A K-MAPPING BASED ENHANCED MACHINE LEARNING SECURITY ALGORITHM FOR BIG DATA MANAGEMENT AND SMART NETWORK APPLICATIONS

J. Gowthama Raja Kumaran and M. Gayathri
Department of Computer Science and Engineering, Mahendra College of Engineering, India

Abstract

In the present era, the continued use of Entries on network systems has not only improved it but also posed many threats to its security. Most networks have a simple network attack such as phishing. These can be easily let go, and even hard attack developments like Ransomware occur on a few complex systems. Most attack events are caused by Entries coming in from outside the network. But sometimes the fact that networking takes place from within has added to the fear. Thus, it is intended to enhance the security improvements of that network. In this paper a machine learning algorithm designed with an advanced artificial intelligence is proposed. This method is designed to track Entries who log in to the network and some Entries who use the networks applications first. It has a security capability of 97.87% for dealing with phishing security threats, 98.88% for cyber data protection and 99.22% for ransomware-controlled defenses. And its user approval is 96.58%. Thus, this proposed method further enhances the key features of the security possibilities. This ensures maximum protection for Big Data modules and its storage systems from hackers and other security vulnerabilities.

Keywords:

Fishing Attacks, Ransomware Attacks, Machine Learning, Network Security, Data Protection, User Authentication

1. INTRODUCTION

The Network security describes the policies and practices implemented by the network administrator to avoid and monitor unauthorized access, exploitation, modification or denial of network and network resources. Well-implemented network security prevents viruses, malware, hackers, etc. from accessing or changing secure information [1]. The first layer of network security is enabled by the username / password mechanism, which allows access only to authorized Entries with customized privileges. When a user is authenticated and grants specific computer access, the built-in firewall enables network policies, which refers to accessible user services. However, firewalls do not always detect viruses or malicious malware, which can lead to data loss. Antivirus Software or Anti-Infiltration System (IPS) is enabled to prevent viruses and / or malicious malware from entering the network [2].

Network security is sometimes confused with information security, which has a different purpose and deals with data integrity of all formats. Network analyzers and network scanners may do similar things in network management, but they work differently. Network analyzer is also sometimes called pocket analyzer or pocket sniffer [3]. By analyzing data packets, the analyzer or sniffer acts to evaluate data in traffic across the network. To help network administrators better understand data traffic and how it moves through the system, the analyst can look at the source data in a pocket, including values or settings in the title. This may include strategies such as pocket capture or intercepting data traffic [4]. The Network scanning is a different type of process that identifies active hosts and network ports for security.

A network security administrator is an individual who manages, monitors, and manages security over one or more computer networks. The network security administrator primarily ensures that a network is protected from any internal or external security threats and incidents. This person is part of the network functions and management team. Network security administrators generally design and implement network security policies throughout the network. They are the basis for advanced experience and skills in network operations, network security threats and vulnerabilities, as well as measures and strategies to mitigate them [5]. They typically work with network administrators and engineers to ensure network-level security. Some of the functions of network security administrators are as follows:

- Designing, implementing and managing nonsense network security policy
- Enable and configure security software and tools such as anti-virus, firewall, intrusion detection, and more
- Identify known and unknown network vulnerabilities and ways to combat them

The Network Scanning Tools can see how IP addresses are linked to network hosts. Tools called vulnerability scanners look for vulnerable points that could be vulnerable to computer hacking. Both of these types of tools can be used by unauthorized parties in order to attack systems [6]. In general, the use of a monitoring tool such as a network analyzer or network scanner can help network adaptation and create better security for the network and monitor threats. Network mapping and network tracking achieve two different goals, sometimes using similar techniques. Network mapping is mainly about visualizing the network, looking at relationships between different components, and broad configuration design. It evaluates network nodes and the ways in which they typically connect to the network [7].

Software limited networking monitoring application (STN monitoring application) is a type of program that monitors and manages network activities and traffic in a software limited network environment. An SDN monitoring application is similar to a standard network monitoring application, but it is designed to operate in SDN-based environments [8]. An SDN monitoring application is primarily a network management and monitoring tool that works in conjunction with a set of network operating systems and APIs. These APIs are integrated across the network across each node, device and network media, enabling network monitoring to extract and deliver specific data to the SDN monitoring application, which records and analyzes it to evaluate

overall network performance [9]. The configuration of the software defined network (STN) is designed with the separation between the control plane and the data plane (user plane). This means that the activation of network functions takes place somewhere other than the physical devices that carry the data packets to distant parts of the world. Centrally managed STN controllers dictate traffic flows and allow efficient and flexible management of the network [10]. The SDN is strongly encouraged by the Open Networking Foundation. The idea is to replace proprietary network equipment with off-the-shelf, white box switches. Linux based servers can be configured with software to create virtual environments [11]. Virtual networking, on the other hand, can refer to many processes. The traditional idea of a virtual network is one that connects wide area network components using virtual connections such as VCs, VLANs or VPNs. In today's evolving IT infrastructure, other developments are heading towards different interpretations for this period. Some vendors have developed virtual service switches or bases that integrate different services and functions [12]. Its purpose is to facilitate network infrastructure through virtualization. One of the key features of virtual networking is the separation of software and hardware. Overlay virtualization is a very common solution. Virtual networking is a form of physics that allows connections between individual devices across isolated network segments without connections [13] [14]. The flexibility of overlay networks enables multiple network traffic between virtual components in cloud computing environments. Virtual machines not only incorporate this framework, but also virtualized switches; routers, firewalls, load balancers, and other network equipment enable network functions (NFV).

2. RELATED WORKS

Zhou et al. [1] introduced cross platform identification of different network Entries. The Different types of network mapping include mapping based on a simple network management protocol, which is used to track devices such as routers, servers, printers and other hardware that supports it. Other methods include active monitoring and path analysis, a new monitoring technique that analyzes routing algorithms through Tier 3 messaging analysis between devices.

Perito et at.[3] discussed about the different privacy terms. The Network monitoring is the evaluation of a network for inactive or low-performance components to detect problems with overloads, malfunctioning servers, or other emergencies. Techniques include regular HTTP requests for servers and other status requests. In network monitoring, administrators can view items such as overtime, service availability, response time, and general reliability of the network.

Liu et al. [4] discussed about unsupervised network approach. A network switch moves data between two network devices, while a router normally directs data between two connected networks. A router is a device that some think of as a sender that sits between two networks and directs data traffic, connecting one network to another. One of the most common examples of router operation is a LAN router that connects to a small home network, often wirelessly.

Zafarani et al. [5] connecting the different Entries across the social media platform. The Network switches, on the other hand, move data efficiently from one network device to another, often by sending a localized signal to only one device, rather than broadcasting to all local devices. Many network switches are cabled, where the switches use MAC addresses or other identifiers to send signals from a dedicated port. Some switches work on multiple levels of the OSI model, and perform special types of data packet control. These are called multilayer switches.

Raad et al. [9] discussed about the profile matching issues in social media networks. The Software-defined networking is the concept of a network where, within the network configuration, the networks control plane or data transfer system is separated from other functions. This creates some abstractions in the software layers that manage the network. In contrast, network virtualization is a broader concept of changing the architectural structure of a network. An Experts talk about network virtualization based on replacing physical hardware structures with logical structures, for example, dividing or partitioning a piece of hardware into different logical functions. Whether it is related to server operations or data storage, the essential concept behind using virtual hardware pieces (sometimes called VMware) is the same.

Cortis et al. [10] who are familiar with software-defined networking and network virtualization describe software-defined networking as a mechanical or practical technique for changing the way a network is created. In a way, the software may serve the overall goal of defined networking network virtualization, which will help control the project as a kind of top-notch principle for design.

Abel et al. [11] use of STN encoding techniques is opposed to the use of virtual machines for early network construction. Others talk about software-limited networking as a type of programming tool; the way individual programming languages work to support hardware and software architecture. These are the building blocks that help build and operate large sophisticated IT systems.

3. PROPOSED METHOD

The Bandwidth monitoring plays an important role in the overall network management. Using bandwidth monitoring tools, administrators assess issues related to internal network traffic and bandwidth usage. One way to enable bandwidth monitoring for administrators is to enable them to see if they are complying with the bandwidth limits set by the Internet Service Provider (ISP) or another vendor [13]. When network bandwidth usage goes beyond these limits, it can trigger all sorts of charges and costs. Administrators use bandwidth monitoring to ensure that the application is within the correct range.

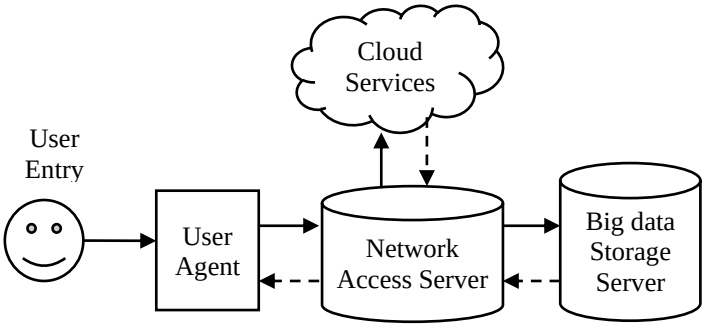

Fig.1. Proposed System Design

In general, bandwidth monitoring can help administrators detect internal sources of traffic congestion. If a network is moving slowly, not because of improper setup or malfunctioning equipment, but because of a large bandwidth balance somewhere in the network, bandwidth monitoring can help identify the problem. Information technology experts refer to this as the discovery of bandwidth pigs. Other types of network monitoring focus more on other types of issues or problems that may be affecting network performance. Some of them help detect malfunctioning equipment or other hardware and software issues within the network, while others detect intrusions, security vulnerabilities, and other potential issues from outside the network [14].

The Karnaugh mapping (K-mapping) is the process of creating a graph map that is used to reduce Boolean expression, resulting in a lower number of characters (logical operations) and variables. K-mapping is similar to drawing a fact Table.so that the position of each variable is displayed on each possible combination with other variables. In this way, common variables can be grouped together to improve the actual equation. It involves compiling expressions with integrated words and spellings, Thus, eliminating unnecessary variables and obtaining the optimal result function. K-mapping is increasingly used where the number of variables involved needs to be reduced. Similarly, you can reduce the number of operations using K-mapping. Exposure depicts a real-time situation issue or case studies. Expressions involving five to six variables are relatively difficult, but sensible, while expressions with seven or more variables are more difficult to improve (if not possible) using K-mapping. The modes of operation of this proposed system based on the K-map are as follows:

- First we can create a K-map based on the Entries of the network modules.
- Calculate the maximum number of Entries and the minimum number of Entries of the network
- Fill in the K-map data for the SOP module in 1 space reserved for the minimum user
- Fill in the K-map data for the POS module in the 0 spaces created by the maximum user
- Create rectangular groups of Entries to enter next and divide the time by that group
- Allocate the priority for each and every user entry

a	b	=	0	1
c	d		1	1

Fig.1. Mapping key calculation

where,

aa = 00; ab = 01; ac = 01; ad = 01; (Key set 1)

bb = 11; ba = 10; bc = 11; bd = 11; (Key set 2)

cc = 11; ca = 10; cb = 11; cd = 11; (Key set 3)

dd = 11; da = 10; db = 11; dc = 11; (Key set 4)

In some ways, the Security Information and Event Management (SIEM) network is different from the normal, average event record management used by businesses to monitor vulnerability and performance. However, as a kind of blanket period for a wide variety of technologies, SIEM is structured in many ways on the core principle of event record management and monitoring. The biggest difference may be the actual techniques and related aspects. Generally, SIEM is a combination of Security Information Management (SIM) and Security Event Management (SEM). This means that SIEM systems combine the general capture of the digital recording record with specific systems that view user events in the environment. For example, a SEM or security event management resource can be set up to capture different types of specific reports on account logins that occur at a specific access level, at a specific time of the day, or in a specific format used by network administrators are sense the risk or to deal with various types of management issues.

a	b	+	e	f	=	10	01
c	d		g	h		11	11

Fig.2. Mapping variables using SIEM

ae	bf	=	10	01
cg	d		11	11

Fig.3. Final K-Map security SIEM

However, a security information management system provides comprehensive reports based on aggregate data collected about network traffic. Some event experts have defined how SIEM violates the average event record tracking tool. For example, some argue that the core value of SIEM lies in very specific statements and specific features that further reveal the implications of developing a network. Event log monitoring and management can provide a general overview of what a recording process is gaining, while SIEM tools can actually provide privacy value to get down to the actual network activity and see what is happening on the network.

4. RESULTS AND DISCUSSION

This proposed algorithm is comparable with two distinct and special user identification algorithms. The first was the friend of a friend (FOAF) algorithm designed by Raad et al. [9] for user identification in social networks and the second was an identity recognition (AIR) algorithm introduced by Cortis et al. [10] for user recognition in different networks.

Table.1. Simulation Parameters

Parameter	Value
Imitation Period	2214 frames
Introduction Period	100 ms
Discovery Connectivity	0.47
Casing Period	150 ms
Tiring State Limitation	6
Unused State Limitation	5
Maximum Interference Ratio	0.8
Permission Connectivity	0.2

4.1 SECURITY CONNECTIVITY CALCULATION

At the same time, all Entries are unable to connect a network. To provide sufficient resources for all devices are too expensive and manage the network traffic without congestion is really difficult. Therefore, the networks are likely to face resource shortage issues at times. Here the blocking helps to filter the familiar devices. Blocking connectivity is the term, to allow the familiar devices and filter the random devices in a network.

For reserved channel scheme, consider that the total available channels area from which h channels are reserved. Then, the blocking Connectivity is given by

$$p_b = \sum_{s=1}^{h} p_j \quad (1)$$

where, p_j - total number of entries

The Table.2 presents the analysis of blocking Connectivity between existing FOAF, AIR and proposed K-MAP.

Table.2. Analysis of Blocking Connectivity

Entries	FOAF	AIR	K-MAP
200	77.12%	85.27%	91.97%
400	80.84%	87.23%	92.88%
600	84.87%	89.47%	94.21%
800	88.68%	91.85%	95.01%
1000	90.87%	93.57%	97.87%

In general, network systems benefit from event log monitoring because these resources and tools help administrators show more about what is happening on a particular network. An Experts point out those even small networks can truly utilize event record tracking, making management more efficient and avoiding serious security issues and other issues. When compared, existing method the proposed method achieves high blocking connectivity because the unfamiliar devices need user passkey. From Table.2, if the device not having user passkey and device passkey then it was sent under security check with network administrator

4.2 DROPPING CONNECTIVITY (FISHING THREATS) CALCULATION

The dropping connectivity is the term to drop all the devices while the network was fully occupied. The random detection of devices based on priority importance and min-max threshold of a network.

- Case 1: Average weight < Min threshold; No dropping connections
  - o If an average weight of a queue is under the min threshold, then no devices will be dropped
- Case 2: Average weight > Max threshold; All connections are dropped
  - o If an average weight of a queue is over the max threshold, then all devices will be dropped

Here the priority provides based on the passkeys. If the Entries have passkeys, then they are allowed to utilize the resources. Otherwise, they are in queue while they are getting passkeys. Assume that the process is ergodic and let $x(t)$ be its stationary distribution. The main QoS indicators of dropping probabilities defined respectively by the following ergodic limit such as,

$$D_x(t)=(\text{dropped calls under time}(x,t))/(\text{non-block arrivals under time }(x,t)) \quad (2)$$

The Table.3 presents the analysis of dropping Connectivity between existing FOAF, AIR and proposed K-MAP.

Table.3. Analysis of Dropping Connectivity

Entries	FOAF	AIR	K-MAP
200	22.88%	14.73%	8.03%
400	19.16%	12.77%	7.12%
600	15.13%	10.53%	5.79%
800	11.32%	8.15%	4.99%
1000	9.13%	6.43%	2.13%

From Table.3, if the device has both user passkey and device passkey then it was paired with the network. When compared, existing method the proposed method achieves less dropping connectivity. The familiar devices no need to get authentication from network admin

4.3 DATA PROTECTION CALCULATION

The network data protection is the amount of the data rates that are distributed to all Entries in a network. It refers the data flow rate of a communication channel. In wireless environment, data protection is an essential measurement while the data are moving without any traffic simultaneously.

$$DP = \sum \frac{P_s - average(P)}{T} \quad (3)$$

where,

DP - Data protection

P_s - Number of successful packets

$average(P)$ - Average packet size and

T - Total Time sent in delivering that amount of data

The Table.4 presents the analysis of data protection between existing FOAF, AIR and proposed K-MAP.

Table.4. Analysis of Data protection

Entries	FOAF	AIR	K-MAP
200	69.05%	85.49%	95.28%
400	69.38%	86.99%	95.87%
600	70.72%	88.1%	96.85%
800	71.86%	88.48%	98.06%
1000	72.91%	89.49%	99.22%

Event log monitoring allows administrators to relate individual cases of cross-sections or issues. For example, network administrators may see cases of RAID errors that can occur when a specific storage disk crashes. They may look at false logins or authentication logs to find out if someone is trying to gain unauthorized access. They check server performance to see if data queries are being handled effectively. They can also run certain

types of security scans and analyzes to detect vulnerabilities in the system. When compared, existing method the proposed method achieves higher data protection because all the familiar Entries utilize the higher bandwidth capacity

4.4 NETWORK DELAY CALCULATION

The time taken for a message to get transferred across a network user group from source user to end user is called End-to-end delay. This is the ratio of total hops (p) essential for routing to the total Entries (tu) in the network which is given by

$$Delay\ (\text{ms}) = p/tu \quad (4)$$

The Table.5 presents the analysis of end to end delay between existing FOAF, AIR and proposed K-MAP.

Table.5. Analysis of End-to-End Delay

Entries	FOAF	AIR	K-MAP
200	80.05%	79.49%	50.28%
400	80.38%	80.99%	50.87%
600	81.72%	82.1%	51.85%
800	82.86%	82.48%	53.06%
1000	83.91%	83.49%	54.2%

Event log tracking is an important way to help, allowing administrators to search for event patterns rather than maintain records that have never been analyzed. This applies to authentication, storage processes, data requests, and more. Instead of simple passive recording of events, event log monitoring helps when any bad events occur on the network. All the familiar Entries access the network, and then the interrupts are extremely low. Then the network speed between the end devices is very easy. If the speed of a network is increased then the delay was automatically decreased. Sometimes it was 0. When compared, existing method the proposed method achieves less end-to-end delay

4.5 BANDWIDTH UTILIZATION CALCULATION

At a given time, the highest quantity of data transferred over a connection is referred the bandwidth. The percentage of consumed bandwidth off the total available bandwidth is called the bandwidth utilization.

$$\text{BU (\%)} = (\text{Total messages transmitted and received})/(\text{speed of transmission}) \times 100 \quad (5)$$

The Table.6 presents the analysis of bandwidth utilization between existing FOAF, AIR and proposed K-MAP.

Table.6 Analysis of Bandwidth Utilization

Entries	FOAF	AIR	K-MAP
200	72.88%	74.73%	98.03%
400	69.16%	72.77%	97.12%
600	65.13%	70.53%	95.79%
800	61.32%	68.15%	94.99%
1000	59.13%	66.43%	92.13%

From Table.6, familiar devices with both user and device passes, that not required any authentication from admin. When compared, existing method the proposed method achieves more bandwidth utilization because all the slots occupied with the familiar devices. The unused bandwidth allocated for other new devices while they are getting approved by admin.

5. CONCLUSION

The detection of different users in different configuration modules can be a daunting task on networks with the most complex settings. This will increase the waiting time of the users of the network. Users' usage time is calculated based on this time. Its progress will be much greater in classifying them based on certain network modules that provide services according to the needs of the users. Thus, classifying those blocks would be a very difficult process. These actions make the process of allowing users even more complicated. Thus, most trips are diverted from the network. These are the points to security vulnerability and problems with validation timing. This method proposed allows different users to log in quickly and easily with the given K-Map token. But here the tokens cannot be logged in by the user who has not traveled properly. So, problems just arise from the user space. The admin is not required to monitor any work. This method completes the user analysis, user testing and allowable work. Its special feature is that it takes less time. So, the user can be easily identified and allowed inside the network.

REFERENCES

[1] X. Zhou, X. Liang, H. Zhang and Y. Ma, "Cross-Platform Identification of Anonymous Identical Entries in Multiple Social Media Networks", *IEEE Transactions on Knowledge and Data Engineering*, Vol. 28, No. 2, pp. 411 424, 2016.

[2] M.M. Mostafa, "More Than Words: Social Networks' Text Mining for Consumer Brand Sentiments", *Expert Systems with Applications*, Vol. 40, No. 10, pp. 4241-4251, 2013.

[3] D. Perito, C. Castelluccia, M.A. Kaafar and P. Manils, "How Unique and Traceable are Usernames?'", *Proceedings of International Symposium on Privacy Enhancing Technologies*, pp. 1-17, 2011.

[4] J. Liu, F. Zhang, X. Song, Y.I. Song, C.Y. Lin and H.W. Hon, "What's in a Name?: An Unsupervised Approach to Link Entries across Communities", *Proceedings of International Conference on Web Search Data Mining*, pp. 495-504, 2013.

[5] R. Zafarani and H. Liu, "Connecting Entries Across Social Media Sites: A Behavioral-Modeling Approach", *Proceedings of International Conference on Knowledge Discovery Data Mining*, pp. 41-49, 2013.

[6] O. Goga, D. Perito, H. Lei, R. Teixeira and R. Sommer, "Large-Scale Correlation of Accounts Across Social Networks", Technical Report, University of California, pp. 1-170, 2013.

[7] N.Ye, L. Zhao, L. Dong, G. Bian, E. Liu and G.J. Clapworthy, "User Identification based on Multiple Attribute Decision Making in Social Networks", *China Communications*, Vol. 10, No. 12, pp. 37-49, 2013.

[8] X.H. Han, L.H. Wang, S.J. Xu, G.Q. Liu and D.W. Zhao, "Linking Social Network Accounts by Modeling user Spatiotemporal Habits", *Proceedings of IEEE International*

Conference on Intelligence Secure Information, pp. 19-24, 2017.

[9] E. Raad, R. Chbeir and A. Dipanda, "User Profile Matching in Social Networks", *Proceedings of IEEE International Conference on Network Based Information System*, pp. 297-304, 2010.

[10] K. Cortis, S. Scerri, I. Rivera and S. Handschuh, "An Ontology-Based Technique for Online Profile Resolution", *Proceedings of IEEE International Conference on Social Informatics*, pp. 284-298, 2013.

[11] F. Abel, E. Herder, G.J. Houben, N. Henze and D. Krause, "Cross-System User Modeling and Personalization on the social Web", *Proceedings of IEEE International Conference on User Model and User-Adaptive Interaction*, Vol. 23, No. 2-3, pp. 190-209, 2013.

[12] Y. Li, Y. Peng, W. Ji, Z. Zhang and Q. Xu, "User Identification based on Display Names across Online Social Networks", *IEEE Access*, Vol. 5, pp. 17342 17353, 2017.

[13] A. Esfandyari, M. Zignani, S. Gaito and G.P. Rossi, "User Identification across Online Social Networks in Practice: Pitfalls and Solutions", *Journal of Information Science*, Vol. 44, No. 3, pp. 377-391, 2016.

[14] Y. Li, Y. Peng, Z. Zhang, H. Yin and Q. Xu, "Matching User Accounts Across Social Networks based on Username and Display Name", *Proceedings of IEEE International Conference on World Wide Web*, pp. 1-23, 2018.