HACKING

AI

BIG AND COMPLETE GUIDE TO HACKING, SECURITY, AI AND BIG DATA

HANS WEBER

Hacking AI

Big and Complete Guide to Hacking, Security, AI and Big Data.

Hans Weber

are the owned by the owners themselves, not affiliated with this document.

Table of Contents

Hacking AI

BOOK 1 - Computer Systems and Networking Guide: A Complete Guide to the Basic Concepts in Computer Systems Networking, IP Subnetting and Network Security.

INTRODUCTION

CHAPTER ONE: AN INTRODUCTION TO COMPUTER SYSTEMS AND NETWORKING

CHAPTER TWO: INSIDE THE COMPUTER SYSTEMS AND NETWORKING CONCEPT

CHAPTER THREE: COMPUTER SYSTEMS NETWORK AND SECURITY

CHAPTER FOUR: COMPUTER SYSTEM NETWORK: SETTING UP YOUR OWN

CHAPTER FIVE: IP AND SUBNETTING EXPLAINED

CHAPTER SIX: APPLYING THE CONCEPTS OF COMPUTER SYSTEMS NETWORK

CHAPTER SEVEN: OUTSIDE THE NETWORK NUMBERS: WHAT STILL MATTERS

CONCLUSION

RESOURCES:

BOOK 2 - Hacking: A Quick and Simple

Introduction to the Basics of Hacking, Scripting, Cybersecurity, Networking and System Penetration.

INTRODUCTION

CHAPTER 1: BLACK HAT HACKING

CHAPTER 2: WHITE HAT HACKING

CHAPTER 3: GREY HAT HACKING

CHAPTER 4: NETWORKS

CHAPTER 5: SCRIPTING AND OTHER TOOLS

CHAPTER 6: THE DIFFERENT TYPES OF HACKING AND HOW THEY WORK

CHAPTER 7: HOW TO PROTECT YOURSELF FROM HACKING

CHAPTER 8: CYBERSECURITY AND HOW IT SAVES YOU FROM BEING HACKED

CONCLUSION

ARTIFICIAL INTELLIGENCE

AND LIFE

BOOK 3 - Artificial Intelligence And Life: A Complete Guide to the Basic Concepts in AI, Neural Networks, Machine Learning and Data Science.

INTRODUCTION

CHAPTER ONE: FOUNDATIONS OF AI

Philosophy and Logic

From Life Itself

Psychology and Cognitive Science

Statistics and Probability

From Computer Engineering

CHAPTER TWO: WHAT IS AI?

The Turing Test Approach: Acting Humanly

The Cognitive Modeling approach: Thinking humanly

The Laws of Thought Approach: Thinking Rationally

The Rational Agent Approach: Acting Rationally

CHAPTER THREE: BASIC CONCEPTS IN AI

Data

Information

Knowledge

Intelligence

Artificial Intelligence

Classifications

Association

Decision Trees

Deep Learning

CHAPTER FOUR: HOW MACHINES LEARN

Forms of Learning

Components to be Improved

Representation and Prior Knowledge

Feedback to Learn From
- Unsupervised Learning
- Reinforcement Learning
- Supervised Learning
- Semi-supervised Learning

CHAPTER FIVE: MACHINE LEARNING

Memory-Based Learning

Case-Based Reasoning

Decision Trees

Data Mining and Decision Trees

EPAM

CLS

ID3

C4.5, CART and Successors

Inductive Logic Programming

Neural Networks
- Input layer
- Output layer
- Hidden layer

Unsupervised Learning

Reinforcement Learning

CHAPTER SIX: BIG DATA

Essentials of Big Data

Sources of Big Data
- Public sources
- Private sources
- Building new data from existing data
- Using existing data sources

Statistics and Machine Learning

Role of Algorithms

Strategies for Algorithms
- Symbolic Reasoning
- Brain Modeling
- Evolutionary Modeling
- Bayesian Inference

Training Data Sets
- Representation
- Evaluation
- Optimization

CHAPTER SEVEN: MODERN AI

Games
- Chess
- Checkers

AI at Home

Advanced Driver Assistance Systems

Route Finding Maps

Recommendation Systems

In Medicine

For Scheduling

For Automated Trading

In Business Practices

In Translating Languages

For Facial Recognition

CONCLUSION

SOURCES

BOOK 4 - Big Data: A Complete Guide to the Basic Concepts in Data Science, Cyber Security, Analytics and Metrics.

INTRODUCTION

CHAPTER ONE: BIG DATA AND DATA SCIENCE

History of Data Science

Definition of Data Science

Who is a Data Scientist?

How Data Scientists Increase the Worth of a Business

The Technique of Data Science

Impacts of Data Science

Importance of Data Science

Programming Languages Every Data Scientist Should Know

- Python
- R Language
- Java
- Scala
- SQL
- Julia

Matlab

The Data Science Process

The Future of Data Science as a Carrier Choice

CHAPTER TWO: CYBER SECURITY

Introduction

Most Common Cybersecurity Threats
- Trojan horses
- Man in the Middle attacks

Impacts of Cyber Attacks on Business
The financial cost of cyber-attack
Reputational Impact of a Cyber-Attack
Legitimate Impact of a Cyber-Attack
Psychological Impact of Cyber-Attack
Physical Impact of a Cyber-Attack
Social Impact of a Cyber-Attack

How can Cyber-Attacks be Reduced?
Creating responsiveness of cyber-security within an organization
Investing in cyber safety and cyber-security backup
Keeping up to date with all of the safety arrangements and testing the security measures regularly

CHAPTER THREE: CYBER TECHNOLOGY

Features that Must be Present in a Cyber Technology Platform

Best Cybersecurity Practices for Businesses

The Future of Cyber Security

How AI (Artificial Intelligence) will shape the future of Cyber Security Methodologies

CHAPTER FOUR: ANALYTICS AND METRICS FOR BIG DATA

Analytics and Metrics of Big Data and Data Science

CYBERSECURITY ANALYTICS AND METRICS

CONCLUSION

BIBLIOGRAPHY

Computer Systems and Networking Guide

A Complete Guide to the Basic Concepts in Computer Systems Networking, IP Subnetting and Network Security.

Hans Weber

Introduction

How much do you know about a computer today? It’s not just about knowing the basics. Computer networking is what drives in the world today, and if you are looking for a career as a computer networking specialist, now is the time.

Computer system networks are the backbone of most companies set by today standards. Even more so, it’s the backbone of all major companies successful in recent times.

Anyone who can understand its importance and knows how to exploit them can safeguard their ventures from most of yesterday’s issues. They provide a seamless way to keep logs, inventories, transactions, and payrolls; these networks are also the primary communications channels right now.

With this guide, you’ll acquire all the knowledge necessary for a first grasp on the subject. That knowledge will be the main pillar for anyone looking to implement similar networks; anyone looking to study and specialize in computer system networks will also get the best starting point for their journey.

What can you expect to understand after finishing the content inside this book?

First, you’ll know the most basic concepts: how these networks started, what makes them up, and how you can identify one. Through the first chapters, you’ll learn the fundamental knowledge about them.

Naturally, the book will deepen the knowledge as the reader continues.

After understanding the concept, you’ll move into the most intricate concepts necessary to set them up: IPs, subnetting, and even how to establish a good security system for keeping your network safe.

The book covers both how to set up a secure network as well as what dangers you could face when establishing your systems.

Finally, you’ll learn how these systems apply to different industries and company models.

In the end, the reader will know all they need to understand the necessary steps and concepts necessary to integrate and exploit these systems to their fullest.

Chapter One: An Introduction to Computer Systems and Networking

How did it start?

If we are going to start talking about computer systems and networks as a career, we have to understand its history and how it became an in-demand job today.

"Information technology," better known as IT, is defined as the technology that involves the development, use and maintenance of computer systems, software and networks for the distribution and processing of data. The term goes way back to 1978.

Computers did exist before 1978, but they were mostly used to perform calculations. Since they started to be used to index and sort written information, the term IT was invented.

Today, IT is a quickly evolving field and that, of course, includes the computer systems and networking career inside this branch.

Computer systems and networking. What exactly is it?

The IT career is based on the upkeep, configurations and reliable operation of the computer system, especially the process of multi-user computers and its networks with other users. Nowadays, companies rely on their networks for a lot of their work, so any issues must be fixed quickly and entirely.

The computer systems and network administrator keeps the organization's workflow and its lines of communication accessible - at all times necessary. Besides identifying and solving network problems, computer systems administrators also make updates to all hardware and software they manage, so they're always ongoing.

A computer systems administrator is the main point of contact for an organization's network users when they suffer some technical issues. The specialist also needs to guarantee that every connection in the office is working well and supervise the performance of the Internet optimizing their hardware and software.

A computer system and network administrator also make sure that the overall performance and security of the computers they supervise, fulfil the needs of the network users, without surpassing the company's budget.

The person in charge of this job needs to be ready to find new issues each day and needs to get their knowledge optimized to approach them efficiently.

The importance of the field in the current era

Today, companies look at networks and computer systems as their foundation to work optimally.

They need their hardware to be highly functional and maintainable, they need their Internet and servers stable, and they need someone to do all the networking and cable installation and check-ups constantly around their offices.

As we said before, they are the ones responsible for the configuration, reliable operation and upkeep of computer systems inside the office. They seek to ensure that the uptime, performance, resources and security of the computers they manage to meet the needs of all the users.

Today, companies are very dependent on this type of professionals. They need to check the performance of their systems and need a person available almost 24/7 to guarantee that everything is functioning as it should. If one of their servers are down, or they lose Internet connection, they start to lose money very quickly mainly because they stop their production for a certain amount of time, and there is no way to monetize the operation without it working efficiently with the connections and servers.

A computer system administrator or specialist is one of the most critical professionals inside every office nowadays.

The computer systems and networking administrator is also the one in charge to guarantee that the Internet connection is working correctly and that the mail server is running and processing emails that are being sent and received by all staff within the organization.

If this task isn't well-executed, it might lead to a lot of expensive problems for the company such as serious mistakes in the production and handling of their job to significant

money problems and losses. Here is a list focused on why it is essential to have a computer system and network specialist:

A specialist in this field maintains the operating system of the serves and applications, such as mail services, web services and more. They troubleshoot

any hardware, OS or application-related problems to ensure the whole operation of the company itself.

A computer system and network administrator is in charge of maintaining the network infrastructure, such as routers and switches and fixing network-related problems. They attend every single detail about the networks and cables going around the office and how to keep them and the security of people working there.

They are also in charge of one of the most critical jobs inside an organization, and that is keeping the database system used by the company. In bigger organizations, this is a highly essential task, to secure all the data and to keep it optimized and safe from third parties, and possibly losses.

This type of specialist is the one that coordinates the daily operation of secure systems. They handle the monitoring systems and the running of regular backups. They set up, delete and manage individual user accounts.

They keep every single system updated and working on its optimal conditions.

An overview of the process to become a computer system network specialist

When wanting to become a computer systems administrator, there is no one single way of learning. Educational requirements commonly include getting a bachelor's degree in computer science, web technology or network administration. Therefore, anyone who wants can become a computer systems administrator by self-learning or on-the-job training.

Some employers may need their administrators on computer systems to hold a certificate or proof of training from some specific software. There are some training methods and certification intended for specific IT fields such as Microsoft training and certification for Microsoft-based systems like Microsoft Windows and SQL. Another certification they might be looking for is a Cisco training certification for Cisco networks.

To become a computer system and network specialist or administrator, the first thing to aim for will be education.

While most employers want their network and computer systems

administrators to have a bachelor's degree, others only need a postsecondary certificate.

Many degree programs focus on computer network and system administration. Because administrators work with computer equipment and hardware, a degree in computer engineering or electrical engineering is adequate as well. Programs in this area frequently include classes in networking, computer programming or systems design.

Because this technology is always changing and evolving, administrators need to keep up with all the recent developments. Many specialists in this field keep taking courses throughout their careers and attend information technology conferences to keep updated with the latest technology as well. Some businesses need that administrators have a master's degree in IT.

Organizations generally want their network and computer systems specialists to be certified in the products they use. Certification programs are usually offered from the vendor or vendor-neutral certification providers. These certifications validate the knowledge and the use of the best methods that are needed of the network computer systems administrators. One of the most standard certifications is the one Microsoft offers.

Network administrators can work and study enough to become computer network architects. They can also advance to managerial jobs in IT departments, like computer and information system managers.

Here is a list of essential qualities everyone who wants to get into this career must have:

● **Analytical Skills**. Every computer systems and network administrator or specialist needs to evaluate networks and policies to ensure that they perform reliably and to prepare to new customer's requirements and changing needs.

● **Multitasking skills**. Every computer systems and network aspirant will have to work on many problems and tasks at the same time.

The ability to deliver an excellent job while multitasking is one of the most important items on this list.

● **Communication skills**. A lot of people overlook this quality when it

comes to this career, but its importance relies on the moment when they have to describe the problems and the solutions to non-IT workers.

● **Responsibility**. This is one mind-changing quality everybody should have, but when it comes to a computer system and network specialist, it is crucial. They manage the whole functioning of an entire company and how their production and work is on its optimal condition. Without responsibility, the company can suffer a lot of losses and issues due to lack of conscience from their IT department.

● **Quick Learning**. As we mentioned many times before, the IT world is always changing and evolving and adding new systems to the list. For a computer system and network administrator, it is crucial to be updated with the software and hardware to ensure the best outcome for the organization.

● **Programming skills**. This is a perfect add-on to every computer system specialist, as sometimes you will need to go a little deeper when it comes to working with a particular server or web technology challenge, and the new knowledge will boost your work.

An overview of what makes an excellent specialist profile

What makes the right specialist for a company these days relies on many things, not only education.

It is not a simple profession, but that is what makes it so demanding. The computer system administrator is a crucial figure in any company, and it has always been a high profile when it comes to knowledge and qualities.

A computer system administrator is an essential part of a big or strong organization IT team. The positions do vary from business to business, but they are all responsible for managing the same things and duties.

These IT professionals must work closely with employees to install updates on computers and provide tech support when the problems appear.

Here are some traits that all excellent computer system and networking specialist must fulfil:

Patience. This is a crucial trait of any good systems administrator. Many times, the employees may be unfamiliar with certain computer functions.

When it comes to optimizing or making changes, administrators need to have the patience to lead employees through the different challenges.

Especially, when it comes to hardware or software issues, it can take a while for employees to do it the right way. In this case, the user is likely to be angry or irritable because they don't understand what to do with their computer or why it is not working as it should. A good system administrator needs to be able to respond with a patient and understanding manner to help and resolve the problem, to also minimize employee frustration. Patience goes first when we talk about managing computer issues.

Flexibility. If you are doing a significant company-wide software upgrade and suddenly one of your primary servers fails, a competent system specialist will be able to quickly prioritize and be flexible when a potential hardware crisis arises. They need to know which upgrades to focus on first and which to pause given the current circumstances .

A flexible approach here can help understand the priority route to take when it comes to dealing with your company's problems. When looking for a candidate to take a job like this, it is essential to pick a multitasker. You need someone who is up to the challenges.

Technical Knowledge. Arguably the essential trait of any good systems administrator is a comprehensive understanding of computer equipment, hardware and software. You need to have strong technical knowledge to be able to figure out the solutions – especially when things aren’t working as they should.

Embedded systems administrators typically have professional experience in an enterprise environment, plus numerous relevant certificates in their field. When hiring or interviewing new candidates, asking them questions related to their technical knowledge will determine how much they know about what they do.

Personable Nature. A computer system and network administrator frequently works with other employees. Operations and network specialist need to know how to communicate clearly and manage difficult personalities while staying calm under pressure and tight deadlines.

How do computer systems and networking improve a company?

When running a growing business, we start to understand that quality IT solutions are crucial to a company efficiency. And computer networks are one of the most critical IT solutions. They help the business grow and let employees share ideas rapidly and work more efficiently. It increases their productivity and creates more income for the company.

Excellent computer systems and network administrators also reduce the amount of money that is spent on hardware by creating a computer network and sharing the equipment you already have.

With a professional on your IT team, you also improve storage efficiency and volume, you have the freedom to choose the best computer networking method for your team to display. Hiring a professional computer systems and network specialist also gives you a lot of flexibility.

Information technology has for a long time dominated the industrial segment. Since the inception of microprocessors, this field hasn't witnessed a dark stage. Each year, we witness a significant change in this domain that brings the real world closer to the virtual one. Smartphones, smart televisions, gaming consoles, motion-sensing devices are all wonders of the IT field. In fact, IT has become an integral part of our lives and it is difficult to imagine a world without it.

Information Technology and Business – What's the connection?

Information technology is required by companies to reduce costs, increase efficiency as well as gain dominance over the market. From website hosting and storage of data to strategy formulation and social networking, IT offers a wide array of corporate solutions. Strong integration of IT is done by businesses leaders to accomplish these goals. Nonetheless, there are certain domains under information technology that are trending and expect to grow exponentially in the near future.

Why should you be keeping an eye on these trends?

It is extremely likely that these trends in Information Technology will be the point of focus in the coming years. Basically, these trends refer to those sectors that allow companies to enhance productivity and make their

consumers aware about their range of products and services. Businesses from all over the world will be looking to exploit the potential of these technologies.

Here are the top technology trends that according to analysts will be the game changers in the near future.

Private Cloud: The private cloud is an excellent alternative to public cloud computing as it resolves all the security issues posed by the latter. Consumers of information technology demand more from the services it provides. Wouldn't it be great if a business could reduce the time-to-market and operate in a cost-effective manner? Every company will want this.

As the private cloud is deployed within the company firewall, all the data can be shared among the employees without having to worry about security breaches.

Cyber Security: This continues to be a cause of serious concern among the IT companies all across the globe. It would be dangerous if someone had access to a company's records and data pertaining to tenders. Our economies, nations, corporations are all interconnected. In fact, most organizations survive on the Internet. Compromising with cyber security can have a devastating impact on the global economy. This has become increasingly essential as Internet-based attacks will increase in the coming years.

Enterprise Social Networking: This is the next big thing in the corporate world. Every company would want to market its new products and enhance brand awareness. As more and more people are getting onto social websites, it has become easy to connect with them on social networks. In the coming years, it is anticipated that people will become more comfortable with using such websites and carry out business transactions over the Internet. Businesses that succeed in becoming social organizations have a better run on the market.

Gamification: This is one of the leading trends in information technology. Companies that focus on enhancing user experience are more successful. Gamification employs gaming mechanics, interactive media and social networking to accomplish this.

Gamification is done to deepen connection with consumers so that they interact well with the company. Gaming has for a long time been a very profitable domain and leading edge companies will also be looking to explore its potential.

Chapter Two: Inside the Computer Systems and Networking Concept

The Computer System

A computer system is the intricate union of physical pieces called hardware, and programs or applications called software. A user, or live ware, uses this machine to find, sort, and manage varied information. It should be able to input data from an external source and process this data. Also, if needed, it should be able to convert this information into a format that can be used externally.

The essential hardware elements that it needs are a monitor, a mouse, a keyboard, and a CPU. Depending on the use that the user intends to give it, it can have a significant number of programs, divided between system software and applications software. The keyboard and mouse are input devices that help to introduce information into the computer and manage that information to give it the format intended.

With a computer system, a user can work with different kinds of information: text, images, sounds, videos, etc. It can transform that information into another format. With that, it creates original projects by mixing the different formats into a final project that can be exported or stored inside the computer or an output device. Examples of output devices are printers, recorded discs, and external hard drives.

Networking and Computer Systems

A computer system network is when two or more computers systems are interconnected, creating a net between them. Through this network, a user can share information with other users without using external output devices. A computer network can be done by physically connecting computers using wires, or through a wireless connection using Bluetooth or WI-FI.

In business, having a computer network is very useful. With all the computers systems linked to each other, supervisors can evaluate their employee's work, and managers can organize the information effectively. It

makes it easier to communicate between different departments, to send memos individually or to all the company members at the same time.

Workgroups can coordinate better between them, sending reports daily so they can all know the advances they have made with the project. They can join through the same programs and add their ideas without being in the same physical place. The final project can be sent to the head of the department for their approval.

It is also essential when managing confidential information only known to those who work in the company. It prevents leaks that could put at risk the future of the business, keeping the relevant information to those who manage the keywords and passwords.

IP and Sub Netting

An IP, or Internet Protocol address, is the number that is assigned to a specific computer system inside a network and allows that computer to communicate with others through the Internet. Four numbers divided by dots form it. Each number have between one and three digits that go from 0 from 255.

This address indicates other computers where it is located around the world. Every address is different from the next one, and depending on the specific set of numbers and digits, it can be known as its exact location. Without IP addresses it would be impossible to communicate by email or peruse through the Internet to search for information.

Sub-netting is the action of dividing a network into small sub networks or subnets. It is advantageous when a company wants to add new subnets but doesn't want to get a different network address. At first, the idea of sub-netting was though as a solution for the shortage of IP addresses.

After many years, sub-netting has proved to be essential as a method of reducing network traffic. The networks come in the classes: Class A, B, and C. When sub-netting, the classes are divided into small portions, the subnets. These subnets can also be separated in even smaller codes if needed.

By doing so, the user is allowed to have a work network in the house, without having to get a new one.

Elements inside a Computer System Network

When working with a computer network, it is essential to take many factors into consideration. The correct use of them can make a difference in any business.

- **Computers**: To create a computer system network, you must have at least two computers that will be working together. They don't need to be in the same room, some of them may even be in houses also if it is a company network. This allows the managers to work from their homes instead of moving to an office.

The characteristics of the computer systems don't have to be the same, although it is recommended. That way, you can be sure that all the computers support the programs that you will need to use for your business.

- **IP**: Once you obtain an IP address for your network, you can divide them into subnets, assigning one for each computer of your system. That way, the computers can share information faster and reduce the volume of the broadcast.

Also, when working with a Local Area Network (LAN), it allows it to manage the constraints, such as the maximum number of hosts permitted in the network.

-Security: A computer network needs to have a good security system. To have many computers linked and sharing information can seem risky, especially if the computer is not inside a secure room.

The security measures can be both for the hardware and the software.

You must install passwords and codes hard to crack by an outsider. You should also install security programs that can protect your information from the inside, with firewalls and virus detectors .

Many safe security systems can be downloaded for free or paid. You can also hire an expert programmer that can build a security system customized to your needs.

-Employees: The use that your employees will give to the computer system network is also important. Depending on the kind of business that you

manage, your employees will spend more or less hours in front of the computer.

One of the essential rules that those who work for you must know is that the computer system is not for personal research and use. It is a company asset, so employees are responsible for their behavior. Improper use of the computer can develop into the reduction of the useful life of the asset.

The employees must get familiar with the programs they will be using and follow the correct instructions. Managers will be able to evaluate their jobs from their computers, keeping organized work.

Computer System Networks and Security: An Overview

A secure computer system network is the goal for every trustful business. Company owners and managers need to go to their houses and don't worry about being robbed or hacked during the night.

Correct network security must be able to prevent unauthorized access to the infrastructure, protect from misuse, modification, and destruction of the information that can be shared through the network. It must create a secure platform for the users and the instalment of safe programs.

Once you have installed a security system in your computer network, it should:

- Protect your programs and the web in general.

- Detect when there is an abnormal alteration in the system, as an unauthorized modification to the network.

- Take fast reaction when discovered the intrusion, to avoid damage to the system.

A security system is not only to stop the attacks but to prevent them altogether, making a strategy to cover every aspect of a possible issue.

Here are a few elements that should create a standard security system:

- Control the level of access to the network. Give passwords to authorized personnel only, and limit the different level of access that they can get depending on the job they do.

- Install anti-malware programs that would protect the network from most

viruses, Trojans, and worms. These programs can detect hidden or dormant bugs that could be incubating and infecting all the system. Firewalls work as a preventive measure, reinforcing the work of the anti-malware.

- Securing your emails with programs that will block unsecured mails that may arrive.

- Use of VPN. By permitting authorized communication between the network and the computer, it prevents the entrance of unknown data, blocking if needed.

- Prevent human error by connecting your system to secure clouds, where the information can be preserved safely. That way, you won't need to use external devices that could be lost or misused.

Currently, there are dangerous threats on the Internet that can affect your computer network, and new ones appear every day. The most prominent are:

- Computer Viruses: They are tough to spot, and spread quickly if you're not careful. They can infect computer after computer, hidden inside the mail and downloaded archives. They can corrupt data, fill your computer with spam, even delete your hard drive without recovery.

- Adware and spyware: Both are programs that track your personal information and preferences on your searches; some adware is "inoffensive" and asks permission to be installed. They fill your screen with pop-ups, and your computer may work slower. Spywares are installed without you noticing it, and can steal personal information, passwords, credit card numbers, and more.

- Phishing: This is a dangerous program that retrieves personal information. They come in emails and instant messages that look "legit", and install a malicious malware when clicking the link in it.

- Trojan Horse: As the term indicates, a Trojan horse is a dangerous attacker disguised as a legitimate archive. They usually hide on emails, those with a familiar name for you, so that you won't doubt it before opening it. The Trojan hijacks your webcam and steals sensitive data that you may have safe on your computer.

- Computer worms: This dangerous malware spreads quickly through the

computer contacts, infecting every other equipment connected to them.

These are just a few of the menace that you can find on the web, so it is crucial to managing good security for your computer network.

Advantages of a Computer Network

We use computer networks for social interactions, shopping and much more. A computer network is a handy and valuable tool for centralizing and dispersing the stored information of a type of organization (companies, institutions, etc.). It is so vital in the contemporary world that we use them regularly without even realizing it.

Thanks to computer networks, we can locate all kinds of operations quickly and over long distances. Some of them are:

- Social interactions, teleconferences, video calls.

- Electronic purchase operations and capital movements.

- Data transmission, email and share resources in real-time.

- Transmission of stored audiovisual content.

- Satellite exploration and other surveillance and military recognition technologies.

Disadvantages of a Computer Network

The weak side of a computer network has to do with cyber-attacks, which violate the confidentiality of the information and can lead to dangerous activities. We talk about malicious software (viruses, adware, etc.) or cyber-terrorists (hackers), whose attacks can cause loss of information (and therefore capital), threats to privacy or damage to equipment and software. The world of networks is diverse and complex.

Computer System Networks: Examples of Computer Networks

Here are some specific examples of computer networks:

- **A home network**: Like the WiFi, networks that anyone can install in their own home to serve a couple of computers or cell phones. Its scope will barely exceed the margins of the department.

- **A cybercafé**: The so-called cyber cafes were very popular with Internet penetration, before the arrival of Smartphones. They contain a series of computers that share their Internet connection and are available for public use. They were all framed in an internal network, whose head was the

computer of the local manager.

- **A university campus network**: Also known as Campus Area Networks, these are actually MAN networks adapted to the various buildings and interests of the university community.

- **Internet. The biggest WAN available today**: communicating multiple technological devices over vast distances, from one side of the world to the other. This massive network involves computers everywhere, operating servers and workstations for millions.

Chapter Three: Computer Systems Network and Security

The Importance of Network Safety

A computer system's network can be the improvement that your business needs to expand, but it can also be risky. If you don't install a computer network correctly, there can be leaks that professional hackers can use to their advantage.

A high number of computers connected to the network means more possibilities to get infected, or have a security breach. If one computer is infected with a computer virus or a Trojan, it can damage and corrupt essential files and data. That corrupted information will also be corrupted in the other computers through the network.

There are other menaces that can affect the computers in your office or your home. The advantages of using a network to keep your computers and devices connected through the Internet can also be the biggest threat. They can share everything, and communication between them is fast, which means the threats spread fast too.

You can avoid most of these issues by installing firewalls and passwords to help prevent the access of malware and other dangerous digital programs. You will also need to be ready to fight them if they enter your network and make rules for every employee that will be using the web. They need to remember that it is for the benefit of all.

The Most Common Dangers for Computer Networks

Here are some of the most common threats you can find, what they are and what can they do to your computer systems.

- Phishing: It is a term used when referring to stolen identity. The malicious programs are hidden inside apparently safe emails, webpages, and more recently through messages. Once the user opens the contaminated archive, the programs are installed secretly on the computer or mobile device.

The e-mail usually looks like a piece of legal and urgent information for the user, disguised as a bank notification or a message from a real company. Once the program is installed, the cyber thief can obtain credit card information, passwords to bank accounts. It can even steal pictures that can be used for blackmail or other dark business.

- DoS and DDoS Attacks: The letters DoS stands for Denial of Service. The intention of these attacks is to crash a network, denying services to the users. It can be done by over-flooding the system with traffic, or the hacker can send a corrupted file that forces the network to shut down for maintenance. The attacks are not meant for stealing relevant data or corrupting information. It is more to annoy the victims, making them lose money and time in fixing the problem. The intended victims are usually webpages of well-known companies, banks, commerce, or government, especially if they have very active online pages.

It also affects the frequent users of these websites that may need to reach some information and can't have access to it because of the attack.
There is another version of these attacks called DDoS (Distributed Denial-of-Service) attacks. They are stronger and harder to determine where it comes from because it uses several computers distributed throughout the globe. They infect the intended network with malware, thus overflooding it. The infected computers are called bot, and the cyber attacker gains control of them from a distance, creating what is called a botnet.

- Spyware: Spyware is malicious digital program or malware that infiltrates in the computer without your knowledge. It is one of the resources used for phishing, although it can be used in other ways. It copies personal data, like passwords and bank accounts. It also installs on your computer's hidden software to make changes in the configuration of your security information.

- Trojan horses: As the name suggest, a Trojan horse is a malicious program that infiltrates your computer disguised as a safe file. It can usually come inside an email with the name of a family member or a friend, to give you a sense of security that it is safe to open it.

The Trojan can also disguise itself as advertising, asking permission to access the computer. Real advertising is not malicious; that is why the Trojan horses hide in advertising that looks safe and legit. Once inside, they can corrupt

your data or act as a spyware, stealing valuable information, and even cloning your webcam.

- Computer viruses: It is one of the most common dangers inside network security. As viruses do, they can incubate inside a computer, and spread quickly to all the computers connected to it. Inside a network, it is something to worry about, because it allows the virus to reach the other computers whenever you are trying to send an important file.

They can over-flood computers with spam, or change the configuration of your security system, making the network vulnerable to external attacks. It can also corrupt files from the inside and steal information. The worst of them are the periods of incubation. They can stay hidden and undetected on the computer until activated, and maybe it will be too late to do something by then.

What tools do you have available for securing your network?

If you want to protect your business network, you can use a varied number of tools available for you. The web offers an extensive range of services, and some companies give a guaranteed product so you and your company can be secured.

- Control of the access: When talking about security, you can't only focus on installing programs. Many security breaches can happen because of human error. An innocent or malicious employee can get access to essential data and use it with or without bad intentions, creating leaks.

If you want your network to be secure of these kinds of leaks, you must create secure passwords. They will allow them access to different levels of information depending on the user.

That way, it reduces the number of people that can access particular data, and leave fewer breaches for cyber attackers .

- Analysis of the system behavior: Whenever an application reacts differently to what it should, or software suddenly changes permissions, it will probably mean that some malware or other malicious programs have infiltrated the network. You can install software that will detect these changes and notify you before the damage is too significant to repair.

These kinds of software may not be able to prevent the attack but will help you and your security team to react faster and avoid irreparable problems inside the network.

- Secure your email box: Most of malware and viruses choose e-mail as their facade to infiltrate inside computers. To prevent the entrance of these dangerous threats, you must install a security software focused on the email box. This application will destroy or label the possible harmful files hidden inside the safe emails so that you can avoid them.

- Firewalls: These programs are one of the most used security tools. A firewall works as a barrier between your computer system network and external networks. It blocks the entrance to all unidentified or suspicious information that intends to enter your network, preventing the undesired infiltration of malware.

- Antimalware and antivirus applications: These programs search and destroy the dangerous malware that may be dormant inside a downloaded file.

They scan through the system continuously and send a notification whenever they detect an anomaly inside the software.Together with the firewall, antivirus and antimalware programs act as your digital soldiers to keep your information safe from thieves and attacks. They are active inside the network, so you don't have to worry when navigating through the Internet or when you open your emails. It doesn't mean that you don’t need to be careful when dealing with unknown emails and attachments .

- Business VPN: It stands for Virtual Private Network, and is one of the safest ways to protect your network. It gives private access only to those who you choose to enter your system, and they can do so from anywhere in the world. It is very secure because it is private, and is cheaper than installing a WAN network.

Some companies can offer you a business VPN service. The information and data sent through a VPN network are encrypted from end to end, which makes it harder to hack or steal. With a business VPN, you can also avoid international censorship applied in some countries, so that you can still have access to your business information from abroad.

Physical safety for your computer systems network

Keeping your hardware in good shape is as vital as giving protection to the software. Dust, heat, and liquids can be as dangerous as malware, or even worse. Information could be recovered from a corrupted folder, but there is nothing you can do with a burned motherboard. Here are a few tips on how to prevent physical damage to your hardware.

Cold areas: Computers generate a significant amount of heat. Most of them come with cooling devices, but these are not enough if the room where you keep them is also hot. Keep closed windows and curtains to avoid the sun's heat, and use air conditioners and fans to cool the room .

- Everything clean: These may sound like silly recommendations, but not many follow them. You must avoid eating and drinking in front of the computer and keep your area clean of dust. Bugs and mice can damage the hardware trying to reach the bits of food that may fall on the keyboards. You can spill your beverage on the computer and burn it. The dust can get inside the buttons of the consoles and damage the sensor, so a letter stops working. It can also get inside the cooling fan and affect it, overheating the CPU.

- Give maintenance to the computers: All the machines should be checked by a professional technician every few months. He can update all programs if needed, and also make cleaning of the hardware or recommend the change of a part if it is necessary. Maintenance is essential; it avoids further problems, which can be detected before they happen.

- Depend on the experts: If a problem appears, you should call a professional technician who knows about computer systems. You can try to do it by yourself and follow a tutorial, but you can jeopardize the computers guarantee. A computer system technician has studied to recognize a problem when it appears and will be able to detect the real issue once it checks the computer.

Some companies offer phone support for simple issues, and they can send someone to check if the problem seems to be more intricate.

The same goes with issues with Internet services and network installation services.

It is better to rely on a specialist rather than making things worse, trying to save time and money. The most probable thing is that you will end up spending a lot more.

Chapter Four: Computer System Network: Setting Up Your Own

What do you need?

Setting up a computer system network is very useful in every business. It makes communication between the computers an easier job, sharing the necessary data quickly and safely. Depending on the size of the network you want to create, you will need specific elements. Here is a list of the essential items you will need to start:

- Computer systems: You will need to have at least two computer systems to start a network. These can be a desktop or laptop, and they do not necessarily need to be in the same room. Other devices such as notebooks and tablets can be included in the network as computer systems. The network also connects the computer systems with computer accessories like printers, scanners, etc.

- Handy Tools: For some of the installations, you will need to have some tools. The most important one will be a screwdriver. If you can, use an antistatic wrist straps or be sure to have rubber shoes to prevent electric shocks. If it is a wired installation, you will probably need to have a drill to open holes and insert the wires so you can link the computer systems through the rooms.

- Modem with a broadband internet connection: This element is essential for a wireless connection, but it can also be handy for all the different installations. It provides Internet to the computer systems linked through the network, making it easier to communicate.

- Wireless Router and Ethernet cables: Depending on the kind of installation, you will need wireless routers or Ethernet cables to link the computer systems. For a wireless network, your computers systems should have a wireless network adapter. Most of the portable computing devices already have it installed on their systems, and even some desktops have one. If not, you need to acquire and install this device on your computer systems.

Types of Network

LAN: A LAN, or Local Area Network, is a network connection that is set in a specific area, like a house, an office or a building. Although computer systems don't need to be close to one another, they have to be inside the specific area. Currently, Virtual LAN is becoming more common when talking about setting up a network; wiring can take time and is more expensive.

Pros

- You can share information with other computers and computers accessories quickly. It makes it easier to send the final product to an output source like an external hard drive for storage, or print it.

- You can keep all the crucial data and information in one computer system. If you need to retrieve data from another computer, you can use a password and log in. It saves space and keeps the information secure.

- You can share a program license without having to buy one for each computer system. You don't even need to install the program on all the computers, connect directly to the main network through your device to use it.

- It is easy to install, and you don't need to be an expert. There are many video tutorials and webpages that explain how to make a Local Area Network.

Cons

- The network is limited to an area. You can’t work from home, or check information outside the building.

- You need to install a particular program to set up a LAN, what you need an administrator to keep everything going. It means it is a significant investment to do.

- If data is corrupted, it will probably be damaged in all the computers. An infected computer can infect the rest of them through the network.

WAN: A Wide Area Network allows to set a network without the limitation of area. It works excellently for companies that have stores in different

regions of the city or other cities. You need to have access to the Internet to set up a WAN.

Pros

- You don't have the space limitation. You can have all the pertinent data in a central office, and manage all the other computer systems around the country. It helps communication between other locals, check inventory, and keep order without leaving your workspace or your home.

- It allows more advanced network technologies, setting exclusive passwords and software to make the net more difficult to hack or corrupt.

Cons

- It is costly because it requires specific connection plans, and the installation must be done in every workspace. You will need to pay monthly to have continuous internet service.

- It can be slow because of Internet traffic, and it can be interrupted if there are problems with the antenna that transmit the information.

- It needs continuous maintenance to secure that the software works correctly, and it has to be done by a professional.

MAN:

MAN refers to Metropolitan Area Network. It is a net compound by smaller local area networks (LANs) and is used more like a network for a large area like universities. They can connect to the LAN of every building into a central system.

Pros

- It uses fiber-optic cable and other high-technology bandwidth to allow faster communication between the computer systems. Different departments can share relevant information in a few seconds; a file can be sent to a printer in another building. A massive memo can arrive in all the computer systems connected to the network at the same time.

- It allows economizing in the cost of Internet and other services, dividing the expenses between the users. At the same time, they all share a high-quality internet. It is not as fast as the LAN network connection but is faster

than the WAN network.

Cons

- It is costly to install, and not all companies can offer this service. The technology it uses is the newest in the market. It can't use older installations, like other networks, so everything has to be installed for the first time, which means breaking walls and changing wires. It is a lot of work.

- You need to install first several LAN networks in the buildings you wish to connect, so it takes more time and investment to be done.

Setting up each type

Depending on the network, you may have to hire a third party to make part of the installation, while others can be done by yourself.

It implies knowledge of hardware installation to the search and setup of specific software.

For LAN:

- Once you have all the elements for the installation mentioned above, you need to select the central computer system. If it is the first time connecting the router to the computer, a "wizard" or installation helper program should appear and ask to create a network. If it is not new, go to Control Panel and look for the Network and Sharing Centre and select the option to set up a new system.

- If you want it to be a virtual LAN or if you're going to share the Internet between the computer settings, you then must set up the Wi-Fi. The manual of the router already has instructions for its installation. If you are not going to share the internet, go to the next step.

- Connect all the computer systems and computer accessories with a wire or through the wireless router. Some of the devices, like the printer, will need further instructions. In Control Panel, you will have to select Devices and Printers and click on the Add Printer.

Now you are ready to start sharing information in your small business or

setting a game night with your friends in the house.

For WAN:

- You will need to have broadband Internet service with a company that provides it. These companies already have WAN plans to offer, so check which one would work for the kind of business you have. They will install the equipment in every building that will be connected to the WAN network.

- Connect the router to the WAN. This step usually is already made by the company when they install the service. In case they don't, you must find a router that can connect with that specific WAN circuit.

- Then it is time to connect all the computer systems to the router. You can do it via Wi-Fi, or use Ethernet cables to link the different devices. Do this process in all the stores or buildings that you want to connect to your WAN network, and you will be ready to start your business in other cities.

For MAN:

- To set up a MAN network, you need to have set a LAN network in every building you want to connect into one interface. You will also need to have broadband Internet service, optical fibers, and router devices to link all the LAN networks to a central system.

- Once you have followed the steps of how to install a LAN network in all the buildings, you must select which building will have a central computer system. There you will set the Internet service and the main router that will then interlink with the other routers in the different buildings.

Choosing the network for you

Each computer system network has its advantages and disadvantages. These reside mostly on the distances they can reach and the speed of the connection between the devices.

LAN: The LAN network is the best option for an office building and home installations. It allows setting many different computer systems, computer accessories, and other mobile devices into the same network as long as the devices are inside the area. You can share information faster than with other networks, link software and programs so the computers can work with it

without having to install them in all of them, nor paying a license for each machine. It can be set up without hiring a professional. It can cover up an area from 100 to 1000 meters.

WAN: The Wide Area Network is perfect for business with multiple stores and branch offices. Whether they are in the same city or many cities around the country, it allows to keep track of the management of each store, so all the offices can offer the same service with the same quality. It can have a slower connection, but the information will arrive, making the communication safer. The range of the WAN can even reach to other countries, as far as 100.000 km.

MAN: The MAN network doesn't reach as far as a WAN, but reaches further than a LAN. It is best for universities or big hospitals, where they need to connect the different buildings and departments into one network. The internet connection is not as good as the LAN, but it is faster than the WAN.

The Metropolitan Area Network is the middle option between the other two networks, but it needs LAN networks to work. The range of the Metropolitan Area Network is between 50 meters to 100 km.

Chapter Five: IP And Subnetting Explained

What is an IP address?

An IP address is also known as an Internet Protocol address. It is a logical numeric direction or address assigned to every single computer, printer, router or any other device that is part of a TCP/IP-based networks.

The IP address is the very core component on which the whole networking architecture is built; no network exists without it. An IP address is known as a logical address that is used to identify every node in the network uniquely.

Because they are logical, they can change and vary. They are very similar to what we know as a town or a city because the IP address is that exactly, an address so you can communicate with other nodes or networks.

An Internet Protocol address is the most critical component in the networking phenomena that works to bind the World Wide Web together. This numeric address is assigned to every unique instance that connects with any computer communication network using the communication protocols, like TCP/IP.

IPv4 and IPv6, what do they mean?

IPv4, one of the core protocols for IP protocols today, was interestingly, also the very first version deployed for production back in 1983. We do see a fair bit of IPV4 traffic even today.

IPv4 uses a 32-bit address space, with a limited number of unique hosts.

We have to remember that Internet Protocol version 4 is a connectionless protocol and it operates on a best-effort delivery model. It does not guarantee delivery, and neither does assure proper sequencing or avoidance of duplicate delivery.

IPv4 addresses are represented in any notation expressing a 32-bit integer value.

IPv6, the newest Internet Protocol version 6, is the communications protocol that computers and networks around the world use for location, when accessing the internet.

As the number of Internet users started growing around the world, there was a need for more identification numbers. The IPv6 was introduced by the Internet Engineering Task Force, mainly to address the issue of a limited number of IPv4 addresses available. The aim, eventually, is to replace IPv4 completely. IPv6 was ratified as an Internet Standard on July 14 of 2017.

This Internet Protocol version 6 provides more technical benefits addressing the different network allocation needs and ensuring that there is optimal route aggregation.

These two versions of Internet protocol supported by manual IP assignment, can provide features of security inbuilt or optionally, and both have a Packet Header part. Moreover, both can transmit fragmented packets; both can have broadcasting and functions related to multicasting.

On the differences side, they tend to seem like two different things; ultimately; it is hard to assume that they are the root of the same tree.

IPv4 has a 32-bit address space while IPv6 has 128-bit address space. Also, IPv4 can provide 4.29×10^9 address while IPv6 can provide 3.4×10^38 addresses. IPv4 can support DHCP Address configuration. On the other hand, IPv6 goes a step further and supports auto and remembering address configuration.

While IPv4 does not provide end to end connection integrity, IPv6 can give purpose to end connection integrity, IPv6 does.

The RIP routing protocol does not support ipv6 while IPv4 is. Also, IPv4 is supported by SNMP protocol while IPv6 is not.

IPv6 doesn't have IP address classes while IPv4 is divided by categories like A, B, C, D, E.

IP address: public & private

A public IP address is one that can be accessed all over the Internet. Think about it like a postal address when it is used to deliver the mail to your home. A public IP address is the device address that everyone can see when they are searching for your device (or trying to locate it) online.

If you want to know what the public IP address is, you need to do a few clicks on your computer.

On the other hand, a private IP address is what devices within your private network uses. If you have multiple computers being used at your home, you may want to use private IP addresses to address each computer within your home.

In this particular scenario, when it comes to private IPs, your router gets the public IP address. Each of the devices connected to it is getting an individual IP address from your router via DHCP protocol.

Private IP addresses can only be guaranteed uniquely to an internal network. You also need a static IP address for the computer. Manually entering IP address will not work either.

To be clear, private IPs cannot be connected directly over the Internet like a computer with a public IP can. The situation mention enables an extra layer of security.

Setting up a computer network

These days, almost every small office has a local network and an Internet connection. To set up any computer network for your home or office, follow the following steps:

1. Wired or Wireless. The first thing you need to do is to choose between a wired and a wireless network. Wired networks use an Ethernet over UTP cable and tend to be faster when compared with a wireless network.

Also, wired systems are known for being secure and reliable. A Wired network doesn't work with devices without an Ethernet port like tablets and smartphones, and it is not easy and fast to set up thanks to all the running cables. Wireless networks, on the other hand, are straightforward to set up from the user perspective, and they allow easy access to mobile devices.

2. Components. Today, most offices networks use a wireless network or a mixed one. The main components required to start building an office network are:

- A router or wireless router that you can put pretty much anywhere in the house,

- A wireless access point,

- You also need an Ethernet HUB or Switch,

- cable cat 5 or cat 6 with RJ45 connectors

- Telephone cables with RJ10 connectors.

Also, broadband filters are needed too. For most networks, the wireless router or the Hub which connects the network to the internet will be the main component of the system and many times the only element. The wireless router has all the things you would want to get connected to the internet – including a wireless access point, an Ethernet switch as well as the DSL modem and router - all in the same box.

3. Router Location. The Wireless router will connect to the telephone line, cable or fiber network access point into the office or home.

So, the router needs to be located close to the main telephone socket.

The router provides wireless access and needs to be placed on a central location to get optimal results. Don't hide your router in a cupboard. Avoid installing it beside electronic machines and devices like motors or microwaves .

4. Test your signal. One of the easiest ways of testing your wireless signal strength in multiple locations is to use an app on your phone made to check your connection in various areas. If the signal is not optimal, try moving your router to a different location.

5. Extending the network. If the components are not enough for the space you want to cover, you will need to buy more parts.

6. Setup. To administer your router, you will have to access it via a web browser and login using a username and password that usually comes in the package where the router came. Before allowing devices to connect to your network, make sure you have edit or at least check the setup parameters of your connections.

7. Connect your devices.

What is sub-netting?

Sub-netting partitions a single physical network into smaller sub-networks or subnets. You get to see two segments, a network segment and a host segment.

Subnets were designed thinking about solving the shortage of IP addresses over the Internet.

An organization can use IP subnets for different reasons. For instance, you can use these to expand your network or address the varied physical requirements.

Sub-netting is also used by routers to make routing choices.

The subnet mask

A subnet mask, just like an IP address, has four bytes, and is what complements it.

To set up a subnet mask, we have to remember that it doesn't works as an IP address. Instead, you will see that subnet masks come with an IP address. Yes, the two work together. For a subnet mask to become valid, its bits must be set to 1 on the left side of the subnet mask.

Setting up a subnet: how can you do it?

Some people may not need to set up a subnet if they only have a few computers in their network. Unless you are a network administrator, this process can seem a bit complex, and it is best to hire a professional. Sub-netting works by using the concept of extended network addresses to individual computer addresses. If a small business plans to use a specific network for its internal hosts, they use a default subnet mask. It allows everyone in the network to access the other device easily.

To subnet this network of more than 24 bits, it must be set to 1 on the left side of the subnet mask.

Exploiting subnets inside computer networks

Using subnets can improve network performance and speed. Sub-netting enables you to ensure that the information will stay in the sub-netted network and at the same time, maximize their speed and effectiveness. Sub-netting also reduces network congestion by ensuring that traffic destined for a device within a subnet stays inside. They also boost network security by splitting or

dividing your network into subnets. You can control the flow of traffic using route-maps, enabling you to identify threats, close entry and target your responses quickly.

Chapter Six: Applying the Concepts of Computer Systems Network

Creating a computer network in your home

Creating or setting up your computer network at home is not as hard as it seems. It is quite easy if you put your mind and hard work to it.

The main reason why you may want to create a computer network at home is that you are looking for a better way to handle the Internet connection of multiple devices or computers.

The first thing you need to evaluate is the best type of network for a home.

You have two options, Wire Network and Wireless Network.

A Wire Network is more secure and reliable, but it doesn't work with devices that don't have an Ethernet port like tablets or smartphones, which is not convenient on a home nowadays.

A Wired Network is mainly used for network backbone, like connecting it to routers, network switches and wireless access points on different levels or floors.

When it comes to Wireless Networks, they work through Wi-Fi and are very easy and quick to install. A Wireless Network is generally slower when compared to wired networks.

They are easy to set up and allow easy access to mobile devices like smartphones and tablets, plus you don't have to run cables around your home .

The best type of network to create or set up inside your home will be the wireless network or creating a mixed network structure.

To set up your wireless network at home you will need a router or Wireless router, a wireless access point, Ethernet HUB or switch, some cables cat 5 or cat 6, a telephone cable and broadband filters.

Later on, you will need to figure out the best Wireless router location. It will

need to be connected to the telephone line or fiber network access point in your home. The wireless router needs to be pretty close to the main telephone socket or you can change the location using a longer WAN cable. Try to keep your router out of cupboards, don't install it behind furniture or next to microwaves or motors.

Once you have your router location ready, you will need to test the signal to make sure the area is optimal. If it isn't, you can try extending your network with an additional wireless access point.

The last and most natural step is to set up your home router. You will need the username and password that usually comes in the box to administer it through a web browser or following the instructions on the box. It depends on the model and brand of the router. Once you are finished, enjoy your brand new home network.

But before we relax, we need to think about security and how much we need when we are working with a home network.

Since we are speaking about our home network, where we handle our most personal information, it is important to protect every connection made inside our house.

You can use a few tips to make your network as secure as possible.

First, change the name of your default home network. If you want to make your home network secure, you should change the name of your Wi-Fi network, better known as the SSID or Service Set Identifier. When changing the name hackers or malicious people out, there won't know what type of router you have, and it will make it a lot harder for them to understand how to attack you.

You should also make sure to set up a unique and robust password to secure your network. Having easy to guess passwords is never a good idea. An excellent wireless password should always be at least 20 characters long and include letters, symbols and numbers. You can search for different guides to set up strong passwords online.

The best thing you can do is activate network encryption to improve your Wi-Fi security.

Nowadays, Wireless networks come with multiple encryption languages, such as WEP, WPA or WPA2 and they encrypt all the traffic on your Wi-Fi network.

Creating a computer network for an office

When it comes to setting up the system for your office, the big choice doesn't rely on a wired or wireless connection, but on what you choose to use: switch or router.

A switch connects multiple devices on the same network inside a building. It enables connected devices like computers and printers to share information. Creating a small office network is not very comfortable without a switch to tie all the tools together.

A router, on the other hand, ties multiple networks together. For your office network, you will need one or more routers to help you connect your computers to the internet. With it, you can also connect computers to share one single internet connection. Think about it as a dispatcher.

If you have a rather small office, we still recommend sticking with a router.

Also, to have sufficient Wi-Fi coverage, you will need to have a wireless access point. Inside a small office is a good idea to have more access points with the signal strength turned down than to have one access point turned up.

For security, when setting your wireless access points, you should know that you also can set up a guest Wi-Fi network. Like this, you will only give your guest access to the internet but not to your internal network. It will ensure that your system will be safe from malicious attacks.

When selecting the right router for your office, keep in mind that you cannot compromise on the firewall. It’s what helps a

router filter incoming cyber-attacks on your system. Plus, it’s a good idea to get a VPN too.

Creating a computer network in a large company

Today, every company works with its unique needs and priorities. We need the right network and network security to help team members to work and exchange information seamlessly.

To set up your computer network for your large company, the first thing you need to do is to define your requirements. Check how many devices you will need to connect to the computer network. Also, check the type of files you will be sending over the networks – file sizes do matter.

Look for the right software applications your team members will be using. Check if your employees need data.

The second but not less important step is to figure out if you should go wired or wireless for your business set up. For a large company, we always recommend a mixed network, where you enjoy the security and speed of a wired connection and the flexibility of a wireless network.

If you choose to go wireless all the way, the connection may drop out if you connect too many computers at the same time, and for a large company that will be happening a lot. That is the main reason why so many companies are choosing both wired and wireless network setups.

Once you have made your choice, the team needs to select the right hardware. Every business computer network, for instance, needs a router and a server.

You can choose a wired router as a getaway or wireless as an access point. You can also choose between two servers: either a cloud-based one that stores all your data online or a physical one that stores the content in-premise. The best option for a big company is a cloud-based server. A cloud-based server has a lot more flexibility, especially if you have big growth plans.

How does a Cloud-Baser Server Work?

‘The cloud’ might be a popular term but when it is combined with ‘computing’ things get big and a little complex. With the expansion of the Internet and portable devices, people now look to take their work everywhere. Cloud computing makes your work portable so that you don’t miss out on important projects.

The concept of Cloud Computing

Suppose you are working on an assignment and wish to send it to a team member for proofreading or checking. You wouldn't like to copy all that data into a pen drive or a portable hard disk and go all the way to deliver it, would you? Email can be a useful tool but fails to work when the data is beyond 25Mb. This is where cloud computing kicks in.

Cloud computing involves use of computer resources that are connected with a localized server of desired specifications using a data connection preferably over a wireless network. In order to gain access to the server, one needs to have a dedicated application with every user possessing a password. Apple's iCloud is an example of cloud storage that allows users accommodate data on a centralized server that can be accessed using any compatible device.

The trend of cloud computing is fast catching up thanks it its flexibility in use and the kind of convenience it offers to the consumers.

What makes cloud computing different from traditional server-based systems?

The architecture of a cloud computing system isn't as easy to implement as it appears. The perfect cloud can only be effective if done by the right team of professionals. So, how does this interconnected system work? Most websites and applications run on massive servers that are capable of handling large amounts of data. What makes websites different from a cloud based system is the fact that the cloud utilizes resources of discrete devices to form a large virtual computer.

Cloud computing reduces the dependence on a single hardware or a software resource. This makes it easier for companies to host websites. For instance, if you are hosting your organization's website from a local server that supports only Windows OS, you are tied up to that OS all the time. On the other hand, if the site is being hosted on a cloud, multiple platform programs can be run without any issue.

Implementing cloud computing network – The Architecture

This network comprises of 2 primary components – the infrastructure and the cloud platform. These can also be termed as the back-end and front-end

layers. The back-end layer of a cloud computing network consists of the hardware, memory as well as software that is used to run it efficiently.

The front-end of the network has the cloud platform which is where the users interact with the entire network. The cloud platform is basically a web-based application with multiple utility options.

Pros and cons of Cloud Computing

Cloud computing provides an excellent platform to share information without having to compromise of resources besides being cost-effective.

However, some IT professionals believe that deployment of cloud computing may cause serious security breaches. Nevertheless, these issues can be dealt effectively with the help of deterrent, preventive and corrective controls.

Cloud Services for your business

There are a truckload of companies offering you cloud storage for free. While many cloud based services like Google Drive and Dropbox offer some GB's of storage for free you can even subscribe to their monthly plans – to suit your business needs. Organizing files, sharing data and working together become a whole lot easier when you use the cloud technology.

When choosing the right routers, opt for the ones that allow you to enable your VPN server and provide employees with safe remote access. Also, you can select one that includes added security features like anti-spam features.

Keep in mind that while you and your team are choosing for a database system, you need to define your looks and remember what type of business is going to use the network setup.

To protect your business network, make sure to use a WPA2 or an encrypted protocol for passwords on the router.

Another way to protect your business network is by disabling or restricting the DHCP. The DHCP defines what IP address the devices on the system will have.

Moreover, make sure to use a VPN or virtual private network to encrypt the internet connections and data transferred through your system.

To have an even more protected network, always update router firmware so your business will not be vulnerable to attacks due to outdated router firmware.

To ensure the security of your network, also disable the file sharing option. This should only be enabled on file servers. If you don't do it like this, the files that your team shares will be seen by every user on the same Wi-Fi connection.

Creating a computer network in a school

When it comes to creating a computer network for a school, we need to think about computer labs and a lot of database systems. It is highly relevant to set up everything right so the educational environment will be secure and working correctly.

If we are going to set up a computer network for a school, the best we can do is choose a wire connection. This way, we can ensure that the connection will be secure and fast, and at the same time, the kids won't be stealing the Wi-Fi passwords and connect to the network with their smartphones.

In these cases, it is also better to use switches instead of routers. So, we can make sure that every device is connected correctly to each other.

When choosing the right database system for a school network, it is essential to check the following points: the integrity of the data, the performance and the ease of maintenance.

If you choose a server-based database system, you will have built-in protection against corruption on your data and files.

Shared-file databases are slower than a server-based system, mainly because each user is reading the whole data over the local area network. So, it is better to go with a server-based system which enables the server to compute and return the answer quickly without pushing large chunks of data over the network.

About the security measurements, make sure to enable the WPA2 or encrypted system, make sure to allow a VPN and protect at maximum the connection of the network.

It is essential to install a web application firewall for extra protection and

always update the router firmware to avoid attacks.

When it comes to school networks, it is a must to disable the file-sharing options and only use it on file servers. So, the information can be secure and not on every single computer inside the network.

Chapter Seven: Outside the Network Numbers: What Still Matters

The outside can still harm your networks

Network security shouldn’t be focused only on the internal threats that can affect the software and corrupt or steal all the relevant data inside your computer systems. There are harmful external elements that can also damage the physical components, especially to the infrastructure of the network.

These threats must not be taken lightly. Virus and malware can be avoided using diverse programs, but a sudden blackout can burn the computers and all-electric elements of the network. A flood on a room can damage the cables inside the walls of the building or house.

External elements can affect either LAN, WAN or MAN networks all the same. Bolts of lightning during a storm can damage the antenna that connects the buildings between them, breaking the link that keeps the network together. Birds on the roof or mice and bugs in the walls can also be a potential threat.

And you must never forget the human factor as a possible problem for your network. Humans make mistakes and can damage the network in unexpected accidents. There are ways to protect your computers system network from most of these possible issues, the same way you can protect it from internal problems.

The location of your network

Currently, most of the LAN networks and all of WAN and MAN networks depend on a considerable measure of Internet connection. Without it, you can’t link computers far away from each other. Many companies can offer this service, and even some of them have special plans for WAN networks. They give exciting prices depending on the number of stores you will connect.

These companies are called Internet Service Providers or ISP. Depending on your country you may have a handful of them, some of which offer other

services like phone service and satellite television connection. The company provides service through ADSL cables that go around the city, or an antenna installed on the roof of your building or house.

The antenna signal can be affected by other signals, the microwaves crashing to each other and making it difficult to connect. Your broadband internet connection can also be changed depending on how far from the primary source you are, or how many repeater antennas are in the area.

The most common problems are slow speed on your Internet and intermittent signal, which makes the network unreliable. If the computer systems are too far from each other, the connection could not be as secure.

Choose the location for your network carefully, and inform yourself about the best ISP companies you can find whether they are local or a widely known company.

Buildings and construction materials

You may have found that the potency of the Internet signal may vary in different rooms of your house or building. These problems are probably because of the materials of the walls and roof of the structure. They prevent the signal to reach effectively to every room. This issue is not suitable for an operative computer network, forcing you to choose specific places to place your computers and mobile devices to work correctly.

One of the most common materials for construction is concrete, which is used for a strong foundation for houses and buildings. But concrete is also the material that blocks Internet signals the most. It is the reason why is hard to have Internet connections in basements and underground garages. Concrete comes in different versions, like reinforced concrete that has metal in their composition. The thicker the concrete is, the harder it will be for the internet signal to reach the computers.

Clay brick wall constructions can also block a certain amount of signal but is not as strong as the different kinds of concrete. Masonry blocks are also used very frequently. Although they don’t block the signal as the concrete, they do prevent the free movement of internet signal.

Lumber walls block less signal than clay bricks, making cabins and other

wood structures a good option too. Glass blocks a little signal and that is why the Wi-Fi connection on mobile devices work better close to a window.

Finally, plywood and drywall don’t block any signal, making them the best materials for walls inside an office building.

The best way to secure a good computer system network is to place the router in a middle point inside the house or building that can reach many rooms at the same time. Also, you can add some routers in the other places that can repeat the signal to those rooms with thick walls and that are located far away from the main router.

For LAN connection this problem can be solved using Ethernet cables instead of a Wi-Fi connection in the interior of the house or building, which is faster. That way, you can put your primary computer on the basement, making an internal cable installation that will allow the computer to connect to the upper floors. This measure allows a safe connection, making it harder for other external problems to reach it.

Weather

ISP companies can offer one of the following services, or both of them: internet by ADSL cable, or by satellite signal. ADSL cable internet is faster than satellite Internet, and it connects the network directly between the computers. At the same time, these cables are more likely to suffer from external elements like rain and strong winds that will make your connection unstable.

Cold and rainy environments can affect the electromechanical switches and breakers, provoking blackouts and damaging the electrical machines. Hot weather can also affect the hardware and electrical element of a computer network, overheating the cables and melting metal components. It can provoke dangerous electrical short circuit.

Although most of the ISP work with copper wire cables, some are starting to use optical fiber cables.

They are safer and avoid environmental damage in a more effective way. But these cables are not cheap so that the service can be pricier than other companies .

One of the most dangerous environmental threats are lightning during a storm that can fry all the circuits. They do not even have to reach the antenna to affect the network. Lightning is attracted to magnetic material, like copper wire cables. The best you can do during a storm is shut down all your electrical elements, unplug them, and wait until it is over. You shouldn't risk it because the physical aspects of your computer network are very delicate, and this kind of problems can give you quite a headache.

The best weather for a computer network is a beautiful cloudy day without a hot burning sun in the sky, neither strong winds that could affect the cables. Storms should be avoided at all costs, taking the correct measures to prevent damage to the network. You can't avoid the possible problems to the wired system in the city, but you can protect your assets the best you can. That way you won't have to buy a new one, and remember that information can't be recovered from a burned motherboard.

Human interaction

Finally, the most common and frequent threat that all computer networks must face is human error. People can be negligent and take everything for granted. Many haven't received correct education on how to interact with computer systems, or don't care about the consequences. Whether it is done unintentionally or with malice, there is danger in the contact of humans with the network.

Forgetting to protect your equipment from environmental problems is considered a human error. In practice, all the issues mentioned before are somehow related to humans. This and many other problems are part of the lack of education that many workers and employees usually have. Some may have never worked with a computer before or are used to be less careful with their computers. Here are some of the most common problems that can happen because of human interaction:

- Downloading viruses and malware: The security system can detect, notify, and put in quarantine the files and data that may be infected with viruses or that have malware attached. It is up to the user if they decide to open the file anyway. Security breaches are frequently done by employees that open e-mails and click on fake links, ignoring the warnings that the computer system

gives.

Some employees use the work computer as their computer, getting in unsafe websites, or opening their mailbox during work hours. This behavior allows the viruses to spread through the network, affecting all the computers connected to the infected one.

- Changing the settings: Some people like to mess with the configuration on the computer trying to make it work as they want without really knowing what they are doing. Changing the settings can affect the functionality of some programs and software, which can affect the data inside the network or provoke leaks and security breaches.

An excellent way to avoid this problem is allowing only administrators to access the settings panel, putting passwords to prevent unauthorized entrance. Still, it is best to give the appropriate warnings to the employees about the settings. A professional technician should do any change and installation of software.

- Liquids accidents: People are so used to computers as a daily tool that they forget how delicate they are. It is a commonplace to see employees eating and drinking in front of their computers, unaware that they could have an accident and spill the drink on the keyboard or worse, in the CPU.

These accidents can result in the loss of a monitor or a keyboard, having to replace the damaged asset. But it can also mean the destruction of relevant data inside the computer system, unable to recuperate it.

- Leave everything on: This error is frequent not only on employees but also managers and owners. Even if it looks like you can save time leaving the computers on once the work hours finish, it is dangerous. A blackout can damage the circuits and burn the networks components. You can use a regulator to avoid most of the damage, but the best to do is turning off and unplugging the computers before leaving.

A good education about computers and how to treat and maintain them should be able to help avoid these problems. It would be a great idea to have a meeting every few months to remember the employees and also the managers and head of departments the importance of taking care of the computers.

Notably, workers must remember that computers are part of the company's network, and not their personal property. They must be extra careful with these essential assets because it is one of the most important tools in the business. Keep the rules in places where they can look at them and keep them in mind. They can even learn to apply these rules in their houses.

Conclusion

The next step is to analyze and digest all the information here provided. Computer system networks can be easy to understand at first; setting up a home network is an easy task—even more so when you're using LAN structures.

However, to fully exploit their advantages, you can apply them to your business, regardless of what it is.

If you have a clothing store, you can use the networks to communicate databases, inventories, store traffic, or even communication between all stores.

Of course, their usefulness increases exponentially as your ventures move towards technology and communications. With that usefulness also comes complexity.

Computer networks provide a plethora of benefits, and these increase with their versatility. That versatility requires a lot of studying and practicing.

Therefore, the best next step is to look for additional resources or a future book delving into the more advanced concepts and procedures. Whether you want to work as a computer systems network expert or simply hire this service, you need to understand how far it can get; then, you need to understand how to get there.

As a last piece of advice, we're leaving additional resources at the end of this book, so you should look them up and internalize all the information contained in them.

For aspiring specialists, the best way to get into the business is to prepare yourself before pursuing a career. That way, the road becomes much simpler, and you can practice beforehand. Entering the professionals market will be less harsh once graduated as well.

For entrepreneurs looking to exploit these concepts, it's also an excellent idea to read further on the topic. That way, you'll know what you want on your personnel; you also skip having to depend entirely on them.

So, that's it: the next step is to learn more and make sure you understand all

the information contained within. Try setting up a network in your house or office as a test!

Resources:

Chapter 1:

https://collegegrad.com/careers/network-and-computer-systems-administrators

https://en.wikipedia.org/wiki/System_administrator

https://www.techopedia.com/definition/25597/computer-network

https://www.genesee.edu/academics/programs/IT/csn/

https://www.bls.gov/ooh/computer-and-information-technology/mobile/network-and-computer-systems-administrators.htm

https://www.careerexplorer.com/careers/computer-systems-administrator/

https://www.careerexplorer.com/careers/computer-systems-administrator/how-to-become/

https://www.careerexplorer.com/careers/computer-systems-administrator/

https://www.bls.gov/ooh/computer-and-information-technology/mobile/network-and-computer-systems-administrators.htm

https://pandorafms.com/blog/how-to-be-a-good-sysadmin/

https://www.business.org/hr/recruitment/traits-look-hiring-system-administrator/

https://pandorafms.com/blog/how-to-be-a-good-sysadmin/

https://www.inspiredtechs.com.au/computer-networking/

https://www.ableone.com/four-benefits-networking-computer-systems-business/

Chapter 2:

https://peda.net/kenya/ass/subjects2/computer-studies/form-1/the-computer-system

https://www.techopedia.com/definition/593/computer-system

https://study.com/academy/lesson/what-is-a-computer-network-types-definition-quiz.html

https://www.britannica.com/technology/computer-network

https://whatismyipaddress.com/ip-address

https://www.techopedia.com/definition/2435/internet-protocol-address-ip-address

https://www.techopedia.com/definition/28328/subnetting

https://www.csoonline.com/article/3285651/what-is-network-security-definition-methods-jobs-and-salaries.html

https://securitytrails.com/blog/top-10-common-network-security-threats-explained

Chapter 3:

https://securitytrails.com/blog/top-10-common-network-security-threats-explained

https://www.theamegroup.com/5-common-network-security-risks/

https://www.cloudflare.com/learning/ddos/what-is-a-ddos-attack/

https://www.csoonline.com/article/2117843/what-is-phishing-how-this-cyber-attack-works-and-how-to-prevent-it.html

https://www.paloaltonetworks.com/cyberpedia/what-is-a-denial-of-service-attack-dos

https://us.norton.com/internetsecurity-how-to-catch-spyware-before-it-snags-you.html

https://www.forbes.com/sites/forbestechcouncil/2018/11/15/what-is-a-business-vpn-and-how-can-it-secure-your-company/

https://business.bt.com/help/guides/vpn-for-business/

https://www.cisco.com/c/en/us/products/security/what-is-network-security.html

Chapter 4:

https://www.quora.com/What-are-the-advantages-and-disadvantages-of-metropolitan-area-networks

http://www.tribuscomputer.com/lan-vs-wan-the-pros-cons-of-each/

https://searchnetworking.techtarget.com/definition/local-area-network-LAN

https://techsolutions.cc/security/guide-wan-vs-lan-vs-man/

http://www.itrelease.com/2018/07/advantages-and-disadvantages-of-local-area-network-lan/

https://www.wikihow.com/Set-up-a-Computer-Network

https://www.broadbandchoices.co.uk/how-to/how-to-set-up-a-local-area-network

https://smallbusiness.chron.com/build-wan-28601.html

https://www.conceptdraw.com/How-To-Guide/metropolitan-area-networks

Hacking

A Quick and Simple Introduction to the Basics of Hacking, Scripting, Cybersecurity, Networking and System Penetration.

Hans Weber

Introduction

If you ask someone what hacking is, they will typically tell you that it is someone that penetrates the security of a system and gains access to it. That is surprisingly not what it always meant. The word "Hacker" was initially used to refer to anyone that was a skilled programmer, but due to popular cultural representations, the definitions have changed over time. So let us dig deeper into some basics of hacking and try to explain the culture, the misunderstandings, and the technicalities.

What is a Security Hacker?

When we refer to a hacker, we are typically talking about a security hacker. A security hacker is a person who can exploit an existing computer or network system and is able to use it for their personal motives. To further understand the motives of hackers, we have to look into the different types of hackers that exist and how they use the information or power that they gain once they have accessed a system.

The Different Kinds of Hackers:

Among the many kinds of hackers that exist, there are three that are popularly depicted through the media and are popularly referred to. They are as follows:

1. **Black hat hacker**: A hacker who has malicious intent
2. **Grey hat hacker**: A hacker who has good intent but hacks without seeking permission
3. **White hat hacker**: A hacker with a good intent that seeks permission before hacking

We will be looking at each of the classifications in detail in further chapters.

Other Classifications of Hackers

There are other ways that security hackers are classified. These include classifications by the skills of a hacker. Some of the following are commonly used skill-based classifications:

1. **Elite Hacker:** The most skilled hacker having extensively exploited systems
2. **Script Kiddie:** A hacker that lacks experience and uses pre-written scripts
3. **Neophyte (Newbie/Noob):** A hacker who lacks both knowledge and experience

How Does a Hacker Hack into a System?

When hacking into a system, a hacker follows a number of steps to ensure that they can enter and use the system as they require. These can be broadly categorized into three separate subheadings, which are as follows:

1) **Network Enumeration/Reconnaissance:** Network enumeration is the first step of hacking. It involves getting acquainted with the system and networks that the victim is using. This usually involves retrieving sensitive data about the network,

including the usernames and information of members that use the network, as well as their email addresses. A hacker typically downloads the entire website. Overt discovery protocols are used for this step of the procedure .

2) **Vulnerability Analysis:** After having carried out a network enumeration, the hacker now knows the people and entities that are a part of a network. The next step is to find the vulnerabilities within the systems that are connected to the network. This allows the hacker to enter the network by exploiting the vulnerabilities that he may have found. Many tools, such as vulnerability or a port scanner, exist to allow a hacker to analyze vulnerabilities within a system quickly. A hacker may also manually test vulnerabilities by looking for automated emails and the email server that is being used by the staff.

3) **Exploitation:** The final step of hacking comes in the form of exploitation. Exploitation refers to overpowering the vulnerabilities to make the software or network act in an inconsistent manner. This is typically the stage that most people refer to as "hacking," and we will look into it in extensive detail when we discuss cybersecurity.

After having actively hacked into a network, a hacker tries to maintain a low

profile. They can do so by accessing accounts that have not been used for a long time, or making an admin account for themselves and trying to blend in. Hackers also typically attack after having changed their IPs and machine codes to ensure that there is no track of their activity left. If no one notices the new staff member on the site,

the hacker has successfully blended in and can continue to do as they wish on the site. That is why it is important to keep track of your staff members and ensure that they aren't "ghost" members.

How to Keep your Network Safe

Now that we know how a hacker manages to access a system or a network, we can logically conclude ways to ensure that the hacker is unable to get into it. The first one is obviously to ensure that your staff members aren't traceable, and vulnerabilities don't exist within the website or network that you run. Unfortunately, that is not as easy as it sounds, and vulnerabilities continue to pop up in all sorts of networks. This includes high-profile tech companies, including Apple and Facebook. That is why it is important always to ensure that you have the updated version of the apps. Updates typically exist to resolve security issues or other bugs.

So you might be wondering, if big-tech firms like those are unable to keep themselves safe, how could you? Well, the answer is pretty simple. While it may not be possible to close off all vulnerabilities, it is possible to train your staff. Ensure that you keep a good check on the members of your network, immediately see the history of any anomalies and take action, and train your staff never to leave their emails vulnerable or ghost accounts standing without informing the management of the network. That way, you will be able to ensure that no hacker manages to exploit any vulnerability that they may find, and all threats are promptly taken care of.

What Can Be Hacked?

It is a common misconception that only systems such as computers can be hacked, and everything else is typically safe. The fact is that anything and everything that is connected to a network can be hacked by exploiting the network itself. While a complete list would be exhaustive, some of the things

that can be hacked include baby monitors, smart TVs, thermostats, printers, and cameras. In a now-famous incident, someone hacked into 50,000 printers and made them rapidly print out messages asking people to subscribe to Pewdiepie.

Can a Network be Un-Hackable?

While there are a number of ways to make your network secure enough to deter hackers from trying to enter it, there are no websites or networks that can prove to be unhackable. Even secure networks such as the NSA have been hacked at one point or another. A popular Reddit thread lists all games that call themselves "unhackable," followed by a general challenge to hack them. All of them eventually got hacked.

Chapter 1: Black Hat Hacking

Now that we know the basics of hacking well, we can dig deeper into the motives of certain types of hackers and what they aim to do by making their way into a system. Any such list obviously starts with the black hat hackers, popularly depicted as villainous typists that work on a black screen, by media.

Motives of Black hat Hackers

To understand the operations of such hackers, it is important that we understand the motives that they have when trying to hack into a system. The motives are usually broadly tagged as personal gain but can be categorized into a number of classes. Typical motives can include the follows:

1) **Blackmailing:** While blackmailing is not an ultimate motive, it is usually what encourages a hacker to hack into the data. By getting access to the data, they are able to blackmail the owner and gain personal benefits. These include financial gains as well as making the person being blackmailed oblige to a request.

2) **Financial gains:** Financial gains can be achieved in a number of ways by black hat hacking. We have already discussed that people can be blackmailed to get money. Other ways of gaining financially include selling the personal that the hacker manages to steal, and working for a third party and getting paid for the hacking services.

Hackers can also hack directly into your bank account and take your funds!

3) **Revenge/Fun:** While typically not the motive of skilled hackers, novice hackers may hack into the data of a person simply for fun, or for some form of revenge.

4) **To practice hacking:** A typical black hacker that is still learning may hack into a website and make it unusable simply for practice, to enable themselves to target bigger and more secure networks in the future.

The Harm of Black hat Hacking

There are a number of harms that we are exposed to when we are targeted by black hat hacking. Contrary to their white hat alters, black hat hackers typically lack moral responsibility and thus take little to no care of how the data that they are hacking is being used.

This means that while a hacker may have hacked the data for an entirely different purpose, having gained access to it, they might just leak it to the internet. Marketplaces in the dark web typically buy such data, making it public knowledge. In a recent attack, someone leaked 773 million emails and 21 million passwords online. With breaches at this massive scale, we should ensure that we keep our passwords secure so that no one can gain unauthorized access to our accounts.

Other than leaking your data online and allowing anyone access to your account,

there are other harms that come with black hat hacking. The first of these would be a financial loss. In the case of being blackmailed, you might choose to pay money against having your privacy being compromised. By hacking into your bank data, they can also easily transfer funds from your account into their own, and thus rip you of your money.

Black hat hacking is particularly dangerous for those who run servers or networks. A black hat hacker would have little care when trying to enter a vulnerable system and can thus use extreme methods to ensure that they get control. That not only means that they might render your site or network useless for the time, but also that they might take over the ownership of the site and use it as they wish, stealing yours, your staff's and your customer's data. That is why it is important that any attempts at hacking are promptly dealt with, and security is tight enough to discourage any hackers.

The Legality of Black hat Hacking

Hacking itself is not illegal. That is because there are a number of ways in which hacking can promote good within society. However, given the work that black hat hackers do, their hacking being non-consensual, their operations are illegal. While laws greatly vary with country, depending on the severity of the case, black hat hackers may find themselves facing years of jail time and thousands of dollars in fine, as well as payment for any

damages, caused. We have seen some cases where hackers have received as many as 90 years of jail time.

Chapter 2: White Hat Hacking

In contrast to the black hat hackers that seek to exploit vulnerabilities of a system for personal gains and break laws in doing so, white hat hackers follow the laws and hack only with consent. That does not, however, mean that the hacking itself is not for their personal gains. Let us explore how these ethical hackers operate and what their motives are.

Why do White hat Hackers Exist?

White hat hackers surprisingly exist because black hat hackers do. If there were no black hat hackers, no white hat hackers would be needed to check the system for flaws and vulnerabilities. White hat hackers primarily get the consent from a network admin and try tap or hack into their network. If they are successfully able to do so, they tell the vulnerabilities within the network to the admin and may offer to help them close any such vulnerabilities. The role of the white hat hacker is the exact opposite of the black hat hacker, and they aim to protect and secure a system. The methodologies that both kinds of hackers use are the same, though.

Motives of White hat Hackers

So you may be wondering why someone would want to hack into a system with consent to find vulnerabilities.

There are a number of reasons that white hat hackers would do so, and we are going to list some of them below:

1) **Securing their Network:** Hackers are typically very skilled programmers and might own their own website or network. They might attack their own network with the motive of finding vulnerabilities and securing them, to ensure that no one else can gain unauthorized access into it.

2) **Financial Gains:** White hat hackers, particularly the good ones, are paid high by companies that wish for the flaws and vulnerabilities in their systems to be closed. Thus by being a white hat hacker, a person can earn a large amount of money through being paid for the services that they offer.

3) **Social Service:** A hacker might wish to provide a social service by raising

awareness about data security. They could thus work with companies for free and show them their vulnerabilities.

4) **Learning:** A newbie hacker may offer their services for free in a bid to learn. If a white hat hacker is able to hack into a complicated and secure system, they are typically offered better payment packages earlier in their careers.

A typical white hat hacker can earn a lot more money than most careers have to offer. Many companies offer thousands of dollars for anyone to identify any vulnerability within their network and might offer more for that person to fix them.

Benefits of White hat Hackers

The benefits of having white hat hackers are many fold, and the most important one is to find vulnerabilities within a system. When a white hat hacker is able to identify the problems within a system, they are quickly fixed. This means that when a black hat hacker tries to hack into the system, he will find that the vulnerabilities no longer exist, and would thus be unable to do much damage.

In an ideal scenario, this should ensure that no black hat hacking occurs. However, there are a number of issues with that. First of all, not many networks hire white hat hackers. That means that they are left open to the vulnerabilities that exist within the system. To add to that, some black hat hackers may be more adept than a white hat hacker and may be able to find vulnerabilities that the ethical hacker missed out on, which means that the system is still open for flaws. That is why we should never let our guard down.

Earnings of a White hat Hacker

A white hat hacker can typically earn a median salary of over 80,000 dollars. Each assignment can earn anything from 15,000 to 20,000 dollars, and the best bounty hunters can manage to earn as much as 500,000 dollars a year.

The Legality of White hat Hackers

White hat hacking is legal, as it does not break any of the laws and is done with consent.

White hat hacking is thus considered a reputable career and can bring you big earnings, and is legal on top of that. If one wants to learn hacking and system penetration, they should try and go for a white hat hacker profession .

Chapter 3: Grey Hat Hacking

Laying between white hat and black hat hacking is grey hat hacking. While typically it was not recognized, it has now become a huge part of the hacking industry and is thus given more notice. Grey hat hacking is technically ethical hacking, and the person engaging in it has no malicious intent and does not intend to steal data or blackmail someone. They also, however, do not seek consent before engaging in the activity and are thus put in a "grey" spot between black and white hat hackers.

Why Would Someone Engage in Grey hat Hacking?

While the motives of most white hat and black hat hackers are clear cut, people in the grey area are harder to read. Nonetheless, there are a number of reasons that we can identify for which someone would want to engage in grey hat hacking.

1) **Financial Gain:** The primary reason that people engage in grey hat hacking is for financial gain. When being unable to find jobs as a white hat hacker, they typically hack into a vulnerable network and tell the company about it, seeking monetary rewards and to be hired to fix the vulnerability. Many companies now tend to report such activities, though, and that means that financial gains have minimalized.

2) **Learning:** When you're a hacker, the whole web is your platform. Learning hackers that wish to engage in ethical hacking in the future might feel that they need to learn by hacking into random sites. While no malicious intent exists on their part, it is still illegal. A number of sites now offer hackers to attempt to hack them for learning purposes, removing the need to learn using illegal means.

3) **Activism:** Perhaps the most widespread usage of grey hat hacking is for activism purposes. This type of hacking is now coined as red hat hacking. We will dedicate a separate section to exploring this form of hacking.

The Legality of Grey hat Hacking

Grey hat hacking is not legal. The legal implications are typically the same as

those in the black hat hacking scenarios. However, given that most grey hat hackers do not seek to harm a site, and at times only wish to help, the incidents are typically not reported at a high rate, which means that grey hat hackers are not always punished. That tends to encourage the activities of such hackers.

The Curious Case of Kevin Mitnick

No story of grey hat hacking is complete without the mention of Kevin Metnick. Having started hacking into systems at a young age of 12, Kevin had managed to hack his way through many high-security systems, including Motorola, Netcom, and Nokia. He had over a hundred spoof cells and codes that he used to hide his location.

He was soon listed as the most wanted hacker and subsequently sentenced to 5 years of jail time.

After having served jail time, Kevin turned over a new leaf and moved from grey to white hat hacking. He wrote multiple books, many of which went on to become bestsellers. He now works with many of the Fortune 500 companies and helps provide them with solutions to their vulnerabilities.

Kevin is now one of the best-known hackers in the world, and his story shows us that grey hat hacking, despite how safe it might seem, is not legal or encouraged. If you feel that you have a knack for hacking, you should go team white!

The Controversial Red Hat Hacking

Red hat hacking is currently branched under the umbrella of grey hat hacking. It is more popularly known as hacktivism and is a prime spotlight of the innovations to hacking in the 21st century.

Hacktivism is typically used as a form of activism, and the hacker uses hacking to draw public interest to a matter of global concern that should be taken note of immediately. Groups of hacktivists exist, including Anonymous.

One of the prime activities that Hacktivists have engaged in is the deletion of sites that display child porn. By gaining access to those sites and deleting their content, they have managed to clean the internet of illegal and immoral content.

WikiLeaks is a depository of documents that were obtained using hacktivism measures, including many state and national secrets that the hackers believe

people have a right to know. As many as 400,000 documents about the US war on Iraq exist on the WikiLeaks server.

Hacktivism is typically illegal since it involves hacking into websites without consent, but it may be considered legal when the hackers hack the deep web to remove problematic content from it.

Chapter 4: Networks

No book about hacking is complete without a mention of networks. If a computer system or data is stored solely in a system with no network connected to it, the hacker would only be able to access it if they were able to access it physically. It is due to networks that hackers are able to enter a system and get access to the data.

What is a Network?

A network is a system in which different entities are connected together. A computer network is no different and consists of interconnected computers that share information with each other.

How are Networks Connected?

Networks are connected and made through a number of ways. Some networks may be connected to each other using the common copper wires that we see being used for the slower internet modems. Other networks are formed through the more solid fiber cables with faster connection speeds. Networks don't need to be hooked together at all and can be formed through all the devices connected to single Wi-Fi.

Types of Network

There are two broad types of networks. The first type is a Local area network, also known as a LAN. The LAN typically connects entities within a small distance and includes simple systems such as those in a company where all the systems are connected to one another. A simple benefit of using the LAN network would be better utilizing the resources. By using a LAN network in a company, you can make your job easier and save space.

How it works is that you would have one computer dedicated to storage. It would store all the files as well as all the software. The other systems can simply use the resources from the storage system, which means that the files and software do not need to be in every computer in the system.

The second type of network is the WAN network, also known as the Wide

Area Network. This is a non-internalized network that connects one entity to outside entities. It thus connects all LANs to one another, letting them pass information between one another. The Internet is a WAN and connects almost all servers in the world in one way or another.

The internet does operate in a way similar to the LAN networks. All the files that you require are contained on the internet, which means that you don't have to go out and locally seek images, text, software, and information. However, access to the internet, or any network for that, does leave you vulnerable to the problems associated with hacking since someone can tap into your network and steal your data.

This is why care needs to be taken with any network that is connected to a network in any way.

Are We Safe if We Never Connect to a Network?

While it is typically not a choice, given how everyone in today's world does need access to resources that you can get from the internet or other sources, let us assume that a computer was hypothetically speaking, never connected to a network. In that case, we would be safe only if the hacker cannot physically access the system. If the hacker is physically able to access the system, he would still be able to enter it using the methods that they typically use.

To demonstrate the point, we'll give you an example of something else that penetrates the system: a virus. The Stuxnet virus was made specifically to target uranium facilities in Iran. While the computers that ran the uranium centrifuges had no active internet connection and were not connected to any outside networks, the virus lay dormant in thousands of devices. Eventually, disaster struck as someone plugged a USB into the system. The dormant virus had affected the USB as well and thus managed to take control of the centrifuges and destroy Iran's uranium and thus nuclear program.

This example shows that the only way to ensure absolute safety is to ensure that the system is never connected to an outside source. With the world now at our fingertips, that definitely doesn't seem to be a viable option!

How does a Hacker Access a Network?

A hacker accesses a network using the steps that we already mentioned in the first chapter. They first find information about the network, then use the information to find vulnerabilities, and then use the vulnerabilities to exploit the system. For clarity purposes, we will give you an example.

Let's assume that there is a server that has 200 active users all logged into the system with access to the files. The hacker would first plan on getting the information on all the users and the system itself. Let's say the system is a website that produces daily content. The hacker would download the site and use a number of tools to analyze who uses the site and what their usernames and emails are.

Once the hacker has that information, he knows that he can now target the users to find a vulnerability in the network. To do this, he might try and crack the password of one of the old unused email accounts that are linked to the site and have access to the files.

After having managed to crack the password (and thus exploited the vulnerability), the hacker would then be into the system and can both get the data or delete it as he wishes. This is why networks are typically very dangerous, and there are a number of security measures employed to ensure that no one can exploit them.

If you have questions over how a hacker can exploit the vulnerability, or how a network is made or can be made secure, then continue reading because we will be digging into the details in the latter half of this book!

Chapter 5: Scripting and Other Tools

Before we dig into the ways that systems can be hacked into in detail, and explain how you can be protected against them, it is important that you know about scripting as well as other tools that hackers typically use. Scripting is one of the main tools in the arsenal of a hacker, and it helps them access the information quicker than they could have otherwise. That makes it important for hackers to be able to script when they are beginning to dig deeper into the world of hacking.

What is Scripting?

Scripting is a way of automating tasks that would otherwise have been had to be written down and coded individually. By running a script, you can basically let the machine do what you would have had to do otherwise. Scripts are typically written in the shell (The black box, so movies didn't get it all wrong!) The help that scripting provides when it comes to hacking is simple. Hackers have to analyze a lot of data to find vulnerabilities or try to exploit them. That would take thousands of lines of code, and mean that cracking into a site would simply not be feasible. With scripts, though, the job becomes way easier. The elite hackers typically make scripts that are used to dig through the system with ease, and with little human input.

What makes Scripts so Dangerous?

What makes scripts so dangerous is how easily available they are. Novice hackers, particularly Script kiddies, typically just use the script to gain access to hacking tools, and can cause DDoS attacks, which can harm companies and cause them thousands in revenue damages. Scripts made to code, or even test, can thus be disastrous in the wrong hands.

Some Script-Based Tools used for Ethical Hacking

Some script-based tools that you can easily find on the internet and use for ethical hacking include:

1) **John the Ripper:** An open-source tool that you can easily download. It is one of the most versatile password hackers and uses intelligent algorithms to

decipher passwords based on the encryption of a system

2) **Metaspoilt:** This tool contains a number of scripts pre-written into it that can scan for and find vulnerabilities within any system, and help you exploit them.

3) **IronWasp:** A multiplatform tool that can search for as many as 25 different web vulnerabilities.

With these kinds of scripts being available for public use, the dangers to any small business sites are typically large. That is why it is recommended that you hire an ethical hacker or a programmer to ensure that all such mainstream vulnerabilities are closed off, and the system does not have to suffer as a consequence.

Is Hacking Easy?

Given that we just told you about a number of tools that can help you hack into systems, find vulnerabilities, and crack passwords, you must be wondering if hacking is easy. The answer to the question is somewhat complicated and mostly depends on a few factors.

If you plan to hack a small recently established site that does not use any security protocols and is completely unprotected, then yes, hacking would ultimately be easy. If, however, you're trying to crack the password of another person on Facebook, Jack the Ripper will be of no use to you. To work around those systems, you can't use pre-made widely available scripts since those vulnerabilities would already have been closed off. You would need to observe the code of the system instead, find any potential vulnerabilities, and custom-make a script to exploit it. So in those cases, hacking is definitely not easy, and we can see why ethical hackers are able to bag so much money.

In the later chapters, we will be discussing how hackers attack and how you can stop them from causing damage to you and your system.

Chapter 6: The Different Types of Hacking and How they Work

Now that we know all about the basics of hacking, we will see how hacking actually works by looking into real-world examples and how they work. This will be an extensive list and should let you know about all the major ways that people can hack into your system or account. In the later chapters, we will also consider how you can protect yourself against all of these methods of hacking, and how systems have been made that work to protect you from them.

Physical Hacking

Perhaps the technique of hacking that we get to hear the least about is physical hacking. That is because, in today's world, physical access points are too secure for most people to be able to break through them. Nonetheless, it remains a valid technique for hacking and one that you need to secure yourself against.

Physical hacking involves physically gaining access to the data. This can be done using a number of methods. If you are a data center owner, hackers can typically climb in through the ceiling or through the air vents that are placed for cooling. Unless these places are carefully secured, they can easily grab the data that they require physically. In cases of a company, hackers can typically masquerade an employee and enter the company, physically hacking into the system and easily being able to intercept and gain access to any data that they wish to.

Another way that a person can physically hack data is by tapping into the lines that connect you to the outside world. While it was much easier to do in the past, and a hacker could have managed to eavesdrop on your phone call had they access to the wire cord, now it has become an increasingly difficult task due to a number of reasons that we will discuss in the later chapters.

Brute Force

While being one of the most common attacks in the past, brute force is no

longer used for a lot of websites. Nonetheless, for those that have not secured themselves, it remains a goldmine. A brute force hacking script works in a simple manner. It has a database to go with it that contains any and all possible combinations that could be used as a password on a website. That's trillions of entries!

The brute force command module then begins to enter the passwords one by one into the system. Each time a password is not accepted, it will move on to the next one until it is eventually able to find the correct password. Such an attack obviously takes time to execute, but given the power of computers today, it can be done relatively quickly.

Another innovation in the brute force technology is the usage of smarter scripts. The scripts now don't check every word on

the list, but rather check only the passwords that would have been considered valid.

Let's consider that a certain website has the following requirements for the password:

- At least 8 characters in length
- At least one capitalized letter
- At least 1 number
- At least 1 symbol

In such a case, the newer and smarter Brute force script would not start from the typical list, but would rather make a specific list of all possible combinations given the conditions. The first item on the list could thus be @@@@@1aA.

Another improvement has come in the form of how the Brute force attacking module starts to input possible combinations. While initially, it used the list in alphabetical sequence, the program now runs the script to check for common passwords first, greatly shortening the time that it takes for a script to find a password.

Phishing

Phishing is a popular form of hacking, and one of the simplest ones for the hacker. Phishing attacks are now taking more sophisticated forms as well and might be given the name of Smishing.

The basic concept of all these attacks is the same: the person that is trying to hack you pretends to be someone that you trust to get your information out of you. This can be done in a number of ways.

The first method comes in the form of appearing to be a website that they aren't. This is done by naming their website's name closely on the website that you actually wished to visit, i.e., facebo0k.com instead of facebook.com.

The second step comes in the form of getting people to visit the website. They might offer in-app advertising which would offer some promo and people might click it. Once people click on the ad, they are redirected to the fake website that the hacker has made. The URLs are pretty similar, and the website design is exactly the same. The coding is very different, though, and once someone enters the password into the fake website that the hacker has made, the password is saved in a database. Thus while people think that they entered information on a legit website, all the websites would have done would be stealing their data.

The second type of Phishing, popularly called Smishing, refers to the method in which the hacker would try to act as if they're your bank or any other service via email or SMS. It is very easy to spoof your email and make it seem legitimate. This is because the email core does not verify the names of the sender, and people can send an email with whatever email they wish in the sender's name.

So in other words, you could compose an email and put whatever you want in the "Sender" field, if you know how to do so. The hackers that use this technique usually know some information about you. They might, for example, know what bank you are a customer of. They would then send an email that would have your bank's official email address in the sender field and ask for information such as your card number and pin. That information is then used for fraud.

The third type of phishing is done over the phone and known as Vishing. There are two particular methods that are used in this. The first one involves

spoofing the caller ID. This is very similar to email spoofing and allows the hacker to show their phone number as that of a legitimate institution. They use them combined with automated answering machines to steal information.

The second method used is easier if the victim is using a phone line. The hacker calls the victim and tells them of fraud and asks them to call their bank to confirm. They then pretend to hang up, playing flat tunes to indicate that the phone has been hung up. In reality, they're still on the other side of the line. The victim then calls the bank, and the hacker pretends to be the bank and obtains sensitive information that is later used for credit card fraud.

Cookie Theft

One of the more complicated ways of hacking into a system involves cookie theft. This is more complicated than most hacking mechanisms. Cookie theft involves stealing the cookies of a system.

The cookies of a system are authenticating the information that is used to authenticate a person for website usage .

This can be done in a number of ways. Some of the ways include session fixation, Sidejacking, malware, and cross-site scripting. Session fixation refers to when the hacker sets the session to an id that is known to him. He does that by sending a specific link. When the user uses that link to log into a session, the hacker is able to steal the cookies. Sidejacking involves stealing the cookies using the Wi-Fi connection. Many websites do not use SSL certificates on their site, and any data sent can thus be sniffed from the Wi-Fi connection. Cross-site scripting involves tricking the victim's computer into running a script that makes the hacker obtain a copy of the cookies. Malware also digs into a system and retrieves the cookies for them.

Cookie logs the active sessions of a victim. This means that if a victim is signed in to a website, a cookie records that and lets them access the website continuously. Once a hacker has access to the cookie session, they are thus able to validate their own server or system and allow them to log in as well.

Using Wi-Fi for Hacking

While the Wi-Fi offers great accessibility for a user, it is also used as a

window to hack into the system or a network. There are a number of ways that this can work. The first one involves a hacker targeting one particular person. They look at their schedules and find a way where they use the internet a lot, for example, at a cafe.

They then set a fake WAP at the Cafe, naming it the same as the WAP that the victim typically uses. Once the victim is connected to their face Wi-Fi access point, they can read all the information that goes through it.

Another way is not to target a specific person but to target all the people in a specific area. So taking the Cafe example again, a hacker would simply set up a Wi-Fi access point in that location and let people connect to it. While people connect to it for free internet, the hacker can then read and access your data that you transmit through your Wi-Fi. This includes sensitive information, including passwords.

Trojan Horse

If you've read the Greek story of the infamous Trojan horse, you already know what you're dealing with here. The Trojans posed the horse as a gift and gave it to the Troy people. It was meant to show that the Trojans had given up, and Troy had won. In reality, though, the Trojan horse housed the army of Greece, which ambushed the city once the horse was pulled inside.

The virus functions in the same manner. It presents itself as software. Typically, Trojans tend to enter your system when you are trying to download software from unauthorized sites. The software isn't actual software and is just pretending to be one. Once you have installed the software, the Trojan can freely roam in your system.

The function of such a virus is to give control to the hacker by installing rootkits.

By installing a Trojan in the system of the victim, the hacker can obtain control of the system of the victim. This means that they can do anything in the system of the victim, including accessing all data and deleting or transferring any data that they wish.

This makes the Trojan horse virus a particularly dangerous one, and antiviruses usually have dedicated functionalities to ensure that no Trojan

goes undetected.

Keylogger

In a lot of ways, a keylogger replicates the behavior of a Trojan. It enters the system in a similar manner, but there are differences that we must take into account. The keylogger has a very simple function and is thus typically easier to develop than a Trojan horse. What that means is that most hackers can easily use it by getting someone to download fake software.

The function of a keylogger is to log keystrokes. What that means is that it records all the keystrokes that a person makes on their system. This allows the hacker to access both personal information that a person may have typed out, as well as any passwords that a person inputs in the system. That way, he can easily gain access to the different accounts of the victim and use them for whatever purposes he wishes. Keyloggers are now also able to carefully analyze the data that is input into them and dig out any passwords for the user, making it way easier to use them. They are thus a simple yet efficient way to hack into a system.

Drive-by Downloads

Drive-by downloads are websites that can force your browser to download a file when you visit them. These are powerful tools if someone is unable to convince other people to download malicious files on their own. Using this technique, the hacker can thus automatically get a file to download on a person's system and leave them vulnerable to Trojans and keyloggers, among other malware.

Social Engineering

One of the more modern and sophisticated techniques of hacking that is now commonly used requires very little technical knowledge. Social engineering is exactly what the name says. It's a method by which the hacker socially engineer their way into your system. They do this by manipulating people on a human level and attempting to get confidential information out of them. The information that can be received includes things such as parent's names, credit card and other document related information, and other things. Once a person is able to access all of this data, they can steal the identity of the

person to hack the system.

They do this by using the information that they have gained to prove that they are the owner of the user ID. Typically, forget password IDs can redirect you to security questions. These include Social Security Numbers and other such information. With the information that the hacker has now gained, they are easily able to cross the system and reset the password by assuming the identity of the person. This is a lengthy technique and often takes time since it would require the hacker to gain the trust of the victim.

Chapter 7: How to Protect yourself from Hacking

Now that you know the different types of hacking that exist, it would be important to know how to protect yourself from them. Without having adequate protection and without knowing how you can ensure that you remain safe from these kinds of attacks, you would let yourself be very vulnerable to hacking. Let us explore the different ways in which hacking can occur and look into how you can keep yourself safe in case of such an attack.

Physical Hacking

Physical hacking is perhaps the easiest to save yourself from. Physical hacking requires that a person has access to the travel channels of your data or the storage servers. To ensure that that doesn't happen, you can do a number of things. The first would be to ensure that all your data is encrypted, and password protected. Numerous services offer you the services to encrypt your data, which would render it useless even if someone was able to steal them. Newer hard drives are also encryptable and can have passwords put over them to ensure that no one manages to enter them. However, it is important to remember that while encryption offers some form of safety, passwords on laptops, computers, and hard drives are typically not as great security. This is because brute force can be used on such a system, and would allow for a person to quickly dig into the data. To secure a premise where your data is contained, a number of additional security measures can be taken. The first thing that you can do is to ensure that there are no vulnerabilities within the floor plan of the place where the confidential data is stored. While vents are necessary for heat sinks from the hard drives, they can be made much more secure by ensuring that better material is used for them. Alarms can also be placed inside the vents to alert the security in case someone tries to hack into them.

Similar measures can be taken in office spaces. A manager should ensure that different profiles are made for each of the employees so that the amount of data they are able to access is limited. This ensures that the lower-level employees that are not required to have access to any confidential data can be shut off from it. Managers should also ensure that the area where the

computers and hard drives are kept is secure. There should be no unauthorized person in the place where the servers are kept. This can be done via the installation of CCTV and hiring some security.

In case of the absolute requirement of confidentiality, newer hard drives are available where the data can automatically be deleted in case someone tries to force their way into them. This ensures that no one can access your data and would render any and all attacks useless unless the person making the attacks is already aware of what the passwords are. When it comes to hackers being able to tap into connecting lines between phones, things can get a lot tougher on the part of the company. It is usually not possible for a company to ensure that no one taps into the line.

However, it is important to remember that landlines offer much higher security than the VOIP options. It is thus important to avoid unsafe connections such as VOIP when you are communicating some confidential information to another person. Moreover, the company should ensure that no one knows when they are making any confidential calls. If the times are not known to the hacker, they would naturally find it much harder to find the information they seek. Eavesdropping on lines all day long is very impractical for a hacker.

If you must absolutely use VOIP or other such vulnerable methods, we recommend using a VPN or a virtual private network. While a number of VPNs now offer free limited services, the services for large amounts of usage would cause you some money. However, they can save you from a lot of hassles. A VPN has a number of features that include identity masking as well as encryption. We will discuss these later.

Brute Force

Brute forcing passwords is one of the easiest things that a hacker can do to get access to your information. While there are a number of measures that internet service providers, as well as websites, have done to ensure that no one can brute force their way through the passwords, there are some cautions that you must take as well.

The first and frankly, the most important cautionary measure that you can take is to make your password difficult to guess. Surprisingly, a large number

of people fail to do so. Millions of accounts are cracked into every year simply because the passwords that they choose are among a few of the following:

- Password
- Qwerty
- Their names
- Their pet's names
- Their crush's names
- Their favorite celebrities names
- Other similar easy to get information

This allows the hackers to quickly be able to get through the system and find your password.

Another thing that helps with brute force attacks is to have different passwords. To demonstrate how this helps, let us give you an example. Consider that a hacker manages to find an application that does not require Captcha clearing before entering passwords again. In that case, they would be able to hack through and get your password on that particular application. Most people keep similar passwords on most applications, and hackers are aware of that. If you do the same, the hacker would now also be able to access anything and everything from your internet banking application to your Facebook account.

If you keep different passwords for all applications, though, you would be able to keep your other accounts safe. There are many accounts that allow you to enter the password a limited number of times, and they would thus be safe.

If it is a hassle to remember a large number of difficult passwords, then you can find a number of applications to help you both find strong passwords as well as to keep them safe and allow you to log in on devices that are already approved automatically.

Phishing

Phishing is one of the easiest ways in which a hacker can access your information. It is luckily also the method that you can easily evade with a little common sense and help. To understand how to ensure that no information is given out using phishing, we have to look into each type of phishing individually.

The first type of phishing comes in the form of websites that are made to replicate other websites that you typically use. These are generally very easy to unmask. First off, you need to ensure that any website that you are linked to via an advertisement or a message is legitimate. For that, you can check the certificate of a site. If a website has an HTTP instead of HTTPS at the start, it is typically a high-risk site, and you should be careful navigating over it. You should ideally carefully read the link to ensure that you are on the right site. There are a number of characters that can easily be confused, including the O and 0, and l and I, so you should pay careful attention to those.

If you have any form of doubt about the validity of the site that is asking you for the password, there is a simple procedure.

You can enter an incorrect username and passwords. Phishing sites do not carry out any validation and simply store the information. Thus it would accept any and all username and passwords that you throw at it. That is an obvious indication of a site being made for Phishing.

The second kind of Phishing comes in the form of emails and SMS. These, although seemingly valid, are typically spoofed. Luckily, the procedure to find out if an email is spoofed is not too difficult. All you have to do is click reply to the email. When you are trying to reply to an email, it is going to be sent to the person that sent the email in the first place, and not to the spoofed address that it showed that the email came from. Thus by clicking a reply, you can easily find out who the email that you will be sending is going to and avoid giving out information if it's anyone not trustable. Another important thing to remember is never to share sensitive information over email. Companies would never ask you to email them your password or any other similar information, so be wary of such requests!

The third type of Phishing comes as the voiced variant. There are two ways in which that occurs. The first method uses call spoofing. It is important to

ensure that you never enter your information on calls that require you to enter them. Your banks would never ask for such information from you via call.

The second type is a scam where the caller pretends to hang up. It is best practice to put the phone down and ensure that the power line is cut before moving on to call the bank.

That ensures that you are safe, and the call was actually held up, and no one is on the other side, trying to listen to your confidential information.

With these simple practices, you can ensure that you are never a victim of phishing attacks and are able to keep yourself safe. These steps are very simple and can save you from a lot of hassles!

Cookie Theft

Your cookies are a piece of information that you should never let hacker access. If a hacker can access your cookies, they can be authorized to enter your accounts and do as they will. That is why you have to ensure that they are unable to get to your cookies in any way.

We previously discussed four different methods in which a hacker can access your cookies. We will now look at how you can save yourself from each one of them.

The first method is session fixation and needs you to click a link to open a session with a particular ID. To ensure that you are not led to a session created to steal your cookies, it is important that you do not use the links that are emailed to you from fishy-looking email addresses. Even legit email addresses should be confirmed by clicking a reply, as we have already mentioned before. These measures ensure that you are never at risk of having a fixated session again.

Sidejacking involves stealing cookies via the Wi-Fi network.

That means that both you and the hacker have to be on the same Wi-Fi network. It is thus recommended for you to keep your Wi-Fi password protected. In case of you being in a public place, you should ensure that any websites that you visit are encrypted. When the websites have SSL or TLS certifications, a hacker would be unable to access the information of the

session and fail to steal the cookies.

This can be done in a number of ways. Some of the ways include session fixation, sidejacking, malware, and cross-site scripting. Session fixation refers to when the hacker sets the session to an id that is known to him. He does that by sending a specific link. When the user uses that link to log into a session, the hacker is able to steal the cookies. Sidejacking involves stealing the cookies using the Wi-Fi connection. Many websites do not use SSL certificates on their site, and any data sent can thus be sniffed from the Wi-Fi connection. Cross-site scripting involves tricking the victim's computer into running a script that makes the hacker obtain a copy of the cookies. Malware also digs into a system and retrieves the cookies for them. You can use certain applications that force the encryption to ensure that you are always safe.

The third method that we discussed was the installation of malware. Malware cannot enter your system until and unless you allow for your system to download it. You should ensure that you do not download anything from fishy sites and that you have a good antivirus available on your system for your protection.

The last kind of cookie stealing is done by running scripts on the system of the victim.

The scripts again, need to be executed by the victim. Trickery is often used for that purpose. Many websites would offer you scripts that they would claim would activate your windows or give you an adobe license. All the scripts do is steal your cookies and send them to the hacker. To ensure this doesn't happen to you, don't download untrusted software and never run scripts that you find on unreliable sources on the internet. If a script sounds too good to be true, it most probably is!

Wi-Fi-Based Hacking

Wi-Fi-based hacking can be scary. However, it is also very easy to ensure that you don't become a victim in this case. The first type of Wi-Fi-based hacking is where a person is specifically targeted. For that to happen, the hacker would have to chase you physically. That means that you should know that someone has been keeping tabs on you if you're simply a little aware of

your surroundings. If you go to the same Cafe to use your internet at all times and find someone always chasing behind you, then you already know that there is something problematic. You can consider changing Cafes or not logging into an active session in their presence.

The second kind of hacking occurs when someone makes their own WAP and gives out free internet to steal data. A number of remedies are available to you in that case. The easiest one is to never log in to any places using the Wi-Fi that you can't trust, as well as to never send out any confidential information using it. However, that is not the best solution since you might need to connect to Wi-Fi for some reason.

Another better resource is available in the form of a VPN. A VPN, aka a virtual private network, ensures that there is another security layer between you and the hacker, and ensures that they are not able to tap into your data.

Trojan Horse

A Trojan horse is pretty similar to any other malware, and the best way to deal with them comes in the form of simple solutions that are applicable to all other such software.

The first way to deal with Trojans comes in the form of precautionary measures. These measures are made to ensure that the Trojan does not enter your system in the first place. These include measures such as ensuring that you never download software from places that you don't trust. Most third party downloading sites are infected with malware and should be avoided at all costs.

The second important thing to note is that you should not download attachments from emails that you do not trust. If someone has emailed you a random file with no explanation, you should avoid it by all means and not download it, for it may contain malware.

If malware has already entered your system, you need to ensure that it is unable to act and take control. It is especially dangerous once it has accessed your BIOS using a rootkit, so immediate action must be taken. For that, you should, first of all, ensure that you are already using a reputable antivirus. Antiviruses typically already have lists of known Trojans and can easily

locate as well as quarantine them.

Another important tool is a sandbox. In case of having to download something that you do not trust, it is recommended that you do so in the sandbox. A sandbox is a virtual container within your system in which you can download any files and test them. If they display abnormal behavior, only the sandbox would be infected and can be deleted. If they are safe, you can then download them directly to your system.

If you are already infected by a Trojan, you should download an anti-malware software such as Malwarebytes to ensure that you can remove it from your system. If the malware has dug into your BIOS as well, you should always perform a BIOS reset to ensure that any rootkits are removed from your system completely.

Keylogger

A keylogger can be a lethal hacking tool. Fortunately, it has its own downs. Any keylogger must be able to transmit data back to the hacker for it to work. That means that there is a strong chance that any firewall would detect the keylogger at work and alert you about it. It is thus important to ensure that your firewall is on at all times to prevent keylogger-based attacks on your system.

A keylogger typically behaves like malware, so the same precautions as a trojan horse apply. You shouldn't download malicious and unknown content, and you should keep your system protected by using a strong antivirus.

For both keyloggers and trojans, one thing that would help is to update your system constantly.

With system updates, you will usually find many of the older exploits are closed down, and the defense is much better, ensuring that your system is kept safe from prying eyes of the hacker.

Drive-by Downloads

Drive-by downloads enable the downloading of malware into your system. There are a number of ways in which you can prevent that from happening.

The first thing that you should consider is to disable auto-downloading options. Most browsers are equipped with an optional turning off of automatic downloads. Any download requests thrown by the website would thus not be processed, and anything that you don't download won't be downloaded.

It is also important to not click links you are sent by a third party or ad sources unless you trust them. The links might often redirect to sites that are made to force downloads of malware.

Having a decent defense system, including a strong antivirus as well as a firewall, can also help. These would immediately quarantine any threat even if it were to be downloaded. It is also helpful to remember that most applications require permission before running. If anything is downloaded automatically, delete it instead of running the script!

Social Engineering

Social engineering is a tricky hacking technique. It is consequently also tricky to avoid. Luckily, with a few precautions, you can generally ensure that you are not a victim of social engineering.

The first thing that you have to remember when it comes to social engineering is to keep your security question and answer safe. Make the answer unique and never tell anyone what it is. Questions like "What's your mother's name" are too easily guessable and should either be avoided completely or should have unique and untrue answers so that no one can engineer their way into your system.

If a sketchy person seems to be taking an unusual interest in particular information that can directly be connected to your bank account, it is usually presumable that the person has malicious intent. You should ensure that no answers are given to such a person, and your personal information remains safe and personal.

Many websites now allow you to set up two-factor authentication and other methods of accessing your information to ensure that you are not made a target of social engineering. You should always enable any such options for added security. It is also usually helpful to have a valid phone number or email address where password reset links can be set. If those exist, most sites will not rely on having to second-guess your identity, potentially letting someone steal it along the way.

You should also ensure that you are aware of the security protocols of your bank. You can normally set limits, transactions above, which would be confirmed from you via the number you provided to the bank. This ensures that no one can pretend to be you and rob you of your money.

Chapter 8: Cybersecurity and How it Saves you from being Hacked

Now that we have gone over some of the basics of hacking and how to prevent yourself from them, it is important to see how systems, as well as websites and applications, aim to save both you and them from hacking. To dig deeper into that information, we have to look at the different ways in which cybersecurity works.

What is Cybersecurity?

Cybersecurity refers to the practice that is used to protect programs, networks, and users from digital or cyber-attacks. They can operate in a number of ways, but the end goal remains the same: to keep the data safe and secure and to ensure that no one can malign, steal, or otherwise destroy it.

We will now explore the different countermeasures that are available against a cyber-attack and how they can help a network evade any form of attack.

Countermeasures by Design

Countermeasures are designed to mean that a system is made to ensure that the maximum amount of security is available to the network and its users. The design elements that can help with security offer a number of features.

One of the main design features to ensure high security is the principle of least privilege. This is a very simple mechanism and offers any user only the minimum authority within a system, as is needed by them. This ensures that even if a hacker assumes the role of anyone within a network, they would be unable to do much and would not be able to access data, alter it, or delete it.

Other design mechanisms that can enhance security include defense in depth. This refers to systems where you need to breach more than one aspect of the system to be able to penetrate it. So, for example, a system might need authorization from both user 1 and user 2 to allow access to anyone to the sensitive data. This design means that the hacker has to hack through multiple security systems, and makes it much harder for a hacker to be able to gain control.

Another important design measure comes in the form of audit trails, which ensures that if any vulnerability is detected, either through black hat hacking attacks or otherwise, it is promptly dealt with, and the system is not left vulnerable in the end. This keeps the system safe from further attacks down the line.

Security Architecture

This form of countermeasure aims to design the system in a way that makes hacking difficult.

This is mostly a design-based system but rather deals with how various entities within a network interact with each other. By limiting the dependence of a system on other systems, we can ensure that the hacker does not gain control of the whole system even if he enters a part of it. Another important role of security architecture is to ensure that any and all entries and vulnerabilities are covered by the security systems in place. It thus dictates where the security measures are placed to ensure that no vulnerabilities remain.

System Penetration Testing

An important way of checking the security design, architecture, and strength of your system is to do a system penetration testing. A penetration testing, which is also popularly known as pen testing is a way of accessing the vulnerabilities within a system to ensure that all of them are closed. There are a number of ways in which a system can conduct penetration testing to ensure that there are no vulnerabilities within the system

If a company, system, or network does not have the resources to hire themselves a hacker, they can usually use software that is already available to conduct such tests. These include tools like MetaSpoilt that we have already discussed. These tools have distinct functions. When they are made to test a network, they will find out all the vulnerabilities. Many tools also exist that can both exploit these vulnerabilities to assess the damage to the system that would occur in case of an attack, as well as provide solutions to these damages to ensure that they do not occur in the future.

For the best service, though, the automated tools are left far behind. The

persons that one should refer to for the best analysis are the white hat hackers.

White hat hackers are usually skilled in penetrating a system and can quickly point out flaws that would not have been pointed out by any other traditional script. However, they can be expensive to hire.

When a white hat hacker is hired, you can make them test the system in two ways. The first type is black box hacking and would mean that the hacker is not told any facts about the network. This is a useful method if you also wish to see how easy it is to find information about your network. Having accessed the information, the hacker would then write scripts to find vulnerabilities, in a much more efficient way than the pre-written scripts could have done. These vulnerabilities are then reported back to the person that hired the white hat hacker to ensure that they can be closed. Black box testing is a good way to see the practical security level that your system holds.

Another type of white hat hacking can be employed. This type revolves around a white box methodology. What that means is that the hacker is already told all the information that he needs to know about the network, and the code and system are transparent before him. This is especially beneficial when you want an in-depth penetration testing since, with the information that the hacker would already have, he can dig much deeper into the vulnerabilities. Most companies would use a mixture of white box and black box based white hat hacking to ensure that the hacker is both able to find the issues in-depth and able to show a realistic strength value of the system.

Elite groups of hackers are usually able to write scripts that other hackers cannot hope to reproduce. They are thus typically paid way higher, and companies tend to employee them to find any flaws where the business model requires high network security. Such models can include government websites as well as websites of cloud storage apps.

If cost is an issue to you and you feel that your system is already impenetrable, a new form of white hat hacking contracts is now becoming popular. These are performance-based. A forum of hackers is offered a bounty for being able to hack into the system of the company. If no hacker succeeds, you already know your system is strong enough. If someone does,

you need to spend the money only in that case. This helps save money if you're already sure of your security.

Two-Factor Authentication

One of the prime ways to ensure that no one is easily able to access the system is to use two-factor authentication. That means that you need two distinct pieces of information to access the data. These typically include one pin or password, and one hardware or biological trait. Typical systems can combine pins with thumbprint scans. Having a system that is based on these traits can mean that it is much more secure. Fingerprints are hard to replicate and typically require access to the victim. Similarly, for any cards using NFC to allow access to a system, the victim's card must be stolen. In such a case, a victim can usually quickly inform the system about a stolen card and thus ensure that the system is not compromised using it.

Having a two-factor authentication can thus make your system much more secure, and make penetration much more difficult.

Data Security: Encryption, Protocols, Packets, and Transmission

Data has to travel around a lot. That leaves data very vulnerable. If data is sent through a simple wire with nothing to hide it, hackers can simply dig right into it. That is where the data security features come in. There are a number of ways in which data is protected. We will explore each of them in detail to see how your data manages to travel from one place to another safely.

If your data is traveling in the form of simple radio waves or through copper lines, anyone with the right software can read the data and thus be able to get your information. Radio waves are particularly unsafe since you don't need to find a line to tap into physically, and the risk of being caught is thus mitigated. Newer technologies now rely on fiber optics, which is more secure. This is because when someone tries to tap into a fiber optic, it will break the glass and possibly trigger alarms. This means that data that travels through that channel is typically much more secure.

However, most of the systems, especially WAN and remote areas, continue

to use radio waves or copper-based wires. Fiber optic is also not impenetrable in any way. That is why data transmission is risky.

To help mitigate the risk, a number of ways are used. These include encryption and packet-based transmissions. We will look into both of them in detail.

Encryption means that data is distorted so that it is not readable. The earliest forms of encryption include the easier Caesar cipher, more popularly known as the shift cipher. The protocol is basically simple. The data in such a cipher is simply encrypted by using a shift. Thus a is transmitted as b, b is transmitted as c, and so on so forth. Of course, given the computing power of today's world, such algorithms are obsolete. We now have more reliable solutions to ensure our security.

There are two basic types of encryption. The first type is symmetric encryption. Data that has been encrypted using this form of encryption can only be decrypted by using the same key that encrypted it. This means that the key has to be transferred as well. This can lead to a security risk since hackers can ultimately intercept the key as well. However, if the key is transmitted in a secure manner, no significant concerns should arise.

Data is typically transmitted as packets. This means that not all the data is channeled at once, but data is rather divided up into small packets that are sent simultaneously or one by one, through a network. All of these packets then combine at the end receiver to make a complete data set, which can be decrypted using the key that was transported by any means. If you are using symmetric encryption, it is thus very important to ensure that the key is kept safe.

In using asymmetric encryption, the data is encrypted using a public key and decrypted using an individual private key. Since the key does not have to be shared, it is typically thought to be more secure. Nonetheless, you have to keep the key safe and ensure that no one can assume your identity or steal your cookies for that to work. Websites that use asymmetric methods of data encryptions typically have the HTTPS tag and have SSL technology, which ensures that any information that you enter on those sites is kept secure. You shouldn't enter sensitive information on websites that lack those protocols.

There are various algorithms that are used within encryptions that are important to know if you wish to keep your data safe. The earlier algorithms include the DES and triple-DES that were found to eventually become vulnerable as technology caught on and was able to brute force the algorithms. Newer algorithms such as the AES can make your data virtually unhackable. Using the current technology, AES 256 encryption would take billions of years to crack!

Protocols define how the algorithms should be utilized, and allow for secure key exchange among other functions. The protocols are thus an important part of the process since they allow for the data to be readable to the intended user.

So to summarize, while data can still be tapped into, it is often kept in the form of cryptic algorithms that can only be decrypted using keys that are transferred using secure protocols such as SSL.

As long as a system sticks to the protocols, encrypts the data, and uses protocols like SSL to keep their data secure, the system itself cannot be brute-forced, and that means that vulnerabilities in terms of data in transit are minimalized.

Encryption is also possible on data at a resting stage, and with those, you can similarly ensure that the data cannot be accessed. However, the decryption keys are usually kept behind simple passwords that can be brute-forced. To ensure that brute force attacks do not succeed, a network can use a number of smart security designing that we will next explore.

Preventing Brute Force Attacks

Brute force attacks can both be used to hack into a system that stores data at rest, as well as to access the data of the users of the network by finding their passwords. With there being a very limited selection of passwords that people typically opt for, and computing being super powerful now, it is hard to prevent an attack that focuses on brute force. Luckily, with a few changes in design, this can be rendered useless.

The important thing to remember here is that a brute force attack aims to mainly get access to the data by having an infinite number of tries at guessing

the password. Systems have now been designed to restrict the number of tries that a person can have at guessing. Most online log-in applications would either send ReCaptcha tests that the brute force algorithm would be unable to solve and thus stop it, or would simply lock the account down after a number of tries. Doing this, they are able to ensure that the brute force mechanism cannot continue to dig into the system.

This is an important aspect of security since by adding this feature, you can ensure that even if someone manages to reach your key and has to guess a password simply, they would be unable to do so quickly .

To make things even more secure, logs can be taken of attempts at logging in to a system. The logs will show when someone tried to log into a system and the passwords that were input. This means that the admin can be altered about a brute force attack, and measures can be taken to ensure that care is taken of it.

By using this measure, a network can ensure that no one is able to gain unauthorized access by way of guessing passwords. It is also a helpful feature to include in case of a website having web portals, and many applications such as Gmail and Dell do make extensive use of it to ensure that the accounts of individuals registered for their services are not compromised.

Firewall

A firewall is an underrated aspect of network security. In simple terms, a firewall monitors all incoming as well as outgoing traffic and governs it via specified rules. This is very important since otherwise, people can simply send packets of malware to your system.

A firewall is a wall between your network and the outside world. It inspects all incoming data packets and reads the protocols, including FTP, HTTP, and DNS.

It then uses the rules that are predetermined to decide if a packet should be allowed into the system or not. Similarly, it also inspects packets going out of the system. This is useful since, in case of a hacker using scripts to gain data or logging your keystrokes and trying to access them, the firewall can inspect the packets and identify there being something wrong, and thus block transmissions.

Given the function of a firewall, it is absolutely crucial for a system to have a good firewall. Firewalls now come with greater functionalities, and features and customizations are pretty simple and straightforward. Companies that require the ultimate security solutions could customize a firewall to their needs and ensure that no data is transmitted in or out of the system without proper authorization.

A firewall is the first defense that a system has, and can be useful by preventing malware from getting into the system and preventing system data from flowing out. A strong firewall is thus protection we should all consider getting.

Anti-Virus and Anti-Malware Software

Anti-virus and anti-malware software typically form a secondary line of defense. The function is simple. They scan the system and try and find any programs that are known for being problematic. Most anti-viruses have huge repositories of codes and can easily identify viruses and trojans by matching them with the codes of other known threats.

The threats are then alerted to the system admin, who can then decide to quarantine or remove the threat. Such removal ensures that the threat is no longer active. They can identify a number of threats, including trojans, rootkits, spam and scams, phishing, and DDoS.

Given the versatility of this software, it is absolutely essential for anyone that works with sensitive data to get one of the premium plans of a reputable anti-virus. By doing so, not only would they be able to identify any trojans or other malware, but they would also be kept safe from phishing attacks and other such hacking methods.

Internet of Things: The Unpatched Cybersecurity Threat

One of the prime things that can cause a system threat is via the internet of things. The internet of things refers to any and all things that require internet or network access to work. These include a number of things, including modern homes, garage doors, cars, printers, refrigerators, and whatnot. With a large number of manufacturers now letting you control your electronics using your cell phone, the internet of things becomes all the more important.

Sadly, we see that vendors typically give very little emphasis to this industry in terms of cybersecurity, and simple patches that would seal vulnerabilities are never made. This means that everything from your printer to your garage door is hackable and can be hacked with ease.

Given the digital world that we are moving to, it is high time for manufacturers to place more emphasis on the security of such systems and to ensure that they are made as secure as the other systems are. Ultimately, all networks should be made secure so that hackers can no longer be a menace, and without a focus on the internet of things, we can expect to continue to see successful cyber-attacks.

How are Cybersecurity Tools used to Secure your Wi-Fi Connection?

Wi-Fi connections are typically very vulnerable. If you are going to be using a public router, in a hacker's eye, you're an easy target. Luckily, with newer system updates, many operating systems include inbuilt protection against such attacks.

Taking the example of windows, when you connect to a network, a system will typically ask you if the network is public or private. In case of you being on a public network, windows will automatically hide your device and make it undiscoverable. This ensures that you are kept safe, even where you are connected to a public network. It is important to be still wary, though, since the network admin can still see the information that you transmit, and WAP attacks are commonplace. That is where a VPN comes in. Let's explore a VPN in more detail.

Virtual Private Network

A VPN or a virtual private network allows you to transmit data using a public network as if you were transmitting it using a private network. There are a number of ways that this is achieved, and we will look into some of them.

To start with, a VPN masks your identity and shows a different IP address. This is typically done by masking your IP with another IP that belongs to the VPN server. This means that anyone that intercepts the data will not know where the data came from, and would rather believe it to have come from a virtual IP address that the VPN sets for you.

Secondly, VPN tunnels past the public servers. This means that the data that is sent through the VPN is delivered through packets. Each of those packets has a protocol and is duly encrypted. The encrypted packets cannot be decrypted in the public network, and would only allow the intended recipient with the key to decrypt them. Thus even where the WAP hacker manages to get a hold of the packages, he would be able to find little use for them. VPN thus ensures that you get the maximum security, and your data is kept safe. It is very important for a firm to use VPNs if they are transmitting highly sensitive data over Wi-Fi since Wi-Fi can be hacked into and allow for someone to sniff the data. That is why VPN, along with other cybersecurity methodologies, becomes an important part of any defense arsenal against cyber-attacks.

Conclusion

Now that we have analyzed all the ways that hacking occurs and the tools that hackers may use to enter a system, as well as mentioned the ways in which you can ensure that you remain safe from hacking, we hope that you have found the answers to all the burning questions that you had.

A few short takeaways include that your data is very vulnerable, and everyone, even newbies, can crack into it unless you take steps to secure it. This is especially true for networks since more people tend to want to break into networks. That is why one should always ensure that their network is able to past the penetration tests and has an adequate design. This would ensure that the data of both the network and the customers are kept safe.

There are a large number of ways in which you can protect yourself, and the best combination for you depends on your individual needs. A website that uses no Wi-Fi-based communication networks in public, for example, would not need a VPN tunneling based security system. Similarly, any network that blocks out all external data and takes no inputs would have no use of a firewall.

Whatever your business model or product may be, irrespective of if you're a simple social media user, a small company, or a large network, hacking is a menace to anyone connected to the internet. That doesn't mean that it has to be that way.

Hacking can be fought against in simple steps that we highlighted in this book, and by using them, you can ensure that your data and identity are never compromised again.
Remember, prevention is better than cure. Once a hacker has your data, it is really hard to be able to retrieve it and get it deleted from the internet. You should thus ensure that you follow the guidelines that are made to ensure that your accounts do not get hacked, and you should ensure that your network has a proper cybersecurity plan that it uses to protect itself against any such attacks. We hope that this book was helpful to you in achieving your ultimate hacker-free dreams and that you will now surf the internet safer (or make a career out of ethical hacking!)

Artificial Intelligence And Life

A Complete Guide to the Basic Concepts in AI, Neural Networks, Machine Learning and Data Science.

Hans Weber

Introduction

Intelligence is a quality that enables an entity to perform certain functions according to the external environment. Artificial Intelligence (AI) is the pursuit of making machines intelligent. If you can relate to this definition, then several things such as animals, humans and certain machines can be considered as intelligent beings.

On one end of this spectrum are machines and animals with their varying degrees of intelligence and on the other end of this continuum are human beings. They can perform numerous tasks, add reason and logic to them, generate and understand languages and perceive as well as react to sensory stimuli. Additionally, they can also play games, prove theorems, create music and art, write history and synthesize large volumes of information. To do this, humans require a wide array of abilities with no discontinuities, and they are heavily dependent on their external environment. It is for these reasons that AI should also be looked into from a larger perspective. This is applicable to all facets of life such as statistics, linguistics, logic, electrical engineering and computer science.

This book has been structured keeping three kinds of readers in mind. The first type of reader it will appeal to is the one who is interested in scientific topics and is curious about AI. The second type of reader shares some similarities

with the first one, but is involved in the technical or professional field and needs to understand the benefits of AI and get a bigger picture of what it entails. The third type of reader essentially consists of students, teachers and researches who need to know about the methods that proved to be effective and the ones that didn’t. Those reading the book must have a general idea about the history of Artificial Intelligence.

For most of us, a machine equates to something that is mechanical and solid. It reminds us of steel parts and gears. The computer has, however, changed our understanding of what a machine is. An operating computer includes both hardware and software components. The software by itself can be considered as a machine. The line that separates software and hardware has become

blurred. Every modern computer has certain programs built into the hardware circuits. These programs make it easier for the system to perform a variety of functions. This book focuses on the different ideas and realizations that have helped reach this point of advancement. You will gather information about what artificial intelligence is and what led to the building of computers with artificial intelligence. You will also learn about the basic concepts of artificial intelligence. I hope the knowledge and abilities gathered in this book provide a wider understanding of the field.

Chapter One: Foundations of AI

The long pursuit of artificial intelligence began with dreams. Numerous people have wanted to develop a machine that has the same abilities as a human being. The idea is to develop a machine that can reason and think like human beings. Such machines have been mentioned throughout history in paintings, scriptures and even literature. An automated chair called a tripod and a golden attendant were mentioned in the Iliad of Homer. The ancient philosopher Aristotle spoke about automation but considered it a fantasy. In the year 1945, Leonardo Da Vinci, whose inventions were always ahead of his time, also designed humanoid robots that could serve as medieval knights. The physical evidence of his invention has not yet been discovered, but it is said that this knight could sit, move its head, move its arms and open its jaw. Isaac Asimov, the famous science fiction writer, wrote stories about robots that were powerful and mighty. His first collection called *I, Robot* had nine stories about positronic robots. He was not into the idea of robots that only caused destruction, like Frankenstein. These positronic robots had the Three Laws of Robotics hardwired into them. These three laws are as follows:

First Law: A robot will not cause any harm or injure a human being.

Second Law: It is essential that a robot obey every command or order given by a human being. It should, however, never act upon a situation that will clash with the First Law of Robotics.

Third Law: A robot is required to protect itself as long as its actions do not adhere to the first two laws of robotics.

Asimov later wrote a fourth law called the Zeroth Law to protect the interest of humans.

Zeroth Law: A robot may not threaten humanity or cause any harm to humanity.

The pursuit for artificial intelligence, unrealistic or not, stems from dreams like these ones. To turn these musings into reality, several clues that tell us how to proceed are required. These clues that are essential to shape this reality are scattered across different subjects like logic, statistics, philosophy, biology, psychology and engineering.

Philosophy and Logic

Throughout the millennia, people reasoned using logic, but it was Aristotle, the Greek Philosopher, who analyzed and coded the process. He recognized a system of reasoning and called it Syllogism. Here is an example of what Syllogism is about:

1. Human beings are all mortal. (Stated)
2. Every Greek is a human being. (Stated)
3. Every Greek is mortal. (Result)

The genius behind Aristotle's theory is in the form of the Syllogism. We do not have to restrict ourselves from applying this to merely humans, mortality or Greeks.

We can use arbitrary symbols instead of humans, mortals and Greeks to rewrite this equation. Rewriting in such a manner will give us:

1. Every Olympic runner is fast. (Stated)
2. Usain Bolt is an Olympic runner. (Stated)
3. Usain Bolt is fast. (Result)

Aristotle's logic bought forward two ideas that would help automate reasoning. The first idea was that the patterns of reasoning, such as syllogisms, could be economically represented as either forms or templates. These generic symbols can stand for anything, and the symbols themselves are irrelevant. Second, after their values substitute the symbols, one can calculate the result. Modern AI reasoning programs utilize the same principle of general symbols and problem solving.

In more modern times, Leibniz tried to mechanize reasoning by designing a language, which can formulate all human knowledge - even metaphysical and philosophical knowledge. Leibniz termed this language as the universal language or *lingua characteristica*. In his system of automatic reasoning, the alphabet played a rather vital role. He theorized that if numbers represented the items in the alphabet, then a complex proposition could be achieved from its primitive constituents by multiplication of its corresponding numbers together. Further mathematical operations could determine if this complex proposition was true or false.

The issue with applying this idea was that it was difficult to break the primitive alphabets into components. Nevertheless, Leibniz's theory provided additional insights into how reasoning can be mechanized. The theory talks about how an alphabet can be created using simple symbols and also talks about how these symbols can be combined to develop complex expressions.

Towards the end of the 18th century, a British politician and scientist, Charles Stanhope, created a device that can solve simple problems in logic and probability. This device can be looked at as the earliest form of an analog computer. This box consisted of slots on its sides used to push in colored slides.

A window on the top could help view the slides that were placed to represent a particular problem. For instance, let us try solving a question in numerical syllogism.

9 of 10 A's are B's; of 10 A's are C's

Therefore, at least 2 B's are C's.

There is a window that represents A, and Stanhope would move a slide with eight units that was painted in red and also a slide with four units that was painted in grey from opposite directions. The slides will overlap in two units, and this represents the minimum number of B's that are also C's.

In 1854, George Boole published a book that looked into the nature and constitution of the human mind. He considered several logical principles of human reasoning and transformed them into a mathematical equation.

For instance, the principle of contradiction states that it is impossible for any being to possess a quality and not possess it at the same time. He then made this principle into an algebraic equation,

$x(1-x) = 0,$

Where x symbolizes an object, (1-x) symbolizes the contrary class of the object, while 0 represents a class that does not exist.

Through Boolean algebra, we know that 0 denotes falsehood and 1 denotes truth. Two of the most important operations in logic, ‘OR’ and ‘AND,’ are symbolized in Boolean algebra by the operations + and X, respectively. For instance, the statement "*either p or q or both*" can be represented by p +

q. To represent the statement *p and q*, we could write *p x q*. Each p and q could either be true or false; therefore, we can calculate the value of *p + q* and *p x q* by how + and x are used and the values are always calculated in terms of whether the value is true or false, that is,

1 + 0 = 1

1 x 0 = 0

1 + 1 = 1

1 x 1 = 1

0 + 0 = 0

0 x 0 = 0

Boolean algebra played a vital role in the design of telephone switching circuits and computers. Although Boole never envisioned computers, his work led us to an important clue about the mechanization of reasoning. A similar approach, but not algebraic, for calculating and manipulating propositions known as propositional calculus, plays an integral part in artificial intelligence.

The main drawback of Boole's logical system was that it does not help you understand anything internal to the propositions. For instance, if we proposed, "Jack is human" by p and "Jack is mortal" by q, there is nothing internally present in p or q to prove that Jack who is human is also the same Jack who is mortal. For this purpose, we need to understand "molecular expressions" which also have internal expressions.

Towards the end of 19th century, philosopher Friedrich Ludwig Ferge invented a system where propositions, along with internal expressions, can be written in a graphical form. Ferge system was the frontrunner for what we now refer to as the "predicate calculus," another vital system in artificial intelligence. It also indicates another graphical form that is used in present-day artificial intelligence: semantic networks. Due to Ferge's work, sentences that included information to be reasoned could be represented in unambiguous and symbolic forms.

From Life Itself

Different aspects of what life have provided insights into what intelligence means.

Since it is the human brain that processes vast amounts of information into action, it is to be expected that works of neurophysiologists and neuroanatomists, those who study the brain, can provide a better understanding of what intelligent behavior is. An understanding of human evolution, which brought forward intelligent life, can also provide hints to help explore the field of artificial intelligence.

The human brain has about 10 billion cells called neuro. They are found in different forms, but their parts are quite similar. The central part is the cell body or soma, incoming fibers are called dendrites and the outgoing fibers are axons. The axon will have a projection called terminal buttons that come close to the dendrites of other neurons. This gap between the terminal button of one neuron and dendrite of another is called a synapse. When a pulse reaches this synapse, as a result of electrochemical actions, it may excite or inhibit the electrochemical activity of the other neuron across the synapse. What the second neuron does will depend on the type of pulses that arrive at the synapses and the efficiency of those synapses in transmitting electrochemical reactions. In the year 1943, a paper by neurophysiologist Warren McCulloch and logician Walter Pitts claimed that the neuron was a logic unit. It proposed a simple demonstration of neurons and showed how networks of these models could perform any computational operations.

A neuropsychologist, Donald O Hebb, also emphasized how neurons in our brains perceived the basic units of thought. He theorized that,

when an axon cell A is close enough to excite B and constantly fires it up, some metabolic or growth process takes place in one or both cells in a way that A's efficiency, as one of the cells firing B, increases. This was later known as Hebb rule and was observed in living animals. Hebb also theorized that a group of neurons that happens to fire together formed a cell assembly. Hebb believed that thinking was the sequential activation of groups of cell assemblies in the brain.

Psychology and Cognitive Science

B.F. Skinner and several psychologists attempted to make psychology more scientific and independent of subjective introspection. They started this by only concentrating on what could be measured objectively, such as, particular behavior as a reaction to a specific stimulus. The behaviorists believed that psychology should be about behavioral science and not of the mind. They did not consider the idea of identifying internal states of mind, such as belief systems, desires, intentions and goals.

Skinner’s theories gave the idea of reinforcing stimulus, which rewards recent behavior and tends to make it happen again under similar circumstances (in the future).

Reinforcement learning is a very popular strategy among AI researchers, even though it does not depend on internal states. Skinner wrote that, for each stimulus, the animal makes a note of the opponent's next move and the next time the stimulus recurs, the animal duplicates the opponent's moves that followed the same stimulus previously.

More repetition of the opponent's move leads to the animal model being conditioned to that particular move. Skinner suggested that verbal behavior in humans could also be explained through his theory.

Noam Chomsky, a linguist, argued against Skinner’s ideas about language. Chomsky raised the question as to how a person can produce an infinite variety of previously unheard and unspoken sentences with complex structures just through experience. These fundamental factors that Skinner omits are, according to Chomsky, linguistic abilities that are inborn and not learned. According to Chomsky, human beings are created to handle data or can formulate the hypothesis of unknown character and complexity. He claimed that humans, by birth, have a universal grammar or a mechanism, which develops our ability to learn and use language.

Psychologist George A Miller focused on internal mental processes and its limitations. He claimed through his experiments that the immediate memory capacity of humans was approximately seven chunks of information. It didn't matter what the chunks represented, and we can only hold seven of them in our immediate memory. This paper was presented at a Symposium on

Information Theory at MIT. Famous AI researchers such as Allen Newell and Herbert Simon gave a paper on a computer program that could solve propositional logic. This Symposium brought together scientist with overlapping interest and is thought to have influenced the founding of Cognitive Science, a discipline devoted to the understanding of the human mind.

Cognitive science tried to give a clearer explanation for internal mental processes through ideas like memory, strategies, goals and task queues. Cognitive science and artificial intelligence have always overlapped since its initial phase. These two fields have helped each other identify new concepts by understanding how the human mind works.

Statistics and Probability

Uncertainty is the common factor in any reasoning and decision-making processes, and dealing with it is an important part when it comes to automating intelligence in machines. The theory of statistics and probability stems from the attempt to quantify the laws of chance and uncertainty. The most important results for artificial intelligence is the Bayes rule that can be described through an example. The most integral application of Bayes rule is in single detection.

Let us assume that a radio receiver is tuned in to a station which broadcasts one of two tones, A or B tone after midnight. We do not know beforehand as to which tone is going to be broadcasted, but we can find out their probabilities. Can we find out by listening to the signal coming to the receiver? Since the station is far away and random noises obscure the tone, listening in cannot solve the matter, but depending on the nature of the obscure noise we can calculate the probability that the tone that night would be A or B. Let's represent the signal with y and the actual tone with x.

The possibility of x = A, provided the evidence for it contained in the incoming signal y, can be formulated as p(x = A| y). It is read as, the probability that x is A, given that the signal is y. In case it’s B, then the equation will be

P(x = B | y).

Decision rule would favor tone A if P(x =A | y) is larger than P(x = B | y) or decide in favor of tone B. These probabilities are not readily calculable; therefore Bayes' rule can be used. It lets you calculate probabilities with respect to other probabilities that can be easily guessed or achievable. Bayes' rule is as follow:

P(x | y) = P(y |x) P(x) / P(y)

Our rule can be reformulated as

P(y | x = A)P(x = A)/ P(y) > P(y | x = B) P(x = B)/ P(y),

If you're in favor of tone A

Otherwise, choose tone B. Since P(y) occurs in both formulas, it does not

affect as to which one is larger.

We assume that we know the probabilities p(x = A) and p (x = B), so P(x | y) is what is left to calculate for x = A and x = B. This can be calculated depending on how the signal y is represented on the statistics of the interfering noise. For computational purposes, time- varying voltage is represented by uniformly spaced points such as,

y(t 1), y(t 2)….y(t I)….y(t N)

When noise alters, the probability of the sequence of them can be calculated using these statistics. These uses of Bayes' rule have been applied in artificial intelligence to mechanize the idea of both visual images and speech sound.

From Computer Engineering

For artificial intelligence to become a reality, we require intelligence and an artifact. The computer has been chosen as the artifact. The modern digital computer was developed independently in three countries at war with each other, i.e., World War II. Alan Turing and his team developed the first operable computer for breaking the German coded messages. The same group developed Colossus, which was a powerful machine based on vacuum tubes. Konrad Zuse invented the first operational programmable computer in 1941. He also developed the floating point numbers and the first programming language. The most influential forerunner for modern computers was the ABC, which was developed by John Atanasoff and Clifford Berry.

From that point, each generation of hardware was better and faster in terms of speed and capacity. There was also a significant decrease in price. Performance soared every 18 months until the year 2005 when the manufacturers increased the number of CPU cores due to power dissipation rather than increasing clock speed. According to current research, in the future, the increase in power can lead to a convergence with the properties of our brain.

Artificial Intelligence is also dependent on the software development side of computer science that includes the programming languages, operating systems and tools for modern programs. Research in AI has brought forward many concepts that have helped computer science advance. Some of these concepts include personal computers with a mouse, the linked list data type, interactive interpreters, time-sharing, automatic storage management and key ideas behind functional, symbolic, declarative and objective programming.

Chapter Two: What is AI?

A deeper understanding of what Artificial Intelligence means can be attained through eight definitions of AI laid out along two dimensions. One dimension is concerned with reasoning and thought processes, while the other is based on behavior. In the figure shown below, we see the eight definitions of AI. The definitions on the left side measure success with respect to human performance. The ones on the right measure rationality based on ideal performance. A system can be considered rational if it does the right thing, considering what it knows.

Thinking Humanly	**Thinking Rationally**
"The exciting new effort to make computers think . . . *machines with minds*, in the full and literal sense." (Haugeland, 1985)	"The study of mental faculties through the use of computational models." (Charniak and McDermott, 1985)
"[The automation of] activities that we associate with human thinking, activities such as decision-making, problem solving, learning . . ." (Bellman, 1978)	"The study of the computations that make it possible to perceive, reason, and act." (Winston, 1992)
Acting Humanly	**Acting Rationally**
"The art of creating machines that perform functions that require intelligence when performed by people." (Kurzweil, 1990)	"Computational Intelligence is the study of the design of intelligent agents." (Poole *et al.*, 1998)
"The study of how to make computers do things at which, at the moment, people are better." (Rich and Knight, 1991)	"AI . . . is concerned with intelligent behavior in artifacts." (Nilsson, 1998)

All four approaches have been followed through in the past, and each approach was used by different people using different methods.

The human centered approach was based on observations and hypotheses about human behavior; therefore, it was more like empirical sciences. The rationalist approach included a mix of engineering and mathematics. The different groups have converged and disparaged each other. We shall look into these approaches in detail.

The Turing Test Approach: Acting Humanly

Alan Turing proposed the Turing test in 1950 that helped to obtain the operational definition of artificial intelligence. A computer would pass the Turing test if it could not identify whether a response to a written question asked by a human being came from a human or another machine. The focus should be on how the computer was programmed to pass the test. The groundwork gives us plenty of insights into how to mechanize intelligence. This computer should have capabilities such as:

- Communicating in English through natural language processing
- Store what it hears or knows through knowledge representation
- Using the stored information to draw conclusions or predictions using automated reasoning
- To detect and predict some new patterns through machine learning

The Turing test intentionally disregarded direct interaction between the computer and the human interrogator, as physical simulation is not required for intelligence. In the total Turing test, a video signal is used to test the subject's perpetual abilities. The computer will need to complete the tasks below to pass this test:

- Vision to observe or see objects
- Use robotics to manipulate and move objects around

The basic composition of most AI is based on the above mentioned six-disciplines. Turing deserves all the credit, since this test is relevant even after 60 years of its making. Researchers, however, pay little attention towards passing the Turing test, as they believe it is more important to understand the internal workings of the intelligence rather than trying to know how to duplicate an archetype.

The Cognitive Modeling approach: Thinking humanly

If we say that a particular computer program thinks like a human, we must first know how the human mind works. We must figure out the internal mechanisms of our brain. There are three ways to find this. One is through introspection- where we try to understand our thoughts as they pass us by. The second method is through psychological experimentations- where an individual is observed in action. The third method is brain imaging- where the brain is observed in action. After gathering precise information about how our brain functions, it is possible to convert it into a computer program. If the program's input-output behavior is similar to its human behavior,

it proves that the program involves certain mechanisms that could also be operating in our brain.

Cognitive science is a discipline that brings together AI computer methods and techniques from psychology that can develop accurate and testable theories of the human mind. A lot of confusion was created in the early days of AI. A researcher would claim that an algorithm performs well on a certain task, and it is, therefore, proof that it is a good model of human performance. Modern-day researchers separated this into two claims, and this distinction has helped in the rapid growth of both fields.

The Laws of Thought Approach: Thinking Rationally

Aristotle's syllogisms gave us patterns for argument structures that drove to correct conclusions in the correct circumstances. These laws of thought were believed to give insights into how our mind operates and the field called logic stemmed from this very school of thought. The 19th-century logicians came up with different laws that added meaning to objects and its relationships with the world around. The logistic approach to artificial intelligence is about trying to build programs based on these relationships to create intelligent systems.

This approach has two crucial obstacles. Firstly, it is difficult to take informal knowledge and create formal terms, especially when this knowledge is not 100% certain. Secondly, there is a difference between solving a problem in practice and solving it in principle. Even if the problem has just around a hundred facts, it will require a lot of computational resources unless it has a framework that provides information on what step to try initially.

The Rational Agent Approach: Acting Rationally

Computer agents perform several functions such as; perceive its environment, operate autonomously, persist over a long period of time, create and achieve goals, and adapt to change. A rational agent comes into use where there is uncertainty; it helps achieve a goal or provides the best-expected outcome. In the laws of thought approach, the emphasis is placed on the right conclusions. Drawing the correct conclusions can always be expected from a rational agent, but correct conclusions do not always sum up to rationality. In certain situations, there is no right thing to be done, but something has to be done. There are also ways where acting rationally does not involve any inferences.

The rational agent approach has two positive points. It is more general than the laws of thought approach because the right conclusion is one of the many possibilities for achieving rationality. It is also quite responsive to scientific development than approaches that are based on human behavior. An important point to keep in mind is that we will find as we go on that reaching the perfect level of rationality - that is always doing the right thing - is not viable in complicated environments.

Chapter Three: Basic Concepts in AI

Data

The static, dynamic, factual, raw, and discrete observations of a particular area of interest are data. The systematic processing of this data produces information. Numerical values within most environments always represent data. Data can be defined as the physical and transactional records of an enterprise’s activities, and this will serve as its basic building blocks of information. They are products of observation and to observe is to sense. Therefore, it has to be processed before utilizing the information. Data to an analyst, problem-solver or investigator might mean sentences, numbers, assumptions or records. It is the known factors irrespective of its form or origin. In other fields, it could mean facts or hypothesis where data is used as an assumption.

Although data is used as evidence, it is not always true. It is difficult to prove whether the given set of data is true or false. It requires advanced processing to gain more information from the existing data. For instance, the temperature of one particular day is a singular atom of data, and it is treated as a fact. There will be many atoms like this one, and they can be combined using standard logical operations. Universal statements like "the daily maximum temperature should be above 30 degrees" should also be considered. These universal statements are much stronger than the atoms of data; therefore, it is difficult to finalize on what the truth is. These types of data should go through another screening to produce the required authentic information.

Data is basically empirical, and it is difficult to value the fictitious non-empirical data.

Information

The structured, organized, processed, and presented data that can be used is information. This processed data is what makes decision making easier. There is, however, some information that is not data. Processing data is about calculations of data, aggregation of data or corrections on data, in such a way that there are generations of information.

You can add value to data in different ways:

- Contextualized- the purpose for which the data was collected is given
- Categorized - the units of analysis or major components of the data is available
- Calculated- if the method used to analyze the data is given, that is, whether it was done mathematically or statistically.
- Corrected- if the errors were removed from the collected data
- Condensed- if the data was concisely summarized

Furthermore, information can be accessed, transmitted, distributed, generated, stored, processed, consumed, produced, searched for, compressed, used, and duplicated. Information can also have different attributes assigned to it. It can be quantitative, qualitative or sensitive information.

Knowledge

Knowledge is defined as human understanding of a subject gained through complete study and experience of the subject. You can consider information and data as a collective mass, but knowledge is based on understanding the problem area, learning, and thinking individually. Knowledge is derived from information the way information is generated from data. It is the fusion of human perceptive processes that results in inferences. It is a true belief system that is associated with human actions and is generated from a flow of information. It is mostly subjective, personal and inherently local. It does not exist objectively.

Intelligence

A deeper understanding of concepts and models leads to a higher level of knowledge called wisdom. An application of principles, morals, and expertise is required to gather and use wisdom. This type of maturity comes with age and experience. This idea of wisdom was taught to us by ancient philosophers like Plato and Aristotle even though there are controversies around it. Further advancement to this wisdom is known as intelligence.

Artificial Intelligence

Artificial intelligence involves a wide array of tools and technologies that can be combined in a variety of ways to perform, sense and cognize along with the capability to learn from experience and adapt with time. Knowledge based systems were the most predominant and practical branch of AI. It was applied in several areas and organizations of the world. They were labeled as expert system shells and came with hundreds of tools. These systems were good enough to create an individual discipline for itself. Other concepts such as rule-based knowledge representation, machine learning, reasoning with uncertainty, verification of domain knowledge were other areas that grew in the field of academics.

Classifications

Classifications refer to the division of data into separate classes by employing models. Training data sets are created for these models, which will be assigned to each class so that the algorithm can learn. Real time information is given to this model that belongs to these classes. This model generates a relationship that exists in the input data depending on what the model knows from the data set. Support vector machines and decision trees are some common classifications. The algorithms will always require a clear-cut definition of these classes; therefore, classification can be considered as a form of supervised machine learning.

Association

Association is a type of analysis that helps identify any existing links between data instances. It can be explained through a market basket analysis, where a link can be found in the data instances of any shopper and what's placed in the basket. The information could be virtual or real, and the algorithm always provides comparisons. Associations come under the unsupervised machine-learning spectrum.

Decision Trees

Decision trees are classifiers. They include two tasks: tree pruning and tree induction. The latter involves pre-classified data instances that are taken as input. Decisions are made based on splitting the data set and recursing on the split data set until every set is categorized. The main aim while building a tree is to create child nodes that are considered pure. This makes sure that the splits for classifications are small in number. A complete decision tree will always be complex and may include a lot of unnecessary structure that will be difficult to distinguish from the other type. Tree pruning refers to the removal of these unnecessary structures and make it more comprehensible. It also helps reduce over-fitting.

Deep Learning

Deep learning involves several hierarchical layers to process the data in a non-linear manner, and some lower-level ideas help define the higher-level ideas.

Deep learning can be considered as a class of machine learning methods that exploit many layers of non-linear information processing for the supervised or unsupervised method of extraction and transformation and pattern recognition and classification.

Chapter Four: How Machines Learn

An intelligent agent is an entity that can perceive its environment using a sensor and act upon its environment through its actuators. For instance, we have our eyes, ears and other sensory organs. Our legs, hands, vocal tract, etc. act as the actuators. A software agent gets input through file contents, keystrokes and network packets. It acts on the environment using a display screen, writing files or by sending network packets. An agent can be considered as learning if its performance improves on future tasks after observing its environment. Learning can range from something as basic as writing down a phone number to complexities, such as theories of the universe.

1. Why do we need an agent to learn?
2. If the design of the agent can be improvised, then why not just program it into the agent?

There are three main reasons for this. First, the designers can't predict all the type of situations the agent will be subjected to at a time. For instance, if a robot is designed to navigate mazes, it has to learn the layout of every maze it encounters. Second, changes are inevitable, and it is difficult for the designer to predict it over time. For instance, if a program is designed to predict future stock market prices, then it should learn to adapt to low and high conditions. Third, it is a difficult task even for human programmers to program a solution by themselves. For instance, people are generally good at recognizing family members, but even the best programmers find it nearly impossible to design a computer to do it without a learning algorithm.

Forms of Learning

Any agent can be improved by learning from data and the techniques used to improve it depend on four major factors:

- The component that is to be improved
- The prior knowledge the agent has
- The type of representation used for the component and data
- The feedback that is available to learn from

Components to be Improved

The agent would include components such as:

- A direct mapping system about current conditions to actions
- A means to draw inferences which are relevant to its environment from its percept sequence
- Information regarding how the environment evolves and the agent's possible responses to it
- Utility information showing the desirability of world conditions
- Action-value information.
- Goals that maximizes the agent's utility

For instance, if an agent is being trained to be a taxi driver, each time the instructor shouts "brake" the agent tends to learn a condition, that is, action rule for when to pull the brake (component 1); it also learns that the instructor does not always shout. The machine is told that there are buses on the road using some images from a camera. The machine will then learn to recognize them (2). By attempting actions and observing the results, like braking hard on a wet road, it learns the results of its actions (3). Lastly, when it does not receive any instruction or tip from the passenger, it can learn about its total utility function (4).

Representation and Prior Knowledge

There are many examples of representations for agent components, such as first-order logic and propositional sentences in a logical agent; Bayesian networks for a decision-theoretic agent and so on. Efficient learning algorithms are used in all these representations. Various other types of learning include inductive learning which is where an agent learns from a general function or rule, analytical or deductive learning is where the agent goes from a general rule to a new rule that is logically generated and is preferred as it leads to more efficient processing.

Feedback to Learn From

Three types of feedback govern three main types of learning

Unsupervised Learning

The agent figures out patterns in the data that is given without any feedback being supplied. Clustering is the most common example of unsupervised learning. Clustering involves detection of useful clusters from the input. For instance, a taxi agent would slowly develop a concept of good days and bad days of traffic without being given any prior knowledge or examples by its instructor.

Reinforcement Learning

In this type of learning the agent uses chains of reinforcements to study. For instance, when the taxi agent is not given any tip at the end of a journey, it is an indication that something wasn't done right. Then the agent decides as to what action was responsible for the negative response.

Supervised Learning

An agent discerns some examples of ‘input-output pairs’ and ascertains a function that links the input to the output. In the example mentioned above (first component), the inputs are percepts and the instructor who says, “Brake”, gives the outputs. In the second component, the inputs are visuals through a camera and the outputs come from the instructor who says, "This is a bus." In component 3, the function is the theory of braking from states and the stopping distance in feet. The output, in this case, is available from the agent's percepts; the teacher will be its environment.

Semi-supervised Learning

We will be given labeled instances and must try to understand what we can from a collection of unlabeled instances in this type of learning. The labels themselves may not be the truths that we think it is. For example, you are trying to build a program that can guess someone's age based on their photograph. You can start by collecting labeled examples like photographs with the person's information. This is supervised learning. In reality, there will be people who must have lied about their age. The inaccuracies might be systematic and to figure it out unsupervised learning should be used. This

will include images, reported ages and true ages. This creates a continuum between supervised and unsupervised learning.

Chapter Five: Machine Learning

The combination of automated data-gathering systems along with inexpensive mass memory storage devices has led to the acquirement and preservation of large amounts of data. It could be temperature readings, pressure readings, financial data, point-of-scale customer purchases, news feed, etc. Efficient data mining techniques will be required to quantify, extract and classify useful information from this large volume of raw data. Machine learning has become vital for the analysis of massive amounts of data. The larger the volume of the data, the better it is.

In most cases, machine-learning techniques build hypotheses from data. For instance, if a large volume of data indicates many instances of swans being white and zero instances of swans being other colors, then the learning algorithm will conclude that “all swans are white.” These conclusions are inductive and not deductive. Deductive learning is about inferences based on logic and the premises, while inductive learning is about constructing hypotheses, which always involves alterations with additional data. That is to say that, there might be an undiscovered island full of black swans. Inductive learning based on large volumes of data has extreme importance. Undeniably, science itself is based on inductive learning and its inferences.

Memory-Based Learning

The most practiced approach by AI when dealing with large volumes of data is to reduce the volume in some way. For instance, a neural network will represent what is fundamental about a large training data set with the networks structure and weight values. Similar to this is the Bayesian network where data is condensed into the network's node structure and conditional probability tables. We now have rapid-access computer memories that help to store and compute large amounts of data without any prior compression or condensation processes. These methods do not condense data unless it is needed for a specific task. All the required reduction is performed when a decision is to be made. Here are some of the memory-based learning methods:

The nearest neighbor method- this is used for classification of a point in a multidimensional space. The k-nearest neighbor rule, for instance, allocates a data point to a similar category as the majority of the k stored data points that are nearest to it. The k-nearest neighbor rule is an ideal example of memory-based learning, and it raises a lot of questions about probable extensions. Firstly, to use the nearest neighbor rule, each datum has to be a list of numbers, that is, a point or vector in a multidimensional space. The first question that can be raised is how should the data be represented for the nearest neighbor method to be applicable? Second, how will the distance be measured between data points? When data is represented as points in multidimensional space, Euclidean distance can be used.

In case the data is not represented as points, then a different technique to measure data closeness must be employed. Third, when considering the closest data points will the points closer to each other influence the result more than the ones that are at a distance from each other. The most basic k-nearest neighbor can be extended by weighing the importance of those data points, based on their closeness. Mostly, a kernel is used to give a gradually diminishing weight to very distant points. Fourth, what should the value of k be?

How many neighbors are going to be employed to make a decision about one particular data? With the right kind of kernel, every data point can be considered. The points that are very far will have zero or negligible influence

on the particular decision. The relevant question now is how distant should a kernel be to influence the outcome. Lastly, after considering all the weighed neighbors, how can a decision be made about assigning numerical values or value? Should it be according to the majority vote of the neighbors, or should it be the average of the weighted members? Different versions of statistical regression methods can be used based on this choice.

Case-Based Reasoning

This is a subfield of AI that can be viewed as a generalized type of memory-based learning. Case-based reasoning involves a library of cases that are used to analyze, interpret and solve new cases. For instance, in the field of medicine, therapeutic and diagnostic records of patients constitute a library of cases; once a new case surfaces, cases that are similar to it can be found from this library, which would help diagnoses and therapy. In the field of law, older legal scenarios are used to interpret and study new cases.

If there are cases that are similar to a new case, they are considered as neighbors in a general space of cases. To find close neighbors, the idea of proximity should be according to the measure of similarities. The most famous pioneer of case-based reasoning is Janet Kolodner, who is a professor of cognitive science and computing. She describes the process in this way: cases that need to be retrieved are the ones that are good cases. They have the potential to construct relevant ideas for a new case. This retrieval is worked out using a feature of the new case as indexes into the library of cases. Cases with subsets of those features labeled to them or features that can be derived from these features will be retrieved. We will now have an array of promising cases to choose from and reason with.

Case-based reasoning has turned out to be the most successfully applied AI technologies in recent years. They have prominent application in commercial and industrial areas. Customer care centers and diagnostic systems are the most common applications.

Decision Trees

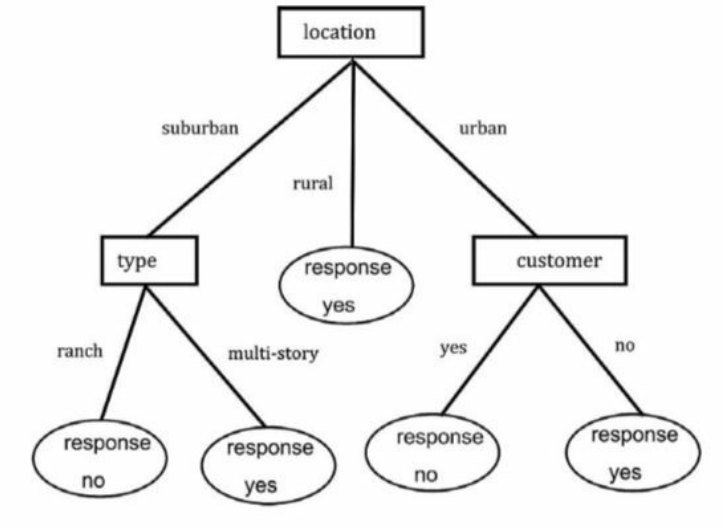
This is a new type of advanced machine learning technique that involves the automatic construction of structures or networks known as decision trees form large sets of data. Decision trees involve a series of test that helps determine a numerical value or category to assign to a data record. They are most applicable for non-numeric as well as numeric data. The methods used to construct them automatically will be described in detail.

Data Mining and Decision Trees

Data mining involves the extraction of useful information from a large volume of information. For instance, think of a database that includes credit card behavior. This will come with payment records, average purchase amounts, average balances, late fee charges and it will go on. By using the right data mining methods, you will be able to find out that people with higher late fee charges, higher average purchases and other similar features are most likely to have high average balance.

One important method of data mining uses data to construct decision trees.

This can be elaborated using a simple example. Consider a company such as Walmart. Imagine that they have maintained a database with information about customers that they have previously mailed regarding a discount coupon. The information would include details such as the location of household, type of house if the customer has shopped with Walmart before or if the customer has ever responded to any of their customer mails before. If this information were to be tabulated, the information in each row would be called a record. The item on top of the column would be the attributes and the items listed in the columns would be its values. After analysis of this database, a decision tree can be constructed.



The figure illustrated above can give you an idea about which household responded to the mails and which ones didn’t. The tests on the values are at the interior nodes of the decision tree (boxes); the results are at the tips of the

tree (the leaves) in ovals. These trees help at predicting the expected response before sending the mails. Several methods have been developed to construct these decision trees. Here are some examples

EPAM

This system was developed in the late 1950s and is the earliest method used to construct decision trees. Edward Feigenbaum developed it as his Ph.D. dissertation. It is named after its acronym, Elementary Perceiver and Memoriser. The goal of his research was to predict and explain the phenomenon of human verbal learning. This program is still considered as a vital contribution towards understanding human intelligence and AI research.

CLS

This program was developed in the 1960s by Carl Hovland and his student Earl B. They developed a computer model of human concept learning. CLS is the acronym for Concept System Learning.

ID3

ID3 is the acronym for Iterative Dichotomizer. J Ross Quinlan developed it in the late 1970s. Using this program data records could be divided into data sets iteratively until they come under two distinct categories. Later versions of this program had more than two categories.

C4.5, CART and Successors

Quinlan continued to develop advanced and applicable decision tree construction systems. He developed the C4.5 system that could compute databases with attributes that had continuous numerical values and categories. During this period, several developments that are of significance to the AI research took place. This was due to the collaboration between statisticians and AI people. CART stands for classification and regression trees. It was developed by Jerome Friedman and is similar to C4.5 when dealing with numeric attributes.

Inductive Logic Programming

Many techniques have been developed to study relational rules from databases and other background knowledge. Quinlan developed the earliest system for studying relational rules called FOIL. Since the rules learned can be applied in computer languages such as PROLOG-which is a language based on logic- a field was dedicated to learning them called Inductive Logic Programming.

Neural Networks

Neural networks help to identify links or patterns between data that has no direct relationship between input and output or instances. The neural nets would identify a pattern between data sets. Let us consider constructing a system that to identify animals based on their attributes or features. If we want the system to identify a tiger, the model that you give should not be of a four-legged animal and warm blood. As the system would not be able to understand whether it is a tiger or a lizard or even a dolphin. A combination of these attributes, however, can help the system identify the right animal. If you divide these attributes by a thousand, you will understand the complexity of neural networks.

A broader understanding of artificial neural networks can be obtained by looking into how the human brain works. There is a vast difference in scale-human neural networks that continue to increase in size as time goes by.

Artificial neural networks might include a million neurons, and the brain has roughly 85 billion neurons. While the neurons in the brain are interconnected, the artificial neural network is unidirectional. These are distributed through the layers listed below:

Input layer

This layer includes neurons that receive data and transfer it to the next layer. The number of neurons depends on how large the attributes in a particular data set are.

Output layer

It involves nodes that are required to build a model. In a classification system, one label will be assigned to a node and in a regression system; there will be one node that will transfer out a value.

Hidden layer

This layer exists between the input and the output layer. The nodes that are a part of this layer transform the input data. As the nodes get trained, they get better at predicting outputs.

Unsupervised Learning

The learning methods, such as neural networks and decision trees, come under examples of supervised learning. In this learning, one tries to learn how to classify data from a large training data set whose classifications are known.

The supervision that takes place in this system refers to informing the system about how each datum in the training set is classified. It is possible to produce classifications from data alone. Such techniques come under the unsupervised learning category.

Let us consider an example where data that is to be classified is represented in a two-dimensional space with points. The coordinates of these points are assigned values of f1 and f2, which are numerically valued features of the data. In the representation, the category can be denoted by small squares for one category and small circles for the other category. If the points are represented in this manner, they can be used as training data for supervised learning techniques.

In case we have unlabeled data points, through visual inspection, we can identify points that are arranged in clusters. Each cluster might be points that come under the same category. Therefore, automatic identification of clusters and their boundaries can be used as techniques for unsupervised learning. This method has been used by AI researchers to identify clusters of training samples. K-means method is the most popular one among them. It is used by the repetition of the steps mentioned below:

1. In the space of samples, install a number, call it k, of cluster seekers at a random location
2. Group the training samples that are close together in each cluster seekers
3. For each of these groups, compute the center of gravity or centroid
4. Transfer these cluster seekers to the corresponding centroid of each group
5. Repeat these processes until none of the cluster seekers have to be moved
6. Once this process has ended, the cluster seekers will be at the centroids of the training samples which can be identified as clusters or separate

categories of data

Now you can classify new data that is not a part of the data set by calculating which cluster seeker it is closer to. This will depend on the ability to predict the number of clusters, k. Techniques for finding the number would involve adjusting the number in such a way that the points within these clusters are closer than the distance between these clusters.

Reinforcement Learning

This type of learning is a combination of both supervised and unsupervised learning. An example of this is a robot, which can execute an ongoing sequence of experiences. Another ideal example would be learning to play chess brilliantly, just with information regarding winning or failing at the end of the game. Although this hasn't been accomplished yet, some programs can play backgammon and learn control flight of helicopters. Reinforcement learning is a subdiscipline of AI that is completely technical and multibranched. In its most basic form, this type of learning is about traversing through a set of states, transferring from one to another until a state where a reward is attained. It is just like the experiments where a rat learns how to navigate through a maze.

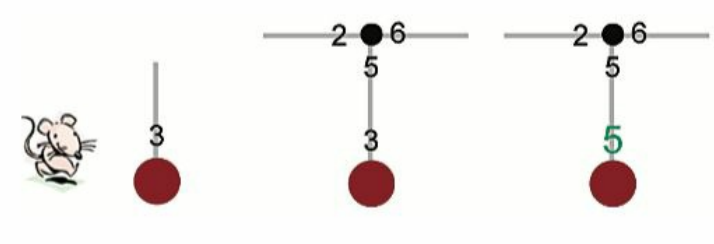
Let us use the maze as an example to understand the basic concepts of reinforcement learning.

The rat is expected to go from point A to point B, where there is a reward (cheese). There are points in the maze that the rat might recognize from before. In technical terms, these are known as states. When the rat reaches each of these locations, it can choose between, let's say, four actions like turning left, right, forward or backward. The actions will depend on the state. Every move takes the rat from one state to the next one in the maze. This set of actions and states can be considered as a graph for better understanding.

Let us now consider a robotic rat, who is learning how to navigate through a maze. Initially, the robot goes through the maze without any prior knowledge about it, that is to say, that it does not know the effects of its actions in any state that it is subjected to. If it were provided with a map or a graph, it would learn to find a path to the endpoint using that information. Trial and error methods are one way to use the graph of states in order to find the links for moving through the maze. A second option and the most commonly used reinforcement learning method is to label all the states the robot has passed through randomly. In technical terms, this is called a policy for navigating the maze that connects to every action with a labeled state. The optimal policy would be the one that connects each state to that action which leads through the shortest path in the maze. Reinforcement learning is about learning this

optimal policy or policies.

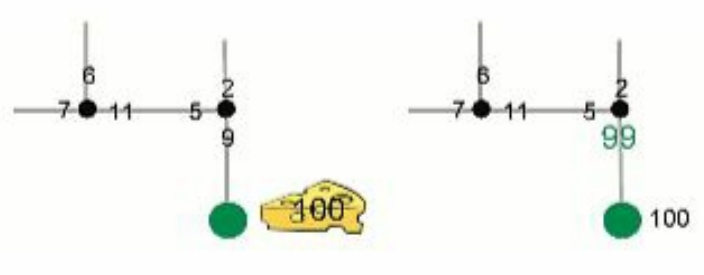
Another method involves assigning a value or a number to every action at each state and then modifying these values until it reaches the endpoint. This technique is known as Q-learning. It starts with the robot allotting a name to a state where it starts and also allotting a value or number to all the actions it takes in that state. We can assume that the robot memorizes all the states that it visited during its learning process and can differentiate between these states and the new ones. During this learning process, the robot takes the action that has been assigned the highest value number.



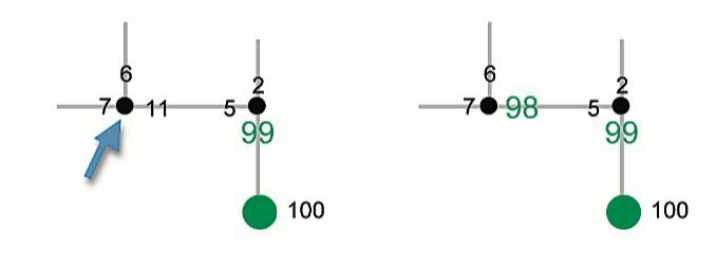
Since there is just a single action during the initial state, which leads the robot to a new state, it allocates a new random value to the actions that are possible in that particular state. This step is illustrated in the figure above. Next comes the important step in learning, as the robot is aware that it can reach a new state that has actions with the highest value, which is 6. It then updates the value of the action named 3 that leads to that state by giving it a value that is consistent with an action that is worth 6. This adjustment of 3 does not go all the way to the highest number, 6, but just to 5.

This is illustrated in the right-hand side sketch where the adjusted value is represented with a larger font and shade than the other fonts. This process is repeated. In every state, the action will depend on the largest valuation number, which is then adjusted to a valuation number that is closer to the largest number in that state. Although this process starts with random numbers being assigned to states, the trial and error method will lead it towards the state where a reward is given. The action that leads to this reward stage will be assigned a valuation number which is almost close or similar to the value given to the reward state. This step is illustrated in the figure below.

The right-hand side figure shows the adjusted value that led to the goal-achieving action. It is important to note that, if the robot ever finds itself in a similar situation, it will use this reward-based value backward to propagate.



Refer to the illustration below to follow what backward propagation is. Let us consider a scenario where the robot finds itself in the state marked by the arrow in the figure. It follows the largest valuation number from that state. And this leads it to another state that is closer to the goal state. The action with the valuation as 99 will be the state that leads to the goal state. Therefore the valuation of the action that was just taken is altered from 11 to 98, as shown in the figure on the right side. Following this method of increasing the value of actions in the states that are closer to the goal using backward propagation is just as rewarding as the goal state themselves.



Most iterations of reinforcement learning can be elaborated as follows:

1. There will be rewards provided in more than one of the states. This means that it is not just about a single goal state but multiple states that will lead to a reward state. Rewards will be denoted by numbers that can be positive, zero, or negative.
2. Instead of finding a policy that correlates with an optimal path that leads to a single goal state, one learns policies that increase the number of rewards that can be expected in the future. Generally, when learning a policy, rewards that are expected in the distant future are not considered as much as the rewards that are expected more immediately.
3. And an action taken at a particular stage might not always lead to the same state. One could try to learn the probabilities that might lead to other states, and this comes under a learning method known as prioritized sweeping. Under the Q-learning process, these probabilities are not considered because they influence the learning process that is assigned to state-action pairs.
4. Another complication that can affect the process is if the sensory apparatus of the robot is not accurate. As the robot will have incorrect knowledge about the state that it is in.

Chapter Six: Big data

Essentials of Big Data

Applications that use various algorithms to perform tasks are used by computers to manage data. In its most basic form, an algorithm is a procedure that helps perform a set of operations. Create, read, update and delete are four basic data operations of an algorithm. Although these operations look simple, they are the basis of everything that is done on a computer. As the dataset becomes, more complex, the computer uses the algorithms found in the application to perform more work. Big data involves complex data sets, and the computer uses pattern recognition in a non-deterministic manner to process it. Therefore, to build a system that can learn, a complex dataset is required for the algorithm to use pattern recognition. Additionally, this pattern recognition also requires a simple subset for statistical analysis or predictions of the whole data set.

The most common example of employing big data is in online databases created by sellers who track consumer purchases. Identifying the right sources of big data helps create machine-learning instances where a machine can learn and produce data in the desired manner. A combination of big data and statistics can create the perfect machine-learning scenario, where the machine can predict the probability of any given event. It is the algorithm, which can affect the learning process and the result. A training process is required for algorithms as it changes how it views big data. This training involves subsets of data that are used for creating patterns that the algorithm has to use to distinguish specific cases from general cases that were given for training.

Although big data implies a lot of data, it is extremely different from any other large database. It is more complex and has depth. When compared to general programming, big data sources can help solve problems more efficiently with just enough details. For instance, consider Google’s self-driven car. Not only should the car have information regarding the hardware and its position in space but also the knowledge as to how it affects human decisions, environmental conditions, road conditions and other vehicles. The data source would, therefore, include numerous variables that have some influence on the workings of the car. With a traditional programming system, data will not be efficiently processed in real-time. If the car is going to crash,

you will need the system to predict it before the event and not five minutes after it has crashed.

The problem is in learning how to control big data. The sheer volume of a data set is not the only thing that has to be considered, but also where it can be stored and transferred for further processing. Generally, it is stored in memory by developers for fast processing. Options such as hard drive are time-consuming and expensive. Another concern when it comes to big data is privacy, but machine learning is all about specifics, but it’s not that helpful. It determines patterns by analyzing training data, and the training algorithm will perform tasks that were not programmed by its developers. Therefore, personal data is not relevant in that environment.

Sources of Big Data

For a machine learning application, you first require a big data source such as corporate-owned databases. In order to create useful sources of big data, new data has to be created. Developers construct existing data sources and depending on the needs of the clients-server architecture, but this may not work for machine learning instances as they are optimized to save space on the hard drive. Therefore, they do not possess the required depth. When you become more familiar with machine learning, you will come across scenarios that cannot be solved using a standard corporate database.

Public sources

Universities, non-profit organizations, governments and other organizations often have an open database that is available to the public. You can use this alone or combine it with other databases to create big data for machine learning. For instance, you can combine various Geographic Information Systems (GIS) to build big data to help with decisions such as the positioning of new factories or stores. This type of machine learning system can consider all kinds of information, such as the amount of taxes that are to be paid and even the elevation of a portion of land. The biggest advantage of using public data is that it is mostly free of cost even for commercial use, and in some cases, you have to provide a nominal fee.

Also, the database would be maintained in perfect condition as the organization has some mandate; they use this data to generate income or utilize the data internally. There are a number of issues you have to consider while using public source data. Firstly, you have to make sure that you can attain some useful information. Listed below are some criteria’s you can follow while making a decision:

1. The format of the data source
2. The fee, if any, for using the data source
3. Permission to use the data source since some sources are copyrighted
4. Potential issues when cleaning the data so that it can be used for machine learning.
5. Provision of proper infrastructure or access to the data source

Private sources

Amazon and Google are examples of some private organizations that maintain a large database, which has all sorts of useful information. In this scenario, you are expected to pay a fee for access to this data, especially when the commercial setting is used. The one major restriction is that you may not be able to download the data to your server and this can affect your machine-learning environment. For instance, certain algorithms work slowly with data that can only be accessed in small pieces. The major benefit you get when using data from a private database is the perfect consistency that you can expect. When compared to the information from a public server, this data will be cleaner. You will also have access to a bigger database with a diverse amount of data types. All this depends on when you get the data.

Building new data from existing data

You can use existing data as a starting point to create a new data source since old data is not best suited for creating new machine learning scenarios. For instance, you might have a customer database that includes all customer orders, but this data is not suitable for machine learning as it lacks the tags or categories required to classify the data into specific types. Machine learning mostly works on 80% of the data, but the remaining 20% of the cases will have to be worked out by professionals who decide how to react to a specific data and then act upon it. In this scenario, you will need people who would include additional information such as tags.

Using existing data sources

Your organization has hidden data in all kinds of places; the problem is in learning how to recognize this data as useful data. For instance, in order to track an assembly line, you may have employed sensors that track how the assembly process works and make sure that the line remains systematic. The visuals from these sensors can be potential input for a machine learning scenario, as they track how the product moves, how it affects customer satisfaction and even the fee you pay for postage. The basic idea is to transform the existing data into a new kind of data that lets you optimize the functioning of your organization.

Statistics and Machine Learning

Although statistics and machine learning are part of very different fields, they have many similarities. Statistics is one among five schools of thoughts that make machine learning attainable. The five schools of thought are:

- Symbolists: This stems from logic and philosophy. It relies on the inverse deduction for problem solving.
- Connectionists: This stems from neuroscience. It depends on back-propagation for solving problems.
- Evolutionaries: It originates from evolutionary biology. It depends on genetic programming for problem solving.
- Bayesians: This originates from statistics and depends on probabilistic inferences
- Analogies: This originates from psychology and depends on kernel machines to do love issues.

The end goal of machine learning is to merge strategies and technologies within the five Schools of thought. The goal is to create a single algorithm or master algorithm that can learn just about anything.

Role of Algorithms

In the field of machine learning, everything pivots around algorithms. It is like a procedure or formula employed for problem solving.

The domain of the problem affects the type of algorithm that is used, but the basic hypothesis is always the same. It could be about any problem, just like driving a car or playing chess. Initially, the problem would look extremely complex, but, in the end, the basic goal is to get the passenger from one position to the next without creating any damage.

You can imagine an algorithm to be like a box or container that stores a procedure which can be used to solve a specific problem. They approach data with the help of a succession of well-defined states. The states will be defined but not necessarily in a deterministic manner. The ultimate goal is to design an algorithm to solve problems. There are cases where inputs that are given to an algorithm serves as the basis for just defining an output. The output will remain the main focus.

Transitioning between states should be expressed using a well-defined and formal language that the computer can process. When solving a problem or processing data, an algorithm tends to define, clean and carry out a function. This function will correspond to the problem that is to be solved by the algorithm.

Strategies for Algorithms

Symbolic Reasoning

In symbolic reasoning, deduction widens the sphere of human knowledge, and induction uplifts the level of human knowledge. An inverse deduction is a term that is commonly referred to as induction.

Induction generally opens up new areas for exploration and deduction helps understand the depth of this field. Therefore, induction is the scientific part of this kind of reasoning, and deduction is the engineering side. These two strategies always work in close association with each other.

It starts by identifying a new area for exploration and then inspecting this area to understand if its problems can be solved. For instance, consider a tree with green leaves, through deduction; we could say that if the tree is green and, green leaves mean that it is alive, then this tree must be alive too. Through induction, we could infer that if the tree is green and if the tree is alive, then green trees are alive.

Brain Modeling

This is the most popular strategy among all. Silicon is used in place of neurons, and the brains functions are reproduced. Every neuron decodes one part of a problem, and when several neurons work in parallel to each other, the problem can be solved as a whole. Backpropagation is the method by which errors are removed from networks that are built to be like the neurons in our brain.

It works by altering the weights and biases of the network. Weights are inputs such as values or numbers that affect the result and biases are the attributes that are finally chosen. The aim is to keep altering the biases and weights until the point where the actual output is similar to the desired output. Since each neuron solves one part of the problem, when they come together, they pass data to the next neuron, thus forming the final output.

Evolutionary Modeling

This system depends on the theory and principles of evolution for problem solving. Which means that this technique is about the survival of the fittest, that is, draw out any aspect that doesn’t fit with the desired output. A fitness

function would regulate the feasibility of each function in problem solving. A tree structure is used to determine the best solution depending on the output function. The function that passes through a level gets to be part of the next level of functions. This next-level will be closer to the end goal but may not be the solution for the problem. This strategy is heavily reliant on languages and recursions that reinforce recursions to decode a problem. The interesting outcome of this technique is algorithms that evolve; that is, one generation builds the next one.

Bayesian Inference

This is a statistical method that is used to solve problems. It is common knowledge that statistical methods can produce multiple correct outputs. The choice would then be about finding the right function with the highest probability of succeeding. For instance, while using these techniques, you can use a set of symptoms as data and connect a specific disease to one symptom, and this would be your output. There will be cases where multiple diseases will have similar symptoms, here the probability is critical as the user might be able to relate to a disease with lower probability, but it could be the right output in that scenario.

Therefore, this strategy is based on the idea that no hypothesis will be completely trusted without the evidence, which led to that output. Upon analysis, you can prove or disapprove the hypothesis. Similarly, it is impossible to establish which disease a person is affected by without testing all the symptoms. The most common use of this technique is in spam filtering.

Training Data Sets

The idea is that an application starts with a function receive data that is the input and processes it into results or output. For instance, when a programmer creates a function -*Add ()*, which accepts values 1 and 2 as input, the result is 3. The result of this procedure is a value. Previously, writing a program was linked to the understanding of the function, which manipulates the data that gives a result. Machine learning changed this approach. At present, you know that your inputs are 1 and 2. You also know that the result is 3, but you may not know which function has to be used to get the required result. With training, a learner algorithm will be provided with different instances of expected inputs and its outputs. The learner creates a function with this input. Therefore, training is a process that lets the learner algorithm map out a viable function for the data. The result will be the probability of a value or a category.

Machine learning is ultimately about generalization. The aim is to generalize the function to an extent where it works on data that is outside the training data set.

For instance, a small dictionary will have around 100,000 words. A training data set with 4000 to 5000-word combinations must build a generalized function which can then be used to filter spam in 2^100,000 iterations that the function will come across while using the actual data. From this perspective, training tends to look complex to learn. Even so, for creating a generalized function, there are three components the learner algorithm depends on:

Representation

First, the learner algorithm builds a model with a function that creates the desired output with the given input. The representation will include a set of models that can be learned by the algorithm. Fundamentally, for the learner algorithm to produce the desired result, a model should be created. In case the learner algorithm doesn’t perform the task, it means that the data is outside the scope of the hypothesis and that it cannot be learned. Representation is used to identify the features that can be used for the learning process.

Evaluation

More than one model cannot be created by the learner, but it is not capable of

distinguishing a good model from a bad one. An evaluation function helps identify which model works in producing the required result from the given inputs. This function also scores the models as more than one model can produce the desired result.

Optimization

At a certain point, the process generates a set of models that can produce the correct results from a set of inputs. The training process looks for a model that would work the best. The model that gets chosen is the final output of the training process.

Chapter Seven: Modern AI

There have been numerous developments in the field of AI in the past few decades, and this advancement came with a range of powerful computational tools. They are very effective when deployed due to the progressing power of the comparatively affordable computers, the convenience of large databases, and the expansion of the World Wide Web. Presently, AI possesses several human-like cognitive abilities and most of them are completely automated. Since AI contributes solutions to several real-world problems, graduates who specialized in AI studies pursue careers in organizations such as Google and Microsoft where they work on improving the cognitive abilities of machines, instead of academic AI research. Just as how there are several branches of engineering, AI has also expanded to some sub-specialties such as:

- Multi-based and agent-based systems
- Logic, Reasoning and Knowledge Representation
- Planning and Scheduling
- Robotics and Vision
- Uncertainty in AI
- Natural Language Processing
- Information Systems that are knowledge-based and web-based
- Machine Learning
- Achievements

Certain AI achievements from recent years are influential milestones of its advancement and there are other developments, which have invisibly become a part of our day-to-day routine. In between these two extremes are the achievements in the fields of science and commerce, but before that let us look into the most prominent systems that became popular during the 20th century, starting with AI game-playing programs.

Games

Even though getting computers to excel at games like chess and checkers are thought as silly distractions from more serious goals, game playing in computers served as a testing ground for exploring new AI techniques such as heuristic learning and reinforcement learning. In a previous chapter, it is explained how the reinforcement learning technique was used to design a machine that could excel at backgammon.

Chess

In 1997, IBM's Deep Blue Chess-playing computer defeated the then world champion Garry Kasparov, and this was big news. This computer used slow and fast evaluation functions to understand and evaluate the position of any piece on the chessboard. There were about 8000 features associated with a few values used in these evaluation functions, and these features were processed by special hardware. These are some of the differences IBM pointed out between Kasparov and Deep Blue.

1. Deep blue can evaluate and inspect around 200,000,000 chess positions per second while Garry Kasparov can evaluate and inspect three chess positions per second.
2. Deep Blue has more computational capacity and less chess knowledge. Garry Kasparov has a wide knowledge of chess and smaller calculation capacity.
3. Garry Kasparov uses his remarkable sense of intuition and feeling to play chess at the world champion level, while Deep Blue is incapable of intuition or feeling.
4. Garry Kasparov is capable of learning and adapting quickly to his mistakes and successes. Deep Blue is not a learning system; therefore, it cannot think and learn from its mistakes or from its opponent.
5. If Deep Blue should change the way it plays chess, the changes will need to be made to the software, and this can be done only by the members of the programming team. Garry Kasparov can change the way he plays as he wishes.

Checkers

In 2007, Professor Jonathan Schaeffer and his team published an article

announcing that the game of checkers has been solved. Schaeffer and his team have been researching it since 1989. They finally concluded that checkers always leads to a draw. They credited their result to the advanced AI algorithms, faster processors, larger memories and disks. It is not at all surprising that it took 18 years of massive effort and skill of computation to solve it,

as there are 500,995,484,682,338,672,639 different positions in the game of checkers.

AI at Home

As time passes, homes are becoming more intelligent. Here is a basic list of what you might find in a smart home:

- Air conditioning systems and thermostats that anticipate temperature changes and requirements of the occupants connect with other home devices and take required action in advance
- Microwave ovens that can scan barcodes on packages to decide how long the item can be cooked
- Smart running shoes which have a computer chip installed that can sense the runner's stride length and directs changes in the sole cushioning through a miniature screw and cable system
- Cameras with visions systems that help identify faces. These devices also have control over exposure, focusing, and framing
- Refrigerators that automatically create an inventory of the contents and inform the owner of required items
- Hearing aids that can adapt to ambient sound levels and block out the chatter
- Vacuum cleaning and floor washing robots
- Robotic pets and toys that can interact with people
- Caretaker robots for the elderly

Advanced Driver Assistance Systems

Modern cars have up to 50 microprocessor chips controlling features like fuel injection system, airbags, automatic transmissions, antilock brakes, security systems, cruise control and more. Not all automobiles have become autonomous; however, there are some that are being equipped with ADAs. Let us look at some features that have been installed in automobiles, and some features that are being designed for the future:

- Adaptive cruise control or ACC, for advanced intelligent control of acceleration, makes the vehicle accelerate or slow down depending on factors such as traffic. The vehicle can perceive traffic by laser sensors or radar
- Intelligent Speed Adaptation or ISA used for examining local speed limits. It warns the driver or automatically slows the vehicle down when it enters a zone with a speed limit
- Lane control systems for observing the presence of hindrances or vehicles in the other lanes. These systems can also be used to monitor whether the driver is moving into the wrong lane or not
- Automatic parking systems for aiding a driver in parallel parking
- Intelligent tire pressure control systems

Route Finding Maps

Most people nowadays have devices installed in their cars that interact with them for navigation purposes.

These devices give instructions for reaching your destination through GPS systems, speech synthesis and map databases. Map databases are graphs with nodes used to mark places and links that depict the connecting roads. These devices look for the link between two or modes to identify the best route. The most commonly used graph searching technique is A*, which is a heuristic search technique that considers both the distance traversed and the remaining distance to the endpoint.

Recommendation Systems

When you use Amazon.com, you might have noticed suggestions like, “we have recommendations for you.” These recommendations include a list of items that are similar to the items you may have purchased in the past or items that you may have looked at in the past but not purchased.

Amazon’s recommendations stem from what is known as social and collaborative filtering. Amazon maintains a database that includes every user's preferences for movies, books or other purchases. If user X’s preferences match with those of user Y, a collaborative filtering system comes up with a recommendation list that includes user X’s purchases to user Y. A very straightforward technique is used to understand the user’s preferences. Other examples of sites that use collaborative filtering include Netflix, iTunes, TiVo, etc.

A different type of recommendation system known as content-based filtering is used to find a user’s preference in books, documents, movies, etc. Content-based filtering, unlike collaborative filtering, does not look for similarities between different users but tries to look for similar content in the same categories.

For instance, in the case of documents, comparisons can be made using vector representations. Content-based filtering is also used for filtering unwanted emails like spam and websites. It can also be used to personalize your web search.

In Medicine

AI technology has been relevant to clinical medical practices since the early 1980s. Several AI-infused systems such as Agilent Acute Cardiac Ischemia Time-Sensitive Predictive Instrument which is an ECG device that predicts the probability of acute cardiac ischemia or a heart attack, MAC 5000 Resting Test System and Marquette 12SL ECG analysis program are just a few examples. OpenClincal maintains a list of websites which provides numerous decision systems such as Athena DSS for hypertension management, Iliad for internal medicine, TherapyEdge HIV for HIV patient management, Gideon for Infectious diseases and many others that are currently in use.

The ATHENA DSS is a system that provides a physician with advice regarding hypertension in a consistent manner that involves guidelines by the U.S Institute of Medicine. ATHENA learns a patient's clinical data against the knowledge for hypertension management in its knowledge base and generates patient-specific recommendations. A newer version called ATHENA-HTN is under evaluation. ATHENA has its roots in programs such as EON, ONCOCIN and MYCIN. Gideon is a program that assists in diagnosing and treating country-specific diseases.

Gideon generates a diagnosis depending on a large database of diseases, signs, lab findings, symptoms and countries. For computing, the probability of a disease, Bayesian analysis is used along with the patient's information as the input.

For Scheduling

Artificial intelligence techniques are commonly used in intelligent scheduling software. For example, AURORA by Stottler Henke Associates, Inc. is a system that specializes in the application of artificial intelligence and other sophisticated software technologies to solve issues that cannot be dealt with using traditional methods or approaches. AURORA is used by companies like the Boeing Company to assist in managing and scheduling the process of manufacturing the Boeing Dreamliner. Using AURORA, a schedule is created, which is then displayed in a graphical format that lets the user understand the scheduled events, resource distributions and the sequential relationships among the activities.

TEMPORIS, a development by United Space Alliance, LLC, is an intelligent spaceflight planning and mission management tool which can be used by the members of the crew onboard during a future space mission. TEMPORIS will assist the crew in scheduling every aspect of their in-flight lives such as day-to-day routines, activities, spacecraft maintenance, housekeeping and performing onboard experiments. At present, you require 50 mission planners and 24/7 for two weeks to schedule one day’s worth of activities on the International Space Station.

TEMPORIS can reduce this to a moment with the click on a command. AURORA software plays a major role in TEMPORIS.

For Automated Trading

For analysis of real-time trading data and news for automatic buy-sell decision-making on currencies, stocks and commodities, artificial intelligence techniques such as data mining, decision methods and texting processing are used. Updated news sources are all available in digital formats like Dow Jones Elementized News Feed and The Reuters News Scope Archive. For instance, Reuters provides its customers with news based programmatic trading techniques with a machine-readable and comprehensive archive of global news. Events are posted exactly at the time of breaking and it will be time-stamped to the second. The information will also be tagged to several metadata fields for easy machine consumption.

An article published in the New York Times on automated trading revealed that there was an association between terms such as depression, anxiety and bankruptcy. These types of correlations can be used to initiate stock trades with the help of stock trading algorithms. Stream base Systems is another firm that provides a product that can track risky market conditions throughout multiple markets and instantly figure out refined strategies to take short-lived trading opportunities.

In Business Practices

Business Rule Management Systems are progenies of the rule-based expert system from the 1980s. BLAZE ADVISOR 6.1, Information Builders Web FOCUS and ILOG's J Rules 6.0 are some examples. These rules express the way a company operates-, its constraints and policies. In BRM's these rules are generally expressed in English-like, computer-readable languages. Since these business rules change from time to time, they have to be updated to reflect the present company's policies.

In Translating Languages

Many commercial natural language translation systems are available nowadays. MASTOR is IBM’s speech-to-speech translator that can translate English speech into Mandarin speech. Similarly, BBN Technologies have also designed several speech-processing systems such as the Broadcast Monitoring System, which helps create a continuous archive of international television that is searchable. The system will be able to translate real-time audio stream into English automatically. The translation and transcript will be in sync with the video

For Facial Recognition

People are generally good at identifying or distinguishing familiar faces whether it’s in a photograph or in person. It is possible to do this, irrespective of the scale, lighting conditions, facial expressions, and pose.

Computers are advancing in this area and can be found in banks, airports and anywhere where your personal identity has to be verified. So, how far has computer facial recognition progressed? A 2007 National Institute of Standards and Technology conducted a facial recognition test and announced that, at low false alarm rates for humans, there was seven automated facial recognition algorithm that was better than or comparable to the human level of facial recognition under various lighting conditions. Additionally, three of the seven algorithms are better than or comparable to humans for the full range of false alarm rates measured.

A variety of algorithms have been created out of which some are based on popular pattern recognition methods. The sample features of the face from the image are pulled out, and the machine compares these features to those that are similar to it in a larger database of identified faces. Certain algorithms use Bayesian methods and HMMs.

Conclusion

You have reached the end of this book. AI relates to numerous fields, as mentioned in the book, and it is for this reason that researchers and engineers have discovered many powerful computational tools. To summarize the ideas and accomplishments of AI, it has been divided into four main groups, such as complete AI systems, architectures, processes and representations. Here are some major accomplishments in each group.

Complete AI systems (systems that do things): DENDRAL, LT, MSYS, speech recognition systems, Genghis, Deep Blue and other game-playing machines, driverless automobiles, CALO and RAX.

Architectures (Organizational principles for AI): Pandemonium, three-level architectures, BDI architectures, production systems, Blackboard architectures, cortical models, behavior-based architectures, ACT-R and SOAR.

Processes (Routines that do the work): Parsing, spreading activation, edge and region finding filters, STRIPS, A* and its progeny, resolution, beam search and the Rete algorithm, case-based reasoning, Bayes rule, clustering, GSAT and DPLL based methods, Q-learning, Kernel computations, Circumscription, genetic algorithms and latent semantic analysis.

Representations (structures that are created modified and accessed by processes):

Vectors, semantic networks, graphical models, decision trees, frames, grammars, blackboards, scripts, logical expressions, programs and augmented transition networks.

Many disciplines have contributed to AI’s successes. As mentioned at the beginning of this book, the early pioneers of AI took inspiration and clues on how to proceed from logic and mathematics, probability and statistics, neurosciences, linguistics, psychology, and computer engineering. Truly, the significant advancement that has occurred in this quest to date has to do with merging and utilizing the ideas from other disciplines.

No predominant theory about AI has surfaced yet and maybe there is not going to be one. As the pursuit continues, no one knows what intelligent

systems of the future will utilize through combinations of methods extended by AI supporting disciplines. There are chances where some of these technologies will be of better use with smarter technologies. There could come a time when books would be written on how two and half millennia after Aristotle’s reveries, we now have the tools that perform tasks either at our will or by themselves depending on the needs.

Sources

https://alex.smola.org/drafts/thebook.pdf

http://ciml.info/dl/v0_8/ciml-v0_8-all.pdf

https://courses.csail.mit.edu/6.034f/ai3/rest.pdf

https://artint.info/AIPython/aipython.pdf

https://www.cs.cmu.edu/afs/cs/academic/class/15381-s07/www/slides/011607comboIntro.pdf

https://stacks.stanford.edu/file/druid:qn160ck3308/qn160ck3308.pdf

http://etext.library.adelaide.edu.au/a/aristotle/a8pra/

http://www.research.ibm.com/journal/rd/021/ibmrd0201B.pdf

http://www.research.ibm.com/journal/rd/033/ibmrd0303H.pdf

http://www.aaai.org/AITopics/html/genalg.html

http://www.asc-cybernetics.org/foundations/timeline.htm

http://www.stat.ucla.edu/history/essay.pdf

Big Data

A Complete Guide to the Basic Concepts in Data Science, Cyber Security, Analytics and Metrics.

Hans Weber

Introduction

Welcome to the world of Big Data. This book will give you a complete insight into the basic concepts in Data Science, Cyber Security, Analytics, and Metrics. The first section of the book will describe the concepts of Big Data. Everyone knows that data and information in the current technical and business world are crucial. Big data refers to large amounts of data sized data in structured and unstructured arrangements that affect business on a daily basis. The amount of data does not matter. Data affects the businesses and organization of that business. Big data is examined for insights that generate better choices and planned corporate changes. Data science is a vast subject that uses technical approaches, procedures, processes, and schemes to get information from many sources.

The second portion of the book will describe the concepts of Cyber Security. Cyber-security is the field of protecting processors, servers, systems, grids, and information from malevolent attacks. Cybersecurity is also known as IT security. This guide will describe the different categories of cybersecurity in detail, the methods, procedures used, and how the systems are protected. The final portion of this guide will deal with analytics and metrics for big data. Analytics and metrics define the methods in which the data is converted into ideas that are applied across businesses and corporations. This book will also discuss performance analytics and metrics of big data, data science, and cybersecurity.

Chapter One: Big Data and Data Science

Nowadays, we all live in a world of information and data. In this digital age, Data is present everywhere. The digital data is increasing at a fast rate. According to statistics, the data gets multiplied by two times every two years (Vaidya, 2019). In a company or business, data is considered as a primary asset. The process of data processing includes concepts of data science, data analysis, and big data.

Some years ago, data was produced only by the work-force working in companies. The data that was generated before used to have some specific structure. It usually consisted of employee records, delivery information, the amount of money generated and hiring information, etc. With the advancements in technology, now, the data comes in large volumes and in many different formats. Data is produced in large volumes by networking, databases, and the new technology of cloud computing. A flood of new information is forming every day. Large amounts of data come in an unstructured format. One common example of unstructured or semi-structured data is the data used during the transaction of payment (Big data Basics). An innovative age of Big Data is evolving. The effects of the corporate sector and government organizations are enormous. Large databases are studied and analyzed to find new patterns of data. This extracted data is used for decision-making capability in business corporations.

Currently, data scientists are searching for new ways to make new uses of large databases.

To forecast the data and information, everybody needs data experts. The fields of Data Science and Big Data hold the vital key to the future of the data field. The field of data science is significant for improved advertising and promotion. Businesses are utilizing data and information to examine their advertising plans and generate better commercial advertisements.

The demand and requirement for data experts are growing so rapidly that it was predicted that by 2018, there would have been a 50% gap in the availability of data experts versus the requirement (Thompson). The

increasing significance of data science has led to the development and prominence of data scientists. These data specialists are currently an essential part of trademarks, industries, community organizations, and non-profit administrations. The data experts work determinedly to organize the enormous quantity of information and regulate appropriate designs and strategies in the data. In this way, the data can be efficiently used to understand future objectives and ideas.

History of Data Science

The beginning of the field of data science can be found around fifty years back when it was used as a substitute for computer science in 1960 by Peter Naur. During the year 1974, Peter printed a book called "Concise Survey of Computer Methods." In this book, he described the term "Data Science" in its review of the modern information and data processing methods (Data Science and Its Growing Importance).

Definition of Data Science

Currently, Data science is a major field. In this field, the data scientists create such processes in which useful knowledge is extracted from larger databases and collections of data. Data Science can be considered as a continuation of fields such as Data mining and analytical analysis. Data science practices numerous concepts and methods that belong to other scientific areas such as information technology, arithmetic calculations, statistics, computer science, and software engineering. Some of the basic techniques used in Data Science include prospect models, concepts of artificial intelligence, signal processing, information mining, arithmetical learning, databases, cloud computing, design recognition, and computer programming. Data science is not limited to the field of big data - big data is a different area because it provides resolutions that are more dedicated to formatting and pre-processing the statistics and information rather than evaluating the data.

Who is a Data Scientist?

A Data scientist is defined as a computer expert, database and program writer, corrective expert, supervisor, and expert annotator. These people are very significant for the effective organization of the latest data assembly.

Their main job is to plan the inventive review and examination for the data to be used in a correct and operative way by organizations. Without the proficiency and knowledge of specialists who turn the latest technology into actionable understanding, big data is nothing. Currently, numerous organizations are adopting big data to unlock its power. This is aggregating the worth of a data scientist who is an expert with the technical know-how to take actionable insights out of a mountain of data. A data scientist is capable of doing three tasks:

● Data Analysis

● Data Modeling.

● The engineering and prototyping of processes.

These three above mentioned tasks describe the working life cycle of a typical data science project.

A Data Scientist also performs the following tasks:

● Linking fresh and dissimilar information to propose products that meet the ambitions and objectives of their marked customers.

● Using sensor information to identify climate surroundings and redirect supply chains.

● Discovering the deceptions and irregularities in the marketplace. Data scientists are skilled in recognizing data that is prominent in some way.

After this, they create geometric net paths and big data practices to project fraud tendency models and use them to generate alerts that will ensure timely replies when uncommon data is found.

● Improve the swiftness at which information sets can be retrieved and combined.

● Recognize the finest and innovative method to utilize the Internet so that

the marketers can use the prospects in a complete way.

How Data Scientists Increase the Worth of a Business

Data specialists or data scientists can add value to a business as well as they will increase the assets of the business. Following are some ways in which a data scientist can add worth to business:

● The managing team can make improved choices:

o A skilled data scientist is expected to be a trustworthy consultant and planning partner to the establishment's higher management by safeguarding that the workforce makes the most out of their methodical abilities.

● Guiding activities based on the latest trends:

o Data scientists inspect and search the establishment's data. After a proper inspection,

they will recommend and suggest definite activities that will aid in improving the establishment's presentation, improve customer involvement, and eventually take the company towards success.

● Challenging the workforce to accept best strategies and emphasize on problems that are important:

o The duties of a data scientist include safeguarding that the workforce is acquainted and familiar with the establishment's operating environment. Data scientists will organize the people of the company to achieve using live demos of the operative use of the computer systems to mine insights and determine the acts that need to be done. When all the people understand the system capabilities, their emphasis could be changed toward addressing important challenges of the company.

● Making choices with the evidence of information:

o When different data scientists are working in an organization, data collection, and investigation from numerous sources does not involve great risks. Data scientists construct models using current data that creates a range of possible activities. It leads to business learning new paths that will get the finest corporate results.

● Refining of target viewers:

o Almost all companies have one or more sources of client data that is collected.

The significance of data science is based on the capability to take current data that is not essentially valuable and uniting it with other information to produce understandings a business can utilize to study about its clients and viewers. Data scientists can assist with the identification of the vital collection of data regarding a company's audience with precision.

● Hiring the right people for the company:

o By extracting the large amount of data that is available, processing for curriculum vitae and applications can be done easily. The employment team can make quicker and more precise choices.

● Testing every decision that the company makes:

o After applying necessary changes to the data and extracting information, it is necessary to test every decision. After this, the data scientists will measure how the changes that they applied changed the business processes.

The Technique of Data Science

Data science highlights the usage of universal approaches for business problems without altering its application. This technique is different from a traditional approach because it is inclined to concentrate on giving solutions that are precise to specific areas of the business. Data science provides a standard solution to all the business's problems. Data science is valued in businesses and the IT industry.

Impacts of Data Science

Currently, data science has its effect in many fields. These fields include both educational and applied engineering domains such as machine learning, language recognition, and economy. It is also used in fields such as medicine, health informatics, and healthcare. The development and expansion of any brand are affected by it. It offers large amounts of intellect about customers and operations by using techniques such as data extraction and data examination.

Importance of Data Science

The importance of data science can be understood from the following points:

● Data science supports different companies to understand their clients in an improved and empowered way. Customers are a vital part of any company, and they have the power of making or breaking a business. Data science provides a personalized way to companies for dealing with their customers. Therefore it ensures the company's power to grow steadily.

● The companies can utilize their customer's data in a proper manner. Companies can share their own stories with their focused viewers to create an enriched company connect.

● Data science is reachable to nearly all divisions of the IT realm. A huge amount of information is present in the world today,

and operating this data in a correct manner can cause success and failure for businesses and organizations.

● The primary advantage of data science is that it finds and produces results. These results can be used in any field regarding tourism, healthcare, and learning, etc.

● Big Data is an innovative field that is frequently rising and developing. Many techniques and tools are being established on an ordered basis. Big data is serving organizations to resolve difficult problems in IT and resource administration in an operational and planned way.

● Data science helps in delivering significant products to clients. One great benefit of data science is that the companies can discover where their end products are in most demand and where they sell best. It can assist in delivering the correct end products at the time when products are needed the most.

● The marketing sector of companies can understand their viewers in detail. This information will help them to generate the best customer experiences.

Programming Languages Every Data Scientist Should Know

Data science is a diverse field to work on. It uses both quantifiable talents and progressive statistical analysis with actual program design ability. Numerous programming languages are required in data science.

Data science involves the usage of scientific procedures and approaches to examine and draw deductions from this information. Particular programming languages are considered for the role of data scientists. Many programming languages are used in the development of software and application development. Software design for Data Science varies from general software development. It supports to pre-practice, examine, and create forecasts from data. These data-focused programming languages have the capability of processing algorithms that are suitable for the particulars of Data Science. It is a necessity for data scientists who own the mandatory skills for becoming an expert in the field of data science. Before attaining proficiency, an ambitious Data Scientist needs to be able to make the correct choice about the kind of programming language necessary for the work of data science. The following is the list of some programming languages every Data Scientist should know about:

Python

Python is a tremendously widespread, generic, and dynamic and extensively used programming language among data science professionals. It is an easy high programming language. The nature of Python language is versatile. It is easier in both reading and writing. Its code is easily readable. While dealing with complex problems, the readability of this language makes it the choice of many data scientists. Implementing solutions is easier in this language. It quickly interfaces with algorithms designed for data science, therefore, making it the first choice of data scientists. Python programmers are in great demand currently in the IT market.

Python has appeared as one of the greatest standard selections for Data Science due to its smooth knowledge arc and valuable libraries.

R Language

It is the most commonly used language in data science. The source code of R language is available. The environment of the R language is used for arithmetical calculations and visuals. R language is maintained by the R Foundation for Statistical Computing (L, 2019). Arithmetic calculations and visuals are in great demand among employers in the field of artificial intelligence and data science. R language brings many numerical models, and many data analysts have created their applications in the R language. The "Community R set library" covers more than 8,000 systems contributed sets. R is perfect for statistical tasks. Linear algebra's complex tasks are easily handled by this programming language. It also provides smooth database connectivity. It is, therefore, a perfect selection for data scientists.

Java

Java is a very general and popular language. It runs on the Java Virtual Machine. This language is supported by Oracle. It permits manageability between different platforms. Java is commonly known as the pillar or foundation of the programming stack. Java Software Engineers and Java System architects are in high demand for data science organizations. Java is considered a good choice by data scientists

Scala

Scala is an extension of Java programming language. It provides functional programming as well as object-oriented programming. It can be used with Apache Spark, which is a big data platform. This conjunction with a big data platform makes it an ideal choice for data scientists when they want to deal with a huge amount of data. Scala is completely operational with the Java language. Scala offers parallel processing, and it is one of the highlighting features of Scala. This language is recommended for experts, but it is not recommended for beginners.

SQL

SQL can be considered as the core of data science. A data scientist must be an expert in this language. It is used to extract data from relational databases by querying the information. SQL is basically an old language. Being a skilled expert in SQL could be the main advantage for Artificial intelligence professionals and data science specialists. SQL is the best-desired skill for

almost all establishments. SQL demonstrates a vital part in Data Science as it supports the following functions such as updating, enquiring, and altering databases. SQL is essential for specific roles in Data Science. Familiarity with SQL is highly essential because the mining and alternation of the data from the database are possible with the help of SQL. SQL has a declarative syntax, and therefore it is easier to read and write.

Julia

For developing the best results regarding scientific computing, Julia is the best option. Julia is a new language. It has fast performance, and it is simple at the same time. Wherever arithmetic operations are required, Julia's language is preferred. It solves highly complex mathematical and arithmetical problems quickly; therefore, it is favored by data scientists. Nowadays, Julia is known as the most used language in artificial intelligence. Julia's language is also used for risk analytics.

Matlab

Matlab is an industrialized language, and it is approved by MathWorks. It is a fast and stable language. Matlab guarantees concrete algorithms for arithmetical calculation. It is used in the academic world and the IT industry. Matlab is considered the best language for arithmeticians and experts that work with particular mathematical requirements such as signal dispensation, image dispensation, and matrix algebra. It is used in statistical applications. Due to its mathematical processing capability, it is used in data science.

The Data Science Process

The data science process involves solving problems related to data science. There is a framework of data science that is followed during the project life cycle. Certain key skills and requirements are present in the complete project life cycle of data science.

Data scientists analyze enormous sets of seemingly dissimilar data to reveal astonishing insights in the field. The procedure they use is a secret to most individuals outside the circle of data science. The following are the detailed steps required to solve a particular problem in the scenario of data science:

- **Consider the problem:**

The first action a data scientist takes before solving a problem is to describe precisely what the problem is. Data scientists have to be able to decipher data queries into something over which action can be taken. A data scientist will get vague responses from individuals who have certain issues. A data scientist will have to get the instinct to change rare responses into actionable results.

- **Gather the data**

The second stage of data science development is very simple. In this step, the data scientists gather the data that they require for data science from obtainable data sources. During this step, the data scientists must query the databases by consuming their practical skills such as MySQL to gather the data. Data experts will obtain data in file arrangements such as Microsoft Excel etc. If the data experts are utilizing programming languages like R or Python, they will have access to precise packages that can read statistics from these data sources right into the data science algorithms and programs. Data experts gather data from databases as well as by connecting to web APIs. The most basic way of getting information is directly from the files. Some websites allow the use of their Web APIs to gather their data.

For the tasks mentioned above, a data expert will need particular expertise. This expertise will include the management of databases. To get big data groups, a data scientist will use distributed storage such as Spark or Apachae Hadoop.

● **Scrubbing the data**

After gaining the data and relevant information, the next instant thing to do is the cleaning and scrubbing of the data. This procedure is to clean and strain the data. If the information is unfiltered and inappropriate, then the effects of the inquiry and analysis will not make any sense.

During this procedure, the data is converted from one arrangement to another. It is recommended to combine everything into one consistent arrangement. For example, if you have data in different Excel files, then it is suggested to combine all the Excel files into one excel file containing all the data. Scrubbing the data also includes the job of mining and substituting values. If it is found that there are lost data groups or the data appears to be non-values, a data expert substitutes the values accordingly. The data needs to be divided, combined, and extracted properly. This procedure is used for cleaning up the data, eliminating what is not required anymore, substituting what is lost, and regulating the format over which all the data is collected.

For doing the above-mentioned processes, a data expert will need a grasp of a good programming language for scrubbing the data. For the management of larger data groups, it is required to have skills in Hadoop and Spark.

● **Exploring the Data**

The third step is the examination of the extracted data. Typically, in a business or corporate environment, the managers just give a group of data to the experts, and they will study the data. The data scientists will figure out the corporate questions and alter the information into a data science query. To achieve this, an exploration of the data is essential. Data experts need to review the data and their characteristics. Different types of data require different actions. In the next step data, experts calculate expressive statistics to get features and examine important variables. Examining important variables frequently is done with the help of the association.

For the above-mentioned task, a data scientist needs to have information and abilities in inferential data and information visualization.

● **Modeling the data**

This is the 4th stage in the data science process, and it is considered the most

important stage in the entire data science project lifecycle. To get to this stage, the cleaning of data and exploration of data are extremely important as those two phases lead to making useful and meaningful models. Firstly, decrease the dimensionality of your information group. All the values are not vital to forecasting the model. Just consider the applicable values that give an estimate of the results. Many tasks can be performed using modeling. Regular expressions and predictions are used for predicting future values. The evaluation total is carried out after the modeling process.

For the above tasks, both supervised and unsupervised algorithms are required.

- **Understanding of the data**

This is the final step of the entire data science process. This step is the most crucial step in the entire process. The prognostic control of a model is present in its skill to generalize. A data scientist will describe a model based on the model's capacity to simplify hidden forthcoming data. Understandable data means the appearance of data to a non-technical person. This final step gives the results to deliver a solution to the business questions that were asked when the project was first started. In this process, the actionable understandings are offered through the data science process. Actionable understandings are key outcomes that demonstrate how data science can deal with projecting analytics and later prescriptive investigations. Data experts need to envision the conclusions and keep them focused in correspondence with the corporate questions. It is vital to present the results in a proper way. The results should be useful to the business.

For the above-mentioned tasks, data scientists need to have a robust business area understanding to show the results in such a way that it provides answers to the business questions.

The above mentioned 5 steps demonstrate the life cycle of a data science project. Every data scientist follows these 5 steps to answer questions related to the field of data science.

The Future of Data Science as a Carrier Choice

The job of data scientist offers a bright career. It has continuing influence in the IT market, and it delivers prospects for individuals who learn data science to offer valued contributions to their businesses and societies at great scale. Following are some reasons why Data Science is a great career choice for the future:

● Businesses Struggle for the management of their Data

Businesses gather statistics and data from clients in the form of electronic transactions, web connections, etc. Data scientists have the opportunity to support companies in order to make an advancement with the information they collect. It makes them make better business decisions.

● Data Privacy Rules increase the demand for Data Scientists

In May 2018, the General Data Protection Regulation (GDPR) was implemented for nations in the European Union (Methews, 2019). The GDPR improved the dependence businesses have on data scientists. It is due to the necessity for actual real-time analysis and storage of the data. It ensures that businesses comprehend how they stock the information and where they store the data properly.

Currently, individuals are reasonably more cautious about providing data to companies than individuals from previous generations. Individuals are knowledgeable that data hacking can happen, and individuals have to face severe penalties. Businesses cannot treat their data carelessly. The GDPR data privacy guidelines are only the start. Data scientists can assist the businesses greatly in the use of data in a proper way along with the coherence of Data Privacy rules given by the GDPR.

● The field of Data Science is continuously Growing

Jobs without development potential remain stagnant. Data science seems to have plentiful prospects to progress over many years. It is continuously evolving, and it is great news for individuals wanting to come into this innovative field. One minor change likely to emerge soon is that data science job titles will get more specific. An individual employed as a data scientist at one business is not essentially doing similar work as another person in that

same job at a different business. Individuals learning for data science jobs can begin to specify and do the tasks that are most significant to them.

- **A Surprising amount of Data Growth**

Data is created on a daily basis. The exponential progress of data that is observed since the start of the digital age is not predicted to slow down. The future will bring an ever-growing flood of data. The new information and data will be utilized for better data science models, and it will give rise to improved and innovative prototypes for data analysis along with original and groundbreaking use cases of data.

Data creation is on an upsurge, and the data scientists should be present to deal with the data efficiently and effectively.

- **High Probability of Career Progression Prospects**

LinkedIn lately chose data scientists as a capable career in 2019. LinkedIn also observed the probability that individuals could get upgrades as data scientists, and it gave this job a career progression score of nine out of 10 (Methews, 2019). LinkedIn's deductions propose that companies are expected to retain data scientists on their IT teams for a long time.

Chapter Two: Cyber Security

Cybersecurity is the training of shielding computer systems, set-ups, and programs from digital hacking attacks. These cyber-attacks are typically aimed at retrieving, altering, and destroying data along with extracting money from computer users. It also includes disturbing usual business procedures.

Introduction

Cybercrime is a global problem. This problem has been ruling the news bulletins for years. It is a danger to the security of an individual, and it is an even larger danger to huge global businesses, banks, and administrations. Cybersecurity is the body of skills, procedures, and practices intended to defend nets, electronic devices, databases, and information from attack, robbery, harm, alteration, or unauthorized entry. Cybersecurity can be defined as the set of values and practices intended to defend our IT resources and virtual material against all kinds of external threats. Nowadays, everybody uses the internet, and the general computer operators are nearly unaware of the fact that how the random bits of 1 and 0 can cause security risks to the computers. This ignorant nature of computer user causes golden opportunities for hackers. With so many opportunities, the hackers, typically known as black hat hackers, create malicious codes, and they explore the vulnerabilities in the computers. Like all other kinds of innovations around us, Cyber-attacks are also evolving and becoming innovative in the nature of the attack. With time, hackers are becoming cleverer and more inventive with malicious software. They improve the methods of bypassing virus scanning software and firewalls. Therefore some kind of procedure to defend users against all of these cyber-attacks should be present. This procedure will ensure that the user data will be protected. Cybersecurity defines a set of methods that are used to defend the reliability of networks, computer programs, and information from external attacks and harm.

Computer security includes both cyber and physical security. These two types of securities are utilized by companies to defend against illicit access to information centers and other electronic systems. Information security is a subdivision of cybersecurity. The usage of cybersecurity can support in avoiding cyber-attacks, information breaches, and incidents such as identity theft. Cybersecurity can also assist in risk management. Usually, computer users save themselves from three different kinds of unauthorized accesses. These include unauthorized access to the data, unauthorized deletion of the data without any warning, and unauthorized modification. Cybersecurity deals with the availability, integrity, and confidentiality of the data. In short, the term cybersecurity prevents the sensitive data of the end-user from being

falling into the wrong hands.

Cybersecurity is the safety of systems connected to the internet, hardware, computer programs, and information from hacking. The cyber-world is connected to technology. The technology comprises computer systems, net, software programs, and information.

The word "security" is related to safety. It further includes systems safety, network safety, application safety, and data security. These skills, procedures, and practices can also be referred to as information technology security. There is heavy dependence on processors in the IT industry that stock and communicate plenty of private and vital information about individuals. Therefore cybersecurity is a vital function, and it is required in almost all kinds of businesses.

Most Common Cybersecurity Threats

Computer threats are persistently ingenious. These are chiefs of disguise and alteration. These cyber-threats continually progress to discover new behaviors to aggravate, steal, and damage. Everyone should know the relevant material and resources needed to defend against difficult and increasing computer safety threats. In this way, individuals can stay safe while connected to the internet. Following are some common types of network and security threats:

● *Computer Viruses*

Everyone knows about computer viruses. For the most frequent Internet users, computer viruses are the greatest danger to cybersecurity. Computer viruses are sections of software programs that are intended to extend from one PC to another. Computer viruses are often showed as electronic mail attachments or transferred from particular websites with the intention of contaminating the computer. Computer viruses also have the capability of transferring to other computers on the contact list. Viruses are famous for showing themselves as spam.

They can disable the security settings, damage the system, and take information from the computer, including private data such as PINs and passwords. They even have the capability of deleting the hard drive entirely.

● *Rogue Safety Software*

Hackers have found a new way of doing internet fraud. Rogue safety software is a malicious piece of software. It misguides the users into believing that there is a virus installed on their computer or their computer is not up to date. In both cases, these scammers offer them to download a piece of software on their computers. They sometimes ask for a specific amount to be paid too. The software that the users download is infected. In this way, the users are tricked into installing malicious software on their computers.

● *Trojan horses*

A Trojan horse is usually a malevolent piece of offensive program or software that hides behind any valid program. Users install it willingly

without knowing about it. One of the ways in which they spread is by email. As soon as the user clicks the email, the Trojan horse is automatically downloaded in their system. Trojan horses also spread in the form of advertisements that are not genuine. Trojan horses have many capabilities that include stealing all the passwords in a computer, stealing sensitive and private information, and they can also track the webcams.

- ***Spyware***

Spyware is similar in its working to adware. The spyware is installed on an individual's PC without his information. It can comprise of key loggers that keep track of individual data comprising of emails, PINs, debit, or credit card information. The most important risk associated with the use of spyware is Identity theft.

- ***Computer Worms***

Computer worms are bits of malware software that duplicate rapidly and extend rapidly between computers. A worm extends from an infected PC by distributing itself to all the contacts of the computer; from there, this cycle continues until all of the computers are affected. The transmission of computer worms is often done by misusing software weaknesses.

- ***Phishing Attack***

Phishing attacks include directed digital e-mails that are transferred to fool individuals into using a link. That link has the capability of installing malware and uncovering important data. Phishing attacks are getting more refined day by day. These kinds of cyber-attacks assist black hat hackers in taking user information regarding login credentials, credit card authorizations, and many types of private monetary information. These attacks also get information from private databases.

- ***Ransomware***

In Ransomware, the hackers organize tools that permit them to steal a discrete database or establishment's databases accurately. After this, they keep all of the data under their control for ransom demands. The growth of cryptocurrencies, such as Bitcoin, is powering ransomware attacks by letting ransom demands to be funded secretly. Different businesses are continuing to

emphasize the construction of resilient fortifications to protect against ransomware.

• *Crypto-jacking*

This is a relatively new term. It is introduced recently. Crypto-jacking is related to cryptocurrency. Crypto-jacking is a new trend that includes cybercriminals taking over the control of other people's home or work PCs to search for cryptocurrency. For utilizing the computing power for this purpose, the hackers use the computing resources of other individuals.

• *State Supported Attacks*

Entire nations are now consuming their cyber abilities to penetrate other managements. They implement attacks on other organizations. Cybersecurity is nowadays a major danger. It is not only a major danger for the private division and individuals, but it is also a danger for the administrations and countries as well.

• *IOT Attacks*

The Internet of Things is getting more universal by each passing day. It includes all the devices attached to the cloud. IOT includes connected electronic devices that are convenient for customers. Many businesses now utilize them to save currency by collecting huge quantities of perceptive information and reorganization of business procedures. The other side of IOT device usage is that more related devices mean larger danger. Therefore it makes IOT systems more susceptible to cyber-attacks and cyber threats. If hackers can gain access to IOT devices, these devices can be utilized to generate chaos, burden systems, and shut down important tools for monetary gain.

• *Third-party Risks*

Third parties, such as retailers and suppliers, pose an enormous danger to companies. Most of the businesses have no safe organizations or devoted groups in place to maintain this third-party personnel. Cyber offenders are becoming progressively refined, and cybersecurity dangers are increasing. The establishments are getting more and more alert of the risks that third parties pose.

● *Shortage of cybersecurity specialists*

Cyber-crime has intensified quickly in the past few years. Businesses and companies have struggled to appoint sufficient capable specialists to protect them against the rising threats.

The great shortage of cybersecurity specialists remains to be a reason for fear. A robust and smart workforce is vital to battle the more recurrent and refined cybersecurity threats originating from all around the world.

● *Man in the Middle attacks*

Man-in-the-middle attacks are also sometimes known as "snooping attacks." They occur when an attacker is present between two-party dealings. In this attack, the hackers tap into other individual's networks and keep an eye on all the confidential transactions and gain access to private and financial information. The hackers set up fake Wi-Fi networks for individuals, or they sometimes install malware on the target's computer. The basic goal is the same. It is to gain valuable confidential information from individuals.

● *Obsolete hardware and software*

The hardware and software of a company should be latest and up-to-date. The latest hardware and software are necessary for the safety of the establishment's system, servers, electric devices, information, and clients. Obsolete technologies no longer offer protection against the latest cyber-attacks made by hackers and scammers. If the cybercriminals are inventing new strategies of hacking the systems and security, then there should be advanced hardware and software available for protection of the computers from unauthorized access.

● *Insufficient Patch Management*

Companies release software patches to deal with the weaknesses and vulnerabilities in the operating systems, computer programs, and other tools. Patch management is essential for the security of companies. If the software vulnerabilities are not covered properly, then they will leave holes in IT safety infrastructure. Preferably the process of patching must be applied as soon as a weakness is realized in software. The software liabilities leave the establishments at the risk of cyber-attacks, downtime, harm, and non-

regulation in cybersecurity standards. Therefore patch management should be sufficient in order to keep the company away from the cyber-attacks.

• *Form Jacking*

This kind of cybersecurity danger includes a hacker taking control of forms on an eCommerce website. Most of the time, the cyber offenders make the use of a malevolent JavaScript program on the "forms page" usually present at the "Check out" page of the eCommerce websites. The main goal is to steal their client's monetary and expense data such as credit card numbers etc.

• *People as a cybersecurity threat*

Sometimes people are intended to harm a system, and sometimes people do it without any intention. Therefore individuals are a major danger to cybersecurity.

These exposures come from workers, retailers, and anybody who can use the computer system within an organization. An information breach or cyber-attack can happen just because of a human mistake or a lack of cybersecurity responsiveness. Workers within an organization can be a substantial cybersecurity danger when they contemplate that they have something to get through their malevolent activities. Sometimes the people within an organization want to get the revenue by marketing or consuming the information they steal. Sometimes people want to get vengeance against a current or previous company for some alleged unfairness. These employers might install malware over the company's computers, download information, or execute other dangerous activities. In some cases, the personnel of retailers can also be considered as a potential risk.

Impacts of Cyber Attacks on Business

As technology is continuously developing, and the use of technology is growing, risk management has an essential part in safeguarding that such expertise does not expose businesses to cyber-attacks. Cyber-attacks can harmfully impact financial security and are now an intrinsic certainty that any maintainable business or industry must be prepared to face (Izuakor, 2016). A cyber-attack can cause a lot of harm to the business. Cyber-attack can affect business and customer reliance. The details of different kinds of impacts of cyber threats are as follows:

The financial cost of cyber-attack

Financial loss is considered as one of the most fundamental impacts of cyber-attacks. Financial loss can be considered as a backbone behind every cyber-attack. Usually, financial loss occurs from the stealing of business information, financial information, funds, and credit card information. One other cost that the businesses face from cyber-attacks is the cost to recover from a cyber-attack and the cost of repairing the systems.

Reputational Impact of a Cyber-Attack

Reliance is a vital component of customer association. Cyber-attacks can harm a business standing and impact the trust of the clients. This reputational impact can lead to the loss of clients, transactions, and incomes. The reputational impact of a cyber-attack can damage the trust between suppliers, investors, and partners within a business.

Legitimate Impact of a Cyber-Attack

Businesses have to defend the data. The confidentiality laws are required to maintain the safety of all individual data that is present. Within a business, this data can be about workers or clients. In some cases, if this information is unintentionally or purposely compromised, then the business has failed to install suitable safety actions, and it may have to face penalties and supervisory authorizations.

Psychological Impact of Cyber-Attack

The psychological impact affects the individuals responsible for maintaining cybersecurity within an organization. After a cyber-attack, the individuals feel unhappy, uncomfortable, disgraced, and confused as they consider the cyber-attack as their fault. This impact can be reduced by the proper training of individuals with the latest developments in mitigation strategies for cyber-attacks.

Physical Impact of a Cyber-Attack

The physical impact of a cyber-attack includes the damage to the physical infrastructure as a result of a cyber-attack. The common example of this impact is wiping off everything from the hard drive and leaving the hard drive entirely useless.

Social Impact of a Cyber-Attack

The social impact of a cyber-attack includes the disturbance to everyday life activities in an IT environment. It also includes the influence on key amenities, a harmful insight of technology, or descent in interior confidence within organizations that are affected by a cyber incident. The social impact of a cyber-attack affects business activities socially.

How can Cyber-Attacks be Reduced?

Cybersecurity breaks can upset even the strongest businesses.

Therefore, it is tremendously vital to manage cybersecurity risks, consequently. In case if a cyber-attack occurs, an active "Cybersecurity event response strategy" must be present. This strategy will help businesses in:

● Reducing the effects of the cyber-attack.

● Reporting the cyber-attack incident to the reporting authority.

● Cleaning and recovering the effected computers and IT infrastructure after the cyber-attack.

● Recovering from the cyber-attack as soon as possible.

Many businesses become victims of cybersecurity attacks, and many businesses are at risk of cyber-attacks due to low-quality cyber-security plans. Here are some ways in which cybersecurity risk can be reduced within a business:

Creating responsiveness of cyber-security within an organization

Companies and Businesses can invest in great cybersecurity infrastructure, but in the end, it all comes in the hands of the employees who will practice the cyber-security practices. Companies have to rely on the employees for making the use of best practices regarding cyber-security. Every individual of the company should be trained properly regarding cybersecurity. Cyber-security education should be universal for every individual of the company. Some key points should be covered in this training. These key points include:

● Rules around satisfactory use of provided technology in both external and internal use.

● Procedure to safeguard that individual and business information is constantly safe from an external data breach.

● Measures on how cyber-attack recovery will be made on the occasion of a cyber-attack.

● Ensuring strong passwords for all the individuals ranging from people in the administration to the general staff.

● General and specific information on how workers are going to utilize the IT infrastructure, networks, and how their levels of access are defined within the organization.

● How to identify ‘suspected' correspondences that the employees will receive during their normal working within an organization.

Investing in cyber safety and cyber-security backup

As the cyber-attacks are becoming more and more inventive every day, therefore the companies should invest in complex safety measures, vigorous backup options, and recovery arrangements to diminish risks associated with cyber-attacks. It is all about being practical and dropping the penalties of a cyber-attack.

Investing in cyber-security measures is better than working on weak IT infrastructure with less protective measurements and undergoing extreme downtime or even paying a ransom for important information to be refunded.

Keeping up to date with all of the safety arrangements and testing the security measures regularly

Security systems should be updated regularly in order to keep the defenses up to date. In this way, the security defenses can deal with the current innovative hacking techniques. All of the IT infrastructures should be regularly updated. Consistent testing of the IT policy should be done to safeguard that the business is never left susceptible to a cyber-attack. The recovery plans should also be tested regularly to ensure proper safety.

Chapter Three: Cyber Technology

The war between cybersecurity professionals and black hat hackers is always continuous. This leads to an ever-developing chain of cybercrimes that continually overcome current cybersecurity technologies. The solution is present in the collective knowledge and the application of progressive cybersecurity skills. Black hat hackers are continuously finding new ways of stealing valuable information and getting money from hacking techniques. Businesses should work with the best practices of cyber technology in order to stay safe from malicious attacks. Identifying and establishing innovative cybersecurity plans to fight cyber threats is the requirement of the current cybersecurity situation. Here is the list of some emerging technologies that will be able to help businesses defend their IT infrastructures against external threats:

- **Hardware verification**

The failures of user credentials, such as names and passwords, are famous. There is a requirement of a more secure form of verification. A new method that is introduced in the use of authentication methods placed in the hardware. Some latest INTEL processors use this form of verification. These processors can syndicate a variation of hardware-enriched features to authenticate a user's distinctiveness. Good verification processes need three things from the end-users. These three things typically include a strong set of passwords, usernames, and typically a token, usually in the form of a security question.

Hardware verification is especially important in the technical area of the "Internet of Things." In IOT technology, the network is important for the verifications of devices that connect to the cloud. Hardware authentication provides better cyber-security for cloud-based solutions.

- **User-conduct analytics**

This technology uses the concepts of big data. Cyber-experts keep a check on the user behavior. Whenever the behavior of compromised user credentials occurs, data scientists figure out that malevolent behavior is present. This

method is based on differentiating legitimate users and fake users by utilizing their credentials. This method also uses the "peer analysis" techniques. Proper training for cyber-security specialists is essential for using the "User conduct analysis."

- **Preventing data damage**

Crucial technologies that prevent data damage are encryption and the use of tokens within IT processes. These techniques can protect the data on all levels. These techniques also provide a number of benefits, including no monetization of data in case of a successful data breach. Analytics on the data can be done in the data protected.

- **Deep learning skills**

Deep learning skills include a variety of technologies, such as the use of machine learning skills and concepts of artificial intelligence. These techniques primarily focus on the identification of irregular behavior.

These techniques help in looking at business entities at all levels. Machine learning techniques can figure out malicious software behavior. Therefore machine-learning models can help in figuring out the malicious code.

- **Cloud technology**

The cloud technology has a transformative influence on security technology. Due to the introduction of cloud technology, different approaches to cloud security have appeared recently. On-premises methods will be transferred to the cloud. The cloud technology introduced the concepts of virtualized hardware, firewalls, and virtual interference detection mechanisms.

- **Blockchain cyber-security technology**

Blockchain cyber safety is one of the newest cyber safety expertise that is attaining recognition. The blockchain cyber technology works on the foundation of credentials among two business parties. Likewise, blockchain cyber safety works on the foundation of blockchain web essentials. Every follower in a blockchain is accountable for authenticating the legitimacy of the information that is added. Furthermore, blockchains generate a very resistant net for hackers, and it defends the information from being compromised. The usage of blockchain with the concepts of AI can create a vigorous authentication arrangement to keep probable cyber threats away.

- **Zero trust Models**

This Zero trust model of cyber safety is founded on a deliberation that a network is currently compromised. It enhances both types of interior and exterior securities in a network. The bottom line here is that together, both inner and outer networks are vulnerable to a hacking attack and need equivalent defense mechanisms. Zero trust models contain recognizing important business data, planning the movement of this information, rational or physical division, and control through mechanization and continuous checking.

Features that Must be Present in a Cyber Technology Platform

Features are the qualities that must be present in a cybersecurity platform. For an effective cyber technology platform, certain capabilities must be present in it to work efficiently and effectively. Following is the list of some features that a cyber technology must have:

● Email security and web security must be covered

According to statistics, at least 90% of cyber-attacks originate from phishing electronic mails, malevolent connections, or faulty URLs. A cybersecurity platform should apply strainers and check to these common dangerous vectors for obstructing malware and giving reflectiveness into irregular, distrustful, and malevolent actions.

● Essential management over all the services

Essential management means alignment organization and strategy organization, along with common management and reporting. Cyber technology platform organization delivers an accumulated substitute to the present condition where organizations function endpoint safety management, grid safety management, and malware management, etc.

● Detection and prevention of Threats

A cybersecurity technology platform must be equipped with detecting a cyber threat. After successful detection, it must be able to provide such functions that can terminate the threat process. In the last step, the cybersecurity platform must be able to roll back onto the original state of the system before the threat attack.

● Cloud-based backend facilities

The cloud can be considered as the backend intelligence of a cyber technology platform. Cloud-based facilities will collect doubtful behaviors across clients. These cloud-based facilities will track these behaviors through progressive and continually refining machine learning procedures. They will deliver tailored analytics and risk intelligence for specific clients and businesses. Therefore, all the clients will benefit from general and tailored

services offered by the cloud.

- **Open API connection support**

Cybersecurity platforms must be equipped with APIs for third-party skill incorporation and developer assistance. This will also reassure the network result where cyber safety technology platform operators share development's finest practices. In this way, cybersecurity experts can make custom solutions by using third party APIs according to their demands.

- **A technology platform that offers numerous deployment choices**

Big establishments incline to use hybrid cyber technology placements. They usually run the security applications at business headquarters while choosing for cloud-based safety representation services to assist distant workplaces and mobile workforces. Cybersecurity skill platforms will provide hybrid provision over all the safety panels with a central organization plan.

Best Cybersecurity Practices for Businesses

The small-sized businesses are primary targets of cybercriminals. In reality, small-sized businesses get attacked by hackers more often than large-sized businesses. According to the statistics, in the US, 71 percent of cyber-attacks occurred at businesses with fewer than 100 workers (Segal, 2019). Small-sized businesses have more security breaches than large-sized businesses. All kinds of cyber-attacks usually have 3 main types of purposes.

These purposes include stealing personal data, getting money, and stealing user credentials. Small size companies have fewer safe networks. Therefore it makes it easier to break the network security. Automated attacks initialized by hackers can breach a number of networks automatically. The network security within a business, irrespective of its size, is immensely important. Within small-sized businesses, an insufficient amount of time, money, and lack of cybersecurity specialists is a top reason for cyber-attacks. Here is a list of some best cyber security practices for all kinds of businesses. These security practices will save the companies from the hacker attacks and unwanted access to their monetary data and sensitive information. The detail of these practices is as follows:

- **Utilizing a firewall for business**

The primary defense mechanism for a cyber-attack is the firewall. This is the first layer of security that is attacked by hackers. Setting up a proper firewall for your network means that there is a defense mechanism staying between sensitive information and cyber hackers. Currently, along with the typical exterior firewall, many businesses are beginning to set up interior firewalls to deliver extra defense. Workers of a company who work remotely must also install firewalls on their computer systems so that the data of the company will always be secure.

- **Documenting the cyber safety guidelines and policies**

Regarding the field of cybersecurity, every company and business should properly document the policies and procedures.

Documentation of all the security protocols is essential. Different types of software are available in the market for documenting the protocols. This software typically includes cybersecurity portals and voluntary cybersecurity

programs for businesses.

- **Proper planning for mobile devices that are used within a company**

With the growing fame of wearable gadgets, such as smart tablets, televisions, watches, and health trackers with wireless competence, it is vital to add all of these devices in a cybersecurity procedure. The automatic security updates must be enabled along with strong password security measures that are required on every remote device that accesses the company's network.

- **Proper cybersecurity training for all the employees**

Every employee who works within a business must be capable of using cybersecurity protocols defined by the company. Proper training is essential for individuals to save themselves from cyber threats. Proper updates should be available on the newer versions of the security protocols within a company.

- **Secure software practices**

Many data breaches occur due to stolen and weak passwords set by company individuals. Any device that accesses the company's network must be password protected. Passwords must be enforced using strong policies, including capital letters, small letters, and special characters.

- **Recurrently back up for the company information**

Backing up all the data deposited on the cloud is essential. It is necessary to ensure that backups are collected in a distinct site in case of fire or natural disaster. Recurrently creating backups is essential to safeguard that there will always be the latest backup always available. It is also important to have a functioning backup all the time.

- **Using anti-malware software**

Regardless of all the security measures that a company takes, a security mistake made by an employee can cripple the entire company's system. Anti-malware software must be installed on the computers so that whenever a phishing link is clicked by a worker by mistake, a defending mechanism is present. Phishing specific tactics are covered by anti-malware software. Therefore for preventing phishing attacks, anti-malware software must be installed on the company's systems.

- **Using Multifactor Identification methods**

Utilizing the multi-factor identification option on the main network and electronic mail products is easy. It also provides an extra layer of protection. Cybersafety is an important target. Cyber offenders get more progressive each day. For the protection of the company's data as much as imaginable, it is very vital that each and every worker take cybersecurity the highest precedence.

- **Protection of the information**

Cyber offenders can generate correspondence addresses and web applications that look genuine. Hackers can forge caller ID materials. Scammers can even take over business's social media accounts and send apparently genuine mails. Therefore, it is significant not to lose your business's information, delicate evidence, or knowledgeable property. Take this example; if an

employee shares an image on the web that demonstrates a whiteboard or processer display in the background, this picture can unintentionally disclose statistics that somebody exterior to the business should not get. A business can defend its personnel, clients, and information by generating and allocating corporate strategies that will use subjects such as "how to abolish information that is not desirable" and "how to recognize and report doubtful correspondences or malware.

● Only a secure WIFI connection should be accepted

Workplace Wi-Fi nets must be safe, encoded, and concealed. If a worker is working remotely, the worker can help to protect the data by using a simulated remote network, if the business is using a VPN connection. A VPN is vital while doing work in the external environment of the office or on a corporate journey. Community Wi-Fi nets can be dangerous and make your information at risk of being interrupted. A good and safe version of a VPN connection must be ensured at the workplace.

● Investing in Cybersecurity systems

Minor industries might hesitate while anticipating the price of capitalizing in a superior safety organization. That typically comprises of defenses such as robust malware and antivirus discovery, exterior hard drives that are responsible for back up information, and successively running steady arrangement checks. Making this kind of investment initially could protect businesses and personnel from the probable monetary and lawful expenses of being breached.

It is significant for a business to deliver data safety in the office. In the case of a security breach, the IT section or Information Safety manager should be informed that a security problem might be present. There might be a fault in the IT organization that the business needs to cover or fix. The earlier a company can report a problem, the better it will be for the safety of the company.

● Enforcing Biometric Security

Biometrics safeguards fast verification, safe contact organization, and exact worker checking.

Confirming users' individualities before giving access to valued possessions is essential for companies. Voice acknowledgment, impression examinations, hand biometrics, facial credit, social biometrics, and posture analysis are flawless choices to classify whether or not the users are genuine and legitimate. Utilizing biometrics delivers more safe verification than passwords and SMS confirmation.

Due to all of these reasons, biometric authentication is an important method of user verification. Verification is not the one used for biometrics. Safety managers get benefits from a wide variety of biometric-based gears that permit them to notice hacked and compromised private accounts in the actual amount of time.

There is another field within Biometrics, and it is called "Behavioral Biometrics." Behavioral biometrics examines the methods by which workers interrelate with the input machines. If irregular conduct is noticed, a device directs a caution warning to safety managers so that they can respond instantly. Different types of Behavioral Biometrics can be placed within a business's IT infrastructure. This Behavioral Biometrics will include the followings:

o Keyboard dynamics:

The keyboard dynamics reflect the typing quickness and the inclination to make distinctive errors indefinite words for generating the end user's performance outlines.

o Mouse Dynamics:

The mouse dynamics measure the amount of time among ticks, beat, and grace of pointer drive.

o Eye Biometrics:

The eye biometrics practices the eye and stare tracking strategies to measure videos of eye association and notice eye patterns.

According to the statistics, the development of the biometrics marketplace increased from $16.8 billion in 2018 to an estimated $41.8 billion by 2023 (12 Best Cybersecurity Practices in 2019, 2019). Therefore, investing in the best cyber biometric cybersecurity practices is the best option for cyber

companies.

- **Using a risk-based approach for security purposes**

Risk assessment is the best tool for companies to assess their risks. Every business has its own particular and concealed risks, so concentrating on compliance and conforming to all the typical guidelines are enough to defend the delicate information. A risk assessment should cover the important assets of the company, the current status of the cybersecurity, and the management of cybersecurity within a company. Appropriate risk calculation permits the company to evade lots of disagreeable things such as penalties for failing to conform to guidelines, remediation charges for possible leakages, security breaches, and the damages from lost or incompetent processes. Adjustments in a company should be made after identifying the weak spots within the cybersecurity practices. A keen eye should be kept on the hacking attempts at all times. Worksheets for risk assessment are also very important for the company's risk assessment.

- **Management of the IOT Devices**

The most puzzling mechanism about IOT devices is their entree to subtle and sensitive information. All of the IOT devices are potential access points for hackers and scammers.

For example, a conceded printer can permit malicious hackers and scammers to get the view of all papers that are passed through the printer or scanned by the printer. Following are some of the business security's best practices:

o Doing the penetration testing for the analysis of real-time risks and planning the safety strategies according to the requirements.

o Encryption strategies must be enforced for all kinds of data.

o Authentication for only giving the connection to the verified endpoints.

o Proper credentials should be followed for every end-user.

o The latest router should be installed along with a proper firewall.

o A scalable safety IT infrastructure should be available for all the departments with the company.

o Proper endpoint security mechanisms should be installed over the entire

business's IT infrastructure.

- **Giving users fewer privileges**

There should not be too many privileged users in any business. Allowing fresh employees all privileges automatically permits them to contact the sensitive information even if they don't essentially require it. It raises the risk of inner cyber threats and permits scammers to get entree to delicate information whenever any of the new worker's account is hacked.

The rule of least privilege is to be always followed. The company should assign every new account the least privileges possible and escalate privileges if necessary.

The Future of Cyber Security

Whenever a user connects to the web from a processer or phone, there is a rising danger of cyber-attack and cyber-threats. If the danger is aimed at a user's workplace, then the whole business around could become susceptible too. A good company, irrespective of its scope or worldwide reach, should ultimately recognize that cybersecurity needs substantial investment.

Many experts consider that an innovative cyber technology founded on the concepts of machine knowledge and AI (Artificial Intelligence) is the future of cyber-technology when it comes to processer, network, and information security. Nowadays, businesses place importance on the safety of their interior systems. If scammers and hackers accomplish to penetrate the network layer of their organization, within a less amount of time, a "minor" break can become a significant hacking attack.

A common method for network defense is a strong firewall. Firewalls can be present either as a software technique or a hardware method that is actually linked with the network. In both of these cases, the firewall's work is to monitor which web connections are permitted on which ports.

The firewall also has the function of blocking all other irrelevant requests. A server administrator is typically responsible for managing these policies regarding firewalls. In a scenario where a hacker has bypassed the firewall and net safety, a business's succeeding line of protection is the antivirus software that is intended to scan hardware and software for malevolent code. The aim is to eliminate the malware before it can distribute itself to other machinery and issue a kind of attack such as ransomware.

In the case of cybersecurity, there is no technique present that is more valued than the training of the workforce regarding cybersecurity. The effective establishments run training sessions on a consistent basis for fresh and existing employees to teach them about threats that exist online and methods in which they can defend themselves and the business.

How AI (Artificial Intelligence) will shape the future of Cyber Security Methodologies

The popular cybersecurity techniques and tools need human communication or configuration at all the levels of cybersecurity. For example, an individual belonging to an IT Team has to fix the firewall strategies and backup timetable. It is the duty of this individual to ensure that everything is running smoothly according to set policies. The progression in the field of Artificial Intelligence will change this entire equation.

In the near future, businesses will be capable of trusting on clever techniques to deal with the bulk of occasion monitoring and event reactions.

The succeeding technology of firewalls will have built-in machine learning tools. These tools will permit the software to identify designs in net requests and robotically block those designs that could be a potential risk or danger to the business.

Specialists also believe that the ordinary language abilities of "Artificial Intelligence" can play a large part in the future of cyber safety techniques. The philosophy is that by skimming big portions of information across the network, AI arrangements can study how cyber-attacks initiate and the AI arrangements can propose answers for cybersecurity specialists within the business.

Security services and products constructed on the AI structure are quite expensive. This high price of such sophisticated systems can pose a dilemma for small-sized and medium-sized companies currently. Appointing and associating a team of machine-learning specialists to shape custom cybersecurity resolutions might not be an instant or even near future selection. Currently, it is highly favorable to invest in hybrid techniques and tools that are present in the marketplace, and implant AI expertise in human functioned products and services.

The mainstream internet users generate their own modified passwords for every website or facility that they contribute to over the internet. This structure can be tiring to preserve, and it is susceptible to cyber-attacks if the users trust easy passwords or utilize the old passwords for numerous sites.

There have been developments in the performance of PIN manager software in current years. Most of this software aims to streamline and fortify online security. This software eliminates a big share of manual exertion from the job through procedures that propose and stock passwords difficult enough to decrease the user's chances of being hacked by hackers. The concepts of artificial intelligence could introduce a new internet world without any passwords. New developments in the world of identity management propose that one day, PINs and passwords might essentially be substituted by intelligent AI-based systems. In this new technology, Artificial Intelligence concepts would track every operator within a business based on roles, rights, and mutual activities. Any abnormality from the standard practices would be identified, and it will need the individual to use an additional kind of verification, such as biometrics, that creates the scans of facial structures and fingerprints.

Financing in cybersecurity resolutions and techniques is an essential job for companies of all kinds, regardless of the size of the company. Businesses with lesser finances may think they can save cash by taking shortcuts, but in reality, they are frequently the major objective for hackers as their cybersecurity defenses are weak. Cybersecurity services prove their value by decreasing the organization's danger and defending it from risky hackers. Now due to progressions in AI technologies, businesses will have no requirement of maintaining large cyber safety teams inside their IT section as in the future, the majority of the tasks regarding cyber-security will be handled by AI processes.

Gears founded on machine learning are very effective at picking up on designs and discovering malicious events before a human can typically recognize. For the current situation, the establishments should train the workers with the advanced tools and techniques in implementing a cybersecurity plan and keep a keen eye on the latest AI developments.

Based on the above-mentioned evidence, it can be presumed that the future of cybersecurity lies in the hands of the latest Artificial Intelligence concepts. By using advanced AI systems, the entire processing abilities of cybersecurity technologies will be modified. It will also provide numerous benefits to the companies and businesses of all sizes.

Chapter Four: Analytics and Metrics for Big Data

Analytics include the analysis and decision making the capability of a system. The metrics include the measurement of performance and the measurement of progress within a company. The metrics provide vital performance pointers. Metrics give the answer to the "what" kind of questions, and the analytics give the answer to "So what" kind of questions. This section will provide a detailed insight into the analytics and metrics of big data, Data Science, and Cybersecurity.

Analytics and Metrics of Big Data and Data Science

Big data analytics is a very composite procedure of inspecting big and diverse data groups and information to expose material such as concealed designs, unidentified associations, marketplace tendencies, and customer favorites that can assist establishments in making knowledgeable corporate choices.

In an overall sense, data analytics skills and techniques deliver a means to examine data groups and draw deductions about the given information, which will help establishments in making knowledgeable business choices. Data analysis provides solutions for elementary questions regarding business processes and presentation. Big data analytics is a procedure of progressive analytics, which includes composite requests with features such as analytical models, arithmetical procedures, and what-if investigation motorized by high-presentation analytics organizations.

Big data analytics are determined by particular analytic arrangements and software along with highly efficient computing schemes; big data analytics provides numerous business benefits. Some of these benefits are as follows:

- Innovative revenue occasions.
- Highly effective advertising.
- Improved client provision.
- Better working competence.
- Competitive benefits over competitors.

Big data analytics requests authentication of big data forecasters, data experts, analytical modelers, mathematicians, and other analytics specialists to examine rising capacities of organized business information. These experts can also analyze other kinds of information that are frequently left unused by predictable Business intelligence operations and analytical processes.

Data analytics work in this way that Hadoop bunches and NoSQL organizations are utilized chiefly as staging parts for big data before it is sent into a data warehouse or logical database for analytical processing.

The data typically processed is in a précised format. This format is more beneficial to relational constructions. When the data is complete, it can be examined with the software normally used for progressive analytic procedures.

Text excavating and arithmetical analysis software also play a great role in the big data analytics process. This software has the capability of mainstreaming business intelligence data and providing data conception gears.

Data scientists use many performance metrics, such as Correctness, memory, accuracy, understanding, etc. Data scientists use performance metrics using pictures so that anyone can understand the predictive models. The most important performance metric used for evaluating the performance is accuracy. This metric describes the model's accuracy to predict a situation. Another performance metric is known as precision. This metric shows the precision of the results in a specific situation. The sensitivity performance metric describes the sensitive content within a model. The specificity performance metric shows the specific criteria in the given information. These 4 performance metrics are used in the field of big data to describe the predictive models for predictive analysis accurately.

Cybersecurity Analytics and Metrics

Establishments currently face a broader range and a larger occurrence of cyber threats. These cyber-threats include all kids of APTs (advanced persistent threats), cyber conflict, unrestrained occurrences via botnet programs, malevolent scripts, malware attacks,

which are as-a-service through the Dark Net, or even interior threat occurrences from persons within the business and organization. The whole thing that includes dispersed denial of service attacks (DDoS), man-in-the-middle cyber-attacks, phishing attacks, crypto-jacking, ransomware, and data breaches hit industries of all dimensions and in all businesses continually and every day.

Security Analytics is a method that is used in cybersecurity. It is fixated on the examination of information to create practical security actions. As an example, observed system traffic could be utilized to recognize pointers of hacking attempts before a real danger happens. Any business, no matter what the type, cannot forecast the future of the company. In case of safety threats, arranging and installing security analytics techniques that are able to examine safety actions can notice a cyber threat before it has an opportunity to affect the IT organization.

The arena of cyber safety analytics is rising. This field is full of prospects, and it offers a vigorous resolution for establishments that are looking to stay away from susceptibilities. The cybersecurity field makes these organizations stay one step ahead of the hackers and scammers. A number of factors are involved in the increasing demand for security analytics. Some of these factors are as follows:

- **Moving between protection and detection**

Scammers use a varied range of cyber-attack tools that achieve numerous susceptibilities.

Some cyberthreats can go unnoticed for many months. Safety analytics gears can keep track of mutual danger designs and send warnings the instant an irregularity is exposed.

- **A cohesive view of the business**

Security analytics arranges data in a way that it proposes both actual time and past view of actions. This analytic technique delivers an integrated vision of threats and safety breaks from a crucial point of view and permits for keener development, quicker determination, and better choice-making ability.

● **Seeing effects and a reoccurrence on investment**

There is rising stress on IT teams to convey the results to high-ranking management and investors. Security analytics delivers results in less time for resolving the metrics and less false positives that permit analysts to classify threats and reply to safety breaks rapidly.

One of the main advantages of cybersecurity analytics is the sheer capacity and variety of data that can be examined in less time according to the requirements. This information typically includes the following types of data:

● End user's behavioral information.

● Information regarding Net traffic.

● Information about business applications and enterprise applications.

● Cloud traffic data.

● Non-IT infrastructure's data.

● Peripheral threat intellect bases.

● Access and individual organization information.

● Proof of agreement during a review.

By examining such a wide variety of information, establishments are able to attach the dots among numerous alerts and actions simply. The outcomes are active security event discoveries and quicker reply times that assist the corporate to defend the reliability of organizations and information. Security analytics techniques also contribute the amenability with business and administration documentations. Rules such as PCI-DSS and HIPAA need organizations to observe information actions and record information collection for forensics and reviewing dedications (What is Security Analytics?).

Security analytics can be applied for a varied diversity of use cases, from operator performance checking to network transportation analysis. A few

common examples of use cases are as follows:

● Examining network traffic to notice designs that designate a possible attack.

● Checking employer behavior, particularly possibly doubtful behavior.

● Noticing insider pressures.

● Noticing information exfiltration.

● Classifying accounts that may have been negotiated.

In a domain where downtime due to any hacking attempt can cause disasters, establishments are required to consider more vigorous, dependable, and responsive methods of danger detection.

Operative management of variable presentation catalogs in IT safety can mean the variance between a real-world and well-organized project and a comprehensive money wastage. In cybersecurity, it is an emerging training to track cyber safety metrics. You cannot manage something if you cannot measure it properly. Cybersecurity metrics are a significant method of keeping track of security efforts. Good cybersecurity metrics are required for efficient and effective communication with business stakeholders. Cybersecurity metrics provide key performance indicators, which are vital for measuring the performance of the cybersecurity practices within an organization. Metrics also give an insight over how the services of cybersecurity are improved over a period of time. The metrics should be very clear and relevant to cybersecurity information. This will help even the non-technical individuals within the company to understand the cybersecurity situation effectively. Some key performance indicator examples are as follows:

● Level of preparation

This includes the number of devices that are attached to the network.

How many devices are fully patched and currently up to date with respect to performance?

● Anonymous devices on the interior network

The workforces carry their devices at work. The company might also be

utilizing the Internet of Things (IoT) services. These are enormous dangers for businesses as these devices are possibly not secure and safe from hacking attempts.

- **Interruption attempts from hackers**

It will include the list of a number of attempts in which hackers have tried to break the company's networks?

- **Number of days required to patch**

Will it track the record of a number of days required by an IT team to apply the security patches? Hackers and scammers often achieve delays between patch issues and the application of patches.

- **How many cyber incidents are reported properly**

Are the workers reporting cybersecurity problems to the IT team within the business? It is a worthy sign. This reporting of cyber issues signifies that the workforces and other shareholders identify cyber issues. This also implies that the cyber training provided to the employees in a company is working.

The cybersecurity metrics that a company will select will be determined by the organization's requirements and its level of risks. Metrics should be clear to anyone who will read the report. The cybersecurity metrics should define the organization's security to the business frontrunners. By monitoring the cybersecurity of the entire business, best cybersecurity decisions can be made regarding the company in the future. The following are some cybersecurity metrics that can be utilized for tracking to safeguard the competence of the safety developments.

- **For any hacking attempt, the meantime to attack and respond**

The poor performance of a company regarding the detection and response of a hacking attempt impacts the company greatly. This impact can be in the form of huge costs required to get back on point from the attack. For any IT security infrastructure, these two factors are very vital. It is necessary for a business to train their company's individuals accordingly for the detection of these two factors.

- **A variety of systems with recognized susceptibilities**

Knowing the variety of susceptible properties in the IT setting is an important cybersecurity metric to defining the risk that your corporate experiences. Handling updates and software patches is a multipart procedure, but it is very significant to evade ambiguities that can be misused in the IT settings.

A susceptibility scan that comprises of all the assets will designate what requests are to be completed in order to progress the safety stance of the business. A susceptibility management package is a necessity in the current situation.

- **SSL credentials configured wrongly**

An SSL certificate is a small-sized folder that confirms the possession of a cryptographic key to the web application or business. The data is being switched with these web applications and businesses, and it assures the genuineness of the business. Checking the safety provisions for each certificate while safeguarding that they are correctly configured on IT systems, stops them from getting into the hands of the criminals. In this way, the business's digital individuality cannot be utilized to take end user's data.

- **Amount of data relocated using the business network**

If your company personnel have unobstructed access to the web through the business network, then checking the capacity of network transportation permits businesses to recognize misappropriation of company possessions. When downloading software, cassettes, pictures, and requests, a user has a chance of encountering the botnets and malware to attack their IT settings. This chance increases greatly if the transfers are from websites that are famous and recognized as dangerous for malware attacks.

- **Users with a high level of access**

Finest practices in data security organization comprise of complete control of an employee's level of access to business resources. It is essential for a worker only to get the information, IT arrangements, and properties that are obligatory to their own work. Recognizing the access levels of all net operators permits the company to regulate them as desirable by delaying any other user or manager that does not require the company's network.

- **Days required to neutralize previous employee's credentials**

By checking these cybersecurity metrics, a company can describe whether the Human Possessions and IT players are working correctly or not. In a perfect situation, the access of operators that are terminated from the business must be disregarded directly. If the credentials of such employees are active, then it is an incredible danger. It can lead to the loss of important company information, and the devices of the company can be hijacked and compromised from obsolete accounts.

- **Checking the communication ports that are opened for connection purposes**

Inbound traffic should not be allowed. Outbound connections should be checked regularly to ensure that the traffic is moving swiftly or not. All the protocols responsible for remote access must be checked regularly to ensure the total protection of the company's assets.

- **Checking the third party access levels regularly**

Sometimes third parties are given access by the company's personnel to complete a project in a given amount of time. It is vital to ensure that after granting access and completing the task, the service of remote access is terminated or not. If the remote access is not terminated on time, then it can pose a great risk as the third party can come back without any alert or warning, and they can steal important data as well. Therefore all the third party accesses must be blocked after the use of service.

- **The proportion of corporate partners with efficient cybersecurity strategies**

There should be a strong control and monitoring of the cybersecurity metrics of the corporations that offer facilities for businesses. Providing access to the company's IT infrastructures to the outsourced corporations can be an enormous danger if it does not have operational policies for its own security first.

Conclusion

The volume of big data at present is enormous. This volume is predicted to rise exponentially as innovative technologies such as the more universal IOT gadgets, drones, and devices, which are a wearable increase in use. According to research, 90 % of the big data in the current IT world has been produced in the last two years (Buttice, 2019). The current progressions in deep learning are playing an important part in assisting businesses to make use of this valuable data. Big data and corporate analytics solutions are considered as a major technology innovation currently. The digital IT processes are further built on the advanced concepts of Artificial intelligence and automation processes.

Big data specialists make the connection between raw information and practical data. Data specialists must have the ability to operate data on the deepest stages. Their knowledge should enable them to understand the data's tendencies and designs in numerous different procedures. The computer programming languages and methods used to attain these objectives are increasing in forte and statistics. Inside a business, data scientists help to resolve big data issues, but typically these issues may be vague. To more confuse the subject, some data specialists work externally to any particular group. One common example of this scenario is the academic research-based field. Big Data Analytics is a security improving device of the future. The quantity of data that can be collected, prepared, and used for the employers in a modified fashion would take a lot of processing time.

Establishments are sinking with the huge volumes of data, and it appears that the businesses need to transfer to completely algorithmic information-driven forecasts to live in an extremely competitive IT domain. This situation makes the challenge of clever usage of these enormous quantities of fresh information. High level and better performance-based predictive analytic models are required for modeling the big data projects. After identification of an issue worth resolving with predictive analytics. It is essential to consult with advisers who make predictive analysis models. A predictive model has four basic stages of solving a problem. These 4 stages include management, planning, delivery, and operation on the problem, along with all of the risk

factors that are involved in the whole process. Data science is a versatile arena that utilizes technical approaches, procedures, processes, and arrangements to mine information and understandings from organized and unorganized information. The main benefit of data science over outdated measurements is that it could create deductions from a garbage pile of the allegedly unconnected data. In the field of data science, the data experts will play with procedures and algorithms for utilizing their understandings and creatively to improve the arithmetical model. A thorough understanding of the business procedures is mandatory for the data scientists to perform data science methodologies.

The profession of data science will increase competence, efficiency, and output within businesses and companies. Data science contributes to the discovery of the data insights and the development and progress of data products. Both of these factors will greatly increase business value. Increasing business value is the optimum goal of the data science field. Establishments are discovering themselves under the burden to respond rapidly to the vigorously growing number of cybersecurity dangers. Cyber threat management strategies are mandatory nowadays for businesses to defend themselves against external threats. The susceptibility organization life cycle is intended to counter the struggles made by the invaders in the fastest and most operative way. Cyber-security complications arise from the intrinsic nature of the IT infrastructure. Complex IT systems and the nature of human judgment all contribute to determining the level of cyber-attack. Threats to cyber-security progress and evolve continuously. The hackers and scammers continuously invest in new techniques to hack the organizations and businesses for material gains.

There are typically three things for which hacking attempts are made. These three things include User credentials, the company's valuable information, and monetary gains. The monetary gains are usually the prime factors that influence hacking attempts. Cybersecurity defenses for a company are an ongoing process. The security procedures are continually getting advance.

As the hackers continually adopt new strategies, therefore the defenders of the companies known as white hat hackers must also frequently use new techniques for effective defense of the company's information. Best cybersecurity practices must be followed for a safe system and up to date

system free from all kinds of cyberattacks.

IT Analytics include the examination and choice-making competence of an IT system. The metrics comprise the extent of performance and the dimension of development and progress within a business. Regarding the big data, data analytics services and methods provide a way to inspect information collections and draw inferences about the specified data, which will support establishments in making well-informed business decisions. Data scientists use many different kinds of performance metrics for evaluating the performance of the company. Data scientists make the use of performance metrics by consuming graphical techniques such as pictures. Graphical representation makes it easy to comprehend the predictive models. The most vital performance metric utilized for assessing the performance of a business in big data is accurate.

Cybersecurity analytics is based on the inspection of data to create real-world security actions. The outcome of cybersecurity analytics is vigorous safety occasion detection and earlier response times. Both of these factors help the business to protect and guard the consistency of organizations. Security analytic techniques also assist in the responsiveness of business and management credentials. The cybersecurity metrics will be determined by the security risks and cybersecurity situation of the company. The cybersecurity metrics for any organization must be clear and consistent for anyone to understand. Some important metrics for cybersecurity include the time of the attack for a hacking attempt and the meantime to respond, third party access levels, the time required to recover from a cyber-attack, third party access levels, etc. Every business should always be prepared for external cyber-attacks, and appropriate mitigation strategies should be present to neutralize the threats.

Bibliography

12 Best Cybersecurity Practices in 2019. (2019, May 30). Retrieved from ekransystem.com: https://www.ekransystem.com/en/blog/best-cyber-security-practices

Big data Basics. (n.d.). Retrieved from sisense.com: https://www.sisense.com/glossary/big-data-basics/

Buttice, C. (2019). Big Data for Big (and Small) Business.

Data Science and Its Growing Importance. (n.d.). Retrieved from https://www.educba.com: https://www.educba.com/data-science-and-its-growing-importance/

Izuakor, C. (2016). Understanding the Impact of Cyber Security Risks on Safety. *2nd International Conference on Information Systems Security and Privacy*.

L, A. B. (2019). Top 8 programming languages every data scientist should master in 2019.

Methews, K. (2019). Why Data Science is The Career of The Future.

Segal, C. (2019). 8 Cyber Security Best Practices For Your Small To Medium-Size Business.

Thompson, R. (n.d.). Understanding Data Science and Why It's So Important. *Behind the data*.

Vaidya, N. (2019, May 22). Data Science vs Big Data vs Data Analytics. p. 1.

What is Security Analytics? (n.d.). Retrieved from Forcepoint.com: https://www.forcepoint.com/cyber-edu/security-