UDC 004.02
IRSTI 20.53.15

SURVEY OF MODERN TRENDS IN COMPUTER SECURITY AND ARTIFICIAL INTELLIGENCE

Mirzakhmet Syzdykov
Satbayev University, Almaty, Kazakhstan
mspmail598@gmail.com
ORCID ID: https://orcid.org/0000-0002-8086-775X

Abstract. In this article we present the modern trends like Explainable Machine Learning (EML) versus classical Machine Learning (ML), the computer security is to be evaluated from the point of view of the hashing and authentication schemes in Hyper-Text Transfer Protocol (HTTP), most of the emphasis of this work is made towards hashing in representing the identity of data sources and making possible to use these hashes as crypto currencies in block-chaining technology, the survey also gives the definition of the basic hashing algorithm in the assumption that P not equals NP, where P is the polynomial class of complexity and NP is a non-polynomial class, since this fact was proved before as equality between P and NP, this is still remains challenging to prove that hashing function can be dual for either input or output, thus giving the possibility of compromising the computer security system and obtain the prior hash in hashing queue in block-chaining and, probably, partial chunks of the hashed data.

Keywords: hashing, algorithm, block-chaining, computer security, artificial intelligence.

Introduction

We present the description of HTTP-protocol authentication schemes like “basic” and “digest” based upon encoding and hashing validation of the user identity.

Even since, the proof of equivalence of complexity classes it doesn’t mean that there could exist the inverse function for hash as this question remains open – we give the strict proof that even since assumption of equality of P and NP-classes (P = NP), this is still impossible to create the inverse function. This proof is based upon definition of the universal hashing algorithmic steps which are dependent on the input. At any time and at each point this becomes evident that data to be obtained from the inverse steps cannot lay within the hash size which should as small as possible with respect to the average size of input data which are hashed on the one-direction function like hash probing.

We also give the notion to the rapidly growing interest of variety of communities for Machine Learning. We give the definition of algorithm of Explainable Machine Learning in order to define the extensible role of the ML in the field of algorithmic approach rather than resorting to the classical approach of neural networks.

HTTP basic authentication

Authentication schemes like basic and digest were presented in the official standard [1]. The basic authentication scheme prerequisites are discussed in [2]. As we operate on HTTP it's necessary to study the initial version of the protocol [3] – to the present day the version “2.0” is in active usage by many software servers like Apache or nginx, we name our project as “Alumni”, version 1.0. The extension to the digest authentication protocol is presented in [4]. The security issues were well studied by author in [5]: these issues commonly are of programmatic case like buffer overflow conditions, when stack is rewritten in the limited space and the return address is thus modified giving the possibility of compromising the system and running arbitrary shell code [5].

According to the server statistics in World Wide Web (WWW) the most popular web servers are IIS, Apache and nginx.

The authentication scheme of the basic authentication is provoked by the server response if “Authorization: “-field is missing and authorization token is followed encoded using base 64 encoding algorithm, the encoded string contains both login and password separated by “:” symbol.

In turn the server sends “WWW-Authenticate: basic[/digest], realm='...'” response if the authentication values are missing and authorization is required for the requested source. In this response

the field “realm” corresponds to the arbitrary identification of secured source.

The Socket Security Layer (SSL) is used in the modern web software and, thus, avoids the evaluation of data by reading data directly on Transfer Control Protocol (TCP) level.

The basic architecture of the web server is based upon the queues and, thus, represents the classical example of the system of mass-servicing.

The basic authorization can be replaced with the usage of hash token after validating the user data.

Hashing as a Modern Trend

The hashing is the term which is about to obtain the short part describing the whole part of the input data, most of these functions are to be one-directional thus giving the protection against the inverse determination of the hashed data.

The example of algorithm to be studied for the case of security level is presented in [6] and is named as a Whirpool algorithm.

In other meaning the hashing functions give the time complexity and performance of O(1) for the hash trees which are widely used in the modern Database Management Systems (DBMS) like Oracle, etc.

The block-chaining based upon the hash algorithm which describes the crypto currency is a task of obtaining the new hash at each transaction with the help of encryption algorithm – we divide these algorithms as past and present, which was elaborated by the author of this work. The classical example of hashing in block-chain is to hash the prior value to the values like additional attributes and, thus, giving the new value for the next crypto currency transaction. We elaborate and give the simple algorithm to obtain these hash values:

1. Generate random hash and encrypt using one-way algorithm with arbitrary key;
2. Decrypt the hash value;
3. Add amount of crypto currency as a big number to the hash;
4. Encrypt the next value;
5. Yield the value from step 4.

From the above algorithm, it follows that we don’t need the classical approach of mining in order to obtain the hash values at each step as this process by itself compromises the security risks of data collection by third parties and, thus, is vulnerable to the hashing attacks.

The last statement is prevalent since the MD5-hash was compromised in a distributed network in year, circa, 2003.

Our algorithm is secure as the key-values evaluation relays on the SSL connection, where the encryption key can be secured and stored in the local storage – this technology is widely used by the modern browsing software like Google Chrome and Microsoft Edge.

The basic hashing algorithm is defined as the number of steps for the data-window of the fixed size to be processed against the larger amount of data:

1. Generate window length less than input;
2. Set current position to zero;
3. For each input symbol module current position increment, set the value in the window array to the bit-mask operation like exclusive-OR.
4. Repeat the step 3 until the end of input.

As we have stated before, the hashing function cannot be inversed – we prove it by using the above example: since the window is less than input it’s obvious that the whole part of data cannot be obtained even if the algorithm is known to the performer of this task.

Explainable Machine Learning

Modern trends of Machine Learning like the TensorFlow and imaginary detection is briefly discussed in [7]. The existing algorithms are described in [8] along with a neural network method. The use of Machine Learning in the present day within the practical experience is presented in [9] – this

research states the question of the evaluation of Machine Learning in different types of application and its impact on the global stage. The modern usage of ML in wireless networks as the application is presented in [10].

After the overview of the research on the latest publications for Machine Learning and its application with pros and cons, we present the algorithmically practical solution for ML, known as Explainable Machine Learning, or, simply, EML.

EML relays on input data and can be represented as the matrix of sorted elements by the category and priority at each row and column, the steps of building this matrix and evaluation of the query are defined as follows:

1. Sort the input data matrix with given priority of each factor;
2. Query the result for each data in the sorted matrix for the short-coming range of lower and upper row;
3. For the number values do the floor or ceil of the value and compare it same way as the word entities in the input matrix.

Neural networks which are based upon sigmoid function can be extended for EML as the algorithm for arbitrary function with respect to the term various probability type which is well-known in modeling theory of mass-servicing queues.

The algorithm for arbitrary function f(x) is as follows:

1. Define the function f(x) and its range L and R;
2. Compute the minimum and maximum values of function f(x) on this range using differential calculus;
3. Train the neural network by applying function f(x) divided by the minimum and maximum, with subtraction of minimum.

The steps above give the novel example of the arbitrary learning without usage of sigmoid-function for neural network training.

Conclusion

Thus, we have presented the novel block-chaining algorithm which is secure and fast and doesn't require expensive mining operation to be performed on the specific devices which utilize Nvidia chipset bundle for this purpose.

The trend for computer security is also given with respect to the hash evaluation after authorizing user using standard methods or methods which involve security token or "cookie" safety.

We have also represented the modern results on the Machine Learning, which is named in this article as Explainable Machine Learning and, thus, is more relevant and doesn't require the neural network to be trained as the EML-decision tree can be expanded by online query also in the linear time as well as the query performance time O(n), where n is the number of factors in the data matrix of the input.

Thus, the modern trends are to be seen from other point of view with respect to the prior results on algorithmic stage like the prior experience of the authors of these methods.

Acknowledgements

The author expresses gratitude to the creators of SharpDevelop programming environment for creating a valuable and free package for effective and professional software development in C# on .NET 2.0 platform.

Funding

This work was partially supported by an educational grant of the Ministry of Education and Sciences of the Republic of Kazakhstan during author's studying at the Satbayev University from 2001 to 2006.

References

[1] Franks John, et al. HTTP authentication: Basic and digest access authentication. 1999. 2617.

[2] Reschke Julian. The 'Basic' HTTP authentication scheme. 2015. 7617.
[3] Fielding Roy, Julian Reschke. Hypertext transfer protocol (http/1.1): Authentication. 2014. 7235.
[4] Franks John, et al. An extension to HTTP: digest access authentication. 1997. 2069.
[5] Сыздыков Мирзахмет. Win32-buffer/heap overrun conditions или секреты написания эксплоитов. Security Lab.ru.
https://www.securitylab.ru/analytics/216279.php?el_id=216279&VOTE_ID=101&view_result=Y. (accessed Jan 25, 2023)
[6] Barreto P. S., Vincent Rijmen. The Whirlpool hashing function. First open NESSIE Workshop, Leuven, Belgium. 2000. 13
[7] Jordan Michael I., Tom M. Mitchell. Machine learning: Trends, perspectives, and prospects. Science 2015. 349.6245. 255-260.
[8] Mahesh Batta. Machine learning algorithms-a review. International Journal of Science and Research (IJSR). 2020. 9. 381-386.
[9] Athey Susan. The impact of machine learning on economics. The economics of artificial intelligence: An agenda. University of Chicago Press. 2018. 507-547.
[10] Chen Mingzhe, et al. Artificial neural networks-based machine learning for wireless networks: A tutorial. IEEE Communications Surveys & Tutorials 21.4. 2019. 3039-3071.

ОБЗОР СОВРЕМЕННЫХ ТЕНДЕНЦИЙ В ОБЛАСТИ КОМПЬЮТЕРНОЙ БЕЗОПАСНОСТИ И ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

Мырзахмет Сыздыков
Казахский национальный исследовательский техический университет имени К.И. Сатпаева, Алматы, Казахстан
mspmail598@gmail.com
ORCID ID: https://orcid.org/0000-0002-8086-775X

Аннотация. В этой статье мы представляем современные тенденции, такие как объяснимое машинное обучение (EML) по сравнению с классическим машинным обучением (ML), компьютерная безопасность должна оцениваться с точки зрения схем хэширования и аутентификации в протоколе передачи гипертекста (HTTP), основной акцент в этой работе сделан на хэширования для представления идентичности источников данных и обеспечения возможности использования этих хэшей в качестве криптовалют в технологии блокчейн-цепочек, в исследовании также дается определение базового алгоритма хэширования в предположении, что P не равно NP, где P - полиномиальный класс сложности, а NP - неполиномиальный класс, поскольку этот факт был доказан ранее как равенство между P и NP, все еще остается сложной задачей доказать, что функция хэширования может быть двойной как для ввода, так и для вывода, что дает возможность поставить под угрозу систему компьютерной безопасности и получить предыдущий хэш в очереди хэширования в блочной цепочке и, возможно, частичные фрагменты хэшированных данных.

Ключевые слова: хеширование, алгоритм, блочная цепочка, компьютерная безопасность, искусственный интеллект.

КОМПЬЮТЕРЛІК ҚАУІПСІЗДІК ПЕН ЖАСАНДЫ ИНТЕЛЛЕКТТІҢ ЗАМАНАУИ ТЕНДЕНЦИЯЛАРЫНА ШОЛУ

Мырзахмет Сыздықов
Қ. И. Сәтбаев атындағы Қазақ ұлттық техникалық зерттеу университеті, Алматы, Қазақстан
mspmail598@gmail.com
ORCID ID: https://orcid.org/0000-0002-8086-775X

Аңдатпа. Бұл мақалада біз классикалық машиналық оқытумен (ML) салыстырғанда түсіндірме машиналық оқыту (EML) сияқты заманауи тенденцияларды ұсынамыз, компьютерлік қауіпсіздікті гипермәтінді беру протоколындағы (HTTP) хэштеу және аутентификация сызбалары тұрғысынан бағалау керек. Бұл жұмыстың негізгі бағыты деректер көздерінің сәйкестігін көрсету және қамтамасыз ету үшін хэштеу болып табылады. Бұл хэштерді блокчейн тізбегі

технологиясында криптовалюта ретінде пайдалану мүмкіндіктері, зерттеу сонымен қатар P NP - ге тең емес деген болжаммен негізгі хэш алгоритмін анықтайды. Мұндағы p-көпмүшелік күрделілік класы, ал NP-көпмүшелік емес класс, өйткені бұл факт бұрын P мен NP арасындағы теңдік ретінде дәлелденген, хэштеу функциясы енгізу үшін де, шығару үшін де қосарлы болуы мүмкін екенін дәлелдеу әлі де қиын міндет болып табылады. Бұл компьютерлік қауіпсіздік жүйесіне қауіп төндіруге және алдыңғысын алуға мүмкіндік береді блокчейндегі хэштеу кезегіндегі хэш және мүмкін болатын хэштелген деректердің ішінара фрагменттері болады.

Кілттік сөздер: хэштеу, алгоритм, блокчейн, компьютерлік қауіпсіздік, жасанды интеллект.

Сведения об авторе:

Анг.: Syzdykov Mirzakhmet - Satbayev University, Almaty, Kazakhstan

Каз.: Сыздықов Мырзахмет- Қ.И. Сәтбаев атындағы Қазақ ұлттық техникалық зерттеу университеті, Алматы, Қазақстан.

Рус.: Сыздыков Мырзахмет- Казахский национальный исследовательский техический университет имени К.И. Сатпаева, Алматы, Казахстан.