Journal of Multimedia Information System VOL. 9, NO. 3, September 2022 (pp. 191-200): ISSN 2383-7632 (Online)
https://doi.org/10.33851/JMIS.2022.9.3.191

Modeling Vulnerability Discovery Process in Major Cryptocurrencies

HyunChul Joh[1], JooYoung Lee[2*]

Abstract

These days, businesses, in both online and offline, have started accepting cryptocurrencies as payment methods. Even in countries like *El Salvador*, cryptocurrencies are recognized as fiat currencies. Meanwhile, publicly known, but not patched software vulnerabilities are security threats to not only software users but also to our society in general. As the status of cryptocurrencies has gradually increased, the impact of security vulnerabilities related to cryptocurrencies on our society has increased as well. In this paper, we first analyze vulnerabilities from the two major cryptocurrency vendors of Bitcoin and Ethereum in a quantitative manner with the respect to the CVSS, to see how the vulnerabilities are roughly structured in those systems. Then we introduce a modified AML vulnerability discovery model for the vulnerability datasets from the two vendors, after showing the original AML dose not accurately represent the vulnerability discovery trends on the datasets. The analysis shows that the modified model performs better than the original AML model for the vulnerability datasets from the major cryptocurrencies.

Key Words: Vulnerability Discovery Model, AML, Bitcoin, Ethereum.

I. INTRODUCTION

After introduction of the first cryptocurrency payment system *eCash* in 1982 [1], more than two decades later, equipped with a sensational blockchain technology, the Bitcoin [2] is considered as a true decentralized cryptocurrency. Its launch in 2009 was the beginning of the entire cryptocurrency movement. Bitcoin and the blockchain technology that works with Bitcoin were invented by an individual or groups of people under the alias of Satoshi Nakamoto, whose identity has yet to be revealed. Today, Bitcoin has been suggested as an alternative to the legal monetary system in some countries (https://www.bbc.com /news/world-africa-61565485).

In the white paper [2], the author argues that the legal monetary system governed by central financial institutions centralized wealth, and it made social and financial mobility difficult. The public savings were eaten away mainly through inflation due to central bank currency issuance. The author says that Bitcoin ties to solve this problem by fixing the number of units issued to eliminate inflation caused by printing money. Introduced P2P blockchain technology used by Bitcoin tells us that financial institutions were not needed to facilitate transactions and verify ownership anymore under the Bitcoin payment system. Today, Bitcoin is recognized as the most popular cryptocurrency, and its fluctuations of price have a substantial impact on the rest of the cryptocurrency market.

Meanwhile, Ethereum [3], which is launched in 2015, is widely acknowledged as the second most popular cryptocurrency after Bitcoin, but it is quite different from the first one. The name of Ethereum represents the blockchain platform for the cryptocurrency, called smart contracts. Ether is actually the name of a cryptocurrency. Ethereum can also be thought as defined rules that can create various dApps (decentralized applications), and since those new attributes are introduced, Ethereum is represented as Blockchain technology version 2.0. Although there have been various smart contract platforms after introducing Ethereum, it has been remained the most popular smart contract platform so far.

Currently, there are more than 18,000 cryptocurrencies available as of March 2022 (https://coinmarketcap.com). Among them, Bitcoin and Ethereum are the two most popular cryptocurrencies with some other major players whose popularities tend to ebb and flow. They are all quite different each other in detail.

Bitcoin's purpose is to be an alternative method for the traditional currencies whereas Ethereum is intended as a payment for the usage for its blockchain platform, although

Manuscript received August 17, 2022; Revised September 6, 2022; Accepted September 9, 2022. (ID No. JMIS-22M-08-025)
Corresponding Author (*): JooYoung Lee, +82-53-600-5844, jjoolee@kiu.kr
[1]School of Smart Industry, Kyungil University, Gyeongsan, Korea, joh@kiu.kr
[2]School of K-Culture Entertainment, Kyungil University, Gyeongsan, Korea, jjoolee@kiu.kr

they are all traded as an asset. Also, Ethereum, which is the second generation Blockchain, introduced dApps which are operating on top of blockchains while the first generation of blockchain technology is virtually only used for cryptocurrencies.

Even though, dApps create rich technical environments on top of the blockchain systems, they increase the security risk as well due to increased complexities [4-6]. Therefore, it is unlikely that entirely safe cryptocurrencies will become possible anytime soon.

One interesting point is that the founder of Ethereum, Vitalik Buterin [3], share his prediction of gloomy future for Bitcoin in 2022. He thinks transaction fees and proof-of-work (PoW) methods will hold back Bitcoin from a long-term perspective. In the interview with economist Noah Smith (https://noahpinion.substack.com/p/interview-vitalik-buterin-creator), Buterin argued that Bitcoin's security level will not be good enough for years to come. He pointed out that the rewards for mining are being converted into transaction fees as Bitcoin's mining rewards are lowered every four years, but the fees are cheaper compared to mining rewards. He predicts that Bitcoin's security will come entirely from fees. Buterin also expressed concern that Bitcoin is still using PoW mining. He said, the PoW method has a lower level of security than the proof-of-stake (PoS) method. However, it is politically impossible for Bitcoin to switch to a PoS method. He also mentioned that if the Bitcoin network is actually attacked, there may be an attempt to switch to a hybrid PoS approach which will be a very painful process.

In spite of the security risks, a significant portion of the public is willing to take the risk of participating cryptocurrency markets because that makes the transactions much more efficient on the Internet and creates enormous and promising virtual marketplaces. Hence, while we are using cryptocurrencies, it is indispensable to allow a certain degree of risk and manage security vulnerabilities with precautionary measures. Of course, it is not easy to determine cryptocurrencies as real money at this stage whatsoever. Something trusted deserves money, but in reality, for it to be used as money, its price must be stable. It is pretty hard for ordinary persons to accept Bitcoin, whose price fluctuates as it is now.

In the paper [7], titled "Is Cryptocurrency Money?," the authors discussed about cryptocurrency with respect to the three core functions of medium of exchange, store of value, and unit of account. Especially, the authors focus on three major cryptocurrencies of Bitcoin, Ether, and Ripple. In the paper, the author conducted three studies of whether cryptocurrencies are valid measurement items or not, whether individuals consider cryptocurrencies as money or not, and among the three cryptocurrencies which one is better in terms of the three core functions introduced above. The results show that individuals recognize that cryptocurrencies are valid measurement items, and they fulfill the core functions of money in positive ways. Also, the Bitcoin is recognized a better medium for exchange and store values while all the three cryptocurrencies are equally recognized in the perception of core function of money.

In the paper [8], the authors give their idea of what cryptocurrency is, what are similarities with money in the forms of bills and coins. They start to discuss the role of the money and what it should be like, how cryptocurrency works, characterizing cryptocurrency, comparing cryptocurrency and money, varieties of cryptocurrency and monetary policy. They utilize technical and philosophical methods together for guiding readers.

Meanwhile, according to the nist.gov (https://csrc.nist.gov/glossary/term/vulnerability), a (software) vulnerability is a "weakness in an information system, system security procedures, internal controls, or implementation that could be exploited or triggered by a threat source." Each year, a tremendous number of Internet users are exposed at great risk due to a laziness of applying security patches on their software systems [9]. Publicly available but not patched vulnerabilities create high alert because those security risks offer invaders the power to have complete control for the system. Also, sensitive data can be leaked.

When we consider about financial importance of cryptocurrencies, potential security vulnerabilities in those systems could trigger a significant negative effect on our society worldwide. Nevertheless, only few quantitative security analyses exist dealing with cryptocurrencies as far as we know. Quantitative security analyses are effective methods due to the statistical risk assessment for the potential intrusions.

In this paper, we are trying to provide a quantitative analysis of security vulnerabilities from the major cryptocurrency vendors of Bitcoin and Ethereum with the respect to the CVSS. Then we introduce a modified vulnerability discovery model which performs well with the vulnerability datasets of the two cryptocurrencies after showing inappropriate model performance from the original AML vulnerability discovery model.

The rest of the paper is organized as follows. Section 2 presents some of the related works and section 3 introduces CVSS and the datasets we are analyzing here. Section 4 reviews AML vulnerability discovery model which requires readers to understand the next section and investigates vulnerability discovery process with AML model in the two cryptocurrency systems. Section 5 introduces modified AML, called LTM, and see whether LTM outperforms AML with the given two vulnerability datasets. Finally, Section 6 concludes this paper.

II. RELATED WORKS

Vujičić et al. [10] give a short introduction about the Bitcoin, Ethereum, blockchain and other related subjects, such as proof-of-work, scalability problem, smart contracts, etc. Hence, to get a quick and brief background about cryptocurrency, this paper could be a good start point. Kushwaha et al. [11] discuss various Ethereum smart contract security vulnerabilities in three categories of Solidity Programming Language, Features of Ethereum Virtual Machine, and Design features of Ethereum Blockchain. The authors provide insights about security challenges and potential research directions via extensive reviews for more than 140 references.

There are quite a lot of survey research papers available relate to the cryptocurrencies. In the paper [12], the authors try to review security vulnerabilities in the Ethereum smart contract systems. The major target for this survey is discussing about security vulnerabilities detection methods, attacks in real life, and preventive methods in the Ethereum smart contract. By considering different elements, they compare several Ethereum smart contract analysis methods. The paper first categorizes the Ethereum blockchain by layers of application, consensus, data and network. Then in each layer, they review papers related to each category over the period of 2016 to 2021. By doing that, they try to provide useful survey results of security concerns and their root causes to researchers, students, and developers. Solidity programming language, Ethereum virtual machine and Ethereum blockchain design are the three main root causes for the major paper survey. In the paper, they also provide some of the guidance for smart contract developers not to generate security vulnerabilities.

Quamara and Singh [13] try to illustrate security concerns in cryptocurrencies systematically, and present state-of-the-art in those research area from various perspectives. They examine various applications, algorithms, technologies, and proper utilizations related to cryptocurrencies in depth. Thoroughly investigate literatures for their contributions with respect to the security aspect in cryptocurrencies, occurred from 2009 to 2020. Furthermore, the paper introduces major ongoing cryptocurrency projects worldwide. Their major strategy for systematic survey has three main steps of, first, deriving research questions, second, extracting literatures, and third, analyzing literatures. Since this paper presents and discusses about fundamentals of cryptocurrencies, consensus of mechanisms, taxonomy of cryptocurrencies, state-of-the-art in depth, this paper should give a fine insight to other researchers.

In the paper [14], the authors try to characterize and identify cryptocurrency related scams, and show the urgency of identifying and publicly announce scammers not to produce more victims. The authors give 300 fake applications and more than 1,500 scam webpages by utilizing *Typosquatting* generation techniques and existing literatures. The authors categorized the scam methods into four classes: fraudulent exchanges, mining scams, Ponzi schemes, and scam wallets. Moreover, they reveal 30 fake app groups and 94 scam domain groups by scrutinizing the relationship among those webpages and apps. To prevent further financial losses, authors publicly released the list of fake domains and apps they identified.

In the paper [15], the authors examine 146 research papers with respect to various characteristics of cryptocurrency transactions, such as bubble and extreme condition, cryptocurrency trading systems, prediction of volatility and return, technical trading, crypto-assets portfolio construction and others. For the methodology point of view, they examine the properties, technologies, and summarizing datasets from the surveyed literatures by categorizing.

Erfani and Ahmadi [16] introduced algorithms, mechanisms, and security services related to some of the distinct Bitcoin security features. In the paper, they suggest a security functional architecture to reduce security risks. In the suggested reference model, there are five layers of i) mathematical module layer, ii) security mechanism layer, iii) security service layer, iv) security management layer, and v) security policy and business requirements layer. The authors claims that the reference model could provide a secure channel to all of cryptocurrencies in a digital wallet when they utilized at P2P network environments.

In the paper [17], the authors suggest a method of swapping assets among different blockchain systems without a reliable third party by utilizing Atomic Cross Chain Swap (ACCS). Currently, if we like to exchange Bitcoins with Ethers, or vice versa, we need a trust third party which provide an exchange platform that will handle the exchange. In order to make it possible to exchange assets between two different blockchains, ACCS based Solidity scripts are implemented by the authors in both Bitcoin and Ethereum sides. The proposed method provides a way to exchange assets between Bitcoin and Ether, but they do not examine a performance analysis.

Christopher et al. [18] analyze volatility stage from the five cryptocurrencies of Bitcoin, Ethereum, Litecoin, BinanceCoin, and DashCoin in the period of January 1st 2018 to April 1st 2021. They conducted quantitative analyses for the cryptocurrencies and the datasets are gathered from https://investing.com. Specifically, they utilize the quantitative manner of Generalized Autoregressive Conditional Heteroscedasticity (GARCH) and Autoregressive Conditional Heteroscedasticity (ARCH) models. The results shows that GARCH and ARCH methods are not suitable for daily life situations in the cryptocurrency products for a calculation of volatility stage, such as forecasting price

movements. Rather, the two methods are more appropriate to annual analysis in general.

In the paper [19], the authors propose a model, called time to ruin of Cramer-Lundberg (CL) model which allows predicting the Bitcoin confirmation time off-the-shelf. In this method, however, with the CL model assumptions, the utilized data might not be conformed completely. The authors try to show that even with small changes, with the model assumptions, the proposed model can be used accurately for predicting the confirmation times heuristically. Findings from this paper can help for dealing with *Mempool* Bitcoin data which deviate from the model assumptions.

In the study [20], the authors try to compare the exchange graphs among Bitcoin, Ethereum, Z-Cash, Litecoin, and Dash in terms of dynamics of their exchange along with the timeline. In the paper, they define a monthly transaction graph and a cumulative monthly transactions graph, and it is observed that the number of cumulative transactions increase with linear pattern and the density is always declining. After that, the authors examine the properties and distinct characters from the cryptocurrencies in a quantitative manner. They found that the price of cryptocurrencies is closely correlated to the edges of the transaction graphs, growth rate of the nodes, and the density of graphs in the paper.

Hu et al. [21] investigate Bitcoin applications running on smartphones to check security stabilities. It is the first study comparing the Bitcoin wallet standards and implementations or operations in real services. They found three types of vulnerabilities of i) leaking Bitcoin wallet user privacy, ii) downloading continuously unwanted Bitcoin transactions in the background, and iii) violating Bitcoin principle of decentralization. Also, they create three proof-of-concept attacks for the identified vulnerabilities. Further, they suggest practical solutions for those problems. The vulnerabilities, they identified, had been reported to the CVE database (https://cve.mitre.org) properly.

Chan et al. [22] systematize security related issues in Ethereum systems according to the three aspects of vulnerabilities, attacks, and defenses. The authors try to serve researchers, practitioners, and students with insights of what are root causes of 40 types of Ethereum vulnerabilities, and what are consequences for 29 attacks against Ethereum. Then the authors precautionary measures 51 defenses to prevent those security problems. They also give guidelines of possible future research topics categorized into three directions: eliminating known Ethereum vulnerabilities, developing Ethereum test tools and environments, and formalizing, analyzing and quantifying Ethereum security. The paper could be a good material for having a simple but comprehensive review on Ethereum related security in details.

Meanwhile, vulnerability discovery models (VDMs) describe the discovery of software vulnerabilities along with the calendar timeline, although some of them are based on the amount of installation based. So far, several VDMs have been proposed. Alhazmi-Malaiya Logistic (AML) model [23], which was originally proposed and validated for operating systems, is one of the most well-known quantitative vulnerability discovery models. Joh and Malaiya [24] compares AML with other S-shaped VDMs based on the skewness for analyzed datasets. The observation tells us that Logistic and Gamma distribution-based model outperform other S-shaped models of Beta, Weibull and Normal distribution-based models.

III. CVSS ANALYSIS

Common Vulnerability Scoring System (CVSS) [25] is providing a method to represent a principal characteristic of a software vulnerability and generate a numerical score values showing its severity. In 2003, National Infrastructure Advisory Council assigned a plan to examine the challenges of incompatible and various software system for generating vulnerability scoring systems. Accordingly, since its original release in 2005, the CVSS has been adopted by numerous software vendors and academics for vulnerability scanning and compliance tools, product risk assessment, and security bulletins [26-28]. Then, significant issues (Version 2 History: https://www.first.org/cvss/v2/history) with the first version of the CVSS was observed, and that brought the project to the following version, released in 2007. Then later, in June 2015, the third version was released with reflecting further considerations.

The score system can be interpreted into a meaningful and qualitative representation such as low, medium, high, or critical for given categories. And that helps security researchers to measure and prioritize for vulnerability management processes properly.

The final CVSS score for each vulnerability ranges from 0.0 to 10.0, and higher scores indicate more vulnerable to exploitation and cause greater severe consequences. The score is composed of three metric groups of Base, Temporal and Environmental. The Base metric signifies the fundamental and essential attributes of a security vulnerability. Scores from the base metric is not changed over time, while Temporal and Environmental metrics are measured dynamically in time and IT environment. Therefore, the base metric is only required for the final CVSS score whereas the other two metrics are optional since they are hard to be measured. For more details about CVSS please refer to the CVSS score user guide [25].

The datasets we are analyzing here are collected on June 2022 at https://www.cvedetails.co*m* where its database is

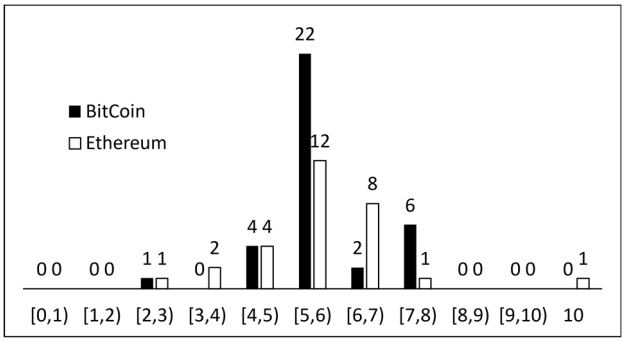


Fig. 1. Ranges of CVSS scores from Bitcoin and Ethereum.

Table 1. Brief stat. from Fig. 1.

	Bitcoin	Ethereum
Count	35	29
Mean	5.3714	5.4482
Median	5	5
Min	2.1	2.1
Max	7.8	10

depending on the National Vulnerability Database (NVD, *nvd.nist.gov*). Before a CVSS score is recorded into NVD database, specialists examine the software vulnerability and allocate one of the letter grades [28]. Because the essential goal of CVSS is delivering analogous vulnerability score system among the software vendors, security experts are only granted to evaluate the vulnerabilities with predefined letters. In the final step, scoring is in the process of merging all the three metric values according to the explicit rules (https://www.first.org/cvss/calculator/3.1).

Fig. 1 shows distributions of CVSS scores from the vulnerabilities in the two cryptocurrencies, and scores are grouped together in the unit of a single point. By and large, the two cryptocurrencies have similar risk levels.

Table 1 shows brief statistics about the datasets. Total 35 and 29 vulnerabilities have been found in Bitcoin and Ethereum, respectively, and the values of mean, median and minimum are about the same. Ethereum has a vulnerability (CVE-2018-15890) having CVSS score of 10.0 published on June 20th 2019. It is related to an unsafe deserialization in *EthereumJ* which is a java implementation of the Ethereum protocol.

IV. AML VULNERABILITY DISCOVERY MODEL

Vulnerability Discovery Models (VDMs) represent the discovery trends of cumulative number of known vulnerabilities. From early 2000s, when software vulnerability datasets became enough to be analyzed, some security and software related researchers have been proposed diverse VDMs [29]. Although there are some VDMs using the number of installations or testing effort as the main factors for modeling [30-31], majority of the VDMs are time-based models, where the calendar time is used for the independent variable factor, due to an ease of use. Each model has different assumptions with different model parameters restricting its performances. Naturally, they represent the vulnerability discovery trends a bit differently with the same datasets.

Among the time-based VDMs, Alhazmi-Malaiya Logistic (AML) model [23], which was initially intended and proven for computer operating systems, is one of the well-recognized software vulnerability discovery models. Fig. 2 describes the S-shaped AML model representing the relationship between the software age (time) and the number of cumulative vulnerabilities found.

The AML model assumes that during the learning phase, very few vulnerabilities are found since software systems do not have a significant number of users testing software systems. During the next linear phase, a steady stream of vulnerabilities is reported due to the gaining popularity. In this stage, the discovery rate achieves the highest value as a result of gaining popularity. And, in the final saturation segment, the discovery rate drops down because of a losing popularity. This is happening since the application users migrate to the next version or an alternative version for the software system.

The durations indirectly rely on features such as market share or hidden vulnerabilities. In Fig. 2, the dashed bell-shaped line expresses the immediate vulnerability discovery rate where the solid S-shaped line signifies the cumulative number of vulnerabilities. Market shares are key element affecting on the effort expended in searching unknown vulnerabilities. A greater market share delivers more incentive to exploit and explore security vulnerabilities. The influence of market share, rise and fall, is implicitly reflected on the AML model [32]. In the figure, a mid-point and the two transition points are well defined by Alhazmi and Malaiya mathematically [33].

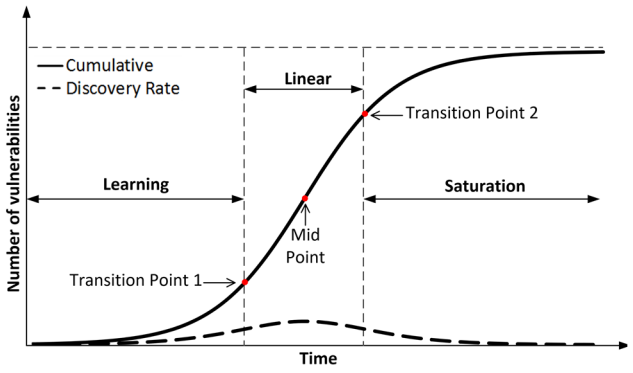


Fig. 2. AML vulnerability discovery model.

Theoretical philosophy of the AML model assumes that the cumulative number of vulnerabilities is controlled by the two components of A and B from equation (1). The first element A increases with the calendar timeline due to the rising of the popularity in usage. The second element B falls down as the number of hidden vulnerabilities reduces. The saturation phenomenon is described by the second element.

$$\omega(t) = A\Omega(B - \Omega). \quad (1)$$

$$\Omega(t) = \frac{B}{BC^{-ABt}+1}. \quad (2)$$

Assuming the discovery rate is specified by equation (1), then equation (2) can be derived by solving the differential equation that model the cumulative number of vulnerabilities. Therefore, $\Omega(t)$ signifies the number of vulnerabilities discovered by time point t. In Fig. 2, the bell-shaped and S-shaped lines are represented by equation (1) and equation (2) respectively.

Factors of A and B are experimental constants which controlled by the recorded data. Constant C is introduced while solving equation (1). The model is defined for time value t from the minus infinity to the plus infinity. Notice that when time t goes to the plus infinity, B becomes the ultimate number of vulnerabilities, in a given software system. There is a possibility that the second phrase of saturation stage might not be indicated since a software system is not presented for a sufficient time enough.

Here, in this section, we examine vulnerability datasets from the Bitcoin and Ethereum. Cryptocurrency is a big topic in these days and the AML model does not perform well with the two datasets due to the relatively high number of vulnerabilities at the beginning of the datasets. And this gives us a good justification to introduce a modified AML model, which well represents the vulnerability discovery trend with the initial peak datasets with a different way from the existing time-based vulnerability discovery models including AML. The modified model will be presented in the next section.

Table 2. AML model fitting from Fig. 3.

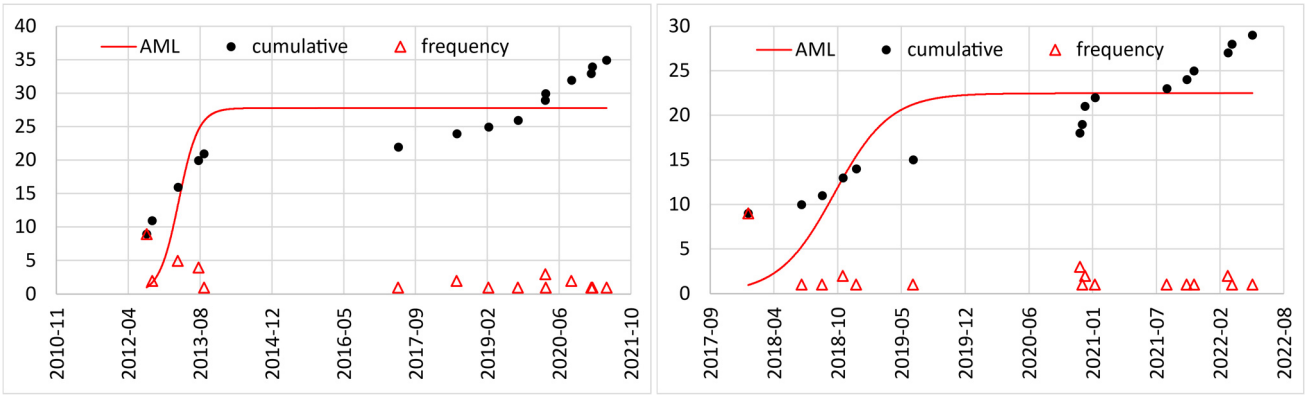
AML Para. / R^2	Bitcoin	Ethereum
A	0.001787454	0.001773956
B	27.79581387	22.47032075
C	1.345032756	1.333266979
R^2	0.851658254	0.791630848

Fig. 3 depicts the AML model fittings and Table 2 shows the AML model fitting parameters from Fig. 3 with correlation coefficient values (R^2). Bitcoin and Ethereum get 0.85 and 0.79 correlation results respectively, which seems not quite good model fittings.

In the figure, for both Bitcoin and Ethereum, the initial adaptation is quick enough, the early learning phases are almost not shown. In this situation, the discovery process may be expressed best with a linear model if the linear growth trend is keep going continually [34]. However, since the vulnerability discovery trend should curve sometime in the future, using a simple linear model seems not a best choice for a long run.

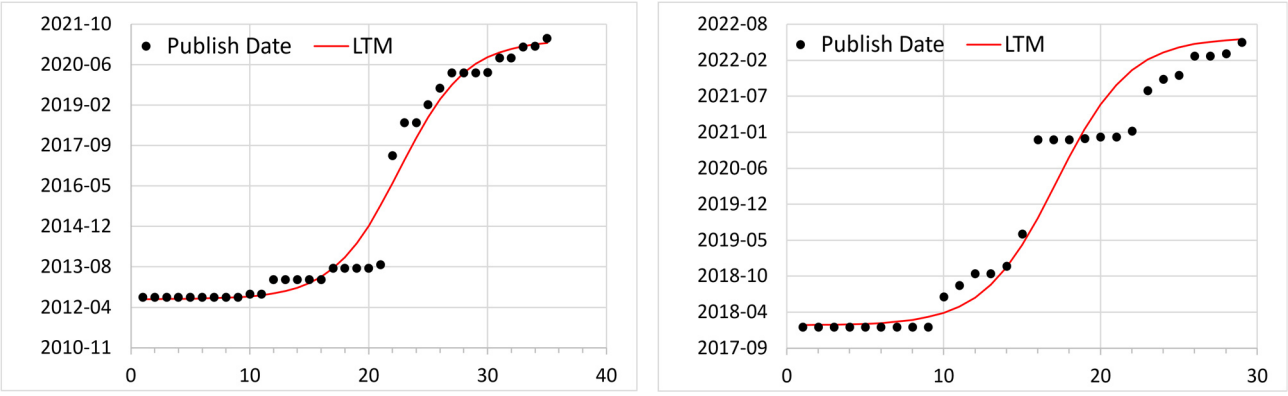
V. MODIFIED AML MODEL

When we transpose the two values between the x-axis of vulnerability publish date and the y-axis of cumulative number of vulnerabilities from Fig. 3, we obtain a graph in Fig. 4. For the model fitting on the transposed graph, a modified AML model will be applied since we are not able to apply the original AML directly into the transposed graph. Here are the two reasons why we are not able to fit the AML



(a) Bitcoin (b) Ethereum

Fig. 3. AML model fittings for the cumulative number of vulnerabilities; x-axis is calendar time, marking in daily basis while y-axis represents the number of vulnerabilities.



(a) Bitcoin (b) Ethereum

Fig. 4. LTM model fittings for the cumulative number of vulnerabilities; y-axis is calendar time, marking in daily basis while x-axis represents the number of vulnerabilities.

model into the transposed graph directly with the given datasets.

First, the values on the y-axis are not integer numbers anymore, but dates after the transposition. So, we need to somehow transfer the dates into consecutive numbers so that regressions can be conducted. Second, since the nature of the logistic model, the starting value should be close to zero, which might not be the case when the values on y-axis are not the cumulative numbers of vulnerabilities.

For the first issue, dates can be converted into consecutive values when we consider a single date as a unit, just like the Unix timestamp system (https://www.unixtimestamp.com), where the number of seconds, elapsed since January 1st 1970 UTC. In our case, the elapse of time should be a day, instead of a second. Then, the date values can be transformed into consecutive values which can be applied into a regression model.

According to the Microsoft (https://support.microsoft.com/en-us/office/datevalue-function-df8b07d4-7761-4a93-bc33-b7471bbff252), the date value in Microsoft Excel represents a date between January 1st, 1900 and December 31st, 9999. Thankfully, we can easily convert the date values into numbers in Excel. Date values are converted into consecutive integer values ranges from 1 (January 1st, 1900) to 29,58,465 (December 31st , 9999).

To solve the second issue, we need to introduce the fourth parameter into AML model, so that the model can start not only from the value of zero, but also from any level of value. The modified AML model is shown in equation (3). Since the parameter D is used to raise the y-intercept level, the initial value for the parameter can be approximately set as the date when the first vulnerability has been published. Mathematically, the only difference between the equations (2) and (3) is existence of parameter D. We call the modified equation as Logistic Transpose Model (LTM) since the model transposes the two axes from the AML model, where the L stands for Logistic.

$$H(n) = \frac{B}{BC^{-ABn}+1} + D. \quad (3)$$

To represent equation (3), the Greek letter H (eta) is used since the word date in Greek is *ημερομηνία*, where the first letter η is lower case of H in Greek. In Fig. 4, we can observe that, for both cases, the model fitting seems quite well. Table 3 shows the LTM model fitting parameters from Fig. 4 with correlation coefficient values. Bitcoin and Ethereum achieve 0.97 and 0.95 of R^2 values, which are better fitting results than the original AML model fittings from Fig. 3.

Moreover, unlike Fig. 3, Fig. 4 shows the clear S-shaped growth trends. Here, when the vulnerability discovery trends follow the S-shaped growth pattern, that indicates for both systems expecting more vulnerabilities in a short period of time in near future.

Now, let us explain how to generate fitting plots (Figs. 3 and Fig. 4) and optimal parameter values in this paper. The bottom line is that we manually utilized Microsoft Excel with the real datasets of vulnerability Publish Dates from the NVD. We first create scatter plots from the datasets, then entered equation (2) for Fig. 3 and equation (4) for Fig. 4 with the required parameter values of A, B, C, D and time factor t in each row in an Excel spreadsheet table. Since we have no clue what the optimal parameter values are, in the beginning, we put some random values. And for the time

Table 3. LTM model fitting from Fig. 4.

AML Para. / R^2	Bitcoin	Ethereum
A	0.726386788	0.79
B	3204.535941	1600
C	1.000153471	1.000339
D	41099.81981	43127.06812
R^2	0.973614882	0.950196869

factor *t*, as we mentioned above, consecutive integer values are used, and the values are converted from the date information, and it is handled by Excel automatically. Then, we achieve ugly fitting plots since the parameter values are not appropriate.

Now, there are total 35 and 29 rows in Bitcoin and Ethereum datasheets respectively since they have 35 and 29 vulnerabilities found each. In each row, for both cryptocurrencies, we calculate the cumulative number of vulnerabilities from the first row up to the current row, to the end. Then, in each row, we calculate the difference between the real cumulative number of vulnerabilities and estimated cumulative numbers by each model.

We add up all the estimation error values (differences), then try to minimize the total add up values to perform a least squares regression method. By doing this, we can achieve the optimal parameter values of A, B, C, and D. To conduct this step, Excel *Solver* is used, which is Excel add-in program, and it is frequently utilized for finding optimal model parameters by model fittings.

VI. CONCLUSION

Here, we investigate the security vulnerabilities from the two most popular cryptocurrencies in a quantitative manner by using NVD dataset. Then the AML vulnerability discovery model and LTM, modified AML version, are applied for model fittings on the two datasets from Bitcoin and Ethereum to see whether the models can represent the vulnerability discovery trends well enough. The results indicate that the LTM outperforms AML in vulnerability discovery model fittings.

In some sense, it is an acceptable result since AML model initially intended and tested only for computer operating systems. Since reasonably LTM describes the vulnerability discovery processes well with the major cryptocurrency software systems, it might be possible that the number of vulnerabilities discovered soon could be estimated. If the estimation is possible in some degree, then the development managers could allocate the resources optimally in advance. Moreover, cryptocurrency investors and consumers have better outlook for their digital assets.

Future research is desirable to assess the impact of evolution of cryptocurrency software systems which go through various versions by clearly reflecting the joint codes among the contiguous software versions, vulnerabilities removed and inserted in the process with the influence on resource allocation for patch development and testing. As we observed that the discovery trends are showing the S-shaped curve, it might be an interesting study to apply various S-shaped distribution based VDM models. Also, further investigation is necessary to see whether other cryptocurrencies show the S-shaped vulnerability discovery pattern or not.

Recently, some researchers have started to apply machine learning techniques into software vulnerability discovery processes [35-36]. A machine learning procedure could be a good alternative method for estimating vulnerability discovery trends in the long run since it does not require source code level analysis nor human interventions. Consequently, applying machine learning practice to a cryptocurrency related software error detection could be a fine future work in any sense.

Additionally, we could examine the seasonality in vulnerability discovery process by using autocorrelation analysis and seasonal index method [37].

ACKNOWLEDGEMENT

This research was supported by the intramural research program in Kyungil University.

REFERENCES

[1] D. Chaum, "Blind signatures for untraceable payments," D. Chaum, R. L. Rivest, and A. T. Sherman (eds.), *Advances in cryptology proceedings of crypto 82*, Plenum, New York, NY: Springer-Verlag,pp.199-203, 1982.

[2] S. Nakamoto, "Bitcoin: A peer-to-peer electronic cash system," *Decentralized Business Review*, 2008.

[3] V. Buterin,"A next-generation smart contract and decentralized application platform," *White Paper*, vol. 3, no. 37, 2014.

[4] P. Daian, Analysis of the DAO Exploit, 2016 https://hackingdistributed.com/2016/06/18/analysis-of-the-dao-exploit/.

[5] S. Palladino, The Paritywallet Hack Explained, 2017. https://blog.openzeppelin.com/on-the-parity-wallet-multisig-hack-405a8c12e8f7/.

[6] L. Poinsignon, BGP Leaks and Cryptocurrencies, 2018. https://blog.cloudflare.com/bgp-leaks-and-crypto-currencies/.

[7] J. Mattke, C. Maier, and L. Reis, "Is cryptocurrency money? Three empirical studies analyzing medium of exchange, store of value and unit of account," in *Proceedings of the 2020 on Computers and People Research Conference,* New York, NY, 2022. pp. 26-35.

[8] A. M. Bailey, B. Rettler, and C. Warmke, "Philosophy, politics, and economics of cryptocurrency I: Money without state," *Philosophy Compass*, vol. 16, no. 11, 2021.

[9] S. Frei, T. Duebendorfer, G. Ollmann, and M. May, "Understanding the web browser threat: Examination of

vulnerable online web browser populations and the insecurity iceberg," *ETH Zurich Tech Report Nr*, vol. 288, 2008.

[10] D. Vujičić, D. Jagodić, and S. Ranđić, "Blockchain technology, bitcoin, and Ethereum: A brief overview, " in *proceedings of the 17th International Symposium Infoteh-Jahorina (Infoteh)*, pp. 1-6, 2018.

[11] S. S. Kushwaha, S. Joshi, D. Singh, M. Kaur, and H. N. Lee, "Systematic review of security vulnerabilities in ethereum blockchain smart contract, *" IEEE Access*, vol. 10, pp. 6605-6621, 2022.

[12] S. S. Kushwaha, S. Joshi, D. Singh, M. Kaur, and H. - N. Lee, "Systematic review of security vulnerabilities in ethereum blockchain smart contract," *IEEE Access*, vol. 10, pp. 6605-6621, 2022.

[13] S. Quamara and A. K. Singh. "A systematic survey on security concerns in cryptocurrencies: State-of-the-art and perspectives," *Computers & Security*, vol. 113, 2022.

[14] P. Xia, H. Wang, B. Zhang, R. Ji, B. Gao, and L. Wu, et al., "Characterizing cryptocurrency exchange scams," *Computers & Security*, vol. 98, 2020.

[15] F. Fang, C. Ventre, M. Basios, L. Kanthan, D. Martinez-Rego, F. Wu, and L. Li, "Cryptocurrency trading: A comprehensive survey," *Financial Innovation*, vol. 8, no. 13, 2022.

[16] S. Erfani and M. Ahmadi, "Bitcoin security reference model: An implementation platform," in *Proceedings of the 2019 International Symposium on Signals, Circuits and Systems*, 2019. pp. 1-5.

[17] L. Lys, A. Micoulet, and M. Potop-Butucaru, "Atomic swapping bitcoins and ethers," in *Proceedings of the 38th Symposium on Reliable Distributed Systems*, 2019. pp. 372-3722.

[18] A. Christopher, K. Deniswara, and B. L. Handoko, "Forecasting cryptocurrency volatility using GARCH and ARCH model," in *Proceedings of the 6th International Conference on E-Commerce, E-Business and E-Government*, New York, NY, pp. 121-128, 2022.

[19] I. Stoepker, R. Gundlach, and S. Kapodistria, "Robustness analysis of bitcoin confirmation times," *ACM SIGMETRICS Performance Evaluation Review*, vol. 48, no. 4, 2021, pp. 20-23.

[20] A. P. Motamed and B. Bahrak, "Quantitative analysis of cryptocurrencies transaction graph," *Applied Network Science*, vol. 4, no. 131, 2019.

[21] Y. Hu, S. Wang, G. H. Tu, L. Xiao, T. Xie, and X. Lei, et al., "Security threats from bitcoin wallet smartphone applications: Vulnerabilities, attacks, and countermeasures," in *Proceedings of the Eleventh ACM Conference on Data and Application Security and Privacy (CODASPY '21)*, New York, NY, 2021. pp. 89-100.

[22] H. Chen, M. Pendleton, L. Njilla, and S. Xu, "A survey on ethereum systems security: Vulnerabilities, attacks, and defenses," *ACM Computing Surveys*, vol. 53, no. 3, pp. 1-43, 2020.

[23] O. H. Alhazmi, and Y. K. Malaiya, "Application of vulnerability discovery models to major operating systems," *IEEE Transactions on Reliability*, vol. 57, no. 1, pp. 14-22, 2008.

[24] H. Joh and Y. K. Malaiya, "Modeling skewness in vulnerability discovery," *Quality and Reliability Engineering International*, vol. 30, no. 8, pp. 1445-1459, 2014.

[25] FIRST.Org, Common Vulnerability Scoring System Version 3.1 User Guide, *While Paper*, 2022. https://www.first.org/cvss/v3.1/user-guide.

[26] A. Stango, N. R. Prasad, and D. M. Kyriazanos, "A threat analysis methodology for security evaluation and enhancement planning, " in *proceedings of the 3rd International Conference on Emerging Security Information, Systems and Technologies*, Washington, DC, pp. 262-267, 2009.

[27] I. Mkpong-Ruffin, D. Umphress, J. Hamilton, and J. Gilbert, "Quantitative software security risk assessment model," in *Proceedings of the 2007 ACM Workshop on Quality of Protection*, New York, NY, 2007. pp. 31-33.

[28] S. H. Houmb, V. N. Franqueira, and E. A. Engum, "Quantifying security risk level from cvss estimates of frequency and impact," *Journal of Systems and Software*, vol. 83, no. 9, pp. 1622-1634, 2010.

[29] F. Massacci and V. H. Nguyen, "An empirical methodology to evaluate vulnerability discovery models," *IEEE Transactions on Software Engineering*, vol. 40, no. 12, pp. 1147-1162, 2014.

[30] O. H. Alhazmi and Y. K. Malaiya, "Quantitative vulnerability assessment of systems software," *Proc. Ann. IEEE Reliability and Maintainability Symposium*, pp. 615-662, 2005.

[31] X. Wang, R. Ma, B. Li, D. Tian, and X. Wang, "E-WBM: An effort-based vulnerability discovery model," *IEEE Access*, vol. 7, pp. 44276-44292, 2019.

[32] S. G. Eick, T. L. Graves, A. F. Karr, J. Marron, and A. Mockus, "Does code decay? assessing the evidence from change management data," *IEEE Transactions on Software Engineering*, vol. 27, no. 1, pp. 1-12, 2001.

[33] O. H. Alhazmi and Y. K. Malaiya, "Prediction capabilities of vulnerability discovery models," in *RAMS '06: Proceedings of the RAMS '06. Annual Reliability and Maintainability Symposium*, Washington, DC, 2006. pp. 86-91.

[34] H. Joh, "Extended linear vulnerability discovery process," *Journal of Multimedia Information System,* vol.

4, no. 2, pp. 57-64, 2017.
[35] H. Hanif, M. H. N. Nasir, M. F. S. Razak, A. Firdaus, and N. B. Anuard, "The rise of software vulnerability: Taxonomy of software vulnerabilities detection and machine learning approaches," *Journal of Network and Computer Applications*, vol. 179, 2021.
[36] X. Li, L. Wang, Y. Xin, Y. Yang, Q. Tang, and Y. Chen, "Automated software vulnerability detection based on hybrid neural network, " *Applied Sciences*, vol. 11, no. 7, 2021.
[37] H. Joh and Y. K. Malaiya, "Periodicity in software vulnerability discovery, patching and exploitation," *International Journal of Information Security*, vol. 16, no. 6, pp. 673-690, 2017.

AUTHORS

HyunChul Joh is an associate professor at the School of Smart Industry in Kyungil University, Korea, since March 2014. He was serving as an executive director at computing information center in Kyungil university from 2018 to 2020. From 2012 to 2014, he was a GIST college laboratory instructor in division of liberal arts and sciences at Gwangju Institute of Science and Technology (GIST) in Korea. His research focuses on modeling the discovery process for software security vulnerabilities and risk metrics. Recently he had started research on A.I. and big data analysis. He received his Ph.D. and M.S. in computer science from Colorado State University, CO, USA, in 2011 and 2007, respectively. He also received a B.E. in information and communications engineering from Hankuk University of Foreign Studies in Korea, 2005.

JooYoung Lee is an associate professor at the School of K-Culture Entertainment in Kyungil University, Korea, since March 2012. From 2010 to 2011, she was a part-time instructor in Department of Fashion Design at Chung-Ang University in Korea. Her research focuses on fashion design and technology. Recently she had started research on wearable computing of fashion design. She completed her doctoral course works at department of clothing in Chung-Ang University, Korea. She received her B.A. and M.A. in Fashion Design and Technology from University of Arts London (LCF) UK, in 2004 and 2009 respectively.