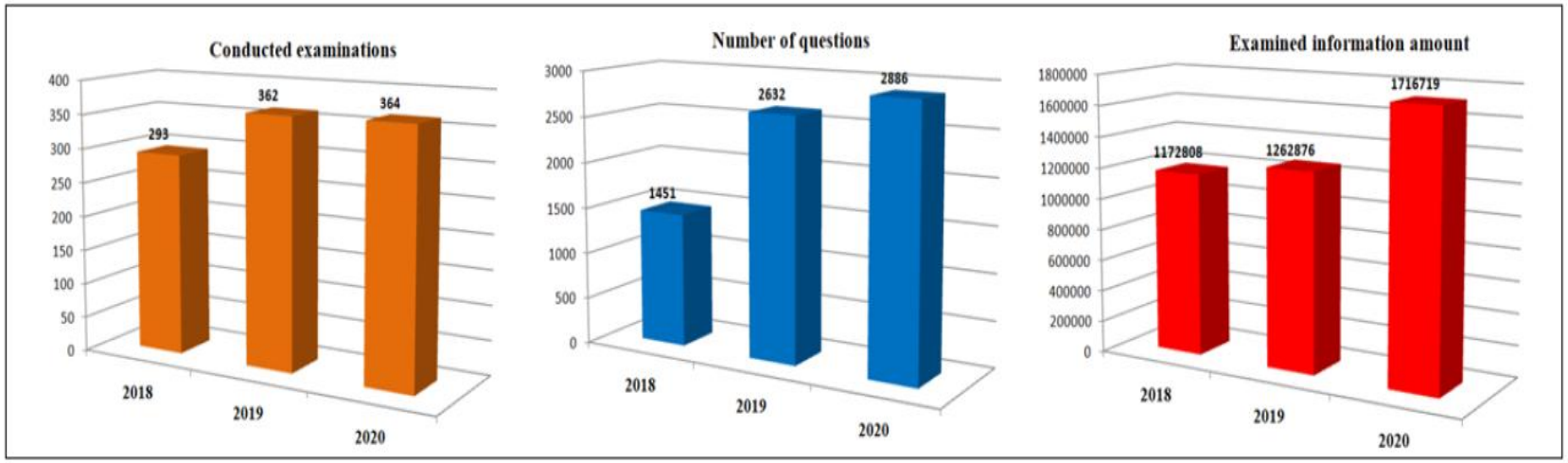
	Darbo pavadinimas	Atlikimo terminai	Pastabos
	2.2. Teorinis tyrimas: 2.2.1. Mašininio mokymosi metodų naudojamų automatizuotame slaptažodžių parinkime tyrimas. 2.2.2. Semantinės slaptažodžių analizės ir šablonų parinkimo metodų tyrimas. 2.2.3.Slaptažodžių parinkimo algoritmų taikant semantinę analizę tyrimas.	2022 m. sausio mėn. – 2022 m. rugsėjo mėn.	Vykdomas apibendrinti rezultatai mokslinėje ataskaitoje.
	2.3. Empirinis tyrimas: 2.3.1. Skirtingų algoritmų palyginimas. 2.3.2. Įgyvendintų algoritmų modifikacijos, ar naujų algoritmų kūrimas, sprendžiant apibrėžtus uždavinius. 2.3.3. Sukurtų modifikacijų eksperimentinis tyrimas analizuojant jų efektyvumą	2022 m. spalio mėn. – 2023 m. gegužės mėn.	Dalinai vykdomas - pradėtas skirtingų algoritmų palyginimas.
	2.4. Gautų rezultatų analizė ir apibendrinimas	2023 m. birželio mėn. – 2023 m. rugsėjo mėn.	

Visų mokslinių tyrimų ir disertacijos rengimo etapai

	Darbo pavadinimas	Atlikimo terminai	Pastabos
3.	Atskirų daktaro disertacijos dalių (tyrimo metodikos, rezultatų, ginamų teiginių, išvadų ir kt.) parengimas: 3.1. Tikslų, uždavinių, tyrimo metodikos, ginamųjų teiginių patikslinimas. 3.2. Analitinės disertacijos dalies parengimas. 3.3. Teorinės disertacijos dalies parengimas. 3.4. Eksperimentinės disertacijos dalies parengimas. 3.5. Bendrųjų išvadų formulavimas.	2023 m. spalio mėn. – 2024 m. gegužės mėn.	
4.	Daktaro disertacijos parengimas ir svarstymas padalinyje	2024 m. birželio mėn.	
5.	Daktaro disertacijos gynimas	2024 m. rugsėjo mėn.	

Ekspertiniai tyrimai

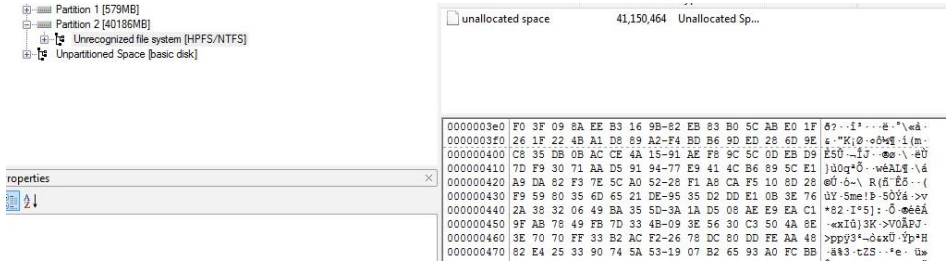
- Teisminės ekspertizės (susijusios su IT) Lietuvoje.
- Pagrindiniai užsakovai.
- Tiriamieji objektai.
- Tyrimų statistika.

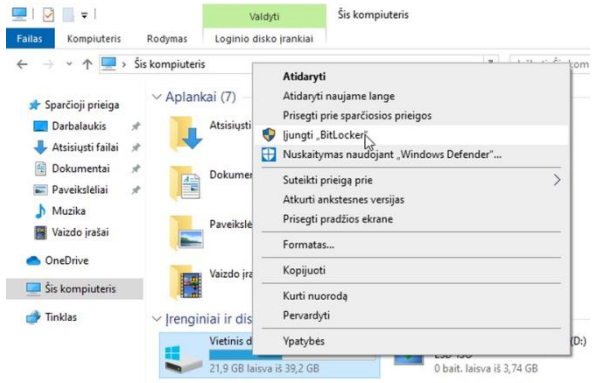


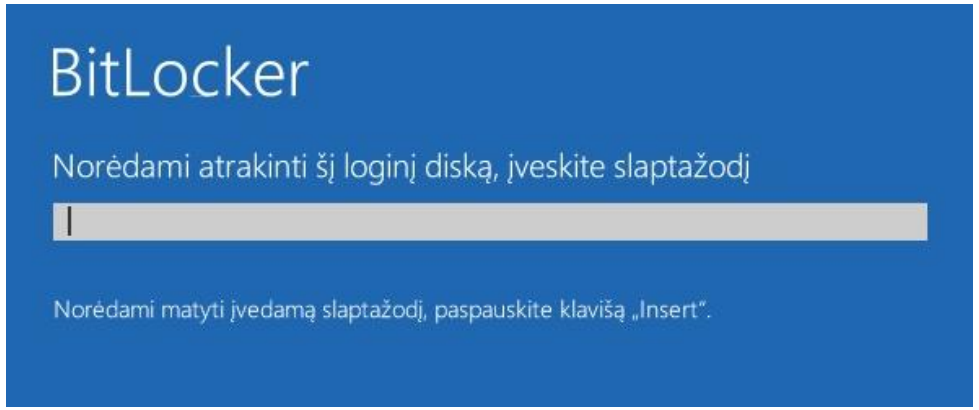
Vilniaus universitetas

Autentifikacija ir duomenų sauga

- **Skaitmeninės informacijos svarba ir saugumas.**
- **Autentifikavimas.**
- **Informacijos šifravimas.**
- **Slaptažodžiai.**
- **Pavyzdžiui** “Bitlocker” apsauga.







Vilniaus universitetas

Problemos

Kaip ištirti šifruotą informaciją?

Slaptažodžių parinkimo metodai:

- **Žodynų taikymas;**
- **Nutekintų slaptažodžių duomenų bazių panaudojimas;**
- **Pilno perrinkimo atakos („brute-force“);**
- **Kombinuotos (mišrios) slaptažodžių parinkimo atakos, skirtinguose etapuose naudojant žodynų ir „brute force“ atakas.**

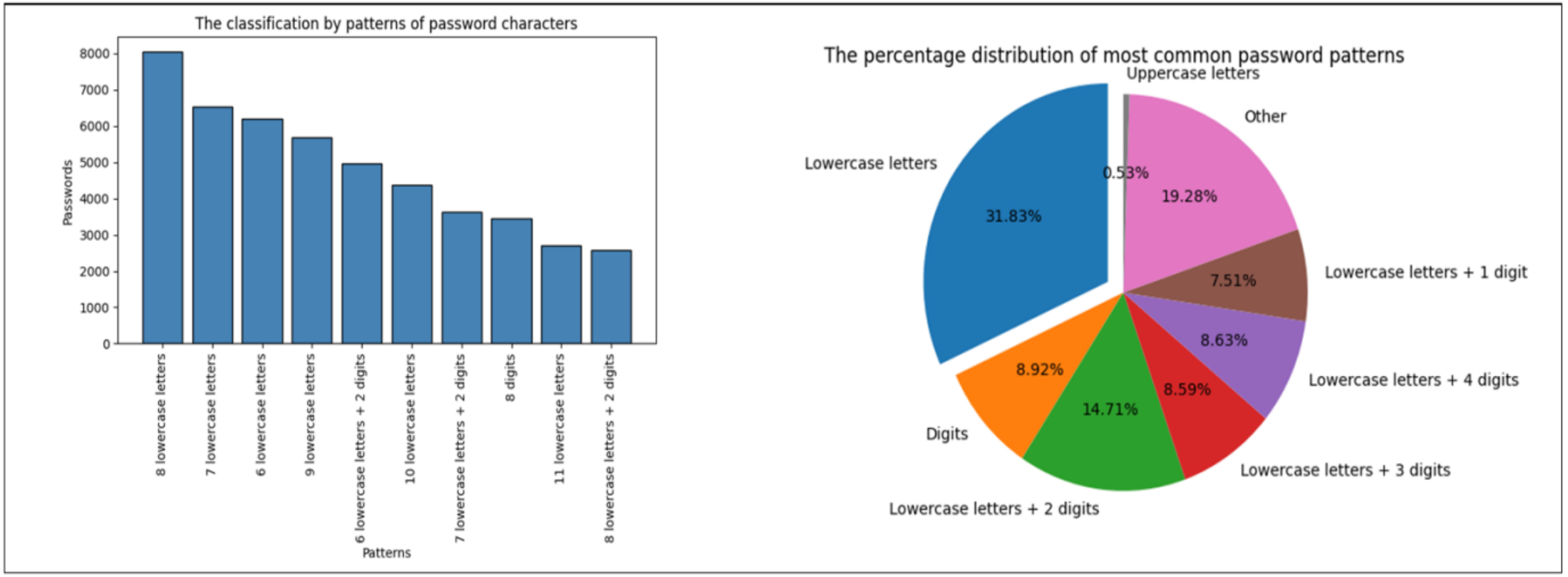
Slaptažodžių parinkimo priemonės:

- **Laboratorijos aparatūrinė įranga;**
- **Laboratorijos programinė įranga.**

Laikas (kiek galime skirti laiko ir resursų parinkti slaptažodį).

Objektai

Pagrindiniai tyrimo objektai yra: lietuviškas **slaptažodis** (ir jį atitinkantis šablonas, būdingas mūsų regiono vartotojams), bei mašininio mokymosi **algoritmas** jo parinkimui.



Tikslai ir uždaviniai

Disertacijos tikslas:

Sukurti naują arba modifikuoti (patobulinti) jau esamą slaptažodžių parinkimo metodą, adaptuotą mūsų regiono vartotojų slaptažodžių ypatybėms, pritaikytą teisminės IT ekspertizės ir tyrimo uždaviniams.

Disertacijos uždaviniai:

•Atlikti naujausių slaptažodžių parinkimo metodų apžvalgą, siekiant nustatyti tinkamiausius lietuviškų slaptažodžių parinkimui.

•Atlikti eksperimentus su empirinėmis duomenimis taikant atrinktus (pirmajame žingsnyje) slaptažodžių parinkimo metodus, siekiant rasti efektyviausius.

•Pasiūlyti metodą, naudojantį kontekstinę vartotojo informaciją, ištirti jo veikimą ir palyginti gautus rezultatus su alternatyviais metodais.

•Pasiūlyti neuroninių tinklų taikymų grįstą metodą, naudojantį kontekstinę naudotojo informaciją ir slaptažodžių šablonus, ištirti jo veikimą ir palyginti gautus rezultatus su alternatyviais metodais.

Slaptažodžių parinkimo metodai

Metodas	Algoritmas/programa
Taisyklėmis pagrįstas slaptažodžių parinkimo metodas	John the Ripper Hashcat
Markovo grandinės	OMEN
PCFG (angl. k. probabilistic context-free grammars) - tikimybiniai gramatikos taisyklių rinkiniai.	PCFG cracker
Pasikartojantys neuroniniai tinklai - Recurrent Neural Networks (RNNs).	RNN
Generatyviniai besivaržantys tinklai - Generative adversarial networks (GAN).	PassGan

Empiriniai duomenys

Duomenų bazės pavadinimas	Bendras slaptažodžių skaičius	Unikalių slaptažodžių skaičius	Šaltinis
LT1	110303	97657	https://raidforums.com
LT2	157617	114861	https://raidforums.com
LT3	645973	470298	https://raidforums.com
Rockyou			http://downloads.skullsecurity.org/passwords

Kito pusmečio darbo planas.

1. Eksperimentų pagal sudarytą teorinį ir empirinį tyrimų planą atlikimas.
2. Gautų rezultatų pristatymas tarptautinėje konferencijoje Euro 2022 (liepos 3-6 dienomis, Helsinkis, Suomija).
3. Gautų rezultatų analizė. Publikacijos (CA WoS) ruošimas.
4. Išlaikyti pasirenkamojo dalyko „Natūralios kalbos apdorojimas“ egzaminą (birželis).
5. Dalyvavimas tarptautinėje mokslinėje – praktinėje konferencijoje „Kriminalistika ir teismo ekspertologija: mokslas, studijos, praktika“ (numatoma, kad vyks 2022 m. rugsėjo mėnesį, Lietuvoje) ir mokslinių tyrimų disertacijos tema apžvalga (konferencijos darbų medžiagoje).

Vilniaus universitetas

Ačiū už dėmesį

Andrius Chaževskas

VU DMSTI doktorantas
Andrius.Chazevskas@mif.stud.vu.lt