BAKI ALİ NEFT MƏKTƏBİ
BAKU HIGHER OIL SCHOOL

Mobile development and security

Session 19

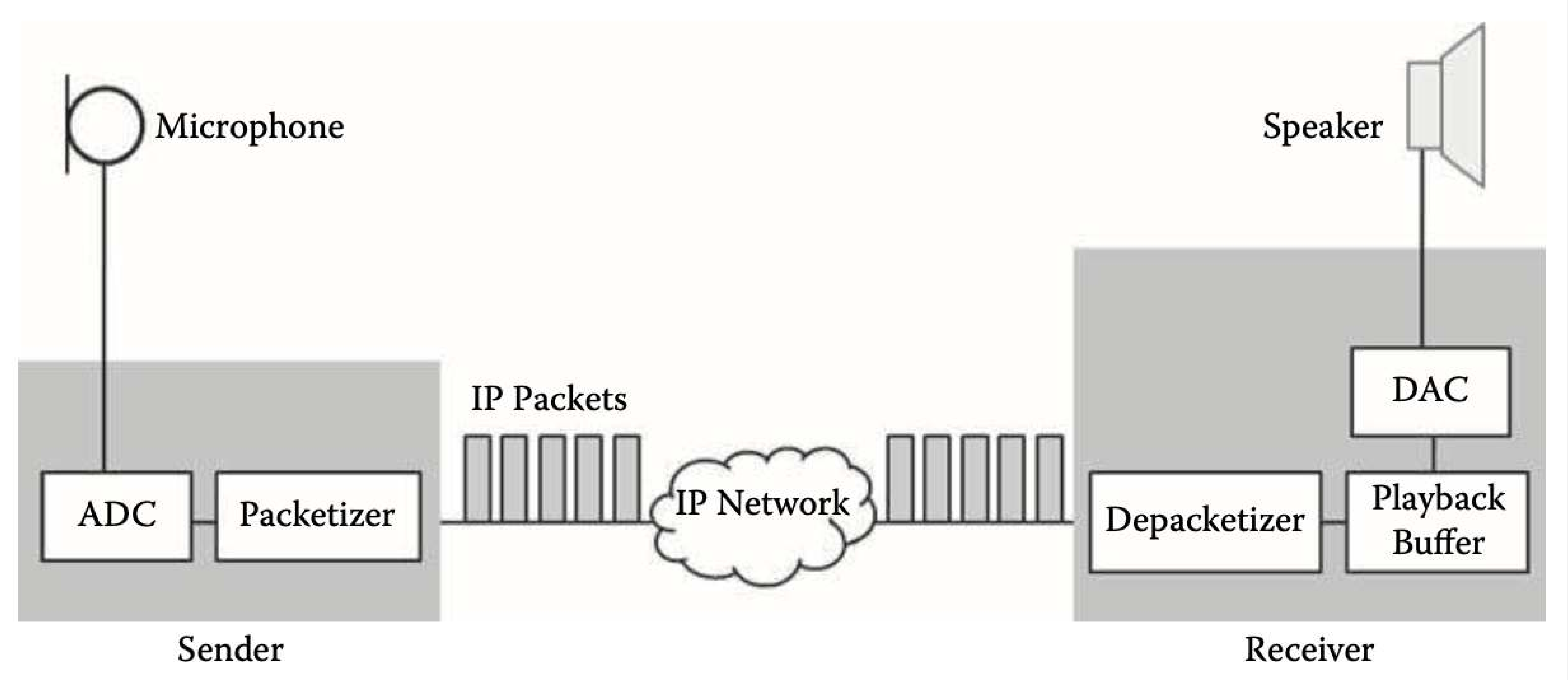
Karim Karimov

Lecturer

Voice communication

Voice communication has been the central service of the mobile communication networks for the last decades. As communication networks are nowadays moving toward packet switched networks, where all types of streams and information are transported by packets and delivered to the mobile users, it comes into vision that providing **circuit-switched services** can no longer be justified since it is inadvisable to have dedicated infrastructures to implement voice services, when their integration can be made on a **packet-switched network** in a scalable and cost-effective manner. In addition, the design, development, deployment, and operation of mobile multimedia services, in general, and voice services, in particular, present several advantages when they are transported on a packet-switched networks, as the networks experience less unnecessarily reservation during service access and their resources are better utilized. In addition, many new services will require more interaction with the user, close synchronization, and better provision of the quality of service.

VoIP


Microphone
Speaker
IP Packets
DAC
ADC
Packetizer
IP Network
Depacketizer
Playback Buffer
Sender
Receiver

VoIP

Recently, a variety of voice-oriented services, where voice is essentially coded into streams of IP packets, have been offered. The technologies used to deploy these services over communication networks are often referred to as **voice over IP (or VoIP)**. A typical VoIP system is depicted in the figure previously and works as follows: At the sender side, the voice sound is first sampled using a microphone. Then, the sampled traffic is translated into a digital representation by an **analog-to-digital converter (ADC)** to obtain a bit stream. The stream is packed into IP packets and sent over an IP network. At the receiver side the samples are first removed from the IP packets. Then, they are put in a playback buffer. This buffer is needed to compensate for the variation of the jitter generated between packets over the network. Finally, a **digital-to-analog converter (DAC)** converts the bit stream back into an analogue signal. To provide a two-way communication, VoIP replicates the system depicted in the figure (previously) reversely between the receiver and the sender.

VoIP problems

In particular, the use of mobile VoIP allows the transmission of all types of streams, **real-time data**, and **non-real-time data** over the same network, enabling the emergence of a wide range of attractive services including voice-enabled electronic commerce and interacting multimedia services. Nevertheless, three major problems have to be addressed:

- **the limited bandwidth and the constraints** on other network resources present in the wireless environment may reduce the performance of mobile VoIP;
- the addition of **security mechanisms** to mobile VoIP systems in order to provide secure communications puts more limitations on the performance of the mobile devices and can even reduce the system flexibility of mobile VoIP; and
- **the management of IP addresses** needs to change the IP address of a mobile subscriber as he moves to a new area. Currently, such situation forces the termination of the VoIP call made by the moving user, since the current VoIP solutions require the hosts involved in a call to have fixed IP addresses.

VoIP signaling protocols

Two signaling protocols are widely used in the VoIP solutions. They are

1. **the SIP (Session Initiation Protocol) protocol**, which is recommended by the Internet Engineering Task Force, and
2. **the H.323**, a protocol recommended by the International Telecommunications Union.

Today, the two protocols are not compatible. The SIP is **simple, scalable, and extensible**. It requires four packets to establish a call, whereas the H323 requires **longer setup time**, needs 12 packets for the call setup, and provides a control within a session providing a way for conferences to coordinate input to the produced media. Two major advantages of SIP can be noticed with respect to H.323. First, the little number of packets needed to process a call setup and the fact that SIP runs on UDP, while H.323 uses both TCP and UDP during the call setup. The TCP is a reliable way to send data, because the data is acknowledged, ordered, and resent if not correctly received. UDP is just used to send data packets with a minimum of protocol overhead, no matter whether the packets arrive or not.

Real-time protocol (RTP)

When using TCP, a packet should be resent if it is not received correctly. This retransmission consumes much time and increases delay unbelievably, because large overhead is generated. **The real-time protocol** (**RTP**) is provided for this purpose. Various VoIP applications have been made available based on H.323 and RTP (e.g., Microsoft NetMeeting, Skype, and MediaRing Talk).

SIP Overview

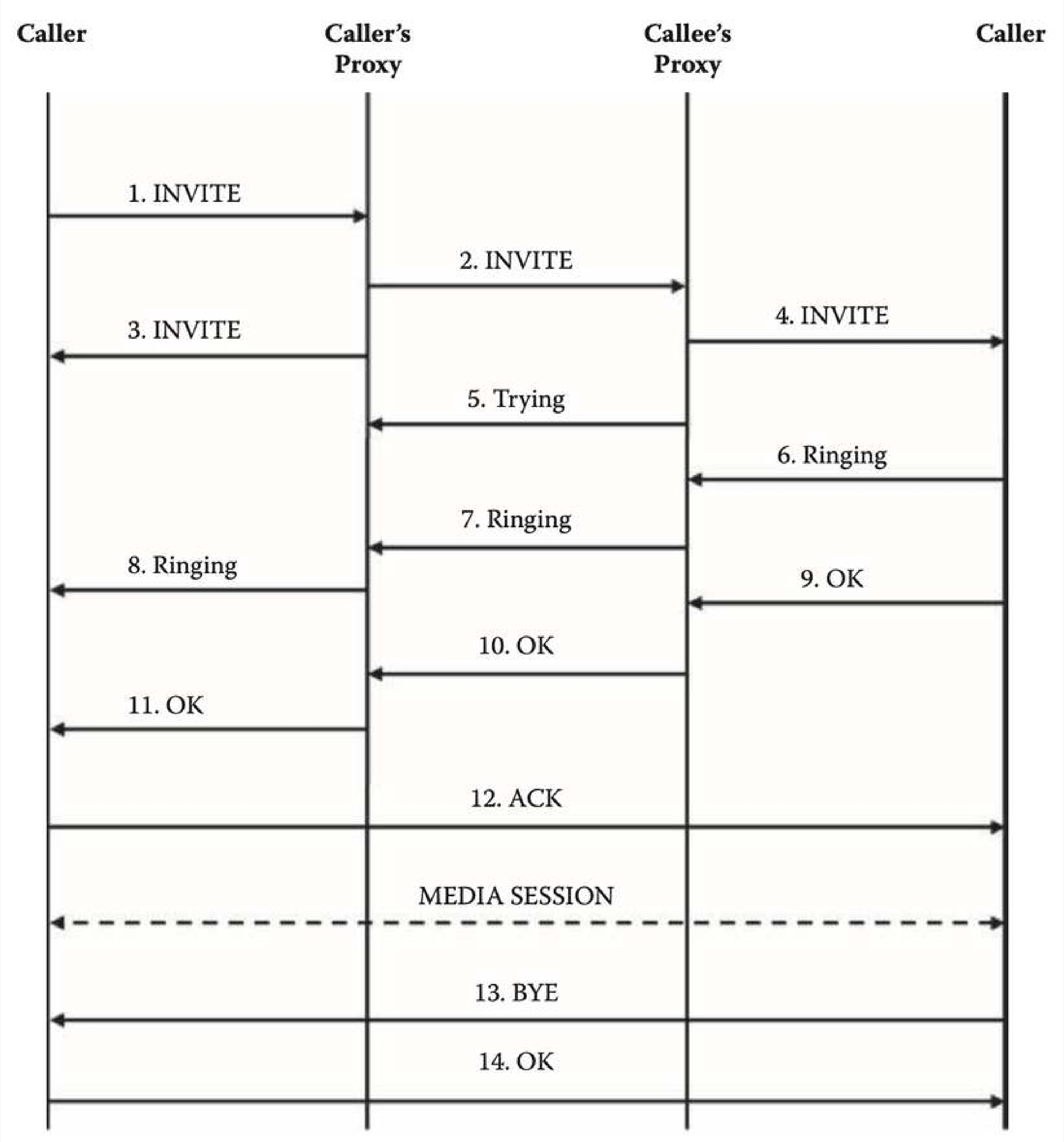
SIP is an application layer control protocol that can **establish**, **modify**, and **terminate** calls. Two main architectural elements can be distinguished in SIP: **the user agent (UA)** and the **network server**. The UA is implemented at the SIP end stations. The SIP user agent has two basic functions: **receiving** incoming SIP messages and **sending** SIP messages upon user actions or incoming messages. It contains a user agent client (UAC) that is responsible for issuing SIP requests and a user agent server (UAS) that is in charge of responding to SIP requests. The SIP UA also starts appropriate applications according to the session that has been established.

SIP Overview

Three network server types can be found in SIP: **the proxy server**, **the redirect server**, and the **registrar**. However, a basic SIP call does not use servers. The SIP proxy server relays SIP messages, so that it is possible to use a domain name to find a user, rather than knowing the IP address or the name of the host. In that way, a SIP proxy can be used to hide the location of the user. A redirect server returns the location of the host; unlike the SIP proxy, it only has to send back a response with the correct location instead of participating in the whole transaction. Both the proxy server and redirect accept registrations from users, in which the current location of the user is given. The location can be stored either locally at the SIP server or in a dedicated location server. Another important function regarding SIP is the registration of the users with their provider's servers. When a SIP-based device (or UA) gets online, it first must get registered with a SIP Registration Server, called the Registrar. This process is handled by sending a **REGISTER** message. Registrations are not permanent; they bind the mobile user's ID with an IP address where the user can be reached.

SIP setup

An example of SIP setup is depicted by Figure right, where the caller calls another user (referred to as the answerer or callee) using his SIP identity, a type of **Uniform Resource Identifier** (**URI**), called **SIP URI**, which is similar to an email address and contains the user name and the host identifier (for example, **caller-ID@host-ID**). The caller sends a request called **INVITE** to the answerer provider SIP server proxy (SPA) via his provider's SIP server. If the answerer accepts the call, the media session is established.



H.323 Overview

The H.323 Standard defines the mechanism by which real-time information can be transmitted over packet-based networks that do not provide a guaranteed quality of service. It defines four major components for a network-based communication system: the **Terminals**, the **Gateways**, the **Gatekeepers**, and the **Multipoint Control Units**. The terminals are client endpoints attached to IP-based networks that provide real-time, two-way communications with other H.323 entities. The H.323 **terminals** implement the following functions:

- ***Signaling and Control*:** The terminals implement a standard for channel usage and capabilities, in addition to a protocol for call signaling, call establishment, and registration/administration/status (RAS) for communication with gatekeepers.
- ***Real-time communication*:** H.323 terminals implement RTP as a protocol for sequencing audio and video packets.
- ***Codecs*:** H323 implements pieces of software, referred to as Codecs, to compress audio/video before transmission. The decompression is operated to get the compressed packets back immediately after their reception.

H.323 Overview

The **gateways** provide the inter-connection between the packet-switched network and the **switched circuit network** (**SCN**). The gateway is not required when there is no connection to other networks. The gateway performs call setup and control on both the packet-switched network and the SCN. It has the responsibility of **translating transmission formats** and **communication** procedures.

The **gatekeepers** perform at least four compulsory functions:

- the **address translation** (typical translations transform the phone numbers into the transport addresses),
- **admission control**,
- **bandwidth control**, and
- **zone management**.

Gatekeepers can also support four optional functions: **call control signaling**, **call authorization**, **call management**, and **bandwidth management**.

H.323 Overview

Finally, the **multipoint control units (MCU)** support conferencing using a set of endpoints. Typically, a MCU consists of a **multipoint controller** (**MC**) and zero or more **multipoint processors** (**MP**). The MC provides control functions such as the negotiation between terminals and the determination of common capabilities for processing audio and video. The MP performs the necessary processing on the media streams for a conference involving audio mixing and audio/video switching.

H.323 supports five types of information streams between endpoints: **Audio**, **Video**, **Data**, **communications control data**, and **call control data**. Audio and Video streams are processed using audio and video ***codecs***; they are transmitted and controlled using the **Real Time Transport Protocol** and the **Real Time Control Protocol (RTCP)** operating over an unreliable transport such as UDP. H.323 uses the concept of channel to structure the information exchange between communicating entities. A channel is a transport-layer connection that can be unidirectional or bi-directional.

Comparing the Signaling Protocols

Functionality

Call Setup and Tear Down

H.323 v2 call setup is based on RTP. The call setup needs a two-phase connection: the TCP connection and call connection. The H.323 v3 supports both TCP and UDP, which simplifies the call setup procedure. SIP call setup procedure is similar to H.323 v3. The tear down procedure is a reverse operation of the call setup. Either caller or callee entity can terminate a call by RELEASE COMPLETE (in H.323) or BYE (in SIP) messages.

Call Forwarding

Call forwarding permits the called party to forward particular pre-selected calls to other addresses. Call forwarding services provided by SIP are usually instantiated with the ***LOCATION*** header fields, which contain the forwarding destination. SIP supports call forwarding busy, call forwarding no response, and selective call forwarding.

Comparing the Signaling Protocols

Quality of Service

The relevant QoS parameters for VoIP flows are the **bandwidth**, **latency**, **the delay jitter**, **the packet loss**, and **the call setup delay**.

Jitter refers to non-uniform packet delays. It is often caused by **low bandwidth** situations in VoIP. Jitter can cause packets to arrive and be processed out of sequence. Jitter can also be controlled throughout the VoIP network by using routers, firewalls, and other network elements that support QoS. **Packet loss** can result from **excess latency**, where a group of packets arrives late and must be discarded in favor of newer ones. It can also be the result of jitter, that is, when a packet arrives after its surrounding packets have been flushed from the buffer, making the received packet useless. The real-time constraints imposed to VoIP do not allow for a reliable protocol, such as TCP, to be utilized to mitigate packet losses. By the time a VoIP packet is reported missing, retransmitted, and received, the time constraints for QoS would be exceeded.

Security Issues in VoIP

Adding security constraints increases significantly the bandwidth usage, generating more latency and jitter, and thus reducing the overall QoS of the network. In addition, these requirements do not explicitly take into account the heterogeneous data flow over the network. Since voice and data streams are sharing the same limited bandwidth, significant amounts of data can generate a congestion in the network and prevent VoIP traffic from reaching its destination in conformance with time constraints imposed in the QoS it needs. However, the following potential risks have been identified for H323 and SIP signaling:

- **The signaling and voice data transport plane can be targeted** by attacks that aim at breaking the integrity, confidentiality, authentication, or non-repudiation of the transported data.
- continues.

Security Issues in VoIP

- The audio payload data and the signaling information, exchanged during a call, are **sensitive** to eavesdropping, jamming, and even active modification. The challenges become even more evident in an open environment where finding, choosing, and using services are subject to competition between service providers.
- **Compromising the identity of an end system or infrastructure component** leads to additional risks even when using standard and non-compromised signaling mechanisms. If a malicious user can register with a H.323 Gatekeeper (or SIP server/registrar), he can potentially gain the identity of the victim. This can cause a potential invasion of privacy (since incoming calls are routed to the attacker and may give him information about the callers) and the possibility of misusing services.

Security Provided by H323

The security service provided by the H323 suite (as given by H235v2) is characterized by the following supports: **the support of elliptic curve cryptography**, **the support for the Advanced Encryption System (AES) standard**, and the **use of several security profiles** to help product interoperability. The Baseline Security Profile relies on symmetric techniques. Shared secrets are used to provide authentication and/or message integrity. Three supported scenarios are available for this profile:

1. the endpoint-to-gatekeeper,
2. the gatekeeper-to-gatekeeper, and
3. the endpoint-to-endpoint.

Security Provided by SIP

Various security services have been made available for SIP including **a digest authentication scheme** that is based on a simple challenge-response paradigm. Using the RTP encryption, it provides confidentiality for media data. RFC 3261 commands the use of TLS for proxy servers, redirect servers, and registrars to protect SIP signaling. In fact, TLS is able to protect the SIP signaling messages against **loss of integrity**, attacks against **confidentiality**, and **replay attacks**. It provides **integrated key-management**, **mutual authentication**, and **secure key distribution**. TLS is applicable on a hop-by-hop fashion between UAs and proxies or between proxies. However, a disadvantage of the TLS use in SIP scenarios can be noticed: TLS requires a reliable transport stack and **cannot** be applied to UDP-based SIP signaling.

The Security threats to Mobile VoIP

A threat to VoIP can be a loss of availability of the VoIP service or a telephone fraud. Typically, four categories of important threats to VoIP can be distinguished. They are related to **the abuse of access**, **denial of service**, **eavesdropping**, and **masquerading**. One can easily believe that the signaling and the media transport planes can be targeted by attacks targeting the integrity, confidentiality, authentication, or non-repudiation of the transported data. In addition, the signaling information exchanged between the components is sensitive to eavesdropping, jamming, and active modification. The challenges generated by these threats become even harder to face in an open environment, where finding, selecting, and running services are subject to competition between service providers.

Abuse of Access

This is a category of threats where malicious mobile users (or programs acting on behalf of mobile users) misuse or abuse their access to the mobile VoIP systems. Abuse of access can have multiple forms. A first form of abuse of access is the so-called **click-to-dial service**, where an enterprise, setting up a specific service, calls back users on his demand, via the regular phone system. The service set up is typically offered through a Web page allowing any user to enter any phone number. An attacker can use this threat to cause some financial losses to the enterprise.

A second example of abuse of access can be provided by phone applications that present an API to other applications allowing them **to initiate calls or insert objects in active calls**. The inserted objects can be malicious. Examples of such applications include the Skype application. Generally, the presented API is protected by an access control mechanism asking the users whether they want to allow the phone application to control the API use. The Skype system, for example, authorizes another application to manage all the aspects of the Skype client.

Denial of Service

This is a category of threats aiming at conducting attacks to deliberately cause loss of availability of the mobile VoIP service. DoS threats against a mobile VoIP can be identified at several communication levels including the **transport/network-level**, **server level**, and **signaling level threats**.

- **The transport level DoS threats:** A DoS attack may be launched by flooding a target. An example is given by the ping of death or Smurf attack.
- **The server level DoS threats:** A server may be made unavailable by simply modifying some stored information in order to prevent authorized users from accessing the service, or by overwhelming the server with a large number of requests.
- **The signaling level DoS threats:** A DoS can make the SIP protocol unavailable to handle legitimate SIP messages by overloading the protocol server with too many messages. Unauthorized mobile users can also generate over-usage problems inducing the degradation of the QoS for the legitimate users. An example is the disruption of network services by collapsing the entire signaling protocol.

Denial of Service

Further DoS attacks that are specific to mobile VoIP include the following attacks that can be launched against SIP.

- A DoS attack on a call can be performed by sending **spoofed bye messages** to the user participating in a SIP call in order to close the call.
- The entities involved in the SIP signaling are **vulnerable to DoS attacks**.They can be simply flooded using ***Register*** or ***Invite*** messages.
- Illegal SIP messages can overload SIP, since SIP ensures that any entity receiving a SIP message must **investigate the entire message** before it can state its validity.

Eavesdropping Threats

This class of threats attempts to list signaling or data packets by copying legitimate messages between the communicating entities. Eavesdropping can be used against privacy, where an attacker can collect information in an unauthorized manner, obtain information about the origin and destination of a call, overhear a private conversation, or intercept personal information related to a mobile client’s account, for example. The eavesdropping attacks in mobile VoIP constitute a real menace as there are packet-sniffers largely available that can be used for eavesdropping on VoIP traffic. In particular, this attack can be easily mounted in a WLAN.

Using the information obtained by eavesdropping on the signaling, an attacker can manipulate fields in the media stream and make fraudulent VoIP calls or inject their own data. This threat can be easily mounted in VoIP as SIP messages and media streams are always sent unencrypted, in practice, to allow interoperability or ease the execution of functions by the wireless network (even though they are encrypted).

Masquerading Threats

These threats allow an entity to pretend being another entity. Masquerading can lead to call charging fraud, violation of privacy, and breaking of integrity. A masquerading attack can be carried out by hijacking a link after the authentication process has been performed or by eavesdropping of authentication information and subsequently replaying it. An attacker can steal the identity of a legitimate user and obtain access by masquerading as the real user. He can gain unauthorized access to mobile VoIP services.

The simplest form of masquerading in SIP is the reuse of username and password that can be obtained through interception. The authentication information can be obtained for the purpose of masquerade by reverse engineering of passwords, in the case of SIP digest authentication. To this end, the attacker may send several false challenges to the SIP user agent in the user’s terminal to generate a list that can be used to break the cryptographic hash of the password as computed by hash functions, such as MD5. A Masquerading attack can then be combined with modification of data to obtain access to services to place an unauthorized call.

Obtaining Control of an End System

When an attacker obtains physical access to an IP telephone, he is able to reset it to its default configuration and can provide backdoors or weak initial passwords. Moreover, the remote management interface that the IP telephone allows it to use is vulnerable to attacks. In addition, since the administrator password is sent in plain-text, this makes the communication vulnerable to sniffing. Finally, the administration password can also be attacked with a series of automated brute-force trials, since it has a limited length and restricted alphabet as it has to be typed in via the telephones keypad. Once the initial access to the device is achieved, its administrative password for both local and remote WWW access can be set to a value known and becomes usable for further malicious operations.

Attacking User’s Privacy

VoIP applications utilize RTP packets transmitted via the UDP protocol to carry audio data streams while using basic mechanisms for symmetrically encrypted audio payloads in RTP packets are described in an appropriate RTP profile. The eavesdropper is able to identify the data streams that form the audio connection(s) thanks to the public availability of VoIP protocol stacks and an in-detail description of protocol mechanisms, and despite the fact the ports used for these streams are typically negotiated in a dynamic manner.

Security Requirements for Mobile VoIP

In order for mobile VoIP to be widely adopted and generalized, a minimum set of security facilities must be provided. In particular, many experts argue that end-to-end authentication between the caller and the callee is not only possible, but it should also establish session keys, which can be used to protect the subsequent voice data stream.

- **A call should only be established with the callee that the caller expects:**

Securing the registration messages will defeat some of the simple redirection attacks. To ensure that the callee is really who the caller is communicating with, end-to-end authentication can be used.

Security Requirements for Mobile VoIP

- **The voice stream should be protected against eavesdropping:** The caller, say A, initiating a secure call to B should expect to be able to request privacy during call establishment and voice data transmission. The need for this service is probably greater for mobile VoIP than for VoIP and regular telephony, because the possibility of eavesdropping of an IP call is greater, in particular since many tools to do this are readily available. A solution to solve this can be based on session keys to encrypt and guarantee the integrity of the audio streams that are generated during a call.
- **Undesired calls should be blocked and VoIP spamming avoided:** Spamming and unwanted calls are mainly caused by the cost of VoIP, which will experience a similar situation as it is currently the case for email spams and unwanted mails. An authentication handshake at call establishment can be used to reject a call automatically based on user preferences.

Security Requirements for Mobile VoIP

- **Charging the call must be done correctly to the caller:** If charging calls is needed, its correctness is essential. Typically, however, one can assume that flat rate will be used for Internet calls (fixed monthly cost or free).
- **The information about caller's identity and who the caller is calling should not be revealed by eavesdropping:** This requires that the communication system must protect against eavesdropping. This can be achieved by encrypting the call setup messages and using TLS transport. The requirement related to callee's identity is hard to meet since, even if the system is able to protect all signaling activity from revealing the identity, there may be many other ways for an attacker to collect this information, by observing other traffic that the user sends and receives, for example.
- **An anonymous call service must be provided:** The caller may require that his identity should be hidden from the callee. The system should allow a caller to be anonymous. However, the callee should be able to reject such calls. In fact, introducing an initial authentication handshake does not exclude the possibility for the caller to remain anonymous.

End of the semester