Towards a SIP-based DDoS Attack to the 4G Network*

Nicola Cibin, Meriem Guerar, Alessio Merlo, Mauro Migliardi, Luca Verderame

Abstract Cellular networks are fundamental infrastructures nowadays, so that any communication problem could affect the user in different ways, from accessing social networks up to personal safety issues. In this work, we explore the feasibility of carrying out a DDoS attack to the Home Subscriber Server of the 4G network through non-3GPP access, i.e. access points that are not specified by the Third Generation Partnership Project, in particular using the SIP register procedure. A previous study on a DDoS attack to UMTS Network showed that injecting 2500 requests in every 4.7s time window is possible to reduce the HLR capability to serve legitimate requests by 93%, and that such an attack can be mounted with a few hundred devices. A limit to that attacking approach is that we would require mobile devices that need to connect to an eNodeB (cellular base station). Instead, in the approach proposed in this paper we carry out a preliminary study to explore the possibility of using devices that are generically connected to the Internet: this means that the population of devices that can be leveraged to mount the attack is wider than in the first case; furthermore, the constraint of having legitimate SIM modules is removed.

1 Introduction

Nowadays, the cellular network is fundamental to support most of the user's everyday activities, spanning from personal to work, as well as for public safety. Furthermore, with the advent of IoT, autonomous driving, and many other innovations, it is

Meriem Guerar, Alessio Merlo, Luca Verderame
DIBRIS - University of Genoa, Via Dodecaneso 35, 16146, Genoa, Italy. e-mail: {meriem.guerar, alessio, luca.verderame}@dibris.unige.it

Nicola Cibin, Mauro Migliardi
DEI - University of Padua, Via Gradenigo, 6/b, 35131, Padua, Italy. e-mail: {nicola.cibin,mauro.migliardi}@unipd.it

* This work is supported by the University of Padua.

increasingly important that data transmission in the cellular network is fast, stable, and reliable. In fact, any malfunction of the network could cause the isolation of thousands of mobile devices and their corresponding services.

The entire architecture of the 3G cellular network is standardized by the 3GPP (Third Generation Partnership Project) collaboration agreement [1]. The choice of a collaborative project was essential in order to achieve the interoperability of the network between different operators and countries. As a consequence, most of the 3G cellular networks in the world are structured in the same way [2].

Basically, in all of them the two main components dedicated to authorizing the access to the network are the Home Location Register (HLR) and the Home Subscriber Server (HSS) databases. The HLR database stores all information relating to subscribers, while the HSS (Home Subscriber Server) is, to simplify, the concatenation of an HLR and an AuC (Authentication Center). The AuC part of the HSS deals with the generation of security keys for the validation of users who request to connect to the network [9].

With the transition from the 3G to the 4G network, new components have been added to the network to allow access from the so-called non-3GPP access points. At the same time, the databases have been converted from HLR to HSS to grant novel ways to access the network, such as the one based on the SIP protocol. With the addition of these new components, Wi-Fi and wireline connections are part of the 4G standard and therefore allow access to the cellular network without the need for a user to connect to an eNodeB (E-UTRAN Node B, base station to which the user must connect to interact with the 4G network) [2]. As it often happens, the combination of well-established and novel technologies may increase the attack surface, thereby leading to new threats that are enabled by the unexpected interaction of components that were not originally designed to cooperate.

One of the most disruptive threats to cellular networks is Denial of Service (DoS). A DoS attack consists of an attempt to deliberately exhaust the resources of an IT system that provides a service to clients. In such an attack, affecting several computing platforms (e.g., [5]), an attacker can keep querying the service to force the IT system to allocate resources to satisfy the incoming requests. As a consequence, the IT system may become unavailable to legitimate users.

A Distributed DoS (DDos) attack is a more disruptive version of a DoS as it is carried out by several attacking entities at the same time and in a distributed way. From a defensive standpoint, this makes much more difficult to distinguish legitimate sources of requests from malicious ones.

SIP-based environments have always been rather weak against DoS and DDoS, albeit some solutions to counteract such attacks have been proposed (e.g., [6]) in the past. However, such attacks are still an open issue for SIP.

In previous research studies, it has already been presented how it is possible to make the HLR databases unreachable through DDoS by leveraging a few thousands of devices equipped with a valid SIM [7]. Furthermore, novel studies [8, 17] proved that the same attack can be carried out with two hundred devices or slightly more than a double number of SIM-less devices. The effectiveness of this type of attacks is partially hampered by the resource-constrained nature of mobile devices, which

does not allow the continuous sending of requests for connection to the network (attachment procedure).

In this work, we carry out a preliminary study on the feasibility of a DDoS attack to the 4G network through non-3GPP access, i.e., access points that do not meet the Third Generation Partnership Project specification. In particular, we focus on the SIP register procedure. The idea is that the exploitation of the SIP register procedure could allow to carry out a more disruptive DDoS attack to the 4G network w.r.t. the SIM or SIM-less based DDoS attacks than the one envisioned in [7], [17] and [8].

It is important to notice that, in this preliminary study, some simplistic assumptions have been made; as an example, given the difficulty of obtaining real HLR and AuC systems to measure their capabilities, we assume the numbers presented in [7] and [8] as the benchmark.

The paper is structured as follows: in Section 2 we provide a very brief introduction to some of the components of the 4G Network Architecture and the SIP register procedure; in Section 3 we describe our attack methodology while in Section 4 we discuss some countermeasures and related work; in Section 5 we describe the results of our simulations in some different scenarios; finally in Section 6 we provide some concluding remarks.

2 The 4G Architecture in a nutshell

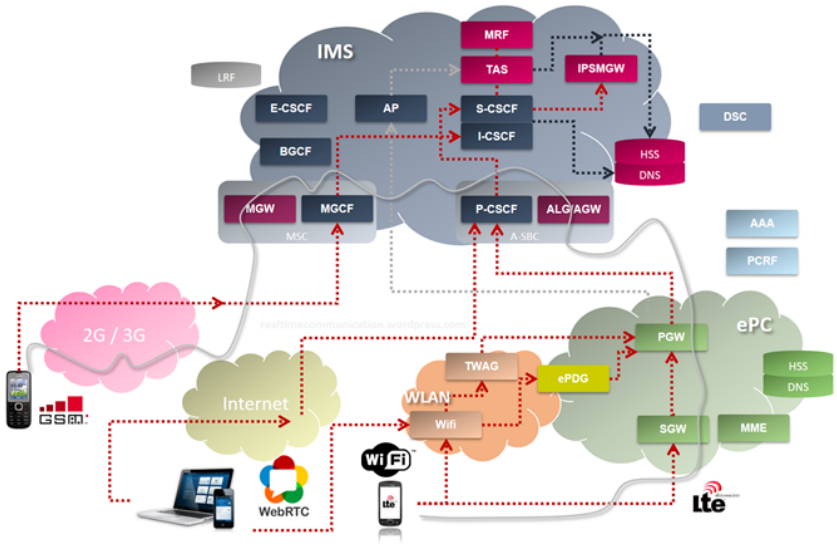

Fig. 1 The 4G architecture.

Figure 1 shows the architecture of the 4G network [2].

The IP Multimedia Subsystem (IMS) network architecture is standardized by the 3GPP not topologically (i.e., as a set of nodes), but functionally (i.e., as a set of functions that interact through standardized interfaces). IMS goal is to foster the access to the network of multimedia and voice applications from wireless and wireline terminals.

The ePC (evolved Packet Core) has been inserted to bring together data and voice traffic on the 4G network on the basis of users' IP addresses. It provides the non-3GPP access defined as Voice Over WiFi.

Our preliminary study focuses on the feasibility of an attack mounted by devices that have non-3GPP access in the form of a generic Internet connection. Hence, the components that are most interesting for our work are the ones shortly described below.

- *P-CSCF* (Proxy Call Session Control Function) [3]. The Proxy-CSCF is the SIP proxy server that allows the user access to the mobile network. All traffic to and from the user must pass through the P-CSCF. As a consequence, this is the entry node for traffic incoming from the Internet. The A-SBC (Access Session Border Controller) is implemented in this module. Among other functions, this module deals with the security of the communication channel (e.g., DDoS attack prevention through blacklisting, see discussion in Section 4).
- *I-CSCF* (Interrogating Call Session Control Function) [3]. Interrogating-CSCF is a SIP proxy server that takes care of i) querying the HSS, ii) identifying the HSS through the Subscriber Location Function in the presence of multiple databases, in order to know and forward subscriber information to other components of the IMS network.
- *S-CSCF* (Serving Call Session Control Function) [3]. The Session-CSCF is a SIP server that manages session control. Also, it acts as a SIP Registrar, i.e., it maintains a link between the user's position and the SIP address associated with it. Another main function of the S-CSCF is to provide SIP routing services.

In the 4G network, the HSS database stores all the subscribers' information and provides the 3G HLR functionalities. The HSS database is the target resource of a DDoS attack in order to prevent access to the cellular network. In previous studies on denial of service attacks to UMTS Network [7] [8], it has been shown that using just 2500 requests injected in every 4.7s time window, it is possible to saturate the HLR component, thereby leading to a successful DDoS attack.

Nevertheless, a limit for this attack is that it relies on a significant amount of mobile devices that must connect to eNodeB (cellular base station). Our idea is that, with the addition of the type of access enabled by the 4G Network architecture, the attacking devices need only to be connected to the Internet; this means that the population of devices that can be recruited into a botnet to carry out the DDoS could be orders of magnitude wider than in the case cited in [8]. It is finally worth to point out that in this case there is no need for any valid SIM module in order to carry out the attack.

3 The Attack

In order to have a set of non-3GPP nodes accessing the HSS, we will leverage the SIP protocol, with a specific focus on the SIP register procedure.

During the SIP register procedure, packets pass through three nodes besides the HSS, namely:

- *P-CSCF* (Proxy Call Session Control Function)
- *I-CSCF* (Internet Call Session Control Function)
- *S-CSCF* (Serving Call Session Control Function)

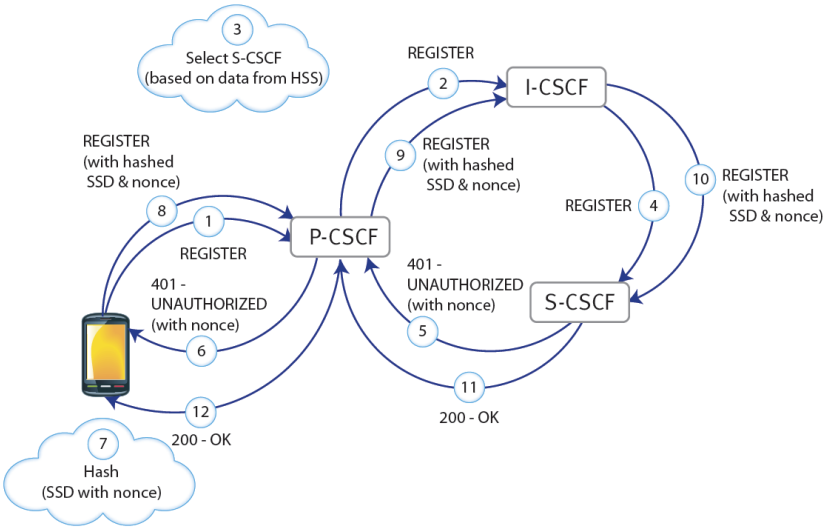

Fig. 2 The SIP register procedure.

The access point for Internet-based non-3GPP connection is the P-CSCF. It acts as a firewall that manages all the traffic from and to the 4G core network, and it can be reached from anybody connected to the Internet by only knowing its IP address. Its IP address can be retrieved through the DHCP-DNS procedure for P-CSCF discovery.

As shown in Figure 2, the SIP register procedure is composed of several steps:

1. The IMS User Equipment (UE, henceforth) attempts to register by sending a REGISTER request to the P-CSCF.
2. The P-CSCF forwards the REGISTER request to the I-CSCF.
3. The I-CSCF polls the HSS for data used to decide which S-CSCF should manage the REGISTER request. The I-CSCF then makes that decision.
4. The I-CSCF forwards the REGISTER request to the appropriate S-CSCF.

5. The S-CSCF typically sends the P-CSCF a *401 - UNAUTHORIZED* response as well as a challenge string in the form of a "nonce".
6. The P-CSCF forwards the *401 – UNAUTHORIZED* response to the UE.
7. Both the UE and the network have stored some Shared Secret Data (SSD), i.e., the UE in its ISIM or USIM and the network in the HSS. The UE uses an algorithm (e.g., AKAv2-MD5) to hash the SSD and the nonce.
8. The UE sends a REGISTER request to the P-CSCF. This time the request includes the result of the hashed nonce and the SSD.
9. The P-CSCF forwards the new REGISTER request to the I-CSCF.
10. The I-CSCF forwards the new REGISTER request to the S-CSCF.
11. The S-CSCF polls the HSS (via the I-CSCF) for the SSD, re-hashes the nonce, and determines whether the UE should be allowed to register. Assuming the hashed values match, the S-CSCF sends *200 – OK* response to the P-CSCF. At this point, an IPSec security association is established by the P-CSCF.
12. The P-CSCF forwards the *200 – OK* response to the UE.

In the first iteration (up to step 6), the UE sends to the network an authentication request, and the HSS will answer sending back the authentication vector just calculated. In the second one, the UE processes the authentication vector received and completes the procedure by retrieving the session information. For the time being, we were not able to complete the procedure because we cannot validate the vector received from the HSS, since we do not have a valid SIM module. However, this is indeed a minor issue, as the HSS processing capacity is mainly stressed during the first part of the connection, namely during the authentication vector calculation. The only private information that we would need for mounting the attack is the IMSI of the SIM card that we would emulate. An IMSI list could be easily obtained using various approaches:

- During the authentication process to an eNodeB, the IMSI is transmitted without any encryption and can be intercepted.
- Build a malicious app able to read the IMSI and force the user to install it through classical phishing attacks on mobile [4].
- Generate the list manually, given that we already know the standard IMSI structure.

4 Countermeasures and Related Work

The P-CSCF implements an A-SBC (Access – Session Border Controller) [1], [10] that, among other features, also deals with DDoS attacks. The DDoS attack detection and prevention adopts a simple blacklisting scheme that compares the number of users who have actually completed the authentication procedure with those who have failed the Response of Challenge in a finite time interval. Indeed, if someone is trying to compromise the system, the difference with respect to the number of legitimate users that successfully authenticate would be considerable.

Once detected, the most commonly used way to face up this threat is the blacklisting technique. The level of defense can be decided by setting the following main parameters:

- *Trigger-Size*: it defines the number of events from the specified source that are allowed before the blacklisting is triggered, and all packets are blocked from the source.
- *Trigger-Period*: it defines the length of the time window in which events are considered.
- *Timeout*: it defines when packets from the source are blocked if the configured limit is exceeded.

Obviously, it is necessary to find the right balance among all these parameters to ensure a good level of security without affecting the Quality of Service. As an example, it is required to avoid blocking a user i) in a low coverage area, or ii) having a malware infection and keep her offline for a long time once the issues are solved. At the same time, a low sensitivity with a very high trigger-size and/or a short trigger-period would allow botnets to go on with just a simple rotation of the attacker set. In past work, more sophisticated approaches have been proposed. In [11], the efficacy of *machine learning* for anomaly detection applied to the detection of SIP DDoS attacks is discussed; however, the false positive rates can reach 40%. In [12] the authors leverage the creation of statistical sketches of the incoming traffic and the measurement of Hellinger Distance; while the approach is promising, it is not capable of coping with slow mounting (a.k.a. stealthy) attacks that poison the traffic statistics. A different approach is described in [13], where the authors propose to introduce a DDoS detection module in every access device in the network. While promising, this approach is best suited to 5G Network, where it is possible to aggressively leverage SD-networking and Network Function Virtualization.

As the only countermeasure currently implemented in the 4G network is the simple blacklisting, we will only consider the effectiveness of the attack in the presence of such a defense.

5 Simulation of the DDoS Attacks

In this section, we describe a simple simulation showing how botnets of different sizes may disrupt the cellular network service by mounting a DDoS attack through the SIP register procedure described in Section 3.

First, it is preferable to choose many different P-CSCFs for accessing the network in order to avoid any bottleneck and maximize the impact of the attack. In this case, all packets will be automatically routed to the target HSS. For estimating the attack effectiveness, we used the results described in [7] and [8] as a capability benchmark; hence, we assumed to need sending 2500 SIP register request in a 4.7 time window, in order to saturate the HSS service capacity. Of course, a full-fledged analysis of the actual computational load of the SIP registration process compared to the generation

of the cryptographic challenge generated during the attach procedure is required to obtain a precise measurement of the efficacy of the attack. However, our simulation shows that a prolonged DDoS attack can still be mounted, even in the case in which the combination of the increase of the HSS service capacity and the complexity of the SIP register procedure is one order of magnitude less demanding than the HLR operation described in [7] and [8].

Regarding the impact of the countermeasures implemented inside the A-SBC component of the 4G architecture, a device will be blocked by the A-SBC blacklisting system if it will cause more authentication failure than the trigger-size in a time interval less than trigger-period. Hence, to avoid blacklisting, we should respect the following inequality:

$$DELAY > TRG\_SIZE / TRG\_PERIOD, \tag{1}$$

where *DELAY* stand for the period in which we send authentication request while TRG_SIZE and TRG_PERIOD are the parameters described in Section 4.

In this first scenario, we model a botnet made by 250,000 devices; this number could appear very large compared with the requirements described in [7] and [8], but this is only an average population of the existing botnet nowadays [14, 15]. We also assume a TRG_SIZE of 2 and a TRG_PERIOD of 300s. We can subdivide our botnet in smaller "sub-botnets" of 2500 devices that will activate sequentially in a 4.7s window, then wait until all the other sub-botnets have fired. In this way, each device will send a request every 470s, almost 8 minutes, a value larger than the TRG_PERIOD: in this way, it will not trigger the blacklisting countermeasure inside the A-SBC. Hence, in this scenario, the attack can ideally continue unchallenged.

We can also assume with the same size for the botnet that the time window used to memorize registration errors is 600s and the TRG_SIZE is still 2. We can calculate the duration of the attack, i.e., after how much time all our devices would be blocked.

$$M = \frac{N_{UE}}{UE_m} * TRG\_PRD * TRG\_SIZE \tag{2}$$

Where:

- M: attack duration in seconds
- N_{UE}: devices in our botnet
- UE_m: number of devices in the m^{th} sub-botnet
- TRG_PRD: time window in which registration errors are accumulated
- TRG_SIZE: value of parameter set in the blacklist parameter

In this scenario, all bots would be blacklisted after 940s, which is about 16 minutes, i.e., in a very short time. However, it is worth noticing that this is indeed an unrealistically strict scenario, as 2 errors in the SIP registration procedure is an event that may happen for many legitimate users for several reasons. Besides, also the timeout value, i.e., the amount of time a device will be kept inside the blacklist, is critical. In fact, if this is comparable to the TRG_PERIOD, bots will turn available to attack again too soon.

Let us analyze now a scenario with a TRG_PERIOD of 600s and a TRG_SIZE of 10. In this case, there will never be a 600s window containing 10 events from a single bot, as each bot will produce an event every 470s. Hence, the attack can, in principle, continue unchallenged. Actually, it is possible to write the following inequality:

$$\frac{N_{UE}}{2500} * 4.7 * TRG\_SIZE > TRG\_PERIOD \tag{3}$$

As long as this inequality can be satisfied, it is possible to mount an attack that goes unchallenged indefinitely. If we consider the case of a botnet with 250,000 bots, this means that it is necessary to have

$$\frac{TRG\_PERIOD}{TRG\_SIZE} > 470 \tag{4}$$

in order to be able to blacklist the bots. As an example, with a TRG_SIZE of 10, it would be necessary for the system to have a memory window of 4700s, in order to be able to blacklist all the bots in a 250000 bots botnet and stop the DDoS attack. At the same time, the memory window also gives the time after which the attack would stop.

6 Concluding Remarks

In this paper, we have presented a preliminary study of the feasibility of mounting a DDoS to the 4G Network through a botnet of generically Internet-connected devices. Given the difficulty of accessing the field apparata to measure their actual service capability, we adopted the measurements in [7] as a benchmark. In such a situation, we showed that a botnet of 250 thousand devices could, in principle, lead to a successful DDoS to the network for an amount of time that ranges from unlimited to 18 minutes, depending on the restrictiveness of the blacklisting parameters adopted in the A-SBC module of the network.

We also provided an inequality that allows calculating the ratio between the blacklisting parameters needed to be resilient to a DDoS attack and the size of the botnet carrying it out. A prolonged disservice due to DDoS would also generate a significant monetary and reputation loss for the MNO hit by such an attack. In fact, according to [16] a DDoS attack could generate a loss from 100'000 to one million dollars for each hour in which the services offered by the company are not available. This would seem to push for the implementation of a very restrictive policy in the A-SBC blacklisting module. At the same time, a very restrictive policy would run the risk to blacklist legitimate users that incur in network errors or other mishaps. Hence, we argue that it is nontrivial to define an optimal policy. The underlying assumption to our work is that the complexity of the SIP registration is comparable to the one of generating the cryptographic challenge required by the 3GPP attach procedure; however, we plan to study in more details the actual computational com-

plexity of the SIP registration procedure so that updated values will be provided, as future work.

References

1. 3GPP The Mobile Global Standard, https://www.3gpp.org/about-3gpp/about-3gpp
2. ETSI TS 136 300 V14.2.0 (2017-04), https://www.etsi.org/deliver/etsi_ts/136300_136399/136300/14.02.00_60/ts_136300v140200p.pdf
3. Technical Specification Group Services and System Aspects, IP Multimedia Subsystem (IMS), Stage 2, TS 23.228, 2006, 3rd Generation Partnership Project., https://www.etsi.org/deliver/etsi_ts/123400_123499/123406/07.01.00_60/ts_123406v070100p.pdf
4. S. Aonzo, A. Merlo, G. Tavella, Y. Fratantonio, "Phishing Attacks on Modern Android", in Proc. of the 2018 ACM SIGSAC Conference on Computer and Communications Security (CCS 2018).
5. M. Ficco and F. Palmieri, "Introducing Fraudulent Energy Consumption in Cloud Infrastructures: A New Generation of Denial-of-Service Attacks," in IEEE Systems Journal, vol. 11, no. 2, pp. 460-470, June 2017. doi: 10.1109/JSYST.2015.2414822
6. F. Palmieri and U. Fiore, Providing true end-to-end security in converged voice over IP infrastructures, Computers & Security, Volume 28, Issue 6, September 2009, Pages 433-449.
7. P. Traynor, M. Lin, M. Ongtang, V. Rao, T. Jaeger, P. McDaniel, and T. La Porta, "On cellular botNets: Measuring the impact of malicious devices on a cellular network core," in Proc. 16th ACM Conf. Comput. Commun. Security, 2009, pp. 223–234.
8. Alessio Merlo, Mauro Migliardi, Nicola Gobbo, Francesco Palmieri, Aniello Castiglione, A Denial of Service Attack to UMTS Networks Using SIM-less Devices, IEEE Transactions on Dependable and Secure Computing, pp. 280-291, Vol. 11, N.3, May-June 2014, http://dx.doi.org/10.1109/TDSC.2014.2315198
9. Home Subscribe Server, https://sites.google.com/site/lteencyclopedia/lte-network-infrastructure-and-elements#TOC-3.2-HSS-Home-Subscriber-Server-
10. Internet Engineering Task Force (IETF), RFC, Requirements from Session Initiation Protocol (SIP) Session Border Control (SBC) Deployments, https://tools.ietf.org/html/rfc5853
11. Z. Tsiatsikas, A. Fakis, D. Papamartzivanos, D. Geneiatakis, G. Kambourakis and C. Kolias, "Battling against DDoS in SIP: Is Machine Learning-based detection an effective weapon?," 2015 12th International Joint Conference on e-Business and Telecommunications (ICETE), Colmar, 2015, pp. 301-308.
12. J. Tang, Y. Cheng, Y. Hao and W. Song, "SIP Flooding Attack Detection with a Multi-Dimensional Sketch Design," in IEEE Transactions on Dependable and Secure Computing, vol. 11, no. 6, pp. 582-595, Nov.-Dec. 2014. doi: 10.1109/TDSC.2014.2302298
13. A. Febro, H. Xiao and J. Spring, "Distributed SIP DDoS Defense with P4," 2019 IEEE Wireless Communications and Networking Conference (WCNC), Marrakesh, Morocco, 2019, pp. 1-8. doi: 10.1109/WCNC.2019.8885926
14. M. Antonakakis et al., Understanding the Mirai Botnet, https://www.usenix.org/system/files/conference/usenixsecurity17/sec17-antonakakis.pdf
15. A. Marzano et al., "The Evolution of Bashlite and Mirai IoT Botnets," 2018 IEEE Symposium on Computers and Communications (ISCC), Natal, 2018, pp. 00813-00818. doi: 10.1109/ISCC.2018.8538636
16. K. Arora, K. Kumar, M. Sachdeva, Impact Analysis of Recent DDoS Attacks, https://pdfs.semanticscholar.org/a097/9ffca5c4669fd90b2e7b56a831a9e2e8d03a.pdf
17. N. Gobbo, A. Merlo, M. Migliardi, "A denial of service attack to GSM networks via attach procedure", Volume 8128 LNCS, 2013, Pages 361-376.