Theory and Practice of Science and Technology
2025, VOL. 6, NO. 8, 7-10
DOI: 10.47297/taposatWSP2633-456901.20250608

Research on Password-Cracking Methods Based on Machine Learning

Zhengyang Chen

Technique University of Denmark, Lyngby, 2800, Denmark

ABSTRACT
Password Security is the first line of defence in the digital world. Traditional password defence relies on preset rules. Attackers use massive leaked data to train models to mine the Semantic Association rules of human passwords deeply; the automatic generation of highly personalised obfuscation rules not only greatly improves the cracking efficiency but also significantly enhances the concealment and adaptability of attacks, making the traditional protection methods based on the blacklist and rule-based increasingly ineffective. Therefore, this paper discusses the technical path of machine learning in password cracking, aiming to provide a reference for constructing a dynamic defence framework and developing a human-computer collaborative password design paradigm.

KEYWORDS
Machine learning; Password cracking; Semantic association; Intelligent defense

1 Introduction

Today, with the acceleration of the digital process, passwords are still the core means to protect personal privacy and enterprise data. However, traditional passwords have been difficult to cope with with the enhancement of computing power and the explosive growth of data resources. Moreover, modern password-cracking systems based on machine learning tap into the laws of human password setting. Such systems have the ability to interact with the environment and can dynamically adjust attack parameters according to real-time signals such as response delay and error locking strategy of the target system, and actively evade risk control detection of behaviour; this "More intelligent attack" makes the defence side into a passive situation. In this context, a deep understanding of the internal logic of machine learning reshaping cryptographic attacks is strategically significant for building the next generation of defence systems. Only by accurately grasping the technological evolution path, capability boundary and evolution law of the attacker can a forward-looking, dynamic protection scheme be designed. Therefore, this paper will analyze the operation paradigm and evolutionary power of intelligent password cracking and propose a strategy and optimization scheme driven by machine learning to provide a methodological reference for reversing the imbalance between attack and defence.

2 The application and technical implementation of machine learning in password cracking

2.1 The basic principles of machine learning in password cracking

The basic principle of a machine learning algorithm in password analysis: by training many password sample data and building a model to realize the prediction and analysis of unknown passwords, the commonly used machine learning algorithms include support vector machine, decision tree, random forest and so on. Machine learning can process large numbers of leaked and generated password samples and gradually discover common features that makeup passwords. The support vector machine algorithm automatically analyzes the mixed patterns of letters and numbers in the data and common word deformation patterns and converts them into mathematical feature vectors; the classification model is established by calculating the weight relationship between different features [1]. The decision tree algorithm can simulate the human decision logic to judge the cryptographic features according to their importance and then convert the human behaviour preference into computable prediction rules. The ensemble algorithms, such as random forest, work in parallel through multiple sets of decision-making units to comprehensively evaluate the elements such as character diversity, adjacent character correlation, keyboard spatial position relationship, and output confidence scores for each candidate password. In the process, the system automatically excludes low-probability combinations, and the samples that fail to predict will be fed back to the training system to readjust the feature weight and optimize the decision boundary, thus significantly improving the efficiency of password cracking.

2.2 The application of generative adversarial networks in password cracking

The generator in generative adversarial networks can mimic the password pattern created by humans and randomly generate character combinations. At the same time, the discriminator can distinguish whether the input content is a real

password sample or an imitation created by the generator. When the discriminator successfully identifies the fake password, the generator will adjust its internal parameters accordingly to make the subsequently generated password closer to human behavioural characteristics, and it also promotes the generator to gradually master the length distribution of the password, the common character replacement mode, and the insertion preference of the digital character. After sufficient training, the generator can efficiently output candidate sequences that conform to the statistical characteristics of real-world ciphers and form a powerful cipher probability model. In the actual cracking scenario, the generator will continuously output high-probability password combinations, and the discriminator will quickly evaluate the authenticity probability of the candidate passwords, guiding the system to skip the inefficient areas that do not conform to human habits. When the target password is not successfully matched, the feedback signal will drive the generator to adjust the generation strategy in real time, analyze the cause of failure, and automatically increase the frequency of attempts for certain special symbol combinations, such as changing the character set, adjusting the length, and trying different modes to generate new candidate passwords, to continuously approach the target until the successful termination condition, to improve the cracking efficiency of complex passwords.

3 Machine learning-driven password cracking strategy and optimization

3.1 Combination of rule-based password cracking and machine learning

Traditional password cracking methods rely on preset dictionary files and simple character replacement modes, which are inefficient and difficult to deal with modern complex passwords. Machine learning can transform fixed rules into dynamically evolving intelligent generation systems that capture the deformation of date formats and popular words that users prefer to use in specific cultural contexts; then, these implicit settings are transformed into dynamic rules to generate candidate password sequences closer to the target user group in real-time. When the cracking system identifies that the target account belongs to a specific industry and age group, it will preferentially output password variants that match the common thinking patterns of the group and predict the character combination that the user is most likely to use [2]. Machine learning models evolve with each crack attempt. Successful passwords are quickly assimilated into new rules and failed attempts help the system eliminate invalid paths and continuously refine the attack. In practice, machine learning is an intelligent commander of password-cracking tactics, scheduling various attack strategies and optimizing resource allocation in real time. Firstly, the system deeply analyzes the characteristics of the target cryptographic system, such as the type of encryption algorithm, the error attempt feedback mechanism, and the log-in frequency restriction rule, and allocates priority and computing resources for different policies based on these characteristics. If some rule combinations fail in succession, resources are dynamically allocated to other policies, showing signs of breaking through. When the logon frequency limit is suddenly tightened, it will automatically switch to a more stealthy, decentralized low-speed penetration mode, significantly improving resilience and password-breaking efficiency in complex security environments.

3.2 Generation and optimization of password obfuscation rules

Traditional password obfuscation uses preset templates for simple transformations, such as replacing the letter"A" with"@" or adding a fixed suffix, which modern defences can easily detect. Machine learning can deeply analyze massive real password databases and leaked data, autonomously discover the hidden semantic association and deformation logic when humans set passwords, and summarize the confusion preferences in different scenarios. When aiming at a specific target, the system will intelligently integrate the target user's professional characteristics, the social platform's language habits, and the keywords of recent hot events to reconstruct a confusion rule set with intense personal colour. Efficient password cracking requires real-time interaction between obfuscation rules and the environment, and machine learning monitors multiple feedback signals simultaneously in each cracking attempt. When the target system is detected as sensitive to special character combinations, it automatically reduces the symbol replacement ratio and uses more subtle obfuscation methods such as case alternation. Suppose the target is found to have a fast response mechanism for long character strings. In that case, it automatically switches to a segmented injection mode and breaks a single complex password into multiple exploratory inputs. Each failed crack attempt will strengthen the matching degree between the rules and the environment, and the successful fragments will be automatically disassembled into effective rule factors and integrated into the generation pool. Finally, password obfuscation can dynamically adapt and optimize the attack strategy [3].

4 The role of machine learning in password defence and security enhancement

4.1 Dynamic threat identification and real-time defence enhancement

The machine learning model can observe the data flow generated by the user's log-in activity around the clock and paint a picture of the normal use of each account. The screen includes the user's usual log-in times, frequently connected locations, keystroke speeds and pause rhythms, and is on high alert for any anomalies that do not fit the picture. When the same account tries to log in from far-flung countries in a matter of minutes, or the rhythm of password entry becomes mechanical, machine learning models can pick up on that signal and discern the type of attack, determine whether the automated program is a brute-force test or whether someone is performing a targeted hack to alert the relevant personnel before the attacker touches the target. Once the abnormal activity is confirmed, the machine learning model automatically activates a multi-layer defence mechanism, which immediately increases the verification strength and limits the permissions of high-risk sessions [4]. When recent concentrated attacks on specific types of passwords are observed, machine learning models automatically enhance the protection level of such cryptographic structures and optimize the key parameters of encryption algorithms in advance. Each successful intercept is used to train the model to identify future threats more accurately, and the new evasion methods used by the attackers push the system to upgrade its defences so that the protection system has always been ahead of the development of attack technology. In the main time, Machine learning can detect the frequency of different kinds of attack so it can use special methods to prevent some frequently used attack methods.

4.2 Intelligent password strategy optimization and risk prediction

The traditional password strategy sets the same complexity and replacement cycle requirements for all users, which not only brings inconvenience to use but also has insufficient protection. Machine learning can deeply analyze the actual behaviour patterns and potential risk levels of user groups, dynamically adjust password strength requirements, and continuously observe how users set and manage passwords; by understanding the different habits and security environments of different user groups, the system can target password protection levels for these accounts. It can also identify common weaknesses in user password design, such as over-reliance on birthdays, to provide real-time, gentle suggestions for improvement as users create passwords and guide them to more complex passwords [5]. Machine learning also anticipates risks by looking at attacks as they occur and identifying emerging trends that have not yet exploded. When a specific type of password structure is frequently detected in recent transactions on the dark web, it will immediately issue a warning and accurately identify which accounts within the internal network are using passwords that may become targets of these emerging attack methods. Based on these predictions, protective measures will be deployed in advance. In addition, the machine learning system will actively scan the entire network for old passwords highly similar to predicted risk patterns, prompting users to change so they can be protected before the attacker acts.

4.3 Intelligent key management and adaptive encryption optimization

In password defence, machine learning systems can continuously analyze the secret key's actual use scenarios, access frequencies, and potential threat environments to dynamically adjust the life cycle and protection strength. It can be sensitive to which secret key is protecting particularly sensitive data, automatically shortening the validity of these secret keys. When a secret key usage pattern is found to be abnormal, such as being called at an unusual time or place or a sudden deviation from the historical pattern of access frequency, a secret key update mechanism is triggered in advance to prevent potential leakage risks. In the whole process, the system quietly completes the generation, distribution, update and revocation of the secret key in the background, ensuring the security of the core data and avoiding the burden and the risk of operational errors caused by frequent manual management. Machine learning can also facilitate the optimization of encryption by enabling systems to analyze the security of the current network comprehensively, the performance load of the device, and the sensitivity of the data being processed, automatically select and apply the most appropriate encryption algorithm and strength. When the system detects that the user is accessing sensitive information through public Wi-Fi, it will instantly increase the encryption level and adopt more complex and harder-to-crack algorithms to wrap the data; otherwise, when processing ordinary data on a highly trusted intranet, the system intelligently reduces encryption overhead to ensure smooth and efficient operation. Once an emerging threat to an encryption method is identified, the system actively and incrementally steers the entire environment towards a more robust algorithm, and it can maintain the optimal, context-dependent protection level in different stages of data storage and transmission and data processing to achieve the intelligent balance between security and efficiency.

5 Conclusion

In summary, machine learning has fundamentally changed the underlying logic of traditional Password cracking, enabling a deep understanding of the behavioural logic of human password-setting to generate highly contextualized candidate passwords and obfuscation rules, significantly improving the accuracy of the attack. Improving attack efficiency also promotes the innovation of defence technology, which can accurately identify dynamic threats, strengthen real-time defence, and optimize intelligent password strategy, adaptive encryption, and risk prediction. The follow-up research can be carried out in two aspects: one is to develop a dynamic defence model based on adversarial learning, which predicts and blocks the intelligent cracking path by simulating the attack behaviour; the other is to deepen the human-computer collaboration mechanism of password design, improve the security of the cryptographic system. Three is use Post-quantum cryptography to let the amount of computational increase exponentially to prevent machine learning to crack the password.

About the Author

Zhengyang Chen, master, research direction: cryptology.

References

[1] Fei Shi. Cracking the code for integrating technological innovation and Industrial Innovation[J].China Informatization,2025,(02):11-12.

[2] Xiaoyun Yu, Zhongtian Jia. Research on cryptographic algorithm recognition in the Internet of things[J].Internet of Things Technologies,2025, 15(03):100-104.

[3] Pengjun Feng, Lin Yang. Research on electronic balance password cheating methods and corresponding cracking methods[J]. China quality supervision,2025,(01):72-73.

[4] Ruiqi Xia, Manman Li, Shaozhen Chen. Recognition and analysis of cryptographic algorithms based on machine learning[J]. Journal of Cyber Security,2025,10(01):143-159.

[5] Peng Wan, Ming Zhu, Chao Zhou.Research on machine learning methods and applications in cryptology[J].China High and New Technology, 2023,(16):120-122.