IIETA International Information and Engineering Technology Association
Advancing the World of Information and Engineering

Mathematical Modelling of Engineering Problems
Vol. 9, No. 2, April, 2022, pp. 397-404
Journal homepage: http://iieta.org/journals/mmep

Network Anomaly Detection in 5G Networks

Atta-ur Rahman[1*], Maqsood Mahmud[2], Tahir Iqbal[2], Linah Saraireh[2], Hisham Kholidy[3], Mohammed Gollapalli[4], Dhiaa Musleh[1], Fahd Alhaidari[5], Dakheel Almoqbil[6], Mohammed Imran Basheer Ahmed[7]

[1] Department of Computer Science, College of Computer Science and Information Technology (CCSIT), Imam Abdulrahman Bin Faisal University, P.O. Box 1982, Dammam 31441, Saudi Arabia
[2] College of Business Administration, Imam Abdulrahman Bin Faisal University, P.O. Box 1982, Dammam 31441, Saudi Arabia
[3] Network & Computer Security Department, College of Engineering, Sunypoly Polytech Institute, Utica, New York NA 13502, USA
[4] Department of Computer Information System, College of Computer Science and Information Technology, Imam Abdulrahman Bin Faisal University, P.O. Box 1982, Dammam 31441, Saudi Arabia
[5] Saudi Aramco Cybersecurity Chair, Department of Networks and Communications, College of Computer Science and Information Technology (CCSIT), Imam Abdulrahman Bin Faisal University, P.O. Box 1982, Dammam 31441, Saudi Arabia
[6] Department of Networks and Communications, College of Computer Science and Information Technology (CCSIT), Imam Abdulrahman Bin Faisal University, P.O. Box 1982, Dammam 31441, Saudi Arabia
[7] Department of Computer Engineering, College of Computer Science and Information Technology (CCSIT), Imam Abdulrahman Bin Faisal University, P.O. Box 1982, Dammam 31441, Saudi Arabia

Corresponding Author Email: aaurrahman@iau.edu.sa

https://doi.org/10.18280/mmep.090213

Received: 13 December 2021
Accepted: 6 April 2022

Keywords:
network anomaly detection, 5G, mmWave, MIMO, machine learning, KNN, K-prototype

ABSTRACT

On the telecommunications front, 5G is the fifth-generation technology standard for broadband cellular networks, which is a replacement for the 4G networks used by most current phones. Hundreds of businesses, organizations, and governments suffer from cyberattacks that compromise sensitive information in which 5G is one of them. Those breaches of the data would not have occurred if there is a way to detect strange behaviors in a 5G network, and this is what this paper presenting. Network Anomaly Detection (NAD) in 5G is a way to observe the network constantly to detect any unusual behavior. However, it is not that straightforward and rather a complex process due to huge, continuous, and stochastic network traffic patterns. In the literature, several approaches and methods have been employed for anomaly detection as well as prediction. This paper illustrates state-of-the-art method to proposed achieve the NAD. For instance, pattern based, machine learning based, ensemble learning based, user intention based, and some integrated methods have been surveyed and analyzed. KNN and K-prototype algorithm were tested together on the dataset and compared with integrated approach. The integrated approach outperformed with respect to the KNN and K-prototype methods. As a conclusion, forecasting of analyst detection of cyber events is presented as a final method for future anomaly prediction.

1. INTRODUCTION

5G technology deals with various technologies like spectrum sharing, Device to device (D2D), Ultra dense network (UDN) and massive Multiple input Multiple output (MIMO). User's experience has high data rate, low latency, energy efficiency, spectrum efficiency, UDN, coverage dependability as well as mobility in the wireless communication network are some of the needs that 5G can meet [1]. Additionally, 5G is expected to provide a wide range of additional services, such as ultra-reliable, massive machine types of information exchange, increased mobile broadband and communication with reduced latency depending on the specific use-case. Interference mitigation, mode selection, device synchronization, security and quality of service are all issues that emerge when using device to device Long Term Evaluation/Advance LTE/LTE-A system [2] LTE is vulnerable to Radio Frequency (RF) jamming, spoofing, and sniffing, and examining the various physical layer vulnerabilities that could disrupt next-generation vital communication networks [3]. A wide range of machine learning algorithms can be used to attack anomaly detection challenges in 5G. In addition, most of them can be used in accordance with the VNO security policies specified in the proposed cyber defense architecture [4]. Network security in 5G is always a critical component in all business fields especially with the widespread cybercrimes in recent years. There are a lot of tools that can detect the pre-known attacks such as antivirus software, firewalls, and honeypots, but these tools are not enough because of the rapid expansion in the unprecedented attack methods and behaviors. This is mainly because of overwhelmingly increasing number of Internet of Things (IoT) devices and applications that can potentially induce the vulnerability. Furthermore, enhanced network

heterogeneity, intensified use of virtualization technologies and distributed architectures make it nearly impossible to cope the situation [5]. Moreover, existing (say for 4G) Network Anomaly Detection (NAD) technique are not that effective for its 5G counterpart [4y]. Situation demands more robust and effective techniques with reduced complexity. This work presents the NAD in 5G networks which provides a high level of security by applying a persistent observation of network to detect any abnormal behavior from potential attackers that they signed up as legit users. There are a lot of methods that can be used to apply the NAD and most of them focus on analyzing the behaviors of the network traffic to apply the detection process. Behavioral analysis and prediction are among the hottest areas of research in many fields of studies like consumer marketing, social network, marketing, information security and many others. Rahman et al. [6] proposed a user behavior classification and prediction using a neuro-fuzzy approach. The analysis was made over the users logs duly obtained by a proxy server continuously monitoring the network user activities related to machine, ethernet and web browsing. Based on the company's rules the restricted actions were mined out of the logs after due preprocessing. A Gaussian Radial Basis Function Neural Network (GRBF-NN) [7, 8] along with a Fuzzy Rule Based System [9, 10] was used to learn the examples provided by an added Differential Evolution (DE) block [11]. The scheme also incorporated user 360 feedback to jointly decide the user's behavior. Similar work was carried out by Rahman et al. [12, 13] by using linear regression and FRBS, respectively. A hybrid approach for network anomaly detection and user behavior prediction was presented by Vadgaonkar [14]. The study investigated a machine learning based approach. The proposed approached claimed as significant accuracy over a predefined dataset in an offline mode. The network performance tools are widely used to continuously analyze the network traffic. However, the perspectives may be different. For example, the overall network performance is represented in terms of Network Availability, CPU and memory utilization, Traffic rate, number of errors and discards and Wide Area Network (WAN) performance. It is worth mentioning that the network anomalies may affect the network availability in terms of denial-of-service attacks (DoS) and Distributed DoS (DDoS) attacks, CPU and memory utilization may be overwhelmed by bots and memory resident viruses attacks duly penetrated by the network. Traffic rate can be compromised by the eavesdroppers and network attackers by inducing more traffic in the regular one. Collectively these factors can result in increase in errors in the data traffic and discards of requests. Thus, the network anomalies encompass all the factors and as a result overall network performance can become vulnerable. Several network analysis tools are available in market. However, their effectiveness is subjective and do not fulfill the requirements for all types of networks. Such as the tool developed for the earlier generations networks may not be that effective due to the variation in standard, data rates and application areas. It should rather be more customized per the organizational needs and type of data and its security [15]. In contrast, Network intrusion detection systems (IDS) have their own pros and cons on the other hand. For instance, the advantages include timely identification of intrusions that results in less security incidents. Similarly, as far as the disadvantages are concerned, the false positive incidents may result in denial of service (DoS) attack as a false alarm. Furthermore, if the concerned organization has poor bandwidth and network resources, it could result in slowing down the overall network traffic [16]. Moreover, IDS are signature based that requires a prior information about the signatures to be detected or identified. However, intrusion is one possible way to compromise the network behavior while there could be more [16].

Rest of the paper is organized as follows: the detailed description of the NAD in 5G background is presented in section 2, six different methods to apply NAD are reviewed section 3, and some methods that can be used in predicting the potential attacks before it occurs in section 4 and section 5 concludes the paper.

2. BACKGROUND

5G networks are digital cellular networks. The service area is further categorized into acute geographical cells. The 5G wireless devices in a cell communicate by radio waves with a local antenna array. Low power automated transceiver (transmitter and receiver) in the cell is one of feature of 5G. The local antennas are further connected to transmission electronics connected to switching centers in the telephone network and routers for Internet access by high-bandwidth optical fiber or wireless backhaul connections. 5G can support up to a million devices per square kilometer, while 4G supports only one tenth of that capacity [17]. The millimeter waves (mmWaves) have a shorter range than microwaves. This is one of the reasons that the cells are constraint to a minute size [18].

2.1 5G fundamentals

Massive multiple input multiple output (mMIMO) was implemented in 4G as early as 2016 and typically comprised of 32 to 128 tiny antennas per cell. Depending on the frequency and arrangement, it can boost performance of the communication system by four to ten times. 5G utilizes sub-6 GHz band and mmWave frequencies and the mMIMO technology which can be created, deployed, and utilized [19, 20]. Several data streams are simultaneously sent out into the world. Using a process known as beamforming, a base station computer will constantly evaluate the optimum path for radio waves to reach each wireless device and organize many antennas to work together as phased arrays to create beams of millimeter waves to reach the device [18]. It is effective as well fault tolerant to failure of any antenna element.

2.2 5G security

5G security include authentication, availability, data confidentiality, integrity, and non-repudiation, etc. Network anomaly detection is a prominent problem in 5G security in the global discussion [21]. Before going through the methods to apply the anomaly detection, three main questions regarding NAD will be discussed in this section, what is NAD, how it works and differences between NAD and other IDS tools.

A variety of approaches have been used to detect and monitor malicious and anomalies through networks users' behaviors. One of these effective approaches is NAD which is continuous monitoring of the network to detect unusual traffics or movements. It is a complementary technology to systems that detect security threats based on packet signatures [22]. NAD is an integral part of Network Behavior Analysis

(NBA), which enhances the security of a network by monitoring the traffic of an active network and collecting data from many data points and devices to give a comprehensive analysis. NAD approach is used in a variety of network and security fields including (i) Log analysis (ii) Packet inspection systems (iii) Flow monitoring systems and (iv) Route analytics [23]. NAD watching every traffic happening in the network, gathering data from many points to support offline analysis. NAD performs this task by grabbing the suspicious packets in anticipation. Then performs the analysis using various approaches like statistical hypothesis test (SHT), machine learning etc. [24]. After establishing an indicator for normal network traffic, the NAD approach monitors network activity and flags anonymous, new, or uncommon patterns that may indicate the presence of a threat. Also, it records any changing that might appears in the bandwidth, or the protocol used. NAD is particularly effective to explore new malware and zero-day attacks. NAD tools continuously monitor the network to find any malicious threat actors. Instead of relying on the perimeter, endpoint, and firewall security systems which can only find security threats that pass-through areas of the network where they are installed, NAD systems watch the entire network for threat actors. It is particularly helpful in detecting threat actors in two cases whereas IDS signature-based cannot detect the new attacks (whose signature is unknown will be treated as normal traffic), and when the threat traffic is encrypted. The main function of a network behavior analysis detection is to minimize the effort and time expended in detecting and solving network issues. It is thus an enhancement to protect the network along with antivirus software, spyware detection tools and software/hardware firewalls [25].

3. NAD METHODS IN 5G

There are various ways that can be used to provide a high level of network security in 5G, this section describes six methods that can be applied to achieve the network's behavior anomaly detection process.

3.1 Pattern-based random walk

Random Walk term refers to the process of randomization to construct a path of n steps (points). For the sake of clarity, assume there is an agent that needs to traverse a path that consists of some points and in each step (point), the agent chooses its direction randomly. Figure 1-a shows a path that the agent would take from the source to the destination without any restricted rules and it may be stuck in a loop (points A and D). Moreover, to predict the walker's behavior in a simple Random Walk we need to calculate the probability of being at a specific point after walking for n steps. This theory used to simulate the random behaviors in computer networks, and it describes only the typical computer system model. In addition, more features should be added to the Random walk to conduct the simulation in more complex computer systems and networks, and that presents the Self-avoiding Walk which is a new type of random walk that simulates the activities in computer networks with some additional rules. The main rule is to limit the walker behavior by disallowing it to revisit the visited points before as described in Figure 1-b, and that would make predicting the walker behavior more accurate. It is important to analyze any network traffic in 5G to identify it to either a pre-known threat then store it in the attack database or classify it as normal traffic. But this is not enough since there is always a new generation of unprecedented threats that must be identified using a pre-known attack database. For the previous reason, Pattern-based Random Walk which is considered as a new customization of Random Walk was presented to detect these kinds of threats and analyze them accurately. The main advantage of this approach is outlier detection that can probably leads to an anomaly in the network. Pattern theory is one of the Self-avoiding Walk that used to determine whether patterns are matched together or not. Every income traffic must be checked and sampled, if it is not detected as one of the pre-known behaviors, the sample should be flagged, and an analysis process should be done on this flagged sample. After that, a self-explained graph will be generated that does not need any more information, and this graph should be compared with some similar pre-known patterns to decide whether there are matched patterns or not. If any, then system should take the proper decision [26].

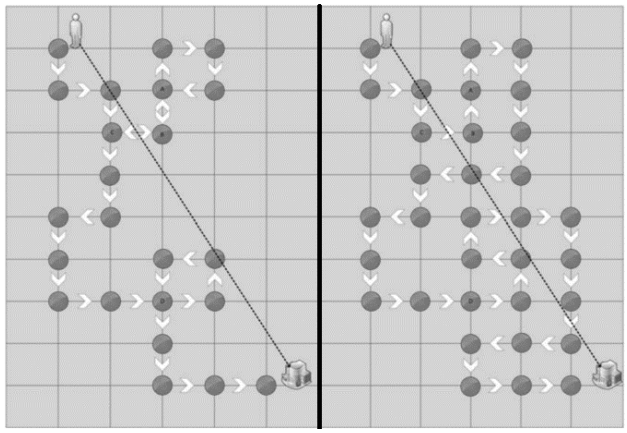

Figure 1. (a) Simple random walk (b) Self-avoiding walk

3.2 Immune network and density peak

It is an unsupervised NAD method in 5G, it was presented to identify the unknown malicious attack behavior from any network flow in the Internet environment. There are some models and algorithms which includes the Artificial Immune Network (aiNet) model, the clustering algorithm based on density peak (CDP), the clustering labeling algorithm (CLA), and the flow anomaly detection algorithm (FAD). It works by obtaining network flows as a starting point. Next, it selects the common features of the network flows since that will help in recognizing anomaly or malicious behaviors easily. Then, clusters the network flows which includes two phases, the coarse-grained phase, and the fine-grained phase. The coarse-grained phase includes using the aiNet model to cluster the samples from the given dataset and this process work by analyzing and filtering the samples. The fine-grained phase includes using the CDP algorithm to cluster the output of the aiNet clustering and this process work by refining the cluster centroids from the coarse-grained phase. After that, it uses the CLA to label the resulted clusters as normal or abnormal detectors and uses them to distinguish between malicious and benign behaviors. The cluster labeling step is very significant since its results will affect its performance. Finally, providing the functionality of anomaly detection for the network flows using FAD since that will play a role in enhancing the network security [27]. It can be deduced from the discussion that there

are certain similarities and difference among these algorithms. As far as similarity is concerned, their ultimate task is to detect the abnormalities in the particular network traffic. While differences include in such a way that one algorithm is best in one way and the other is in another way. For instance, the aiNet implies a coarse-grained clustering algorithm to extract the abstract internal patterns of network flows, and the CDP signifies a fine-grained clustering algorithm to obtain more precise cluster number and cluster centroids. A study [27] proposed a NAD approach coined as Artificial Immune network and Density peak (ADAID), that combines aiNet and CDP for better results in terms of improved accuracy rates and reduced false alarm rates.

3.3 Hybrid data optimization based on ML algorithms

Intrusion detection system (IDS) can also detect anomaly behaviors in the 5G network. This work produced an effective IDS by using data optimization which consists of two main sections: extract samples from dataset and select common features from these data, named DO-IDS. The Isolation Forest (iForest) is used in extract random sample to eliminate outliers in dataset and reduce the negative effect of unbalanced data, after generating the sample genetic algorithm (GA) it used to optimize the sampling ratio, and the Random Forest (RF) classifier as the evaluation criteria to gain the optimal training dataset. After optimizing the extracted sample feature selection is used to search for features that reflect the difference between anomalous behaviors and normal behaviors and delete unrelated features to enhance the detection performance, GA and RF are used again to obtain the optimal feature subset. More specifically GA and RF complements each other in the weak areas to counter the issue of unbalanced dataset examples and fine tuning the significant features. After the optimum training dataset and the feature subset are selected, those will be occupied into the classifier training phase which uses RF algorithm. The whole process is shown in Figure 2 [28].

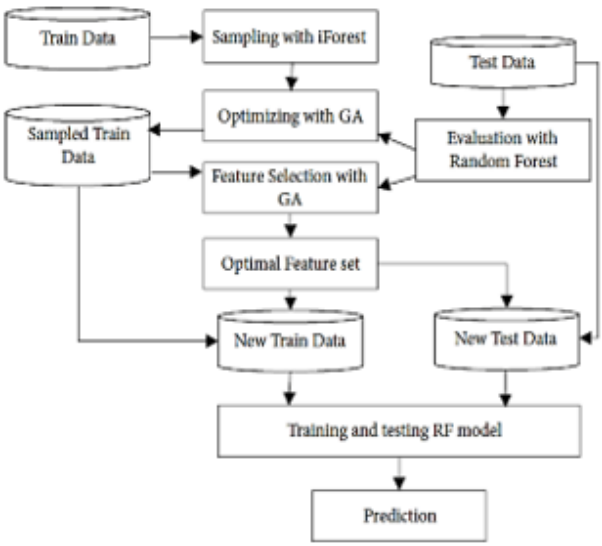

Figure 2. Process of DO IDS

3.4 Multi-view ensemble learning

This work proposes a malware detection scheme based on ensemble learning which is a concept that been used a lot in recent years. This concept can also be used for malware detection in 5G. It is multiple models, such as classifiers or experts combined to solve a particular computational intelligence. Ensemble learning is primarily used to enhance prediction, classification performance of a model, or reduce the likelihood of errors and low accuracy results. problem [29-32]. In this work, three ensemble learning methods, and different ensemble strategies are employed to fuse multi-view features which they were:

Byte n-grams: It is syntactic patterns features that cover the entire binary executable program and without explicit semantic information.

Opcode n-grams: It is explicit semantic information were used as features for detecting unknown malware it is extracted only from the code section of portable executable (PE) file.

Format feature: It is explicit semantic information extracted from PE header, section header, import, and resource sections.

The benefits of the integrated learning methods over the other are that they adequately combine their strengths to complement the weak areas among them to overall increase the precision rate in NAD. Those methods partially capture the different information between malicious and non-malicious applications. Also, they are complementary for each other in extracting features. The procedure of this experiment was as follows: (i) First is to find a data representation which captures the information within a given feature view, obtain and select the optimal parameter of each feature subset. (ii) Design the scheme to incorporate three feature views. The schema uses each of feature views to construct different classifiers which are then combined using three ensemble methods. (iii) Two datasets are used to evaluate new malware detection performance and the generalization performance of the proposed scheme. Finally, the results show that the proposed scheme does enhance the performance of new malware detection [33].

4. RESULTS AND DISCUSSION

This section discusses the results related to the network anomaly detection in 5G. This section benchmark KNN, K-Prototype and Integrated method. The integrated method is one of the effective methods for an anomaly detection in massive system logs especially mMIMO in 5G networks. Logs are used for recording systems' detailed operations; these logs are useful for detecting anomalies within the systems. Anomaly means unusual behavior can affect the system negatively. There are problems in detecting those anomalies with logs: (i). Previous knowledge is needed in anomaly detection, which is not possible for it to expose new anomalies. (ii). The large size of logs might require more complex computations. A new integrated method is used for anomaly detection is proposed to solve the previous problems; this method merges clustering method K-prototype, k-NN classification method, and novel clustering-filtering-refinement. The approach is like those conducted for user behavior classification and prediction by using the proxy servers log as presented in the studies [5, 11, 12].

4.1 Feature elicitation

The analysis of the logs' abnormal behavior characteristics and logs' data are done using login activity and session

statistics. Ten features were extracted based on previously mentioned aspects.

4.2 Clustering

K-prototype clustering algorithm was used to create clusters out of the extracted features. This step is critical to classify which event was normal and which was not. It also helps in extracting the abnormal events as candidates for further examination which decreases complex computations.

4.3 Refinement

This step is for analyzing abnormal candidates in deeper manner to identify if they were anomalies or not by performing two steps: 1. Distance measurement and it has two types: a) from each data dot to the center of the clusters. b) from each data dot and its closest neighbors. 2. Measure local and global anomaly degree by applying k-NN classification method.

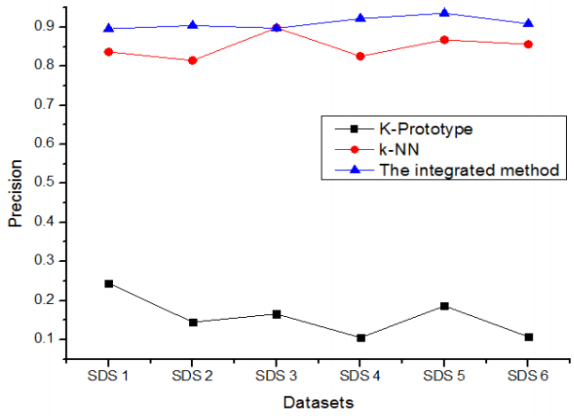

Figure 3. Precision rate evaluation

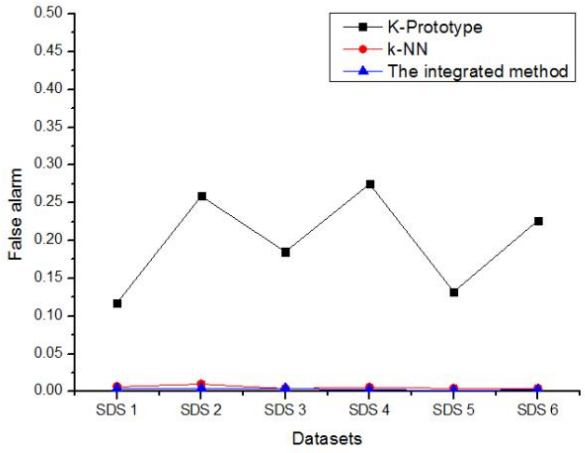

Figure 4. False alarm evaluation

After the experiment, it turns out that the integrated method is more accurate and stable than stand-alone K-prototype and k-NN algorithms in the context of 5G (Figure 3). Without refinement step, the K-prototype had the worst performance, so it is not the preferred option in anomaly detection since it has the highest false alarm rate (lowest precision) as shown in Figure 4. The k-NN algorithm has close results from integrated in terms of precision. In terms of time, K-prototype has better results than k-NN algorithm, and it has close results as integrated method per the Figure 5. Finally, the experiment summarizes that integrated method is the best in detecting anomaly for massive-size logs in the perspective of 5G [34].

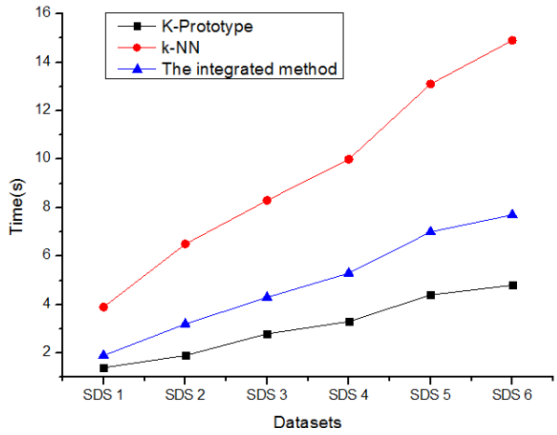

Figure 5. Computational complexity evaluation

4.4 User intention-based traffic dependence analysis for NAD in 5G

This work describes a Causal Relation Miner (CR-Miner) framework that detects anomaly in 5G network traffic based on user's activities [35]. The framework emphasizes in the mining the relationship in the outbound network traffic by means of a traffic dependency graph (TDG). Consequently, a novel breadth first search algorithm is proposed to traverse this graph to infer the potential dependencies. This relation is valuable to detect the anomaly and to know whether the anomaly happened due to software failure or malicious code. CR-Miner algorithm can detect user intention-based traffic with accurate results. It also can detect HTTP-based spyware. Bayesian can be used for anomaly detection, but it does not work for complicated and diverse traffics. The proposed solution is used for detecting dependencies among users' inputs (clicking on a link) and systems' events (HTTP requests). The power of the algorithm it does not require previous knowledge of user's intentions. Also, the algorithm can detect vagabond events which they mean the reaction of the application based on user's actions; those vagabond events can be identified as anomalous events. The presented work has three main steps:

4.4.1 Find dependencies

Finding user behavior traffic dependencies is difficult since we browsers generate actions without users' involvement. Hence, it is important to identify explicit users' actions by a) Dependencies in Browser Traffic. b) Applications and threat model [35].

4.4.2 Building TDG

Detecting anomalous based on dependencies is the aim of CR-Miner and it can be achieved through Traffic-dependency graph (TDG). TDG can be built incrementally by feeding anonymous dependencies of traffics' events. TDG is used to exclude previous knowledge bases [35].

4.4.3 Security analysis

Since CR-Miner has data gathering and data analysis phases;

the critical phase is data gathering (integrity), so CR-Miner should protect against: Forgery attacks (data modification) and Piggybacking attacks (gaining unauthorized access through pre-established another user's session). It turns out CR-Miner can protect against attacks [35].

4.4.4 Evaluation

CR-Miner works fine with legit websites, time-efficient, 99.6% accurate, and it can detect real-world spyware [35].

Cyberattacks in 5G networks can jeopardize institutions, businesses, and governments. That is why they initiated defense mechanisms against those attacks which they are called time-series forecasting, such as automated intrusion detection/prevention, honeypot, and network telescopes. This study suggests analyst-detected forecasting which has dataset that does not rely on automated systems, contains cyber events for seven years, and validated by analysts, so it does not have false positive alarms. The strength of this study is its limitation since it has a unique dataset validated by analysts and with minimum automated systems involvement. Bayesian State Space Model (BSSM) was used for forecasting, the results showed the model anticipated one week in advance events with high precision. The main finding is the number of next week's cyberattacks is predictable based on the number of previous week's cyberattacks in 5G, this prediction will help improving threat awareness. It may empower organization for better planning and protection of their domains in more effective way [36, 37].

4.5 Security implications in 5G

With threats and attack, it will be a major challenge the rapid and improved expansion of Internet-based devices and edge-based computing. They will not necessarily be connected to a central network. With so many devices connected through a networked edge environment, any one of them can become the weakest link between them in the security path [38-41]. To keep communicating and the devices are as secure as possible following steps can be considered [42].

- Everything connected to the enterprise ecosystem needs to be identified, criticality rated, and their state confirmed. Then, all requests for access to network resources will need to be verified, validated, and authenticated.
- Security should conjointly support elastic, edge-to-edge hybrid systems combining well-tried ancient ways with new approaches. While network segmentation may be a well-tried technique for holding cybersecurity risks and protect sensitive resources, recent ways might not be best suited to a 5G world. New segmentation ways can navigate native and remote resources that blend segments that organizations might or might not have management. IT groups can value a way to manage the complexity of multiple co-managed systems as they implement 5G networks and public cloud services.
- Sharing threat intelligence, correlating event information, and supporting automatic incident response would require security technologies to be deeply integrated [43]. This will need the event and adoption of a comprehensive, fabric-based security design. Machine learning [44], computing and automation [45] are going to be key to fast decision-making, thereby closing the gap between detection and mitigation.

These are just a few of the security implications resulting from the adoption and deployment of 5G networks. But that's simply the beginning of the impact of this new era of networking and computing. Security will also need to address the following scenarios [42]:

- Automated network application lifecycle management would require security tools to not solely be high activity however additionally extremely adjustive to confirm that constant innovation includes consistent protection. It will additionally need organizations to transition from a DevOps model to a DevSecOps model to confirm that security is integrated directly into the event strategy.
- Support for cloud-optimized distributed network applications would force security to maneuver seamlessly between and across fully completely different network ecosystems whereas not losing track of workflows or dropping security utility.
- Digital transformation can generate Brobdingnagian amounts of latest knowledge, most of which can be encrypted. Encrypted data currently constitutes more than 70% of network traffic. That share can solely grow as cryptography is employed to shield knowledge moving through open network environments. This will need superior security tools in IoT and different edge devices that may examine encrypted traffic at each speed and scale.
- New methods, like network slicing, can change organizations to a lot of with efficiency consume resources moving through large information settings.
- This will additionally need segmentation and edge-based micro-segmentation to safeguard crucial resources whereas uninflected them from open and fewer secure environment.

5. CONCLUSION

5G sets a better standard for commandability and speed on phones and other devices, as well as on the internet. Many of the security issues that plague today's 4G, 3G, and 2G networks have been addressed in 5G's design. Mutual authentication and identity protection are among the new measures that have been implemented. An unparalleled level of network and service security may be achieved with the introduction of 5G. Log Anomaly Detection may be savior. Logs are a terrific way to keep track of how your system is running. With the rapid development of technology in 5G as increases of using mobile devices which make societies open and connected to each other all day which makes the world a global village. Moreover, Intrusion detection has attracted the attention of many researchers in identifying the risks can affect network users. In this work, NAD was presented in detailed description which is a high level of security in 5G by applying a persistent observation of network to detect any abnormal behavior from potential attackers that they signed up as a legit user. Primarily, the current study can be useful for the scholars and researchers in the field of network intrusion and anomaly detection, as a head start in 5G. On the dataset, the KNN and K-prototype algorithms were put to the test jointly and compared to an integrated method. The integrated approach outperforms the KNN and K-prototype techniques in terms of predictive power. The higher security of 5G will lead to its higher suitability globally.

ACKNOWLEDGMENT

The authors would like to acknowledge SAUDI ARAMCO Cybersecurity Chair, Imam Abdulrahman Bin Faisal University (IAU), Dammam, Saudi Arabia for funding this research.

REFERENCES

[1] Shafi, M., Jha, R.K., Sabraj, M. (2020). A survey on security issues of 5G NR: Perspective of artificial dust and artificial rain. Journal of Network and Computer Applications, 160: 102597. https://doi.org/10.1016/j.jnca.2020.102597

[2] Noura, M., Nordin, R. (2016). A survey on interference management for device-to-device (D2D) communication and its challenges in 5G networks. Journal of Network and Computer Applications, 71: 130-150. https://doi.org/10.1016/j.jnca.2016.04.021

[3] Lichtman, M., Jover, R.P., Labib, M., Rao, R., Marojevic, V., Reed, J.H. (2016). LTE/LTE-A jamming, spoofing, and sniffing: Threat assessment and mitigation. IEEE Communications Magazine, 54(4): 54-61. https://doi.org/10.1109/MCOM.2016.7452266

[4] Maimó, L.F., Gómez, Á.L.P., Clemente, F.J.G., Pérez, M.G., Pérez, G.M. (2018). A self-adaptive deep learning-based system for anomaly detection in 5G networks. IEEE Access, 6: 7700-7712. https://doi.org/10.1109/ACCESS.2018.2803446

[5] Doan, M., Zhang, Z. (2020). Deep learning in 5G wireless networks - anomaly detections. 2020 29th Wireless and Optical Communications Conference (WOCC), pp. 1-6. https://doi.org/10.1109/WOCC48579.2020.9114924

[6] Rahman, A., Dash, S., Luhach, A.K., Chilamkurti, N., Baek, S., Nam Y. (2019). A neuro-fuzzy approach for user behavior classification and prediction. Journal of Cloud Computing, 8(17): 1-15. https://doi.org/10.1186/s13677-019-0144-9

[7] Rahman, A. (2020). GRBF-NN based ambient aware realtime adaptive communication in DVB-S2. Journal of Ambient Intelligence and Humanized Computing, 12: 1-11. https://doi.org/10.1007/s12652-020-02174-w

[8] Rahman, A., Qureshi, I.M., Malik, A.N., Naseem, M.T. (2014). A real time adaptive resource allocation scheme for OFDM systems using GRBF-neural networks and fuzzy rule base system. International Arab Journal of Information Technology (IAJIT), 11(6): 593-601.

[9] Rahman, A. (2019). Optimum information embedding in digital watermarking. Journal of Intelligent and Fuzzy Systems, 37(1): 553-564. https://doi.org/10.3233/JIFS-162405

[10] Rahman, A., Qureshi, I.M., Malik, A.N., Naseem M.T. (2016). QoS and rate enhancement in DVB-S2 using fuzzy rule base system. Journal of Intelligent & Fuzzy Systems (JIFS), 30(1): 801-810. https://doi.org/10.3233/IFS-151802

[11] Rahman, A. (2019). Memetic computing based numerical solution to Troesch problem. Journal of Intelligent and Fuzzy Systems, 37(1): 1545-1554. https://doi.org/10.3233/JIFS-18579

[12] Rahman, A, AlRashed, S.A., Abraham, A. (2017). User behavior classification and prediction using FRBS and linear regression. Journal of Information Assurance and Security, 12(3): 86-93.

[13] Rahman, A., Zaidi, D.N., Salam, M.H., Jamil, S. (2013). User behavior classification using fuzzy rule based system. 13th International Conference on Hybrid Intelligent Systems (HIS'13), pp. 118-123.

[14] Vadgaonkar, P.H. (2020). Network anomaly detection and user behavior analysis using machine learning. International Journal of Computer Applications, 175(13): 47-53.

[15] Saqib, N.A., Abdus Salam, A., Rahman, A., Dash, S. (2021). Reviewing risks and vulnerabilities in web 2.0 for matching security considerations in web 3.0. Journal of Discrete Mathematical Sciences and Cryptography, 24(3): 1-17. https://doi.org/10.1080/09720529.2020.1857903

[16] Wang, J., Rossell, D., Cassandras, C.G., Paschalidis, I.C. (2013). Network anomaly detection: A survey and comparative analysis of stochastic and deterministic methods. 52nd IEEE Conference on Decision and Control, pp. 182-187. https://doi.org/10.1109/CDC.2013.6759879

[17] Khraisat, A., Gondal, I., Vamplew, P., Kamuzzaman, J. (2019). Survey of intrusion detection systems: Techniques, datasets and challenges. Cybersecur, 2(1): 1-22. https://doi.org/10.1186/s42400-019-0038-7

[18] Rappaport, T.S., Sun, S., Mayzus, R., Zhao, H., Azar, Y., Wang, K., Wong, G.N., Schulz, J.K., Samimi, M., Gutierrez, F. (2013). Millimeter Wave Mobile Communications for 5G Cellular: It Will Work! IEEE Access, 1: 335-349. https://doi.org/10.1109/ACCESS.2013.2260813

[19] Nordrum, A., Clark, K. (2017). Everything you need to know about 5G. IEEE Spectrum magazine. Institute of Electrical and Electronic Engineers. Archived from the original on January 20, 2019. Retrieved January 23, 2019. ""I am crazy about Massive MIMO," Kitihara of Softbank ordering 1,000's of Massive MIMO bases". wirelessone.news. Retrieved March 27, 2020.

[20] Yu, H., Lee, H., Jeon, H. (2017). What is 5G? Emerging 5G Mobile Services and Network Requirements. Sustainability, 9(10): 1848. https://doi.org/10.3390/su9101848.

[21] Bjornson, E., Van der Perre, L., Buzzi, S., Larsson, E.G. (2019). Massive MIMO in sub-6 GHz and mmWave: Physical, practical, and use-case differences. IEEE Wireless Communications, 26(2): 100-8. https://doi.org/10.1109/MWC.2018.1800140

[22] Park, J.H., Rathore, S., Singh, S.K., Salim, M.M., Azzaoui A.E., Kim T.W., Pan Y., Park J.H. (2021). A Comprehensive Survey on Core Technologies and Services for 5G Security: Taxonomies, Issues, and Solutions. Human-centric Computing and Information Sciences, 11(3). https://doi.org/10.22967/HCIS.2021.11.003

[23] Hein, D. (2019). Network Behavior Analysis and Anomaly Detection: The Basics. Retrieved from https://solutionsreview.com/network-monitoring/network-behavior-analysis-and-anomaly-detection-the-basics/, accessed on 12 November 2019.

[24] What is network behavior anomaly detection (NBAD)? - Definition from WhatIs.com. (2019). Retrieved from https://searchsecurity.techtarget.com/definition/network-behavior-anomaly-detection, accessed on 10 November

2019.
[25] Reese, B. (2019). Intrusion detection systems vs. network behavior analysis: Which do you need? Retrieved from https://www.networkworld.com/article/2346145/intrusion-detection-systems-vs--network-behavior-analysis--which-do-you-need-.html, accessed on 10 November 2019.
[26] Nia, M.A., Atani, R.E., Fabian, B., Babulak, E. (2016). On detecting unidentified network traffic using pattern-based random walk. Security and Communication Networks, 9(16): 3509-3526. https://doi.org/10.1002/sec.1557
[27] Shi, Y., Shen, H. (2020). Anomaly detection for network flow using immune network and density peak. International Journal of Network Security, 22(2): 337-346. https://doi.org/10.6633/IJNS.202003 22(2).18
[28] Ren, J., Guo, J., Qian, W., Yuan, H., Hao, X., Jingjing, H. (2019). Building an effective intrusion detection system by using hybrid data optimization based on machine learning algorithms. Security and Communication Networks. https://doi.org/10.1155/2019/7130868
[29] Rahman, A., Sultan, K., Naseer, I., Majeed, R., Musleh, D. (2021). Supervised machine learning-based prediction of COVID-19. Computers, Materials and Continua, 69(1): 21-34. https://doi.org/10.32604/cmc.2021.013453
[30] Alotaibi, S.M., Rahman, A., Basheer, M.I., Khan, M.A. (2021). Ensemble machine learning based identification of pediatric epilepsy. Computers, Materials & Continua, 68(1): 149-165. https://doi.org/10.32604/cmc.2021.015976
[31] Zagrouba, R., Khan, M.A., Rahman, A., Saleem, M.A., Mushtaq, M.F. (2021). Modelling and simulation of COVID-19 outbreak prediction using supervised machine learning. Computers, Materials & Continua, 66(3): 397-2407.
[32] Alhaidari, F., Almotiri, S.H., Ghamdi, M.A., Khan, M.A., Rehman, A., Abbas, S., Khan, K.M., Rahman, A. (2021) Intelligent software-defined network for cognitive routing optimization using deep extreme learning machine approach. Computers, Materials & Continua, 67(1): 1269-1285. https://doi.org/10.32604/cmc.2021.013303
[33] Bai, J., Wang, J. (2019). Improving malware detection using multi-view ensemble learning. Security and Communication Networks, 9(17): 4227-4241. https://doi.org/10.1002/sec.1600
[34] Liu, Z., Qin, T., Guan, X., Jiang, H., Wang, C. (2019). An integrated method for anomaly detection from massive system logs. IEEE Access, 6: 30602-30611. https://doi.org/10.1109/ACCESS.2018.2843336
[35] Zhang, H., Banick, W., Yao, D., Ramakrishnan, N. (2012). User intention-based traffic dependence analysis for anomaly detection. 2012 IEEE Symposium on Security and Privacy Workshops, pp. 104-112. https://doi.org/10.1109/SPW.2012.15
[36] Bakdash, J.Z., Hutchinson, S., Zaroukian, E.G., Marusich, L.R., Thirumuruganathan, S., Sample, C., Das, G. (2018). Malware in the future? Forecasting of analyst detection of cyber events. Journal of Cybersecurity, 4(1): tyy007. https://doi.org/10.1093/cybsec/tyy007
[37] Rehman, S.U., Mahmud, M., Rahman, A., Haq, I.U., Safdar, M. (2021). Information Security in Business: A Bibliometric Analysis of the 100 Top Cited Articles. Library Philosophy and Practice, 1-49.
[38] Rahman, A., Dash, S., Ahmad, M., Iqbal, T. (2021). Mobile cloud computing: A green perspective. Intelligent Notes in Networks and Systems, 185: 523-533.
[39] Alhaidari, F., Rahman, A., Zagrouba, R. (2020). Cloud of things: Architecture, applications and challenges. Journal of Ambient Intelligence and Humanized Computing, pp. 1-19. https://doi.org/10.1007/s12652-020-02448-3
[40] Ahmad, M., Qadir, M.A., Rahman, A. (2020). Enhanced query processing over semantic cache for cloud based relational databases. Journal of Ambient Intelligence and Humanized Computing, pp. 1-19. https://doi.org/10.1007/s12652-020-01943-x
[41] Dash, S., Biswas, S., Banerjee, D., Rahman, A. (2019). Edge and fog computing in healthcare – A review. Scalable Computing: Practice and Experience, 20(2): 191-206. https://doi.org/10.12694/scpe.v20i2.1504
[42] Nguyen-Duy, J. (2019). What Are the Security Implications for 5G and IoT? Fortinet.
[43] Zagrouba, R., AlAbdullatif, A., AlAjaji, K., Al-Serhani, N., Alhaidari, F., Almuhaideb, A., Rahman, A. (2021). Authenblue: A new authentication protocol for the industrial Internet of Things. Computers, Materials & Continua, 67(1): 1103-1119. https://doi.org/10.32604/cmc.2021.014035
[44] Ahmed, M.I.B., Rahman, A., Farooqui, M., Alamoudi, F., Baageel, R., Alqarni, A. (2021). Early identification of COVID-19 using dynamic fuzzy rule based system. Mathematical Modelling of Engineering Problems, 8(5): 805-812. https://doi.org/10.18280/mmep.080517
[45] Dilawari, A., Khan, M.U.G., Al-Otaibi, Y.D., Rehman, Z., Rahman, A., Nam, Y. (2021). Natural language description of videos for smart surveillance. Applied Sciences, 11(9): 3730. https://doi.org/10.3390/app11093730