Session Initiation Protocol: Security Issues Overview

Bruno Cruz   *brunof.cruz@protonmail.com*
Department of Informatics Engineering
CoimbraUniversity
Coimbra, Portugal

Rui Filipe Pereira   *rui.pereira@protonmail.ch*
Lab UbiNET – Computer Science
Security and Cybercrime
Polytechnic Institute of Beja
Beja, Portugal

Abstract

The leading method of correspondence is clearly through voice trade. There are essentially two different ways through which voice can be effortlessly communicated on an organization: PSTN (Public Switched Telephone Network) and VoIP (Voice over Internet Protocol).

Mainly represented by SIP, VoIP protocols and implementations contain several vulnerabilities, particularly related to their complexities and in the face of interoperability of telephony equipment's.

It was by identifying a lack of literature with focus in security and potential vulnerabilities of the SIP Protocol that we propose in this document. We attempt to provide a theoretical analysis from security aspects used by one of the signaling call protocols, Session Initiation Protocol (SIP).

It is intended to lucidly illustrate and identify threats, vulnerabilities, security mechanisms, developed methods and protocols and, finally over time improvements.

Keywords: Session Initiation Protocol (SIP), SIP Security, Voice over IP (VoIP).

1. INTRODUCTION

Voice over IP (VoIP) has been a rapidly growing technology that delivers voice communications, by allowing the transportation of both audio and video as data packets over a private or a public IP network.

This provides significant benefits for users, companies, and service providers alike. Therefore, allowing location independence, simplicity, and low costs. These types of protocols are required to allow components to work smooth and accordingly in communication services around the globe, but before audio or video can be transmitted between devices it must be employed a way to find the remote device and negotiate data transmission.

Session Initiation Protocol (SIP) is an application-layer signaling protocol and a text-based client-server protocol used for communication sessions (calls). With one or more users (participants), working with both IPv4 and IPv6 (Schooler, Rosenberg, Schulzrinne, Johnston, Camarillo, Peterson & Handley, 2002).

The protocol was developed by Internet Engineering Task Force (IETF) and documented in (Schooler, Rosenberg, Schulzrinne, Johnston, Camarillo, Peterson & Handley, 2002). With the adoption of Session Initiation Protocol (SIP) telephony increases and, its considered as the dominant signaling protocol for calls over the Internet. SIP, like other internet protocols is

vulnerable to known internet attacks, introducing new security issues, concerns and risks to systems confidentiality, integrity, and availability in VoIP systems.

Furthermore, according to (Schooler, Rosenberg, Schulzrinne, Johnston, Camarillo, Peterson & Handley, 2002), it runs on top of several different transport protocols, these being: Transmission Control Protocol (TCP), User Datagram Protocol (UDP), Transport Layer Security (TLS) and Stream Control Transmission Protocol (SCTP). As a result, each protocol vulnerabilities are inherited.

This paper intends to sensitize the development and deployment of recent security mechanisms in IP communications. As such it is divided as follows: Section 2 provides information for the understanding of SIP, on how it works and operates. Section 3 shows a brief explanation of the most common experienced types of attacks, which attack methods, are identified and crossed information with the three principles of information security: confidentiality, integrity, and availability, to display the affected principles. Finally, in Section 4 we have SIP existing defense mechanisms and some developed methods. The paper is concluded in Section 5.

2. SIP PROTOCOL
SIP was designed with two-way communication sessions in mind and by utilizing sessions as a tool. For this to occur and, before any session can take place the developers have integrated a standard named Session Description Protocol (SDP), which defines multimedia sessions accordingly for the purpose of initiating and exchanging data.

This protocol is described in RFC 2327 and updated by RFC 4566. SIP is a text-based protocol simpler than other signaling protocols, for example like H.323, becoming similar to Hypertext Transfer Protocol (HTTP) or even Simple Mail Transfer Protocol (SMTP), consisting of headers and a message body.

3. SIP ARCHITECTURE
Following (Schooler, Rosenberg, Schulzrinne, Johnston, Camarillo, Peterson & Handley, 2002), SIP requires the handling of multiple elements for transactions to take place, so its behavior is set accordingly with each processing stage.

A SIP transaction can be defined as a conversation between a client and a server, before two clients start communicating. A generic architecture of the protocol is composed by user agents, proxy servers, redirect and registrar servers.
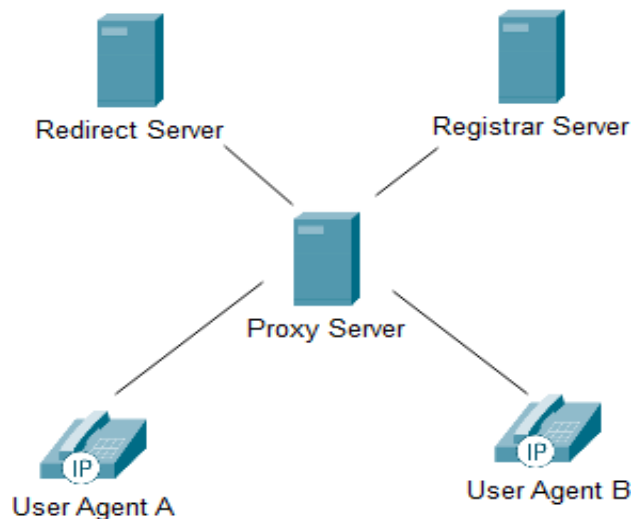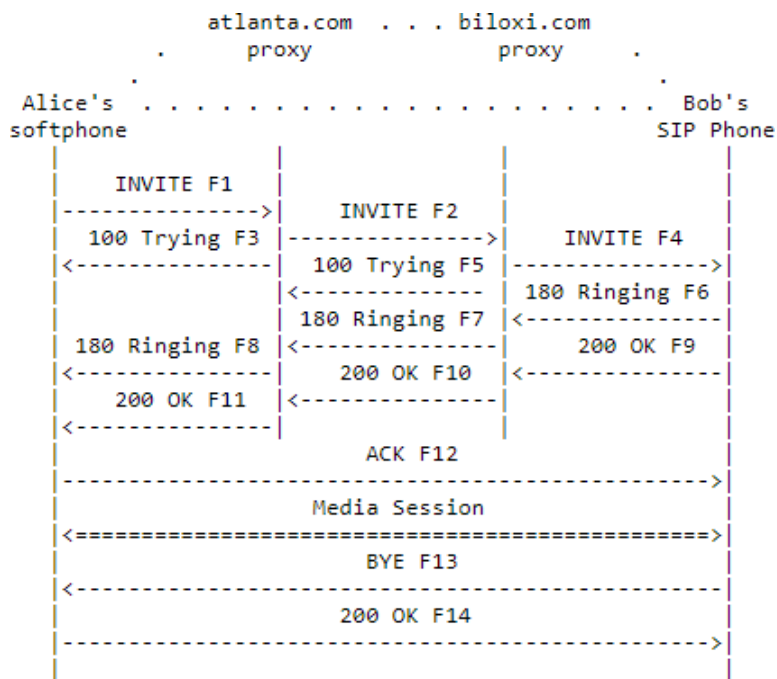

FIGURE 1: Generic SIP physical Architecture.

User Agent (UA) is a logical network endpoint that sends requests and receives responses, its communication lasts for as long as the transaction is taking place. Redirect Server only routes information requests by helping proxy servers locate their target and establish communications.

Registrar Server accepts registration requests and helps maintain information on the user agents by forwarding it into the handling domain. Proxy Server behaves like an intermediary, acting as both server and the client with the intent of making requests on behalf of those clients, forwarding them so that a connection can be established between caller and callee.

SIP is based on the HTTP-like request and response transaction model (Schooler, Rosenberg, Schulzrinne, Johnston, Camarillo, Peterson & Handley, 2002).Meaning that each request will always invoke at least one response from a server or client. The image below provided by (Schooler, Rosenberg, Schulzrinne, Johnston, Camarillo, Peterson & Handley, 2002) intends to demonstrate SIP basic communication functions via handshake: the desire to communicate, endpoint location, session (call) negotiation and session teardown (end of communication).

```
           atlanta.com  . . .  biloxi.com
               proxy                proxy        .
          .                                    .
Alice's  . . . . . . . . . . . . . . . . . . .  Bob's
softphone                                       SIP Phone
   |            |              |           |
   |  INVITE F1 |              |           |
   |----------->|  INVITE F2   |           |
   | 100 Trying F3 |---------->|  INVITE F4 |
   |<-----------|  100 Trying F5 |--------->|
   |            |<------------ | 180 Ringing F6 |
   |            | 180 Ringing F7 |<---------|
   | 180 Ringing F8 |<----------|  200 OK F9 |
   |<-----------|  200 OK F10  |<---------|
   | 200 OK F11 |<-------------|           |
   |<-----------|              |           |
   |            ACK F12        |           |
   |------------------------------------------>|
   |            Media Session  |           |
   |<=========================================>|
   |            BYE F13        |           |
   |<------------------------------------------|
   |            200 OK F14     |           |
   |------------------------------------------>|
   |            |              |           |
```

FIGURE 2: SIP session initiation example.

Considering that each proxy has incorporated both registrar and redirect functions in Figure 2, where Alice wants to contact Bob. Firstly, Alice will try to establish a connection by sending a request to initiate a conversation. Secondly, the proxy server (atlanta.com) will receive this request and start searching for Bob on behalf of the caller (Alice). Thirdly, after finding Bob through another proxy (biloxi.com) the request is then transmitted by the latter, which in turn redirects the request to Bob so that he can reply. Fourthly, Bob then replies to Alice's request and this information will travel through the same path (going from proxy to proxy) until it arrives to Alice. Lastly, after receiving confirmation, Alice acknowledges the reception, and a conversation is established directly from caller to callee. To terminate the communication a message is sent to inform the end of session.

SIP holds two different types of messages, requests, and responses. Each request represents a different functionality and is the foundation that allows to start a "dialog", Typically this communication is only between clients and servers which will then reach other clients when requested.

A response is the feedback provided by an action in relation to an answered request. Table 1 and Table 2 from (Schooler, Rosenberg, Schulzrinne, Johnston, Camarillo, Peterson & Handley, 2002) consist of requests and responses related to the original development. Later, several more requests were implemented through extensions.

REGISTER	Registers contact information
INVITE	Sets up session via invitation
ACK	Confirms the reception of a request response
BYE	Ends session by closure
CANCEL	Cancels the establishment of a call
OPTIONS	Queries server for information

TABLE 1: SIP Request Messages.

1xx	Provisional response, server is conducting action to determine a definitive answer.
2xx	Request successful
3xx	Redirects or provides alternative ways to establish a call
4xx	Provides failure response, from a particular server depending on the failure origin
5xx	Failure response originated from server errors
6xx	User (callee) related failures

TABLE 2: SIP Response Messages.

Failure is a type of response to a request with information related to an event reply. The difference between 4xx and 5xx is that 4xx is oriented for the user end-to-end communication while 5xx is aimed at server processing.

3.1 Layer Stack Model
According to (Schooler, Rosenberg, Schulzrinne, Johnston, Camarillo, Peterson & Handley, 2002), SIP is structured as a layered stack protocol and is divided into four layers (Transaction User, Transaction Layer, Transport Layer, and Syntax and Encoding) allowing the isolation of each layer based on their functionalities and to display their behaviors in different processing stages, as described in (Schooler, Rosenberg, Schulzrinne, Johnston, Camarillo, Peterson & Handley, 2002). This section was analyzed in conjunction with (Rehman & Abbasi, 2015) and (Rehman & Abbasi, 2014) which followed their own the oretical and practical analysis.

Transaction User Layer
Transaction Layer
Transport Layer
Syntax and Encoding Layer

TABLE 3: Stacked Model.

Transaction User (TU) Layer contains all SIP entities except the stateless proxy.

It allows the creation and cancelation of transactions at will (Schooler, Rosenberg, Schulzrinne,

Johnston, Camarillo, Peterson & Handley, 2002). Transaction Layer handles application-layer retransmissions, matches and compares responses to requests and handles application-layer timeouts. This layer does not exist in stateless proxies (Schooler, Rosenberg, Schulzrinne, Johnston, Camarillo, Peterson & Handley, 2002).

Transport Layer defines the mechanisms of processing how a client sends requests and receives responses.

Also, how a server receives requests and sends responses over the internet (Schooler, Rosenberg, Schulzrinne, Johnston, Camarillo, Peterson & Handley, 2002). Syntax and Encoding Layer is responsible for encoding the protocol data via a context-free grammar named Backus-Naur Form grammar (BNF) (Schooler, Rosenberg, Schulzrinne, Johnston, Camarillo, Peterson & Handley, 2002).

4. SIP SECURITY

Being a text-based message implies that this protocol is very susceptible to having worrying vulnerabilities that affect the three areas of information security: confidentiality, integrity and, availability.

Attacks to Integrity: VoIP applications make use of encryption, hash algorithms and message digests, to decrease the threat of attacks. Although, vulnerabilities are still existent within signaling messages and media packets. These attacks consist in intercepting and modifying network traffic, where the modifications can be deleted, inject and/or replace certain pieces of information inside the VoIP message or media. A typical example of attacks against the integrity of the signaling traffic is call rerouting (Lazzez, 2013).

Attacks to Confidentiality: Data must be protected from being read by an unauthorized user whenever he tries to capture media, identities, patterns, and credentials for the purpose of establishing unauthorized connections and practice deceptive actions.

Typical confidentiality attacks are eavesdropping, call pattern tracking, data mining and reconstruction (Lazzez, 2013).

Attacks to Availability: Systems, applications and stored information must be always available and accessible to users for when they require access, but certain attacks can affect their availability in a way that becomes inaccessible to everyone. These attacks are denial of service, and they can cause for instance, resource exhaustion or congestion through the consumption of all bandwidth. Commonly known availability attacks are Call Hijacking, Call Flooding and Malformed Messages (Lazzez, 2013).

SIP Attacks

Denial of Service (DOS) is a type of attack that floods a system to disrupt an internet host service, rendering it unavailable (Tas, Unsalver & Baktir, 2020). There is also, Distributed Denial of Service (DDOS) which is a similar type of attack, but the difference is that the latter sends traffic to flood a host from multiple machines quite often already infected with malware (Tas, Unsalver & Baktir, 2020).

Eavesdropping occurs when an attacker intercepts and decodes messages between a caller and a callee by capturing their traffic and converting the data packages into a conversation (Rehman & Abbasi, 2014).

Man-in-The-Middle (MITM) takes place when an attacker actively places himself between a session where a conversation is taking place or can even initiate a conversation. For example, replay attacks, sessions tear down, caller ID spoofing, toll fraud and message tempering. Here he can work as a proxy by intercept, decipher and forward the content of each message to the

respective user accordingly altered or unaltered, considering that he can forge the message himself (Rehman & Abbasi, 2014).

Spam over Internet Telephony (SPIT) relates to the transmission of unsolicited messages, in this case calls in VoIP systems, where a spammer will attempt to initiate a conversation and can even play a recorded message in the caseof the call being answered. These problems are more deeply analyzed in (Atkinson & Kent, 1998) and (Rehman & Abbasi, 2014).

Table 4 is based in the work of McGann and Sicker during SIP technologies security threats tests from 2005 which is then complemented with a sum of their students (McGann, 2005), here are listed some of VoIP vulnerabilities that affect this protocol and which CIA Triad principle is affected following the communication system of the TCP/IP model, although some of them may cross layers.

TCP/IP Model Layers	Vulnerability	Confidentiality	Integrity	Availability
Network Access	Physical Attacks	x		x
	ARP cache	x	x	x
	ARP flood			x
	MAC Spoofing	x	x	x
Internet	IP spoofing			
	Registration server, IP phone, MGCP, DNS, etc	x	x	x
	Redirect via IP spoofing	x	x	x
	Malformed packets	x	x	x
	IP fragmentation	x	x	x
	Jolt			x
Transport	Transport TCP/ flood			x
	TCP/UDP	x	x	
Application	TFTP server insertion		x	
	DHCP server insertion (redirect)		x	
	DHCP IP address starvation			x
	ICMP flood			x
	SIP			
	Registration Hijacking	x	x	x
	Call Hijacking (MGCP			
	Notified Entity parameter	x	x	x
	Message body modification	x	x	
	RTP insertion			

	Spoof via header	x	x	x
	Cancel/bye attack			x
	Malformed method			
	Redirect method			x
	RTP			
	SDP redirect			x
	RTP payload			x
	RTP message tampering	x	x	x
	Encryption	x	x	x
	Default settings/passwords	x	x	x
	Disable unnecessary	x	x	x
	services HTTP, FTP, etc			
	Buffer overflow	x	x	x
	Legacy Network Interaction	x	x	x

TABLE 4: VoIP attacks based on layers and CIA Triad Principle.

Network Access Layer is vulnerable to ARP corruption which can allow attackers to intercept data packages in traffic although some of these attacks require access to another layer.

Another threat is an exploit that can lead to Media Access Control (MAC) spoofing, allowing an attacker to impersonate a server, compromising conversations, and even allowing illegitimate VoIP phone calls.

Two primary attacks are ARP flood and ARP cache. While ARP flood attacks consist in using spoofed ARP replies, attempting to overwhelm and be able to overflow the ARP cache. In ARP cache attacks the attacker will try to manipulate this cache in a device by forging ARP replies to redirect communications (McGann, 2005).

Internet Layer holds prevalent vulnerabilities from the Internet protocol (IP). Attackers can impersonate devices, i.e., registration servers, proxy servers, IP phones and Domain Name Servers (DNS), by IP spoofing. Through impersonation, unauthorized VoIP calls can be established. IP addresses can be obtained by data packets sniffing. IP phones are a weakness in the systems and should be complemented with a firewall and by having their default passwords changed. This is also important for telecommunication service operators to avoid toll fraud. Attackers, who spoof IP addresses can in the absence of higher authentication layers, alter data. Two common attacks are IP fragmentation and Jolt. Both can lead a receiver to become unstable or even crash from data processing (McGann, 2005).

Transport Layer weaknesses primarily rely on Real Time Protocol (RTP) which mostly uses UDP since TCP favor's reliability and not timeliness. Despite this, both protocols are vulnerable to replay, flooding and fragmentation attacks. Here packets can be intercepted and sniffed allowing the conversion of captured voice data into an audio file. Encryption can greatly reduce risks to attacks in this layer but not in its fullness. Meanwhile, weaknesses such as exchanging keys in endpoints can lead to the interception of those keys, capture, and decrypt packets or even run a MITM attack. In TCP flood attacks SIP's three-way, hand shake fails in a manner that from it results a half open session and, if several of these are created similarly to a loop then a server memory can be exhausted causing a crash. UDP flood is similar but instead packets are used to

generate ICMP unreachable destinations packets in a loop consuming memory until the system crashes (McGann, 2005).

Application Layer is where a great number of the threats to VoIP systems resides since there are numerous exploitable applications. User Agents (UA) can be intercepted and modified, but this can be mitigated by authentication usage. Although it has authentication via HTTP Digest there is a weakness, the mechanism tends to have poor key management and lacks third party authority.

SIP uses S/MIME for encryption, but it is vulnerable to MITM attacks because keys can be intercepted during exchanges. REGISTER messages can be targeted, and its headers modified, allowing malicious registrations through REINVITE messages enabling rerouting. Manufacturers default configurations can also be exploited due to existing documented information, such as logins and passwords. Configurations should be altered accordingly to prevent or at least mitigate these exploits.

Applications and operative systems should be kept up to date with the latest patches and weak or non-existing passwords must be avoided.

In this layer attackers can also perform DOS attacks and impersonate users, where DOS attacks are derived from RTP, SDP and DHCP exploits. DHCP exploits particularly, can lease a server IP address causing legitimate request addresses to be vanquished (McGann, 2005).

5. PROTOCOL DEFENSE MECHANISMS

Through the analysis of (Schooler, Rosenberg, Schulzrinne, Johnston, Camarillo, Peterson & Handley, 2002) SIP does not provide security features on its own, instead uses existing protection mechanisms to complement itself and provide defenses, but they are not flawless. These methods are required to help preserving the confidentiality, integrity and at the same time authentication for users and administrators alike. In the publications (Belmekki, Raouyane, Belmekki & Bellafkih, 2014),(Sawda & Urien, 2006),(Rehman & Abbasi, 2015) and (Rehman & Abbasi, 2014) the authors also reflect over these same mechanisms based on SIP analysis.

IP Security (IPSec) makes it possible to provide end-to-end security by protecting exchangeable data traffic "paths" but only when two entities have established a trust relationship between each other (Atkinson & Kent, 1998).

Transport Layer Security (TLS) allows clients and servers to negotiate authentication by using cryptographic keys and encryption algorithm for data traffic, while also providing a message integrity check method (Chown, 2002) and (Allen & Dierks, 1999).

Secure/Multipurpose Internet Mail Extensions(S/MIME) makes use of public key encryption to signal MIME data, ensuring end-to-end messages encapsulation and protection (Ramsdell, 1999).

HTTP Digest is an authentication mechanism for credentials negotiation through tunneling using algorithms to harden data transfers (Franks, Hallam-Baker, Stewart, Hostetler, Lawrence, Leach, & Luotonen, 1999).

SIP Uniform Resource Identifier (URI) allow send-to-end protection by making use of tokens for identification and authentication, user (address) and password, securing communications (Audet, 2009), (Berners-Lee & R. F 1998).

6. DEVELOPED PROTECTION MECHANISMS AND PROTOCOLS IMPROVEMENTS

SIP was created with protocol versions prior to 2002 so some updated mechanisms that provide new benefits to it deserve to be of note. Together with recent developments on how to improve

security measures with demonstrated results. The work in this section was followed by an analysis of recent developments, each with its own meaning.

Rehman and Abbasi in (Rehman & Abbasi, 2015) devised a solution that employs the addition of a fifth layer entitled Security Layer that is subdivided into two other layers, Endorsement Layer and Cherry-Pick Layer on SIP stacked model, mentioned on Section2, Part-B. The idea is to provide a way to protect users during VoIP sessions from attacks like Man-in-The-Middle (MITM), Eavesdropping and Spam over Internet Telephony (SPIT). The first layer is entrusted with tokens and certificates management that will contain information about the user´s identification, timestamp, digital signature and generating authentication tokens before establishing a session. Cherry-Pick Layer provides two options for communication, standard and secure. Standard communication uses public keys to guarantee the establishment of a session, while Secure Communication encrypts data with Sessions Keys, where only users that will be establishing communication will receive information about this key to allow the encryption and decryption of messages during the call.

Gupta and Prajapati in (Gupta & Prajapati, 2019), came up with a four phases model (Setup Phase, Registration Phase, Login and Authentication Phase and password Change Phase) to elevate authentication security for the establishment of a call session. By focusing entirely on authentication, it allows to work on ways that improve the security of initiation, focusing only on user and server identification and becoming intangible. This method consists of performing multiple checks to make sure that there are no impersonators during communications and each end target is legitimate. In this method if a user wants to communicate, he will have to register first, where a hash function will be assigned, then if he wishes to establish a connection the user has to "present himself" before the system, to verify his own identity before the server tries to establish a call. If both users are not a fraud, then a session key will be generated to secure and allow communication.

Farley and Wang in (Farley & Wang, 2012), created a solution that was named VoIP Shield that acts as a gateway for both server and user, offering protection against SIP package manipulation from MITM attacks during data transfers. Also, its development focuses on providing a practical and lightweight solution. This method involves pairing a single server with its users by using a pre-shared key, provided by the proxy server and a cryptographic hash function to generate a message authentication code. Each will send a message by UA and, will be tagged with an authentication code that has a pre-shared key provided by the shield before proceeding to the proxy. Before a data package arrives at his destination it will have to go through the gateway inspection, acting has a shield to verify legitimacy before a valid exchange. To verify this, the proxy shield or gateway performs calculations over the received has hand pre-shared key to determine if the encrypted packet message authentication code matches the provided key. If it is a match, then the package will be forward to the server if not the package will be considered invalid and will be dropped.

Biondi, Bognanni and Bellain (2020) devised and conducted an experiment in a small environment using both VoIP and IoT (Internet of Things), by making use of a Raspberry Pi. They demonstrated that with just such device it is possible to counter two types of attacks through scripting: DoS and Eavesdropping. Although it was conducted in a small scale and would need to be tested on a larger scale environment to prove its effectiveness in bandwidth processing. But the concept promotes the usage of scripts, one to analyze received packages and discard them in case a possible DoS attack is detected. Whilst the other, makes use of encryption to protect a stream with little overhead and latency.

Tas, Unsalver and Baktir in (Tas, Unsalver & Baktir, 2020), introduced a new mechanism by exploring IP spoofing, reflection based and DDOS attacks, followed by an investigation of several other attack scenarios. This new defense mechanism was named DRDoS and is made of three modules/layers. Each one investigates a passing by data package on its own. These three modules are: Statistics, Inspection and Action. Statistics module oversees daily data traffic

storage according to timestamp (date and time), creating patterns based on the operating network, packet specifications and calculates bandwidth usage. Inspection Module oversees the comparison of suspicious traffic and creation of IP rules that are used by the Action Module. Basically, Inspection only becomes active when a traffic threshold is reached. When this occurs, the module becomes suspicious of an attack, and will start comparing what it considers to be normal traffic to the currently received traffic and starts inspecting its headers. If suspicious activity is found, then the packets are either dropped or blocked by the Action Module. This mechanism allows to quickly limit a server load, avoiding a system crash.

Rescorla in (Rescorla, 2018), upgraded Transport Layer Security (TLS) to a newer version (1.3) considering new improvements and when in comparison with the initial Version 1.0 (Allen & Dierks, 1999), it is a much more optimized and hardened version of the protocol. Big changes were firstly seen in the release of Version 1.2 where newer and more recent cryptographic hash functions were introduced with also better encryption support. One of the newer version's big changes were that it stopped supporting legacy and already obsolete algorithms allowing for a "rebuild". Improvements were also made to secret keys.

Schaad, Ramsdell and Turner in (Schaad, Ramsdell, & Turner, 2019) developed Secure/Multipurpose Internet Mail Extensions (S/MIME) Version 4.0 introducing updates related with new cryptography authentication methods, removed the older ones, and improved its hash functions algorithms, making a big improvement when compared to Version 3 (Ramsdell, 1999) which is now obsolete.

P. Segeþ, M. Moravþík, J. Hrabovský, J. Papán and J. Uramová in (2017), focused their research in securing RTC communication, which consists of two parts – signalization and multimedia data transfer. Within the signalization part, the SIP (Session Initialization Protocol) protocol became dominant. In the data transfer part, there is the RTP (Realtime Transport Protocol) protocol. From the security point of view, on one hand both protocols use several built-in security mechanisms. On the other hand, there exists a whole autonomous infrastructure specially built for the provisioning and support of secure communication. The infrastructure it uses is known as the Public Key Infrastructure (PKI). The work developed by the authors focuses on the present state analysis of use and cooperation of RTC and PKI to provide forms of better and more complex RTC security. The authors demonstrate, there are mechanisms for securing the signalization (SIP) as well as multimedia streams that can use the asymmetric encryption and infrastructure of public key.

Pereira D. and Oliveira R. in (Pereira & Oliveira, 2021), propose a deep learning-based approach to detect possible attacks. Their solution is based on the definition of an orthogonal space capable of representing the sampling space of flowing data through incremental time changes, which is then used to train a recurrent neural network to classify the type of SIP dialog for the sequence of packets observed. When a sequence of observed SIP messages is unknown, this represents possible exploitation of a vulnerability and in that case, it should be classified accordingly.

Golait and Hubballi in (Golait & Hubballi, 2017), the authors consider the SIP operation sequence as a Discrete Event System (DES). They developed a probabilistic timed transition model (PCDTA) to characterize SIP event sequences and their timings. They proposed learning transition and delay probabilities of events of state transition diagram from a collection of known non malicious SIP events. Therefore, making learning automatic instead of manually. They identify a range of anomalies that can happen in anytime and map these anomalies to various DoS attacks in SIP. The authors used the timed transition model as an anomaly detection system, rising because of the volume of illegal transitions, which in turn will help detect different SIP attacks.

The authors were able to modulate different SIP dialogues and transactions as discrete event systems and proposed a probabilistic state transition machine to describe these dialogues and

transactions. They describe algorithms to detect various DoS attacks using the proposed state transition model.

7. CONCLUSIONS

In conclusion, VoIP vulnerabilities are a real worldwide threat that must be faced seriously alongside with the implementation of efficient defensive mechanisms to mitigate or in an ideal scenario stop the attack completely. During defensive mechanisms development, something that must be kept in mind is that systems are not only vulnerable from inside but mostly from the outside attacks. Hackers have all the internet at their disposal and sometimes security developments can be misleading mainly due to the kind of situations above described. And, because of this it can lead to a false sense of trust, safety, and security. One of the biggest weaknesses in SIP that represents a huge threat and can lead telecommunication operators to heavy costs caused by a hacker, are physical devices. The main reasons being that IP phones can easily be hacked if not properly protected by a firewall, and default configurations can leak critical information details if their authentication credentials remain standard and are not changed. As time goes by, these systems are becoming more and more consistent, and so are the efforts in trying to secure them. In the real-world an application encryption mechanism alone cannot be considered the best possible practice to create secure communications between devices. They can bring latency related issues but what is worse is that not all mechanisms consider external system threats. Following the work done by Keromytis in (Keromytis, 2010) which demonstrates that at least during this period most attacks on these systems were DoS (58%), followed by eavesdropping and hijacking (20%) then social threats (18%) and at last service abuse (4%) shows that DoS attacks are the most worrying type followed by MITM types. But recent developments seem to help. Tas, Unsalver and Baktirin in (Tas, Unsalver & Baktir, 2020) proposed a solution that greatly mitigates DDoS attacks automatically, and Farley and Wang in (Farley & Wang, 2012) developed a lightweight solution for MITM attacks. All this demonstrates that the community is taking an interest into VoIP and willingly directing their efforts into protecting this technology and keeping their user's and their user's information evermore protected.

8. REFERENCES

A. D. Keromytis, "Voice-over-IP Security: Research and Practice," in *IEEE Security & Privacy*, vol. 8, no. 2, pp. 76-78, March-April 2010, doi: 10.1109/MSP.2010.87.

Allen, C., & Dierks, T. (1999, Januarie). *The TLS Protocol Version 1.0*. doi:10.17487/RFC2246

Atkinson, R., & Kent, S. (1998, November). *Security Architecture for the Internet Protocol*. doi:10.17487/RFC2401

Audet, F. (2009, Oktober). *The Use of the SIPS URI Scheme in the Session Initiation Protocol (SIP)*. doi:10.17487/RFC5630

B. B. Gupta and V. Prajapati, "Secure and efficient Session Initiation Protocol authentication scheme for VoIP Communications," 2019 International Conference on Communication and Electronics Systems (ICCES), 2019, pp. 866-871, doi: 10.1109/ICCES45898.2019.9002125.

Chown, P. (2002, Julie). *Advanced Encryption Standard (AES) Ciphersuites for Transport Layer Security (TLS)*. doi:10.17487/RFC3268

D. Golait and N. Hubballi, "Detecting Anomalous Behavior in VoIP Systems: A Discrete Event System Modeling," in IEEE Transactions on Information Forensics and Security, vol. 12, no. 3, pp. 730-745, March 2017, doi: 10.1109/TIFS.2016.2632071.

E. Belmekki, B. Raouyane, A. Belmekki and M. Bellafkih, "Secure SIP signalling service in IMS network," 2014 9th International Conference on Intelligent Systems: Theories and Applications (SITA-14), 2014, pp. 1-7, doi: 10.1109/SITA.2014.6847291.

Franks, P. J., Hallam-Baker, P., Stewart, L. C., Hostetler, J. L., Lawrence, S., Leach, P. J., &Luotonen, A. (1999, Junie). *HTTP Authentication: Basic and Digest Access Authentication*. doi:10.17487/RFC2617

I. M. Tas, B. G. Unsalver and S. Baktir, "A Novel SIP Based Distributed Reflection Denial-of-Service Attack and an Effective Defense Mechanism," in IEEE Access, vol. 8, pp. 112574-112584, 2020, doi: 10.1109/ACCESS.2020.3001688.

Lazzez, A. (2013). VoIP Technology: Security Issues Analysis. ArXiv, abs/1312.2225.

McGann, S. (2005). An Analysis of Security Threats and Tools in SIP-Based VoIP Systems.

P. Biondi, S. Bognanni and G. Bella, "VoIP Can Still Be Exploited - Badly," 2020 Fifth International Conference on Fog and Mobile Edge Computing (FMEC), 2020, pp. 237-243, doi: 10.1109/FMEC49853.2020.9144875.

Pereira D., Oliveira R. (2021) Detection of Signaling Vulnerabilities in Session Initiation Protocol. In: Camarinha-Matos L.M., Ferreira P., Brito G. (eds) Technological Innovation for Applied AI Systems. DoCEIS 2021. IFIP Advances in Information and Communication Technology, vol 626. Springer, Cham. https://doi.org/10.1007/978-3-030-78288-7_20

P. Segeč, M. Moravčík, J. Hrabovský, J. Papán and J. Uramová, "Securing SIP infrastructures with PKI — The analysis," 2017 15th International Conference on Emerging eLearning Technologies and Applications (ICETA), 2017, pp. 1-8, doi: 10.1109/ICETA.2017.8102525.

Ramsdell, B. C. (1999, Junie). *S/MIME Version 3 Message Specification*. doi:10.17487/RFC2633

Rescorla, E. (2018, Augustus). *The Transport Layer Security (TLS) Protocol Version 1.3*. doi:10.17487/RFC8446

R. Farley and X. Wang, "VoIP Shield: A transparent protection of deployed VoIP systems from SIP-based exploits," 2012 IEEE Network Operations and Management Symposium, 2012, pp. 486-489, doi: 10.1109/NOMS.2012.6211937.

Schooler, E., Rosenberg, J., Schulzrinne, H., Johnston, A., Camarillo, G., Peterson, J., … Handley, M. J. (2002, Julie). *SIP: Session Initiation Protocol*. doi:10.17487/RFC3261

Schaad, J., Ramsdell, B. C., & Turner, S. (2019, April). *Secure/Multipurpose Internet Mail Extensions (S/MIME) Version 4.0 Message Specification*. doi:10.17487/RFC8551

S. El Sawda and P. Urien, "SIP Security Attacks and Solutions: A state-of-the-art review," 2006 2nd International Conference on Information & Communication Technologies, 2006, pp. 3187-3191, doi: 10.1109/ICTTA.2006.1684926.

T. Berners-Lee, R. F. (1998, Agust). *Uniform Resource Identifiers (URI): Generic Syntax, RFC 2396*. Retrieved from IETF Tools: https://www.ietf.org/rfc/rfc2396.txt

U. U. Rehman and A. G. Abbasi, "Secure Layered Architecture for Session Initiation Protocol Based on SIPSSO: Formally Proved by Scyther," 2015 12th International Conference on Information Technology - New Generations, 2015, pp. 185-190, doi: 10.1109/ITNG.2015.35.

U. U. Rehman and A. G. Abbasi, "Security analysis of VoIP architecture for identifying SIP vulnerabilities," 2014 International Conference on Emerging Technologies (ICET), 2014, pp. 87-93, doi: 10.1109/ICET.2014.7021022.