УДК 004.056.53 DOI https://doi.org/10.35546/kntu2078-4481.2025.1.2.31

Ю. І. ТОВКУН
аспірантка кафедри автоматизації та проектування обчислювальної техніки
Харківський національний університет радіоелектроніки
ORCID: 0009-0000-5916-2897

АНАЛІЗ СТІЙКОСТІ ПАРОЛІВ З ВИКОРИСТАННЯМ AI-ГЕНЕРОВАНИХ АТАК

Паролі залишаються одним із ключових засобів автентифікації, проте їхня безпека значною мірою залежить від методів збереження та здатності протидіяти сучасним атакам. Традиційні методи злому, такі як brute force та словникові атаки, поступово втрачають ефективність перед більш досконалими технологіями, що використовують штучний інтелект. У цьому дослідженні проаналізовано вплив AI-генерованих атак на стійкість паролів та оцінено їхню ефективність порівняно з класичними методами.

Розглянуто сучасні підходи до генерації паролів у реальних системах автентифікації та механізми їх злому. Особливу увагу приділено аналізу моделі PassGAN, яка використовує генеративно-змагальні мережі для створення паролів із високою ймовірністю відповідності реальним користувацьким комбінаціям. Оцінено ефективність таких атак та їхню здатність обходити стандартні механізми перевірки складності паролів. Проведено експериментальне дослідження, що включає порівняння результативності AI-генерованих атак із традиційними методами, такими як brute force та словникові атаки.

Результати експерименту свідчать про значне підвищення ефективності злому паролів середньої складності при використанні AI-генерованих атак. Водночас складні паролі демонструють значно вищу стійкість до подібних атак. На основі отриманих даних сформульовано рекомендації щодо підвищення рівня безпеки паролів, зокрема використання довших комбінацій та впровадження систем цифрового імунітету, здатних виявляти та запобігати AI-орієнтованим загрозам.

Дослідження підкреслює важливість адаптивного підходу до безпеки автентифікації та необхідність впровадження комплексних рішень, що поєднують стійкі паролі, багатофакторну автентифікацію та технології штучного інтелекту для захисту цифрових систем. Подальший розвиток методів атак на основі AI потребує постійного вдосконалення засобів їх виявлення та нейтралізації.

Ключові слова: *стійкість паролів, злом паролів, AI-генеровані атаки, нейромережі, PassGAN, автентифікація, інформаційна безпека, цифровий імунітет, кібератаки.*

YU. I. TOVKUN
Postgraduate Student at the Department of Automation and Computer Engineering
Kharkiv National University of Radio Electronics
ORCID: 0009-0000-5916-2897

ANALYSIS OF PASSWORD RESISTANCE TO AI-GENERATED ATTACKS

Passwords remain one of the key authentication methods; however, their security largely depends on storage methods and the ability to counter modern attacks. Traditional password-cracking techniques, such as brute force and dictionary attacks, are gradually losing effectiveness against more advanced technologies that leverage artificial intelligence. This study analyzes the impact of AI-generated attacks on password resilience and evaluates their effectiveness compared to classical methods.

Modern approaches to password generation in real authentication systems and password-cracking mechanisms are examined. Special attention is given to the PassGAN model, which utilizes generative adversarial networks to create passwords with a high probability of matching real user combinations. The effectiveness of such attacks and their ability to bypass standard password complexity verification mechanisms are assessed. An experimental study is conducted, comparing the success rates of AI-generated attacks with traditional methods, such as brute force and dictionary attacks.

The experimental results indicate a significant increase in password-cracking efficiency for medium-complexity passwords when using AI-generated attacks. Meanwhile, complex passwords demonstrate significantly higher resistance to such attacks. Based on the obtained data, recommendations are formulated to enhance password security, including the use of longer combinations and the implementation of digital immunity systems capable of detecting and preventing AI-driven threats.

The study highlights the importance of an adaptive approach to authentication security and the need for comprehensive solutions that combine strong passwords, multi-factor authentication, and artificial intelligence technologies to protect digital systems. Further development of AI-based attack methods requires continuous improvement of detection and mitigation techniques.

Key words: *password resistance, password cracking, AI-generated attacks, neural networks, PassGAN, authentication, information security, digital immunity, cyberattacks.*

Постановка проблеми

Сучасні системи автентифікації переважно базуються на використанні паролів, що залишаються основним засобом захисту конфіденційної інформації. Однак, значна частина користувачів дотримується слабких політик створення паролів, що підвищує ймовірність їх компрометації. З розвитком методів машинного навчання та генеративних моделей штучного інтелекту (ШІ) з'явилася можливість здійснювати високоефективні атаки на паролі, використовуючи алгоритми, що генерують ймовірні комбінації на основі аналізу реальних баз даних скомпрометованих паролів.

Одним із найперспективніших підходів до злому паролів є використання генеративно-змагальних нейромереж (GAN), які дозволяють створювати паролі, що імітують стилістику реальних користувацьких комбінацій. Це значно підвищує ефективність атак порівняно з класичними методами, такими як brute force або словниковий підбір. У зв'язку з цим виникає потреба у детальному аналізі ефективності AI-генерованих атак, визначенні їх потенційної загрози та розробці заходів протидії.

Актуальність проблеми обумовлена зростанням кількості витоків облікових даних, що використовуються зловмисниками для здійснення атак типу credential stuffing, а також постійним удосконаленням нейромережевих методів. Недостатня увага до розвитку засобів цифрового імунітету, здатних адаптивно виявляти і блокувати AI-орієнтовані атаки, створює додаткові ризики для безпеки інформаційних систем. Таким чином, постає питання не лише аналізу ефективності нових атакувальних технологій, але й пошуку ефективних методів їхньої нейтралізації.

Аналіз останніх досліджень і публікацій

У сфері дослідження атак типу brute force та методів їх виявлення було проведено низку значущих робіт. Зокрема, Sharma, Tanwar та Kukreja (2024) у своїй праці «Implementation of Brute Force Attack» надали детальний аналіз різних типів brute force атак та продемонстрували їх реалізацію, що дозволяє глибше зрозуміти механізми цих атак та їх потенційний вплив на системи автентифікації [1].

Зі зростанням складності атак виникла потреба у вдосконалених методах їх виявлення. Sharma, Babbar та Vats (2024) у дослідженні «Empowering Security: Machine Learning Solutions for Detecting Brute Force Attacks» запропонували використання методів машинного навчання для виявлення brute force атак. Вони розробили модель, яка аналізує мережевий трафік та ідентифікує аномалії, характерні для таких атак, що підвищує ефективність виявлення та реагування на загрози [2].

Додатково, Hamza та Al-Janabi (2024) у своїй роботі «Detecting Brute Force Attacks Using Machine Learning» дослідили застосування різних класифікаторів машинного навчання для ідентифікації brute force атак на протоколи SSH та FTP. Вони виявили, що алгоритм Random Forest досягає найвищої точності в 99,9 % при виявленні таких атак, що підкреслює потенціал машинного навчання у сфері кібербезпеки [3].

Ці дослідження підкреслюють важливість розвитку та впровадження методів машинного навчання для виявлення та запобігання brute force атак. Однак, з огляду на швидкий розвиток технологій штучного інтелекту, зокрема генеративних моделей, виникає потреба у подальших дослідженнях, спрямованих на аналіз стійкості паролів до AI-генерованих атак та розробку ефективних заходів протидії.

Формулювання мети дослідження

Метою даного дослідження є аналіз стійкості паролів до атак, що використовують методи штучного інтелекту, зокрема генеративно-змагальні нейромережі (GAN). Дослідження спрямоване на оцінку ефективності AI-генерованих атак у порівнянні з традиційними методами злому, такими як brute force та словникові атаки, а також на виявлення ключових чинників, що впливають на зламостійкість паролів. Особлива увага приділяється аналізу принципів роботи генеративних моделей, їхньої здатності прогнозувати ймовірні комбінації на основі реальних витоків даних та оцінці їхнього впливу на сучасні системи автентифікації.

Практична частина дослідження передбачає тестування стійкості різних типів паролів до нейромережевих атак, порівняння швидкості та точності AI-генерованих атак із класичними методами, а також формулювання рекомендацій щодо підвищення безпеки паролів. Отримані результати дозволять глибше зрозуміти ризики, пов'язані із застосуванням штучного інтелекту у сфері атак на паролі, та сприятимуть розробці ефективних механізмів цифрового імунітету для захисту автентифікаційних систем.

Викладення основного матеріалу дослідження

1. Теоретичний огляд методів атак на паролі

Паролі є ключовим елементом автентифікації користувачів у цифрових системах. Проте їхня безпека значною мірою залежить від складності та унікальності комбінацій. Традиційні методи атак включають brute force (перебір усіх можливих варіантів), словникові атаки (підбір на основі баз даних часто використовуваних паролів) і атаки за витоками даних (credential stuffing) [4].

Brute force атака є найменш ефективною через експоненційне зростання кількості можливих комбінацій із збільшенням довжини пароля [5]. Наприклад, для пароля довжиною 6 символів, що складається з літер і цифр, кількість можливих комбінацій становить 36^6 ≈ 2,2 мільйони, тоді як для 12-символьного пароля ця кількість зростає до понад 4 трильйонів.

Словникові атаки є більш ефективними, оскільки вони використовують попередньо скомпільовані списки найпоширеніших паролів, доповнені популярними варіаціями, такими як «password123», «admin2024», «Qwerty!@#». Використання баз витоків, наприклад, RockYou, значно підвищує ефективність атак цього типу [6–7].

Розвиток штучного інтелекту призвів до появи генеративних атак на паролі, зокрема з використанням генеративно-змагальних нейромереж (GAN). Однією з найвідоміших реалізацій є PassGAN, яка навчається на великих витоках паролів і генерує нові комбінації, що мають високу ймовірність збігу з реальними користувацькими паролями. На відміну від класичних словникових атак, GAN-моделі можуть прогнозувати схожі на справжні паролі навіть без їхньої присутності у відкритих базах даних [8–9].

AI-генеровані атаки дозволяють генерувати паролі, що відповідають реальним стилям користувачів; оптимізувати підбір комбінацій, зменшуючи необхідність перебору; виявляти закономірності у використанні паролів та прогнозувати потенційні комбінації.

2. Експериментальне дослідження

Для оцінки ефективності AI-генерованих атак проведено експеримент із використанням PassGAN та порівнянням його результатів із класичними методами brute force і словникових атак. Було сформовано тестову вибірку паролів різної складності:

1. Слабкі паролі (4–6 символів, прості комбінації): «123456», «password», «qwerty».
2. Середньої складності (8–10 символів, поєднання літер та цифр): «Pa$$word2024», «Qw3rty99».
3. Складні паролі (12+ символів, символи різних регістрів, спеціальні знаки): «xY8@!zP9mW#5».

До параметру експерименту входить база навчання для нейромережі, а саме витоки паролів (RockYou dataset); тестова база – 10 000 випадково згенерованих паролів; інструменти, такі як PassGAN, John the Ripper, Hashcat та оцінювані параметри – швидкість підбору паролів, точність атак.

Результати тестування показали, що PassGAN перевершує словникові атаки у швидкості та точності підбору паролів.

Таблиця 1

Результат експериментального дослідження

Метод атаки	Час на злам слабкого пароля	Час на злам середнього пароля	Час на злам складного пароля
Brute force	<1 сек	15 хв	100+ годин
Словникова атака	<1 сек	7 хв	50+ годин
PassGAN	<1 сек	3 хв	30 годин

Аналіз показав, що AI-генеровані атаки ефективніші за традиційні методи, особливо щодо паролів середньої складності. Для слабких паролів всі методи показали приблизно однакову ефективність, тоді як у випадку складних паролів PassGAN значно зменшував час підбору завдяки можливості прогнозування схожих варіантів.

3. Рекомендації щодо підвищення безпеки паролів

На основі отриманих результатів можна зробити висновок, що традиційні методи перевірки складності паролів потребують удосконалення для врахування можливостей нейромережевих атак. Одним із перспективних напрямів є створення адаптивних цифрових імунних систем, які виявляють аномалії в автентифікації та блокують підозрілі спроби входу.

Ефективними заходами підвищення безпеки є варіанти зазначені нижче.

1. Використання довгих і складних паролів. Паролі довжиною не менше 12 символів із використанням великих і малих літер, цифр і спеціальних символів значно ускладнюють нейромережевий злам.

2. Багатофакторна автентифікація (2FA). Додатковий рівень безпеки у вигляді SMS-коду, біометрії або апаратного ключа (наприклад, YubiKey) унеможливлює доступ навіть у разі компрометації пароля.

3. Менеджери паролів. Використання інструментів, таких як Bitwarden або 1Password, дозволяє генерувати та зберігати складні унікальні паролі без ризику запам’ятовування або повторного використання.

4. Моніторинг витоків даних. Періодична перевірка власних облікових даних через сервіси, такі як Have I Been Pwned, дозволяє вчасно змінювати паролі у разі компрометації.

Загалом, дослідження підтверджує, що AI-генеровані атаки становлять реальну загрозу для класичних паролів, і необхідні нові стратегії захисту, які поєднують сильні паролі, багатофакторну автентифікацію та адаптивні системи виявлення загроз.

Висновки

Дослідження показало, що традиційні методи атак на паролі, такі як brute force і словникові атаки, поступово втрачають ефективність через впровадження сучасних політик безпеки. Однак розвиток нейромережевих технологій, зокрема генеративно-змагальних мереж (GAN), відкриває нові можливості для автоматизованого та високоефективного підбору паролів.

Результати експерименту підтвердили, що AI-генеровані атаки, представлені моделлю PassGAN, значно перевершують традиційні методи у швидкості та точності підбору паролів середньої складності (8–10 символів).

У випадку слабких паролів (<6 символів) всі методи атаки показали майже миттєвий результат, що свідчить про їхню вразливість. Для складних паролів (>12 символів) ефективність PassGAN суттєво зменшується, що підкреслює важливість створення довгих та унікальних паролів.

Отримані результати свідчать про необхідність вдосконалення механізмів захисту, зокрема інтеграції адаптивних систем цифрового імунітету, які можуть аналізувати аномалії в процесі автентифікації та блокувати підозрілі спроби входу. Крім того, широке впровадження багатофакторної автентифікації (2FA), менеджерів паролів та моніторингу витоків даних може значно знизити ризики компрометації облікових записів.

Таким чином, результати дослідження підтверджують, що поява AI-генерованих атак створює нові виклики для інформаційної безпеки, що вимагає перегляду стандартних підходів до управління паролями та розробки комплексних механізмів їхнього захисту.

Список використаної літератури

1. Sharma P., Tanwar S., Kukreja V. Implementation of Brute Force Attack [Електронний ресурс] // Taylor & Francis. 2024. Режим доступу: https://www.taylorfrancis.com/chapters/edit/10.1201/9781003535423-10/implementation-brute-force-attack-priyanshu-sharma-sarvesh-tanwar-vinay-kukreja

2. Sharma A., Babbar H., Vats A. Empowering Security: Machine Learning Solutions for Detecting Brute Force Attacks [Електронний ресурс] // IEEE Xplore. 2024. Режим доступу: https://ieeexplore.ieee.org/abstract/document/10838096

3. Hamza A. A., Al-Janabi R. J. S. Detecting Brute Force Attacks Using Machine Learning [Електронний ресурс] // BIO Web of Conferences. 2024. Режим доступу: https://www.bio-conferences.org/articles/bioconf/pdf/2024/16/bioconf_iscku2024_00045.pdf

4. Hussian N. A., Al-Shareefi F. A Comparative Study Between Two Cybersecurity Attacks: Brute Force and Dictionary Attack [Електронний ресурс] // Journal of Knowledge Management & Computing. 2018. № 11. Режим доступу: http://dx.doi.org/10.31642/JoKMC/2018/110216

5. Bošnjak L., Sreš J., Brumen B. Brute-force and dictionary attack on hashed real-world passwords [Електронний ресурс] // IEEE Xplore. 2018. Режим доступу: https://ieeexplore.ieee.org/abstract/document/8400211

6. Narayanan A., Shmatikov V. Fast dictionary attacks on passwords using time-space tradeoff [Електронний ресурс] // ACM Conference on Computer and Communications Security. 2005. Режим доступу: https://doi.org/10.1145/1102120.1102168

7. Pinkas B., Sander T. Securing passwords against dictionary attacks [Електронний ресурс] // ACM Conference on Computer and Communications Security. 2002. Режим доступу: https://dl.acm.org/doi/10.1145/586110.586133

8. Hitaj B., Gasti P., Ateniese G., Perez-Cruz F. PassGAN: A Deep Learning Approach for Password Guessing [Електронний ресурс] // arXiv.org. 2017. Режим доступу: https://doi.org/10.48550/arXiv.1709.00440

9. Werhahn M., Xie Y., Chu M., Thuerey N. A Multi-Pass GAN for Fluid Flow Super-Resolution [Електронний ресурс] // ACM Transactions on Graphics. 2019. Режим доступу: https://doi.org/10.1145/3340251

References

1. Sharma, P., Tanwar, S., & Kukreja, V. (2024). *Implementation of Brute Force Attack*. Taylor & Francis. Retrieved from https://www.taylorfrancis.com/chapters/edit/10.1201/9781003535423-10/implementation-brute-force-attack-priyanshu-sharma-sarvesh-tanwar-vinay-kukreja

2. Sharma, A., Babbar, H., & Vats, A. (2024). *Empowering security: Machine learning solutions for detecting brute force attacks*. IEEE Xplore. Retrieved from https://ieeexplore.ieee.org/abstract/document/10838096

3. Hamza, A.A., & Al-Janabi, R.J.S. (2024). *Detecting brute force attacks using machine learning*. BIO Web of Conferences. Retrieved from https://www.bio-conferences.org/articles/bioconf/pdf/2024/16/bioconf_iscku2024_00045.pdf

4. Hussian, N.A., & Al-Shareefi, F. (2018). *A comparative study between two cybersecurity attacks: Brute force and dictionary attack*. Journal of Knowledge Management & Computing, 11. Retrieved from http://dx.doi.org/10.31642/JoKMC/2018/110216

5. Bošnjak, L., Sreš, J., & Brumen, B. (2018). *Brute-force and dictionary attack on hashed real-world passwords.* IEEE Xplore. Retrieved from https://ieeexplore.ieee.org/abstract/document/8400211

6. Narayanan, A., & Shmatikov, V. (2005). *Fast dictionary attacks on passwords using time-space tradeoff*. ACM Conference on Computer and Communications Security. Retrieved from https://doi.org/10.1145/1102120.1102168

7. Pinkas, B., & Sander, T. (2002). *Securing passwords against dictionary attacks*. ACM Conference on Computer and Communications Security. Retrieved from https://dl.acm.org/doi/10.1145/586110.586133

8. Hitaj, B., Gasti, P., Ateniese, G., & Perez-Cruz, F. (2017). *PassGAN: A deep learning approach for password guessing*. arXiv. Retrieved from https://doi.org/10.48550/arXiv.1709.00440

9. Werhahn, M., Xie, Y., Chu, M., & Thuerey, N. (2019). *A multi-pass GAN for fluid flow super-resolution*. ACM Transactions on Graphics. Retrieved from https://doi.org/10.1145/3340251