I

Name.....

Roll No. .....Year 20.....20.....

Exam Seat No.....

COMPUTER GROUP | SEMESTER – VI | DIPLOMA IN ENGINEERING AND TECHNOLOGY

LEARNING MANUAL

FOR

Emerging Trends in Computer Engineering and Information Technology (22618)

MAHARASHTRA STATE BOARD OF TECHNICAL EDUCATION, MUMBAI

(Autonomous) (ISO 9001 : 2015) (ISO / IEC 27001 : 2013)

EMERGING TRENDS IN COMPUTER ENGINEERING AND INFORMATION TECHNOLOGY GROUP

INDEX

Unit Number	Topics and Sub-topics	Page Number
Unit I	1.1 Introduction of AI ● Concept ● Scope of AI ● Components of AI ● Types of AI ● Application of AI	**01**
	1.2 Data Visualization ● Data types in data visualization ● Scales map of data values in aesthetics ● Use of coordinate system in data visualization ● Use of colors to represent data values ● Representing - Amounts, Distribution, and Proportions	**11**
	1.3 Data Storytelling ● Introduction ● Ineffectiveness of Graphical representation of data ● Explanatory Analysis ○ Who ○ What ○ How	**22**
	1.4 Concept of machine learning and deep learning.	**25**
Unit II	2.1 Internet of Things (IoT) ● Definition ● Characteristics of IoT ● Features and Applications of IoT ● Advantages and Disadvantages of IoT	**29**
	2.1.2 Design of IoT ● Physical design of IoT ● Logical design of IoT	**34**
	2.1.3 IoT Protocols	**36**
	2.1.4 Sensors and actuators used in IoT	**39**
	2.2 Introduction to 5G Network • 5-G characteristics and application areas. • NGN architecture: Features, Functional block diagram, Network components: Media Gateway, Media Gateway Controller, and Application Server. • NGN Wireless Technology: Telecom network Spectrum: Types [licensed and unlicensed], Mobile Network Evolution (2G to 5G), Comparative features, • NGN Core: Features, Multi-Protocol Label Switching (MPLS): Concepts, Features and Advantages.	**46**

Unit III	3.1 Introduction to Blockchain ● Backstory of Blockchain ● What is Blockchain? 3.2 Centralize versus Decentralized System 3.3 Layers of Blockchain ● Application Layer ● Execution Layer ● Semantic Layer ● Propagation Layer ● Consensus Layer 3.4 Importance of Blockchain ● Limitations of Centralized Systems ● Blockchain Adoption So Far 3.5 Blockchain Use and Use Cases	**67** **70** **72** **74** **75**
Unit IV	4.1 Digital forensics ● Introduction to digital forensic ● Digital forensics investigation process ● Models of Digital Forensic Investigation - o Abstract Digital Forensics Model (ADFM) o Integrated Digital Investigation Process (IDIP) o An extended model for cybercrime investigation 4.2 Ethical issues in digital forensic ● General ethical norms for investigators ● Unethical norms for investigation 4.3 Digital Evidences ● Definition of Digital Evidence ● Best evidence rule ● Original Evidence 4.4 Characteristics of Digital Evidence ● Locard's Exchange Principle ● Digital Stream of bits 4.5 Types of evidence ● Illustrative, Electronics, Documented, Explainable, Substantial, Testimonial 4.6 Challenges in evidence handling ● Authentication of evidence ● Chain of custody ● Evidence validation 4.7 Volatile evidence	**78** **80** **80** **81** **82** **85** **87** **87** **92**
Unit V	5.1 Ethical Hacking ● How Hackers Beget Ethical Hackers ● Defining hacker, Malicious users ● Data Privacy and General Data Protection and Regulation(GDPR) 5.2 Understanding the need to hack your own systems 5.3 Understanding the dangers your systems face ● Non Technical attacks	**102** **112** **114**

	● Network-infrastructure attacks ● Operating-system attacks ● Application and other specialized attacks 5.4 Obeying the Ethical hacking Principles ● Working ethically ● Respecting privacy ● Not crashing your systems 5.5 The Ethical Hacking Process ● Formulating your plan ● Selecting tools ● Executing the plan ● Evaluating results ● Moving on 5.6 Cyber Security act	**115** **116** **120**
Unit VI	6.1 Network Hacking Network Infrastructure: ● Network Infrastructure Vulnerabilities ● Scanning-Ports ● Ping sweep ● Scanning SNMP ● Grabbing Banners ● MAC-daddy attack Wireless LANs: ● Wireless Network Attacks 6.2 Operating System Hacking ● Introduction of Windows and Linux Vulnerabilities ● Buffer Overflow Attack 6.3 Applications Hacking Messaging Systems: ● Vulnerabilities ● E-Mail Attacks- E-Mail Bombs ● Banners ● Best practices for minimizing e-mail security risks Web Applications: ● Web Vulnerabilities ● Directories Traversal and Countermeasures ● Google Dorking Database system ● Database Vulnerabilities ● Best practices for minimizing database security risks	**124** **125** **133** **135** **139**

Unit-1 Artificial Intelligence

Contents

1.1 Introduction of AI

- Concept
- Scope of AI
- Components of AI
- Types of AI
- Application of AI

1.2 Data Visualization

- Data types in data visualization
- Scales map of data values in aesthetics
- Use of coordinate system in data visualization
- Use of colors to represent data values
- Representing - Amounts, Distribution, and Proportions

1.3 Data Storytelling

- Introduction
- Ineffectiveness of Graphical representation of data
- Explanatory Analysis
    - Who
    - What
    - How

1.4 Concept of machine learning and deep learning.

1.1 Introduction of AI

A branch of Computer Science named Artificial Intelligence (AI)pursues creating the computers / machines as intelligent as human beings. John McCarthy the father of Artificial Intelligence described AI as, **"The science and engineering of making intelligent machines, especially intelligent computer programs".** Artificial Intelligence (AI) is a branch of science which deals with helping machines find solutions to complex problems in a more human-like fashion.

Artificial is defined in different approaches by various researchers during its evolution, such as "Artificial Intelligence is the study of how to make computers do things which at the moment, people do better."

There are other possible definitions "like AI is a collection of hard problems which can be solved by humans and other living things, but for which we don't have good algorithms for solving." e. g., understanding spoken natural language, medical diagnosis, circuit design, learning, self-adaptation, reasoning, chess playing, proving math theories, etc.

- **Data:** Data is defined as symbols that represent properties of objects events and their environment.
- **Information:** Information is a message that contains relevant meaning, implication, or input for decision and/or action.
- **Knowledge:** It is the (1) cognition or recognition (know-what), (2) capacity to act(know-how), and (3) understanding (know-why) that resides or is contained within the mind or in the brain.
- **Intelligence:** It requires ability to sense the environment, to make decisions, and to control action.

1.1.1 Concept:

Artificial Intelligence is one of the emerging technologies that try to simulate human reasoning in AI systems The art and science of bringing learning, adaptation and self-organization to the machine is the art of Artificial Intelligence. Artificial Intelligence is the ability of a computer program to learn and think. Artificial intelligence (AI) is an area of computer science that emphasizes the creation of intelligent machines that work and reacts like humans. AI is built on these three important concepts

Machine learning: When you command your smartphone to call someone, or when you chat with a customer service chatbot, you are interacting with software that runs on AI. But this type of software actually is limited to what it has been programmed to do. However, we expect to soon have systems that can learn new tasks without humans having to guide them. The idea is to give them a large number of examples for any given chore, and they should be able to process each one and learn how to do it by the end of the activity.

Deep learning: The machine learning example I provided above is limited by the fact that humans still need to direct the AI's development. In deep learning, the goal is for the software to use what it has learned in one area to solve problems in other areas. For example, a program that has learned how to distinguish images in a photograph might be able to use this learning to seek out patterns in complex graphs.

Neural networks: These consist of computer programs that mimic the way the human brain processes information. They specialize in clustering information and recognizing complex patterns, giving computers the ability to use more sophisticated processes to analyze data.

1.1.2 Scope of AI:

The ultimate goal of artificial intelligence is to create computer programs that can solve problems and achieve goals like humans would. There is scope in developing machines in robotics, computer vision, language detection machine, game playing, expert systems, speech recognition machine and much more.

The following factors characterize a career in artificial intelligence:

- Automation
- Robotics
- The use of sophisticated computer software

Individuals considering pursuing a career in this field require specific education based on the foundations of math, technology, logic and engineering perspectives. Apart from these, good communication skills (written and verbal) are imperative to convey how AI services and tools will help when employed within industry settings.

AI Approach:

The difference between machine and human intelligence is that the human think / act rationally compares to machine. Historically, all four approaches to AI have been followed, each by different people with different methods.

	like Humans	Well
Think	GPS	Rational Agent
Act	Eliza	Heuristic Systems

Figure 1.1 AI Approaches

Think Well:

Develop formal models of knowledge representation, reasoning, learning, memory, problem solving that can be rendered in algorithms. There is often an emphasis on a system that are provably correct, and guarantee finding an optimal solution.

Act Well:

For a given set of inputs, generate an appropriate output that is not necessarily correct but gets the job done.

- A heuristic (heuristic rule, heuristic method) is a rule of thumb, strategy, trick, simplification, or any other kind of device which drastically limits search for solutions in large problem spaces.
- Heuristics do not guarantee optimal solutions; in fact, they do not guarantee any solution at all:
- all that can be said for a useful heuristic is that it offers solutions which are good

enough most of the time

Think like humans:

Cognitive science approach. Focus not just on behavior and I/O but also look at reasoning process. The Computational model should reflect "how" results were obtained. Provide a new language for expressing cognitive theories and new mechanisms for evaluating them. GPS (General Problem Solver): Goal not just to produce humanlike behavior (like ELIZA), but to produce a sequence of steps of the reasoning process that was similar to the steps followed by a person in solving the same task.

Act like humans:

Behaviorist approach-Not interested in how you get results, just the similarity to what human results is.

Example:

ELIZA: A program that simulated a psychotherapist interacting with a patient and successfully passed the Turing Test. It was coded at MIT during 1964-1966 by Joel Weizenbaum. First script was DOCTOR. The script was a simple collection of syntactic patterns not unlike regular expressions. Each pattern had an associated reply which might include bits of the input (after simple transformations (my →your) Weizenbaum was shocked at reactions: Psychiatrists thought it had potential. People unequivocally anthropomorphized.

1.1.3 Components of AI

The core components and constituents of AI are derived from the concept of logic, cognition and computation; and the compound components, built-up through core components are knowledge, reasoning, search, natural language processing, vision etc.

Level	Core	Compound	Coarse components
Logic	Induction Proposition Tautology Model Logic	Knowledge Reasoning Control Search	Knowledge based systems, Heuristic Search Theorem Proving
Cognition	Temporal Learning Adaptation Self-organization	Belief Desire Intention	Multi Agent system Co-operation, Co-ordination AI Programming
Functional	Memory Perception	Utterance	Vision Natural Language Speech Processing

The core entities are inseparable constituents of AI in that these concepts are fused at atomic level. The concepts derived from logic are propositional logic, tautology, predicate calculus, model and temporal logic. The concepts of cognitive science are of two types: one is functional which includes learning, adaptation and self-organization, and the other is memory and perception which are physical entities. The physical entities generate some functions to make the compound components.

The compound components are made of some combination of the logic and cognition stream. These are knowledge, reasoning and control generated from constituents of logic such as predicate calculus, induction and tautology and some from cognition (such as learning and adaptation). Similarly, belief, desire and intention are models of mental states that are predominantly based on cognitive components but less on logic. Vision, utterance (vocal) and expression (written) are combined effect of memory and perceiving organs or body sensors such as ear, eyes and vocal. The gross level contains the constituents at the third level which are knowledge-based systems (KBS), heuristic search, automatic theorem proving, multi-agent systems, Al languages such as PROLOG and LISP, Natural language processing (NLP). Speech processing and vision are based mainly on the principle of pattern recognition.

AI Dimension: The philosophy of Al in three-dimensional representations consists in logic, cognition and computation in the x-direction, knowledge, reasoning and interface in the y-direction. The x-y plane is the foundation of AI. The z-direction consists of correlated systems of physical origin such as language, vision and perception as shown in Figure.1.2

Figure 1.2 Three-dimensional model of AI

The First Dimension (Core)

The theory of logic, cognition and computation constitutes the fusion factors for the formation of one of the foundations on coordinate x-axis. Philosophy from its very inception of origin covered all the facts, directions and dimensions of human thinking output. Aristotle's theory of syllogism, Descartes and Kant's critic of pure reasoning and contribution of many other philosophers made knowledge-based on logic. It was Charles Babbage and Boole who demonstrated the power of computation logic. Although the modern philosophers such as Bertrand Russell correlated logic with mathematics but it was Turing who developed the theory of computation for mechanization. In the 1960s, Marvin Minsky pushed the logical formalism to integrate reasoning with knowledge.

Cognition:

Computers has become so popular in a short span of time due to the simple reason that they adapted and projected the information processing paradigm (IPP) of human beings: sensing organs as input, mechanical movement organs as output and the central nervous system (CNS) in brain as control and computing devices, short-term and long-term memory were not distinguished by computer scientists but, as a whole, it was in conjunction, termed memory.

In further deepening level, the interaction of stimuli with the stored information to produce new information requires the process of learning, adaptation and self-organization. These functionalities in the information processing at a certain level of abstraction of brain activities demonstrate a state of mind which exhibits certain specific behavior to qualify as intelligence. Computational models were developed and incorporated in machines which mimicked the functionalities of human origin. The creation of such traits of human beings in the computing devices and processes originated the concept of intelligence in machine as virtual mechanism. These virtual machines were termed in due course of time artificial intelligent machines.

Computation

The theory of computation developed by Turing-finite state automation—was a turning point in mathematical model to logical computational. Chomsky's linguistic computational theory generated a model for syntactic analysis through a regular grammar.

The Second Dimension

The second dimension contains knowledge, reasoning and interface which are the components of knowledge-based system (KBS). Knowledge can be logical; it may be processed as information which is subject to further computation. This means that any item on the y-axis is correlated with any item on the x-axis to make the foundation of any item on the z-axis. Knowledge and reasoning are difficult to prioritize, which occurs first: whether knowledge is formed first and then reasoning is performed or as reasoning

is present, knowledge is formed. Interface is a means of communication between one domain to another. Here, it connotes a different concept then the user's interface. The formation of a permeable membrane or transparent solid structure between two domains of different permittivity is termed interface. For example, in the industrial domain, the robot is an interface. A robot exhibits all traits of human intelligence in its course of action to perform mechanical work. In the KBS, the user's interface is an example of the interface between computing machine and the user. Similarly, a program is an interface between the machine and the user. The interface may be between human and human, i.e. experts in one domain to experts in another domain. Human-to- machine is program and machine-to-machine is hardware. These interfaces are in the context of computation and AI methodology.

The Third Dimension

The third dimension leads to the orbital or peripheral entities, which are built on the foundation of x-y plane and revolve around these for development. The entities include an information system. NLP, for example, is formed on the basis of the linguistic computation theory of Chomsky and concepts of interface and knowledge on y-direction. Similarly, vision has its basis on some computational model such as clustering, pattern recognition computing models and image processing algorithms on the x-direction and knowledge of the domain on the y-direction.

The third dimension is basically the application domain. Here, if the entities are near the origin, more and more concepts are required from the x-y plane. For example, consider information and automation, these are far away from entities on z-direction, but contain some of the concepts of cognition and computation model respectively on x-direction and concepts of knowledge (data), reasoning and interface on the y-direction.

In general, any quantity in any dimension is correlated with some entities on the other dimension. The implementation of the logical formalism was accelerated by the rapid growth in electronic technology, in general and multiprocessing parallelism in particular.

1.1.4 Types of AI

Artificial Intelligence can be divided in various types, there are mainly two types of main categorization which are based on capabilities and based on functionally of AI. Following is flow diagram which explain the types of AI.

Figure 1.3 Types of AI

Types of AI

AI type-1: Based on Capabilities

1. Weak AI or Narrow AI:

- Narrow AI is a type of AI which is able to perform a dedicated task with intelligence. The most common and currently available AI is Narrow AI in the world of Artificial Intelligence.
- Narrow AI cannot perform beyond its field or limitations, as it is only trained for one specific task. Hence it is also termed as weak AI. Narrow AI can fail in unpredictable ways if it goes beyond its limits.
- Apple Siri is a good example of Narrow AI, but it operates with a limited pre-defined range of functions.
- IBM's Watson supercomputer also comes under Narrow AI, as it uses an Expert system approach combined with Machine learning and natural language processing.
- Some Examples of Narrow AI are playing chess, purchasing suggestions on e-commerce site, self-driving cars, speech recognition, and image recognition.

2. General AI:

- General AI is a type of intelligence which could perform any intellectual task with efficiency like a human.
- The idea behind the general AI to make such a system which could be smarter and think like a human by its own.
- Currently, there is no such system exist which could come under general AI and can perform any task as perfect as a human.
- The worldwide researchers are now focused on developing machines with General AI.
- As systems with general AI are still under research, and it will take lots of efforts and time to develop such systems.

3. Super AI:

- Super AI is a level of Intelligence of Systems at which machines could surpass human intelligence, and can perform any task better than human with cognitive properties. It is an outcome of general AI.
- Some key characteristics of strong AI include capability include the ability to

think, to reason, solve the puzzle, make judgments, plan, learn, and communicate by its own.

●Super AI is still a hypothetical concept of Artificial Intelligence. Development of such systems in real is still world changing task.

Artificial Intelligence type-2: Based on functionality

1. Reactive Machines

- Purely reactive machines are the most basic types of Artificial Intelligence.
- Such AI systems do not store memories or past experiences for future actions.
- These machines only focus on current scenarios and react on it as per possible best action.
- IBM's Deep Blue system is an example of reactive machines.
- Google's AlphaGo is also an example of reactive machines.
- Limited Memory
- Limited memory machines can store past experiences or some data for a short period of time.
- These machines can use stored data for a limited time period only.
- Self-driving cars are one of the best examples of Limited Memory systems. These cars can store recent speed of nearby cars, the distance of other cars, speed limit, and other information to navigate the road.

2. Theory of Mind

- Theory of Mind AI should understand the human emotions, people, beliefs, and be able to interact socially like humans.
- This type of AI machines are still not developed, but researchers are making lots of efforts and improvement for developing such AI machines.

3. Self-Awareness

- Self-awareness AI is the future of Artificial Intelligence. These machines will be super intelligent, and will have their own consciousness, sentiments, and self-awareness.
- These machines will be smarter than human mind.
- Self-Awareness AI does not exist in reality still and it is a hypothetical concept.

1.1.5 Application of AI

AI has been dominant in various fields such as –

●**Gaming:** AI plays crucial role in strategic games such as chess, poker, tic-tac-toe, etc., where machine can think of large number of possible positions based on heuristic knowledge.

●**Natural Language Processing:** It is possible to interact with the computer that understands natural language spoken by humans.

●**Expert Systems:** There are some applications which integrate machine,

software, and special information to impart reasoning and advising. They provide explanation and advice to the users.

●**Vision Systems:** These systems understand, interpret, and comprehend visual input on the computer. For example,

- A spying aeroplane takes photographs, which are used to figure out spatial information or map of the areas.
- Doctors use clinical expert system to diagnose the patient.
- Police use computer software that can recognize the face of criminal with the stored portrait made by forensic artist.

●**Speech Recognition:** Some intelligent systems are capable of hearing and comprehending the language in terms of sentences and their meanings while a human talks to it. It can handle different accents, slang words, noise in the background, change in human's noise due to cold, etc.

●**Handwriting Recognition:** The handwriting recognition software reads the text written on paper by a pen or on screen by a stylus. It can recognize the shapes of the letters and convert it into editable text.

●**Intelligent Robots:** Robots are able to perform the tasks given by a human. They have sensors to detect physical data from the real world such as light, heat, temperature, movement, sound, bump, and pressure. They have efficient processors, multiple sensors and huge memory, to exhibit intelligence. In addition, they are capable of learning from their mistakes and they can adapt to the new environment.

1.2 Data Visualization

1.2.1 Introduction –

Data visualization is the graphical representation of information and data. By using visual elements like charts, graphs, and maps, data visualization tools provide an accessible way to see and understand trends, outliers, and patterns in data. It also provides an excellent way to present data to non-technical audiences without confusion. The first and foremost objective of data visualization is to convey data correctly. Whenever we visualize data, we take data values and convert them in a systematic and logical way into the visual elements that make up the final graphic. Even though there are many different types of data visualizations, and on first glance a scatterplot, a pie chart, and a heatmap don't seem to have much in common, all these visualizations can be described with a common language that captures how data values are turned into blobs of ink on paper or colored pixels on a screen. The key insight is the following: all data visualizations map data values into quantifiable features of the resulting graphic. We refer to these features as aesthetics.

1.2.2 Data types in data visualization –

When we consider types of data in data visualization, we consider various types of data in use as well as aesthetics too. Aesthetics describe every aspect of a given graphical element. For example, in Figure 1.4 -

A critical component of every graphical element is of course its position, which describes where the element is located. In standard 2D graphics, we describe positions by an x and y value, but other coordinate systems and one- or three-dimensional visualizations are possible. Next, all graphical elements have a shape, a size, and a color. Even if we are preparing a black-and-white drawing, graphical elements need to have a color to be visible: for example, black if the background is white or white if the background is black. Finally, to the extent we are using lines to visualize data, these lines may have different widths or dash–dot patterns. There are many other aesthetics may encountered in a data visualization. For example, if we want to display text, we may have to specify font family, font face, and font size, and if graphical objects overlap, we may have to specify whether they are partially transparent.

Figure 1.4 Commonly used aesthetics in data visualization: position, shape, size, color, line width, line type. Some of these aesthetics can represent both continuous and discrete data (position, size, line width, color), while others can usually only represent

All aesthetics are categorized into two groups:

(1) *Those that can represent continuous data* and

(2) *those that cannot represent continuous data*.

Continuous data values are values for which arbitrarily fine intermediates exist. For example, time duration is a continuous value. Between any two durations, say

50 seconds and 51 seconds, there are arbitrarily many intermediates, such as 50.5 seconds, 50.51 seconds, 50.50001 seconds, and so on. By contrast, number of persons in a room is a discrete value. A room can hold 5 persons or 6, but not 5.5. For the examples in Figure 1.4, position, size, color, and line width can represent continuous data, but shape and line type can usually only represent discrete data. Next, we'll consider the types of data we may want to represent in our visualization. You may think of data as numbers, but numerical values are only two out of several types of data we may encounter. In addition to continuous and discrete numerical values, data can come in the form of discrete categories, in the form of dates or times, and as text (Table 1.1). When data is numerical, we also call it quantitative and when it is categorical, we call it qualitative. Variables holding qualitative data are factors, and the different categories are called levels. The levels of a factor are most commonly without order (as in the example of dog, cat, fish in Table 1.1 given below, but factors can also be ordered, when there is an intrinsic order among the levels of the factor (as in the example of good, fair, poor in Table 1.1).

Table 1.1 Types of variables encountered in Data Visualization Scenario

Types of Variables	Example	Appropriate Scale	Description
Quantitative/ numerical continuous	1.3, 5.7, 83, 1.5 × 10–2	Continuous	Arbitrary numerical values. These can be integers, rational numbers, or real numbers.
Quantitative/ numerical discrete	1, 2, 3, 4	Discrete	Numbers in discrete units. These are most commonly but not necessarily integers. For example, the numbers 0.5, 1.0, 1.5 could also be treated as discrete if intermediate values cannot exist in the given dataset.
Qualitative/ categorical unordered	dog, cat, fish	Discrete	Categories without order. These are discrete and unique categories that have no inherent order. These variables are also called factors.
Qualitative/ categorical ordered	good, fair, poor	Discrete	Categories with order. These are discrete and unique categories with an order. For example, "fair" always lies between "good" and "poor." These variables are also called ordered factors.
Date or time	Jan. 5 2018, 8:03am	Continuous or Discrete	Specific days and/or times. Also, generic dates, such as July 4 or Dec. 25 (without year).

Text	The quick brown fox jumps over the lazy dog.	None, or discrete	Free-form text. Can be treated as categorical if needed.

Let's consider an example, the below Table 1.2 shows the first few rows of a dataset providing the daily temperature normal (aver- age daily temperatures over a 30-year window) for four US locations. This table contains five variables: month, day, location, station ID, and temperature (in degrees Fahrenheit). Month is an ordered factor, day is a discrete numerical value, location is an unordered factor, station ID is similarly an unordered factor, and temperature is a continuous numerical value.

Table 1.2 First 8 rows of a dataset listing daily temperature normal for four weather stations

Month	Day	Location	Section ID	Temperature (F)
Jan	1	Chicago	USW00014819	25.6
Jan	1	San Diego	USW00093107	55.2
Jan	1	Houston	USW00012918	53.9
Jan	1	Death Valley	USC00042319	51.0
Jan	2	Chicago	USW00014819	25.5
Jan	2	San Diego	USW00093107	55.3
Jan	2	Houston	USW00012918	53.8
Jan	2	Death Valley	USC00042319	51.2

Data source: National Oceanic and Atmospheric Administration (NOAA).

Figure 1.5 The data representation of the above-mentioned example

Figure 1.6 Monthly normal mean temperatures for the same example

1.2.3 Use of coordinate system in Data Visualization

To make any sort of data visualization, we need to define position scales, which deter- mine where in graphic different data values are located. We cannot visualize data without placing different data points at different locations, even if we just arrange them next to each other along a line. For regular 2D visualizations, two numbers are required to uniquely specify a point, and therefore we need two position scales. These two scales are usually but not necessarily the x and y axes of the plot. We also have to specify the relative geometric arrangement of these scales. Conventionally, the x axis runs horizontally and the y axis vertically, but we could choose other arrangements. For example, we could have the y axis run at an acute angle relative to the x axis, or we could have one axis run in a circle and the other run radially. The combination of a set of position scales and their relative geometric arrangement is called a coordinate system.

- **Cartesian coordinates –**

The most widely used coordinate system for data visualization is the 2D Cartesian coordinate system, where each location is uniquely specified by an x and a y value. The x and y axes run orthogonally to each other, and data values are placed in an even spacing along both axes. The two axes are continuous position scales, and they can represent both positive and negative real numbers. To fully specify the coordinate system, we need to specify the range of numbers each axis covers. Any data values between these axis limits are placed at the appropriate respective location in the plot.

Any data values outside the axis limits are discarded.

Figure 1.7 Cartesian coordinate system's sample example

Figure 1.8 Daily temperature normals for Huston using different aspect ratio

A Cartesian coordinate system can have two axes representing two different units. This situation arises quite commonly whenever we're mapping two different types of variables to x and y. For example, consider below image, if we plot temperature versus days of the year. The y axis of is measured in degrees Fahrenheit, with a grid line every at 20 degrees, and the x axis is measured in months, with a grid line at the first of every third month. Whenever the two axes are measured in different units, we can stretch or compress one relative to the other and maintain a valid visualization of the data. Which version is preferable may depend on the story we want to convey. A tall and narrow figure emphasizes change along the y axis and a short and wide figure does the opposite. Ideally, we want to choose an aspect ratio that ensures that any important differences in position are noticeable.

- **Nonlinear Axes –**

In a Cartesian coordinate system, the grid lines along an axis are spaced evenly both in data units and in the resulting visualization. We refer to the position scales in these

coordinate systems as *linear*. While linear scales generally provide an accurate representation of the data, there are scenarios where nonlinear scales are preferred. In a nonlinear scale, even spacing in data units corresponds to uneven spacing in the visualization, or conversely even spacing in the visualization corresponds to uneven spacing in data units. The most commonly used nonlinear scale is the logarithmic scale, or log scale for short. Log scales are linear in multiplication, such that a unit step on the scale corresponds to multiplication with a fixed value. To create a log scale, we need to log- transform the data values while exponentiating the numbers that are shown along the axis grid lines. This process is demonstrated in Figure 3-4, which shows the numbers 1, 3.16, 10, 31.6, and 100 placed on linear and log scales. The numbers 3.16 and 31.6 may seem like strange choices, but they were selected because they are exactly half-way between 1 and 10 and between 10 and 100 on a log scale.

Figure 1.9 Representation of Linear and logarithmic scales

- **Coordinate systems with curved axes –**

All the coordinate systems we have encountered so far have used two straight axes positioned at a right angle to each other, even if the axes themselves established a nonlinear mapping from data values to positions. There are other coordinate systems, however, where the axes themselves are curved. In particular, in the *polar* coordinate system, we specify positions via an angle and a radial distance from the origin, and therefore the angle axis is circular. Polar coordinates can be useful for data of a periodic nature, such that data values at one end of the scale can be logically joined to data values at the other end. For example, consider the days in a year. December 31st is the last day of the year, but it is also one day before the first day of the year. If we want to show how some quantity varies over the year, it can be appropriate to use polar coordinates with the angle coordinate specifying each day. Let's apply this concept to the temperature normals. Because temperature normals are average temperatures that are not tied to any specific year, Dec. 31st can be thought of as 366 days later than Jan. 1st (temperature normals include Feb. 29th) and also 1 day earlier.

By plotting the temperature normals in a polar coordinate system, we emphasize this cyclical property they have. The polar version highlights how similar the temperatures

are in Death Valley, Houston, and San Diego from late fall to early spring. In the Cartesian coordinate system, this fact is obscured because the temperature values in late December and in early January are shown in opposite parts of the figure and therefore don't form a single visual unit.

Figure 1. 10 Representation of data on curved axes

1.2.4 Use of colors to represent data values

There are three fundamental use cases for color in data visualizations:

i. to distinguish groups of data from each other,
ii. to represent data values, and
iii. to highlight.

The types of colors we use and the way in which we use them are quite different for these three cases.

i. <u>Color as a tool to distinguish –</u>

We frequently use color as a means to distinguish discrete items or groups that do not have an intrinsic order, such as different countries on a map or different manufactures of a certain product. In this case, we use a qualitative color scale.

Such a scale contains a finite set of specific colors that are chosen to look clearly distinct from each other while also being equivalent to each other. The second condition requires that no one color should stand out relative to the others. Also, the colors should not create the impression of an order, as would be the case with a sequence of colors that get successively lighter. Such colors would create an apparent order among the items being colored, which by definition have no order. Many appropriate qualitative color scales are readily available. Figure 4-1 shows three representative examples. In particular, the ColorBrewer project provides a nice selection of qualitative color scales, including both fairly light and fairly dark colors [Brewer 2017].

Figure 1.11. Example qualitative color scales. The Okabe Ito scale is the default scale used throughout this book [Okabe and Ito 2008]. The ColorBrewer Dark2 scale is provided by the ColorBrewer project [Brewer 2017]. The ggplot2 hue scale is the default qualitative scale in the widely used plotting software ggplot2.

Figure 1.12 Representing data using various colors to distinguish regions

ii. Color as a tool to represent data values –

Color can also be used to represent quantitative data values, such as income, temperature, or speed. In this case, we use a sequential color scale. Such a scale contains a sequence of colors that clearly indicate which values are larger or smaller than which other ones, and how distant two specific values are from each other. The second point implies that the color scale needs to be perceived to vary uniformly across its entire range. Sequential scales can be based on a single hue (e.g., from dark blue to light blue) or on multiple hues (e.g., from dark red to light yellow). Multihued scales tend to follow color gradients that can be seen in the natural world, such as dark red, green, or blue to light yellow, or dark purple to light green. The reverse (e.g., dark yellow to light blue) looks unnatural and doesn't make a useful sequential scale.

Figure 1.13. Example sequential color scales. The ColorBrewer Blues scale is a monochromatic scale that varies from dark to light blue. The Heat and Viridis scales are multihue scales that vary from dark red to light yellow and from dark blue via green to light yellow, respectively.

iii. Color as a tool to highlight –

Color can also be an effective tool to highlight specific elements in the data. There may be specific categories or values in the dataset that carry key information about the story we want to tell, and we can strengthen the story by emphasizing the relevant figure elements to the reader. An easy way to achieve this emphasis is to color these figure elements in a color or set of colors that vividly stand out against the rest of the figure. This effect can be achieved with accent color scales, which are color scales that contain both a set of subdued colors and a matching set of stronger, darker, and/or more saturated colors.

Figure 4-7. Example accent color scales, each with four base colors and three accent col- ors. Accent color scales can be derived in several different ways: (top) we can take an existing color scale (e.g., the Okabe Ito scale) and lighten and/or partially desaturate some colors while darkening others; (middle) we can take gray values and pair them with colors; (bottom) we can use an existing accent color scale (e.g., the one from the ColorBrewer project).

1.2.5 Representing - Amounts, Distribution, and Proportions

Commonly used plots and charts to visualize different types of data -

i. Amounts

The most common approach to visualizing amounts (i.e., numerical values shown for some set of categories) is using bars, either vertically or horizontally. However, instead of using bars, we can also place dots at the location where the corresponding bar would end.

If there are two or more sets of categories for which we want to show amounts, we can group or stack the bars. We can also map the categories onto the *x* and *y* axes and show amounts by color, via a heatmap.

ii. Distributions

Histograms and density plots provide the most intuitive visualizations of a distribution,

but both require arbitrary parameter choices and can be misleading. Cumulative densities and quantile-quantile (q-q) plots always represent the data faithfully but can be more difficult to interpret.

Boxplots, violin plots, strip charts, and since plots are useful when we want to visualize many distributions at once and/or if we are primarily interested in overall shifts among the distributions. Stacked histograms and overlapping densities allow a more in-depth com- parison of a smaller number of distributions, though stacked histograms can be difficult to interpret and are best avoided. Ridgeline plots can be a useful alternative to violin plots and are often useful when visualizing very large numbers of distributions or changes in distributions over.

iii. Proportions

Proportions can be visualized as pie charts, side-by-side bars, or stacked bars. As for amounts, when we visualize proportions with bars, the bars can be arranged either vertically or horizontally. Pie charts emphasize that the individual parts add up to a whole and highlight simple fractions. However, the individual pieces are more easily compared in side-by-side bars. Stacked bars look awkward for a single set of proportions, but can be useful when comparing multiple sets of proportions.

When visualizing multiple sets of proportions or changes in proportions across conditions, pie charts tend to be space-inefficient and often obscure relationships. Grouped bars work well as long as the number of conditions compared is moderate, and stacked bars can work for large numbers of conditions. Stacked densities are appropriate when the proportions change along a continuous variable.

When proportions are specified according to multiple grouping variables, mosaic plots, tree maps, or parallel sets are useful visualization approaches. Mosaic plots assume that every level of one grouping variable can be combined with every level of another grouping variable, whereas tree maps do not make such an assumption. Tree maps work well even if the subdivisions of one group are entirely distinct from the subdivisions of another. Parallel sets work better than either mosaic plots or tree maps when there are more than two grouping variables.

1.3 Data Storytelling

1.3.1 Introduction

Data storytelling is a methodology for communicating information, tailored to a specific audience, with a compelling narrative. It is the last ten feet of your data analysis and arguably the most important aspect. Data storytelling is the concept of building a compelling narrative based on complex data and analytics that help tell your story and influence and inform a particular audience.

- The benefits of data storytelling
- ✓ Adding value to your data and insights.
- ✓ Interpreting complex information and highlighting essential key points for the audience.
- ✓ Providing a human touch to your data.
- ✓ Offering value to your audience and industry.
- ✓ Building credibility as an industry and topic thought leader.

1.3.2. Ineffectiveness of Graphical representation of data

Data visualization plays a significant role in determining how receptive your audience is to receiving complex information. Data visualization helps transform boundless amounts of data into something simpler and digestible. Here, you can supply the visuals needed to support your story. Effective data visualizations can help:

- Reveal patterns, trends, and findings from an unbiased viewpoint.
- Provide context, interpret results, and articulate insights.
- Streamline data so your audience can process information.
- Improve audience engagement.

1.3.3. The three key elements of data storytelling

Through a structured approach, data storytelling and data visualization work together to communicate your insights through three essential elements: narrative, visuals, and

data. As you create your data story, it is important to combine the following three elements to write a well-rounded anecdote of your theory and the resulting actions you'd like to see from users.

1. Build your narrative

As you tell your story, you need to use your data as supporting pillars to your insights. Help your audience understand your point of view by distilling complex information into informative insights. Your narrative and context are what will drive the linear nature of your data storytelling.

2. Use visuals to enlighten

Visuals can help educate the audience on your theory. When you connect the visual assets (charts, graphs, etc.) to your narrative, you engage the audience with otherwise hidden insights that provide the fundamental data to support your theory. Instead of presenting a single data insight to support your theory, it helps to show multiple pieces of data, both granular and high level, so that the audience can truly appreciate your viewpoint.

3. Show data to support

Humans are not naturally attracted to analytics, especially analytics that lack contextualization using augmented analytics. Your narrative offers enlightenment, supported by tangible data. Context and critique are integral to the full interpretation of your narrative. Using business analytic tools to provide key insights and understanding to your narrative can help provide the much-needed context throughout your data story.

By combining the three elements above, your data story is sure to create an emotional response in your audience. Emotion plays a significant role in decision-making. And by linking the emotional context and hard data in your data storytelling, you're able to influence others. When these three key elements are successfully integrated, you have created a data story that can influence people and drive change.

1.3.4. Explanatory Analysis

Exploratory analysis is what you do to understand the data and figure out what might be noteworthy or interesting to highlight to others. When it comes to explanatory analysis, there are a few things to think about and be extremely clear on before visualizing any data or creating content. First, to whom are you communicating? It is important to have a good understanding of who your audience is and how they perceive you. This can help you to identify common ground that will help you ensure they hear your message. Second, what do you want your audience to know or do? You should be clear how you want your audience to act and take into account how you will communicate to them and the overall tone that you want to set for your communication. It's only after you can concisely answer these first two questions that you're ready to move forward with the third: How can you use data to help make your point?

1.3.4.1. <u>Who -</u>

- *Your audience* - The more specific you can be about who your audience is, the better position you will be in for successful communication. Avoid general audiences, such as "internal and external stakeholders" or "anyone who might be interested"—by trying to communicate to too many different people with disparate needs at once, you put yourself in a position where you can't communicate to any one of them as effectively as you could if you narrowed your target audience. Sometimes this means creating different communications for different audiences. Identifying the decision maker is one way of narrowing your audience. The more you know about your audience, the better positioned you'll be to understand how to resonate with them and form a communication that will meet their needs and yours.
- *You -* It's also helpful to think about the relationship that you have with your audience and how you expect that they will perceive you. Will you be encountering each other for the first time through this communication, or do you have an established relationship? Do they already trust you as an expert, or do you need to work to establish credibility? These are important considerations when it comes to determining how to structure your communication and whether and when to use data, and may impact the order and flow of the overall story you aim to tell.

1.3.4.2. What -

- *Action* - What do you need your audience to know or do? This is the point where you think through how to make what you communicate relevant for your audience and form a clear understanding of why they should care about what you say. You should always want your audience to know or do something. If you can't concisely articulate that, you should revisit whether you need to communicate in the first place.
- *Mechanism* - How will you communicate to your audience? The method you will use to communicate to your audience has implications on a number of factors, including the amount of control you will have over how the audience takes in the information and the level of detail that needs to be explicit. We can think of the communication mechanism along a continuum, with live presentation at the left and a written document or email at the right, as shown in Figure 1.1. Consider the level of control you have over how the information is consumed as well as the amount of detail needed at either end of the spectrum.

1.3.4.3. How -

Finally—and only after we can clearly articulate who our audience is and what we need them to know or do—we can turn to the data and ask the question: What data is available that will help make my point? Data becomes supporting evidence of the story you will build and tell.

1.4 Concept of machine learning and deep learning

1.4.1 Machine Learning:

- Machine learning is a branch of science that deals with programming the systems in such a way that they automatically learn and improve with experience. Here, learning

means recognizing and understanding the input data and making wise decisions based on the supplied data.

- It is very difficult to cater to all the decisions based on all possible inputs. To tackle this problem, algorithms are developed. These algorithms build knowledge from specific data and past experience with the principles of statistics, probability theory, logic, combinatorial optimization, search, reinforcement learning, and control theory.

The developed algorithms form the basis of various applications such as:

- Vision processing
- Language processing
- Forecasting (e.g., stock market trends)
- Pattern recognition
- Games
- Data mining
- Expert systems
- Robotics

Machine learning is a vast area and it is quite beyond the scope of this tutorial to cover all its features. There are several ways to implement machine learning techniques, however the most commonly used ones are **supervised** and **unsupervised learning**.

1.4.2. Supervised Learning: Supervised learning deals with learning a function from available training data. A supervised learning algorithm analyzes the training data and produces an inferred function, which can be used for mapping new examples. Common examples of supervised learning include:

- classifying e-mails as spam,
- labeling webpages based on their content, and
- voice recognition.

There are many supervised learning algorithms such as neural networks, Support Vector Machines (SVMs), and Naive Bayes classifiers. Mahout implements Naive Bayes classifier.

1.4.3. Unsupervised Learning: Unsupervised learning makes sense of unlabeled data without having any predefined dataset for its training. Unsupervised learning is an extremely powerful tool for analyzing available data and look for patterns and trends. It is most commonly used for clustering similar input into logical groups. Common approaches to unsupervised learning include:

- k-means
- self-organizing maps, and
- hierarchical clustering

1.4.4. Deep Learning

Deep learning is a subfield of machine learning where concerned algorithms are inspired by the structure and function of the brain called artificial neural networks.

All the value today of deep learning is through supervised learning or learning from labelled data and algorithms.

Each algorithm in deep learning goes through the same process. It includes a hierarchy of nonlinear transformation of input that can be used to generate a statistical model as output. Consider the following steps that define the Machine Learning process

- Identifies relevant data sets and prepares them for analysis.
- Chooses the type of algorithm to use
- Builds an analytical model based on the algorithm used.
- Trains the model on test data sets, revising it as needed.
- Runs the model to generate test scores.

Deep learning has evolved hand-in-hand with the digital era, which has brought about an explosion of data in all forms and from every region of the world. This data, known simply as big data, is drawn from sources like social media, internet search engines, e-commerce platforms, and online cinemas, among others. This enormous amount of data is readily accessible and can be shared through fintech applications like cloud computing.

However, the data, which normally is unstructured, is so vast that it could take decades for humans to comprehend it and extract relevant information. Companies realize the incredible potential that can result from unraveling this wealth of information and are increasingly adapting to AI systems for automated support.

1.4.5. Applications of Machine Learning and Deep Learning

- Computer vision which is used for facial recognition and attendance mark through fingerprints or vehicle identification through number plate.
- Information Retrieval from search engines like text search for image search.
- Automated email marketing with specified target identification.
- Medical diagnosis of cancer tumors or anomaly identification of any chronic disease.
- Natural language processing for applications like photo tagging. The best example to explain this scenario is used in Facebook.
- Online Advertising.

References:

- https://www.tutorialspoint.com/artificial_intelligence/artificial_intelligence _overview. htm
- https://www.javatpoint.com/introduction-to-artificial-intelligence
- https://www.tutorialspoint.com/tensorflow/tensorflow_machine_learning_d eep_learni ng.htm
- Story telling with data by Cole Nissbuamer Knafilc – Wiley Publication - ISBN 9781119002253
- Fundamentals of Data Visualization, A primer on making informative and compelling figures by Claus O Wilke - O'Reilly Publication – March 2019

Sample Multiple Choice Questions

1. _____is a branch of Science which deals with helping machines find solutions to complex problems in a more human-like fashion
   a. Artificial Intelligence
   b. Internet of Things
   c. Embedded System
   d. Cyber Security
2. In_____the goal is for the software to use what it has learned in one area to solve problems in other areas.
   a. Machine learning
   b. Deep learning
   c. Neural networks
   d. None of these
3. Computer programs that mimic the way the human brain processes information is called as
   a. Machine learning
   b. Deep learning
   c. Neural networks
   d. None of these
4. The core components and constituents of AI are derived from
   a. concept of logic
   b. cognition
   c. computation
   d. All of above
5. _____ is the graphical representation of information and data.
   a. Data visualization
   b. Data Analytics
   c. Data mapping
   d. Data storytelling
6. ____ type of variables is used to represent whole integers
   a. Numerical continuous
   b. Numerical discrete
   c. Categorical ordered
   d. Numerical integers
7. Sequential color scale is used when _____.
   a. colors are used to distinguish discrete items.
   b. colors are used to represent data values.
   c. colors are used to highlight.
   d. colors are used to represent descriptive data
8. In data storytelling, internal and external stakeholders are ____ .
   a. Targeted audience

b. General audience
c. Specific audience
d. Data specific audience

9. _____ is a methodology for communicating information, tailored to a specific audience, with a compelling narrative.
   a. Data visualization
   b. Data storytelling
   c. Data mining
   d. Data Analysis
10. Chomsky's linguistic computational theory generated a model for syntactic analysis through
   a. regular grammar
   b. regular expression
   c. regular word
   d. none of these`
11. These machines only focus on current scenarios and react on it as per possible best action
   a. Reactive Machines
   b. Limited Memory
   c. Theory of Mind
   d. Self-Awareness

Unit-2 Internet of Things

Content

2.1 Internet of Things (IoT)
- Definition
- Characteristics of IoT
- Features of IoT
- Advantages and Disadvantages of IoT

2.1.2 Design of IoT
- Physical design of IoT
- Logical design of IoT

2.1.3 IoT Protocols

2.1.4 Sensors and actuators used in IoT

2.2 Introduction to 5G Network
- 5-G characteristics and application areas.
- NGN architecture: Features, Functional block diagram, Network components: Media Gateway, Media Gateway Controller, and Application Server.
- NGN Wireless Technology: Telecom network Spectrum: Types [licensed and unlicensed], Mobile Network Evolution (2G to 5G), Comparative features,
- NGN Core: Features, Multi-Protocol Label Switching (MPLS): Concepts, Features and Advantages.

.

2.1 Internet of Things (IoT)

IoT Definition:
- The internet of things (IoT) is a computing concept that describes the idea of everyday physical objects being connected to the internet and being able to identify themselves to other devices.
- Internet of Things (IoT) refers to physical and virtual objects that have unique identities and are connected to the internet to facilitate intelligent applications that make energy, logistics, industrial control, retail, agriculture and many other domains "smarter".
- Internet of things (IoT) is a new revolution in which endpoints connected to the internet and driven by the advancements in sensor networks, mobile devices, wireless communications, networking and cloud technologies.

Characteristics of IoT:
- **Dynamic & Self-Adapting:**IoT devices and systems may have the capability to dynamically adapt with the changing contexts and take actions based on their operating conditions, user's context, or sensed environment. For example, the surveillance cameras can adapt their modes (to normal or infra-red night modes) based on whether it is day or night.
- **Self-Configuring:**IoT devices may have self-configuring capability, allowing a large number of devices to work together to provide certain functionality (such as weather monitoring).

- **Interoperable Communication Protocols:**IoT devices may support a number of interoperable communication protocols and can communicate with other devices and also with the infrastructure.
- **Unique Identity:** Each IoT device has a unique identity and a unique identifier (such as an IP address or a URI). IoT device interfaces allow users to query the devices, monitor their status, and control them remotely, in association with the control, configuration and managementinfrastructure.
- **Integrated into Information Network:** IoT devices are usually integrated into the information network that allows them to communicate and exchange data with other devices and systems.
- **Enormous scale:** The number of devices that need to be managed and that communicate with each other will be at least an order of magnitude larger than the devices connected to the current Internet.

Features of IoT:

- **Connectivity:** Connectivity refers to establish a proper connection between all the things of IoT to IoT platform it may be server or cloud.
- **Analyzing:** After connecting all the relevant things, it comes to real-time analyzing the data collected and use them to build effective business intelligence.
- **Integrating:** IoT integrating the various models to improve the user experience as well.
- **Artificial Intelligence:** IoT makes things smart and enhances life through the use of data.
- **Sensing:** The sensor devices used in IoT technologies detect and measure any change in the environment and report on their status.
- **Active Engagement:** IoT makes the connected technology, product, or services to active engagement between each other.
- **Endpoint Management:** It is important to be the endpoint management of all the IoT system otherwise; it makes the complete failure of the system.

Applications of IoTs

a. Home Automation:

- **Smart Lighting:** helps in saving energy by adapting the lighting to the ambient conditions and switching on/off or diming the light when needed.
- **Smart Appliances:** make the management easier and also provide status information to the users remotely.
- **Intrusion Detection:** use security cameras and sensors (PIR sensors and door sensors) to detect intrusion and raise alerts. Alerts can be in the form of SMS or email sent to the user.
- **Smoke/Gas Detectors:** Smoke detectors are installed in homes and buildings to detect smoke that is typically an early sign of fire. Alerts raised by smoke detectors can be in the form of signals to a fire alarm system. Gas detectors can detect the presence of harmful gases such as CO, LPG etc.,

b. Cities:

- **Smart Parking:** make the search for parking space easier and convenient for drivers. Smart parking are powered by IoT systems that detect the no. of empty parking slots and send information over internet to smart application back ends.
- **Smart Lighting:** for roads, parks and buildings can help in saving energy.
- **Smart Roads:** Equipped with sensors can provide information on driving condition, travel time estimating and alert in case of poor driving conditions, traffic condition and accidents.
- **Structural Health Monitoring:** uses a network of sensors to monitor the vibration levels in the structures such as bridges and buildings.
- **Surveillance:** The video feeds from surveillance cameras can be aggregated in cloud based scalable storage solution.
- **Emergency Response:**IoT systems for fire detection, gas and water leakage detection can help in generating alerts and minimizing their effects on the critical infrastructures.

c. Environment:

- **Weather Monitoring:** Systems collect data from a no. of sensors attached and send the data to cloud based applications and storage back ends. The data collected in cloud can then be analyzed and visualized by cloud based applications.
- **Air Pollution Monitoring:** System can monitor emission of harmful gases (CO2, CO, NO, NO2 etc.,) by factories and automobiles using gaseous and meteorological sensors. The collected data can be analyzed to make informed decisions on pollutions control approaches.
- **Noise Pollution Monitoring:** Due to growing urban development, noise levels in cities have increased and even become alarmingly high in some cities. IoT based noise pollution monitoring systems use a no. of noise monitoring systems that are deployed at different places in a city. The data on noise levels from the station is collected on servers or in the cloud. The collected data is then aggregated to generate noise maps.
- **Forest Fire Detection:** Forest fire can cause damage to natural resources, property and human life. Early detection of forest fire can help in minimizing damage.
- **River Flood Detection:** River floods can cause damage to natural and human resources and human life. Early warnings of floods can be given by monitoring the water level and flow rate. IoT based river flood monitoring system uses a no. of sensor nodes that monitor the water level and flow rate sensors.

d. Retail:

- **Inventory Management:** IoT systems enable remote monitoring of inventory using data collected by RFID readers.
- **Smart Payments:** Solutions such as contact-less payments powered by technologies such as Near Field Communication(NFC) and Bluetooth.

- **Smart Vending Machines:** Sensors in a smart vending machines monitors its operations and send the data to cloud which can be used for predictive maintenance.

e. Logistics:

- **Route generation & scheduling:**IoT based system backed by cloud can provide first response to the route generation queries and can be scaled upto serve a large transportation network.
- **Fleet Tracking:** Use GPS to track locations of vehicles in real-time.
- **Shipment Monitoring:** IoT based shipment monitoring systems use sensors such as temp, humidity, to monitor the conditions and send data to cloud, where it can be analyzed to detect food spoilage.
- **Remote Vehicle Diagnostics:** Systems use on-board IoT devices for collecting data on Vehicle operation's (speed, RPMetc.,) and status of various vehicle sub systems.

f. Agriculture:

- **Smart Irrigation:** to determine moisture amount in soil.
- **Green House Control:** to improve productivity.

g. Industry:

- Machine diagnosis and prognosis
- Indoor Air Quality Monitoring

h. Health and Life Style:

- Health & Fitness Monitoring
- Wearable Electronics

Advantages and Disadvantages of IoT:

a. Advantages of IoT

- **Efficient resource utilization:** If we know the functionality and the way that how each device work we definitely increase the efficient resource utilization as well as monitor natural resources.
- **Minimize human effort:** As the devices of IoT interact and communicate with each other and do lot of task for us, then they minimize the human effort.
- **Save time:** As it reduces the human effort then it definitely saves out time. Time is the primary factor which can save through IoT platform.
- **Improve security:** Now, if we have a system that all these things are interconnected then we can make the system more secure and efficient.
- **Reduced Waste**: IoT makes areas of improvement clear. Current analytics give us superficial insight, but IoT provides real-world information leading to more effective management of resources.
- **Enhanced Data Collection**: Modern data collection suffers from its limitations and its design for passive use. IoT breaks it out of those spaces, and places it

exactly where humans really want to go to analyze our world. It allows an accurate picture of everything.

b. Disadvantages of IoT

- **Security:** As the IoT systems are interconnected and communicate over networks. The system offers little control despite any security measures, and it can be lead the various kinds of network attacks.
- **Privacy**: Even without the active participation on the user, the IoT system provides substantial personal data in maximum detail.
- **Complexity:** The designing, developing, and maintaining and enabling the large technology to IoT system is quite complicated.
- **Flexibility:** Many are concerned about the flexibility of an IoT system to integrate easily with another. They worry about finding themselves with several conflicting or locked systems.
- **Compliance**: IoT, like any other technology in the realm of business, must comply with regulations. Its complexity makes the issue of compliance seem incredibly challenging when many consider standard software compliance a battle.

2.1.2 Design of IoT

a. Physical design of IoT:

- The "Things" in IoT usually refers to IoT devices which have unique identities and can perform remote sensing, actuating and monitoring capabilities.
- IoT devices can exchange data with other connected devices and applications (directly or indirectly), or collect data from other devices and process the data either locally or send the data to centralized servers or cloud-based application back-ends for processing the data, or perform some tasks locally and other tasks within the IoT infrastructure, based on temporal and space constraints (i.e., memory, processing capabilities, communication latencies and speeds, and deadlines).

Fig 2.1 Generic Bock Diagram of an IoT Device

- An IoT device may consist of several interfaces for connections to other devices, both wired and wireless. These include (i) I/O interfaces for sensors, (ii) interfaces for Internet connectivity, (iii) memory and storage interfaces and (iv) audio/video interfaces as shown in Fig.2.1
- An IoT device can collect various types of data from the on-board or attached sensors, such as temperature, humidity, light intensity. The sensed data can be communicated either to other devices or cloud-based servers/storage.
- IoT devices can be connected to actuators that allow them to interact with other physical entities (including non-IoT devices and systems) in the vicinity of the device. For example, a relay switch connected to an IoT device can turn an appliance on/off based on the commands sent to the IoT device over the Internet.
- IoT devices can also be of varied types, for instance, wearable sensors, smart watches, LED lights, automobiles and industrial machines.
- Almost all IoT devices generate data in some form or the other which when processed by data analytics systems leads to useful information to guide further actions locally or remotely.
- For instance, sensor data generated by a soil moisture monitoring device in a garden, when processed can help in determining the optimum watering schedules.

b. Logical design of IoT:

Logical design of an IoT system refers to an abstract representation of the entities and processes without going into the low-level specifics of the implementation.

IoT functional blocks:

An IoT system comprises of a number of functional blocks that provide the system the capabilities for identification, sensing, actuation, communication, and management as shown in Figure 2.2.

Fig. 2.2 Fundamental block of IoT

- **Device:** An IoT system comprises of devices that provide sensing, actuation, monitoring and control functions.

- **Communication:** The communication block handles the communication for the IoT system.
- **Services:** An IoT system uses various types of IoT services such as services for device monitoring, device control services, data publishing services and services for device discovery.
- **Management:** Management functional block provides various functions to govern the IoT system.
- **Security:** Security functional block secures the IoT system and by providing functions such as authentication, authorization, message and content integrity, and data security.
- **Application:** IoT applications provide an interface that the users can use to control and monitor various aspects of the IoT system. Applications also allow users to view the system status and view or analyze the processed data.

2.1.3 IoT Protocols

- Following Figure 2.3 shows different IoT protocols.

Fig. 2.3 IoT Protocols

a. Link Layer Protocols:

- Link layer protocols determine how the data is physically sent over the network's physical layer or medium e.g., copper wire, coaxial cable, or a radio wave.
- Link layer determines how the packets are coded and signaled by the hardware device over the medium to which the host is attached (such as a coaxial cable).
  1. **802.3-Ethernet:** IEEE 802.3 is a collection of wired Ethernet standards for the link layer. For example, 802.3 is the standard for 10BASE5 Ethernet that uses coaxial cable as a shared medium, 802.3.i is the standard for 10BASE-T Ethernet over copper twisted-pair connections, 802.3.j is the standard for 10BASE-F Ethernet over fiber optic connections, 802.3ae is the standard for 10 Gbit/s Ethernet over fiber, and so on.
  2. **802.11- WiFi:** IEEE 802.11 is a collection of wireless local area network (WLAN) communication standards, including extensive description of the link layer. 802.11a operates in the 5 GHz band, 802.11b and 802.11g operate in the

2.4 GHz band, 802.11n operates in the 2.4/5 GHz bands, 802.11ac operates in the 5 GHz band and 802.11ad operates in the 60 GHz band. These standards provide data rates from 1 Mb/s to upto 6.75 Gb/s.

3. **802.16-WiMax:** IEEE 802.16 is a collection of wireless broadband standards, including extensive descriptions for the link layer (also called WiMax). WiMaxstandards provide data rates from 1.5 Mb/s to 1 Gb/s. The recent update (802.16m) provides data rates of 100 Mbit/s for mobile stations and 1 Gbit/s for fixed stations.
4. **802.15.4-LR-WPAN:** IEEE 802.15.4 is a collection of standards for low-rate wireless personal area networks (LR-WPANs). These standards form the basis of specifications for high level communication protocols such as ZigBee. LR-WPAN standards provide data rates from 40 Kb/s 250 Kb/s. These standards provide low-cost and low-speed communication for power constrained devices.
5. **2G/3G/4G - Mobile Communication:** There are different generations of mobilecommunication standards including second generation (2G including GSM and CDMA), third generation (3G - including UMTS and CDMA2000) and fourth generation (4G - including LTE). IoT devices based on these standards can communicate over cellular networks. Data rates for these standards range from 9.6 Kb/s (for 2G) to upto 100 Mb/s (for 4G) and are available from the 3GPP websites.

b. Network/Internet Layer Protocols:

The network layers are responsible for sending of IP datagrams from the source network to the destination network. This layer performs the host addressing and packet routing. The datagrams contain the source and destination addresses which are used to route them from the source to destination across multiple networks. Host identification is done using hierarchical IP addressing schemes such as IPv4 or IPv6.

1. **IPv4:** Internet Protocol version 4 (IPv4) is the most deployed Internet protocol that is used to identify the devices on a network using a hierarchical addressing scheme. IPv4 uses a 32-bit address scheme that allows total of 232 or 4,294,967,296 addresses. IPv4 has been succeeded by IPv6. The IP protocols establish connections on packet networks, but do not guarantee delivery of packets. Guaranteed delivery and data integrity are handled by the upper layer protocols (such as TCP).
2. **IPv6:** Internet Protocol version 6 (IPv6) is the newest version of Internet protocol and successor to IPv4, IPv6 uses 128-bit address scheme that allows total of 2128 or 3.4 x 1038 addresses.
3. **6LOWPAN:** 6LOWPAN (IPv6 over Low Power Wireless Personal Area Networks) brings IP protocol to the low-power devices which have limited processing capability. 6LOWPAN operates in the 2.4 GHz frequency range and provides data transfer rates of 250 Kb/s. 6LOWPAN works with the 802.15.4 link layer protocol and defines compression mechanisms for IPv6 datagrams over IEEE 802.15.4-based networks.

c. Transport Layer Protocols:

The Transport layer protocols provide end-to-end message transfer capability independent of the underlying network. The message transfer capability can be set up on connections, either using handshakes (as in TCP) or without handshakes/acknowledgements (as in UDP). The transport layer provides functions such as error control, segmentation, flow control and congestion control.

1. **TCP:** Transmission Control Protocol (TCP) is the most widely used transport layer protocol, that is used by web browsers (along with HTTP, HTTPS application layer protocols), email programs (SMTP application layer protocol) and file transfer (FTP). TCP is a connection oriented and stateful protocol. TCP ensures reliable transmission of packets in-order and also provides error detection capability so that duplicate packets can be discarded and lost packets are retransmitted.
2. **UDP:** User Datagram Protocol (UDP) is a connectionless protocol. UDP is useful for time-sensitive applications that have very small data units to exchange and do not want the overhead of connection setup. UDP is a transaction oriented and stateless protocol. UDP does not provide guaranteed delivery, ordering of messages and duplicate elimination. Higher levels of protocols can ensure reliable delivery or ensuring connections created are reliable.

d. Application Layer Protocols:

Application layer protocols define how the applications interface with the lower layer protocols to send the data over the network. The application data, typically in files, is encoded by the application layer protocol and encapsulated in the transport layer protocol which provides connection or transaction oriented communication over the network. Port numbers are used for application addressing (for example port 80 for HTTP, port 22 for SSH, etc.). Application layer protocols enable process-to-process connections using ports.

1. **HTTP:** Hypertext Transfer Protocol (HTTP) is the application layer protocol that forms the foundation of the World Wide Web (WWW). HTTP includes commands such as GET, PUT, POST, DELETE, HEAD, TRACE, OPTIONS, etc. The protocol follows a request-response model where a client sends requests to a server using the HTTP commands. HTTP is a stateless protocol and each HTTP request is independent of the other requests. An HTTP client can be a browser or an application running on the client (e.g., an application running on an IoT device, a mobile application or other software). HTTP protocol uses Universal Resource Identifiers (URIs) to identify HTTP resources.
2. **CoAP:** Constrained Application Protocol (CoAP) is an application layer protocol for machine-to-machine (M2M) applications, meant for constrained environments with constrained devices and constrained networks. Like HTTP, COAP is a web transfer protocol and uses a request-response model, however it

runs on top of UDP instead of TCP. COAP uses a client-server architecture where clients communicate with servers using connectionless datagrams. COAP is designed to easily interface with HTTP. Like HTTP, COAP supports methods such as GET, PUT, POST, and DELETE. COAP draft specifications are available on IEFT Constrained environments (CORE) Working Group website.

3. **WebSocket:** WebSocket protocol allows full-duplex communication over a single socket connection for sending messages between client and server. WebSocket is based on TCP and allows streams of messages to be sent back and forth between the client and server while keeping the TCP connection open. The client can be a browser, a mobile application or an IoT device.
4. **MQTT:** Message Queue Telemetry Transport (MQTT) is a light-weight messaging protocol based on the publish-subscribe model. MQTT uses a client-server architecture where the client (such as an IoT device) connects to the server (also called MQTT Broker) and publishes messages to topics on the server. The broker forwards the messages to the clients subscribed to topics. MQTT is well suited for constrained environments where the devices have limited processing and memory resources and the network bandwidth is low.
5. **XMPP:** Extensible Messaging and Presence Protocol (XMPP) is a protocol for real-time communication and streaming XML data between network entities. XMPP powers wide range of applications including messaging, presence, data syndication, gaming, multi-party chat and voice/video calls. XMPP allows sending small chunks of XML data from one network entity to another in near real-time. XMPP is a decentralized protocol and uses a client-server architecture. XMPP supports both client-to-server and server-to-server communication paths. In the context of IoT, XMPP allows real-time communication between IoT devices.
6. **DDS:** Data Distribution Service (DDS) is a data-centric middleware standard for device-to-device or machine-to-machine communication. DDS uses a publish-subscribe model where publishers (e.g. devices that generate data) create topics to which subscribers (e.g., devices that want to consume data) can subscribe. Publisher is an object responsible for data distribution and the subscriber is responsible for receiving published data. DDS provides quality-of-service (QoS) control and configurable reliability.
7. **AMQP:** Advanced Message Queuing Protocol (AMQP) is an open application layer protocol for business messaging. AMQP supports both point-to-point and publisher/subscriber models, routing and queuing. AMQP brokers receive messages from publishers (e.g., devices or applications that generate data) and route them over connections to consumers (applications that process data). Publishers publish the messages to exchanges which then distribute message copies to queues. Messages are either delivered by the broker to the consumers which have subscribed to the queues or the consumers can pull the messages from the queues.

2.1.4 Sensors and actuators used in IoT

- **Sensors:** A **sensor** is an electronic instrument that is able to measure the physical quantity and generate a considerate output. These output of the sensors are usually in the form of electrical signals. Sensors are placed as such they can directly interact with the environment to sense the input energy with the help of sensing element. This sensed energy is converted into a more suitable form by a transduction element. There are various types of sensors such as position, temperature, pressure, speed sensors, but fundamentally there are two types – analog and digital. The different types come under these two basic types. A digital sensor is incorporated with an Analog-to-digital converter while analog sensor does not have any ADC.
- **Actuators:** An **actuator** is a device that alters the physical quantity as it can cause a mechanical component to move after getting some input from the sensor. In other words, it receives control input (generally in the form of the electrical signal) and generates a change in the physical system through producing force, heat, motion, etcetera. An actuator can be interpreted with the example of the stepper motor, where an electrical pulse drives the motor. Each time a pulse given in the input accordingly motor rotates in a predefined amount. A stepper motor is suitable for the applications where the position of the object has to be controlled precisely, for example, robotic arm.

a. Sensors used in IoT

1. **Temperature sensors:** These devices measure the amount of heat energy generated from an object or surrounding area. They find application in air-conditioners, refrigerators and similar devices used for environmental control. They are also used in manufacturing processes, agriculture and health industry.Temperature sensors include thermocouples, thermistors, resistor temperature detectors (RTDs) and integrated circuits (ICs) as shown in Figure 2.4

Fig. 2.4 Temperature Sensors

2. **Humidity sensors:** The amount of water vapour in air, or humidity, can affect human comfort as well as many manufacturing processes in industries. So monitoring humidity level is important. Most commonly used units for humidity measurement are relative humidity (RH), dew/frost point (D/F PT) and parts per million (PPM) as shown in Figure 2.5.

Fig. 2.5 Humidity Sensor

3. **Motion sensors:** Motion sensors are not only used for security purposes but also in automatic door controls, automatic parking systems, automated sinks, automated toilet flushers, hand dryers, energy management systems, etc. You use these sensors in the IoT and monitor them from your smartphone or computer. HC-SR501 passive infrared (PIR) sensor is a popular motion sensor for hobby projects as shown in Figure 2.6.

Fig. 2.6 Motion Sensor

4. **Gas sensors:** These sensors are used to detect toxic gases shown in Fig.2.7. The sensing technologies most commonly used are electrochemical, photo-ionisation and semiconductor. With technical advancements and new specifications, there are a multitude of gas sensors available to help extend the wired and wireless connectivity deployed in IoT applications.

Fig. 2.7 Gas Sensor

5. **Smoke sensors:** Smoke detectors have been in use in homes and industries for quite a long time as shown in Fig 2.8. With the advent of the IoT, their application has become more convenient and user-friendly. Furthermore, adding a wireless connection to smoke detectors enables additional features that increase safety and convenience.

Fig. 2.8 Smoke Sensor

6. **Pressure sensors:** These sensors are used in IoT systems to monitor systems and devices that are driven by pressure signals. When the pressure range is beyond the threshold level, the device alerts the user about the problems that should be fixed. For example, BMP180 shown in Fig. 2.9 is a popular digital pressure sensor for use in mobile phones, PDAs, GPS navigation devices and outdoor equipment. Pressure sensors are also used in smart vehicles and aircrafts to determine force and altitude, respectively. In vehicle, tyre pressure monitoring system (TPMS) is used to alert the driver when tyre pressure is too low and could create unsafe driving conditions.

Fig. 2.9: Pressure Sensor

7. **Image sensors:** These sensors are found in digital cameras, medical imaging systems, night-vision equipment, thermal imaging devices, radars, sonars, media house and biometric systems as shown in Fig. 2.10. In the retail industry, these sensors are used to monitor customers visiting the store through IoT network. In offices and corporate buildings, they are used to monitor employees and various activities through IoT networks.

Fig. 2.10 Image Sensor

8. **Accelerometer sensors:** These sensors are used in smartphones, vehicles, aircrafts and other applications to detect orientation of an object, shake, tap, tilt, motion, positioning, shock or vibration as shown in Fig.2.11. Different types of accelerometers include Hall-effect accelerometers, capacitive accelerometers and piezoelectric accelerometers.

Fig. 2.11 Accelerator Sensors

9. **IR sensors:** These sensors can measure the heat emitted by objects as shown in Fig. 2.12. They are used in various IoT projects including healthcare to monitor blood flow and blood pressure, smartphones to use as remote control and other functions, wearable devices to detect amount of light, thermometers to monitor temperature and blind-spot detection in vehicles.

Fig. 2.12 IR Sensor

10. **Proximity sensors:** These sensors detect the presence or absence of a nearby object without any physical contact. Different types of proximity sensors are inductive, capacitive, photoelectric, ultrasonic and magnetic as shown in Fig.2.13. These are mostly used in object counters, process monitoring and control.

IR Proximity Sensor | Inductive proximity sensor | Capacitive Sensor | Reed Switch

Fig.2.13: Proximity Sensors

b. Actuators used in IoT

1. **Servo motors:** A Servo is a small device that incorporates a two wire DC motor as shown in Fig.2.14, a gear train, a potentiometer, an integrated circuit, and a shaft (output spine). The shaft can be positioned to specific angular positions by sending the servo a coded signal. Of the three wires that stick out from the servo casing, one is for power, one is for ground, and one is a control input line. It uses the position-sensing device to determine the rotational position of the shaft, so it knows which way the motor must turn to move the shaft to the commanded position.

Fig. 2.14 Servo Motor

2. **Stepper Motor:** Stepper motors as shown in Fig.2.15 are DC motors that move in discrete steps. They have multiple coils that are organized in groups called "phases". By energizing each phase in sequence, the motor will rotate, one step at a time. With a computer controlled stepping, you can achieve very precise positioning and/or speed control.

Fig. 2.15 Stepper Motor

3. **DC motors:** Direct Current (DC) motor as shown in Fig.2.16 is the most common actuator used in electronics projects. They are simple, cheap, and easy to use. DC motors convert electrical into mechanical energy. They consist of permanent magnets and loops of wire inside. When current is applied, the wire loops generate a magnetic field, which reacts against the outside field of the static magnets.

Fig. 2.16 DC Motor

4. **Linear Actuator:** A linear actuator as shown in Fig. 2.17 is an actuator that creates motion in a straight line, in contrast to the circular motion of a conventional electric motor. Linear actuators are used in machine tools and industrial machinery, in computer peripherals such as disk drives and printers, in valves and dampers, and in many other places where linear motion is required.

Fig. 2.17 Linear Actuator

5. **Relay:** A relay is an electrically operated switch as shown in Fig.2.18. Many relays use an electromagnet to mechanically operate a switch, but other operating principles are also used, such as solid-state relays. The advantage of relays is that it takes a relatively small amount of power to operate the relay coil, but the relay itself can be used to control motors, heaters, lamps or AC circuits which themselves can draw a lot more electrical power.

Fig. 2.18 Relay

6. **Solenoid:** A solenoid as shown in Fig.2.19 is simply a specially designed electromagnet. Solenoids are inexpensive, and their use is primarily limited to on-off applications such as latching, locking, and triggering. They are frequently used in home appliances (e.g. washing machine valves), office equipment (e.g. copy machines), automobiles (e.g. door latches and the starter solenoid), pinball machines (e.g., plungers and bumpers), and factory automation.

Fig. 2.19 Solenoid

2.2 Introduction to 5G Network

Mobile networks having a history of 40-year history which is parallels with the Internet's, have undergone significant change. The first two generations i.e. 1G and 2G supports voice and text. Then 3G defining the transition to broadband access, supporting data rates with hundreds of kilobits-per-second. Now a day, the industry is at 4G supporting data rates with the few megabits-per-second and now transitioning to 5G, with the increase in data rates.

5G is the 5th generation mobile network which is a new global wireless standard after 1G, 2G, 3G, and 4G networks. 5G is a new kind of network that is designed to connect virtually everyone and everything together including machines, objects, and devices.

5G wireless technology is used to deliver higher multi-Gbps peak data speeds, ultra-low latency, more reliability, massive network capacity, increased availability, and a more uniform user experience to more users. Higher performance and improved efficiency empower new user experiences and connects new industries.

a. Characteristics of 5G

- **High Data Transfer speed** is a comprehensive upgrade of upload and download speed and the download speed is theoretically more than 10 times higher than that of 4G. A data transfer rate of up to 10Gbps which is 10 to 100 times better than the 4G networks
- **Low latency** is a great improvement in the feedback speed of the network. Low latency of 1 millisecond will be used in remote control, games and the Internet of Things.
- **Massive devices** are the Internet of Things. Compared with traditional 4G networks, 5G networks can accommodate more devices.
- **Broadband** 1000 times faster per unit area
- **Availability** of 99.999%
- Reduction of 90% in the **energy** consumption of the network
- Up to 10 years of battery life on low-power IoT (Internet of Things) devices

b. Application areas of 5G Network

1. High-speed mobile network:

- Can support up to 10 to 20 GBPS of data download speed which is equivalent to a fiber-optic and internet connection accessed wirelessly.
- Voice and high-speed data can be simultaneously transferred efficiently in 5G.
- Due to low latency of 5G technology, it is significant for autonomous driving and mission-critical applications.
- The small cell concept used in 5G will have multiple advantages of better cell coverage, maximum data transfer, low power consumption, cloud access network, etc.…

2. Entertainment and multimedia

- High-speed streaming of 4K videos and crystal clear audio clarity.
- Live events streaming via a wireless network with high definition.
- HD TV channels can be accessed on mobile devices without any interruptions.
- Augmented reality and virtual reality with HD video with low latency.
- HD virtual reality games.

3. Internet of Things that Connecting everything

- Internet of Things will connect every object, appliance, sensor, device, and application to the Internet.
- IoT applications will collect huge amounts of data from millions of devices and sensors. It requires an efficient network for data collection, processing, transmission, control, and real-time analytics.
- 5G is the most efficient technology for the Internet of Things due to its flexibility, unused spectrum availability, and low-cost solutions for deployment.
- IoT can benefit from 5G networks in many areas like:

i. Smart Home

- 5G wireless network will be utilized by smart appliances which can be configured and accessed from remote locations, closed-circuit cameras will provide high-quality real-time video for security purposes.

ii. Logistics and shipping

- The logistics and shipping industry can be use smart 5G technology for goods tracking, fleet management, centralized database management, staff scheduling, and real-time delivery tracking and reporting.

(i) Efficient use of RFID tags
(ii) Accelerated packing and labeling
(iii) Use of smart tracking devices for accurate monitoring of temperature, shock, light exposure, humidity, etc…
(iv) Realtime GPS location tracking and reporting
(v) Efficient monitoring minimizes theft risk and misplacing of items
(vi) Realtime delivery tracking and reporting
(vii) Self-driving cars and drones for future goods delivery

iii. Smart cities

- Smart city applications like traffic management, Instant weather update, local area broadcasting, energy management, smart power grid, smart lighting of the

street, water resource management, crowd management, emergency response, etc. can use a reliable 5G wireless network for its functioning.

iv. Industrial IoT

- Future industries can be used 5G technology for efficient automation of equipment, predictive maintenance, safety, process tracking, smart packing, shipping, logistics, and energy management.
- Smart sensor technology compromises unlimited solutions for industrial IoT for smarter, safe, cost-effective, and energy-efficient industrial operations.

v. Smart farming

- 5G technology can be used for agriculture and smart farming in the future. Using smart RFID sensors and GPS technology, farmers can track the location of livestock and manage them easily. Smart sensors can be used for irrigation control, access control, and energy management.

vi. Fleet management

- Many companies are using smart tracking devices for fleet management, 5G technology will provide much better solutions for location tracking and fleet management.

vii. Healthcare and mission-critical applications

5G technology can support

- Medical practitioners to perform advanced medical procedures with a reliable wireless network connected to another side of the globe.
- Connected classrooms will help students to attend seminars and important lecturers.
- People with chronic medical conditions will benefit from smart devices and real-time monitoring.
- Doctors can connect with patients from anywhere anytime and advise them when necessary.
- Scientists are working on smart medical devices which can perform remote surgery.
- Smart medical devices like wearables will continuously monitor a patient's condition and activate alerts during emergencies using which hospitals and ambulance services will get alerts during critical situations and they can take necessary steps to speed up diagnosis and treatment.
- Patients with special needs can be tracked using special tags and precise location tracking devices.
- Healthcare databases can be accessed from any location collected data analysis can be used for research and improvement of treatments.
- Medical practitioners will be able to share large files like an MRI report which is often larger than 1GB size within seconds using a 5G network.

viii. Autonomous Driving

- In the future, cars can communicate with smart traffic signs, surrounding objects, and other vehicles on the road where every millisecond is important for self-driving vehicles, decision has to be made in split second to avoid collision and make sure passenger safety.

ix. Drone Operation

- Drones can be used for multiple operations ranging from entertainment, video capturing, medical and emergency access, smart delivery solutions, security, surveillance, etc using 5G network which provide strong support with high-speed wireless internet connectivity for drone operation in a wide range of applications.
- During emergencies like natural calamities, humans have limited access to many areas where drones can reach out and collect useful information.

x. Security and surveillance

- 5G wireless technology can be used for security and surveillance due to its higher bandwidth and unlicensed spectrum.

4. Satellite Internet

- Satellite internet technology with High-speed 5G network connectivity offers connectivity in urban and rural areas across the globe with the help of a constellation of thousands of small satellites.

c. Introduction to NGN (Next Generation Network)

Next Generation Network (NGN) is a new concept and becoming more and more important for future telecommunication networks. Next Generation Network (NGN) is a packet-based network able to provide telecommunication services and able to make use of multiple broadband, Quality of service (QoS)-enabled transport technologies. It supports mobility. In NGN service related functions are independent from underlying transport-related technologies. It enables unfettered access for users to networks and to competing service providers. It supports generalized mobility which will allow consistent and ubiquitous provision of services to users.

The general idea behind the NGN is that one network transports all information and services (voice, data, and all sorts of media such as video) by encapsulating these into IP packets, similar to those used on the Internet NGNs are commonly built around the Internet, and therefore the term all IP is also sometimes used to describe the transformation of formerly telephone-centric networks toward NGN.

d. NGN Features:

1) NGN Support for a wide range of converged services between fixed/mobile and ability to deliver a wide variety of services including voice, video, audio and visual data, via session and interactive based services in unicast, multicast and broadcast modes.
2) NGN provides End-to-end QoS (Quality of Service): The NGN aims to provide high quality broadband communication by controlling the quality of service (QoS) on end to end basis.

   The NGN support a wide range of QoS enabled services as it define:

   a) Bearer service QoS classes;
   b) QoS control mechanisms;
   c) QoS control functional architecture;
   d) QoS control/signaling.

3) NGN provides *e*nd-to-end Packet-based Transfer (machine-to-machine, human-to-human, human-to-machine) with Broadband capabilities approx. > 2Mbps
4) Technologies used at access and transport are broadband in nature. 3G WCDMA, 4 G LTE Advance is used at wireless access whereas FTTH, xDSL technology used at wire line access
5) MPLS use at transport technology
6) Separation of control functions (routing of payload) among bearer (payload: voice, data) capabilities, call/session
7) Decoupling of service provision: One of the main characteristics of NGN is the decoupling of services and transport, allowing them to be offered separately and to evolve independently. Therefore, in the NGN architectures, there shall be a clear separation between the functions for the services and the functions for the transport. NGN allows the provisioning of both existing and new services independently of the network and the access type used**.**
8) The separation is represented by two distinct blocks or stratum of functionality. The transport functions reside in the transport stratum and the service functions related to applications reside in the service stratum as shown in Fig. 2.20

Fig. 2.20: Separation of services from transport in NGN

[**Courtesy: (T-REC-Y.2011-200410-I!!PDF-E%20**ITU-T Y.2001]

9) Interworking with legacy (old) networks via open interfaces**:** Many services have to be operated across a hybrid combination of NGN and non NGN technologies. In such cases interworking between NGNs of different operators and between NGN and existing networks such as PSTN (Public Switched Telephone Network), ISDN (Integrated Services Digital Network) and GSM (Global System for Mobile communications) is provided by means of gateways.
10) NGN Provides Generalized mobility is the ability for the user or other mobile entities to communicate and access services irrespective of changes of the location (anywhere in world) or technical environment. (technology independent)
11) Unrestricted access by users to different service providers**,** User can access services of different service providers along with own service provider.
12) Variety of identification schemes: Since the NGN consists of interconnected heterogeneous networks, using heterogeneous user access and heterogeneous user devices and that the NGN should provide a seamless capability, independent of access method and network, the NGN should address Numbering, Naming and Addressing. NGN can identify different naming and numbering scheme such as: telephone scheme, IPV4, IPV6, E.164, Enum

13) NGN compliant with all regulatory requirements, for example concerning emergency communications, security, privacy, lawful interception, etc.
14) Reliability: To improve reliability, every communication device is highly reliable. Provision of redundant configuration for communication circuits and equipment's is maintained
15) NGN is Layered Architecture having four layers:
    a) Access Layer
    b) Transport /Core Layer
    c) Control Layer
    d) Service Layer

e. NGN Architecture

Fig. 2.21 shows the functional block diagram of the NGN architecture. The NGN functional architecture supports the UNI, NNI, ANI and SNI reference points. The NGN architecture supports the delivery of multimedia services and content delivery services, including video streaming and broadcasting. An aim of the NGN is to serve as aPSTN and ISDN replacement. The NGN architecture defines a Network-Network Interface (NNI), User-Network Interface (UNI), and an Application Network Interface (ANI).

The NGN functions are divided into:

- Service stratum (layer) functions
- Transport stratum (layer) functions.

To provide services, several functions in both the service stratum and the transport stratum are needed. The delivery of services/applications to the end-user is provided by utilizing the application support functions and service support functions, and related control functions.

Fig 2.21 NGN Architecture

i. Transport stratum functions

The transport stratum functions include transport functions and transport control functions.

1) **Transport functions**: The transport functions provide the connectivity for all components and physically separated functions within the NGN. These functions provide support for unicast and/or multicast transfer of media information, as well as the transfer of control and management information. Transport functions include access network functions, edge functions, core transport functions, and gateway functions.
2) **Transport control functions**: The transport control functions include resource and admission control functions, network attachment control functions as well as mobility management and control functions.
3) **Access network functions**: The access network functions take care of end-user's access to the network as well as collecting and aggregating the traffic coming from these accesses towards the core network. These functions also perform QoS control mechanisms dealing directly with user traffic, including buffer management, queuing and scheduling, packet filtering, traffic classification, marking, policing, and shaping. In addition, the access network provides support for mobility. The access network includes access-technology dependent functions, e.g., for W-CDMA technology and xDSL access.

   Depending on the technology used for accessing NGN services, the access network includes functions related to:

   a) Cable access;
   b) xDSL access;

c) Wireless access (e.g., [b-IEEE 802.11] and [b-IEEE 802.16] technologies and 3G RAN access)
d) Optical access.

4) **Gateway functions**: The gateway functions provide capabilities to interwork with end-user functions and/or other networks, including other types of NGN and many existing networks, such as the PSTN/ISDN, the public Internet, and so forth. Gateway functions can be controlled either directly from the service control functions or through the transport control functions
5) **Resource and admission control functions (RACF):** Within the NGN architecture, the resource and admission control functions (RACF) act as the arbitrator between service control functions and transport functions for QoS. The RACF provides an abstract view of transport network infrastructure to service control functions (SCF) and makes service stratum functions agnostic to the details of transport facilities, such as network topology, connectivity, resource utilization and QoS mechanisms/technology.
6) **Network attachment control functions (NACF):** The network attachment control functions (NACF) provide registration at the access level and initialization of end-user functions for accessing NGN services. They also announce the contact point of NGN functions in the service stratum to the end user.

   The NACF provides the following functionalities:
   a. Dynamic provisioning of IP addresses and other user equipment configuration parameters.
   b. By endorsement of the user, auto-discovery of user equipment capabilities and other parameters.
   c. Authentication of end user and network at the IP layer (and possibly other layers). Regarding the authentication, mutual authentication between the end user and the network attachment is performed.
   d. Authorization of network access, based on user profiles.
   e. Access network configuration, based on user profiles.

ii. Service stratum functions

Functional grouping in the service stratum includes:

a) The service control and content delivery functions including service user profile functions.
b) The application support functions and service support functions.
   1) **Service control functions (SCF):** The service control functions (SCF) include resource control, registration and authentication- authorization functions at the service level for both mediated and non-mediated services. They can also include functions for controlling media resources, i.e., specialized resources and gateways at the service-signaling level.
   2) **Content delivery functions (CDF):** The content delivery functions (CDF) receive content from the application support functions and service support

functions, store, process, and deliver it to the end-user functions using the capabilities of the transport functions, under control of the service control functions.

3) **Application support functions and service support functions (ASF&SSF)**: The application support functions and service support functions (ASF&SSF) include functions such as the gateway, registration, authentication and authorization functions at the application level. These functions are available to the "applications" and "end-user" functional groups. Through the UNI, the application support functions and service support functions provide reference points to the end-user functions. Application interactions with the application support functions and service support functions are handled through the ANI reference point.
4) **End-user functions**: No assumptions are made about the diverse end-user interfaces and end-user networks that may be connected to the NGN access network. End-user equipment may be either mobile or fixed.
5) **Management functions:** These functions provide the capabilities to manage the NGN in order to provide NGN services with the expected quality, security, and reliability. All the NGN components are centralized controlled.

f. NGN Network components

Components of NGN Network is shown in Fig. 2.22

Fig. 2.22: NGN Network Components [TEC]

1) **Media Gateway Controller (MGC)**: Media gateway controllers are also known as Soft switches and Call controllers, Wireless Call Server or Call Agents. The MGC is located in the service provider's network in control layer and provides call logic and call control functions, typically maintaining call state for every call in the network. Many MGCs interact with application servers to supply services that are not directly hosted on MGCs.
2) **Media Gateway**: Media Gateways are located in access layer of NGN. Media Gateway performs following functionality:
    i. Access gateway (AG)
    ii. Trunk Media gateway (TMG)

iii. Signaling gateway (SG)
iv. Border Gateway (BGW)/ Session Border Controller(SBC)

3) **Access gateway (AG)**: The AG is located in the service provider's network. It supports the line side interface to the core IP network for use by phones, devices, and PBXs. This element provides functions such as media conversion (circuit to Packet, Packet to circuit) and echo control.
4) **Trunk Media gateway (TMG)**: The TMG supports a trunk side interface to the PSTN and/or IP routed flows in the packet network. It supports functions such as packetisation, echo control etc.
5) **Signaling gateway (SG):** The SG provides the signaling interface between the VoIP network and the PSTN signaling network. It terminates SS7 links and provides Message Transport Part (MTP) Level 1 and Level 2 functionality. Each SG communicates with its associated CA to support the end-to-end signaling for calls.
6) **Border Gateway (BGW)/ Session Border Controller (SBC)**: It is deployed at the edge and core of a service provider's network to control signaling and media streams as they enter and exit the network.
   i The "edge" is any IP-IP network border such as between a service provider and a customer or between a service provider and an enterprise network.
   ii The "core" is any IP-IP network border such as those between two service providers.

   SBC provides functions such as security, denial of Service attacks, overload control, Network Address Translation and Firewall Traversal, Lawful Interception, Quality of Service (QoS) management, Protocol Translation, call accounting etc.
7) **Access network (AN)**: The access network provides connectivity between the customer premises equipment and the access gateways in the service provider's network. There are various access methods:
   i TDM direct access,
   ii Switched TDM,
   iii Broadband access (cable, DSL),
   iv IP managed Internet service, etc.
8) **IP core network**: The primary function of the IP core network is to provide routing and transport of IP packets. The IP core also has the added value of architecturally isolating the gateways, and their associated access networks, from the MGC and associated service intelligence.
9) **Media Server:** The Media Server is located in the service provider's network and uses a control protocol such as H.248 or SIP, under the control of the MGC or application server, to provide announcements and tones, and collect user information.
10) **Application Server:** The Application Server is located in the service provider's network and provides the service logic and execution for one or more applications

or services that are not directly hosted on the MGC. Typically, the MGC routes calls to the appropriated AS for features the MGC does not support.

g. NGN Wireless Technology

1. Telecom network Spectrum - Licensed and Unlicensed Radio Bands

The radio spectrum is the part of the electromagnetic spectrum with frequencies from 30 hertz to 300 GHz. Electromagnetic waves in this frequency range, called radio waves, are widely used in modern technology, particularly in telecommunication. Radio Spectrum, in general, can be categorized into two types, licensed radio bands and unlicensed radio bands.

i. **Licensed radio bands**-To use this radio bands, a license must be obtained from a government agency. This requirement is true of all users of these radio spectrums. A few of the uses of licensed radio bands are are given in Table 2.1:

Table 2.1: Licensed Radio Bands

Sr.No	Type	Frequency Range
1	AM Broadcast a. Short Wave b. Medium Wave c. Long Wave	1.711MHz-30.0 MHz 520MHz – 1610 MHz 148.5KHz-283.5KHz
2	FM Broadcast	87.5 MHz-108.0 MHz
3	Cellular Phone	840 MHz, 900 MHz

ii. **Unlicensed radio bands**: Unlicensed radio bands have been allocated to certain users by the government or any individual can use it, but to be able to use and broadcast on these bands, you do not need to have a license; you only need to create compliant devices that are to be used. Regulations exist around these bands. Some of the types of unlicensed radio bands are given in Table 2.2:

Table 2.2: Unlicensed Radio Bands

Sr.No	Type	Frequency
1	Industrial, Scientific, Medical [ISM] Includes several Medical monitors and Industrial devices.	900 MHz, 2.4 GHz, 5 GHz.
2	Unlicensed National Information Infrastructure (U-NII) Defines specifications for the use of wireless devices such as WLAN access points and routers	5GHz Band

Standard Bodies: IEEE 802.11 networks have several choices of wireless bands that are available to them to use, without the requirement to lease the frequencies from the government. Following groups and standards bodies have helped to develop standards so that all users can be good neighbors with others who use those radio bands.

Table 2.3: Standard Organization Bodies

S No	Abbreviation	Full Form	Function
1	FCC	Federal Communications Commission	Manages and sets standards with regard to the spectrum use
2	IEEE	Institute of Electrical and Electronics Engineers	A leading standards organization which publishes standards that are adopted across industries
3	ETSI	European Telecommunications Standards Institute	Another standards organization that has contributed many worldwide standards
4	ITU-R	International Telecommunication Union, Radio communication Sector	With the FCC, defines how WLANs should operate from a regulatory perspective, such as operating frequencies, antenna gain, and transmission power
5	WLANA	WLAN Association	Provides information resources related to WLANs with regard to industry trends and usage.
6	WPC	The Wireless Planning and Coordination	National Radio Regulatory Authority responsible for Frequency Spectrum Management, including licensing and caters for the needs of all wireless users (Government and Private) in the India. It exercises the statutory functions of the Central Government and issues licenses to establish, maintain and operate wireless stations

h. Mobile Network Evolution (1G to 5G):

In the last few decades, Mobile Wireless Communication networks have experienced a remarkable change. The mobile wireless Generation (G) generally refers to a change in the nature of the system, speed, technology, frequency, data capacity, latency etc. Each generation have some standards, different capacities, new techniques and new features which differentiate it from the previous one.

1. The first generation (1G) mobile wireless communication network was analog used for voice calls only.
2. The second generation (2G) is a digital technology and supports text messaging.
3. The third generation (3G) mobile technology provided higher data transmission rate, increased capacity and provide multimedia support.

4. The fourth generation (4G) integrates 3G with fixed internet to support wireless mobile internet, which is an evolution to mobile technology and it overcome the limitations of 3G. It also increases the bandwidth and reduces the cost of resources.
5. 5G stands for 5th Generation Mobile technology and is going to be a new revolution in mobile market which has changed the means to use cell phones within very high bandwidth. User never experienced ever before such high value technology which includes all type of advance features and 5G technology will be most powerful and in huge demand in near future.

i. 1G – First generation mobile communication system

The first generation of mobile network was deployed in Japan by Nippon Telephone and Telegraph company (NTT) in Tokyo during 1979. In the beginning of 1980s, it gained popularity in the US, Finland, UK and Europe. This system used analogue signals and it had many disadvantages due to technology limitations.

Most popular 1G system during 1980s

- Advanced Mobile Phone System (AMPS)
- Nordic Mobile Phone System (NMTS)
- Total Access Communication System (TACS)
- European Total Access Communication System (ETACS)

Key features (technology) of 1G system

- Frequency 800 MHz and 900 MHz
- Bandwidth: 10 MHz (666 duplex channels with bandwidth of 30 KHz)
- Technology: Analogue switching
- Modulation: Frequency Modulation (FM)
- Mode of service: voice only
- Access technique: Frequency Division Multiple Access (FDMA)

Disadvantages of 1G system

- Poor voice quality due to interference
- Poor battery life
- Large sized mobile phones (not convenient to carry)
- Less security (calls could be decoded using an FM demodulator)
- Limited number of users and cell coverage
- Roaming was not possible between similar systems

ii. 2G – Second generation communication system GSM

Second generation of mobile communication system introduced a new digital technology for wireless transmission also known as Global System for Mobile communication (GSM). GSM technology became the base standard for further development in wireless standards later. This standard was capable of supporting up to 14.4 to 64kbps (maximum) data rate which is sufficient for SMS and email services.

Code Division Multiple Access (CDMA) system developed by Qualcomm also introduced and implemented in the mid-1990s. CDMA has more features than GSM in terms of spectral efficiency, number of users and data rate.

Key features of 2G system

- Digital system (switching)
- SMS services is possible
- Roaming is possible
- Enhanced security
- Encrypted voice transmission
- First internet at lower data rate

Disadvantages of 2G system

- Low data rate
- Limited mobility
- Less features on mobile devices
- Limited number of users and hardware capability

iii. 3G – Third generation communication system

Third generation mobile communication started with the introduction of UMTS – Universal Mobile Terrestrial / Telecommunication Systems. UMTS has the data rate of 384kbps and it support video calling for the first time on mobile devices. After the introduction of 3G mobile communication system, smart phones became popular across the globe. Specific applications were developed for smartphones which handles multimedia chat, email, video calling, games, social media and healthcare.

Key features of 3G system

- Higher data rate
- Video calling
- Enhanced security, more number of users and coverage
- Mobile app support
- Multimedia message support
- Location tracking and maps
- Better web browsing
- TV streaming
- High quality 3D games

Disadvantages of 3G systems

- Expensive spectrum licenses
- Costly infrastructure, equipment and implementation
- Higher bandwidth requirements to support higher data rate
- Costly mobile devices

- Compatibility with older generation 2G system and frequency bands

iv. 4G – Fourth generation communication system

4G systems are enhanced version of 3G networks developed by IEEE, offers higher data rate and capable to handle more advanced multimedia services. LTE and LTE advanced wireless technology used in 4th generation systems. Furthermore, it has compatibility with previous version thus easier deployment and upgrade of LTE and LTE advanced networks are possible.

Simultaneous transmission of voice and data is possible with LTE system which significantly improve data rate. All services including voice services can be transmitted over IP packets. Complex modulation schemes and carrier aggregation is used to multiply uplink / downlink capacity.

Wireless transmission technologies like WiMax are introduced in 4G system to enhance data rate and network performance.

Key features of 4G system

- Higher data rate up to 1Gbps
- Improved security and mobility
- Reduced latency for mission critical applications
- High definition video streaming and gaming
- Voice over LTE network VoLTE

Disadvantages of 4G system

- Expensive hardware and infrastructure
- Costly spectrum in most countries, frequency bands are too expensive.
- High end mobile devices compatible with 4G technology required, which is costly
- Wide deployment and upgrade is time consuming

v. 5G – Fifth generation communication system

5G network is using advanced technologies to deliver ultra-fast internet and multimedia experience for customers. Existing LTE advanced networks will transform into supercharged 5G networks in future.

In earlier deployments, 5G network will function in non-standalone mode and standalone mode. In non-standalone mode both LTE spectrum and 5G-NR spectrum will be used together. Control signaling will be connected to LTE core network in non-standalone mode.

Key features of 5G technology

- Ultra-fast mobile internet up to 10Gbps
- Low latency in milliseconds significant for mission critical applications
- Total cost deduction for data
- Higher security and reliable network

- Uses technologies like small cells, beam forming to improve efficiency
- Forward compatibility network offers further enhancements in future
- Cloud based infrastructure offers power efficiency, easy maintenance and upgrade of hardware

Table 2.4: Comparison of All Generations of Mobile Technologies:

Technology	1G	2G	3G	4G	5G
Start/ Development	1970-80	1990-2004	2004-10	2010-12	Soon (probably by 2020)
Data bandwidth	2 Kbps	64 Kbps	2 Mbps	1 Gbps	More than 1 Gbps
Technology	Analog	Digital	CDMA-2000, UMTS,EDGE	Wi max, Wi-Fi, LTE	WWW
Core network	PSTN	PSTN	Packet Network	Internet	Internet
Multiplexing	FDMA	TDMA/ CDMA	CDMA	CDMA	CDMA
Switching	Circuit	Circuit Packet	Packet (except Air interference)	All Packet	All Packet
Primary service	Analog phone calls	Digital phone calls and Messaging	Phone calls and Messaging, data. Integrated high quality audio, video and data	Dynamic integration of access, variable Devices. All IP service (including voice messages)	Dynamic integration of access, variable devices with air interference. High speed, High capacity and provide large broadcasting of data in Gbps
Key differentiator	Mobility	Secure, Mass adoption	Better Internet experience	Faster broadband Internet , lower Latency	Better coverage and no Dropped calls ,much lower latency, Better performance

NGN Core

MPLS (Multi-protocol label switching) is used at core transport layer in NGN network. MPLS provides faster switching, propagation delay is less.

a. MPLS Concept

In NGN, a packet of a connectionless network layer protocol travels from one router to the next, each router makes an independent forwarding decision for that packet. That is, each router analyzes the packets header, and each router runs a network layer routing algorithm. Each router independently chooses a next hop for the packet, based on its analysis of the packet's header and the results of running the routing algorithm.

Packet headers contain considerably more information than is needed simply to choose the next hop. Choosing the next hop can therefore be thought of as the composition of two functions.

i. The first function partitions the entire set of possible packets into a set of Forwarding Equivalence Classes (FECs).
ii. The second maps each FEC to a next hop.

b. MPLS Basics

1) **Labels:** A label is a short, fixed length, locally significant identifier which is used to identify a FEC. The label which is put on a particular packet represents the Forwarding Equivalence Class to which that packet is assigned. Most commonly, a packet is assigned to a FEC based on its network layer destination address. Each IP destination network has a different label which has local significance. Label for a destination network changes in each hop.

2) **Label Switch Router**: A label switch router (LSR) is a router that supports MPLS. It is capable of understanding MPLS labels and of receiving and transmitting a labeled packet on a data link. Three kinds of LSRs exist in an MPLS network:

Fig. 2.23 MPLS Architecture

a. **Provider Edge [PE] Router/Ingress LSRs**: Ingress LSRs receive a packet that is not labeled yet, insert a label (stack) in front of the packet, and send it on a data link.

b. **Provider Edge [PE] Router/Egress LSRs**: Egress LSRs receive labeled packets, remove the label(s), and send them on a data link. Ingress and egress LSRs are edge LSRs.

c. **Core Edge [CE] Router/Intermediate LSRs**: Intermediate LSRs receive an incoming labeled packet, perform an operation on it, switch the packet, and send the packet on the correct data link.

d. **Upstream and Downstream LSRs**:
An LSR can do the three operations: PUSH/adding labels, POP/remove labels, or SWAP/interchanging labels.
   i. It must be able to pop one or more labels (remove one or more labels from the top of the label stack) before switching the packet out.
   ii. An LSR must also be able to push one or more labels onto the received packet.
   iii. If the received packet is already labeled, the LSR pushes one or more labels onto the label stack and switches out the packet.
   iv. If the packet is not labeled yet, the LSR creates a label stack and pushes it onto the packet. An LSR must also be able to swap a label. This simply means that when a labeled packet is received, the top label of the label stack is swapped with a new label and the packet is switched on the outgoing data link.
   v. An LSR that pushes labels onto a packet that was not labeled yet is called an imposing LSR because it is the first LSR to impose labels onto the packet.
   vi. One that is doing imposition is ingress LSR. An LSR that removes all labels from the labeled packet before switching out the packet is a disposing LSR.
   vii. One that does disposition is an egress LSR.

e. **Label Switch path**:A label switched path (LSP) is a sequence of LSRs that switch a labeled packet through an MPLS network or part of an MPLS network. Basically, the

LSP is the path through the MPLS network or a part of it that packets take. The first LSR of an LSP is the ingress LSR for that LSP, whereas the last LSR of the LSP is the egress LSR. All the LSRs in between the ingress and egress LSRs are the intermediate LSRs. LSP is unidirectional. The flow of labeled packets in the other direction right to left between the same edge LSRs would be another LSP.

MPLS Features:

1 QoS enabled MPLS transport network will provide real time and data transport application.
2 MPLS increases operator revenue.
3 MPLS offers high transport efficiency by using hybrid technology (packetswitching and circuit switching)
4 MPLS offers high reliability and operational simplicity.
5 MPLS provide control and deterministic usage of network resources, end-to-end control to engineer network paths and to efficiently utilize network resources.
6 In MPLS network management is simple
7 It ensures smooth interworking of the packet transport network with other existing/legacy packet networks.
8 MPLS header length is of 32 bits, in which label length is 20 bits, 3 bits for service quality and 8bits for Timeto live (no of hops in network) and 1 bit for stacking the labels.

MPLS Advantages

1. **Consistent Network Performance**: The MPLS allows different Class of Service classifications to be applied to packets, to help ensure that data loss (packet loss), transmission delays (latency) and variations in transmission delay (jitter) are kept without appropriate limits.
2. **Obscures Network Complexity**: MPLS can effectively hide the underlying complexity of the network from devices and users that don't need to know about it.
3. **Reduced Network Congestion**: The MPLS can support Traffic Engineering having many uses, including the re-routing of comparatively delay-tolerant traffic to slower, circuitous, under-utilised routes, freeing up capacity on quicker (lower-latency) overcrowded paths.
4. **Increased Uptime**: MPLS is that it has the potential to increase uptime as Fast Reroute that enables traffic to be switched to an alternative path very rapidly
5. **Scalability**: According to the needs, the MPLS can be scaled up and down
6. **Efficiency:** MPLS offers much higher quality connections without packet loss and jitter using it along with VoIP may lead to increased efficiency
7. **Reliability:** Since MPLS uses labels for forwarding packets, it can be assured that the packets will be delivered to the right destination.

References:

- https://data-flair.training/blogs/iot-applications/
- https://kainjan1.files.wordpress.com/2018/01/chapter-1_iot.pdf
- https://www.tutorialspoint.com/internet_of_things/internet_of_things_tutorial.pdf
- https://www.iare.ac.in/sites/default/files/lecture_notes/IOT%20LECTURE%20NOTES_IT.pdf

- https://www2.deloitte.com/content/dam/insights/us/articles/iot-primer-iot-technologies-applications/DUP_1102_InsideTheInternetOfThings.pdf
- https://techdifferences.com/difference-between-sensors-and-actuators.html
- https://electronicsforu.com/technology-trends/tech-focus/iot-sensors
- https://iotbytes.wordpress.com/basic-iot-actuators/
- https://www.rfpage.com/evolution-of-wireless-technologies-1g-to-5g-in-mobile-communication /#:~:text=In%20order%20to%20support%20higher ,to%20473.6kbps%20(maximum).
- https://www.itu.int/rec/dologin_pub.asp?lang=e&id=T-RECY.2012-200609-S!!PDF-E&type=items
- https://pdfs.semanticscholar.org/4dfd/40cc3a386573ee861c5329ab4 c6711210819.pdf
- www.tec.gov.in
- www.trai.gov.in

Sample Multiple Choice Questions:

1. IoT stands for
   a. Internet of Technology
   b. Intranet of Things
   c. Internet of Things
   d. Information of Things

2. Which is not the Feature of IoT
   a. Connectivity
   b. Self-Configuring
   c. Endpoint Management
   d. Artificial Intelligence

3. Devices that transforms electrical signals into physical movements.
   a. Sensors
   b. Actuators
   c. Switches
   d. display

4. IPv4 uses a _____address scheme.
   a. 8-bit
   b. 16-bit
   c. 32-bit
   d. 64 bit

5. Data Distribution Service (DDS) is a _____standard for device-to-device or machine-to-machine communication
   a. data-centric middleware
   b. data-centric hardware
   c. data-centric Software
   d. None of above

6. The _____provide the connectivity for all components and physically separated functions within the NGN
    a. Transport functions
    b. Access network functions
    c. Gateway functions
    d. Resource and admission control functions (RACF)

7. The _____ take care of end-user's access to the network as well as collecting and aggregating the traffic coming from these accesses towards the core network.
    a. Transport functions
    b. Access network functions
    c. Gateway functions
    d. Resource and admission control functions (RACF)

8. _____ functions also perform QoS control mechanisms dealing directly with user traffic, including buffer management, queuing and scheduling, packet filtering, traffic classification, marking, policing, and shaping.
    a. Transport functions
    b. Access network functions
    c. Gateway functions
    d. Resource and admission control functions (RACF)

9. The _____ provide capabilities to interwork with end-user functions and/or other networks, including other types of NGN and many existing networks, such as the PSTN/ISDN, the public Internet, and so forth.
    a. Transport functions
    b. Access network functions
    c. Gateway functions
    d. Resource and admission control functions (RACF)

Unit-3 Blockchain Technology

Content

3.1 Introduction to Blockchain
- Backstory of Blockchain
- What is Blockchain?

3.2 Centralize versus Decentralized System

3.3 Layers of Blockchain
- Application Layer
- Execution Layer
- Semantic Layer
- Propagation Layer
- Consensus Layer

3.4 Importance of Blockchain
- Limitations of Centralized Systems
- Blockchain Adoption So Far

3.5 Blockchain Use and Use Cases

3.1 Introduction to Blockchain

Backstory of Blockchain

When the Internet was made available to the public using the World Wide Web (WWW) in the early 1990s, it was supposed to be more open and P2P peer-to- peer because it was built atop the open and decentralized TCP/IP.

The blockchain technology was defined in 1991 by the research scientist Stuart Haber and W. Scott Stornetta. They wanted to introduce a computationally practical key for time-stamping digital documents so that they could not be backdated or tampered by any one. They develop a system using the concept of cryptographically secured chain of blocks to store the time-stamped documents.

- **1991:** In 1991, researcher scientists named Stuart Haber and W. Scott Stornetta introduce Blockchain Technology and required some Computational Practical Solution for time-stamping the digital documents so that they couldn't be tempered or misdated. Hence both scientists together developed a system using Cryptography where, the time-stamped documents are stored in a Chain of Blocks.
- **1992**: Iin 1992, Merkle Trees established a legal corporation by using a system developed by Stuart Haber and W. Scott Stornetta with additional features. Hence, Blockchain Technology became efficient to store several documents to be collected into one block. Merkle used a Secured Chain of Blocks that stores multiple data records in a sequence.
- **2000:** In the year 2000, Stefan Konst published theory of cryptographic secured chains, plus ideas for implementation.
- **2004:** In the year 2004, Cryptographic activist Hal Finney presented a system for digital cash known as "Reusable Proof of Work" which was the game-changer in the history of Blockchain and Cryptography. This System helps to solve the Double

Spending Problem by keeping the ownership of tokens registered on a trusted server.

- **2008**: In the year 2008, Satoshi Nakamoto conceptualized the concept of "Distributed Blockchain" called as "A Peer to Peer Electronic Cash System". Satoshi Nakamoto modified the model of Merkle Tree and developed a system which is more secure and contains the secure history of data exchange. This System follows a peer-to-peer network of time stamping and became so useful that Blockchain become the backbone of Cryptography.
- **2009:** In the year 2009, the evolution of Blockchain is steady and promising and required in various fields. Blockchain technology is so secure that the following surprising news will give proof of that. James Howells was an IT worker in the United Kingdom, started mining bitcoins which are part of Blockchain in 2009 and stopped this in 2013. He spends $17,000 on it and after he stopped he sells the parts of his laptop on eBay and keep the drive with him so that when he needs to work again on bitcoin he will utilize it but while cleaning his house in 2013, he thrashed his drive with garbage and now his bitcoins cost nearly $127 million. This money now remains unclaimed in the Bitcoin system.
- **2014:** The year 2014 is marked as the turning point for blockchain technology. Blockchain technology is separated from the currency and Blockchain 2.0 is born. Financial institutions and other industries started shifting their focus from digital currency to the development of blockchain technologies.
- **2015:** In 2015, Ethereum Frontier Network was launched, thus enabling developers to write smart contracts and dApps that could be deployed to a live network. In the same year, the Linux Foundation launched the Hyperledger project.
- **2016:** The word Blockchain is accepted as a single word instead of two different concepts as they were in Nakamoto's original paper. The same year, a bug in the Ethereum DAO code was exploited resulting in a hard fork of the Ethereum Network. The Bitfinex bitcoin exchange was hacked resulting in 120,000 bitcoins being stolen.
- **2017:** In the year 2017, Japan recognized Bitcoin as a legal currency. Block.one company introduced the EOS blockchain operating system which was designed to support commercial decentralized applications.
- **2018:** Bitcoin turned 10 in the year 2018. The bitcoin value continued to drop, reaching the value of $3,800 at the end of the year. Online platforms like Google, Twitter, and Facebook banned the advertising of cryptocurrencies.
- **2019:** In the year 2019, Ethereum network transactions exceeded 1 million per day. Amazon announced the general availability of the Amazon Managed Blockchain service on AWS.
- **2020:** Stablecoins were in demand as they promised more stability than traditional cryptocurrencies. The same year Ethereum launched Beacon Chain in preparation for Ethereum 2.0.

What is Blockchain?

Blockchain is a common, unchallengeable digital ledger that allows the process of recording transactions and tracking assets in a business network. Where an *asset* may be tangible such as a house, car, cash, land or intangible such as intellectual property, patents, copyrights, branding. Virtually everything of value can be tracked and traded on a blockchain network, decreasing risk and cutting costs for all involved.

Business runs on information can says better where faster it's received and the more accurate it is. Blockchain is perfect which provides immediate, shared and completely transparent information stored on an immutable digital ledger that can be accessed only by permissioned network members. A blockchain network can track orders, payments, accounts, production etc. Hence members share a single view of the fact, other can see all details of a transaction end to end, giving greater sureness, as well as new efficiencies and opportunities.

- Blockchain is a peer-to-peer system of transacting values with no trusted third parties in between.
- It is a shared, decentralized, and open ledger of transactions. This ledger database is replicated across a large number of nodes.
- This ledger database is an append-only database and cannot be changed or altered. It means that every entry is a permanent entry. Any new entry on it gets reflected on all copies of the databases hosted on different nodes.
- There is no need for trusted third parties to serve as intermediaries to verify, secure, and settle the
- transactions.
- It is another layer on top of the Internet and can coexist with other Internet technologies.
- Just the way TCP/IP was designed to achieve an open system, blockchain technology was designed to enable true decentralization. In an effort to do so, the creators of Bitcoin open-sourced it so it could inspire many decentralized applications.

A typical blockchain may look as shown in Figure 3.1.

Fig. 3.1 The blockchain data structure

The genesis block is the first block in any blockchain referred as Block 0 and there is no previous block for reference. The Genesis Block has previous hash value set to 0 to indicate no data was processed before the Genesis Block. The hash of genesis

block is added to all new transactions in a new block and used to create its unique hash. This process is repeated until all the new blocks are added to a blockchain.

Every node on the blockchain network has an identical copy of the blockchain shown in Figure 3.1, where every block is a list of transactions. There are two major parts in every block. The “header” part pointer links back to the previous block in the chain and every block header contains the hash of the previous block so that no one can modify any transaction in the previous block. The other part of a block contents the validated list of transactions, their amounts, timestamp, the addresses of the parties involved, and some another details.

3.2 Centralize versus Decentralized System

Blockchain is designed to be decentralized design, instead of the centralized design. Whether a system is centralized or decentralized, it can still be distributed. A centralized distributed system is one in which there is a master node also sometime called as server responsible for breaking down the tasks or data and distribute the load across nodes. On the other hand, a decentralized distributed system is one where there is no “master” node as such and yet the computation may be distributed. Blockchain is one such example.

Centralized System

Figure 3.2 shows representation of a centralized system.

Fig.3.2 Centralized System

A centralized system has a centralized control with all administrative rights and are easy to design, maintain, enforce trust, and administrate, but suffer from many intrinsic limitations, as given below:

1. A centralized system has a central point of failure; hence they are less stable.
2. A centralized system is more vulnerable to attack and hence less secured.
3. Centralization of power may lead to unethical operations.
4. Most of the time, scalability is difficult.

Decentralized System

Figure 3.3 shows representation of a decentralized system.

Fig.3.3 Decentralized System

A decentralized system does not have a centralized control and every node has equal authority. Such systems are difficult to design, maintain, govern, or impose trust. But, decentralized system do not suffer from the limitations of conventional centralized systems. Decentralized systems offer the following advantages:

1. A decentralized system is more stable and fault tolerant as they do not have a central point of failure.
2. Attack resistant, as no central point to easily attack and hence more secured
3. Symmetric system with equal right to all nodes, so less scope of unethical operations and usually independent in nature

A distributed system may also be decentralized system, for example, blockchain. But, the task is not subdivided and delegated to nodes unlike common distributed systems, because there is no master that in blockchain. A typical decentralized and distributed system is efficiently a peer-to-peer system as shown in Figure 3.3.

Fig.3.3 A decentralized and peer-to-peer system

3.3 Layers of Blockchain

Blockchain have a collection of distributed nodes where immutable transactions are repeated again and again. The blockchain technology is made of a layered architecture where all these layers are present on all the nodes as shown in Figure 3.5.

Fig.3.5 Layers of Blockchain

1. Application Layer

In the application layer, you can find smart contracts, decentralized applications (DApps), user interfaces (UIs), and chain code, which is the fifth layer of the blockchain layers. Hence, this layer consists of services and application programming interfaces (APIs), client-side programming constructs, scripting, development frameworks that offer other apps with access to the blockchain network. Application Layer acts as the front end tool of the blockchain and is basically what a user would typically encounter when operating within a blockchain network.

2. Execution Layer

The Execution Layer execute the instructions of application in the Application Layer on all the nodes in a blockchain network. The instructions could be simple instructions or a set of multiple instructions in the form of a smart contract. In either case, a program or a script desires to be executed to confirm the correct execution of the transaction. All the nodes in a blockchain network have to execute the programs or scripts autonomously. Deterministic execution of programs or scripts on the same set of inputs and conditions always produces the same output on all the nodes, which helps to avoid inconsistencies.

3. Semantic Layer

Semantic Layer layer also called as logical layer of blockchain layer and deals in validation of the transactions done in the blockchain network and also validating the blocks being created in the network. When a transaction initiated by a node, the set of instruction are executed on the execution layer and validated by the semantic layer. Semantic layer is also accountable for the linking of the blocks created in the network. Each block in the blockchain contains the hash of the earlier block except the Start block and this linking of block needs to be defined on Semantic layer.

4. Propagation Layer

A propagation layer is used in the peer-to-peer communications between the nodes that allow them to discover each other and get synchronized with another node in a network. When a transaction is done, then gets broadcasted to all other nodes in the network. Also, when a node advises a block, it is immediately get broadcast in the entire network so that other nodes can use this newly created block and work on it. Hence, the propagation of the block or a transaction in the network is defined in this layer and confirms the stability of the complete network. In the asynchronous Internet network, there are often latency issues for transaction or block propagation. However, depending upon the network capacity or bandwidth, the propagation may occur instantly or it may take a longer.

5. Consensus Layer

Consensus layer is the first layer for most of the blockchain systems and main purpose is to make sure that all the nodes must get approve on a common state of the shared ledger. Consensus layer also deals with the safety and security of the blockchain. There are many consensus algorithms which can be applied for generation of cryptocurrencies like Bitcoin and Ethereum, they use proof-of-work mechanism to select a node randomly out of various nodes present on the network that can propose a new block. Once a new block is created, the block is propagated to all the other nodes to check if the new block is valid or not with the transactions in it and based on the consensus from all other nodes the new block gets added on to the blockchain.

3.4 Importance of Blockchain

1. Blockchain helps in the verification and traceability of multistep transactions needing verification and traceability.
2. Blockchain can provide secure transactions, reduce compliance costs, and speed up data transfer processing.
3. Blockchain technology can help contract management and audit the origin of a product.
4. Blockchain can be used in voting platforms and managing titles and deeds.
5. The transactions are done instantly and transparently, as the ledger is updated automatically.
6. As Blockchain is a decentralized system, no intermediary fee is required
7. In Blockchain, the authenticity of a transaction is verified and confirmed by participants
8. Blockchain is an immutable public digital ledger, which means when a transaction is recorded, it cannot be modified

Limitations of Centralized Systems

1. Trust issues
2. Security issue
3. Privacy issue i.e. data sale privacy is being undermined
4. Cost and time factor for transactions

5. Can't scale up vertically after a certain limit.
6. Bottlenecks can appear when the traffic spikes.

Advantages of decentralized systems over centralized systems:

1. Elimination of intermediaries
2. Easier and genuine verification of transactions
3. Increased security with lower cost
4. Greater transparency
5. Decentralized and immutable

Blockchain Adoption So Far

1. Blockchain came along with Bitcoin, a digital cryptocurrency, in 2009 via a simple mailing list.
2. Some companies started with various flavors of blockchain offerings such as Ethereum, Hyperledger, etc.
3. Microsoft and IBM came up with SaaS (Software as a Service) offerings on their Azure and Bluemix cloud platforms, respectively.
4. Various start-ups and established companies took blockchain initiatives that motivated on solving some of the business problems that were not easily solved before.
5. Financial market, media and entertainment, energy trading, prediction markets, retail chains, loyalty rewards systems, insurance, logistics and supply chains, medical records, government and military applications have adopted Blockchain technology.

3.5 Blockchain Uses and Use Cases

Now, we will see some of the initiatives that are already being taken across industries such as finance, insurance, banking, healthcare, government, supply chains, IoT (Internet of Things), and media & entertainment. Some of the existing use cases are given below.

1. Any type of property or asset may be physical or digital, for example laptops, mobile phones, diamonds, automobiles, real estate, e-registrations, digital files, etc. can be registered on blockchain. Blockchain technology can enable these asset transactions from one person to another, maintain the transaction log, and check validity or ownerships. Also, notary services, proof of existence, tailored insurance schemes use cases can be developed.
2. Many financial use cases being developed on blockchain for example cross-border payments, share trading, loyalty and rewards system, Know Your Customer (KYC) among banks, etc. Initial Coin Offering (ICO) is one of the most trending use cases and is the greatest way of crowdsourcing nowadays by using cryptocurrency as digital assets which is very easy to buy and trade.
3. Blockchain can be used to allow "The Wisdom of Crowds" to take the lead and shape businesses, economies, and various other national phenomena by using collective wisdom. Financial and economic forecasts based on the wisdom of

crowds, decentralized prediction markets, decentralized voting, as well as stocks trading can be possible on blockchain.

4. The process of determining music royalties has always been complicated. The Internet-enabled music streaming services facilitated higher market penetration, but made the royalty determination more complex and can be addressed by blockchain by maintaining a public ledger of music rights ownership information as well as authorised distribution of media content.
5. In IoT –Internet of things, with billions of IoT devices everywhere and many more to join the pool. A whole bunch of different makes, models, and communication protocols makes it difficult to have a centralized system to control the devices and provide a common data exchange platform. This is also an area where blockchain can be used to build a decentralized peer-to- peer system for the IoT devices to communicate with each other. ADEPT (Autonomous Decentralized Peer-To-Peer Telemetry) is a joint initiative from IBM and Samsung that has developed a platform that uses elements of the Bitcoin's underlying design to build a distributed network of devices—a decentralized IOT. ADEPT uses three protocols: BitTorrent for file sharing, Ethereum for smart contracts, and TeleHash for peer-to-peer messaging in the platform. The IOTA foundation is another such initiative.
6. In the government sectors as well, blockchain has gained momentum. There are use cases where technical decentralization is necessary, but politically should be governed by governments: land registration, vehicle registration and management, e-Voting, etc. are some of the active use cases. Supply chains are another area where there are some great use cases of blockchain. Supply chains have always been prone to disputes across the globe, as it was always difficult to maintain transparency in these systems.

References:

- https://www.horizen.io/blockchain-academy/technology/advanced/blockchain-as-a-data-structure/
- https://www.ibm.com/in-en/topics/what-is-blockchain
- https://www.javatpoint.com/blockchain-tutorial

Sample Multiple Choice Questions.

1. What does P2P stand for?

a. Password to Password
b. Peer to Peer
c. Product to Product
d. Private Key to Public Key

2. What is a blockchain?

a. A Currency
b. A centralized ledger
c. A type of cryptocurrency
d. A distributed ledger on a peer to peer network

3. Who first proposed a blockchain-like protocol?
   a. David Chaum
   b. Dave Bayer
   c. W. Scott Stornetta
   d. Stefan Konst

4. Blockchain is a peer-to-peer _____ distributed ledger technology that makes the records of any digital asset transparent and unchangeable.
   a. Secure
   b. Popular
   c. Demanding
   d. Decentralized

5. What is a node?
   a. A Blockchain
   b. An exchange
   c. A type of cryptocurrency
   d. A computer on a Blockchain network

6. Who created Bitcoin?
   a. Elon Musk
   b. Warren Buffett
   c. Satoshi Nakamoto
   d. Mark Zuckerberg

7. A blockchain is a type of?
   a. Database
   b. Object
   c. Table
   d. View

8. What are the benefits of blockchain technology?
   a. Security & Speed
   b. No hidden fees
   c. Fraud control & Access levels
   d. All of the Above

9. What is a dApp?
   a. A type of Cryptocurrency
   b. A condiment
   c. A type of blockchain
   d. A decentralized application

10. What is a genesis block?
   a. The first block of a Blockchain
   b. A famous block that hardcoded a hash of the Book of Genesis onto the blockchain
   c. The first block after each block halving
   d. The 2nd transaction of a Blockchain

Unit-4 Digital Evidences

Content

4.1 Digital forensics
- Introduction to digital forensic
- Digital forensics investigation process
- Models of Digital Forensic Investigation –
  - Abstract Digital Forensics Model (ADFM)
  - Integrated Digital Investigation Process (IDIP)
  - An extended model for cybercrime investigation

4.2 Ethical issues in digital forensic
- General ethical norms for investigators
- Unethical norms for investigation

4.3 Digital Evidences
- Definition of Digital Evidence
- Best Evidence Rule
- Original Evidence

4.4 Characteristics of Digital Evidence
- Locard's Exchange Principle
- Digital Stream of bits

4.5 Types of Evidence : Illustrative, Electronics, Documented, Explainable, Substantial, Testimonial

4.6 Challenges in evidence handling
  - Authentication of evidence
  - Chain of custody
  - Evidence validation

4.7 Volatile evidence

4.1 Digital Forensics

4.1.1 Introduction to Digital Forensics

Forensics science is a well-established science that pays vital role in criminal justice systems. It is applied to both criminal and civil action. Digital forensics sometimes known as digital forensic science, is a branch of forensic science encompassing the recovery and investigation of material found in digital devices, often in relation to computer crime.

Digital forensics includes the identification, recovery, investigation, validation, and presentation of facts regarding digital evidence found on computers or similar digital storage media devices.

4.1.2 History of Forensic

1. Field of pc forensics began in 1980s when personal computers became a viable possibility for the buyer.

2. In 1984, an associate Federal Bureau of Investigation program was created, which was referred to as magnet media program.
3. It is currently referred to as Computer Analysis and Response Team (CART).
4. Michael Anderson, the Father of Computer Forensics, came into limelight during this period.
5. International Organization on Computer Evidence (IOCE) was formed in 1995.
6. In 1997, the great countries declared that law enforcement personnel should be trained and equipped to deal with sophisticated crimes.
7. In 1998, INTERPOL Forensic Science symposium was apprehended.
8. In 1999, the FBI CART case load goes beyond 2000 case examining, 17 terabytes of information.
9. In 2000, the first FBI Regional Computer Forensic Laboratory was recognized.
10. In 2003, the FBI CART case load exceeds 6500 cases, examining 782 terabytes of information.

4.1.3 Rule of Digital Forensics

While performing digital forensics investigation, the investigator should follow the given rules:

Rule 1. An examination should never be performed on the original media.

Rule 2. A copy is made onto forensically sterile media. New media should always be used if available.

Rule 3. The copy of the evidence must be an exact, bit-by-bit copy. (Sometimes referred to as a bit-stream copy).

Rule 4. The computer and the data on it must be protected during the acquisition of the media to ensure that the data is not modified.

Rule 5. The examination must be conducted in such a way as to prevent any modification of the evidence.

Rule 6. The chain of custody of all evidence must be clearly maintained to provide an audit log of whom might have accessed the evidence and at what time.

4.1.4 Definition of Digital Forensics

Digital forensics is a series of steps to uncover and analyze electronic data through scientific method. Major goal of the process is to duplicate original data and preserve original evidence and then performing the series of investigation by collecting, identifying and validating digital information for the purpose of restructuring past events.

4.1.5 Digital Forensic Investigation

Digital forensic investigation (DFI) is a special type of investigation where the scientific procedures and techniques used will be allowed to view the result- digital evidence- to be admissible in a court of law.

4.1.6 Goal of Digital Forensic Investigation:

The main objective of computer forensic investigation is to examine digital evidences and to ensure that they have not been tampered in any manner. To achieve this goal investigation must be able to handle all below obstacles:

1. Handle and locate a certain amount of valid data from a large amount of files stored in the computer system.
2. It is viable that the information has been deleted, in such a situation searching inside the file is worthless.
3. If the files are secured by some passwords, investigators must find a way to read the protected data in an unauthorized manner.
4. Data may be stored in damaged device but the investigator searches the data in working devices.
5. Major obstacle is that each and every case is different. Identifying the techniques and tools will take a long time.
6. The digital data found should be protected from being modified. It is very tedious to prove that data under examination is unaltered.
7. Common procedures for investigation and standard techniques for collecting and preserving digital evidence are desired.

4.1.7 Models of Digital Forensics:

I. Abstract Digital Forensic Model (ADFM)

Reith, Carr, Gunsh proposed the Abstract Digital Forensic model in 2002.

Fig.4.1: Abstract Digital Forensic Model (ADFM)

- **Phases of ADFM model are as follows:**
  1. **Identification** –it recognizes an incident from indicators and determines its type.
  2. **Preparation** –it involves the preparation of tools, techniques, search warrants and monitoring authorization and management support
  3. **Approach strategy** –formulating procedures and approach to use in order to maximize the collection of untainted evidence while minimizing the impact to the victim
  4. **Preservation**–it involves the isolation, securing and preserving the state of physical and digital evidence
  5. **Collection** –This is to record the physical scene and duplicate digital evidence using standardized and accepted procedures
  6. **Examination** –An in-depth systematic search of evidence relating to the suspected crime. This focuses on identifying and locating potential evidence.
  7. **Analysis** –This determines importance and probative value to the case of the examined product
  8. **Presentation -**Summary and explanation of conclusion
  9. **Returning Evidence** –Physical and digital property returned to proper owner

II. Integrated Digital Investigation Process (IDIP)

DFPM along with5 groups and 17 phases are proposed by Carrier and Safford. DFPM is named the Integrated Digital Investigation Process (IDIP). The groups are indexed as shown in following Figure 2.3.

Fig. 4.2: An Integrated Digital Investigation Process

- **The phases of IDIP are as follows:**
  1. **Readiness phase** The goal of this phase is to ensure that the operations and infrastructure are able to fully support an investigation. It includes two phases:
     - Operations Readiness phase
     - Infrastructure Readiness phase
  2. **Deployment phase** The purpose is to provide a mechanism for an incident to be detected and confirmed. It includes two phases:

- Detection and Notification phase; where the incident is detected and then appropriate people notified.
- Confirmation and Authorization phase; which confirms the incident and obtains authorization for legal approval to carry out a search warrant.

3. **Physical Crime Investigation phase** The goal of these phases is to collect and analyze the physical evidence and reconstruct the actions that took place during the incident.
   It includes six phases:

- Preservation phase; which seeks to preserve the crime scene so that evidence can be later identified and collected by personnel trained in digital evidence identification.
- Survey phase; that requires an investigator to walk through the physical crime scene and identify pieces of physical evidence.
- Documentation phase; which involves taking photographs, sketches, and videos of the crime scene and the physical evidence. The goal is to capture as much information as possible so that the layout and important details of the crime scene are preserved and recorded.
- Search and collection phase; that entails an in-depth search and collection of the scene is performed so that additional physical evidence is identified and hence paving way for a digital crime investigation to begin
- Reconstruction phase; which involves organizing the results from the analysis done and using them to develop a theory for the incident.
- Presentation phase; that presents the physical and digital evidence to a court or corporate management.

4. **Digital Crime Investigation phase**The goal is to collect and analyze the digital evidence that was obtained from the physical investigation phase and through any other future means. It includes similar phases as the Physical Investigation phases, although the primary focus is on the digital evidence. The six phases are:
   - Preservation phase; which preserves the digital crime scene so that evidence can later be synchronized and analyzed for further evidence.
   - Survey phase; whereby the investigator transfers the relevant data from a venue out of physical or administrative control of the investigator to a controlled location.
   - Documentation phase; which involves properly documenting the digital evidence when it is found. This information is helpful in the presentation phase.
   - Search and collection phase; whereby an in-depth analysis of the digital evidence is performed. Software tools are used to reveal hidden, deleted, swapped and corrupted files that were used including the dates, duration, log

file etc. Low-level time lining is performed to trace a user's activities and identity.

- Reconstruction phase; which includes putting the pieces of a digital puzzle together, and developing investigative hypotheses.
- Presentation phase; that involves presenting the digital evidence that was found to the physical investigative team.

It is noteworthy that this DFPM facilitates concurrent execution of physical and digital investigation.

5. **Review phase** this entails a review of the whole investigation and identifies areas of improvement. The IDIP model does well at illustrating the forensic process, and also conforms to the cyber terrorism capabilities which require a digital investigation to address issues of data protection, data acquisition, imaging, extraction, interrogation, ingestion/normalization, analysis and reporting. It also highlights the reconstruction of the events that led to the incident and emphasizes reviewing the whole task, hence ultimately building a mechanism for quicker forensic examinations.

III. An Extended Model of Cybercrime Investigation (EMCI)

The DFPM proposed by S. O. Ciardhuain- an Extended Model of Cybercrime Investigation (EMCI) - is more likely the most comprehensive till date.

- **Phases of EMCI:** The EMCI follows the waterfall model as every activity occurs in sequence. The sequence of examine, hypothesis, present, and prove/defend are bound to be repeated as the evidence heap increases during the investigation.

1. Awareness is the phase during which the investigators are informed that a crime has taken place; the crime is reported to some authority. An intrusion detection system may also trigger such awareness.
2. Authorization is the stage where the nature of investigation has been identified and the unplanned authorization may be required to proceed and the authorization is obtained internally or externally.
3. Planning is impacted by information from which and outside the organization that will affect the investigation. Internal factors are the organization policies, procedures, and former investigative knowledge while outside factors consist of legal and other requirements not known by the investigators.

Figure 4.3: An Extended Model of Cybercrime Investigation

4.2 Ethical issues in Digital Forensic

Ethics in the digital forensic field can be defined as a set of moral principles that regulate the use of computers. Ethical decision making in digital forensic work comprises of one or more of the following:

1. Honesty towards the investigation
2. Prudence means carefully handling the digital evidences
3. Compliance with the law and professional norms.

4.2.1 General ethical norms for investigator

Investigator should satisfy the following points:

1. Should contribute to the society and human being
2. Should avoid harm to others
3. Should be honest and trustworthy
4. Should be fair and take action not to discriminate
5. Should honor property rights, including copyrights and patents
6. Should give proper credit to intellectual property

7. Should respect the privacy of others
8. Should honor confidentiality

4.2.2 Unethical norms for Digital Forensic Investigation

Investigator should not:

1. Uphold any relevant evidence
2. Declare any confidential matters or knowledge
3. Express an opinion on the guilt or innocence belonging to any party
4. Engage or involve in any kind of unethical or illegal conduct
5. Deliberately or knowingly undertake an assignment beyond him or her capability
6. Distort or falsify education, training, credentials
7. Display bias or prejudice in findings or observation
8. Exceed or outpace authorization in conducting examination

4.3 Digital Evidences:

The field of computer security includes events that provide a successful courtroom experience, which are both worthwhile and satisfactory. Investigation of a computer security incident leads to legal proceeding, such as court proceeding, where the digital evidence and documents obtained are likely used as exhibits in the trial.

To meet the requirements of the judging body and to withstand or face any challenges, it is essential to follow the evidence-handling procedure. Also, it is necessary to ensure that the evidence-handling procedures chosen are not difficult to implement at your organization as this can sometimes become an overhead for an organization.

While investigating a computer security incident, we are sometimes unsure and indecisive whether an item(viz. a chip, floppy disk, etc)should be considered as an evidence or an attachment or an addendum.

Digital devices are everywhere in today's world, helping people communicate locally and globally with ease. Most people immediately think of computers, cell phones and the Internet as the only sources for digital evidence, but any piece of technology that processes information can be used in a criminal way. For example, hand-held games can carry encoded messages between criminals and even newer household appliances, such as a refrigerator with a built-in TV, could be used to store, view and share illegal images. The important thing to know is that responders need to be able to recognize and properly seize potential digital evidence.

4.3.1 Digital Evidences: (Electronic evidence)

- **Evidence:** Any information that can be confident or trusted and can prove something related to a case in trial that is, indicating that a certain substance or condition is present.
- **Relevant Evidence**: An information which has a positive impact on the action occurred, such as the information supporting an incident.

- **Digital Evidence:** Digital evidence is any information or data that can be confident or trusted and can prove something related to a case trial, that is, indicating that a certain substance or condition is present. It is safe to use to use such information as evidence during an investigation.

Digital evidence or **Electronic evidence** is any probative information stored or transmitted in digital form that a party to a court case may use at trial. Before accepting digital evidence a court will determine if the evidence is relevant, whether it is authentic, if it is hearsay and whether a copy is acceptable or the original is required.
Digital evidence is also defined as information and data of value to an investigation that is stored on, received or transmitted by an electronic device. This evidence can be acquired when electronic devices are seized and secured for examination. Digital evidence:

- Is latent (hidden), like fingerprints or DNA evidence
- Crosses jurisdictional borders quickly and easily
- Can be altered, damaged or destroyed with little effort
- Can be time sensitive

There are many sources of digital evidence; the topic is divided into three major forensic categories of devices where evidence can be found: Internet-based, stand-alone computers or devices, and mobile devices. These areas tend to have different evidence-gathering processes, tools and concerns, and different types of crimes tend to lend themselves to one device or the other.

Some of the popular electronic devices which are potential digital evidence are: HDD, CD/DVD media, backup tapes, USB drive, biometric scanner, digital camera, smart phone, smart card, PDA, etc.

4.3.2 Forms of digital evidence: Text message, emails, pictures, videos and internet searches are the most common types of Digital evidences.

The digital evidence is used to establish a credible link between the attacker, victim, and the crime scene. Some of the information stored in the victim's system can be potential digital evidence, such as IP address, system log-in & remote log-in details, browsing history, log files, emails, images, etc.

Digital Evidences may be in the form:

- Email Messages (may be deleted one also)
- Office file
- Deleted files of all kinds
- Encrypted file
- Compressed files
- Temp files
- Recycle Bin

- Web History
- Cache files
- Cookies
- Registry
- Unallocated Space
- Slack Space
- Web/E-Mail server access Logs
- Domain access Logs

4.3.3 Best Evidence Rule:

The original or true writing or recording must be confessed in court to prove its contents without any expectations. An original copy of the document is considered as superior evidence.

One of the rules states that if evidence is readable by sight or reflects the data accurately, such as any printout or data stored in a computer or similar devices or any other output, it is considered as "original".

It states that multiple copies of electronic files may be a part of the "original" or equivalent to the "original". The collected electronic evidence is mostly transferred to different media. Hence, many computer security professionals are dependent on this rule.

Best Evidence: The most complete copy or a copy which includes all necessary parts of evidence, which is closely related to the original evidence.

Example-A client has a copy of the original evidence media.

The "Best Evidence Rule" says that an original writing must be offered as evidence unless it is unavailable, in which case other evidence, like copies, notes, or other testimony can be used. Since the rules concerning evidence on a computer are fairly reasonable (what you can see on the monitor is what the computer contains, computer printouts are best evidence) computer records and records obtained from a computer are best evidence.

4.3.4 Original Evidence:

The procedure adopted to deal with a situation or case takes it outside the control of the client/victim. A case with proper diligence or a case with persistence work will end up in a judicial proceeding, and we will handle the evidences accordingly.

For this purpose original evidence as the truth or real (original) copy of the evidence media which is given by victim/client.

We define best incidence as the most complete copy, which includes all the necessary parts of the evidence that are closely related to the original evidence. It is also called as duplication of the evidence media. There should be an evidence protector which will

store either the best evidence or original evidence for every investigation in the evidence safe.

4.4 Characteristics of Digital Evidence:

Characteristics of digital evidences can help and challenge investigators during an investigation.The main goals in any investigation are to follow the trails that offenders leave during the commission of a crime and to tie perpetrators to the victims and crime scenes. Although witnesses may identify a suspect, tangible evidence of an individual's involvement is usually more compelling and reliable. Forensic analysts are employed to uncover compelling links between the offender, victim, and crime scene.

4.4.1 Locard's Exchange Principle:

According to Edmond Locard's principle, when two items make contact, there will be an interchange. The Locard principle is often cited in forensic sciences and is relevant in digital forensics investigations.

When an incident takes place, a criminal will leave a hint evidence at the scene and remove a hint evidence from the scene. This alteration is known as the Locard exchange principle. Many methods have been suggested in conventional forensic sciences to strongly prosecute criminals. Techniques used consists of blood analysis, DNA matching and fingerprint verification. These techniques are used to certify the existence of a suspected person at a physical scene. Based on this principle, Culley suggests that where there is communication with a computer system, clues will be left.

According to Locard's Exchange Principle, contact between two items will result in an exchange. This principle applies to any contact at a crime scene, including between an offender and victim, between a person with a weapon, and between people and the crime scene itself. In short, there will always be evidence of the interaction, although in some cases it may not be detected easily (note that absence of evidence is not evidence of absence). This transfer occurs in both the physical and digital realms and can provide links between them as depicted in Figure 1. In the physical world, an offender might inadvertently leave fingerprints or hair at the scene and take a fiber from the scene. For instance, in a homicide case the offender may attempt to misdirect investigators by creating a suicide note on the victim's computer, and in the process leave fingerprints on the keyboard. With one such piece of evidence, investigators can demonstrate the strong possibility that the offender was at the crime scene. With two pieces of evidence the link between the offender and crime scene becomes stronger and easier to demonstrate. Digital evidence can reveal communications between suspects and the victim, online activities at key times, and other information that provides a digital dimension to the investigation.

Figure 4.4: Evidence transfer in the physical and digital dimensions helps investigators establish connections between victims, offenders, and crime scenes.

In computer intrusions, the attackers will leave multiple traces of their presence throughout the environment, including in the fi le systems, registry, system logs, and network-level logs. Furthermore, the attackers could transfer elements of the crime scene back with them, such as stolen user passwords or PII in a file or database. Such evidence can be useful to link an individual to an intrusion.

In an e-mail harassment case, the act of sending threatening messages via a Web-based e-mail service such as Hotmail can leave a number of traces. The Web browser used to send messages will store fi les, links, and other information on the sender's hard drive along with date-time–related information. Therefore, forensic analysts may find an abundance of information relating to the sent message on the offender's hard drive, including the original message contents. Additionally, investigators may be able to obtain related information from Hotmail, including Web server access logs, IP addresses, and possibly the entire message in the sent mail folder of the offender's e-mail account.

4.4.2 Digital Stream of Bits :

Cohen refers to digital evidence as a bag of bits, which in turn can be arranged in arrays to display the information. The information in continuous bits will rarely make scene and tools are needed to show these structures logically so that it is readable.

The circumstance in which digital evidence are found also helps the investigator during the inspection. Metadata is used to portray data more specifically and is helpful in determining the background of digital evidence.

4.5 Types of Evidences:

There are many types of Evidences, each with their own specific or unique characteristics. Some major types of evidences are :

1. Illustrative evidence: Illustrative evidence is also called as demonstrative evidence. It is generally a representation of an object which is a common form of proof. For example , photographs , videos , sound recordings , X-rays , maps , drawing , graphs , charts , simulations , sculptors , and models.

2. Electronic Evidence: Electronic evidence is nothing but digital evidence. As we know, the use of digital evidence in trials has greatly increased .The evidences or proof that can be obtained from the electronic source is called the digital evidence.(viz. Email , hard drives etc.)

3. Documented Evidence: Documented evidence is same as demonstrative evidence. However, in documentary evidence , the proof is presented in writing (Viz. Contracts , wills , invoices etc.).

4. Explainable Evidence: This type of evidence is typically used in criminal cases in which it supports the dependent, either partially or totally removing their guilt in the case.It is also referred to as exculpatory.

5.Substantial Evidence: A proof that is introduced in the form of a physical object, whether whole or in part, is referred to as substantial evidence. It is also called physical evidence. Such evidence might consist of dried blood, fingerprint, and DNA samples, casts of footprints or tries at the scene of crime.

6. Testimonial: It is the kind of evidence spoken by the spectator under the oath , or written evidence given under the oath by an official declaration that is affidavit. This is the common forms of evidence in the system.

4.6 Challenges in Evidence handling:

While responding to a computer security incident, a failure to adequately document is one of the most common mistakes made by computer security professional’s .Analytical data might never be collected, critical data may be lost or data's origin or meaning may become unknown. As there are many evidences collected based on technical complexity is the fact that the properly retrieved evidence requires a paper trial.

Such documentations give an impression of having a certain quality against the natural instincts of the technical practical knowledge of individuals, who often investigate computer security incidents.

The challenges faced in the evidence handling must be properly understood by all the investigators. They should also understand how to meet these challenges. Therefore, it is essential for every organization to have formal evidence handling procedures that support computer security investigation. The most difficult task for an evidence handler is to substantiate the collected evidence at the judicial proceedings. Maintaining the chain of custody is also necessary. You must have both power and skill to validate your evidence.

4.6.1 Authentication of Evidence:

The laws of many state jurisdictions define data as Written Works and Record keeping. Before introducing them as evidence, documents and recorded material must be authenticated.

The evidence that are collected by any person/investigator should be collected using authenticate methods and techniques because during court proceedings these will become major evidences to prove the crime. In other words, for providing a piece of evidence of the testimony, it is necessary to have an authenticated evidence by a spectator who has a personal knowledge to its origin.

For an evidence to be admissible, it is necessary that it should be authenticated, otherwise the information cannot be presented to judging only. The matter of record is that the evidence collected by any person should meet the demand of authentication. The evidence collected must have some sort of internal documentation that records the manner of collected information.

4.6.2 Chain of Custody:

What Is the Chain of Custody in Computer Forensics?

The chain of custody in digital forensics can also be referred to as the forensic link, the paper trail, or the chronological documentation of electronic evidence. It indicates the collection, sequence of control, transfer, and analysis. It also documents each person who handled the evidence, the date/time it was collected or transferred, and the purpose for the transfer.

Why Is It Important to Maintain the Chain of Custody?

It is important to maintain the chain of custody to preserve the integrity of the evidence and prevent it from contamination, which can alter the state of the evidence. If not preserved, the evidence presented in court might be challenged and ruled inadmissible.

Importance to the Examiner:

Suppose that, as the examiner, you obtain metadata for a piece of evidence. However, you are unable to extract meaningful information from it. The fact that there is no meaningful information within the metadata does not mean that the evidence is insufficient. The chain of custody in this case helps show where the possible evidence might lie, where it came from, who created it, and the type of equipment that was used. That way, if you want to create an exemplar, you can get that equipment, create the exemplar, and compare it to the evidence to confirm the evidence properties.

Importance to the Court:

It is possible to have the evidence presented in court dismissed if there is a missing link

in the chain of custody. It is therefore important to ensure that a wholesome and meaningful chain of custody is presented along with the evidence at the court.

What Is the Procedure to Establish the Chain of Custody?

In order to ensure that the chain of custody is as authentic as possible, a series of steps must be followed. It is important to note that, the more information a forensic expert obtains concerning the evidence at hand, the more authentic is the created chain of custody. Due to this, it is important to obtain administrator information about the evidence: for instance, the administrative log, date and file info, and who accessed the files. You should ensure the following procedure is followed according to the chain of custody for electronic evidence:

- **Save the original materials:** You should always work on copies of the digital evidence as opposed to the original. This ensures that you are able to compare your work products to the original that you preserved unmodified.
- **Take photos of physical evidence**: Photos of physical (electronic) evidence establish the chain of custody and make it more authentic.
- **Take screenshots of digital evidence content:** In cases where the evidence is intangible, taking screenshots is an effective way of establishing the chain of custody.
- **Document date, time, and any other information of receipt**. Recording the timestamps of whoever has had the evidence allows investigators to build a reliable timeline of where the evidence was prior to being obtained. In the event that there is a hole in the timeline, further investigation may be necessary.
- **Inject a bit-for-bit clone of digital evidence content into our forensic computers.** This ensures that we obtain a complete duplicate of the digital evidence in question.
- **Perform a hash test analysis to further authenticate the working clone**. Performing a hash test ensures that the data we obtain from the previous bit-by-bit copy procedure is not corrupt and reflects the true nature of the original evidence. If this is not the case, then the forensic analysis may be flawed and may result in problems, thus rendering the copy non-authentic.

The procedure of the chain of custody might be different. depending on the jurisdiction in which the evidence resides; however, the steps are largely identical to the ones outlined above.

What Considerations Are Involved with Digital Evidence?

A couple of considerations are involved when dealing with digital evidence. We shall take a look at the most common and discuss globally accepted best practices.

1. **Never work with the original evidence to develop procedures:** The biggest consideration with digital evidence is that the forensic expert has to make a complete copy of the evidence for forensic analysis. This cannot be overlooked because, when errors are made to working copies or comparisons are required, it

will be necessary to compare the original and copies.

2. **Use clean collecting media:** It is important to ensure that the examiner's storage device is forensically clean when acquiring the evidence. This prevents the original copies from damage. Think of a situation where the examiner's data evidence collecting media is infected by malware. If the malware escapes into the machine being examined, all of the evidence can become compromised.
3. **Document any extra scope:** During the course of an examination, information of evidentiary value may be found that is beyond the scope of the current legal authority. It is recommended that this information be documented and brought to the attention of the case agent because the information may be needed to obtain additional search authorities. A comprehensive report must contain the following sections:
    - Identity of the reporting agency
    - Case identifier or submission number
    - Case investigator
    - Identity of the submitter
    - Date of receipt
    - Date of report
    - Descriptive list of items submitted for examination, including serial number, make, and model
    - Identity and signature of the examiner
    - Brief description of steps taken during examination, such as string searches, graphics image searches, and recovering erased files
    - Results/conclusions
4. **Consider safety of personnel at the scene**. It is advisable to always ensure the scene is properly secured before and during the search. In some cases, the examiner may only have the opportunity to do the following while onsite:
    - Identify the number and type of computers.
    - Determine if a network is present.
    - Interview the system administrator and users.
    - Identify and document the types and volume of media, including removable media.
    - Document the location from which the media was removed.
    - Identify offsite storage areas and/or remote computing locations.
    - Identify proprietary software.
    - Determine the operating system in question.

The considerations above need to be taken into account when dealing with digital evidence due to the fragile nature of the task at hand.

Chain of custody prevents evidence from being tainted; it thus establishes trustworthiness of items brought into evidence. The U.S. legal system wants the

proponent of evidence to be able to demonstrate an unbroken chain of custody for items he wants to have admitted.

Often, there is a stipulation,for example, when there is an agreement between the parties or a concession by the opponent of the evidence that allows it to be admitted without requiring testimony to prove the foundational elements. The purpose of stipulation is to move the trial quickly forward, without pondering idle questions.

If there is a break in the chain of custody brought to the attention of the court, then the court has to decide whether the breach is so severe as to meet exclusion of the item from trial. Alternatively, the court can decide that the Trier (trial judge or jury) need to decide the value of the evidence. To prevent a breach, a forensic investigation should follow a written policy, so that necessary deviations of the policy can be argued. The policy itself should take all reasonable (or arguably reasonable) precautions against tampering.

For example, assume that a PDA is seized from a suspected drug dealer. In the case of an PDA, there is no hard drive image to mirror, that is, the examination will have to be done on the powered-on original. The PDA can lose data, for example by disconnecting it from its battery. On seizure, the device should not be switched on. If it is seized switched on, it should be switched off in order to preserve battery power. It needs to be put into an evidence bag that does not allow access to the PDA without breaking the seal (no clear plastic bag!). The evidence needs to be tagged with all pertinent data, including the serial number of the PDA and the circumstances of the seizure. The PDA should never be returned to the accused at the scene, because the device can lose data if reset. To maintain the data in the PDA, it needs to be kept in a continuously charged mode. It should only be used to extract evidence by a competent person who can testify in court. As long as the PDA could be evidence, it needs to be kept in an evidence locker, with check-out logs, so that it can be determined who had access to the PDA at any time.

4.6.3 Evidence Validation: The challenge is to ensure that providing or obtaining the data that you have collected is similar to the data provided or presented in court. Several years pass between the collection of evidence and the production of evidence at a judiciary proceeding, which is very common. To meet the challenge of validation, it is necessary to ensure that the original media matches the forensic duplication by using MD5 hashes. The evidence for every file is nothing but the MD5 hash values that are generated for every file that contributes to the case.

The verify function within the Encase application can be used while duplicating a hard drive with Encase. To perform a forensic duplication using dd , you must record MD5 hash for both the original evidence media and binary files or the files which compose the forensic duplication.

Note: Evidence collection calculated by MD5 after 6 months may not be helpful.MD5 hashes should be performed when the evidence is obtained.

4.7 Volatile Evidence: Not all the evidence on a system is going to last very long. Some evidence is residing in storage that requires a consistent power supply; other evidence may be stored in information that is continuously changing. When collecting evidence, you should always try to proceed from the most volatile to the least. Of course, you should still take the individual circumstances into account—you shouldn't waste time extracting information from an unimportant/unaffected machine's main memory when an important or affected machine's secondary memory hasn't been examined.

You need to respond to the target system at the console during the collection of volatile data rather than access it over the network. This way the possibility of the attacker monitoring your responses is eliminated, ensuring that you are running trust commands. If you are creating a forensic duplication of the targeted system, you should focus on obtaining the volatile system data before shutting down the system.

To determine what evidence to collect first, you should draw up an Order of Volatility—a list of evidence sources ordered by relative volatility. An example an Order of Volatility would be:

1. Registers and cache
2. Routing tables
3. Arp cache
4. Process table
5. Kernel statistics and modules
6. Main memory
7. Temporary file systems
8. Secondary memory
9. Router configuration
10. Network topology

Note: Once you have collected the raw data from volatile sources you may be able to shutdown the system.{Matthew Braid, "Collecting Electronic Evidence After A System Compromise," Australian Computer Emergency Response Team}

Registers, Cache: The contents of CPU cache and registers are extremely volatile, since they are changing all of the time. Literally, nanoseconds make the difference here. An examiner needs to get to the cache and register immediately and extract that evidence before it is lost.

Routing Table, ARP Cache, Process Table, Kernel Statistics, Memory: Some of these items, like the routing table and the process table, have data located on network devices. In other words, that data can change quickly while the system is in operation,

so evidence must be gathered quickly. Also, kernel statistics are moving back and forth between cache and main memory, which make them highly volatile. Finally, the information located on random access memory (RAM) can be lost if there is a power spike or if power goes out. Clearly, that information must be obtained quickly.

Temporary File Systems: Even though the contents of temporary file systems have the potential to become an important part of future legal proceedings, the volatility concern is not as high here. Temporary file systems usually stick around for awhile.

Disk: Even though we think that the data we place on a disk will be around forever, that is not always the case (see the SSD Forensic Analysis post from June 21). However, the likelihood that data on a disk cannot be extracted is very low.

Remote Logging and Monitoring Data that is Relevant to the System in Question: The potential for remote logging and monitoring data to change is much higher than data on a hard drive, but the information is not as vital. So, even though the volatility of the data is higher here, we still want that hard drive data first.

Physical Configuration, Network Topology, and Archival Media: Here we have items that are either not that vital in terms of the data or are not at all volatile. The physical configuration and network topology is information that could help an investigation, but is likely not going to have a tremendous impact. Finally, archived data is usually going to be located on a DVD or tape, so it isn't going anywhere anytime soon. It is great digital evidence to gather, but it is not volatile.

Case Studies :

Case-1: Credit Card Fraud

State : Tamil Nadu

City : Chennai

Sections of Law : Section of Law: 66 of Information Technology Act 2000 & 120(B), 420,467,468,471 IPC.

Background:

The assistant manager (the complainant) with the fraud control unit of a large business process outsourcing (BPO) organization filed a complaint alleging that two of its employees had conspired with a credit card holder to manipulate the credit limit and as a result cheated the company of INR 0.72 million.

The BPO facility had about 350 employees. Their primary function was to issue the

bank's credit cards as well as attend to customer and merchant queries. Each employee was assigned to a specific task and was only allowed to access the computer system for that specific task. The employees were not allowed to make any changes in the credit-card holder's account unless they received specific approvals.

Each of the employees was given a unique individual password. In case they entered an incorrect password three consecutive times then their password would get blocked and they would be issued a temporary password.

The company suspected that its employees conspired with the son (holding an add-on card) of one of the credit card holders. The modus operandi suspected by the client is as follows.

The BPO employee deliberately keyed in the wrong password three consecutive times (so that his password would get blocked) and obtained a temporary password to access the computer system. He manually reversed the transactions of the card so that it appeared that payment for the transaction has taken place. The suspect also changed the credit card holder's address so that the statement of account would never be delivered to the primary card holder.

Investigation: A procedure to find the Digital Evidence

The investigating team visited the premises of the BPO and conducted detailed examination of various persons to understand the computer system used. They learnt that in certain situations the system allowed the user to increase the financial limits placed on a credit card. The system also allowed the user to change the customer's address, blocking and unblocking of the address, authorisations for cash transactions etc.

The team analysed the attendance register which showed that the accused was present at all the times when the fraudulent entries had been entered in the system. They also analysed the system logs that showed that the accuser's ID had been used to make the changes in the system.

The team also visited the merchant establishments from where some of the transactions had taken place. The owners of these establishments identified the holder of the add-on card.

Current status: The BPO was informed of the security lapse in the software utilised. Armed with this evidence the investigating team arrested all the accused and recovered, on their confession, six mobile phones, costly imported wrist watches, jewels, electronic items, leather accessories, credit cards, all worth INR 0. 3 million and cash INR 25000.

The investigating team informed the company of the security lapses in their software so that instances like this could be avoided in the future.

This case won the second runner-up position for the India Cyber Cop Award, for its investigating officer Mr S. Balu, Assistant Commissioner of Police, Crime, Chennai Police. The case was remarkable for the excellent understanding displayed by the investigating team, of the business processes and its use in collecting digital evidence.

Case-2: Hosting Obscene Profiles

State	:	Tamil Nadu
City	:	Chennai
Sections of Law	:	67 of Information Technology Act 2000 469, 509 of the Indian Penal code

Background: The complainant stated that some unknown person had created an e-mail ID using her name and had used this ID to post messages on five Web pages describing her as a call-girl along with her contact numbers.
As a result she started receiving a lot of offending calls from men.

Investigation: A procedure to find the Digital Evidence
After the complainant heard about the Web pages with her contact details, she created a username to access and view these pages.

Using the same log-in details, the investigating team accessed the Web pages where these profiles were uploaded. The message had been posted on five groups, one of which was a public group. The investigating team obtained the access logs of the public group and the message to identify the IP addresses used to post the message. Two IP addresses were identified.

The ISP was identified with the help of publicly available Internet sites. A request was made to the ISPs to provide the details of the computer with the IP addresses at the time the messages were posted. They provided the names and addresses of two cyber cafes located in Mumbai to the police.

The investigating team scrutinised the registers maintained by the cyber cafes and found that in one case the complainant's name had been signed into the register.

The team also cross-examined the complainant in great detail. During one of the meetings she revealed that she had refused a former college mate who had proposed marriage.

In view of the above the former college mate became the prime suspect. Using this information the investigating team, with the help of Mumbai police, arrested the suspect and seized a mobile phone from him. After the forensic examination of the SIM card and the phone, it was observed that phone had the complainant's telephone number that was posted on the internet. The owner of the cyber cafes also identified the suspect as the one who had visited the cyber cafes.

Based on the facts available with the police and the sustained interrogation the suspect confessed to the crime.

Current status: The suspect was convicted of the crime and sentenced to two years of imprisonment as well as a fine.

Case - 3: Illegal money transfer

State : Maharashtra

City : Pune

Sections of Law : 467,468, 471, 379,419, 420, 34 of IPC & 66 of IT ACT

Background: The accused in the case were working in a BPO, that was handling the business of a multinational bank. The accused, during the course of their work had obtained the personal identification numbers (PIN) and other confidential information of the bank's customers. Using these the accused and their accomplices, through different cyber cafes, transferred huge sums of money from the accounts of different customers to fake accounts.

Investigation: A procedure to find the Digital Evidence

On receiving the complaint the entire business process of the complainant firm was studied and a systems analysis was conducted to establish the possible source of the data theft.

The investigators were successful in arresting two people as they laid a trap in a local bank where the accused had fake accounts for illegally transferring money.

During the investigation the system server logs of the BPO were collected. The IP addresses were traced to the Internet service provider and ultimately to the cyber cafes through which illegal transfers were made.

The registers maintained in cyber cafes and the owners of cyber cafes assisted in identifying the other accused in the case. The e-mail IDs and phone call print outs were also procured and studied to establish the identity of the accused. The e-mail accounts

of the arrested accused were scanned which revealed vital information to identify the other accused. Some e-mail accounts of the accused contained swift codes, which were required for internet money transfer.

All the 17 accused in the case were arrested in a short span of time. The charge sheet was submitted in the court within the stipulated time. In the entire wire transfer scam, an amount to the tune of about INR 19 million was transferred, out of this INR 9 million was blocked in transit due to timely intimation by police, INR 2 million was held in balance in one of the bank accounts opened by the accused which was frozen. In addition the police recovered cash, ornaments, vehicles and other articles amounting to INR 3 million.

During the investigation the investigating officer learned the process of wire transfer, the banking procedures and weakness in the system. The investigating officer suggested measures to rectify the weakness in the present security systems of the call centre. This has helped the local BPO industry in taking appropriate security measures.

Current status: Pending trial in the court.
This case won the India Cyber Cop Award, for its investigating officer Mr Sanjay Jadhav, Assistant Commissioner of Police, Crime, Pune Police. The panel of judges felt that this case was the most significant one for the Indian IT industry during 2005 and was investigated in a professional manner, with substantial portion of the swindled funds being immobilised, a large number of persons were arrested and the case was sent to the court for trial within 90 days.

Case-4: Fake Travel Agent

State	: Maharashtra
City	: Mumbai
Sections of Law	: 420, 465, 467, 468, 471, 34 of IPC r/w 143 of Indian Railway Act 1989.

Background: The accused in this case was posing to be a genuine railway ticket agent and had been purchasing tickets online by using stolen credit cards of non residents. The accused created fraudulent electronic records/ profiles, which he used to carry out the transactions.The tickets so purchased were sold for cash to other passengers. Such events occurred for a period of about four months.

The online ticket booking service provider took notice of this and lodged a complaint with the cyber crime investigation cell.

Investigation: A procedure to find the Digital Evidence

The service provider gave the IP addresses, which were used for the fraudulent online bookings, to the investigating team. IP addresses were traced to cyber cafes in two locations.

The investigating team visited the cyber cafŽs but was not able to get the desired logs as they were not maintained by the cyber cafŽ owners. The investigating team was able to short list the persons present at cyber cafes when the bookings were made. The respective owners of the cyber cafes were able to identify two persons who would regularly book railway tickets.

The investigating team then examined the passengers who had travelled on these tickets. They stated that they had received the tickets from the accused and identified the delivery boy who delivered the tickets to them. On the basis of this evidence the investigating team arrested two persons who were identified in an identification parade.

Current status: The charge sheet has been submitted in the court.

Case-5: Creating Fake Profile

State	: Andhra Pradesh
City	: Hyderabad
Sections of Law	: 67 Information Technology Act 2000 507, 509 of the Indian Penal Code

Background: The complainant received an obscene e-mail from an unknown e-mail ID. The complainant also noticed that obscene profiles along with photographs of his daughter had been uploaded on matrimonial sites.

Investigation: A procedure to find the Digital Evidence

The investigating officer examined and recorded the statements of the complainant and his daughter. The complainant stated that his daughter was divorced and her husband had developed a grudge against them due to the failure of the marriage.

The investigating officer took the original e-mail from the complainant and extracted the IP address of the same. From the IP address he could ascertain the Internet service provider.

The IP address was traced to a cable Internet service provider in the city area of Hyderabad. The said IP address was allotted to the former husband sometime back and his house was traced with the help of the staff of ISP.

A search warrant was obtained and the house of the accused was searched. During the search operation, a desktop computer and a handicam were seized from the premises. A forensic IT specialist assisted the investigation officer in recovering e-mails (which

were sent to the complainant), using a specialised disk search tool as well as photographs (which had been posted on the Internet) from the computer and the handicam respectively. The seized computer and the handicam were sent to the forensic security laboratory for further analysis.

The experts of the forensic security laboratory analysed the material and issued a report stating that: the hard disk of the seized computer contained text that was identical to that of the obscene e-mail; the computer had been used to access the matrimonial websites on which the obscene profiles were posted; the computer had been used to access the e-mail account that was used to send the obscene e-mail; the handicam seized from the accused contained images identical to the ones posted on the matrimonial Websites. Based on the report of the FSL it was clearly established that the accused had: created a fictitious e-mail ID and had sent the obscene e-mail to the complainant; posted the profiles of the victim along with her photographs on the matrimonial sites.

Current status: Based on the material and oral evidence, a charge sheet has been filed against the accused and the case is currently pending for trial.

References

1. http://www.forensicsciencesimplified.org/digital/
2. http://www.forensicsciencesimplified.org/digital/
3. https://www.helpnetsecurity.com/2007/07/20/the-rules-for-computer-forensics/ as on 28 August 2019
4. Digital Evidence and Computer Crime, Third Edition © 2011 Eoghan Casey. Published by Elsevier Inc.
5. www.cse.scu.edu/~tschwarz/COEN252_13/LN/legalissues.html

Sample Multiple Choice Questions

1. Digital forensics is all of them except:
   a) Extraction of computer data
   b) Preservation of computer data
   c) Interpretation of computer data
   d) Manipulation of computer data

2. IDIP stands for
   a) Integrated Digital Investigation Process
   b) Integrated Data Investigation Process
   c) Integrated Digital Investigator Process
   d) Independent Digital Investigator Process

3. The digital evidence are used to establish a credible link between____
    a. Attacker and victim and the crime scene
    b. Attacker and the crime scene
    c. victim and the crime scene
    d. Attacker and Information

4. Digital evidences must follow the requirements of the _____.
    a. Ideal Evidence rule
    b. Best Evidence Rule
    c. Exchange Rule
    d. All of the mentioned

5. From the two given statements 1 and 2 , select the correct options from a-d.
   1): Original media can be used to carry out digital investigation process.
   2): By default, every part of the victim's computer is considered unreliable.
    a. a and b both are true
    b. a is true and b is false
    c. a and b both are false
    d. a is false and b is true

6. The evidences or proof that can be obtained from the electronic source is called the_____
    a. digital evidence
    b. demonstrative evidence
    c. Explainable Evidence
    d. substantial evidence

7. Which of the following is not a type of volatile evidence?
    a. Routing Tables
    b. Main Memory
    c. Log files
    d. Cached Data

Unit-5 Basics of Hacking

Contents

5.1 Ethical Hacking
- How Hackers Beget Ethical Hackers
- Defining hacker, Malicious users
- Data Privacy and General Data Protection and Regulation(GDPR)

5.2 Understanding the need to hack your own systems

5.3 Understanding the dangers your systems face
- Nontechnical attacks
- Network-infrastructure attacks
- Operating-system attacks
- Application and other specialized attacks

5.4 Obeying the Ethical hacking Principles
- Working ethically
- Respecting privacy
- Not crashing your systems

5.5 The Ethical hacking Process
- Formulating your plan
- Selecting tools
- Executing the plan
- Evaluating results
- Moving on

5.6 Cyber Security act

5.1 Ethical Hacking

History:

Hacking developed alongside "Phone Phreaking", a term referred to exploration of the phone network without authorization, and there has often been overlap between both technology and participants.

Ethical hacking is the science of testing computers and network for security vulnerabilities and plugging the holes found before the unauthorized people get a chance to exploit them.

Social Engineering Cycle **Social Engineering Counter Measures**

Fig. 5.1 Social Engineering Cycle and Counter Measures

- **Gather Information**: This is the first stage, the learns as much as he can about the intended victim. The information is gathered from company websites, other publications and sometimes by talking to the users of the target system.
- **Plan Attack**: The attackers outline how he/she intends to execute the attack
- **Acquire Tools**: These include computer programs that an attacker will use when launching the attack.
- **Attack**: Exploit the weaknesses in the target system.
- **Use acquired knowledge**: Information gathered during the social engineering tactics such as pet names, birthdates of the organization founders, etc. is used in attacks such as password guessing.

Most techniques employed by social engineers involve manipulating human biases. To counter such techniques, an organization can;

- ✓ **To counter the familiarity exploit**
- ✓ **To counter intimidating circumstances attacks**
- ✓ **To counter phishing techniques**
- ✓ **To counter tailgating attacks**
- ✓ **To counter human curiosity**
- ✓ **To counter techniques that exploit human greed**

Summary

- Social engineering is the art of exploiting the human elements to gain access to un-authorized resources.
- Social engineers use a number of techniques to fool the users into revealing sensitive information.
- Organizations must have security policies that have social engineering countermeasures.

Hacker's attitude:

A hacker-cracker separation gives more emphasis to a range of different categories, such as white hat (ethical hacking), grey hat, black hat and script kiddies. The term cracker refers to black hat hackers, or more generally hackers with unlawful intentions.

Hackers are problem solvers. They get extract from understanding a problem and sorting out a solution. Their motivation to meet challenges is internal. Hackers do what

they do because it's extremely satisfying to solve puzzles and fix the up-until-now unfixable. The pleasure derived is both intellectual and practical. But one doesn't have to be a geek to be a hacker. Being a hacker is a mind-set. In Raymond's dissertation, "How to Become a Hacker", he describes the fundamentals of a hacker attitude.

These are very same principles apply to being innovative which are explained as below:

The world is full of fascinating problems waiting to be solved.

Innovation happens because hackers like to solve the problem rather than complaining. If one happen to find these problems fascinating and exciting, then it won't even feel like hard work.

No Problem should ever have to be solved twice.

Hackers are perfectionists for clarifying the problem before they start generating ideas. It's easy to jump to solutions, but sometimes that means wrong problems are solved. A little bit of accuracy on the front end of a problem solving process means one tackles the right and real problem, so one only have to do it once.

Boredom and drudgery(more and more work) are evil.

The best way to lose touch with innovation is to become too repetitive. Innovation requires constant and vigilant creativity. It may not be broken enough to fix, but there's no reason not to squeeze it and cut boredom off at the pass.

Freedom is good.

Hackers need freedom to work upon their ideas.

Attitude is no substitute for competence.

They are open-minded and they see problems as interesting opportunities. Innovators are seeking to understand a problem more deeply, puzzling at how an unworkable idea might become workable, increasing their skill set so that they are better problem solvers and can better execute their ideas.

Hackers are the innovators of the Internet. Overall hackers are who have got that relentless, curious, problem-solving attitude.

Computer Hacking:

Computer Hackers have been in existence for more than a century. Originally, "hacker" did not carry the negative implications. In the late 1950s and early 1960s, computers were much different than the desktop or laptop systems most people are familiar with. In those days, most companies and universities used mainframe computers: giant, slow-moving hunks of metal locked away in temperature-controlled glass cages. It cost thousands of dollars to maintain and operate those machines, and programmers had to fight for access time.

Because of the time and money involved, computer programmers began looking for ways to get the most out of the machines. The best and brightest of those programmers created what they called "hacks" - shortcuts that would modify and improve the performance of a computer's operating system or applications and allow more tasks to be completed in a shorter time.

Still, for all the negative things hackers have done, they provide a necessary (and even valuable) service, which is elaborated on after a brief timeline in the history of computer hacking

5.1.1 How Hackers Beget Ethical Hackers:

Hacker is a word that has two meanings:

- ✓ Traditionally, a hacker is someone who likes to tamper with software or electronic systems. Hackers enjoy exploring and learning how computer systems operate. They love discovering new ways to work electronically.
- ✓ Recently, hacker has taken on a new meaning — someone who maliciously breaks into systems for personal gain. Technically, these criminals are crackers (criminal hackers). Crackers break into (crack) systems with malicious intent. They are out for personal gain: fame, profit, and even revenge. They modify, delete, and steal critical information, often making other people miserable.

The good-guy (white-hat) hackers don't like being in the same category as the bad-guy (black-hat) hackers. Whatever the case, most people give hacker a negative meaning. Many malicious hackers claim that they don't cause damage but instead are selflessly helping others. In other words, many malicious hackers are electronic thieves.

Hackers go for almost any system they think they can compromise. Some prefer prestigious, well-protected systems, but hacking into anyone's system increases their status in hacker circles.

If one need protection from hacker troubles; one has to become as savvy as the guys trying to attack systems. A true security assessment professional possesses the skills, mind-set, and tools of a hacker but is also trustworthy. He or she performs the hacks as security tests against systems based on how hackers might work.

- **Ethical hacker's attitude** encompasses formal and methodical penetration testing, white hat hacking, and vulnerability testing ,which involves the same tools, tricks, and techniques that criminal hackers use, but with one major difference: Ethical hacking is performed with the target's permission in a professional setting.

The intent of ethical hacking is to discover vulnerabilities from a malicious attacker's viewpoint to better secure systems. Ethical hacking is part of an overall information risk management program that allows for on-going security improvements. Ethical hacking can also ensure that vendors' claims about the security of their products are genuine.

- **Ethical hacking versus auditing**

Many people confuse security testing via the ethical hacking approach with security auditing, but there are big differences, namely in the objectives.

Security auditing involves comparing a company's security policies (or compliance requirements) to what's actually taking place. The intent of security auditing is to validate that security controls exist using a risk-based approach.

Auditing often involves reviewing business processes and, in many cases, might not be very technical. Security audits are usually based on checklists.

Equally, security assessments based around ethical hacking focus on vulnerabilities that can be exploited. This testing approach validates that security controls do not exist or are incompetent at best.

Ethical hacking can be both highly technical and nontechnical, and although one can use a formal methodology, it tends to be a bit less structured than formal auditing.

- **Policy considerations**

If it is chosen to make ethical hacking an important part of business's information risk management program, one really need to have a documented security testing policy. Such a policy outlines who's doing the testing, the general type of testing that is performed, and how often the testing takes place.

- **What is Hacking?**

Hacking is identifying weakness in computer systems or networks to exploit its weaknesses to gain access.

- **Example of Hacking:**

Computers have become mandatory to run successful businesses. It is not enough to have isolated computers systems; they need to be networked to facilitate communication with external businesses.

- ✓ Using password cracking algorithm to gain access to a system.
- ✓ This exposes them to the outside world and hacking. Hacking means using computers to commit fraudulent acts such as fraud, privacy invasion, stealing corporate/personal data, etc.
- ✓ Cybercrimes cost many organizations millions of dollars every year. Businesses need to protect themselves against such attacks.

Ethical Hacking is identifying weakness in computer systems and/or computer networks and coming up with countermeasures that protect the weaknesses.

Ethical hacking is a branch of information security or information assurance which tests an organization's information systems against a variety of attacks. Ethical hackers are also sometimes known as **White Hats.**

Many people are confused when the terms "Ethical" and "Hacking" are used together. Usually the term "hacker" has a negative connotation due to media reports using incorrect terminology.

Ethical hackers must abide by the following rules:

- ✓ Get **written permission** from the owner of the computer system and/or computer network before hacking.

- ✓ **Protect the privacy of the organization** been hacked.
- ✓ **Transparently report** all the identified weaknesses in the computer system to the organization.
- ✓ **Inform** hardware and software vendors of the **identified weaknesses**.

Definition:

Ethical hacking

- ✓ Refers to the act of locating weaknesses and vulnerabilities of computer and information systems by duplicating the intent and actions of malicious hackers.
- ✓ Known as **penetration testing, intrusion testing** or **red teaming.**

An ethical hacker is a security professional who applies their hacking skills for defensive purposes on behalf of the owners of information systems.

By conducting penetration tests, an ethical hacker looks to answer the following four basic questions:

1. What information/locations/systems can an attacker gain access?
2. What can an attacker see on the target?
3. What can an attacker do with available information?
4. Does anyone at the target system notice the attempts?

An ethical hacker operates with the knowledge and permission of the organization for which they are trying to defend. In some cases, the organization will neglect to inform their information security team of the activities that will be carried out by an ethical hacker in an attempt to test the effectiveness of the information security team. This is referred to as a double-blind environment. In order to operate effectively and legally, an ethical hacker must be informed of the assets that should be protected, potential threat sources, and the extent to which the organization will support an ethical hacker's efforts.

5.1.2 Defining hacker, Malicious users:

Definition of Hacker: A **Hacker** is a person who finds and exploits the weakness in computer systems and/or networks to gain access. Hackers are usually skilled computer programmers with knowledge of computer security.

An **Ethical Hacker**, also known as a whitehat hacker, or simply a whitehat, is a security professional who applies their hacking skills for defensive purposes on behalf of the owners of information systems.

Nowadays, certified ethical hackers are among the most sought after informationsecurity employees in large organizations such as Wipro, Infosys, IBM, Airtel and Reliance among others.

What Is a Malicious User?

Malicious users (or internal attackers) try to compromise computers and sensitive information from the inside as authorized and “trusted” users. Malicious users go for systems they believe they can compromise for fraudulent gains or revenge.

- ✓ Malicious attackers are, generally known as both, hackers and malicious users.
- ✓ Malicious user means a rogue employee, contractor, intern, or other user who abuses his or her trusted privileges .It is a common term in security circles.

Users search through critical database systems to collect sensitive information, e-mail confidential client information to the competition or elsewhere to the cloud, or delete sensitive files from servers that they probably do not have access.
There's also the occasional ignorant insider whose intent is not malicious but who still causes security problems by moving, deleting, or corrupting sensitive information. Even an innocent "fat-finger" on the keyboard can have terrible consequences in the business world.

Malicious users are often the worst enemies of IT and information security professionals because they know exactly where to go to get the goods and don't need to be computer savvy to compromise sensitive information. These users have the access they need and the management trusts them, often without question. In short they take the undue advantage of the trust of the management.
Hackers are classified according to the intent of their actions.

The following list classifies hackers according to their intent.

Symbol	Description
	Ethical Hacker (White hat): A hacker who gains access to systems with a view to fix the identified weaknesses. They may also perform penetration Testing and vulnerability assessments.
	Cracker (Black hat): A hacker who gains unauthorized access to computer systems for personal gain. The intent is usually to steal corporate data, violate privacy rights, transfer funds from bank accounts etc.
	Grey hat: A hacker who is in between ethical and black hat hackers. He/she breaks into computer systems without authority with a view to identify weaknesses and reveal them to the system owner.

Symbol	Description
	Script kiddies: A non-skilled person who gains access to computer systems using already made tools.
	Hacktivist: A hacker who use hacking to send social, religious, and political, etc. messages. This is usually done by hijacking websites and leaving the message on the hijacked website.
	Phreaker: A hacker who identifies and exploits weaknesses in telephones instead of computers.

Figure 5.2: classification of hackers according to their intent

Why Ethical Hacking?

- Information is one of the most valuable assets of an organization. Keeping information secured can protect an organization's image and save an organization a lot of money.
- Hacking can lead to loss of business for organizations that deal in finance such as PayPal. Ethical hacking puts them a step ahead of the cyber criminals who would otherwise lead to loss of business.

Legality of Ethical Hacking

Ethical Hacking is legal if the hacker abides by the rules stipulated as above. The International Council of E-Commerce Consultants (EC-Council) provides a certification program that tests individual's skills. Those who pass the examination are awarded with certificates. The certificates are supposed to be renewed after some time.

Figure 5.3: Penetration Testing Stages

5.1.3 Data Privacy and General Data Protection and Regulation (GDPR):

5.1.1 Data Privacy:

Data privacy is a guideline for how data should be collected or handled, based on its sensitivity and importance. Data privacy is typically applied to personal health information (PHI) and personally identifiable information (PII). This includes financial information, medical records, social security or ID numbers, names, birthdates, and contact information.

Data privacy concerns apply to all sensitive information that organizations handle, including that of customers, shareholders, and employees. Often, this information plays a vital role in business operations, development, and finances.

Data privacy helps ensure that sensitive data is only accessible to approved parties. It prevents criminals from being able to maliciously use data and helps ensure that organizations meet regulatory requirements.

5.1.2 Data Protection:

Data protection is a set of strategies and processes you can use to secure the privacy, availability, and integrity of your data. It is sometimes also called data security.

A data protection strategy is vital for any organization that collects, handles, or stores sensitive data. A successful strategy can help prevent data loss, theft, or corruption and can help minimize damage caused in the event of a breach or disaster.

Data privacy defines who has access to data, while data protection provides tools and policies to actually restrict access to the data.

5.1.3 Data Protection Principles:

Data protection principles help protect data and make it available under any circumstances. It covers operational data backup and business continuity/disaster recovery (BCDR) and involves implementing aspects of data management and data availability.

Here are key data management aspects relevant to data protection:

- Data availability—ensuring users can access and use the data required to perform business even when this data is lost or damaged.
- Data lifecycle management—involves automating the transmission of critical data to offline and online storage.
- Information lifecycle management—involves the valuation, cataloging, and protection of information assets from various sources, including facility outages and disruptions, application and user errors, machine failure, and malware and virus attacks.

5.1.4 GDPR:

The GDPR is a legal standard that protects the personal data of European Union (EU) citizens and affects any organization that stores or processes their personal data, even if it does not have a business presence in the EU.

Because there are hundreds of millions of European Internet users, the standard affects almost every company that collects data from customers or prospects over the Internet. GDPR non-compliance carries severe sanctions, with fines up to 4% of annual revenue or €20 million.

GDPR legislators aimed to define data privacy as a basic human right, and standardize the protection of personal data while putting data subjects in control of the use and retention of their data.

There are two primary roles in the GDPR: the **GDPR Data Controller** is an entity that collects or processes personal data for its own purposes, and a **GDPR Data Processor** is an entity that holds or processes this type of data on behalf of another organization.

Finally, the Data Protection Officer is a role appointed by an organization to monitor how personal data is processed and ensure compliance of the GDPR.

What is personal data according to the GDPR?

“Personal data”, according to the legal definition of the GDPR legislation, is any information about an identified or identifiable person, known as a **data subject**.

Personal data includes any information that can be used, alone or in combination with other information, to identify someone.

This includes: name, address, ID or passport number, financial info, cultural details, IP addresses, or medical data used by healthcare professionals or institutions.

Other special data you may not process or store: Race or ethnicity, sexual orientation, religious beliefs, political beliefs of memberships, health data (unless the explicit concern is granted or there is substantial public interest).

5.1.5 GDPR data privacy rights:

The GDPR aims to protect the following rights of data subjects with respect to their personal data.

Data subjects have the following basic rights under the GDPR:

- **Collecting data from children** — requires parental consent until children are between 13-16 years old.
- **Data portability and access** — data subjects must be able to access their data as stored by the Data Controller, know-how and why it is being processed, and where it is being sent.
- **Correcting and objecting to data** — data subjects should be able to correct incorrect or incomplete data, and data controllers must notify all data recipients of the change. They should also be able to object to the use of their data, and Data Controllers must comply unless they have a legitimate interest that overrides the data subject's interest.
- **Right to erasure** — data subjects can ask data controllers to "forget" their personal data. Organizations may be permitted to retain the data, for example, if they need it to comply with a legal obligation or if it is in the public interest, for example in the case of scientific or historical research.
- **Automated decision-making** — data subjects have the right to know that they were subject to an automated decision based on their private information, and can request that the automated decision is reviewed by a person, or contest the automated decision.
- **Notification of breaches** — if personal data under the responsibility of a data controller is exposed to unauthorized parties, the controller must notify the Data Protection Authority in the relevant EU country within 72 hours, and in some cases also needs to inform individual data subjects.
- **Transferring data outside the EU** — if personal data is transferred outside the EU, the data controller should ensure there are equivalent measures to protect the data and the rights of data subjects.

5.1.6 Data Protection Technologies and Practices to Protect Your Data:

When it comes to protecting your data, there are many storage and management options you can choose from. Solutions can help you restrict access, monitor activity, and respond to threats. Here are some of the most commonly used practices and technologies:

1. **Data discovery**—a first step in data protection, this involves discovering which data sets exist in the organization, which of them are business critical and which contains sensitive data that might be subject to compliance regulations.
2. **Data loss prevention (DLP)**—a set of strategies and tools that you can use to prevent data from being stolen, lost, or accidentally deleted. Data loss prevention solutions often include several tools to protect against and recover from data loss.
3. **Storage with built-in data protection**—modern storage equipment provides built-in disk clustering and redundancy.
4. **Backup**—creates copies of data and stores them separately, making it possible to restore the data later in case of loss or modification. Backups are a critical strategy for ensuring business continuity when original data is lost, destroyed, or damaged, either accidentally or maliciously.

5. **Snapshots**—a snapshot is similar to a backup, but it is a complete image of a protected system, including data and system files. A snapshot can be used to restore an entire system to a specific point in time.
6. **Replication**—a technique for copying data on an ongoing basis from a protected system to another location. This provides a living, up-to-date copy of the data, allowing not only recovery but also immediate failover to the copy if the primary system goes down.
7. **Firewalls**—utilities that enable you to monitor and filter network traffic. You can use firewalls to ensure that only authorized users are allowed to access or transfer data.
8. **Authentication and authorization**—controls that help you verify credentials and assure that user privileges are applied correctly. These measures are typically used as part of an identity and access management (IAM) solution and in combination with role-based access controls (RBAC).
9. **Encryption**—alters data content according to an algorithm that can only be reversed with the right encryption key. Encryption protects your data from unauthorized access even if data is stolen by making it unreadable.
10. **Endpoint protection**—protects gateways to your network, including ports, routers, and connected devices. Endpoint protection software typically enables you to monitor your network perimeter and to filter traffic as needed.
11. **Data erasure**—limits liability by deleting data that is no longer needed. This can be done after data is processed and analyzed or periodically when data is no longer relevant. Erasing unnecessary data is a requirement of many compliance regulations, such as GDPR.
12. **Disaster recovery**—a set of practices and technologies that determine how an organization deals with a disaster, such as a cyber attack, natural disaster, or large-scale equipment failure. The disaster recovery process typically involves setting up a remote disaster recovery site with copies of protected systems, and switching operations to those systems in case of disaster.

5.2 Understanding the need to hack your own systems:

To catch a thief, think like a thief. That's the basis for ethical hacking.

The law of averages works against security. With the increased numbers and expanding knowledge of hackers combined with the growing number of system vulnerabilities and other unknowns, the time will come when all computer systems are hacked or compromised in some way. Protecting your systems from the bad guys and not just the generic vulnerabilities that everyone knows about is absolutely critical. When the hacker tricks are known, one can see how vulnerable the systems are.

Hacking targets on weak security practices and undisclosed vulnerabilities. Firewalls, encryption, and virtual private networks (VPNs) can create a false feeling of safety. These security systems often focus on high-level vulnerabilities, such as viruses and traffic through a firewall, without affecting how hackers work. Attacking your own systems to discover vulnerabilities is a step to making them more secure.

This is the only proven method of greatly hardening your systems from attack. If weaknesses are not identified, it's a matter of time before the vulnerabilities are exploited.

As hackers expand their knowledge, one should also gain the required knowledge of it. You must think like them to protect your systems from them. As the ethical hacker, one must know activities hackers carry out and how to stop their efforts. One should know what to look for and how to use that information to spoil hackers' efforts.
One cannot protect the systems from everything. The only protection against everything is to unplug computer systems and lock them away so no one can touch them , not even you.
That's not the best approach to information security. What's important is to protect your systems from known vulnerabilities and common hacker attacks. It's impossible to support all possible vulnerabilities on all systems. One can't plan for all possible attacks, especially the ones that are currently unknown.
However, the more combinations you try — the more you test whole systems instead of individual units ,the better your chances of discovering vulnerabilities that affect everything as a whole.

Building the Foundation for Ethical Hacking:

One should not forget about insider threats from malicious employees. One's overall goals as an ethical hacker should be as follows:

- ✓ Hack your systems in a non-destructive fashion.
- ✓ Enumerate vulnerabilities and, if necessary, prove to upper management that vulnerabilities exist.
- ✓ Apply results to remove vulnerabilities and better secure your systems.

5.3 Understanding the dangers your systems face

Systems are generally under fire from hackers around the world. It's another to understand specific attacks against your systems that are possible.
There are some well-known attacks. Many information-security vulnerabilities aren't critical by themselves. However, exploiting several vulnerabilities at the same time can take its toll.
For example, a default Windows OS configuration, a weak SQL Server administrator password, and a server hosted on a wireless network may not be major security concerns separately. But exploiting all three of these vulnerabilities at the same time can be a serious issue as:

- Nontechnical attacks
- Network-infrastructure attacks
- Operating-system attacks
- Application and other specialized attacks

5.3.1 Nontechnical attacks:

Exploits that involve manipulating people or end users and even yourself are the greatest vulnerability within any computer or network infrastructure. Humans are trusting by nature, which can lead to social-engineering exploits. Social engineering is defined as the exploitation of the trusting nature of human beings to gain information for malicious purposes.

Other common and effective attacks against information systems are physical. Hackers break into buildings, computer rooms, or other areas containing critical information or property. Physical attacks can include dumpster diving (searching through trash cans and dumpsters for intellectual property, passwords, network diagrams, and other information).

5.3.2 Network-infrastructure attacks:

Hacker attacks against network infrastructures can be easy, because many networks can be reached from anywhere in the world via the Internet.

Here are some examples of network-infrastructure attacks:

- ✓ Connecting into a network through a rogue modem attached to a computer behind a firewall
- ✓ Exploiting weaknesses in network transport mechanisms, such as TCP/IP and NetBIOS.
- ✓ Flooding a network with too many requests, creating a Denial of Service (DoS) for legitimate requests
- ✓ Installing a network analyzer on a network and capturing every packet that travels across it, revealing confidential information in clear text
- ✓ Piggybacking onto a network through an insecure wireless configuration.

5.3.3 Operating-system attacks Hacking

Operating Systems (OSs) is a preferred method of the bad guys(hackers). Operating systems comprise a large portion of hacker attacks simply because every computer has one and so many well-known exploits can be used against them.

Occasionally, some operating systems that are more secure out of the box, such as Novell NetWare and the flavor's of BSD UNIX are attacked, and vulnerabilities turn up.

But hackers prefer attacking operating systems like Windows and Linux because they are widely used and better known for their vulnerabilities.

Here are some examples of attacks on operating systems:

- ✓ Exploiting specific protocol implementations
- ✓ Attacking built-in authentication systems
- ✓ Breaking file-system security
- ✓ Cracking passwords and encryption mechanisms

5.3.4 Application and other specialized attacks:

Applications take a lot of hits by hackers. Programs such as e-mail server software and Web applications often are beaten down:

- ✓ Hypertext Transfer Protocol (HTTP) and Simple Mail Transfer Protocol (SMTP) applications are frequently attacked because most firewalls and other security mechanisms are configured to allow full access to these programs from the Internet.
- ✓ Malicious software (malware) includes viruses, worms, Trojan horses, and spyware. Malware clogs networks and takes down systems.
- ✓ Spam (junk e-mail) is wreaking havoc on system availability and storage space. And it can carry malware. Ethical hacking helps reveal such attacks against computer systems.

5.4 Obeying the Ethical Hacking Commandments:

Every ethical hacker must abide by a few basic commandments. If not, bad things can happen.

5.4.1 Working ethically:

The word ethical in this context can be defined as working with high professional morals and principles. While performing ethical hacking tests against own systems or for someone who has hired for, everything one need to do as an ethical hacker must be above board and must support the company's goals. No hidden agendas are allowed. Trustworthiness is the ultimate principle. The misuse of information is absolutely forbidden. That's what the bad guys or hackers do.

5.4.2 Respecting privacy:

Treat the information gathered with the greatest respect. All information obtained during testing from Web-application log files to clear-text passwords must be kept private. This information shall not be used to watch into confidential corporate information or private lives. If you sense or feel that someone should know there's a problem, consider sharing that information with the appropriate manager.

Involve others in process. This is a "watch the watcher" system that can build trust and support ethical hacking projects.

5.4.3 Not crashing your systems:

One of the biggest mistakes seen when people try to hack their own systems is inadvertently crashing their systems. The main reason for this is poor planning. These testers have not read the documentation or misunderstand the usage and power of the security tools and techniques.

DoS-Denial of Service conditions on the systems are easily created when testing. Running too many tests too quickly on a system causes many system lockups. Things should not be rushed and assumed that a network or specific host can handle the beating that network scanners and vulnerability assessment tools can be useless.

Many security-assessment tools can control how many tests are performed on a system at the same time. These tools are especially handy if one needs to run the tests

on production systems during regular business hours. One can even create an account or system lockout condition by social engineering, changing a password, not realizing that doing so might create a system lockout condition.

5.5 The Ethical Hacking Process:

Like practically any IT or security project, ethical hacking needs to be planned in advance. Strategic and tactical issues in the ethical hacking process should be determined and agreed upon. Planning is important for any amount of testing from a simple password-cracking test to an all-out penetration test on a Web application.

5.5.1 Formulating your plan:

Approval for ethical hacking is essential. What is being done should be known and visible at least to the decision makers. Obtaining sponsorship of the project is the first step. This could be the manager, an executive, a customer, or even the boss. Someone is needed to back up and sign off on the plan. Otherwise, testing may be called off unexpectedly if someone claims they never authorized one to perform the tests.

The authorization can be as simple as an internal memo from the senior-most person or boss if one is performing these tests on own systems. If the testing is for a customer, one should have a signed contract in place, stating the customer's support and authorization. Get written approval on this sponsorship as soon as possible to ensure that none of the time or effort is wasted. This documentation works as a proof as what one is doing when someone asks or demands.

A detailed plan is needed, but that doesn't mean that it needs volumes of testing procedures. One slip can crash your systems.

A **well-defined scope** includes the following information:

- ✓ Specific systems to be tested
- ✓ Risks that are involved
- ✓ When the tests are performed and your overall timeline
- ✓ How the tests are performed
- ✓ How much knowledge of the systems you have before you start testing
- ✓ What is done when a major vulnerability is discovered
- ✓ The specific deliverables — this includes security-assessment reports and a higher-level report outlining the general vulnerabilities to be addressed, along with countermeasures that should be implemented.
- ✓ When selecting systems to test, start with the most critical or vulnerable systems. For instance, one can test computer passwords or attempt social engineering attacks before drilling down into more detailed systems.

  What if one is assessing the firewall or Web application, and one takes it down? This can cause system unavailability, which can reduce system performance or employee productivity. Even worse, it could cause loss of data integrity, loss of data, and bad publicity.

Handle social-engineering and denial-of-service attacks carefully. Determine how they can affect the systems you're testing and entire organization.

Determining when the tests are performed is something that one must think long and hard about. Do the tester test during normal business hours? How about late at night or early in the morning so that production systems aren't affected? Involve others to make sure they approve tester's timing.

The best approach is an unlimited attack, wherein any type of test is possible. The hackers aren't hacking the systems within a limited scope. Some exceptions to this approach are performing DoS, social engineering, and physical-security tests.

One should not stop with one security hole. This can lead to a false sense of security. One should keep going to see what else he/she can discover. It's not like to keep hacking until the end of time or until one crash all his/ her systems. Simply pursue the path he/she is going down until he//she can't hack it any longer.

One of the goals may be to perform the tests without being detected.

For example, one may be performing his/her tests on remote systems or on a remote office, and he/she doesn't want the users to be aware of what they are doing. Otherwise, the users may be on to him/her and be on their best behaviour.

Extensive knowledge of the systems is not needed for testing . Just a basic understanding is required to protect the tested systems.

Understanding the systems which are being tested shouldn't be difficult if one is hacking his/her own in-house systems. If hacking a customer's systems, one may have to dig deeper. In fact, Most people are scared of these assessments. Base the type of test one will perform on his/her organization's or customer's needs.

5.5.2 Selecting tools:

If one don't have the right tools for ethical hacking, to accomplish the task is effectively difficult. just using the right tools doesn't mean that all vulnerabilities will be discovered.

Know the personal and technical limitations.

Many security-assessment tools generate false positives and negatives (incorrectly identifying vulnerabilities). Some tools may miss vulnerabilities. Many tools focus on specific tests, but no one tool can test for everything. This is why a set of specific tools are required that can call on for the task at hand. The more are the tools , the easier ethical hacking efforts are.

Make sure the right tool is being used for the task :

- To crack passwords, one needs a cracking tool such as LC4, John the Ripper, or pwdump.

A general port scanner, such as SuperScan, may not crack passwords.

- For an in-depth analysis of a Web application, a Web-application assessment tool (such as Whisker or WebInspect) is more appropriate than a network analyzer (such as Ethereal).

When selecting the right security tool for the task, ask around. Get advice from the colleagues and from other people online. A simple Groups search on Google (www.google.com) or perusal of security portals, such as SecurityFocus.com, SearchSecurity.com, and ITsecurity.com, often produces great feedback from other security experts.

Some of the widely used commercial, freeware, and open-source security tools:

- Nmap
- EtherPeek
- SuperScan
- QualysGuard
- WebInspect
- LC4 (formerly called L0phtcrack)
- LANguard Network Security Scanner
- Network Stumbler
- ToneLoc

Here are some other popular tools:

- Internet Scanner
- Ethereal
- Nessus
- Nikto
- Kismet
- THC-Scan

The capabilities of many security and hacking tools are often misunderstood. This misunderstanding has shed negative light on some excellent tools, such as **SATAN** (**S**ecurity **A**dministrator **T**ool for **A**nalysing **N**etworks) and **Nmap** (**N**etwork **map**per).

Some of these tools are complex. Whichever tools are being used, one should be familiarized with them before starting to use them.

Here are ways to do that:

- ✓ Read the readme and/or online help files for tools.
- ✓ Study the user's guide for commercial tools.
- ✓ Consider formal classroom training from the security-tool vendor or another third-party training provider, if available.
- ✓ One should Look for these characteristics in tools for ethical hacking:
- ✓ Adequate documentation.

- ✓ Detailed reports on the discovered vulnerabilities, including how they may be exploited and fixed.
- ✓ Updates and support when needed.
- ✓ High-level reports that can be presented to managers or non-techie types.
- ✓ These features can save time and effort when writing the report.

5.5.3 Executing the plan:

Ethical hacking can take persistence. Time and patience are important. One should be careful when performing ethical hacking tests. A hacker in network or a seemingly gentle employee looking over one's shoulder may watch what's going on. This person could use this information against tester.

It's not practical to make sure that no hackers are on one's systems before starting. Just one has to make sure to keep everything as quiet and private as possible. This is especially critical when transmitting and storing own test results. If possible, one should encrypt these e-mails and files using Pretty Good Privacy (PGP) or something similar. At a minimum, password-protect them.

In an investigation mission, attach as much information as possible about the organization and systems, which is what malicious hackers do.

Start with a broad view and narrow down the focus:

1. Search the Internet for own organization's name, computer and network system names, and the IP addresses.
   Google is a great place to start for this.
2. Narrow the scope, targeting the specific systems to be tested or being tested.
   Whether physical-security structures or Web applications, a casual assessment can turn up much information about the systems.
3. Further narrow down focus with a more critical eye. Perform actual scans and
   other detailed tests on the systems.
4. Perform the attacks, if that's what one choose to do.

5.5.4 Evaluating results:

Assess the results to see what has been uncovered, assuming that the vulnerabilities haven't been made obvious before now. This is where knowledge counts. Evaluating the results and correlating the specific vulnerabilities discovered is a skill that gets better with experience. One will end up knowing his/her own systems as well as anyone else. This makes the evaluation process much simpler moving forward.

Submit a formal report to upper management or to the customer, outlining results. Keep these other parties in the loop to show that efforts and their money are well spent.

5.5.5 Moving on:

When finished with ethical hacking tests, one still need to implement his/her analysis and recommendations to make sure that the systems are secure.

New security vulnerabilities continually appear. Information systems constantly change and become more complex. New hacker exploits and security vulnerabilities are regularly uncovered. Security tests are a snapshot of the security posture of the systems.

At any time, everything can change, especially after software upgrades, adding computer systems, or applying patches. Plan to test regularly (for example, once a week or once a month).

5.6 Cyber Security act:

Cyber law Act 2000:

The Act provides a legal framework for electronic governance by giving recognition to electronic records and digital signatures. It also defines cyber crimes and prescribes penalties for them. The Act directed the formation of a Controller of Certifying Authorities to regulate the issuance of digital signatures.

Amendments:

A major amendment was made in 2008. It introduced Section 66A which penalized sending "offensive messages". It also introduced Section 69, which gave authorities the power of "interception or monitoring or decryption of any information through any computer resource". Additionally, it introduced provisions addressing - pornography, child porn, cyber terrorism and voyeurism. The amendment was passed on 22 December 2008 without any debate in Lok Sabha. The next day it was passed by the Rajya Sabha. It was signed into law by President Pratibha Patil, on 5 February 2009.

Offences:

List of offences and the corresponding penalties

Section	Offence	Penalty
65	Tampering with computer source documents	Imprisonment up to three years, or/and with fine up to ₹200,000
66	Hacking with computer system	Imprisonment up to three years, or/and with fine up to ₹500,000
66B	Receiving stolen computer or communication device	Imprisonment up to three years, or/and with fine up to ₹100,000
66C	Using password of another person	Imprisonment up to three years, or/and with fine up to ₹100,000

66D	Cheating using computer resource	Imprisonment up to three years, or/and with fine up to ₹100,000
66E	Publishing private images of others	Imprisonment up to three years, or/and with fine up to ₹200,000
66F	Acts of cyberterrorism	Imprisonment up to life.
67	Publishing information which is obscene in electronic form.	Imprisonment up to five years, or/and with fine up to ₹1,000,000
67A	Publishing images containing sexual acts	Imprisonment up to seven years, or/and with fine up to ₹1,000,000
67C	Failure to maintain records	Imprisonment up to three years, or/and with fine.
68	Failure/refusal to comply with orders	Imprisonment up to 2 years, or/and with fine up to ₹100,000
69	Failure/refusal to decrypt data	Imprisonment up to seven years and possible fine.
70	Securing access or attempting to secure access to a protected system	Imprisonment up to ten years, or/and with fine.
71	Misrepresentation	Imprisonment up to 2 years, or/and with fine up to ₹100,000
72	Breach of confidentiality and privacy	Imprisonment up to 2 years, or/and with fine up to ₹100,000
72A	Disclosure of information in breach of lawful contract	Imprisonment up to 3 years, or/and with fine up to ₹500,000
73	Publishing electronic signature certificate false in certain particulars	Imprisonment up to 2 years, or/and with fine up to ₹100,000
74	Publication for fraudulent purpose	Imprisonment up to 2 years, or/and with fine up to ₹100,000

References

- https://www.dynamicchiropractic.com/mpacms/dc/article.php?id=18078)
- Hacking For Dummies, 5th Edition By Kevin Beaver
- http://cdn.ttgtmedia.com/searchNetworking/downloads/hacking_for_dummies
- http://wiki.cas.mcmaster.ca/index.php/Ethical_Hacking
- https://www.dummies.com/programming/networking/what-is-a-malicious-user/
- https://www.guru99.com/what-is-hacking-an-introduction.html#2
- http://cdn.ttgtmedia.com/searchNetworking/downloads/hacking_for_dummies.pdf
- 2600 — The Hacker Quarterly magazine (www.2600.com)
- (IN)SECURE Magazine (www.net-security.org/insecuremag.php)

- Hackin9 (http://hakin9.org)
- PHRACK (www.phrack.org/archives/)
- https://learning.oreilly.com/library/view/hacking-for-dummies/9781118380956/06_9781118380956-ch02.html
- https://www.quora.com/What-knowledge-is-required-to-become-an-ethical-hacker

Sample Multiple Choice Questions:

1) Ethical Hacking is also known as_____
   a. Black Hat hacking
   b. White hat hacking
   c. Encrypting
   d. None of these

2) Tool/s used by ethical hackers _____
   a. Scanner
   b. Decoder
   c. Proxy
   d. All of these

3) Vulnerability scanning in Ethical hacking finds_____
   a. Strengths
   b. Weakness
   c. Strengths & Weakness
   d. None of these

4) Ethical hacking will allow to_____ all the massive security breaches.
   a. remove
   b. measure
   c. reject
   d. None of these

5) Sequential steps hackers uses are: __, ___, __, __
   A) Maintaining Access
   B) Reconnaissance
   C) Scanning
   D) Gaining Access
   a. B, C, D, A
   b. B, A, C, D
   c. A, B, C, D
   d. D, C, B, A

6) The Information Technology Act 2000 is an Act of Indian Parliament notified on_____.

a. 27th October 2000
b. 15th December 2000
c. 17th November 2000
d. 17th October 2000

7) The offense “Receiving stolen computer or communication device” comes under ____ of Cyber security Act 2000.

a. 66B
b. 67A
c. 66E
d. 66C

8) _____is similar to a backup, but it is a complete image of a protected system, including data and system files.

a. Replication
b. Backup
c. Snapshots
d. DPLR

9) Data subjects can ask data controllers to “forget” their personal data is _____.

a. Right to erasure
b. Automated decision making
c. Transferring data outside the EU
d. Right to Control

Unit-6 Types of Hacking

Contents

6.1 Network Hacking

- Network Infrastructure
- Network Infrastructure Vulnerabilities
- Scanning-Ports
- Ping sweeping
- Scanning SNMP
- Grabbing Banners
- MAC-daddy attack

Wireless LANs:

- Wireless Network Attacks

6.2 Operating System Hacking

- Introduction of Windows and LinuxVulnerabilities
- Buffer Overflow Attack

6.3 Applications Hacking

Messaging Systems

- Vulnerabilities
- E-Mail Attacks- E-Mail Bombs,
- Banners
- Best practices for minimizing e-mail security risks

Web Applications:

- Web Vulnerabilities
- Directories Traversal and Countermeasures
- Google Dorking

Database system

- Database Vulnerabilities
- Best practices for minimizing database security risks

-

6.1 Network Hacking:

- Network is one of the most fundamental communications systems in your organization. Network consists of such devices as routers, firewalls, and even generic hosts (including servers and workstations) that you must assess as part of the ethical hacking process.
- There are thousands of possible network vulnerabilities, equally as many tools, and even more testing techniques. We don’t need to test our network for every possible vulnerability, using every tool available.
- We can eliminate many well-known network vulnerabilities by simply patching your network hosts with the latest vendor software and firmware patches. We can eliminate many other vulnerabilities by following some security best practices on our network.

6.1.1 Network Infrastructure Vulnerabilities:

- Network infrastructure vulnerabilities are the foundation for all technical security issues in your information systems. These lower-level vulnerabilities affect everything running on your network. That's why you need to test for them and eliminate them whenever possible.
- Your focus for ethical hacking tests on your network infrastructure should be to find weaknesses that others can see in your network so you can quantify your level of exposure.
- Many issues are related to the security of your network infrastructure. Some issues are more technical and require you to use various tools to assess them properly. You can assess others with a good pair of eyes and some logical thinking. Some issues are easy to see from outside the network, and others are easier to detect from inside your network
- Network infrastructure security involves assessing such areas as
  - ✓ Where such devices as a firewall or IDS (intrusion detection system) are placed on the network and how they are configured
  - ✓ What hackers see when they perform port scans and how they can exploit vulnerabilities in your network hosts
  - ✓ Network design, such as Internet connections, remote-access capabilities, layered defenses, and placement of hosts on the network
  - ✓ Interaction of installed security devices
  - ✓ Protocols in use
  - ✓ Commonly attacked ports that are unprotected
  - ✓ Network host configuration
  - ✓ Network monitoring and maintenance
- If any of these network security issues is exploited, such things can happen:
  - ✓ A DoS attack can take down your Internet connection or even your entire network.
  - ✓ A hacker using a network analyzer can steal confidential information in e-mails and files being transferred.
  - ✓ Backdoors into your network can be set up.
  - ✓ Specific hosts can be attacked by exploiting local vulnerabilities across the network.
- Always remember to do the following:
  - ✓ Test your systems from both the outside in and the inside out.
  - ✓ Obtain permission from partner networks that are connected to your network to check for vulnerabilities on their ends that can affect your network's security, such as open ports and lack of a firewall or a misconfigured router.

6.1.2 Network Testing and port Scanning tools:

- **Sam Spade** for Windows for network queries from DNS lookups to trace routes
- **SuperScan** for ping sweeps and port scanning

- **NetScan**Tools Pro for dozens of network security-assessment functions, including ping sweeps, port scanning, and SMTP relay testing
- **Nmap or NMapWin** as a happy-clicky-GUI front end for host-port probing and operating-system fingerprinting
- **Netcat** the most versatile security tool for such security checks as port scanning and firewall testing
- **WildPacketsEtherPeek** for network analysis.

6.1.3 Scanning-Ports:

- A port scanner is a software tool that basically scans the network to see who's there. Port scanners provide basic views of how the network is laid out. They can help identify unauthorized hosts or applications and network host configuration errors that can cause serious security vulnerabilities.
- The big-picture view from port scanners often uncovers security issues that may otherwise go unnoticed. Port scanners are easy to use and can test systems regardless of what operating systems and applications they're running. The tests can be performed very quickly without having to touch individual network hosts, which would be a real pain otherwise.
- Port-scan tests take time. The length of time depends on the number of hosts you have, the number of ports you scan, the tools you use, and the speed of your network links. Also, perform the same tests with different utilities to see whether you get different results. Not all tools find the same open ports and vulnerabilities. This is unfortunate, but it's a reality of ethical hacking tests.
- If your results don't match after you run the tests using different tools, you may want to explore the issue further. If something doesn't look right such as a strange set of open ports it probably isn't. Test it again; if you're in doubt, use another tool for a different perspective.
- As an ethical hacker, you should scan all 65,535 UDP and 65,535 TCP ports on each network host that's found by your scanner. If you find questionable ports, look for documentation that the application is known and authorized. For speed and simplicity, you can scan commonly hacked ports.

Port Nos.	Service	Protocols
7	Echo	TCP, UDP
19	Chargen	TCP, UDP
20	FTP data (File Transfer Protocol)	TCP
21	FTP control	TCP
22	SSH	TCP

23	Telnet	TCP
25	SMTP (Simple Mail Transfer Protocol)	TCP
37	Daytime	TCP, UDP
53	DNS (Domain Name System)	UDP
69	TFTP (Trivial File Transfer Protocol)	UDP
79	Finger	TCP, UDP
80	HTTP (Hypertext Transfer Protocol)	TCP
110	POP3 (Post Office Protocol version 3)	TCP
111	SUN RPC (remote procedure calls)	TCP, UDP
135	RPC/DCE end point mapper for Microsoft networks	TCP, UDP
137, 138, 139	NetBIOS over TCP/IP	TCP, UDP
161	SNMP (Simple Network Management Protocol)	TCP, UDP
220	IMAP (Internet Message Access Protocol)	TCP
443	HTTPS (HTTP over SSL)	TCP
512, 513, 514	Berkeley r commands (such as rsh, rexec, and rlogin)	TCP
1214	Kazaa and Morpheus	TCP, UDP
1433	Microsoft SQL Server	TCP, UDP
1434	Microsoft SQL Monitor	TCP, UDP
3389	Windows Terminal Server	TCP
5631, 5632	pcAnywhere	TCP
6346, 6347	Gnutella	TCP, UDP

12345, 12346, 12631, 12632, 20034, 20035	NetBus	TCP
27444	Trinoo	UDP
27665	Trinoo	TCP
31335	Trinoo	UDP
31337	Back Orifice	UDP
34555	Trinoo	UDP

Table 5.: Commonly hacked ports

Countermeasures (Port Scanning)

- You can implement various countermeasures to typical port scanning.

✓ Traffic restriction

- Enable only the traffic you need to access internal hosts preferably as far as possible from the hosts you're trying to protect. You apply these rules in two places: External router for inbound traffic & Firewall for outbound traffic .
- Configure firewalls to look for potentially malicious behavior over time (such as the number of packets received in a certain period of time), and have rules in place to cut off attacks if a certain threshold is reached, such as 100 port scans in one minute. Most firewalls, IDSs, andIDPs detect port scanning and cut it off in real time.

✓ **Gathering network information**

- NetScanTools Pro is a great tool for general network information, such as the - number of unique IP addresses, NetBIOS names, and MAC addresses found.
- The following report is an example of the NetScanner (network scanner) output of NetScanTools Pro 2000:
- Scan completion time = Sat, 7 Feb 2004 14:11:08
- Start IP address: 192.168.1.1
- End IP address: 192.168.1.254
- Number of target IP addresses: 254
- Number of IP addresses responding to pings: 13
- Number of IP addresses sent pings: 254
- Number of intermediate routers responding to pings: 0
- Number of successful NetBIOS queries: 13
- Number of IP addresses sent NetBIOS queries: 254
- Number of MAC addresses obtained by NetBIOS queries: 13
- Number of successful Subnet Mask queries: 0

- Number of IP addresses sent Subnet Mask queries: 254
- Number of successful Whois queries: 254

✓ **Traffic denial**

- Deny ICMP traffic to specific hosts you're trying to protect. Most hosts don't need to have ICMP enabled especially inbound ICMP requests unless it's needed for a network management system that monitors hosts using this protocol.
- You can break applications on your network, so make sure that you analyze what's going on, and understand how applications and protocols are working, before you disable such network traffic as ICMP.

6.1.4 Ping sweeping:

- Port sweeping is regarded by certain systems experts to be different from port scanning.
- They point out that port scanning is executed through the searching of a single host for open ports. However, they state that port sweeping is executed through the searching of multiple hosts in order to target just one specific open port.
- While Port scanning and sweeping have legitimate uses with regard to network management, unfortunately, they are used almost as frequently for the purpose of criminal activity.

A Serious Threat

- Any times there are open ports on one's personal computer, there is potential for the loss of data, the occurrence of a virus, and at times, even complete system compromise.
- It is essential for one to protect his or her virtual files, as new security risks concerning personal computers are discovered every day.
- Computer protection should be the number one priority for those who use personal computers.
- Port scanning is considered a serious threat to one's PC, as it can occur without producing any outward signs to the owner that anything dangerous is taking place.

Firewall Protection

- Protection from port scanning is often achieved through the use of a firewall. A firewall monitors incoming and outgoing connections through one's personal computer.
- One technique used by firewall technology is the opening of all the ports at one time. This action stops port scans from returning any ports. This has worked in many situations in the past, however, most experts agree it is best to have all open ports investigated individually.
- Another approach is to filter all port scans going to one's computer. An individual can also choose to port scan his or her own system, which enables one to see the personal computer through the eyes of a hacker.

- Firewalls are the best protection one can invest in with regard to port scanning. Firewalls deny outside access to an individual's personal computer. With this type of protection, a personal computer is essentially hidden from unwelcome visitors and is also protected from a variety of other hacking techniques. With firewall software, an individual is assured that his or her sensitive and personal information remains protected.
- A ping sweep of all your network subnets and hosts is a good way to find out which hosts are alive and kicking on the network.
- A ping sweep is when you ping a range of addresses using Internet Control Message Protocol (ICMP) packets.
- Dozens of Nmap command-line options exist, which can be overwhelming when you just want to do a basic scan.
- You can just enter nmap on the command line to see all the options available.
- These command-line options can be used for an Nmap ping sweep:
- **sP** tells Nmap to perform a ping scan.
- **n**tells Nmap to not perform name resolution. You may want to omit this if you want to resolve hostnames to see which systems are responding. Name resolution may take slightly longer, though.
- -T 4 option tells Nmap to perform an aggressive (faster) scan.
- 192.168.1.1-254 tells Nmap to scan the entire 192.168.1.x subnet.

6.1.5 SNMP(Simple Network Management Protocol) scanning:

- Networks are the of every business. Even in small or enterprise-level businesses, the loss of productivity during a network outage can result in hefty damages.
- Network monitoring helps you anticipate potential outages and address network problems proactively. This helps in maintaining a congestion-free network that keeps your business up and running.
- A network monitoring software helps you to monitor the performance of any IP-based device and helps businesses remotely visualize their system performance and monitor network services, bandwidth utilization, switches, routers and traffic flow.

Vulnerabilities (SNMP)

- The problem is that most network hosts run SNMP that isn't hardened or patched to prevent known security vulnerabilities. The majority of network devices have SNMP enabled and don't even need it.
- If SNMP is compromised, a hacker can gather such network information as ARP tables and TCP connections to attack your systems. If SNMP shows up in port scans, you can bet that a hacker will try to compromise the system.
- **Countermeasures (SNMP)**
- Preventing SNMP attacks can be as simple as A-B-C:
- Always disable SNMP on hosts if you're not using it period.
- Block the SNMP port (UDP port 161) at the network perimeter.

- Change the default SNMP community string from public to another value that's more difficult to guess. This makes SNMP harder to hack.

6.1.6 Banner grabbing:

- Banner grabbing is the act of capturing the information provided by banners, configurable text-based welcome screens from network hosts that generally display system information. Banners are intended for network administration
- Banner grabbing is often used for white hathacking endeavors like vulnerability analysis and penetration testing as well as gray hat activities and black hat hacking. Banner screens can be accessed through Telnet at the command prompt on the target system's IP address.
- Other tools for banner grabbing include Nmap, Netcat and SuperScan. A login screen, often associated with the banner, is intended for administrative use but can also provide access to a hacker. Meanwhile, the banner data can yield information about vulnerable software and services running on the host system.
- For the sake of security, if banners are not a requirement of business or other software on a host system, the services that provide them may be disabled altogether. Banners can also be customized to present disinformation or even a warning message for hackers
- Banners are the welcome screens that divulge software version numbers and other host information to a network host. This banner information may identify the operating system, the version number, and the specific service packs, so hackers know possible vulnerabilities. **You can grab banners by using either plain old telnet or Netcat.**
- Telnet

✓ You can telnet to hosts on the default telnet port (TCP port 23) to see whether you're presented with a login prompt or any other information.

✓ Just enter the following line at the command prompt in Windows or UNIX:

✓ telnet ip_address

- Netcat

✓ Netcat can grab banner information from routers and other network hosts, such as a wireless access point or managed Ethernet switch.

- Countermeasures (Banner Grabbing)

✓ The following steps can reduce the chance of banner-grabbing attacks:

- If there is no business need for services that offer banner information, disable those unused services on the network host.
- If there is no business need for the default banners, or if you can customize the banners displayed, configure the network host's application or operating system to either disable the banners or remove information from the banners that could give an attacker a leg up.

6.1.7 The MAC-daddy attack:

- Hackers can use ARP Protocol that is running on the network to make their systems seem as your system or another allowed host on your network.
- A too much number of ARP (Address Resolution Protocol) requests can be a sign of an ARP poisoning or spoofing attack on your network. Anyone can run a program, such as dsniff tool or Cain & Abel tool, can modify the ARP tables, which are responsible for saving IP addresses to media access control (MAC) address mappings on network hosts.
- That makes the victim machines to think they require to forward traffic to the hacker's computer rather than to the correct destination machine when communicating on the network. And this is a type of man-in-the-middle (MITM) attacks. Spoofed ARP responses can be sent to a switch, which returns the switch to broadcast mode and basically turns it into a hub. When this happens, a hacker can sniff every packet going through the switch and capture anything and everything from the network.

ARP spoofing

- ✓ An excessive amount of ARP requests can be a sign of an *ARP poisoning* attack (or *ARP spoofing*) on your network.
- ✓ What happens is that a client running a program such as the UNIX-based dsniff or the UNIX- and DOS/Windows-based ettercap can change the ARP tables the tables that store IP addresses to *media access control (MAC)* mappings on network hosts.
- ✓ This causes the victim computers to think they need to send traffic to the attacker's computer, rather than the true destination computer, when communicating on the network. This is often referred to as a Man-in-the-Middle (MITM) attack.

MAC-address spoofing

- ✓ MAC-address spoofing tricks the *switch* into thinking you (actually, your computer) are someone else. You simply change your MAC address and masquerade as another user
- ✓ You can use this trick to test such access control systems as your IDS, fire-wall, and even operating-system login controls that check for specific MAC addresses.

Countermeasures(MAC-daddy attack)

- ✓ A few countermeasures on your network can minimize the effects of a hacker attack against ARP and MAC addresses on your network.

- You can prevent MAC-address spoofing if your switches can enable port security to prevent automatic changes to the switch MAC address tables.
- No realistic countermeasures for ARP poisoning exist. The only way to prevent ARP poisoning is to create and maintain static ARP entries in your switches for every host on the network. This is definitely something that no network administrator has time to do.

Detection

- ✓ You can detect these two types of hacks through either an IDS or a stand-alone MAC address monitoring utility.
- ✓ Arp watch is a UNIX-based program alerts you via e-mail if it detects changes in MAC addresses associated with specific IP addresses on the network.

6.1.8 Wireless LAN:

- A wireless LAN (or WLAN) is one in which a mobile user can connect to a local area network (LAN) through a wireless (radio) connection. The IEEE 802.11 group of standards specify the technologies for wireless LANs 802.11 standards used the Ethernet Protocol and CSMA/CA (carrier sense multiple access with collision avoidance) for path sharing and include an encryption method, the Wired Equivalent Privacy algorithm.

- **Implications of Wireless Network Vulnerabilities**

✓ WLANs are very susceptible to hacker attacks even more so than wired networks are.

✓ They have vulnerabilities that can allow a hacker to bring your network to its knees and allow your information to be gleaned right out of thin air.

✓ If a hacker comprises your WLAN, you can experience the following problems:

1. Loss of network access, including e-mail, Web, and other services that can cause business downtime
2. .Loss of confidential information, including passwords, customer data, intellectual property, and more.
3. Legal liabilities associated with unauthorized users

- Most of the wireless vulnerabilities are in the 802.11 protocol and within wireless access points the central hub like devices that allow wireless clients to connect to the network. Wireless clients have some vulnerability as well.
- Various fixes have come along in recent years to address these vulnerabilities, but most of these fixes have not been applied or are not enabled by default.
- You may also have employees installing rogue WLAN equipment on your network without your knowledge; this is the most serious threat to your wireless security and a difficult one to fight off. Even when WLANs are hardened and all the latest patches have been applied, you still may have some serious security problems, such as DoS and man-in-the-middle attacks (like you have on wired networks), that will likely be around for a while.
- Common Wireless Threats
  - There are a number of main threats that exist to wireless LANS, these include:

✓ Rogue Access Points/Ad-Hoc Networks

✓ Denial of Service

✓ Configuration Problems (Mis Configurations/Incomplete Configurations)

✓ Passive Capturing

Wireless Network Attacks

- Wi-Fi networks can be vulnerable to a variety of different attacks. Because of this, it's important to be aware of them so you can take the necessary steps to prevent and reduce their impact.
- Different kinds of attacks are Encrypted traffic, Rogue networks, Physical security problems, Vulnerable wireless workstations, Default configuration settings

✓ **Encrypted traffic**

- Wireless traffic can be captured directly out of the airwaves, making this communications medium susceptible to malicious eavesdropping.
- Unless the traffic is encrypted, it's sent and received in clear text just like on a standard wired network.
- On top of that, the 802.11 encryption protocol, Wired Equivalent Privacy (WEP), has its own weakness that allows hackers to crack the encryption keys and decrypt the captured traffic.

✓ **Rogue networks**

- Watch out for unauthorized Access Points and wireless clients attached to your network that are running in ad-hoc mode.
- Using NetStumbler or your client manager software, you can test for Access Points that don't belong on your network.
- You can also use the network monitoring features in a WLAN analyzer such as AiroPeek.
- Walk around your building or campus to perform this test to see what you can find.
- Physically look for devices that don't belong a well-placed Access Point or WLAN client that's turned off won't show up in your network analysis tools.
- Search near the outskirts of the building or near any publicly accessible areas.
- Scope out boardrooms and the offices of upper level managers for any unauthorized devices. These are places that are typically off limits but often are used as locations for hackers to set up rogue Access Points.

✓ **Physical-security problems**

- Various physical-security vulnerabilities can result in physical theft, the reconfiguration of wireless devices, and the capturing of confidential information.
- You should look for the security vulnerabilities when testing your systems such as Access Points mounted on the outside of a building and accessible to the public,Poorly mounted antennas or the wrong types of antennas that broadcast too strong a signal and that are accessible to the public.

- You can view the signal strength in NetStumbler or your wireless client manager.

✓ **Vulnerable wireless workstations**

- Wireless workstations have tons of security vulnerabilities from weak passwords to unpatched security holes to the storage of WEP(Wired Equivalent Privacy) keys locally.
- One serious vulnerability is for wireless clients using the Orinoco wireless card.
- The Orinoco Client Manager software stores encrypted WEP keys in the Windows Registry even for multiple networks.

✓ **Default configuration settings**

- Similar to wireless workstations, wireless Access Points have many known vulnerabilities.
- The most common ones are default SSIDs (Service Set IDentifier) and admin passwords. The more specific ones occur only on certain hardware and software versions that are posted in vulnerability databases and vendor Web sites.
- The one vulnerability that stands out above all others is that certain Access Poinits, including Linksys, D-Link, and more, are susceptible to a vulnerability that exposes any WEP key(s), MAC(Media Access Control) address filters, and even the admin password! All that hackers have to do to exploit this is to send a broadcast packet on UDP port 27155 with a string of gstsearch.

6.2 Operating system Hacking:

- An operating system is a program that acts as an interface between the software and the computer hardware. It is an integrated set of specialized programs used to manage overall resources and operations of the computer. It is specialized software that controls and monitors the execution of all other programs that reside in the computer, including application programs and other system software. Many Operating systems are available now days.
- Many security flaws in the headlines aren't new. They're variants of vulnerabilities that have been around for a long time in UNIX and Linux, such as the Remote Procedure Call vulnerabilities that the Blaster worm used.
- You've heard the saying "the more things change, the more they stay the same." That applies here, too.
- Most Windows attacks are prevented if the patches were properly applied. Thus, poor security management is often the real reason.

6.2.1 Windows:

✓ The Microsoft Windows OS is the most widely used OS in the world.

✓ It's also the most widely hacked, because Microsoft doesn't care as much about security as other OS vendors? The answer is no.Numerous security mistakes were unnoticed especially in the Windows NT days but because Microsoft products are so pervasive throughout networks. Microsoft is the easiest vendor to pick on, and often its Microsoft products that end up in the

crosshairs of hackers. This is the same reason for many vulnerability alerts on Microsoft products. The one positive about hackers is that they're driving the requirement for better security!

- ✓ There are variants of vulnerabilities that have been around for a long time in UNIX and Linux, such as the RPC vulnerabilities that the Blaster worm used. Most Windows attacks are prevented if the patches were properly applied. Thus, poor security management is often the real reason Windows attacks are successful
  - Much vulnerability have been published for windows operating system.
  - Some of the common vulnerabilities found in all versions of windows are: DoS, Remote Code Execution, Memory Corruption, Overflow, Sql Injection, XSS, Http Response Splitting, Directory Traversal, Bypass something Gain Information /Privileges, CSRF File Inclusion etc.
  - The maximum number of vulnerabilities detected were of Gaining Privileges by which the confidentiality and integrity was highly impacted.

- **Windows Vulnerabilities**
  - ✓ Due to the ease of use of Windows, many organizations have moved to the Microsoft platform for their networking needs.
  - ✓ Many businesses especially the small to medium sized ones depend solely on the Windows OS for network usage.
  - ✓ Many large organizations run critical servers such as Web servers and database servers on the Windows platform.
  - ✓ If security vulnerabilities aren't addressed and managed properly, they can bring a network or an entire organization to its knees.
  - ✓ When Windows and other Microsoft software are attacked especially by a widespread Internet-based worm or virus hundreds of thousands of organizations and millions of computers are affected.
  - ✓ Many **well-known attacks against Windows** can lead to

- Leakage of confidential information, including files being copied and credit card numbers being stolen
- Passwords being cracked and used to carry out other attacks
- Systems taken completely offline by DoS attacks
- Entire databases being corrupted or deleted when insecure Windows-based systems are attacked, serious things can happen to a tremendous number of computers around the world.
- Autoplay feature came in Windows XP. This feature checks removable media/ devices then identifies and launches appropriate application based on its contents. This feature is useful for authentic users but is a gateway for an attacker.
- Clipboard vulnerability can allow attacker to get access to the sensitive clipboard data. In windows clipboard is common for all applications. This may lead to

access and modification in the clipboard of all applications in the operating system.

- MS-Windows stores its configuration settings and options in a hierarchical database which is known as windows Registry. Registry is used for low level operating system settings and for settings of applications running on the platform.

6.2.2 LINUX:

- ✓ It is the latest flavor of UNIX that has really taken off in corporate networks.
- ✓ It is the competitor Operating System for Microsoft.
- ✓ A common misunderstanding is that Windows is the most insecure operating system. However, Linux and most of its sister variants of UNIX are prone to the same security vulnerabilities as any other operating system.
- ✓ Hackers are attacking Linux because of its popularity and growing usage in today's network environment, because some versions of Linux are free.
- ✓ Many organizations are installing Linux for their Web servers and e-mail servers in expectations of saving money.
- ✓ Linux has grown in popularity for other reasons, including the following:

- Ample resources available, including books, Web sites, and consultant expertise.
- Perception that Linux is more secure than Windows.
- Unlikeliness that Linux will get hit with as many viruses (not necessarily worms) as Windows and its applications do. This is an area where Linux excels when it comes to security, but it probably won't stay that way.
- Increased buy-in from other UNIX vendors, including IBM and Sun Micro systems
- Growing ease of use.

- **Linux Vulnerabilities**
  - ✓ Vulnerabilities and hacker attacks against Linux are affecting a growing number of organizations especially e-commerce companies and ISPs that rely on Linux for many of their systems.
  - ✓ When Linux systems are hacked, the victim organizations can experience the same side effects as if they were running Windows, including:
    - Leakage of confidential intellectual property and customer information
    - Passwords being cracked
    - Systems taken completely offline by DoS attacks
    - Corrupted or deleted databases

6.2.3 Buffer Overflow Attack:

A **buffer** is a temporary area for data storage. When more data (than was originally allocated to be stored) gets placed by a program or system process, the extra data overflows. It causes some of that data to leak out into other buffers, which can corrupt or overwrite whatever data they were holding.
In a **buffer-overflow attack,** the extra data sometimes holds specific instructions for actions intended by a hacker or malicious user; for example, the data could trigger a response that damages files, changes data or unveils private information. Attacker would use a buffer-overflow exploit to take advantage of a program that is waiting on a user's input.
Every program contains a buffer, but an attacker can follow one of two methods to take it over and begin an attack.
A buffer overflow attack can be:

- **Stack-based.** Your attacker sends data to a program, and that transmission is stored in a too-small stack buffer. Your hacker could choose a "push" function and store new items on the top of the stack. Or the hacker could choose a "pop" function and remove the top item and replace it. That means the hacker has officially inserted malicious code and taken control.
- **Heap-based.** Your hacker corrupts data within the heap, and that code change forces your system to overwrite important data.

Countermeasure & Prevention:

Hackers have been leaning on buffers for years, but new opportunities for their work are appearing. For example, experts say connected devices (including Internet-of-Things elements like refrigerators and doorbells) could be susceptible to these attacks.

Protect your company by following these common-sense steps:

- **Use new operating systems.** It's time to throw out legacy programs with expired support systems. Newer code comes with more protections.
- **Watch the language.** Programs written in COBOL, Python, and Java are likely safer than others.
- **Add space.** Some programs allow for executable space protections. When enabled, a hacker can't execute code inserted via an overflow attack.
- **Lean on developers.** System administrators often complain that developers ignore their bug reports. Be persistent. When you spot a problem leading to a buffer overflow, keep talking about it until someone fixes it.
- **Apply your patches.** When developers do find out about buffer overflow problems, they fix them with code.

6.3 Applications Hacking:

6.3.1 Messaging System:

Messaging systems are those e-mail and instant messaging (IM) applications that we depend on are often hacked within a network. Why? Because messaging software both at the server and client level is vulnerable because network administrators forget about securing these systems, believe that antivirus software is all that's needed to keep trouble away, and ignore the existing security vulnerabilities.

- **Messaging system Vulnerabilities**
  - ✓ E-mail and instant-messaging applications are hacking targets on your network.
  - ✓ In fact, e-mail systems are some of the most targeted.
  - ✓ A ton of vulnerabilities are inherent in messaging systems.
  - ✓ The following factors can create **weaknesses**:
    - Security is rarely integrated into software development.
    - Convenience and usability often outweigh the need for security.
    - Many of the messaging protocols were not designed with security in mind.
    - Especially those developed several decades ago, when security wasn't nearly the issue it is today.
- ✓ Many hacker attacks against messaging systems are just minor nuisances. Others can inflict serious harm on your information and your organization's reputation. The **hacker attacks against messaging systems** include these:
  - Transmitting malware
  - Crashing servers
  - Obtaining remote control of workstations
  - Capturing and modifying confidential information as it travels across the network
  - Perusing e-mails in e-mail databases on servers and workstations
  - Perusing instant-messaging log files on workstation hard drives
  - Gathering messaging trend information, via log files or a network analyzer, that can tip off the hacker about conversations between people and organizations
  - Gathering internal network configuration information, such as hostname and IP addresses
- ✓ Hacker attacks like these can lead to such problems as lost business, unauthorized and potentially illegal disclosure of confidential information and loss of information.

Email Attacks

- Many people rely on the Internet for many of their professional, social and personal activities. But there are also people, who attempt to damage our

Internet-connected computers, violate our privacy and render inoperable the Internet services.

- Email is a universal service used by number of people worldwide. As one of the most popular services, email has become a major vulnerability to users and organizations.
- The following e-mail attacks use the most common e-mail security vulnerabilities. Some of these attacks require the basic hacking methodologies, gathering public information, scanning and enumerating your systems, and attacking. Others can be carried out by sending e-mails or capturing network traffic.
- Different email attacks are email bomb, banner etc.

- **Email Bombs**
  - ✓ E-mail bombs can crash a server and provide unauthorized administrator access.
  - ✓ They attack by creating DoS conditions against your e-mail software and even your network and Internet connection by taking up so much bandwidth and requiring so much storage space.
  - ✓ An email bomb is a form of Internet abuse which is perpetrated through the sending of massive volumes of email to a specific email address with the goal of overflowing the mailbox and overwhelming the mail server hosting the address, making it into some form of denial of service attack.
  - ✓ An email bomb is also known as a **letter bomb.**
  - ✓ Different email bomb attacks are as attachment overloading attack, connection attack, autoresponder attack.

1. Attachment Overloading Attack

- An attacker can create an **attachment-overloading attack** by sending hundreds or thousands of e-mails with very large attachments.
- Attachment overloading attacks may have a couple of different goals
- The whole e-mail server may be targeted for a complete interruption of service with these failures like **storage overload and bandwidth blocking.**

a. Storage overload

- Multiple large messages can quickly fill the total storage capacity of an e-mail server. If the messages aren't automatically deleted by the server or manually deleted by individual user accounts, the server will be unable to receive new messages.
- This can create a serious DoS problem for your e-mail system, either crashing it or requiring you take your system offline to clean up the junk that has accumulated. E.g. 100MB file attachment sent ten times to 80 users can take 80GB of storage space.

b. Bandwidth blocking

- An attacker can crash your e-mail service or bring it to a crawl by filling the incoming Internet connection with junk. Even if your system automatically identifies and discards obvious attachment attacks, the bogus messages eat resources and delay processing of valid messages

Countermeasures (Attachment-Overloading Attack)

These countermeasures can help prevent attachment-overloading attacks:

- **Limit the size of either e-mails or e-mail attachments.** Check for this option in e-mail server configuration options, e-mail content filtering, and e-mail clients. This is the best protection against attachment overloading.
- **Limit each user's space on the server**. This denies large attachments from being written to disk. Limit message sizes for inbound and even outbound messages if you want to prevent a user from launching this attack inside your network.

2. Connection Attack

✓ A hacker can send a huge amount of e-mails simultaneously to addresses on your network.

✓ These connection attacks can cause the server to give up on servicing any inbound or outbound TCP requests.

✓ This can lead to a complete server lockup or a crash, often resulting in a condition where the attacker is allowed administrator or root access to the system!

✓ This attack is often carried out as spam attack.

Countermeasures (Connection Attacks)

✓ Many e-mail servers allow you to limit the number of resources used for inbound connections.

✓ It can be impossible to completely stop an unlimited amount of inbound requests.

✓ However, you can minimize the impact of the attack. This setting limits the amount of server processor time, which can help prevent a DoS attack.

✓ Even in large companies, there's no reason that thousands of inbound e-mail deliveries should be necessary within a short time period.

3. Autoresponders Attack

✓ This is an interesting attack to find two or more users on the same or different e-mail systems that have autoresponder configured.

✓ Autoresponder is that annoying automatic e-mail response you often get back from random users when you're subscribing to a mailing list.

✓ A message goes to the mailing list of subscribers and then users have their e-mail configured to automatically respond back, saying they're out of the office or, on vacation.

- ✓ An autoresponder attack is a pretty easy hack.
- ✓ Many unsuspecting users and e-mail administrators never know what hit them!
- ✓ The hacker sends each of the two (or more) users an e-mail from the other simply by masquerading as that
- ✓ This attack can create a never-ending loop that bounces thousands of messages back and forth between users.
- ✓ This can create a DoS condition by filling either the user's individual disk space quota on the e-mail server or the e-mail server's entire disk space.

Countermeasures (Autoresponder attack)

- ✓ The best countermeasure for an autoresponder attack is to make policy that no one sets up an autoresponder message.
- ✓ Prevent e-mail attacks as far considering perimeter of your network.
- ✓ The more traffic or malicious behavior you keep off, your e-mail servers and clients are better.

- **Banners**
  - ✓ One of the first orders of business for a hacker when hacking an e-mail server is performing a basic banner grab to see whether he can tell what e-mail server Software is running
  - ✓ This is one of the most critical tests to find out what the World knows about your SMTP, POP3, and IMAP servers.
  - ✓ Gathering Information
    - when a basic telnet connection is made on port 25 (SMTP) banner displayed on an e-mail server.
    - To do this, at a command prompt, Simply enter telnet IP or hostname.
    - From that we get what e-mail software type and version of the server is running. This information can give hackers some ideas about possible attacks, especially if they search a vulnerability database for known vulnerabilities of that software version.
    - If you've changed your default SMTP banner, don't think that no one can figure out the version.
    - One Linux-based tool called smtpscan determines e-mail server version information based on how the server responds to malformed SMTP requests.

Countermeasures (Banners)

There is not a 100 percent secure way of disguising banner information.

Following are some banner security tips for SMTP, POP3, and IMAP servers:

- Change your default banners to cover up the information.
- Make sure that you're always running the latest software patches.
- Harden your server as much as possible by using well-known best practices

General Best Practices for minimizing email security risk

The following countermeasure helps to keep email messages as secure as possible: -

- ✓ Use of right software can neutralize many threats such as - Use malware protection software on the e-mail server better, Apply the latest operating system and e-mail application security patches consistently.
- ✓ Use of encrypted messages or messaging system.
- ✓ Put your e-mail server behind a firewall, preferably in a DMZ that's on a different network segment from the Internet and from your internal LAN.
- ✓ Disable unused protocols and services on your e-mail server.
- ✓ Run your e-mail server on a dedicated server, if possible, to help keep hackers out of other servers and information if the server is hacked.
- ✓ Log all transactions with the server in case you need to investigate malicious use in the future.
- ✓ If your server doesn't need e-mail services running (SMTP, POP3, and IMAP) disable them immediately.
- ✓ Email monitoring can detect and block messages sent from compromised accounts.
- ✓ Email filtering can block certain types of attachments that are known to carry malicious content.
- ✓ Secure email client configurations can also reduce the risk of malicious email.
- ✓ Checking to see if the email address of a questionable message matches the reply-to email address.
- ✓ Verifying that URLs in an email go to legitimate websites.

6.3.2 Web Applications:

- Web applications, like e-mail are common hacker targets because they are everywhere and often open for anyone to poke around in.
- Basic Web sites used for marketing, contact information, document downloads and so on are a common target for hackers especially the script-kiddie's types to deface.
- However, for criminal hackers, Web sites that store valuable information, like credit-card and Social Security numbers, are especially attractive.
- Why are Web applications so vulnerable? The general consent is they're vulnerable because of poor software development and testing practices. Sound familiar? It should, because this is the same problem that affects operating systems and practically all computer systems.
- This is the side effect of relying on software compilers to perform error checking, lack of user demand for higher-quality software and emphasizing time-to-market instead of security and stability.

- **Web application Vulnerabilities**

✓ Hacker attacks against insecure Web applications via Hypertext Transfer Protocol (HTTP) make up the majority of all Internet-related attacks.

✓ Most of these attacks can be carried out even if the HTTP traffic is encrypted (via HTTPS or HTTP over SSL) because the communications medium has nothing to do with these attacks.

✓ The security vulnerabilities actually lie within either the Web applications themselves or the Web server and browser software that the applications run on and communicate with.

✓ Many attacks against Web applications are just minor nuisances or may not affect confidential information or system availability.

✓ However, some attacks can cause destruction on your systems. Whether the Web attack is against a basic brochure ware site or against the company's most critical customer server, these attacks can hurt your organization.

✓ Some other **web application security vulnerabilities** are as follows

SQL Injection

- Injection is a security vulnerability that allows an attacker to alter backend SQL statements by manipulating the user supplied data.
- Injection occurs when the user input is sent to an interpreter as part of command or query and trick the interpreter into executing unintended commands and gives access to unauthorized data.

Cross site scripting

- Cross Site Scripting is also shortly known as XSS.
- XSS vulnerabilities target scripts embedded in a page that are executed on the client side i.e. user browser rather than at the server side. These flaws can occur when the application takes untrusted data and send it to the web browser without proper validation.
- Attackers can use XSS to execute malicious scripts on the users in this case victim browsers. Since the browser cannot know if the script is trusty or not, the script will be executed, and the attacker can hijack session cookies, deface websites, or redirect the user to an unwanted and malicious websites.
- XSS is an attack which allows the attacker to execute the scripts on the victim's browser.

Security Misconfiguration

- Security Configuration must be defined and deployed for the application, frameworks, application server, web server, database server, and platform. If these are properly configured, an attacker can have unauthorized access to sensitive data or functionality.
- Sometimes such flaws result in complete system compromise. Keeping the software up to date is also good security.

Directory Traversals

- ✓ A directory traversal is a really basic attack, but it can turn up interesting information about a Web site.
- ✓ This attack is basically browsing a site and looking for clues about the server's directory structure.
- ✓ Properly controlling access to web content is crucial for running a secure web server.
- ✓ Directory traversal or Path Traversal is an HTTP attack which allows attackers to access restricted directories and execute commands outside of the web server's root directory.
- ✓ Web servers provide two main levels of security mechanisms

Access Control Lists (ACLs)

- An Access Control List is used in the authorization process.
- It is a list which the web server's administrator uses to indicate which users or groups are able to access, modify or execute particular files on the server, as well as other access rights

Root directory

- The root directory is the top-most directory on the server file System.
- User access is confined to the root directory, meaning users are unable to access directories or files outside of the root

Countermeasures (Directory Traversal Attack)

- ✓ There are two main countermeasures to having files compromised via Malicious directory traversals:
- o **Don't store old, sensitive, or otherwise nonpublic files on your Web server.**
- The only files that should be in your /htdocs or Document Root folder are those that are needed for the site to function properly.
- These files should not contain confidential information that you don't want the world to see.
- o **Ensure that your Web server is properly configured to allow public access only to those directories that are needed for the site to function.**
- Minimum necessary privileges are key here, so provide access only to the bare-minimum files and directories needed for the Web application to perform properly.

Google Dorking:

Google dorking is a hacking technique that makes use of Google's advanced search services to locate valuable data or hard-to-find content. Google dorking is also known as "Google hacking."

At the surface level, Google dorking involves using specific modifiers to search data. For example, instead of searching the entire Web, users can click on tags like "image" or "site" to collect images or find information about a specific site. Users can utilize other commands like "filetype" and "datarange" to get other specific search results.

Although benign types of Google dorking simply use the resources that are available from Google, some forms of it are concerning to regulators and security specialists because they could indicate hacking or cyber attack reconnaissance. Hackers and other cyber-criminals can use these types of Google Dorking to obtain unauthorized data or to exploit security vulnerabilities in websites, which is why this term is gaining a negative connotation from the security community.

Understanding Google Dorks and How Hackers Use Them:

The idea of using Google as a hacking tool or platform certainly isn't novel, and hackers have been leveraging this incredibly popular search engine for years. Google Dorks had their roots in 2002 when a man named Johnny Long started using custom queries to search for elements of certain websites that he could leverage in an attack. At its core, that's what Google Dorks are – a way to use the search engine to pinpoint websites that have certain flaws, vulnerabilities, and sensitive information that can be taken advantage of. As a side note, some people refer to Google Dorks as Google Hacking (they're more or less synonymous terms).

Google Dorking is a technique used by hackers to find the information exposed accidentally to the internet. For example, log files with usernames and passwords or cameras, etc. It is done mostly by using the queries to go after a specific target gradually. We start with collecting as much data as we can using general queries, and then we can go specific by using complex queries.

Google Dorks can uncover great information such as email addresses and lists, login credentials, sensitive files, website vulnerabilities, and even financial information (e.g., Payment card data).

Google Dorks Operators:

Just like in simple math equations, programming code, and other types of algorithms, Google Dorks has several operators, some of the most common are:

- intitle – This allows a hacker to search for pages with specific text in their HTML title. So intitle: "login page" will help a hacker scour the web for login pages.
- allintitle – Similar to the previous operator, but only returns results for pages that meet all of the keyword criteria.

- inurl – Allows a hacker to search for pages based on the text contained in the URL (i.e., “login.php”).
- allinurl – Similar to the previous operator, but only returns matches for URLs that meet all the matching criteria.
- filetype – Helps a hacker narrow down search results to specific files such as PHP, PDF, or TXT file types.
- ext – Very similar to filetype, but this looks for files based on their file extension.
- intext – This operator searches the entire content of a given page for keywords supplied by the hacker.
- allintext – Similar to the previous operator but requires a page to match all of the given keywords.
- site – Limits the scope of a query to a single website.

6.3.3 Database System Vulnerabilities:

- ✓ Database management systems are nearly as complex as the operating systems on which they reside.
- ✓ As a security professional, there is need to assess and manage any potential security problems.
- ✓ Following are the Vulnerabilities in database management systems
  - ➢ **Loose access permissions.** Like applications and operating systems, database management systems have schemes of access controls that are often designed far too loosely, which permits more access to critical and sensitive information than is appropriate. This can also include failures to implement cryptography as an access control when appropriate.
  - ➢ **Excessive retention of sensitive data.** Keeping sensitive data longer than necessary increases the impact of a security breach.
  - ➢ **Aggregation of personally identifiable information.** The practice known as aggregation of data about citizens is a potentially risky undertaking that can result in an organization possessing sensitive personal information. Sometimes, this happens when an organization deposits historic data from various sources into a data warehouse, where this disparate sensitive data is brought together for the first time. The result is a gold mine or a time bomb, depending on how you look at it.

Best practices for minimizing database security risks

- ✓ While some attackers still focus on denial of service attacks, cyber criminals often target the database because that is where the money is.
- ✓ The databases that power web sites hold a great deal of profitable information for someone looking to steal credit card information or personal identities.
- ✓ Database security on its own is an extremely in-depth topic that could never be covered in the course of one article; however there are a few best practices that

can help even the smallest of businesses secure their database enough to make an attacker move on to an easier target.

Separate the Database and Web Servers

- Keep the database server separate from the web server.
- When installing most web software, the database is created for you. To make things easy, this database is created on the same server where the application itself is being installed, the web server. Unfortunately, this makes access to the data all too easy for an attacker to access.
- If they are able to crack the administrator account for the web server, the data is readily available to them.
- Instead, a database should reside on a separate database server located behind a firewall, not in the DMZ (Demiltarized Zone) with the web server. While this makes for a more complicated setup, the security benefits are well worth the effort.

Encrypt Stored Files

- Encrypt stored files.
- White Hat security estimates that 83 percent of all web sites are vulnerable to at least one form of attack.
- The stored files of a web application often contain information about the databases the software needs to connect to.
- This information, if stored in plain text like many default installations do, provide the keys an attacker needs to access sensitive data.

Encrypt Your Backups Too

- Encrypt back-up files.
- Not all data theft happens as a result of an outside attack. Sometimes, it’s the people we trust most that are the attackers.

Use a WAF

- Employ web application firewalls.
- The misconception here might be that protecting the web server has nothing to do with the database.
- Nothing could be further from the truth. In addition to protecting a site against cross-site scripting vulnerabilities and web site vandalism, a good application firewall can thwart SQL injection attacks as well.
- By preventing the injection of SQL queries by an attacker, the firewall can help keep sensitive information stored in the database away from prying eyes.

Keep Patches Current

- Keep patches current. This is one area where administrators often come up short.

- Web sites that are rich with third-party applications, widgets, components and various other plug-ins and add-ons can easily find themselves a target to an exploit that should have been patched.

Minimize Use of 3rd Party Apps

- Keep third-party applications to a minimum.
- We all want our web site to be filled with interactive widgets and sidebars filled with cool content, but any app that pulls from the database is a potential threat.
- Many of these applications are created by hobbyists or programmers who discontinue support for them.

Don't Use a Shared Server

- Avoid using a shared web server if your database holds sensitive information.
- While it may be easier, and cheaper, to host your site with a hosting provider you are essentially placing the security of your information in the hands of someone else.
- If you have no other choice, make sure to review their security policies and speak with them about what their responsibilities are should your data become compromised.

Enable Security Controls

- Enable security controls on your database.
- While most databases nowadays will enable security controls by default, it never hurts for you to go through and make sure you check the security controls to see if this was done.
- Keep in mind that securing your database means you have to shift your focus from web developer to database administrator. In small businesses, this may mean added responsibilities and additional buy in from management.
- However, getting everyone on the same page when it comes to security can make a difference between preventing an attack and responding to an attack.

References:

1. Hacking for Dummies (5th Edition), Kevin Beaver CISSP, Wiley Publishing Inc.
   ISBN: 978-81-265-6554-2
2. CISSP for Dummies(5th Edition),Lawrence C. Miller, Peter H. Gregory, ISBN: 978-1-119-21023-8
3. http://www.applicure.com/blog/database-security-best-practice
4. https://thecybersecurityplace.com/database-hacking-its-prevention
5. https://www.valencynetworks.com/blogs/cyber-attacks-explained-database-hacking
6. https://www.acunetix.com/websitesecurity/directory-traversal
7. https://www.veracode.com/security/directory-traversal
8. https://www.hackingloops.com/google-dorks

Sample Multiple Choice Questions:

1. SNMP stands for
   a. Simple Network Messaging Protocol
   b. Simple Network Mailing Protocol
   c. Simple Network Management Protocol
   d. Simple Network Master Protocol

2. Which of the following tool is used for Network Testing and port Scanning
   a. NetCat
   b. SuperScan
   c. NetScan
   d. All of Above

3. Banner grabbing is often used for
   a. White Hat Hacking
   b. Black Hat Hacking
   c. Gray Hat Hacking
   d. Script Kiddies

4. An attacker can create an……attack by sending hundreds or thousands of e-mails with very large attachments.
   a. Connection Attack
   b. Auto responder Attack
   c. Attachment Overloading Attack
   d. All of the above

5) “allintitle “ Google dork operator returns

a. results for pages that meet all of the keyword criteria
b. pages with specific text in their HTML title
c. matches for URLs that meet all the matching criteria
d. specific files containing title

6) _____ is a technique used by hackers to find the information exposed accidentally to the internet.

a. Buffer overflow
b. Google Dorking
c. Google Shadow
d. GDPR

7) In _____, your hacker corrupts data within the ____, and that code change forces your system to overwrite important data.

a. Stack Based, heap
b. Stack Based, stack
c. Heap-based, heap
d. Heap-based, stack