Preverjanje moči gesel: primerjava rešitve na osnovi pametnih algoritmov in markovskih verig z obstoječimi spletnimi merilniki

Viktor Taneski, Marko Hölbl
Univerza v Mariboru, Fakulteta za elektrotehniko, računalništvo in informatiko, Koroška cesta 46, 2000 Maribor
viktor.taneski@um.si, marko.holbl@um.si

Izvleček

Gesla kljub razvoju alternativnih metod ostajajo najpogostejši način overjanja, zato je njihova moč še vedno ključnega pomena. Ob pomanjkljivostih obstoječih merilnikov moči gesel članek obravnava razvoj naprednega pristopa za njihovo preverjanje moči. Predstavljeno je spletno orodje za preverjanje moči moči gesel, ki temelji na markovskih verigah in uporablja tri modele, trenirane na različnih podatkovnih množicah. Orodje je bilo testirano na 10.000 geslih in primerjano z devetimi obstoječimi spletnimi merilniki. Za dodatno preverjanje je bila izvedena tudi analiza odpornosti na napade z uporabo verjetnostne kontekstno proste slovnice. Rezultati kažejo konkurenčno delovanje razvitega modela, a z rahlim zaostankom za najboljšimi rešitvami. Članek tako potrjuje uporabnost verjetnostnih modelov pri praktičnem izboljševanju varnosti gesel.

Ključne besede: Markovske verige, preverjanje moči gesel, varna gesla, odpornost na razbijanje gesel, verjetnostni modeli

Password Strength Evaluation: Comparison of a Solution Based on Smart Algorithms and Markov Chains with Existing Online Meters

Abstract

Despite the rise of alternative authentication methods, passwords remain the most widely used mechanism, making their strength evaluation critical. Due to the limitations of current password strength meters, this paper explores a more advanced, probabilistic approach. We present a web-based password strength evaluation tool based on Markov chains. The tool combines three models, each trained on a different dataset, and calculates the average score as the final strength estimate. The system was tested on a dataset of 10,000 passwords and benchmarked against nine popular online meters. Additionally, password cracking resistance was assessed using probabilistic context-free grammars. Results show that the proposed tool performs competitively, although it still lags slightly behind top-performing solutions. This confirms the practical value of statistical modelling in enhancing password security.

Keywords: Markov chains, password cracking resistance, secure passwords, password strength estimation, probabilistic models

1 UVOD

Kljub alternativnim metodam, kot sta biometrija in uporaba žetonov, gesla še vedno ostajajo prevladujoč mehanizem overjanja za lokalne in spletne storitve [1] in je zato razvijanje natančnejših metod za preverjanje moči gesel še vedno aktualno raziskovalno vprašanje [2], [3]. Njihov glavni namen je zaščita podatkov in sistemov, pri čemer se moč gesel meri

predvsem z odpornostjo proti nepooblaščenemu razkritju gesla [4]. V nadaljevanju uporabljamo izraz moč gesla (angl. password strength), ki označuje stopnjo odpornosti gesla proti nepooblaščenemu razkritju in ponazarja, kako zahtevno je geslo razbiti z metodami, kot so napadi z grobo silo, slovarski napadi ali napadi, ki temeljijo na verjetnostnih modelih. Moč gesla je odvisna od dolžine, kompleksnosti in nepredvidljivosti znakov, vendar uporabniki pogosto izbirajo gesla, ki si jih je preprosto zapomniti, kar povečuje njihovo predvidljivost in ranljivost za napade. Posledično napadalci pri napadih na gesla najprej preizkušajo kombinacije, ki jih uporabniki najverjetneje uporabljajo [1], [2], [3], [4], [5], [6].

Za izboljšanje varnosti gesel številne storitve uvajajo različne zahteve, kot so minimalna dolžina gesla ali uporaba posebnih znakov [2]. Kljub temu te omejitve pogosto ne zagotavljajo zadostne zaščite. Natančnejše preverjanje moči gesel omogočajo različne metrike, ki ocenijo odpornost gesla in pomagajo pri oblikovanju pravil za njihovo kreiranje. Takšne metrike so vključene tudi v merilnike moči gesel (angl. Password Strength Meter – PSM), ki vizualno predstavljajo oceno moči gesla in lahko preprečijo uporabo prešibkih gesel. Obstoječe metode ocenjevanja moči gesel delimo na tri glavne skupine: metode, ki temeljijo na napadih na gesla, hevristične metode in statistične oziroma verjetnostne metode. Prvi pristop meri odpornost gesla glede na čas, potreben za njegovo razkritje, medtem ko hevristične metode ocenjujejo gesla na podlagi pravil, kot je izračun entropije. Statistične metode, ki temeljijo na verjetnostnih modelih, se osredotočajo na pomanjkljivosti prejšnjih dveh pristopov in pogosto uporabljajo markovske verige za preverjanje moči gesla na podlagi verjetnosti pojavljanja določenih nizov znakov. Vsaka izmed teh metod ima svoje prednosti in slabosti, zato se pogosto kombinirajo za izboljšanje natančnosti [1], [4].

Kljub široki uporabi gesel njihova moč ostaja izziv, saj uporabniki pogosto dajejo prednost preprostosti in zapomnljivosti pred kompleksnostjo. Pravilniki za ustvarjanje gesel so večinoma zasnovani za zaščito pred napadi z grobo silo (angl. brute-force), vendar pogosto ne zagotavljajo zadostne zaščite. Prav tako metode ocenjevanja moči gesel postajajo manj učinkovite zaradi napredka v napadih. Ena od možnosti za izboljšanje ocenjevanja gesel je kombinacija različnih metod, ki združujejo prednosti posameznih pristopov in omogočajo večjo prilagodljivost [1], [4].

Namen raziskave je prispevati k izboljšanju moči gesel z razvojem lastne metode, ki temelji na markovskih verigah. Razvili smo izpostavljen API, ki bo za oceno moči gesel uporabljal tri različne markovske modele, vsakega treniranega na drugačnem naboru podatkov, pri čemer je končna ocena določena kot povprečje izračunov vseh treh modelov. Da bi omogočili hitrejši in enostavnejši dostop do storitve, smo aplikacijo preoblikovali v spletno storitev ter razvili Google Chrome razširitev, ki uporabnikom omogoča preverjanje moči gesel neposredno v brskalniku. Našo rešitev tudi primerjamo z več uveljavljenimi spletnimi orodji za preverjanje moči gesel ter analiziramo njeno učinkovitost v primerjavi z obstoječimi metodami. Pri razvoju in testiranju smo upoštevali tudi relevantno literaturo, ki obravnava različne pristope k preverjanju moči gesel in orodja, ki so že na voljo.

Na osnovi identificiranih problemov smo oblikovali naslednja raziskovalna vprašanja:

- **RV 1**: Kako se sodobne metode in orodja za preverjanje gesel (kot so merilniki moči gesla, preverjalniki gesel) soočajo in prilagajajo naprednim napadom na gesla?
- **RV 2**: Kakšna je funkcionalnost razvite rešitve v smislu natančnosti ocenjevanja moči gesel (brez napak) in hitrosti računanja?
- **RV 3**: Kako se učinkovitost naše rešitve primerja z drugimi uveljavljenimi merilniki moči gesel na spletu?

S to raziskavo želimo pridobiti vpogled v učinkovitost različnih pristopov k preverjanju moči gesel in razviti rešitev, ki bo izboljšala natančnost ter dostopnost tovrstnih orodij.

V nadaljevanju je članek strukturiran kot sledi: v drugem poglavju predstavimo teoretično ozadje ter pregled sorodnih del na področju ocenjevanja moči gesel. V tretjem poglavju opišemo zasnovo razvite rešitve, vključno z arhitekturo spletne storitve in razširitvijo za brskalnik. Poglavje vsebuje tudi opis eksperimentalne validacije. V četrtem poglavju primerjamo učinkovitost naše rešitve z izbranimi uveljavljenimi spletnimi merilniki moči gesel. V petem poglavju sledi diskusija dobljenih rezultatov glede na zastavljena raziskovalna vprašanja. Članek zaključujemo s šestim poglavjem, kjer podamo ključne ugotovitve in predloge za nadaljnje delo.

2 OZADJE IN SORODNA DELA

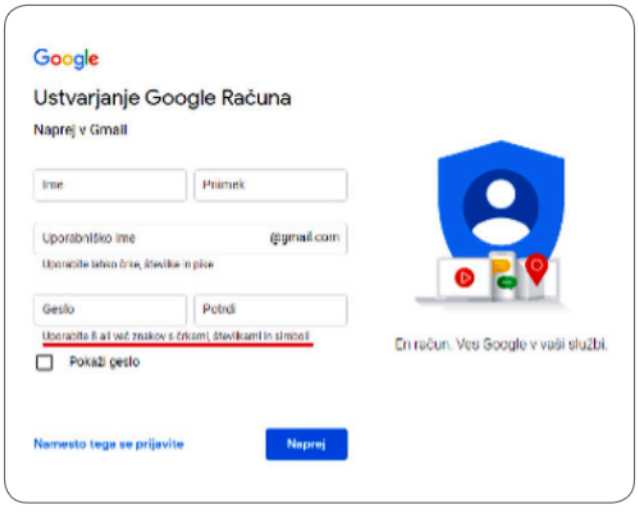
Gesla že več kot pet desetletij predstavljajo temeljni mehanizem overjanja identitete v okviru informacijske varnosti. Kljub razvoju številnih sodobnejših metod, kot so biometrični sistemi (npr. prepoznavanje glasu, prstnega odtisa ali obraza), overjanje z uporabo žetonov (npr. pametne kartice, mobilne naprave) ter uvedba večfaktorskega overjanja, gesla še naprej ostajajo prevladujoča izbira. Njihova priljubljenost izhaja predvsem iz nizkih tehničnih zahtev, enostavne implementacije ter uporabniške dostopnosti, kar jim v praksi daje izrazito prednost pred alternativnimi rešitvami. Slednje se pogosto soočajo z izzivi, kot so višji stroški uvedbe, kompleksnost vzdrževanja ter zmanjšana uporabniška sprejemljivost. Poleg tega na strani uporabnikov pomembno vlogo igrata občutek domačnosti in udobje pri uporabi – na primer, stalna uporaba fizičnih žetonov lahko sčasoma postane moteča. Ob upoštevanju navedenih prednosti ter široke razširjenosti gesel je mogoče sklepati, da bodo ta še naprej igrala pomembno vlogo v sodobnih sistemih informacijske varnosti [2], [7], [8], [9].

2.1 Trenutno stanje uporabniških gesel

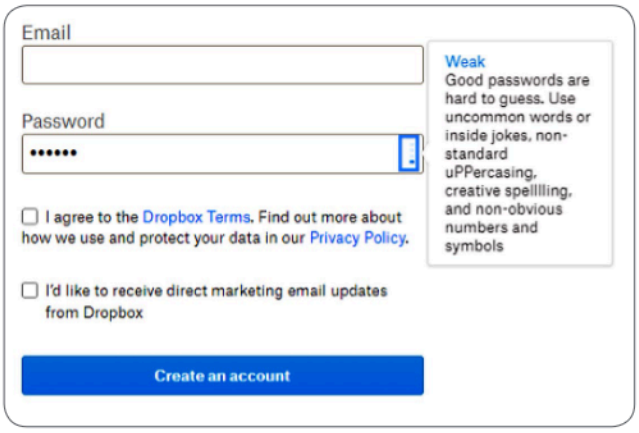
Z izjemno hitro rastjo spletnih tehnologij se je močno povečala tudi uporaba spletnih storitev in družbenih omrežij tako v rekreativne kot poslovne namene. Overjanje z gesli tako ostaja ključen mehanizem za zaščito in identifikacijo velikega števila uporabniških računov. Stopnja varnosti gesel pa je v veliki meri odvisna od same izbire gesla s strani uporabnika, kar že vrsto let predstavlja temeljno varnostno ranljivost. Uporabniki se pri oblikovanju gesel pogosto odločajo med preprostostjo, ki omogoča lažje pomnjenje, in kompleksnostjo, ki pa zagotavlja višjo raven varnosti. Kljub zavedanju o pomenu močnih gesel, številni posamezniki še vedno izbirajo krajša, enostavna gesla, ki pogosto vključujejo osebne podatke, kot so imena, rojstni datumi, telefonske številke ali običajne slovarske besede. Dodatno ranljivost prinašajo tudi slabe navade, kot so ponovna uporaba enakega gesla za več različnih računov, zapisovanje gesel na lahko dostopna mesta (npr. list papirja ali računalniško namizje) ter neredno posodabljanje gesel [2], [5], [8], [9], [10], [11].

Po podatkih raziskave iz leta 2021 kar 53 % anketirancev v svoja gesla vključuje osebne informacije, ki napadalcem omogočajo lažje ugibanje. Poleg tega kar 74 % uporabnikov isto geslo ponovno uporabi za več različnih storitev, gesla pa redko posodabljajo. Tovrstne prakse sicer olajšajo upravljanje z gesli, vendar hkrati bistveno povečujejo njihovo predvidljivost in dovzetnost za napade [9].

Z namenom zmanjšanja vpliva človeške malomarnosti so številni ponudniki spletnih storitev uvedli določene varnostne politike oziroma zahteve za oblikovanje gesel, npr. minimalna dolžina gesla, obvezna uporaba velikih črk, številk in posebnih znakov. Primer tovrstnih zahtev je prikazan na sliki 1, kjer Google ob ustvarjanju računa zahteva geslo z vsaj osmimi znaki, ki vključuje črke, številke in simbole [12]. Kljub temu te zahteve pogosto ne zadostujejo za doseganje ustrezne varnostne ravni, saj uporabniki še vedno pogosto sledijo predvidljivim vzorcem. Kot dopolnitev teh pravil se zato uvajajo tudi merilniki moči gesel (angl. password strength meter), ki uporabnikom nudijo vizualno povratno informacijo o ocenjeni moči njihovega gesla. Na sliki 2 je prikazan primer takega sistema pri ustvarjanju računa na platformi Dropbox, kjer je moč gesla ponazorjena s štirimi kvadrati, ki se obarvajo glede na moč vnešenega gesla. V prikazanem primeru se obarva le en kvadrat, kar pomeni, da sistem geslo oceni kot šibko [13]. Algoritmi za vrednotenje moči gesel, ki te ocene izračunajo, so podrobneje predstavljeni v tretjem poglavju. V primeru, da geslo ne ustreza določenim kriterijem ponudnika, se le-to preprosto zavrne, pri čemer nekatere platforme celo prepovedujejo uporabo določenih gesel [8], [10], [14], [15].



Slika 1: **Primer zahtev za izbiro gesla pri ustvarjanju Google računa. Uporabnik mora ustvariti geslo, ki vsebuje vsaj 8 znakov s črkami, številkami in simboli [12].**

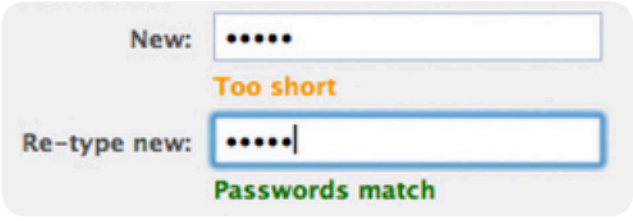


Slika 2: **Primer merilnika moči gesel pri ustvarjanju Dropbox računa [13].**

Kot že omenjeno, uporabniki pogosto predstavljajo najšibkejši člen v procesu overjanja. Na izbiro gesla vpliva več dejavnikov, med katerimi so ključna pravila za oblikovanje gesel, ki jih določijo ponudniki storitev. Uporabniki se morajo tem zahtevam podrediti, pogosto pa so tudi aktivno spodbujeni ali celo prisiljeni k spremembi gesel. Čeprav ta pravila sama po sebi ne zagotavljajo visoke ravni varnosti, dokazano pozitivno vplivajo na oblikovanje varnejših gesel. Pomemben vpliv ima tudi sam uporabnik – njegova osebnost, navade in razumevanje varnosti. Vsak posameznik si različno razlaga, kaj pomeni močno geslo, kar lahko zaradi pomanjkanja znanja vodi v suboptimalne izbire. Eden ključnih razlogov za izbiro šibkejših gesel je tudi človeški spomin. Povprečen uporabnik ima lahko do osem različnih gesel, katerih si je pogosto težko zapomniti. Posledično uporabniki pogosteje posegajo po lažje zapomljivih in s tem manj varnih geslih ter jih ponovno uporabljajo na različnih platformah [8].

2.2 Merilniki moči gesel

Številne sodobne spletne storitve za izboljšanje kakovosti gesel uporabnikov vključujejo tako imenovane merilnike moči gesel. Veliki ponudniki, kot je Google, pogosto uporabljajo enoten merilnik za več svojih platform, medtem ko drugi, kot na primer Ebay, uporabljajo različne merilnike glede na geografsko lokacijo storitve. Osnovni namen teh mehanizmov je spodbujanje uporabnikov k izbiri močnejših gesel s pomočjo vizualnih indikatorjev, ki podajajo oceno trenutne kakovosti gesla in priporočila za njegovo izboljšavo. Vizualne povratne informacije se lahko realizirajo v obliki barvnih lestvic, obarvanih besed (slika 3), kljukic (slika 4), indikatorjev napredka (slika 5) ali kontrolnih seznamov (slika 6). Uporabniška izkušnja se ob tem izboljša, če je prikaz ocene hiter in interaktiven; nasprotno pa lahko zamik ali nedoslednost med vnosom gesla in prikazom ocene zmanjša uporabnikovo zaupanje in motivacijo.



Slika 3: **Primer merilnika moči gesel, ki vrača povratne informacije le z različno pobarvanimi besedami [14].**

Slika 4: **Primer merilnika moči, ki vrača povratne informacije s kljukicami. Prav tako je v pomoč zelena lestvica [14].**

Slika 5: **Primer merilnika moči, ki vrača oceno gesla s pomočjo barvne lestvice. Geslo umesti v enega izmed petih razredov [14].**

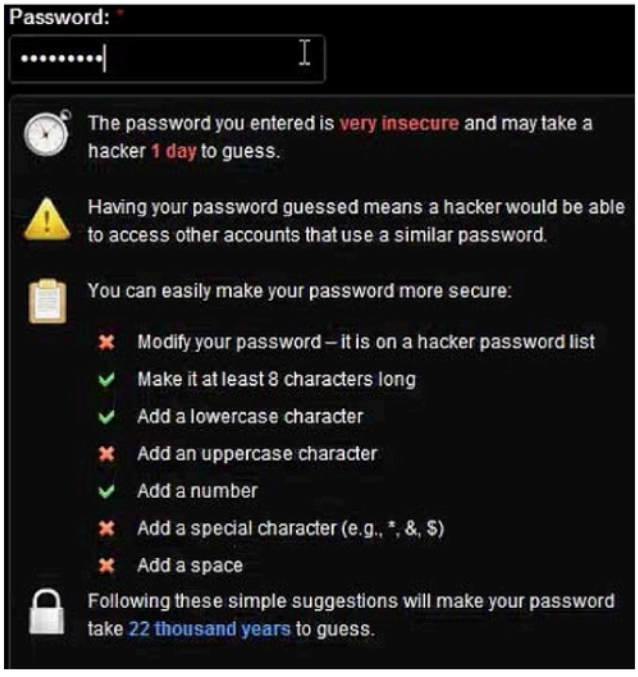
Slika 6: **Primer merilnika moči, ki vrača povratne informacije s pomočjo seznama [14].**

V večini primerov merilniki gesel razvrščajo gesla v tri (npr. šibko, srednje močno, močno) ali pet kakovostnih razredov (npr. zelo šibko, šibko, srednje, dobro, zelo dobro), ki so pogosto označeni z barvami (npr. rdeča za šibko, zelena za močno). Le redko pa merilniki posredujejo konkretno številčno oceno ali podatke o metodologiji izračuna, ki običajno poteka v ozadju. Nekateri sicer uporabijo odstotkovno oceno ali oceno časa, potrebnega za razbitje gesla glede na uporabljeni vrednotenjski sistem [7], [14], [15], [16].

Čeprav prisotnost merilnikov na splošno spodbuja uporabnike k izbiri daljših in kompleksnejših gesel, imajo največji učinek strožji merilniki, ki temeljijo na natančnih in restriktivnih kriterijih. Takšni merilniki pogosto vključujejo zahteve po uporabi raznovrstnih znakov (številke, simboli, velike črke) in vodijo k daljšemu premisleku uporabnika o ustreznosti izbranega gesla. Po drugi strani pa pretirano strogi ali slabo zasnovani merilniki lahko povzročijo frustracije, zmanjšajo pripravljenost uporabnika za sodelovanje in vodijo k izbiri enostavnih, a formalno skladnih gesel. Uporabniki so nagnjeni k večjemu trudu pri ustvarjanju gesel, kadar merilnik omogoča vizualno povratno informacijo, ki je jasna, razumljiva in motivacijska.

Kljub razširjenosti merilnikov številne spletne strani še vedno uporabljajo neustrezne in nenatančne mehanizme, ki vračajo nedosledne rezultate. Tako se lahko zgodi, da eno geslo pri eni storitvi velja za močno, pri drugi pa za šibko [7], [14], [15], [16]. Učinkovitost merilnikov je pogosto odvisna od konteksta, in sicer v primeru manj pomembnih računov so uporabniki nagnjeni k uporabi že obstoječih gesel, zaradi česar merilniki nimajo bistvenega vpliva. Za naslavljanje teh izzivov nekateri merilniki vključujejo elemente socialnega vpliva, kot so primerjave z drugimi uporabniki (slika 7), ali opozorila na varnostna tveganja, ki spremljajo slaba gesla, skupaj s predlogi za izboljšave (slika 8) [7], [16].

Slika 7: **Primer merilnika moči, ki gesla primerja z drugimi uporabniki [7].**

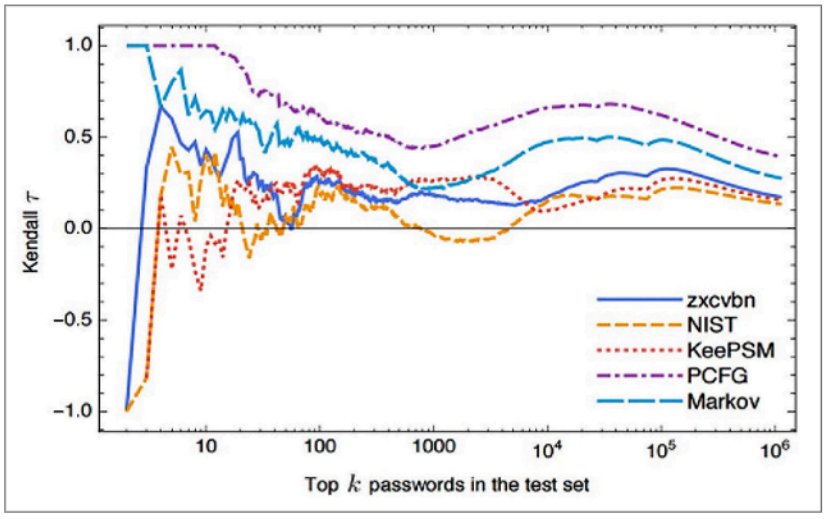


Slika 8: **Primer sporočil pri interaktivnem merilniku moči, ki opominjajo na nevarnost [16].**

Empirične študije so pokazale, da zgolj interaktivnost merilnikov ni dovolj za doseganje varnejšega vedenja uporabnikov – pomembno je tudi, da merilnik jasno predstavi resnost tveganja in učinkovitost uporabniškega odziva nanj [7], [14], [15], [16].

Večina spletnih merilnikov moči gesel še vedno temelji na hevrističnih pristopih, predvsem zaradi njihove enostavne implementacije in nizkih stroškov delovanja, čeprav so rezultati teh metod pogosto nezadovoljivi. Znotraj hevrističnih metod se običajno ločujeta dva temeljna pristopa. Prvi temelji na vnaprej določenih pravilih oziroma zahtevah, ki jih mora geslo izpolnjevati za pridobitev višje ocene in najpogosteje gre za pravila, skladna s smernicami NIST. Drugi pristop vključuje prepoznavanje vzorcev v geslu, pri čemer sistem išče ujemanja s preddefiniranimi semantičnimi ali zaporednimi strukturami.

V zadnjih letih se razvijajo sodobnejši, bolj prilagodljivi merilniki, ki temeljijo na simulaciji napadov na gesla. Ti merilniki se v grobem delijo na tri skupine: (1) merilniki, ki temeljijo na orodjih za ugibanje gesel, (2) merilniki, ki uporabljajo verjetnostne modele, ter (3) merilniki, ki temeljijo na nevronskih mrežah. V nadaljevanju se osredotočamo na verjetnostne pristope, ki najpogosteje temeljijo na markovskih verigah ali na verjetnostnih kontekstno prostih slovnicah (angl. Probabilistic Context-Free Gram-



Slika 9: **Primerjava raznih tipov merilnikov moči gesel [18].**

mars - PCFG). Ti merilniki praviloma dosegajo boljšo natančnost pri ocenjevanju ranljivosti gesel (slika 9), vendar kljub temu vodilni ponudniki spletnih storitev še vedno uporabljajo klasične hevristične rešitve. Eden od razlogov za to je pomanjkanje obsežnih empiričnih primerjav med najnovejšimi akademskimi modeli in uveljavljenimi industrijskimi orodji, kot sta Zxcvbn in KeePass. Zaradi tega številne prednosti in pomanjkljivosti sodobnih pristopov ostajajo nepoznane širši skupnosti, kar zavira njihovo širšo uveljavitev. Pomembna omejitev verjetnostnih modelov je predpostavka, da uporabniki vedno generirajo nova gesla, pri čemer zanemarjajo pogosto prakso ponovne uporabe gesel. Na to težavo opozarja več novejših raziskav, vključno z deli [17] in [18], ki predstavljajo izboljšane verjetnostne modele, upoštevajoč ta pomemben dejavnik [3], [6], [17], [18], [19], [20].

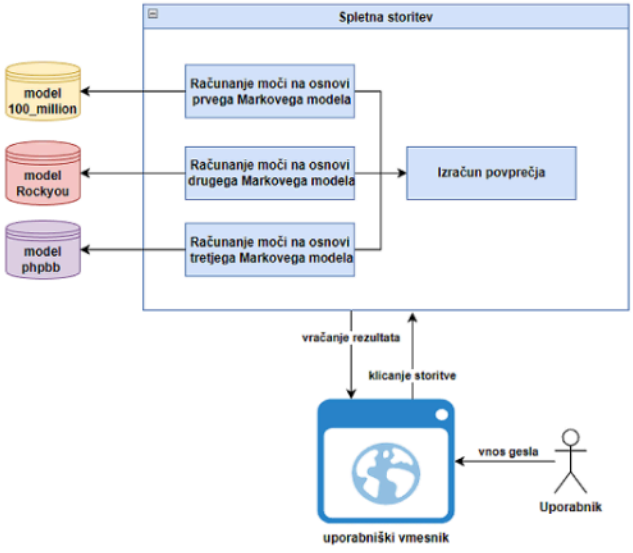
3 PREVERJANJE MOČI GESEL Z UPORABO MARKOVSKIH VERIG

Na področju varnosti gesel želimo prispevati z razvojem novega verjetnostnega ocenjevalnika moči gesel, ki temelji na kombinaciji več modelov. Naša rešitev temelji na pristopu predstavljenem v [11], kjer so bile razvite markovske verige na podlagi dvanajstih različnih podatkovnih zbirk. Za potrebe naše raziskave smo izbrali tri modele z najvišjo uspešnostjo: model, osnovan na zbirki *10_million*, model na zbirki *rockyou* ter model na zbirki *phpbb*. Ocenjevalnik gesel izračuna moč gesla na podlagi vsakega izmed teh treh modelov, končni rezultat pa predstavlja aritmetično povprečje vseh treh izračunov.

Razvoj rešitve smo razdelili na dva ključna dela (slika 10). Prvi del predstavlja spletna storitev, ki v ozadju izvaja izračune moči gesla na podlagi izbranih markovskih modelov. Drugi del predstavlja razširitev za spletni brskalnik Chrome, ki služi kot uporabniški vmesnik. Ta omogoča uporabniku vnos gesla ter komunikacijo s spletno storitvijo v ozadju. Po posredovanju gesla se izvede analiza s pomočjo treh modelov, uporabniku pa se vrne povprečna ocena varnosti gesla.

3.1 Spletna storitev

Za namen izračuna moči gesel v ozadju smo razvili spletno storitev v obliki aplikacijskega programskega

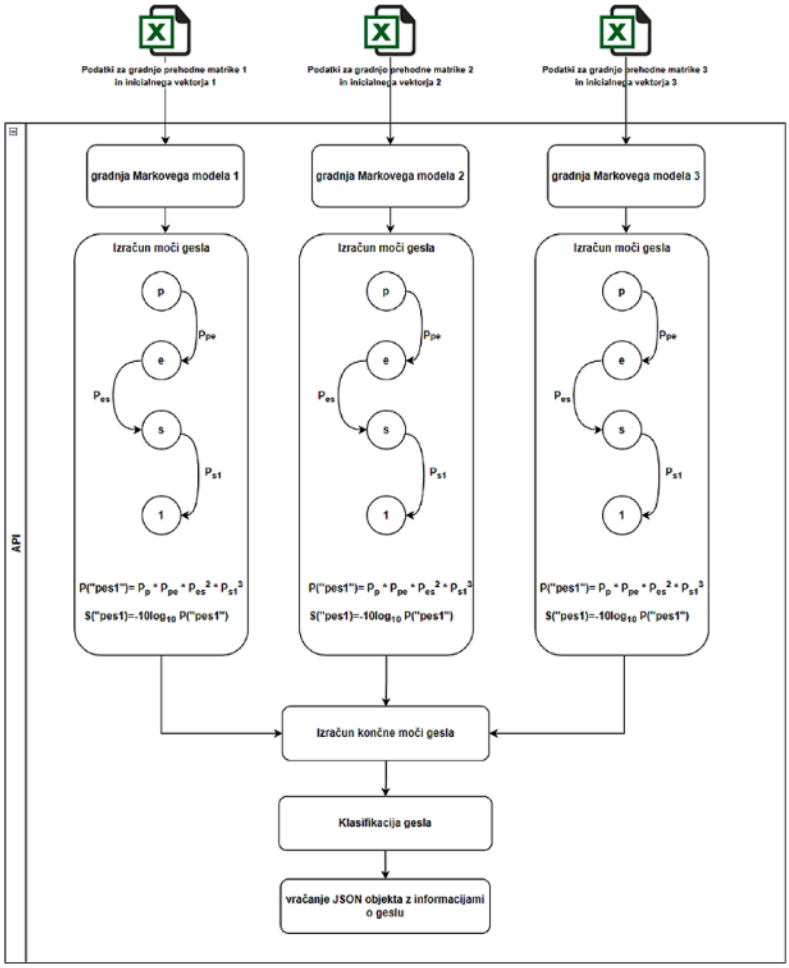


Slika 10: **Preprost prikaz strukture aplikacije.**

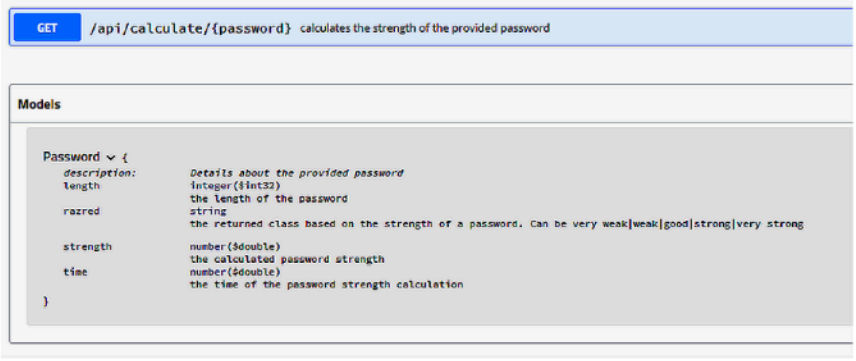
vmesnika (angl. Application Programming Interface - API). Storitev je implementirana z uporabo ogrodja Spring Boot za programski jezik Java, ki omogoča hitro in učinkovito vzpostavitev REST API-jev. Za uporabo jave smo se odločili, ker izhajamo iz obstoječe namizne aplikacije za preverjanje moči gesel opisane v [11], ki smo jo želeli nadgraditi v spletno storitev in s tem omogočiti enostavnejši dostop do funkcionalnosti. Razvoj je potekal v okolju IntelliJ IDEA.

Osrednja funkcionalnost API-ja obsega gradnjo markovskih verig in izračun verjetnosti gesla glede na vsak posamezen model. Izračun verjetnosti gesla temelji na začetnem vektorju in prehodni matriki, ki opisuje verjetnosti prehodov med znaki v geslu. Začetni vektor določa verjetnosti, da se posamezen znak pojavi na začetku gesla, prehodna matrika pa opisuje verjetnost, da določen znak sledi drugemu znaku. Na ta način markovska veriga omogoča izračun skupne verjetnosti gesla kot produkta začetne verjetnosti prvega znaka in verjetnosti prehodov za vse nadaljnje znake. Višja kot je ta verjetnost, večja je predvidljivost gesla, posledično pa je geslo šibkejše. Podrobnejši matematični opis metode je na voljo v [11]. Tako je vsak model definiran z začetnim vektorjem in prehodno matriko, katerih dimenzije so odvisne od nabora znakov, ki jih obravnavamo pri ocenjevanju moči gesla. V našem primeru obravnavamo 95 ASCII znakov, kar vključuje velike in male črke, številke ter posebne znake. Prehodna matrika je zato definirana kot dvodimenzionalno polje velikosti 95 × 95, začetni vektor pa kot enodimenzionalno polje dolžine 95. Vse vrednosti začetnih in verjetnosti prehodov za tri izbrane modele so shranjene v ločenih Excel dokumentih.

API vsebuje en glavni klic, ki kot vhodni parameter prejme geslo. Ob prejemu klica se najprej zgradijo markovski modeli iz pripadajočih Excelovih datotek, nato pa se za vsakega izmed modelov izvede izra-



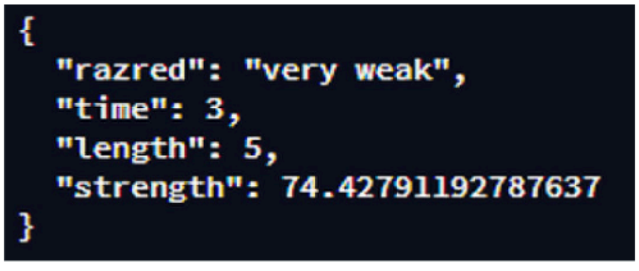
Slika 11: **Vizualna predstavitev delovanja spletne storitve za izračun moči gesla »pes1«.**



Slika 12: **Primer dokumentacije za klic izračuna moči gesla in razred Password.**

čun moči gesla, skladno s postopki predstavljenimi v [11]. Proces se izvede vzporedno za vse tri modele, končni rezultat pa predstavlja povprečje vseh treh izračunov. Celoten potek je prikazan na sliki 11. Na podlagi dobljene vrednosti se geslo nadalje klasificira v enega izmed petih varnostnih razredov: zelo šibko, šibko, dobro, močno in zelo močno. Uporabniku se vrne strukturiran JSON objekt, ki vsebuje razred gesla, dolžino gesla, čas izračuna in samo oceno moči.

Za lažje razumevanje in testiranje smo razvili tudi dokumentacijo API-ja z uporabo orodja SwaggerUI. Dokumentacija podaja opise vseh dostopnih klicev ter pojasnila ključnih razredov. Na sliki 12 je prikazana dokumentacija klica za izračun moči gesla, ki kot parameter prejme le geslo. Ob izbiri klica se prikažejo tudi podrobnejši podatki o možnih odgovorih in strukturi vrnjenega objekta. Na sliki je predstavljen tudi razred Geslo, katerega instanca se vrača v obliki JSON objekta. Primer konkretnega odgovora je prikazan na sliki 13: za geslo 'drevo' je bila ocena moči 74, izračun pa je trajal tri sekunde, kar je sistem uvrstil v razred »zelo šibko«.



Slika 13: **Primer vrnjenega JSON objekta za geslo »drevo«.**

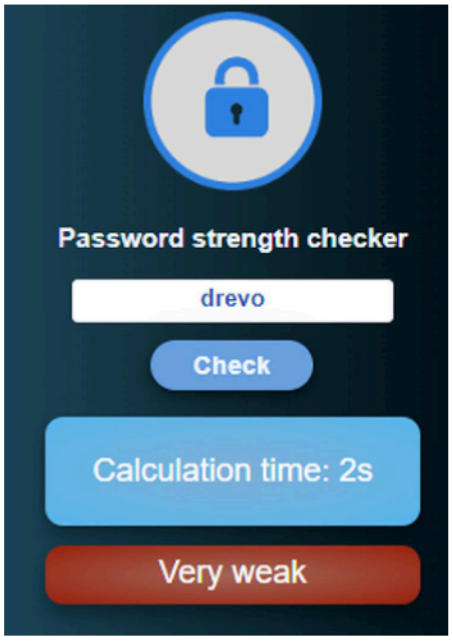
3.2 Uporabniški vmesnik

Za zagotavljanje boljše uporabniške izkušnje in lažji dostop do razvitega API-ja smo implementirali preprost uporabniški vmesnik v obliki razširitve za spletni brskalnik Google Chrome. Razširitev je bila razvita v programskem jeziku JavaScript, pri čemer smo za razvoj uporabljali okolje Visual Studio Code. Glavni namen razširitve je delovati kot merilnik moči gesel: uporabnik vpiše geslo v vnosno polje in z enim klikom sproži klic API-ja.

Po izvedbi klica razširitev prejme JSON objekt in izpiše izbrane informacije – konkretno: razred moči gesla ter čas, potreben za izračun. Samo numerične vrednosti ocene moči ne prikazujemo, saj uporabniku ta podatek ne prinaša dodatne vrednosti. Namesto tega je razred moči predstavljen tudi z barvno oznako, s čimer smo želeli izboljšati vizualno razumljivost rezultata: šibko geslo je označeno z rdečo, dobro z rumeno, in močno geslo z zeleno barvo.

Na sliki 14 je prikazan primer rezultata za zelo šibko geslo. Poleg osnovnega delovanja razširitev vključuje tudi odzive na posebne dogodke, s katerimi dodatno informira uporabnika. V primeru, da je vnosno polje prazno, uporabnika opozori na potrebo po vnosu gesla pred izvedbo preverjanja. Prav tako razširitev sproži obvestilo v primeru daljšega trajanja izračuna, s čimer obvešča uporabnika o morebitnem čakanju na rezultat.

Razvita Chrome razširitev je namenjena raziskovalni in pedagoški uporabi ter je dostopna na zahtevo. Gesla se obdelujejo izključno v pomnilniku za čas ocenjevanja, zato rešitev ne shranjuje uporabniških podatkov in ne predstavlja varnostnega tveganja. Takšne odprtokodne akademske rešitve prispevajo k ozaveščanju o kakovosti gesel ter predstavljajo varnejšo alternativo komercialnim pristopom, kjer obstaja nevarnost zlorabe shranjenih podatkov.



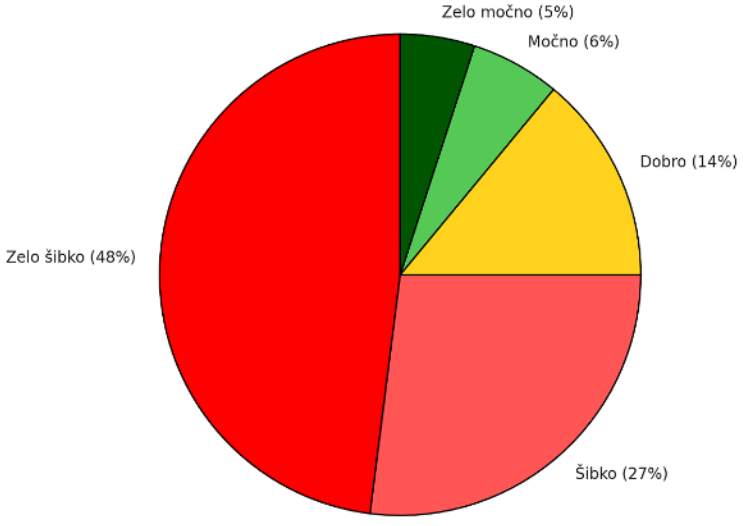
Slika 14: **Primer prikaza rezultatov za šibko geslo.**

3.3 Eksperimentalna validacija funkcionalnosti rešitve

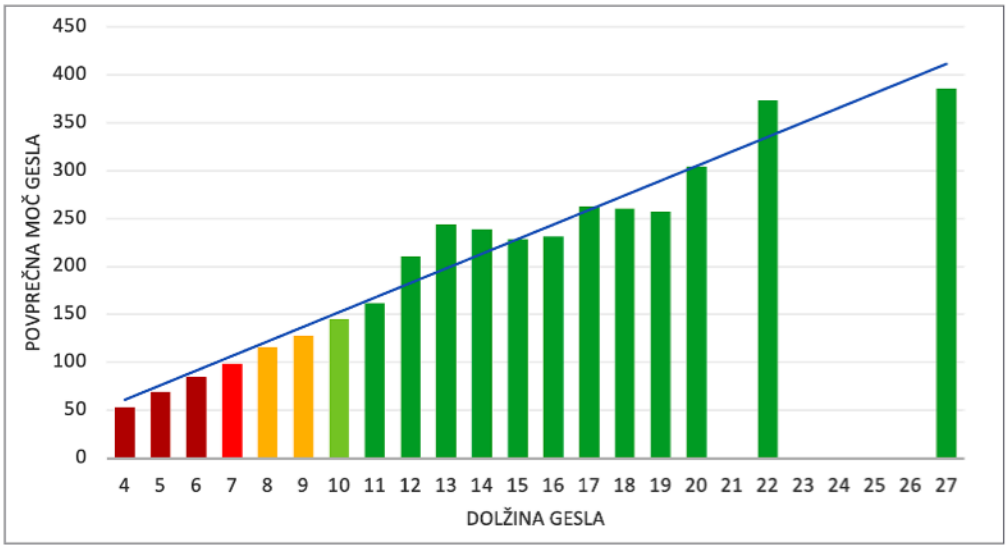
Delovanje razvite rešitve smo preverili s pomočjo zbirke testnih gesel. V ta namen smo izvedli testiranje na 10.000 geslih, ki smo jih pridobili iz javno dostopnih podatkovnih zbirk Rockyou in Phpbb. Gesla smo shranili v besedilni obliki v ustrezno datoteko. Zaradi strojnih omejitev naprave, uporabljene za testiranje (kot opisano v uvodnem delu), smo se odločili za manjši obseg testne zbirke. Procesiranje celotnega nabora 10.000 gesel je trajalo približno 27 ur, pri čemer je bil povprečni čas obdelave enega gesla okoli 9 sekund. Sistem je uspešno izračunal moč in vrnil pripadajoče podatke za vsa gesla. Rezultati klasifikacije gesel glede na njihovo ocenjeno moč so prikazani na sliki 15. Analiza kaže, da je bila kar polovica gesel ocenjena kot zelo šibka (48 %), dodatnih 27 % pa kot šibka. Le 14 % gesel je bilo označenih kot dobra, medtem ko je 6 % gesel doseglo oznako močna in 5 % zelo močna.

V okviru testiranja smo izvedli tudi dodatno analizo, s katero smo želeli preučiti razmerje med dolžino gesla, ocenjeno močjo in časom potrebnim za izračun. Podatki te analize so predstavljeni v tabeli 1. Prikazani rezultati predstavljajo povprečne vrednosti, dobljene pri testiranju naše rešitve na zbirki 10.000 gesel. Za vsako dolžino gesla so izračunani povprečni časi obdelave, povprečne ocene moči in število gesel v posamezni skupini.

Dolžina testnih gesel je variirala med 4 in 27 znaki. Največ gesel je vsebovalo med 6 in 8 znakov, število daljših gesel pa se je hitro zmanjševalo, pri čemer je bilo zelo malo primerov gesel z več kot 15 znaki. Rezultati so pokazali pričakovano korelacijo med dolžino gesla, njegovo močjo in časom obdelave. Gesla, dolga med 4 in 7 znaki, so bila večinoma ocenjena kot šibka, medtem ko so gesla z dolžino 8 ali 9 znakov dosegla oceno dobra. Gesla z 10 ali več znaki so bila večinoma uvrščena v kategorijo močnih. Povezava med dolžino in ocenjeno močjo gesla je dodatno ponazorjena na sliki 16. Analiza časa obdelave je pokazala, da se trajanje izračuna moči gesla z naraščajočo dolžino hitro povečuje. Že za gesla z dol-



Slika 15: **Klasifikacija 10 000 gesel glede na izračunano moč.**



Slika 16: **Moč gesel glede na dolžino gesel. Obarvanost posameznih stolpcev predstavlja moč gesla (rdeča barva-šibko geslo, rumena barva-dobro geslo, zelena barva-močno geslo).**

žino 8 znakov je povprečni čas presegel 10 sekund. Najkrajši časi obdelave so bili zabeleženi pri geslih dolžine med 4 in 6 znakov, kjer je bil čas obdelave v povprečju krajši od 10 sekund. Najdaljše geslo v testni zbirki, dolgo 27 znakov, je sistem obdeloval kar 58 sekund.

Tabela 1: **Prikaz podatkov za dolžino gesel, povprečni čas izračuna, povprečno moč in število gesel.**

Dolžina gesla	Povprečen čas izračuna	Povprečna moč	Število gesel
4	2	53,3	785
5	4,9	69,2	794
6	7	84,6	2907
7	9,6	98,4	1938
8	12	115,8	2479
9	14,2	128,04	434
10	17	145,4	252
11	19	161,7	124
12	22	210,5	118
13	24	243,9	57
14	26,7	239,3	37
15	29	228,9	31
16	31,1	231,5	15
17	34	262,7	12
18	36	260,3	6
19	39	257	4
20	41	304,5	3
21			0
22	46	373	3
23			0
24			0
25			0
26			0
27	58	386	1

4 PRIMERJAVA REŠITVE Z OBSTOJEČIMI MERILNIKI

Za preverjanje učinkovitosti razvite spletne rešitve za preverjanje moči gesel smo izvedli empirično primerjavo z več uveljavljenimi spletnimi merilniki. Cilj primerjave je bil oceniti, kako različni pristopi – tako heuristični kot probabilistični – razvrščajo gesla glede na njihovo domnevno varnost, v kolikšni meri se njihove ocene skladajo z dejansko odpornostjo gesel proti napadom ter kakšna je stopnja ujemanja med posameznimi orodji.

V raziskavo smo vključili naslednje merilnike: Password Monster [21], UIC ocenjevalec gesel [22], BitWarden [23], Password Meter [24], Nordpass [25], Kaspersky [26], Lastpass [27], Haveibeenpwned [28] in Zxcvbn [29]. Vsak od teh merilnikov uporablja svoj mehanizem za preverjanje moči – nekateri temeljijo na enostavnih pravilih (npr. dolžina gesla, uporaba velikih črk, številk in simbolov), drugi na ocenah časa razbijanja, spet tretji na podatkovno podprtih modelih. Rešitev, ki smo jo razvili temelji na povprečju treh markovskih modelov (1-gram, 2-gram, 3-gram), ki so bili trenirani na treh ločenih realnih podatkovnih zbirkah gesel (model, osnovan na zbirki *10_million*, model na zbirki *rockyou* ter model na zbirki *phpbb*), kar omogoča robustnejšo oceno na osnovi dejanske pogostosti zaporedij znakov.

Za analizo smo uporabili 10.000 testnih gesel, pri čemer je vsak merilnik podal svojo oceno moči. Osem izbranih merilnikov razvršča gesla v ustrezne razrede na podlagi ocenjene moči. Edina izjema je Have I Been Pwned, ki namesto ocene varnosti poda število pojavitev gesla v javno dostopnih podatkovnih zbirkah. V tej raziskavi smo za potrebe enotne klasifikacije uvedli naslednjo razvrstitev: gesla, ki se ne pojavijo, so označena kot zelo močna; gesla z največ dvema pojavitvama kot močna; gesla z 2 do 5 pojavitvami kot dobra; gesla z več kot 5 in do 50 pojavitvami kot šibka; in gesla z več kot 50 pojavitvami kot zelo šibka. Zaradi različnega načina preverjanja moči med našim orodjem (preverjanje moči na osnovi verjetnosti) in izbranimi merilniki (preverjanje moči na osnovi hevristik) so lahko rezultati primerjave nezadovoljivi. V namen natančnejše analize smo odpornost gesel v zbirki testirali tudi s simuliranim napadom z uporabo PCFG, ki velja za enega izmed učinkovitejših pristopov k razbijanju gesel.

V tabeli 2 je prikazana porazdelitev gesel po merilnikih. Rezultati kažejo, da je v primerjavi z ostalimi izbranimi merilniki naša rešitev dosledno zaznala nižji delež gesel kot šibkih oziroma zelo šibkih. Skupno je bilo 75 % vseh gesel s strani našega orodja klasificiranih kot šibkih ali zelo šibkih, medtem ko so ostali merilniki v povprečju kar 91 % gesel uvrstili v to kategorijo. Najstrožji med njimi so bili merilniki Nordpass, Kaspersky in Lastpass, saj so vsi trije kot šibka označili več kot 95 % analiziranih gesel, kar kaže na zelo konservativen kriterij ocenjevanja.

Obraten trend pa je bil zaznan pri klasifikaciji močnih ali zelo močnih gesel. Naša rešitev je kot takšna prepoznala 11 % gesel, kar je občutno več kot pri večini ostalih merilnikov. Povprečno so drugi merilniki kot močna označili le 4 % gesel, pri čemer so Nordpass, Kaspersky in Lastpass ponovno izstopali z najnižjim deležem – zgolj 1 % gesel so uvrstili med močne. Izmed vseh orodij sta največji delež gesel, označenih kot močna (6 %), zabeležila Zxcvbn in PasswordMonster, ki imata nekoliko bolj uravnoteženo porazdelitev.

Tabela 2: **Delež gesel pri tri delitvi na tri razrede.**

	šibko	**dobro**	**močno**
Naša rešitev	75%	14%	11%
PasswordMonster	89%	5%	6%
UIC	89%	7%	4%
BitWarden	94%	4%	2%
PasswordMeter	89%	7%	4%
Nordpass	96%	3%	1%
Kaspersky	97%	2%	1%
Lastpass	95%	4%	1%
Haveibeenpwned	90%	3%	7%
Zxcvbn	83%	11%	6%
Povprečje	91%	5%	4%

Zanimala nas je tudi podobnost klasifikacije za vsak posamezen razred (npr. koliko gesel, ki jih je naša rešitev označila za šibka je tudi druga rešitev označila za šibka). Rezultati te primerjave so predstavljeni v tabeli 3.

Tabela 3: **Enakost ocenjevanja za posamezen razred.**

	šibko	dobro	močno
password monster	0,979	0,1	0,42
UIC	0,959	0,11	0,25
bitwarden	0,99	0,04	0,14
passwordmeter	0,959	0,11	0,25
nordpass	0,99	0,01	0,05
kaspersky	0,99	0,01	0,04
lastpass	0,99	0,04	0,12
haveibeenpwned	0,95	0,03	0,3
zxcvbn	0,93	0,22	0,41
povprečje	0,970777778	0,074444444	0,22

Rezultati, prikazani v tabeli 4, kažejo, da med izbranimi merilniki in razvito rešitvijo obstaja visoka stopnja soglasja pri razvrščanju šibkih gesel. V povprečju kar 97 % gesel, ki jih naš model oceni kot šibka, enako ovrednotijo tudi ostali merilniki. To kaže na precejšnje soglasje glede kriterijev za identificiranje najšibkejših gesel. Povsem drugačna slika pa se pojavi pri razvrščanju gesel v razreda dobra in močna, kjer so razlike med merilniki bistveno izrazitejše.

V povprečju le 22 % tistih gesel, ki jih naša rešitev klasificira kot močna, enako oceni tudi katerikoli drug merilnik. To nakazuje na precejšnje odstopanje v kriterijih za najvišjo stopnjo varnosti. V tej zvezi ponovno izstopata merilnika PasswordMonster in Zxcvbn, saj imata z našim merilnikom najvišjo stopnjo ujemanja pri ocenjevanju močnih gesel – nad 40 %. Najnižjo skladnost pa beležimo pri orodjih Nordpass (5 %) in Kaspersky (4 %), kar potrjuje, da je njihova ocena varnosti v tem razredu bistveno bolj restriktivna oziroma temelji na drugačnih kriterijih.

Kljub obsežnim podatkom o stopnjah ujemanja med posameznimi orodji, ti rezultati še ne omogočajo zanesljive ocene učinkovitosti posameznega merilnika. Ujemanje z drugimi orodji ne pomeni nujno višje točnosti. Da bi ocenili dejansko zanesljivost razvrstitev, smo zato izvedli dodatno analizo, ki temelji na dejanski odpornosti gesel proti napadu. Za preverjanje dejanske odpornosti gesel smo izvedli simuliran napad z uporabo modela PCFG, treniranega na isti učni množici. Rezultati napada so podani v številu potrebnih poskusov za ugotovitev gesla. Glede na število poskusov smo gesla klasificirali v pet razredov, kot je prikazano v tabeli 4.3. Skozi napad smo kar 63 % gesel identificirali za šibke. Za močne smo prepoznali 7 % gesel. To so gesla, ki so potrebovala za zlom nad 100 milijonov poskusov. Napad prav tako ni odkril 4 % gesel, ki seveda prav tako veljajo za zelo močna.

Table 4: **Rezultati napada na gesla.**

zelo šibko	< 103	438	4,38 %
šibko	< 106	6262	62,62 %
dobro	< 108	1152	11,52 %
močno	< 1010	727	7,27 %
zelo močno	> = 1010	1018	10,18 %
neodkrito		403	4,03 %

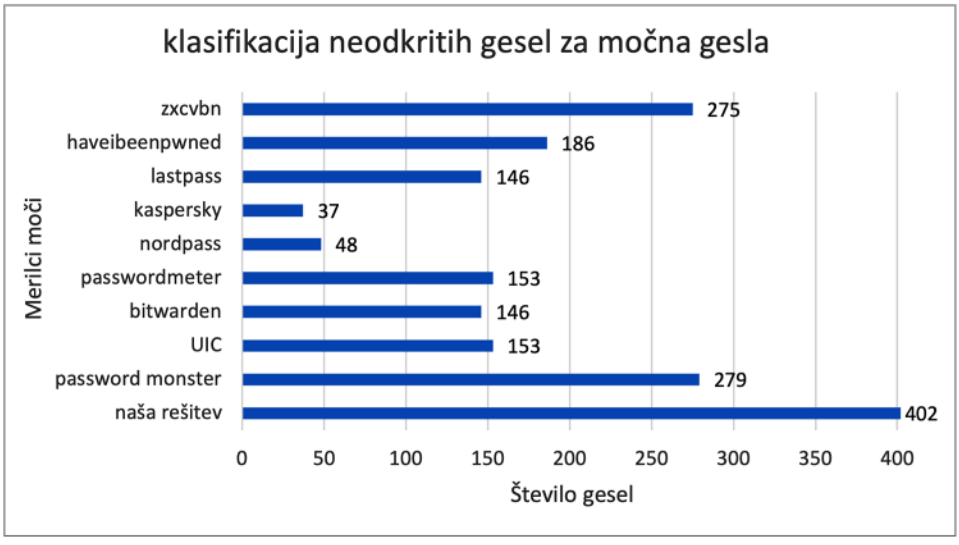


Figure 17: **Število neodkritih gesel v napadu označene za močne.**

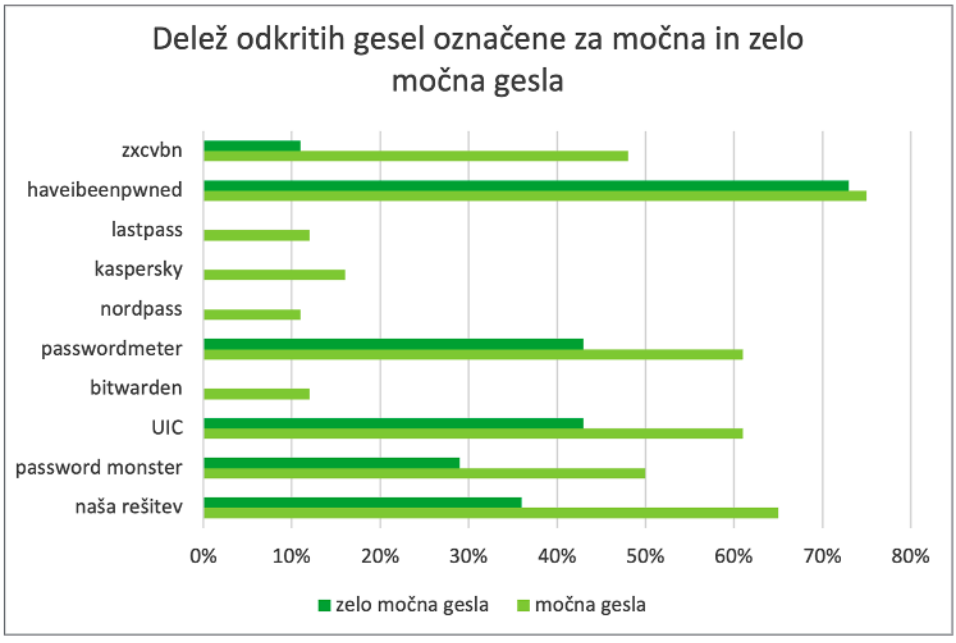


Figure 18: **Delež odkritih gesel označene za močna in zelo močna gesla.**

Rezultate napada smo primerjali z rezultati merilnikov. Sprva nas je zanimalo koliko neodkritih gesel so merilniki klasificirali za močna gesla in koliko močnih gesel (identificiranih iz strani merilnikov) je bilo odkritih v napadu. Od 403 neodkritih gesel naša rešitev le eno ni označila za močno. Naslednja merilnika, ki sta okoli 70 % neodkritih gesel označila za močna sta PasswordMonster in Zxcvbn. Najmanj neodkritih gesel sta za močne označila Nordpass in Kaspersky in sicer 10 %. Deleži drugih merilnikov se gibajo nekje okoli 35 %. Ti podatki so podrobneje predstavljeni na sliki 17. V pravilnosti klasifikacije neodkritih gesel naša rešitev prekaša vse druge.

Pri analizi ujemanja med ocenami moči gesel in njihovo dejansko odpornostjo proti napadu smo zaznali, da je naša rešitev kljub sicer visoki natančnosti pri klasifikaciji šibkih gesel nekoliko manj uspešna pri razpoznavanju najvarnejših gesel. Kar 65 % gesel, ki jih je naš model ocenil kot močna, je bilo uspešno razbitih z napadom s PCFG. Če pa analiziramo zgolj podskupino zelo močnih gesel, ta delež pade na 35 %, kar nakazuje, da je razvrščanje v najvišji razred bolj dosledno, vendar še vedno ne optimalno. V obeh primerih gre za najvišji delež razbitih močnih gesel med vsemi preizkušenimi merilniki, kar kaže na možnost precenjevanja varnosti določenih gesel znotraj našega modela.

Nasprotno pa so se merilniki Nordpass, Kaspersky in Lastpass izkazali kot bolj previdni. Med gesli, ki so jih ta orodja označila kot močna, je bilo le 20 % dejansko razbitih – kar pomeni, da so njihova merila za klasifikacijo močnih gesel strožja in bolj usklajena z realno odpornostjo.

Ob upoštevanju petstopenjske razdelitve gesel, kjer smo analizirali tudi razred zelo močnih gesel, se razlike med orodji še dodatno poglobijo. Merilniki UIC in PasswordMeter so imeli najvišji delež razbitih gesel v tem najvišjem varnostnem razredu – pri obeh je bilo kar 43 % zelo močnih gesel uspešno razbitih. Še slabše rezultate je pokazal Have I Been Pwned, saj je bilo kar 73 % gesel, ki so bila po tej storitvi označena kot zelo močna, dejansko prepoznanih v napadu.

Ti rezultati opozarjajo, da visoka ocena moči gesla s strani merilnika še ne pomeni nujno dejanske varnosti – zlasti če model temelji na statičnih ali heurističnih kriterijih brez preverjanja verjetnostne porazdelitve gesel v realnih podatkovnih bazah. Podrobne vrednosti za vse obravnavane primere so predstavljene na sliki 18, kjer so prikazani deleži razbitih gesel znotraj posameznih razredov glede na pripadajoči merilnik.

V nadaljevanju smo posebno pozornost namenili natančnosti klasifikacije močnih gesel pri posameznih merilnikih glede na rezultate dejanskega napada. Analiza je bila usmerjena v ugotavljanje, kolikšen delež gesel, ki jih je napad prepoznal kot močna, so ustrezno kot takšna prepoznali tudi merilniki. Re-

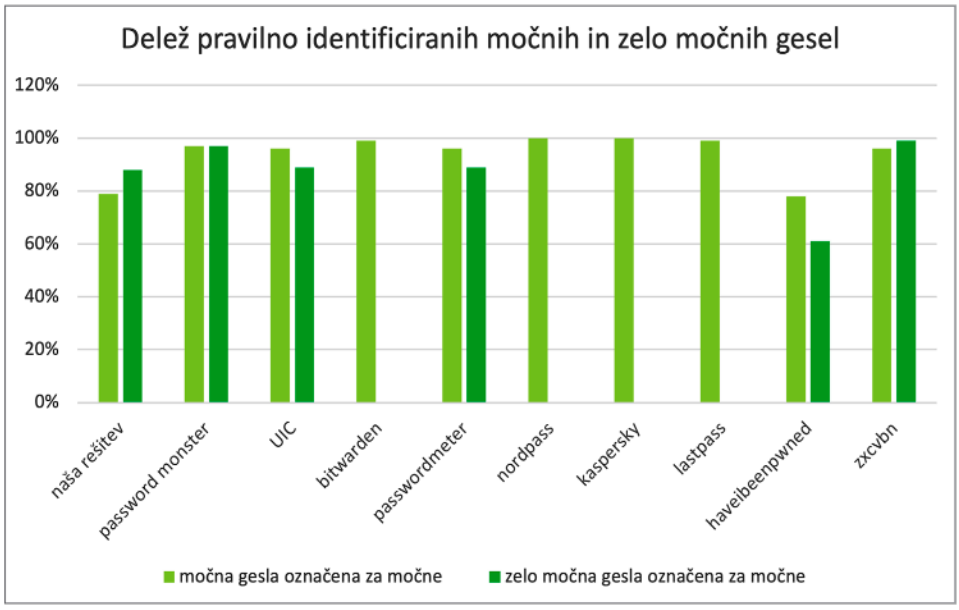


Figure 19: **Delež pravilno identificiranih močnih gesel.**

zultati kažejo, da tudi po tem kriteriju naša rešitev nekoliko zaostaja. Med vsemi gesli, ki so bila s strani napada ocenjena kot močna, jih je naša rešitev pravilno klasificirala 79 %, kar pomeni, da se 21 % močnih gesel po naši oceni ne bi uvrstilo v najvišji razred – kljub njihovi visoki odpornosti.

Vsi ostali merilniki, z izjemo Have I Been Pwned, so pri tem dosegli boljše rezultate. Vsak od njih je pravilno identificiral vsaj 96 % gesel, ki so bila po napadu ocenjena kot močna. Najbolje sta se izkazala merilnika Kaspersky in Nordpass, ki sta dosegla 100 % skladnost – torej sta vsa gesla, ki so se po rezultatih napada izkazala kot močna, tudi sama ustrezno označila.

Pri analizi klasifikacije zelo močnih gesel so se razlike med merilniki nekoliko zmanjšale. Naša rešitev je pri tem dosegla 88 % pravilno klasificiranih zelo močnih gesel, kar jo približa ostalim orodjem in kaže na večjo zanesljivost pri prepoznavanju najbolj odpornih gesel.

Natančni podatki po posameznem merilniku in razredu so prikazani na slki 19, ki omogoča neposredno primerjavo uspešnosti klasifikacije močnih in zelo močnih gesel med vsemi vključenimi rešitvami.

5 DISKUSIJA

S pomočjo preučene literature in izvedenih eksperimentov smo uspeli pridobiti naslednje odgovore na zastavljena raziskovalna vprašanja.

RV1: Kako se sodobne metode in orodja za preverjanje gesel soočajo in prilagajajo naprednim napadom na gesla?

Kljub izboljšanemu zavedanju glede varne uporabe gesel v primerjavi s stanjem pred desetletjem, še vedno ostajajo uporabniki najšibkejši člen v verigi informacijske varnosti. Z neprestano rastjo števila spletnih storitev in posledično večjim številom uporabniških računov se pričakuje, da bo problematika gesel v prihodnje še bolj izrazita. Raziskave s področja varnosti priporočajo uvedbo strožjih meril za preverjanje moči gesel – gesla, ki teh standardov ne izpolnjujejo, naj bodo zavrnjena že ob vnosu.

Raziskovalna skupnost vse bolj podpira razvoj metod za preverjanje moči gesel, ki temeljijo na verjetnostnih modelih (npr. markovske verige, PCFG) in umetnih nevronskih mrežah, saj te ponujajo višjo natančnost v primerjavi s klasičnimi heurističnimi pristopi. Pri naši primerjavi smo se osredotočili na merilnike moči gesel, ki so javno dostopni in enostavni za uporabo brez posebnih tehničnih zahtev, saj so ti orodja, s katerimi se srečujejo povprečni uporabniki. Zavedamo se, da obstajajo neposredno primerljivi verjetnostni pristopi, kot so OMEN, PCFG-cracker ter novejši modeli SMMl in uLMM, ki dosegajo visoko natančnost pri ocenjevanju. Vendar zaradi zahtevne namestitve in nižje uporabnosti za širšo skupnost teh rešitev nismo vključili v empirično primerjavo. Kljub temu menimo, da bi razširitev raziskave tudi

na ta orodja pomenila zanimivo smer za prihodnje delo, saj bi omogočila neposredno primerjavo različnih verjetnostnih pristopov in celovitejšo oceno njihove praktične uporabnosti.

Obenem se kaže trend v smeri razvoja prilagodljivih in kombiniranih rešitev, ki za preverjanje moči uporabljajo več metod hkrati in tako bolje pokrivajo različne vidike varnosti. Velik poudarek je tudi na uporabniški izkušnji, zlasti pri reševanju izzivov, povezanih z zapomnitvijo kompleksnih gesel. Orodja, kot so generatorji gesel in upravljalniki gesel, postajajo vse bolj razširjena.

Klasični spletni merilniki moči, ki uporabniku podajajo le vizualno povratno informacijo, pogosto ne dosegajo svojega namena – uporabniki takšne ocene ne jemljejo resno in se zadovoljijo z gesli, ki komaj dosegajo osnovne standarde. V ta namen so se pričele razvijati alternativne oblike povratne informacije, kot so družbeno primerjalni merilniki, ki ocenjujejo moč gesla v primerjavi z gesli drugih uporabnikov in tako ustvarjajo dodaten psihološki pritisk za izboljšanje izbire.

RV2: Kakšna je funkcionalnost razvite rešitve (računanje in klasifikacija moči testiranih gesel brez napak, hitrost računanja)?

V okviru testiranja je razvita rešitev uspešno analizirala 10.000 gesel, pri čemer je za vsako geslo pravilno izračunala oceno moči in ga razvrstila v ustrezen razred. Med testiranjem ni bilo zaznanih napak v delovanju. Poleg natančnosti nas je zanimal tudi časovni vidik, saj je odzivnost ključna pri uporabi spletnih orodij. Povprečni čas obdelave celotne množice gesel je znašal približno 9 sekund na geslo. Za gesla dolžine med 8 in 9 znaki, kar je najpogostejša dolžina uporabniških gesel, se je povprečni čas izračuna podaljšal na 13 sekund. Z višanjem dolžine gesla pa čas linearno narašča in lahko preseže 30 do 60 sekund, kar presega običajna pričakovanja uporabnikov.

Iz tega razloga menimo, da trenutna rešitev ni optimalno primerna za neposredno vključitev v spletna okolja, kjer se od ocenjevanja gesel pričakuje skoraj takojšen odziv. Časovna zahtevnost torej predstavlja eno od ključnih omejitev naše implementacije.

RV3: Kako se učinkovitost naše rešitve primerja z drugimi uveljavljenimi merilniki moči gesel na spletu?

Razvito rešitev smo primerjali z devetimi uveljavljenimi spletnimi merilniki moči gesel. V primerjavi s konkurenco je naša rešitev manj stroga pri ocenjevanju šibkih in močnih gesel, saj zazna manj šibkih in hkrati več močnih gesel kot večina drugih merilnikov. Ugotovili smo, da se ocene naše rešitve v povprečju ujemajo s 76 % ocen drugih merilnikov. Največje ujemanje je bilo zaznano pri šibkih geslih, kjer je 97 % ocen enakih, medtem ko se pri močnih geslih ocene ujemajo le v 4 % primerov.

Primerjava z rezultati simuliranega napada je pokazala, da naša rešitev pravilno klasificira 99 % neodkritih gesel kot močna, kar jo uvršča med najzanesljivejše glede na ta vidik. Vendar pa v nadaljnjih primerjavah zaostaja za ostalimi, saj je 65 % gesel, ki jih je označila kot močna, bilo uspešno razbitih. To je eden najvišjih deležev med primerjanimi orodji. Poleg tega je zaznano tudi razmeroma visoko število šibkih gesel, ki so bila napačno označena kot močna.

Po drugi strani je bila uspešnost naše rešitve pri nepravilnem označevanju močnih gesel za šibka boljša – napačno je bilo klasificiranih le 10 % takšnih primerov, kar je ugodnejši rezultat v primerjavi z večino drugih orodij. Pri natančnosti ocenjevanja močnih gesel pa naša rešitev zopet nekoliko zaostaja – le 79 % gesel, ki jih je označila za močna, je bilo v napadu dejansko potrjenih kot takšnih. Vsi ostali merilniki so v tem pogledu presegli prag 95 % natančnosti.

Najboljše rezultate so v celotni primerjavi dosegli merilniki Nordpass, Kaspersky in Lastpass. Ti so uspešno klasificirali največ gesel pravilno ter označili najmanj šibkih gesel za močna. Napad je razkril, da so ta tri orodja med vsemi najmanjkrat napačno ocenila šibka gesla kot močna (manj kot 20 %). Vendar pa imajo ta orodja tudi nasproten učinek – pogosto ne prepoznajo dejansko močnih gesel, kar pomeni, da njihovi kriteriji morda presegajo realne zahteve. Naša rešitev je bolj popustljiva, saj v najvišji razred uvrsti več gesel, vendar s tem nekoliko ogrozi zanesljivost.

Pri interpretaciji rezultatov pa je treba biti previden, saj lahko premehki kriteriji razvrščanja povzročijo, da so šibka gesla ocenjena kot dobra ali celo močna. Takšne ocene so za uporabnika zavajajoče in lahko pomenijo resno varnostno tveganje v praksi.

Na podlagi teh rezultatov ocenjujemo, da trenutna rešitev ne predstavlja boljše alternative v primerjavi z najboljšimi obstoječimi merilniki – tako z vidika hitrosti kot z vidika natančnosti. Visoko število napačno ocenjenih močnih gesel lahko pripišemo omejitvam v markovskih modelih, predvsem pomanjkanju verodostojnih verjetnostnih porazdelitev

za nekatera zaporedja znakov. Potencialno izboljšavo predstavlja dodatno treniranje modelov ter vključitev drugih pristopov. Ker se končna ocena trenutno izračuna kot povprečje treh modelov, bi alternativna strategija – kot je uporaba najnižje posamezne ocene – morda vodila do večje previdnosti pri dodeljevanju najvišjih ocen.

Čeprav trenutna različica še ne predstavlja konkurenčne alternative, ima zaradi 80 % natančnosti pri prepoznavanju močnih gesel potencial, ki ga je mogoče nadgraditi. Nadaljnji razvoj bi lahko vključil nov markovski model, treniran na drugem jeziku (npr. nemščini), kar bi omogočilo lokalizirane rešitve, prilagojene posameznim uporabniškim skupinam.

Ključni prispevek našega dela je v inovativni uporabi verjetnostnega pristopa, temelječega na markovskih verigah, in njegovi praktični integraciji v okolje spletnega brskalnika. S tem smo pokazali, da je mogoče metode, ki so bile doslej večinoma obravnavane v raziskovalnem okolju, uspešno prenesti v uporabniški kontekst ter jih približati vsakodnevni rabi. Poleg tega smo izvedli obsežno primerjalno analizo z devetimi obstoječimi spletnimi merilniki moči gesel, kar omogoča celovitejše razumevanje razlik med heurističnimi in verjetnostnimi pristopi ter osvetljuje njihove prednosti in omejitve v praksi. Naše delo tako združuje aplikativni in raziskovalni vidik: prvi zagotavlja uporabniško relevantnost, drugi pa prispeva k nadaljnjemu razvoju metod za preverjanje moči gesel.

6 ZAKLJUČEK

Gesla že več kot pet desetletij ostajajo prevladujoča metoda overjanja, kar je posledica njihove enostavne implementacije, nizkih stroškov ter minimalnih tehničnih zahtev. Kljub razvoju naprednih pristopov overjanja ostaja največja ranljivost na področju gesel človeška malomarnost, kar ob naraščujočem številu spletnih storitev in uporabniških računov povečuje kompleksnost upravljanja gesel. Glede na trajnostno vlogo gesel v informacijski varnosti, tudi ob uvedbi novih mehanizmov overjanja, je nujno spodbujati oblikovanje močnejših gesel. To področje večajo strožji kriteriji za ustvarjanje gesel, vpeljava merilnikov moči gesel, vizualnih povratnih indikatorjev ter uporaba upraviteljev gesel.

V članku smo najprej pregledali obstoječo strokovno literaturo o metodah ocenjevanja moči gesel in se nato osredotočili na razvoj verjetnostnega ocenjevalnika moči, ki temelji na markovskih verigah. Učinkovitost in funkcionalnost razvite rešitve smo ovrednotili z eksperimentalnim preizkusom na testni zbirki gesel ter primerjavo z devetimi obstoječimi spletnimi ocenjevalci moči. Zaradi zmogljivostnih omejitev testne strojne opreme smo uporabili omejeno zbirko gesel, sestavljeno iz znakovnega nabora 95 ASCII znakov v angleško govorečem okolju. Sistem je neprekinjeno izračunal moč in razred za vsako geslo, vendar je povprečni čas obdelave enega gesla znašal približno 13 sekund, kar presega sprejemljive čakalne dobe v spletnem okolju.

Pri primerjavi z drugimi spletnimi merilniki smo ugotovili, da naša rešitev ne ponuja bistvene prednosti: poleg počasnejše obdelave je tudi ocenjevalna politika premalo stroga pri prepoznavanju močnih gesel. Kljub temu smo zaznali, da ima razvita metoda s svojimi markovskimi modeli potencial za visok nivo natančnosti pri identifikaciji močnih gesel. Z nadaljnjim treniranjem obstoječih modelov ali vključitvijo dodatnih, na primer jezikovno specifičnih modelov (npr. za nemško jezikovno območje), bi lahko izboljšali konkurenčnost naše rešitve v primerjavi z uveljavljenimi merilniki.

LITERATURA

[1] J. Galbally, I. Coisel, and I. Sanchez, »A probabilistic framework for improved password strength metrics,« in *2014 international carnahan conference on security technology (ICCST)*, 2014, pp. 1–6. doi: 10.1109/CCST.2014.6986985.

[2] B. Ye, Y. Guo, L. Zhang, and X. Guo, »An empirical study of mnemonic password creation tips,« *Computers & Security*, vol. 85, pp. 41–50, 2019, doi: 10.1016/j.cose.2019.04.009.

[3] D. Pereira, J. F. Ferreira, and A. Mendes, »Evaluating the accuracy of password strength meters using off-the-shelf guessing attacks,« in *2020 IEEE international symposium on software reliability engineering workshops (ISSREW)*, 2020, pp. 237–242. doi: 10.1109/ISSREW51248.2020.00079.

[4] J. Galbally, I. Coisel, and I. Sanchez, »A new multimodal approach for password strength estimation—part i: Theory and algorithms,« *IEEE Transactions on Information Forensics and Security*, vol. 12, no. 12, pp. 2829–2844, 2017, doi: 10.1109/TIFS.2016.2636092.

[5] M. M. Taha, T. A. Alhaj, A. E. Moktar, A. H. Salim, and S. M. Abdullah, »On password strength measurements: Password entropy and password quality,« in *2013 international conference on computing, electrical and electronic engineering (ICCEEE)*, 2013, pp. 497–501. doi: 10.1109/ICCEEE.2013.6633989.

[6] Q. Dong, C. Jia, F. Duan, and D. Wang, »RLS-PSM: A robust and accurate password strength meter based on reuse, leet and separation,« *IEEE Transactions on Information Forensics and Security*, vol. 16, pp. 4988–5002, 2021, doi: 10.1109/TIFS.2021.3107147.

[7] M. Dupuis and F. Khan, »Effects of peer feedback on password strength,« in *2018 APWG symposium on electronic*

crime research (eCrime), 2018, pp. 1–9. doi: 10.1109/ECRIME.2018.8376210.
[8] R. Dubey and M. V. Martin, »Fool me once: A study of password selection evolution over the past decade,« in *2021 18th international conference on privacy, security and trust (PST)*, 2021, pp. 1–7. doi: 10.1109/PST52912.2021.9647823.
[9] I. Mannuela, J. Putri, Michael, and M. S. Anggreainy, »Level of password vulnerability,« in *2021 1st international conference on computer science and artificial intelligence (ICCSAI)*, 2021, pp. 351–354. doi: 10.1109/ICCSAI53272.2021.9609778.
[10] C. Brill and A. Olmsted, »Password reinforcement leveraging social media,« in *2016 international conference on information society (i-society)*, 2016, pp. 135–136. doi: 10.1109/i-Society.2016.7854197.
[11] V. Taneski, M. Heričko, and B. Brumen, »Analysing real students' passwords and students' passwords characteristics received from a questionnaire,« in *2016 39th international convention on information and communication technology, electronics and microelectronics (MIPRO)*, 2016, pp. 1436–1441. doi: 10.1109/MIPRO.2016.7522365.
[12] Google, »Create your google account.« [Online]. Available: https://accounts.google.com/lifecycle/steps/signup/name
[13] Dropbox, »Login – dropbox.« [Online]. Available: https://www.dropbox.com/register
[14] B. Ur *et al.*, »How does your password measure up? The effect of strength meters on password creation,« in *Proceedings of the 21st USENIX security symposium*, USENIX Association, 2012, p. 5. Available: https://www.usenix.org/conference/usenixsecurity12/technical-sessions/presentation/ur
[15] S. Egelman, A. Sotirakopoulos, I. Muslukhov, K. Beznosov, and C. Herley, »Does my password go up to eleven? The impact of password meters on password selection,« in *Proceedings of the SIGCHI conference on human factors in computing systems*, in CHI '13. New York, NY, USA: ACM, 2013, pp. 2379–2388. doi: 10.1145/2470654.2481329.
[16] A. Vance, D. Eargle, K. Ouimet, and D. Straub, »Enhancing password security through interactive fear appeals: A web-based field experiment,« in *2013 46th hawaii international conference on system sciences*, 2013, pp. 2988–2997. doi: 10.1109/HICSS.2013.196.
[17] D. He, B. Zhou, X. Yang, S. Chan, Y. Cheng, and N. Guiana, »Group password strength meter based on attention mechanism,« *IEEE Network*, vol. 34, no. 4, pp. 196–202, 2020, doi: 10.1109/MNET.001.1900482.
[18] D. Wang, D. He, H. Cheng, and P. Wang, »fuzzyPSM: A new password strength meter using fuzzy probabilistic context-free grammars,« in *2016 46th annual IEEE/IFIP international conference on dependable systems and networks (DSN)*, 2016, pp. 595–606. doi: 10.1109/DSN.2016.60.
[19] B. Pal, T. Daniel, R. Chatterjee, and T. Ristenpart, »Beyond credential stuffing: Password similarity models using neural networks,« in *2019 IEEE symposium on security and privacy (SP)*, 2019, pp. 417–434. doi: 10.1109/SP.2019.00056.
[20] X. Cui, C. Li, Y. Qin, and Y. Ding, »A password strength evaluation algorithm based on sensitive personal information,« in *2020 IEEE 19th international conference on trust, security and privacy in computing and communications (TrustCom)*, 2020, pp. 1542–1545. doi: 10.1109/TrustCom50675.2020.00211.
[21] PasswordMonster, »How secure is your password?« [Online]. Available: https://www.passwordmonster.com/
[22] U. of Illinois Chicago, »Password strength test.« [Online]. Available: https://www.uic.edu/apps/strong-password/
[23] Bitwarden, »Password strength testing tool.« [Online]. Available: https://bitwarden.com/password-strength/
[24] PasswordMeter, »How secure is your password?« [Online]. Available: https://passwordmeter.com/
[25] NordPass, »How secure is my password?« [Online]. Available: https://nordpass.com/secure-password/
[26] Kaspersky, »How secure is my password?« [Online]. Available: https://password.kaspersky.com/
[27] LastPass, »How secure is your password?« [Online]. Available: https://lastpass.com/howsecure.php
[28] T. Hunt, »Pwned websites.« [Online]. Available: https://haveibeenpwned.com/PwnedWebsites
[29] Dropbox, »Low-budget password strength estimation (zxcvbn).« [Online]. Available: https://github.com/dropbox/zxcvbn

▪

Viktor Taneski je asistent na Fakulteti za elektrotehniko, računalništvo in informatiko Univerze v Mariboru, kjer je zaposlen od leta 2016. Doktoriral je iz računalništva na Univerzi v Mariboru, kjer je opravil tudi dodiplomski. Njegovo raziskovalno delo je usmerjeno v metode overjanja, preizkušanje varnosti sistemov (penetracijsko testiranje), kibernetsko varnost in spletni razvoj. Sodeloval je pri več evropskih raziskovalnih projektih, med drugim CyberSec4Europe in Smart Villages, ter pomembno prispeval k razvoju mobilnih in spletnih aplikacij. Poleg tega je organiziral in sodeloval na delavnicah, usposabljanjih in znanstvenih konferencah s področja kibernetske varnosti, kot sta OTS in SQAMIA. Kot recenzent je ocenjeval različne izobraževalne materiale, povezane z razvojem programske opreme in varnostjo. Z bogatim znanjem s področij programiranja, modeliranja sistemov in varnostnih praks aktivno prispeva k prenosu znanja skozi poučevanje in raziskovanje.

▪

Raziskave **Marka Hölbla** na področju kibernetske varnosti in zasebnosti zajemajo širok spekter tem, od kriptografije do uporabniških vidikov informacijske varnosti in zasebnosti. Trenutno je prodekan za raziskovalno dejavnost na Fakulteti za elektrotehniko, računalništvo in informatiko Univerze v Mariboru (FERI UM), aktiven član in generalni sekretar CEPIS LSI, član delovne skupine ECSO (WG6), član izvršnega odbora sekcije slovenskega društva Informatika ter član Sekcije za kibernetsko varnost pri Gospodarski zbornici Slovenije.