Novel Intrusion Detection System for 5G[1]

Maksim Iavich[a], Avtandili Gagnidze[b], Giorgi Iashvili[a], Sergei Simonov[b], Razvan Bocu[c]

[a] *Caucasus University, 1 Paata Saakadze St, Tbilisi, 0132, Georgia*
[b] *East European University, 4 Shatili St., 0178 Tbilisi, Georgia*
[c] *Transilvania University of Brasov, 500036 Brasov, Romania*

Abstract

The telecommunication industry is majorly transforming towards 5G networks. It needs to satisfy the needs of the new and existing users. The users and the customers need much better quality of the corresponding services and they need the corresponding security in order to secure the transmitting data and other services. Therefore, the mobile leading networks must provide much better quality of an experience and security, and they must improve the performance for the services they provide. Novel services envisioned by 5G, new networking, service deployment, the new processing technologies and the storage is required. The mentioned technologies will involve the new security problems for the 5G systems. The world scientists are seriously working on analyzing the 5G security. The researchers have identified the existing problems of 5G systems. Our analysis illustrates the basics reasons of security problems in 5G. The researchers have found the vulnerabilities in 5G, which give the attackers opportunity to integrate the malicious code and to run it. MiTM, MNmap and Battery drain attacks can be successfully implemented on 5G.

Our paper analyzes an existing security problems of 5G. As the result, we offer the new Intrusion Detection System using machine-learning approaches. The paper offers an integration of this intrusion detection systems into an existing 5G architecture. We offer the methodology and a pseudo code of this system.

Keywords

5G, security, intrusion detection systems, ids.

1. Introduction

The scale of the traffic needed for the wireless networks and the quantity of mobile and IoT devices are enhancing very fast, because of the different factors. The telecommunication industry is majorly transforming towards 5G networks and it has to satisfy the requirements of the target users. The 5G wireless networks must provide very high data rates and much higher coverage by means of dense base station deployment. It must have high capacity, much better QoS and it must have the very small latency. All this will involve the new technologies, and, as expected, these technologies cause new security problems for the 5G systems. The scientists are analyzing the security of 5G technology and they have successfully discovered some of the security problems.

5G will deal with the critical infrastructure, which requires rather high security level to ensure security of the critical infrastructure and of the whole society. For the instance, a security attack on the hypothetical online power supply system is critical for the electronic and electrical systems. As we can see, it is very important to analyze and emphasize the existing security problems in 5G networks and to offer the novel solution, which will make them secure. We have analyzed the difference between 5G and 4G security architecture

These architectures are pretty much similar. The network nodes in 5G and 4G that are needed for security, and the communication links have very much in common. The security mechanisms of these systems can be grouped like following: One group contains the mechanisms used to offer users the safe access to the different services and to make the system safe against the vulnerabilities connected with an air interface, that can be located between some device and the radio node. The second group contains other safety mechanisms for the network access. These mechanisms are needed to transfer the signaling and user's data from the nodes of radio to the core network. The security architecture of 5G and 4G is shown on Fig.1

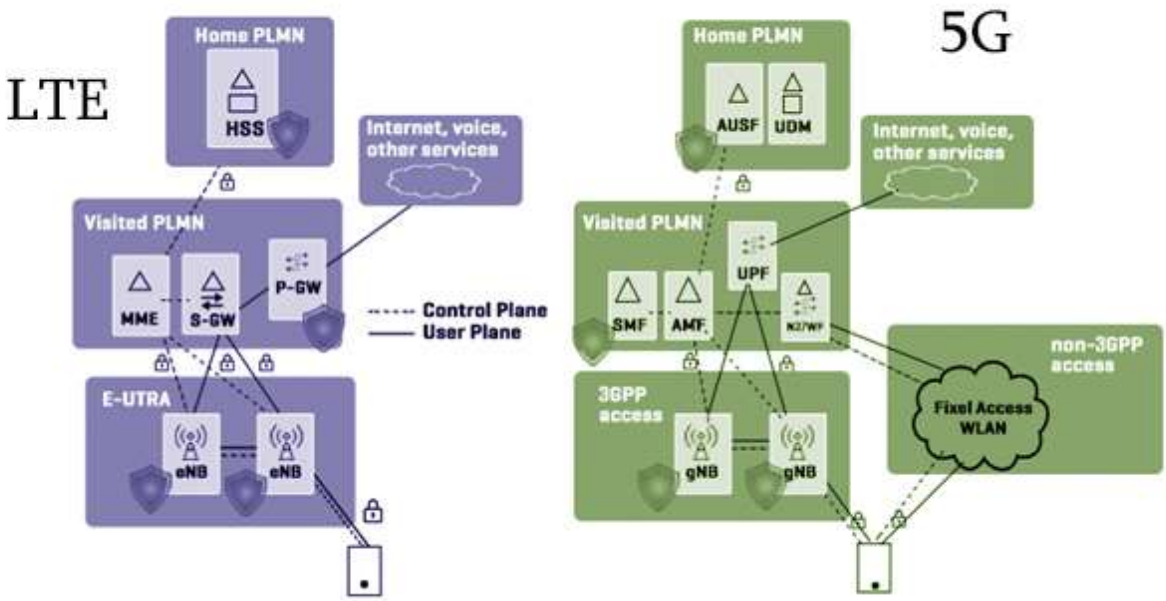

Figure 1: Security architecture of 5G and 4G

2. Intrusion detection systems

Nowadays, ten years ago and even at the very beginning of "IT revolution", cyber-defense systems held special place and that's logical: nobody really wants to be hacked. Security specialists have used several different tools to provide users protection. Some of these tools are: antiviruses, firewalls, IDS (intrusion detection systems) and IPS (intrusion prevention systems). Let's take a brief look at each of them. Antivirus is a software which is installed on the host machine and is checking system processes for a suspicious activity. Firewall is a software which is, the most often, installed between two

networks and analyses a network traffic blocking the suspicious one. As for the intrusion detection systems, these pieces of software are targeted on analyzing network traffic and checking if there is any suspicious activity in there. Intrusion detection systems are different from Firewall. Firewall controls traffic between two networks but if someone performs attack inside the network, the firewall will not be able to identify it. For a purpose of fortifying our network, we can add an intrusion detection system inside that network, which will sniff for traffic and send out an alert that something is wrong. By itself, an intrusion prevention system does not perform any action, it only can notify us about malicious traffic and log the incident to the file. The one who takes care "about" bad traffic is an IPS [], which stands for intrusion prevention system. This piece of software receives the alarm and perform corresponding action whether it will be dropping a packet or letting it to the quarantine. In addition, it is not necessary to have two separate devices for IDS and IPS, both of them can be integrated in the router. Therefore, as we can see, IDS and IPS are just a specific software, which help us extending our security. As for the IDS role in the 5G infrastructure, it can help us detecting DoS and software defined (brute force for example) attacks on the fly, even before the data is delivered to its destination. Of course, there will be some problems during the implementation, because such a system will require computing power, which is not really common if we speak about non-core segment of the system. A lot of concurrent traffic has to be analyzed at the one unit of time, that's why a deployment of the IDS is so important.

3. Novel Security mechanism

Our idea is to integrate the new server software with the IDS embedded into it to the 5G base station. The visualization of the approach is illustrated on the figure 2.

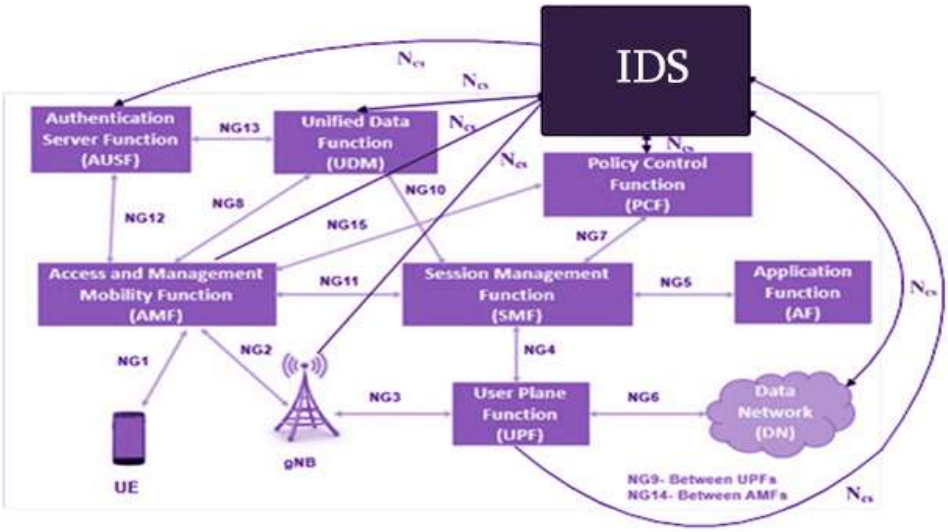

Figure 2: Secure architecture

Modern 5G intrusion detection systems are well known for using a KDD99 [4-7] dataset and machine learning algorithms [8-11]. A few words about this technology: Machine learning is a methodology of receiving processed output based on the input, for example: we supply the model (model is specific

neural network which processes the input and gives us an output) with images of animals and want it to guess if what is the animal it has received. Machine learning algorithms use datasets [12,13]. Datasets are files that contain specific information about something, animals for example. Each line must contain parameters of this animal and, the most important, the conclusion about what that animal is. Using datasets, we can train our model to recognize and process information. An example of the dataset can be found on the figure 3 (each and every line contains an information about the attack and a name of the attack itself)

Figure 3: KDD99 dataset example

To improve over existing approach, we offer adding a CIC-IDS datasets along with KDD99, which will help us to get protection not only from KDD99 attacks, but also from different denial of service attacks. Our intrusion detection system is trained to use two CIC-IDS style datasets. The first dataset contains the information about the following DOS[12,13] attacks: 'MSSQL', 'LDAP', 'Syn', 'NetBIOS', 'UDPLag', 'UDP' and weights 341 MB, let us call it DOS1. As for the second database, it only contains the information about the reflected 'Portmap' attack and its size is 89 MB, let us call it DOS2.

Also, KDD99 was separated in two different datasets, one for training and one for testing. The training one stores 90% of the information while the test dataset includes the remaining 10%. DOS1 and DOS2 datasets were also divided into two separate datasets, each of them. The training dataset stores 80% of the data, while the test dataset contains remaining 20%. Both of these datasets are split using the same method, which was chosen because of the best accuracy results after the training. The model itself is being trained with each dataset separately. In the case of the KDD99 dataset, model accuracy is 0.96670491252916389, in the case of DOS1 is 0.9942894736842107 and in the case of DOS2 is 0.9998966703182065.

When the training process is finished, the system asks for input, which is being extracted from the network sniffer. First, the input is checked for the attacks contained in the KDD99 dataset. If corresponding attack pattern is found, it sends a signal to the intrusion protection system. In the case when finding pattern for the attack fails, the system then trying to check data relative to the attack patterns which are stored in the DOS1 dataset. If an attack pattern is identified, a program sends a signal to the intrusion protection system. The same process is applied to the DOS2 dataset, if the attack pattern was not

determined. If the attack pattern is failed to be identified, the intrusion detection system gives us an output that the traffic is legitimate and goes after processing the next input.

The algorithmic core of the intrusion detection system is explained by the following pseudo code.

The pseudo code of the idea is shown below:

```
Class NOVELIDS():
        Initialization of variables
        Def __init__(self):
                Preprocessing the date
        Def new_model(self, type):
                Creating the model
                Return the new model
        Def train (self, model_type):
                Training the model
                Return the trained model
        Def test (self, model, type):
                Testing
                    Measuring accuracy
                Return the score
        Def predict(self, data):
                Predicting data by means of the module
                Return the result
        Def accuracy(self, type):
                Return NOVELIDS.test ()
IDS_real = NOVELIDS (df1)     # making the dos prediction model
IDS_real2 = NOVELIDS (df2)   # making KDD99 prediction model
IF IDS_real2 (df) == 'KDD_attack'
        Passing the information
Elif IDS_concrette.predict(df) == ‘DOS’:
        Passing the information

Else:
        Print “not vulnerable”
Processing the new traffic
```

4. Relevance of the research

Our analyses have shown us the concrete reasons, which can be the security concern for 5G networks [1-3]. These reasons are:

- 5G system has a very large exposure to the different software attacks and it has a lot of entry points for the attackers, because of the virtualization 5G systems are mostly based on the software mechanisms. The software security attacks can be implemented on 5G.
- Because of the new functionality, the parts of some network equipment and some network functions are very sensitive to the different attacks. Different base stations and the key management functions in the network are sensitive to the attacks.

- Because network operators rely on concrete suppliers, the new attacks can be implemented.
- The great number of IT applications need 5G network, it makes 5G sensitive to attacks, which influence on integrity and availability.
- Because of large number of devices DoS and DDoS attacks are much more relevant.
- Because of the network slicing, attackers can attack the different slices.

5. Experiments

We have created a small test laboratory using 20 RASPBERRY PI devices, we have also used 20 modems with sim cards. We installed our IDS software on the server. Attacks replicated by us are the following: BACK, LAND, POD, SMURF, NEPTUNE, NMAP, TEARDROP, BUFFER_OVERFLOW, LOADMODULE, ROOTKIT FTP_WRITE, GUESS_PASSWD, IMAP, MULTIHOP, PHF, SPY, PROBE, IPSWEEP, PORTSWEEP, SATAN, MSSQL, Portmap and LDAP.

It must be emphasized that 5G system can be vulnerable to these types of attacks.

We have simulated the attacks and wrote the traffic down using a network sniffer, then all the traffic was examined. parsed all the parameters that are relevant for our KDD99 and DOS [8, 9] samples were parsed, using the Python programming language. The output was transformed to the format of the original datasets. After this, we all this information was passed to our IDS system which performed the analysis.

Attack type	Number of attacks	Identified attacks
BACK	50	48
LAND	50	100
NEPTUNE	50	98
POD	50	100
SMURF	50	84
TEARDROP	50	82
BUFFER_OVERFLOW	50	76
FTP_WRITE	50	86
LOADMODULE	50	91
ROOTKIT	50	62
GUESS_PASSWD	50	100
MULTIHOP	50	91
SPY	50	51
PROBE	50	98
IPSWEEP	50	92
NMAP	50	95
PORTSWEEP	50	98
SATAN	50	82
LDAP	50	81
MSSQL	50	99
Portmap	50	97

These results prove that the IDS is rather useful and it can be used as the prototype version of the future real-world IDS system.

6. Results

As the result, we received the novel IDS oriented on 5G attacks. The IDS is trained using machine learning algorithms. It is trained using KDD99 dataset and the DOS/DDOS attacks dataset.

The IDS is trained using the attacks vectors, which are vulnerable for 5G. These attacks vectors were identified based on our research.

7. Discussion

After conducting the corresponding experiments, we have identified that IDS is doing its job rather well, but, unfortunately, it still has some efficiency problems. We are working on improving the efficiency, optimization and creating our own training patterns for IDS.

8. Conclusion

5G networks give us more bandwidth and speed, which can also be a huge downside: Think about what hackers can do. DoS, DDoS, reflected DDoS and other volumetric attacks will become even stronger. Also, some critical infrastructure will hardly depend on the 5G, smart cars, hospitals, power plants, this means that any attack can on these can be critical: taking a digital hostage in the face of the hospital`s inner network is a joke no more.

5G is a modern, rapid technology that requires corresponding security systems to protect users and the critical infrastructure. Using neural networks and machine learning to create a smart and flexible software can help us in attacks mitigation and prevention

The offered IDS is aiming to provide a good level of security and accuracy, but it still has some efficiency problems. Certain work must be conducted in order to achieve the secure 5G services.

9. Acknowledgments

The work was finances by Shota Rustaveli National Science Foundation and was conducted in the frame of CARYS-19-121 grant.

10. References

[1] Huawei 5G Security White Paper, https://www-file.huawei.com/-/media/corporate/pdf/trust-center/huawei-5g-security-white-paper-4th.pdf, 2019.
[2] 5G Americas: The evolution of Security in 5G, https://www.5gamericas.org/files/4715/6450/22-67/5G_Security_White_Paper_07-26-19_FINAL.pdf, 2019.
[3] Report on EU coordinated risk assessment of 5G, https://ec.europa.eu/commission/presscorner/detail/en/IP_19_6049, 2019.
[4] Kumar V., Das A.K., Sinha D. (2020) Statistical Analysis of the UNSW-NB15 Dataset for Intrusion Detection. In: Das A., Nayak J., Naik B., Pati S., Pelusi D. (eds) Computational Intelligence in Pattern Recognition. Advances in Intelligent Systems and Computing, vol 999. Springer, Singapore. https://doi.org/10.1007/978-981-13-9042-5_24.

[5] Ullah I., Mahmoud Q.H. (2020) A Scheme for Generating a Dataset for Anomalous Activity Detection in IoT Networks. In: Goutte C., Zhu X. (eds) Advances in Artificial Intelligence. Canadian AI 2020. Lecture Notes in Computer Science, vol 12109. Springer, Cham. https://doi.org/10.1007/978-3-030-47358-7_52.

[6] Luo C., Wang L., Lu H. (2018) Analysis of LSTM-RNN Based on Attack Type of KDD-99 Dataset. In: Sun X., Pan Z., Bertino E. (eds) Cloud Computing and Security. ICCCS 2018. Lecture Notes in Computer Science, vol 11063. Springer, Cham. https://doi.org/10.1007/978-3-030-00006-6_29.

[7] Y. I. . Ibrahim, F. M. . Remo, and Y. S. Younis, “Design a Hybrid Algorithm Based on Tabu Search and Misuse Detection for Intrusion Dataset (KDD99 10%)”, *JUBPAS*, vol. 27, no. 5, pp. 337-351, Dec. 2019.

[8] Jacobs, M., Pradier, M.F., McCoy, T.H. *et al.* How machine-learning recommendations influence clinician treatment selections: the example of the antidepressant selection. *Transl Psychiatry* 11, 108 (2021). https://doi.org/10.1038/s41398-021-01224-x.

[9] Sipper, M., Moore, J.H. Conservation machine learning: a case study of random forests. *Sci Rep* 11, 3629 (2021). https://doi.org/10.1038/s41598-021-83247-4.

[10] Garg, S., Sinha, S., Kar, A.K. and Mani, M. (2021), "A review of machine learning applications in human resource management", International Journal of Productivity and Performance Management, Vol. ahead-of-print No. ahead-of-print. https://doi.org/10.1108/IJPPM-08-2020-0427

[11] Mullainathan, Sendhil, and Jann Spiess. 2017. "Machine Learning: An Applied Econometric Approach." *Journal of Economic Perspectives*, 31 (2): 87-106.

[12] Zhang, Y., Ling, C. A strategy to apply machine learning to small datasets in materials science. *npj Comput Mater* 4, 25 (2018). https://doi.org/10.1038/s41524-018-0081-z

[13] Kohli, M.D., Summers, R.M. & Geis, J. Medical Image Data and Datasets in the Era of Machine Learning—Whitepaper from the 2016 C-MIMI Meeting Dataset Session. *J Digit Imaging* 30, 392–399 (2017). https://doi.org/10.1007/s10278-017-9976-3

[14] Monowar H. Bhuyan, H. J. Kashyap, D. K. Bhattacharyya, J. K. Kalita, Detecting Distributed Denial of Service Attacks: Methods, Tools and Future Directions, *The Computer Journal*, Volume 57, Issue 4, April 2014, Pages 537–556, https://doi.org/10.1093/comjnl/bxt031

[15] V. Gupta, S. Krishnamurthy and M. Faloutsos, "Denial of service attacks at the MAC layer in wireless ad hoc networks," *MILCOM 2002. Proceedings*, Anaheim, CA, USA, 2002, pp. 1118-1123 vol.2, doi: 10.1109/MILCOM.2002.1179634.