RESEARCH ARTICLE OPEN ACCESS

ON CONFIDENTIALITY OF ENCRYPTED VOICE TELEPHONY

K.R.SRIVIDYA[1]

----------**********----------

Abstract:

Silence suppression, an essential feature of speech communications over the Internet, saves bandwidth by disabling voice packet transmissions when silence is detected. However, silence suppression enables an adversary to recover talk patterns from packet timing. In this paper, we investigate privacy leakage through the silence suppression feature. More specifically, we propose a new class of traffic analysis attacks to encrypted speech communications with the goal of detecting speakers of encrypted speech communications. These attacks are based on packet timing information only and the attacks can detect speakers of speech communications made with different codec's. We evaluate the proposed attacks with extensive experiments over different type of networks including commercial anonymity networks and campus networks. The experiments show that the proposed traffic analysis attacks can detect speakers of encrypted speech communications with high accuracy based on traces of 15 minutes long on average.

.----------**********----------

I. INTRODUCTION

In this paper, we propose a class of passive traffic analysis attacks to compromise privacy of encrypted speech communications. the procedure of the proposed attacks is as follows: first an adversary collects traces of encrypted speech communications made by a speaker, say alice. the adversary then extracts application-level features of alice's speech communications and trains a hidden markov model (hmm) with the extracted features. to test whether one speech communication of interest is made by alice, the adversary can extract features from the trace of interest and calculate the likelihood of the speech communication being made by alice. the proposed attacks can detect speakers of encrypted speech communications with high probabilities. in comparison with traditional traffic analysis attacks, the proposed traffic analysis attacks are different in the following aspects:

1) the proposed traffic analysis attacks do not require simultaneous accesses to one traffic flow of interest at both sides.

2) the attacks can detect speakers of encrypted speech communications made with different codecs.

The major difference between the new attacks and previous traffic analysis attacks on speech communications over the internet is: previous traffic analysis attacks are based on packet size information and the new attacks are based on packet timing information only. in this paper, we assume packet size information is not available for traffic analysis attacks because 1) voice packets generated by constant bit rate (cbr) codecs are of the same size, 2) encryption can pad voice packets to the same size during the encryption process, and 3) packets in anonymity networks such as tor are of the same size to prevent traffic analysis attacks based on packet size information. the contributions made in this paper are summarized as follows:

We propose a class of traffic analysis attacks to compromise privacy of encrypted speech communications. the attacks are passive and based

on the hmm, a powerful tool to model temporal data. we also propose a method to extract application-level features from traffic flows for application-level traffic analysis attacks.

we evaluate the proposed traffic analysis attacks through extensive experiments over the internet and commercial anonymity networks. for most of speech communications made in the experiments, the two communication parties are at least 20 hops away and the end-to-end delay is at least 80 ms. our experiments show that the traffic analysis attacks are able to detect speakers of encrypted speech communications with high probabilities based on only a small amount of encrypted traffic. we propose intersection attacks to improve the effectiveness of the proposed traffic analysis attacks. . we discuss possible countermeasures to mitigate the proposed traffic analysis attacks and analyze the effect of the countermeasures on the quality of speech communications.

SYSTEM ANALYSIS

Existing System:

The increasing popularity of speech communications over the Internet has brought a lot of attention and concern over security and privacy issues of these speech communications. To protect confidentiality of speech communications, tools and protocols such as Zfone, a tool capable to encrypt voice packets, and SRTP, the secure version of the realtime transport protocol (RTP), are developed or implemented. To further protect privacy of speech communications, advanced users are using anonymity networks to anonymize speech communications. For this purpose, low latency anonymity networks such as Tor and JAP can be used. One of the common anonymizing techniques used in anonymity networks is rerouting which usually routes packets through a randomly selected and usually longer path instead of the shortest path.

Proposed System:

We propose a new class of traffic analysis attacks to encrypted speech communications with the goal of detecting speakers of encrypted speech communications. These attacks are based on packet timing information only and the attacks can detect speakers of speech communications made with different codecs. We evaluate the proposed attacks with extensive experiments over different type of networks including commercial anonymity networks and campus networks. The experiments show that the proposed traffic analysis attacks can detect speakers of encrypted speech communications with high accuracy**.**

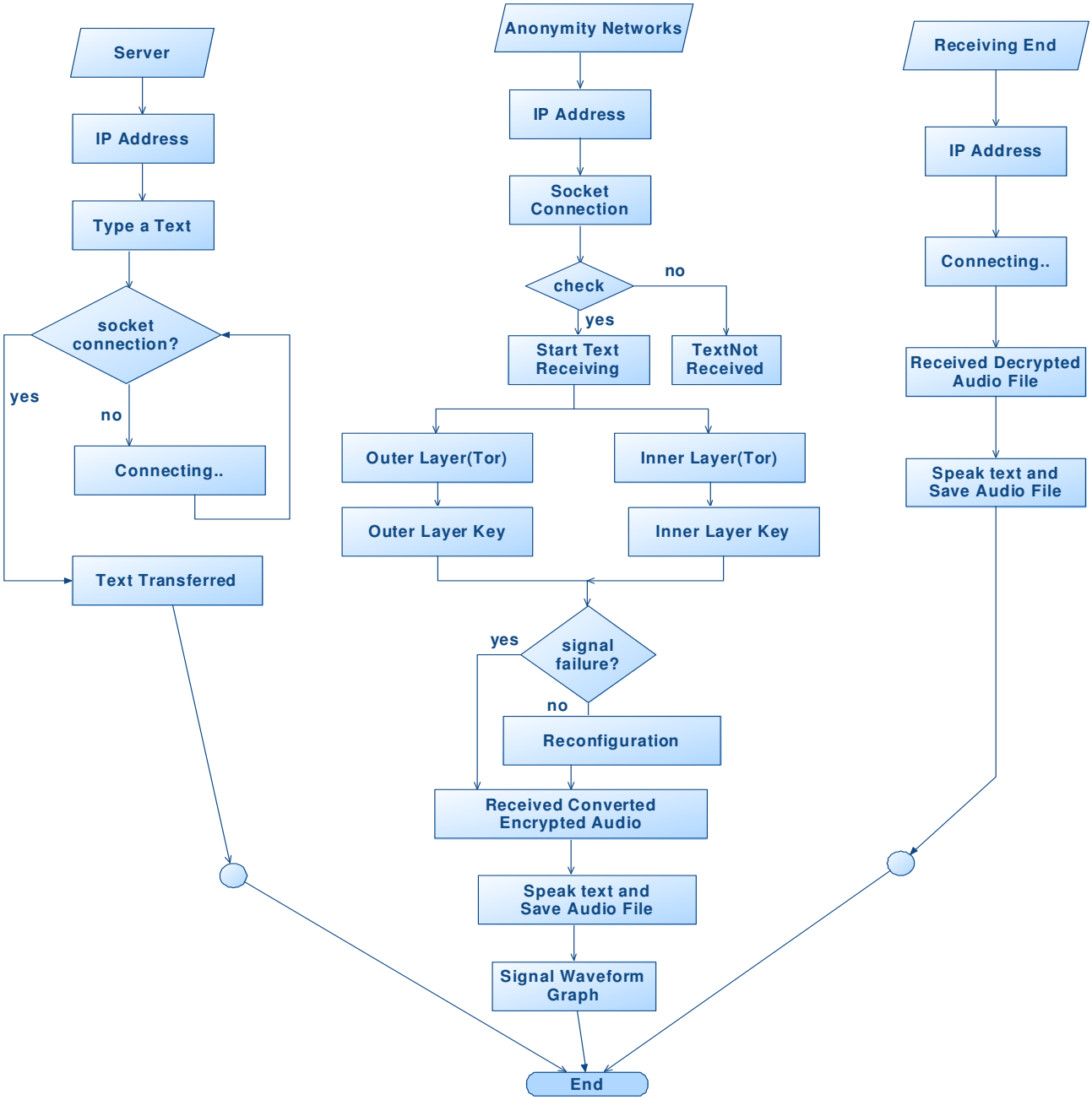
Advantages:

1. Packet size information.
2. Packet timing information.
3. Encryption of voice packets.

SYSTEM DESIGN

Data Flow Diagram / Use Case Diagram / Flow Diagram

The DFD is also called as bubble chart. It is a simple graphical formalism that can be used to represent a system in terms of the input data to the system, various processing carried out on these data, and the output data is generated by the system.



INPUT DESIGN AND OUTPUT DESIGN

INPUT DESIGN

The input design is the link between the information system and the user. It comprises the developing specification and procedures for data preparation and those steps are necessary to put transaction data in to a usable form for processing can be achieved by inspecting the computer to read data from a written or printed document or it can occur by having people keying the data directly into the system. The design of input focuses on controlling the amount of input required,

controlling the errors, avoiding delay, avoiding extra steps and keeping the process simple. The input is designed in such a way so that it provides security and ease of use with retaining the privacy. Input Design considered the following things:'

- What data should be given as input?
- How the data should be arranged or coded?
- The dialog to guide the operating personnel in providing input.
- Methods for preparing input validations and steps to follow when error occur.

OBJECTIVES

1.Input Design is the process of converting a user-oriented description of the input into a computer-based system. This design is important to avoid errors in the data input process and show the correct direction to the management for getting correct information from the computerized system.

2. It is achieved by creating user-friendly screens for the data entry to handle large volume of data. The goal of designing input is to make data entry easier and to be free from errors. The data entry screen is designed in such a way that all the data manipulates can be performed. It also provides record viewing facilities.

3.When the data is entered it will check for its validity. Data can be entered with the help of screens. Appropriate messages are provided as when needed so that the user will not be in maize of instant. Thus the objective of input design is to create an input layout that is easy to follow

OUTPUT DESIGN

A quality output is one, which meets the requirements of the end user and presents the information clearly. In any system results of processing are communicated to the users and to other system through outputs. In output design it is determined how the information is to be displaced for immediate need and also the hard copy output. It is the most important and direct source information to the user. Efficient and intelligent output design improves the system's relationship to help user decision-making.

1. Designing computer output should proceed in an organized, well thought out manner; the right output must be developed while ensuring that each output element is designed so that people will find the system can use easily and effectively. When analysis design computer output, they should Identify the specific output that is needed to meet the requirements.

2.Select methods for presenting information.

3.Create document, report, or other formats that contain information produced by the system.

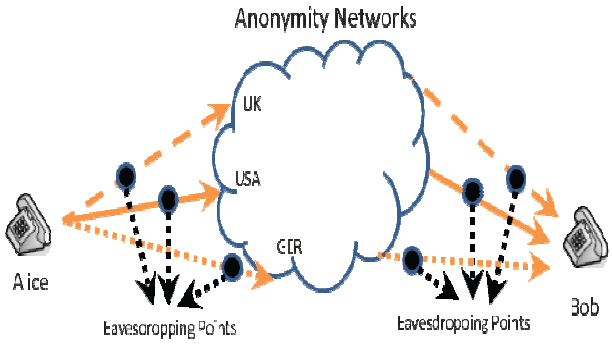
The output form of an information system should accomplish one or more of the following objectives.

- Convey information about past activities, current status or projections of the
- Future.
- Signal important events, opportunities, problems, or warnings.
- Trigger an action.
- Confirm an action.

1) IMPLEMENTATION

Implementation is the stage of the project when the theoretical design is turned out into a working system. Thus it can be considered to be the most critical stage in achieving a successful new system and in giving the user, confidence that the new system will work and be effective.

The implementation stage involves careful planning, investigation of the existing system and it's constraints on implementation, designing of ethods to achieve changeover and evaluation of changeover methods.



Algorithm

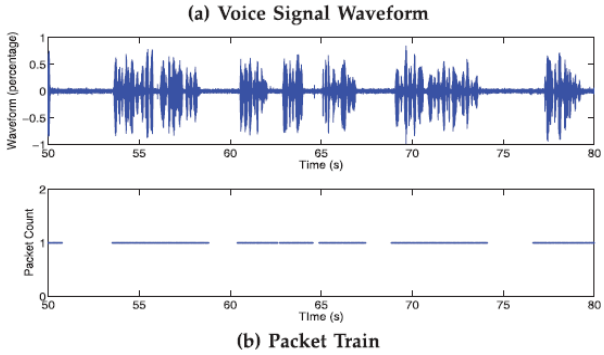
Encrypted Speech Communication A class of passive traffic analysis attacks to compromise privacy of encrypted speech communications. The proposed attacks can detect speakers of encrypted speech communications with high probabilities. In comparison with traditional traffic analysis attacks, the proposed traffic analysis attacks are different in the following aspects: 1) The proposed traffic analysis attacks do not require simultaneous accesses to one traffic flow of interest at both sides. 2) The attacks can detect speakers of encrypted speech communications made with different codecs.

Module Description:

1. Speech Coding:

In speech communications, an analog voice signal is first converted into a voice data stream by a chosen codec. Typically in this step, compression is used to reduce the data rate. The voice data stream is then packetized in small units of typically tens of milliseconds of voice, and encapsulated in a packet stream over the Internet. In this paper, we focus on constant bit rate codecs since most codecs used in current speech communications are CBR codecs.

2. Voice Signal Waveform:



Above Fig. shows the packet train generated by feeding the voice signal to X-Lite, a popular speech communication tool. From above Fig., we can easily observe the correspondence between the silence periods in the voice signal and the gaps in the packet train. The length of a silence period is different from the length of the corresponding gap in the packet train

3. Speaker Detection:

The inputs to this step are the Alice's HMM trained in the previous step and the feature vectors generated from a pool of raw speech communication traces of interest. The output of this step is the intermediate detection result, i.e., Ktop speakers from the candidate pool with talk patterns closest to Alice's talk pattern. The detection step can be divided into two phases: 1) First, the likelihood of each feature vector is calculated with the trained HMM. 2) The trace with the highest likelihood is declared as Alice's trace if the intersection step is not used. To improve the detection accuracy, the intermediate detection results can be fed into the optional intersection attack step.

4. Cross – Codec Detection:

In this set of experiments, the training traces and the traces to be detected are generated with different codecs. We believe this set of experiments is important because: 1) Practically training traces and the traces to be detected can be collected from speech communications made with different codecs. 2) Since speech packets are encrypted and possibly padded to a fixed length, adversaries may not be able to differentiate speech communications made with different codecs.

Literature Survey

Literature survey is the most important step in software development process. Before developing the tool it is necessary to determine the time factor, economy n company strength. Once

these things r satisfied, ten next steps is to determine which operating system and language can be used for developing the tool. Once the programmers start building the tool the programmers need lot of external support. This support can be obtained from senior programmers, from book or from websites. Before building the system the above consideration r taken into account for developing the proposed system.

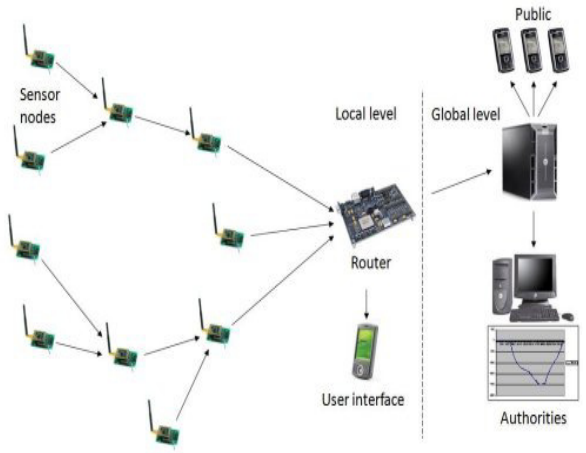
We have to analysis the ***Secure Computing***:

Networking:

In the world of computers, networking is the practice of linking two or more computing devices together for the purpose of sharing data. Networks are built with a mix of computer hardware and computer software.

Networks consist of the computers, wiring, and other devices, such as hubs, switches and routers that make up the network infrastructure. Some devices, such as network interface cards, serve as the computer's connection to the network. Devices such as switches and routers provide traffic- control strategies for the network. All sorts of different technologies can actually be employed to move data from one place to another, including wires, radio waves, and even microwave technology.

Network architecture:



Asynchronous Transfer Mode:

Asynchronous Transfer Mode (ATM) is a switching technique for telecommunication networks. It uses asynchronous time-division multiplexing and encodes data into small, fixed-sized cells. This differs from other protocols such as the Internet Protocol Suite or Ethernet that use variable sized packets or frames. ATM has similarity with both circuit and packet switched networking. This makes it a good choice for a network that must handle both traditional high-throughput data traffic, and real-time, low-latency content such as voice and video. ATM uses a connection-oriented model in which a virtual circuit must be established between two endpoints before the actual data exchange begins.

Network topology

Common layouts

A network topology is the layout of the interconnections of the nodes of a computer network. Common layouts are:

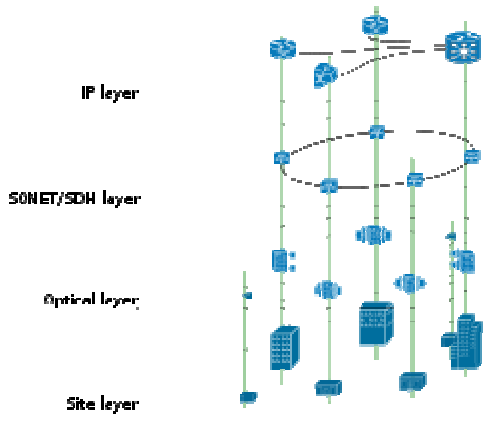
- A bus network: all nodes are connected to a common medium along this medium. This was the layout used in the original Ethernet, called 10BASE5 and 10BASE2.
- A star network: all nodes are connected to a special central node. This is the typical layout found in in a Wireless LAN, where each wireless client connects to the central Wireless access point.
- A ring network: each node is connected to its left and right neighbor node, such that all nodes are connected and that each node can reach each other node by traversing nodes left- or rightwards. The Fiber Distributed Data Interface (FDDI) made use of such a topology.
- A mesh network: each node is connected to an arbitrary number of neighbors in such a way that there is at least one traversal from any node to any other.

- A fully connected network: each node is connected to every other node in the network.

Note that the physical layout of the nodes in a network may not necessarily reflect the network topology. As an example, with FDDI, the network topology is a ring (actually two counter-rotating rings), but the physical topology is a star, because all neighboring connections are routed via a central physical location.

Overlay network

An overlay network is a virtual computer network that is built on top of another network. Nodes in the overlay are connected by virtual or logical links, each of which corresponds to a path, perhaps through many physical links, in the underlying network. The topology of the overlay network may (and often does) differ from that of the underlying one.



A sample overlay network: IP over SONET over Optical

For example, many peer-to-peer networks are overlay networks because they are organized as nodes of a virtual system of links run on top of the Internet. The Internet was initially built as an overlay on the telephone network .[14]

The most striking example of an overlay network, however, is the Internet itself: At the IP layer, each node can reach any other by a direct connection to the desired IP address, thereby creating a fully connected network; the underlying network, however, is composed of a mesh-like interconnect of subnetworks of varying topologies (and, in fact, technologies). Address resolution and routing are the means which allows the mapping of the fully-connected IP overlay network to the underlying ones.

Overlay networks have been around since the invention of networking when computer systems were connected over telephone lines using modems, before any data network existed.

Another example of an overlay network is a distributed hash table, which maps keys to nodes in the network. In this case, the underlying network is an IP network, and the overlay network is a table (actually map) indexed by keys.

Overlay networks have also been proposed as a way to improve Internet routing, such as through quality of service guarantees to achieve higher-quality streaming media. Previous proposals such as IntServ, DiffServ, and IP Multicast have not seen wide acceptance largely because they require modification of all routers in the network. On the other hand, an overlay network can be incrementally deployed on end-hosts running the overlay protocol software, without cooperation from Internet service providers. The overlay has no control over how packets are routed in the underlying network between two overlay nodes, but it can control, for example, the sequence of overlay nodes a message traverses before reaching its destination.

Routers

A router is an internetworking device that forwards packets between networks by processing information found in the datagram or packet (Internet protocol information from Layer 3 of the OSI Model). In many situations, this information is processed in conjunction with the routing table (also known as forwarding table). Routers use routing tables to determine what interface to

forward packets (this can include the "null" also known as the "black hole" interface because data can go into it, however, no further processing is done for said data).

Network security

In the field of networking, the area of **network security**[20] consists of the provisions and policies adopted by the network administrator to prevent and monitor unauthorized access, misuse, modification, or denial of the computer network and network-accessible resources. Network Security is the authorization of access to data in a network, which is controlled by the network administrator. Users are assigned an ID and password that allows them access to information and programs within their authority. Network Security covers a variety of computer networks, both public and private that are used in everyday jobs conducting transactions and communications among businesses, government agencies and individuals. Networks can be private, such as within a company, and others which might be open to public access. Network Security is involved in organization, enterprises, and all other type of institutions. It does as its titles explains, secures the network. Protects and oversees operations being done.

SYSTEM STUDY

FEASIBILITY STUDY

The feasibility of the project is analyzed in this phase and business proposal is put forth with a very general plan for the project and some cost estimates. During system analysis the feasibility study of the proposed system is to be carried out. This is to ensure that the proposed system is not a burden to the company. For feasibility analysis, some understanding of the major requirements for the system is essential.

Three key considerations involved in the feasibility analysis are

- ECONOMICAL FEASIBILITY
- TECHNICAL FEASIBILITY
- SOCIAL FEASIBILITY

1. ECONOMICAL FEASIBILITY

This study is carried out to check the economic impact that the system will have on the organization. The amount of fund that the company can pour into the research and development of the system is limited. The expenditures must be justified. Thus the developed system as well within the budget and this was achieved because most of the technologies used are freely available. Only the customized products had to be purchased.

2. TECHNICAL FEASIBILITY

This study is carried out to check the technical feasibility, that is, the technical requirements of the system. Any system developed must not have a high demand on the available technical resources. This will lead to high demands on the available technical resources. This will lead to high demands being placed on the client. The developed system must have a modest requirement, as only minimal or null changes are required for implementing this system.

3. SOCIAL FEASIBILITY

The aspect of study is to check the level of acceptance of the system by the user. This includes the process of training the user to use the system efficiently. The user must not feel threatened by the system, instead must accept it as a necessity. The level of acceptance by the users solely depends on the methods that are employed to educate the user about the system and to make him familiar with it. His level of confidence must be raised so that he is also able to make some constructive criticism, which is welcomed, as he is the final user of the system.

Software Environment

4.1Features OF. Net:

Microsoft .NET is a set of Microsoft software technologies for rapidly building and integrating XML Web services, Microsoft Windows-based applications, and Web solutions. The .NET Framework is a language-neutral platform for writing programs that can easily and securely interoperate. There's no language barrier with .NET: there are numerous languages available to the developer including Managed C++, C#, Visual Basic and Java Script.

SYSTEM TESTING

The purpose of testing is to discover errors. Testing is the process of trying to discover every conceivable fault or weakness in a work product. It provides a way to check the functionality of components, sub assemblies, assemblies and/or a finished product It is the process of exercising software with the intent of ensuring that the Software system meets its requirements and user expectations and does not fail in an unacceptable manner. There are various types of test. Each test type addresses a specific testing requirement.

TYPES OF TESTS

Unit testing

Unit testing involves the design of test cases that validate that the internal program logic is functioning properly, and that program inputs produce valid outputs. All decision branches and internal code flow should be validated. It is the testing of individual software units of the application .it is done after the completion of an individual unit before integration. This is a structural testing, that relies on knowledge of its construction and is invasive. Unit tests perform basic tests at component level and test a specific business process, application, and/or system configuration. Unit tests ensure that each unique path of a business process performs accurately to the documented specifications and contains clearly defined inputs and expected results.

Integration testing

Integration tests are designed to test integrated software components to determine if they actually run as one program. Testing is event driven and is more concerned with the basic outcome of screens or fields. Integration tests demonstrate that although the components were individually satisfaction, as shown by successfully unit testing, the combination of components is correct and consistent. Integration testing is specifically aimed at exposing the problems that arise from the combination of components.

Functional test

Functional tests provide systematic demonstrations that functions tested are available as specified by the business and technical requirements, system documentation, and user manuals.

Functional testing is centered on the following items:

- **Valid Input: identified classes of valid input must be accepted.**
- **Invalid Input: identified classes of invalid input must be rejected.**
- **Functions: identified functions must be exercised.**
- **Output: identified classes of application outputs must be exercised.**
- **Systems/Procedures: interfacing systems or procedures must be invoked.**

Organization and preparation of functional tests is focused on requirements, key functions, or special test cases. In addition, systematic coverage pertaining to identify Business process flows; data fields, predefined processes, and successive processes must be considered for testing. Before functional testing is complete, additional tests are identified and the effective value of current tests is determined.

System Test

System testing ensures that the entire integrated software system meets requirements. It tests a

configuration to ensure known and predictable results. An example of system testing is the configuration oriented system integration test. System testing is based on process descriptions and flows, emphasizing pre-driven process links and integration points.

White Box Testing

White Box Testing is a testing in which in which the software tester has knowledge of the inner workings, structure and language of the software, or at least its purpose. It is purpose. It is used to test areas that cannot be reached from a black box level.

Black Box Testing

Black Box Testing is testing the software without any knowledge of the inner workings, structure or language of the module being tested. Black box tests, as most other kinds of tests, must be written from a definitive source document, such as specification or requirements document, such as specification or requirements document. It is a testing in which the software under test is treated, as a black box .you cannot "see" into it. The test provides inputs and responds to outputs without considering how the software works.

Unit Testing:

Unit testing is usually conducted as part of a combined code and unit test phase of the software lifecycle, although it is not uncommon for coding and unit testing to be conducted as two distinct phases.

Test strategy and approach

Field testing will be performed manually and functional tests will be written in detail.

Test objectives

- All field entries must work properly.
- Pages must be activated from the identified link.
- The entry screen, messages and responses must not be delayed.

Features to be tested

- Verify that the entries are of the correct format
- No duplicate entries should be allowed
- All links should take the user to the correct page.

Integration Testing

Software integration testing is the incremental integration testing of two or more integrated software components on a single platform to produce failures caused by interface defects.

The task of the integration test is to check that components or software applications, e.g. components in a software system or – one step up – software applications at the company level – interact without error.

Test Results: All the test cases mentioned above passed successfully. No defects encountered.

Acceptance Testing

User Acceptance Testing is a critical phase of any project and requires significant participation by the end user. It also ensures that the system meets the functional requirements.

Test Results: All the test cases mentioned above passed successfully. No defects encountered.

System Specification

HARDWARE & SOFTWARE EQUIREMENTS:
HARDWARE REQUIREMENTS:

System	:	Pentium IV 2.4 GHz.
Hard Disk	:	40 GB.
Floppy Drive	:	1.44 Mb.
Monitor	:	15 VGA Color.
Mouse	:	Logitech.
Ram	:	512 MB.

SOFTWARE REQUIREMENTS:

Operating system	:	Windows XP Professional.
Coding Language	:	C#.NET

CONCLUSION

In this paper, we propose a class of passive traffic analysis attacks to compromise privacy of speech communications. The proposed attacks are based on application-level features extracted from speech communication traces. We evaluated the proposed attacks by extensive experiments over different types of networks including commercial anonymity networks and the campus network. The experiments show that the proposed traffic analysis attacks can detect speakers of encrypted speech communications with high detection rates based on speech communication traces of 15 minutes long on average.

Reference:

[1] S. Casner and S. Deering, "First ietf Internet Audiocast,"SIGCOMM Computer Comm. Rev., vol. 22,pp.92-97, http://doi.acm.org /10.1145/142267 .142338, July 1992.

[2] P. Zimmermann, A. Johnston, and J. Callas, "Zrtp: Media PathKey Agreement for Secure rtp Draft-Zimmermann-Avt-Zrtp-11,"RFC, United States, 2008.

[3] M. Baugher, D. McGrew, M. Naslund, E. Carrara, and K.Norrman, "The Secure Real-Time Transport Protocol (srtp)," 2004.

[4] R. Dingledine, N. Mathewson, and P. Syverson, "Tor: The Second-Generation Onion Router," Proc. 13th USENIX Security Symp.,pp. 303-320, Aug. 2004.

[5] O. Berthold, H. Federrath, and S. Ko¨ psell, "Web MIXes: A System for Anonymous and Unobservable Internet Access," Proc. Designing Privacy Enhancing Technologies: Workshop Design Issues in Anonymity and Unobservability, H. Federrath, ed., pp. 115-129, July 2000.

[6] J.M. Valin, "Speex: A Free Codec for Free Speech," Proc. Australian Nat'l Linux Conf., 2006.

[7] P.T. Brady, "A Technique for Investigating on-off Patterns of Speech," The Bell System Technical J., vol. 44, pp. 1-22, 1968.

[8] speex.org, "The Speex Projectpage ,"http://www.speex.org, 2005.

[9] S. Henning, "Voice Communication Across the Internet: ANetwork Voice Terminal," COINS technical report, Dept. OfComputer and information Science, Univ. of Massachusetts atAmherst, 1992.