DETECTION OF IDENTITY THEFT OVER SOCIAL MEDIAL NETWORKS USING FUZZY LOGICS

by

Muhammad Saim Jehan Khan

Supervisor: Dr. Mian Muhammad Waseem Iqbal

A thesis submitted to the faculty of Information Security Department,

Military College of Signals, National University of Sciences and Technology,

Islamabad, Pakistan, in partial fulfillment of the requirements for the degree of MS in Information Security

AUGUST, 2022

ABSTRACT

Social media networks (SMNs) have become a norm in everyone’s life and significance for the interrelationship amongst people all over the world. This is the reason theft of individuals’ identity has become very common. In order to detect identity theft, my research proposes fuzzy logic-based system and decides based on results if a person is being an identity theft victim. To do so, the natural language and numerical values of each user is considered, with the help of values such as the total frequency of posts, messages, new recipients, and geo locations. Fuzzy logic idea has been proposed before to detect identity theft in SMNs but with no actual implementations and idea was taken forward with limited number of parameters. In our paper the accuracy increased by taking more parameters into consideration.

ACKNOWLEDGMENTS

I am grateful to God Almighty who has bestowed me with the strength and the passion to accomplish this thesis and I am thankful to Him for His mercy and benevolence. Without his consent I could not have indulged myself in this task.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Table Number   Page

INTRODUCTION

1.1 Introduction

A social media platform is an online platform that users use to create social networks or relationships with others that have similar personally or professionally interests, hobbies, backgrounds, or real-life connections. Social media has a significant impact on young people. It's becoming evident that social media has become an integral component of people's life. Many teenagers check Tweets and face book updates by their relatives and friends on their notebooks, personal computers, and mobile phones. People are being driven to accept alternative lifestyles as a result of technological advancements. Social networking platforms can help young individuals improve their social skills. Social media is a type of web-based data transmission. Users can engage in conversation, exchange intelligence, and produce web content on social sites.

Other forms of social media include blogs, micro blogs, wikis, social media platforms, photo-sharing sites, face book messenger, video-sharing sites, podcasts, gadgets, virtual worlds, and others. Huge numbers of individuals throughout the world use social media to communicate and build connections. On a personal basis, social networking enables us to keep in touch with old friends, discover new things, explore our interests, and be entertained.

Social media is an original concept with tremendous potential for growth. Many firms are utilizing social media to improve their procedures as a result of its advancement. We can promote or interact more effectively with the help of social networking. Similarly, individuals do not need to rely on the media or television for their regular quota of news; everything can be found on a social media platform. Social media provides a forum for people all over the world to communicate their concerns and thoughts. The engagement amongst individuals or groups wherein the participants generate, publish, and occasionally trade ideas, photographs, videos, and other media through the internet and in digital environments is referred to as social media Children are growing up in a world surrounded by portable devices and participatory social networking sites like Face book, Twitter and MySpace, as well as Google plus, which has made social media a significant part of their lives. The way young people interact with their parents and friends, but also how they utilize technology, is changing as a result of social media.

Individual engagement on a broad scale become relatively easy to make than ever before as a result of the emergence of internet and smart phones, ushering in a new media age in which responsiveness was positioned at the heart of modern media operations. One person could now communicate with a large number of people, and fast reaction was a potential. Citizens and customers used to have restricted and rather subdued voices, but now they can share their thoughts with a large number of people. Individuals now have the chance to explore information from a variety of references and to converse with each other about the information posted via message forums, thanks to the low cost and accessibility of new technology. Instead of only a few news outlets, individuals now have the ability to seek information from a variety of sources and to converse with each other via message forums well about information published.

For starters, they enable employees to work on their identities. When someone creates a profile for themselves, they must reflect on who they are. Individuals will examine themselves in a new light when they see reactions to their online social existence, and they will note that online engagement allows users to comfortable about their views, beliefs, and queries, for good or worse. Secondly, social media encourages customers to take alternative approaches to their interactions. Even while public discourse frequently portrays social media platforms like Face book and Twitter as selfish and superficial, research demonstrates that they enable people who might not be able to communicate to do so. Through computer-mediated communication systems, people have met some of their closest friends and even spouses.

Third, social media enables people to carry out work-related tasks. When a prominent blog or someone with a huge social network circle is paid to advertise an event, the social media is sometimes their work. Others communicate with coworkers using social media sites or, more commonly, email, and manage their professional communication through the social media platform. Fourth, users can use social media to find facts or share ideas. This information can cover a wide range of topics, including electoral candidates, local problems, emergency relief, and where to buy plus-size clothing. Furthermore, sometimes in line with knowledge transfer, people can use social media to express their thoughts or examine the perspectives of many others. Finally, such sites can provide enjoyment to users.

Connectivity is the first and most important advantage of social media. People from all over the world can communicate with each other. Regardless of where you are or what religion you practice. The allure of media platforms is that it allows you to connect with anyone in order to learn and express your thoughts. The fundamental advantage of web-based networking sites is that it allows you to keep up with the most recent events on the planet. The majority of the time, print and television media are one-sided and do not convey the true message. You can receive the facts and real data while doing some research with the use of internet social media sites.

Students and teachers benefit from social networking in a variety of ways. It is incredibly simple to learn from other professionals and experts via social media. Anyone can learn from as well as enhance their performance in any field by following them. We can educate ourselves without paying for it, regardless of our location or educational background.

You can share your problems with the group for support and energy. Irrespective of whether you need financial assistance or counsel, you can get anything from the group with which you are affiliated.

We can reach the widest possible audience with our message. You have the entire globe at your disposal, and you may promote yourself to them. It will assist in increasing profits and achieving business objectives.

Social media can be utilized for noble causes as well. The public is using social media to contribute to those in need, and it can be a simple way to aid them.

People from other communities can interact to talk and share similar topics because our globe is diverse in terms of religions and philosophies. Traditional marketing methods such as radio, television commercials, and print advertisements are now entirely outmoded and cost thousands of dollars. Businesses may engage with their targeted clients for free through social media; the only costs are energy and effort. Bloggers, content producers, and creators are increasingly turning to social media sites like Twitter, Face book, and LinkedIn as a viable communication option. These long distance interpersonal communication venues have enabled all bloggers to connect for their well clientele and share their expertise.

Learners and professionals are capable of sharing and exchanging information with like-minded others, as well as soliciting comments and opinions on a certain topic. People who have never met outside of social media forums can benefit from social media. Social media facilitates the exchange of ideas across geographical borders. It gives all authors and publishers the ability to engage with their customers. It brings individuals together on a massive platform to pursue specific aims. This has a good impact on society.

It is easier to comprehend the needs of customers while using social media. Social media aids in the promotion of a company all over the world. Through standard connection and favorable client benefit, Social Media helps to develop deals and maintain client relationships.

Customers can have a lot of fun with social media. You may learn a lot about your competition by following social media. Sharing contents about the business is faster and easier with the help of social media. By providing a variety of services, social networking sites aid in the acquisition of new clients. With online networking, you may have a better understanding of the market and go out beyond your competitors. It also aids in raising client awareness for a better knowledge of items.

- Prescriptions from doctors are shared with friends, families, and coworkers. Health researchers will have more data to work with.
- Doctors can be consulted online at any time and from anytime.
- Suggestions regarding various diseases and their symptoms are shared among friends, families, and coworkers.
- In undeveloped countries, there is a lack of information.
- On online health forums, there is a lot of support and mutual responsibility.
- Assistance for causes relating to health.
- Assisting health-care providers in prioritizing essential situations.
- Consumer responsibility is now more important than ever.

Wide uses of Social Media:

Communication: In the communication business, the tools are essentially a well-known sort of social networking site. Sites and blogs are examples of these tools, which allow you to produce articles on the internet in order to interact, engage, inform, and motivate your audiences. Readers of your articles will be allowed to move comments. Other applications include social networking sites like Twitter, Face book, and Instagram, which leverage personal information, comments, images, video uploads, and more to boost the odds of connection between people.

Synergy: It is, after all, a knowledge service that allows users to update their perspectives. It's essentially an online encyclopedia that anyone can access and update. Another example appears in the form of Google Docs, which allows users to modify and share documents online. You'll be option to access and share documents using this interactive tool, which is referred to as social media because it allows several people to share a single platform.

Reviews and Opinions: If you're still not convinced, go to Amazon, and read the reviews and ratings to assist you choose the things you desire. If it isn't helpful, we're not sure what is. Businesses can request that consumers share their comments or testimonials on social media platforms, which will instantly persuade more people to buy from them. On your Twitter or Face book posts, you can start a discussion or ask your users to offer their reviews or opinions. You can promote any product or solicit feedback from your event attendees on your social media page so that you can provide them with better services in the future.

Brand Observation: The instruments for brand maintenance are not well-known, yet they are quite crucial. All consumer brands and retailers that interact with the public utilize brand management tools to keep track of what is being said about their products and services. With the support of social media, this form of online presence is now possible. When it concerns to summarizing all of the feedback and remarks on a specific firm, these tools are invaluable.

Media Exchange: YouTube is, after all, one of the most famous and well-known sites for sharing material. With over 500 million subscribers, this website has already made a name for itself as a result of the incredible services it provides to its users. Vimeo is another platform that facilitates the sharing of media. These are the sites that aid in the creation of channels and interactivity among users. There are also various websites that can assist in the sharing of music.

Paid Promotion: Running paid ads on various social media portals is also a very efficient use of social media. You may run paid ads on social media channels like Face book, LinkedIn, Snap chat, Instagram, Pinterest, and others. Social media sites already have a large user base that you may base on demographic approach to improve your brands, product's, and services' online exposure. You may also track the outcomes of your paid advertising strategies and make changes as needed to improve the campaign's success.

Cybercriminals/ Hackers: Cybercrime is defined as illegal behaviors that use an electronic tool or communications network as a tool, a targeted, or a combination of both. Cybercrime is also known as computer crime, electronically crime, e-crime, higher crime rates, information society crime, informatics crime, device crime, or digital crime.

While the definition of "hacker" has changed over time, the operations of this class are still typically seen as dark and wicked, conducted out in secret locations, and with the goal of harming society's communication systems. Hackers are the primary perpetrators of cybercrime. Their motivations range from simple personal amusement, such as script kids defacing websites and cracking passwords, to the joy of being acknowledged as an elite hacker by breaching information security and thieving from Fortune 500 companies. There are numerous hackers' categories, each with its own terminology and iconography, causing confusion about the terminologies used by computer attackers. Hackers are a term used by the media and the general public to describe those who are accountable for accessing and damaging computer systems. However, employing the term hacker to describe a remote attacker or a technology vandal disrespects both the phrase and the notion. Cybercrime is defined as crimes performed utilizing a computer as a tool or a targeted victim over the internet. Many crimes evolve on a daily basis, making it impossible to categories them into separate groupings. Even in the actual world, crimes such as murder and theft do not have to be separated. All cybercrimes, however, make both the computer and the person behind it victims; the difference is which of the two the primary target is. A small handful of criminals are responsible for these crimes. Unlike crimes that use a computer as a tool, these crimes necessitate the offenders' technical knowledge. These offenses are fairly recent, having only existed for as much as computers have been used to conduct crimes of this sort on a regular basis on the internet.

Most hacker online actions are totally legal; the distinction between hackers, hackers who conduct crimes, and cybercriminals is based on how a hacker interprets the activity and the motivations. Based on past studies, the SANS Institute has identified several classifications and subcategories of hackers:

White hats: These people work as security specialists or as hackers who follow the hacker ethic (do no harm). Lopht, one of the most well-known old school hacking organizations, invented the phrase "grey hats." These hackers are now working as security consultants as rehabilitated Black Hats.

Hackers who are driven by power, rage, or hatred are known as black hats. They have no qualms about stealing or destroying network data once they have gained access.

1.2 Problem Statement

The notion of "identity theft" can be portrayed in a variety of ways, but they all boil down to one basic definition, which is the unapproved use of another person's private details for personal gain. Identity theft has been in the news recently, but it has long been a problem that existed before the Internet. Traditionally, cybercriminals had to physically go through garbage bins looking for private details, such as discarded bills and records that recognized an individual, a practice known as "dumpster diving." It is simply defined as a collection of Internet-based applications that allow people to spread and share data. Its capacity to bring societies from all over the world together on a single platform drew an increasing number of individuals to it.

Traditionally, social media websites such as LinkedIn, Twitter, and Face book were assumed to be used just for socializing by the younger generation, and that they could not aid businesses in any way. However, as the adage says, "not everything that glitters are gold," and these WebPages have their own set of drawbacks. The study's principal objectives are twofold. First, to comprehend the many sorts of identity theft that occur on social media networks; and second, to comprehend how identity theft impacts people on social media and to what degree they are aware of the significance, as well as prevention methods to avoid it. This research also intends to raise awareness among Social Media users about the influence of personal details and its general information.

People put a lot of faith in Face book because it is the most popular social network. This is frequently because they believe their Face book 'Friends' are individuals they can trust. According to reports, one-third of social media users supply at least three pieces of information that could lead to identity theft. Names, dates of birth, pet names, contact information, and the mother's middle name are examples of data.

YouTube is another most popular social media platform, with over 30 million daily visitors and over a billion eligible participants. Identity theft is not as common on Face book as it is on similar social media platforms, but it is still a problem. For example, a YouTube channel with over 12 million subscribers called "The Diamond Mine cart" has over 1000 accounts with the same name. The goal of these bogus identities is to basically the original material in order to increase the number of views.

With 700 million monthly users, Instagram is the third most widely used social networking platform. However, not all of these are legitimate users; according to an Italian security assessment, more over 8% of the identities are false accounts.

Twitter is most important platform among its peers, which means it has been subjected to a number of security breaches. In 2013, Twitter stated that the credentials of approximately 250,000 users may have been hacked as a result of a phishing attempt on their network aimed at stealing personal data. This includes all Twitter users' passwords, email addresses, and identities.

Frauds and scams on Social media: The goal of a scam is to deceive people into splitting with money or giving highly confidential information such as email addresses, accounts, and personally identifiable information in order to assist identity theft (also known as phishing), which could then be utilized to make money. Approaches, on the other hand, can differ.

It should go without mentioning that you'll be cautious about who you connect with on social networks. People approaching you as a friend and begging money appears to be something that. The fraudster could even impersonate one of your colleagues or offer you a fraudulent link that leads to a harmful website.

Be wary of abbreviated URLs that conceal the entire address of a website. They're fairly frequent on Twitter, and although they may well take you to the correct website, there's always the risk that they'll redirect you to one that installs malware.

Hackers can open credit cards or apply for loans in your name if they obtain your Social Security number, name, birthdates, and address. "Hackers collecting personal information, such as Social Security numbers, might enable someone to impersonate their victim and acquire credit or loans that they never repay," says Steven J.J. Weisman, a lawyer and author of "Identity Theft."

People that acquire your information don't just use it to purchase high-priced items. They can even receive medical treatment using your National Insurance number and health coverage account details.

"In the overwhelming majority of instances, (cybercriminals) use your actual address, contact information, as well as other private information," says Justin Lavelle, chief liaison officer of background investigation firm BeenVerified.com. "Their objective is to earn treatment and/or meds, then simply vanish, leaving less time for them to be discovered."

Between October 2009 and December 2016, there were about 1,800 instances of medical data breaches involving patients' information, according to Michigan State University research.

Hackers can gain access to your airline miles by using your email and passwords to book trips or perhaps even redeem for cash. "Converting airline miles to cash is as simple as visiting to sites that buy miles," Lavelle explains.

According to the Federal Trade Commission (FTC), phone and utility accounts accounted for 13% of all fraud instances in 2016. In these circumstances, hackers may have used a stolen Social Security card to set up an account with just an electric, gas, or cellular company.

There's also another con to be aware of. Identity thieves may phone you and claim to be from the utilities company, attempting to shut off your power.

Already techniques used to combat identity theft:

Without a question, social media has fostered an environment and culture of unnecessarily revealing and broadcasting personal information that should be kept private most of the time.

As a result, identity theft thieves have been able to access these data and create financial losses. Although sites such as Face book, Twitter, and others have taken significant steps to combat online fraud and includes concerns' privacy, it remains a problem for these businesses to allow users to freely communicate and connect without exposing them to fraud. Because the trend of exploitation of sensitive data is on the rise, it is critical for users to take care to prevent being detected and becoming victims of online identity theft.

With increased use of internet, particularly social media, user's exposure to cybercriminals has been increased exponentially. Cybercriminals are individuals or teams of people who use technology to commit malicious activities on digital systems or networks with the intention of stealing sensitive company information or personal data and generating profit.

The core factor of what makes a crime a cybercrime is that it's directed at a computer or other devices and/or these technologies are used to commit the crime. Cybercrime is prevalent because the Internet has become a major part of people's lives. In 2014, the FBI's Internet Crime Complaint Center (IC3) reported they received 269,422 complaints from people with an adjusted dollar loss of $800,492,0731 (Internet Crime Complaint Center, 2014). These numbers of course only reflect reported crimes, and not the numerous others who fall victim but never report it because they're too embarrassed or for other reasons.

Identity theft – defined as the intentional, unauthorized use of a person's identifying information for unlawful purposes – is a growing public health problem. While identity theft is not a new crime, the magnitude of the problem has increased with society's growing reliance on the electronic transfer and storage of personal information across all forms of commerce and services. Identity theft occurs when someone uses another person's personally identifiable information, such as their name, identification number or credit card number, without their permission, to commit fraud or other crimes. Identity theft can include theft of a person's name, date of birth, bank or credit card account number, electronic signatures, fingerprints, or any other type of password that identifies the person. Among the most common purposes of identity theft are those related to money, such as credit card or bank account fraud, or those related to social networks, a means of reaching many people, which can spread viruses, malware or catch new victims by taking advantage of the identity of stolen.

This paper has focused on the detection of identity theft in social networks since its use has increased in recent years in society as the data shows that, as of January 2022 there are 3.96 billion total social media users across all platforms. The average person bounces between seven different social networks per month. The average daily time spent by an individual on all such platforms has reached around 95 minutes.

To solve this problem, our proposed system, composed of fuzzy logic and behavioural analysis, will oversee making the final decision on whether a user's account is being compromised, allowing to improve the quality of service in social networks. To do this, the system considers the average values of linguistic variables usage, non-linguistic variables like display picture, type of device used to log-in and average geolocation from which users are performing such activities.

1.3 Contribution

The following contributions are made through this research:

- This work addresses to increase trust issues in social media networks by doing human like decision making to detect Identity theft through a novel approach.
- This paper utilizes fuzzy logic to decide whether a case of one's identity or account compromise in ore humanistic way.
- The proposed system is implemented, and a comprehensive performance analysis is performed. It has proven to be practical and effective for identifying the identity theft cases.

1.4 Thesis Outline

Chapter 2 discusses the existing work done, by briefly discussing their pros and cons. In Chapter 3 preliminaries related to fuzzy logic and our proposed methodology are presented to enhance the understanding of experimental framework. In Chapter 4 Experimental Setup is discussed which is used to achieve the results. Chapter 5 discusses the complete analysis and decision-making steps including comparative analysis against the state-of-the-art. In Chapter 6 conclusion and future work is drawn towards the end.

LITERATURE REVIEW

2.1 Existing Work

It's no coincidence that social media platforms have grown in popularity not just among youngsters, but also among adults and experts in our period. This is due to a variety of factors, depending on each person's priorities. Its appeal among youths and non-tech people may be attributed to the reason that they are knowledge that enables and do not requires special skills to operate. In fact, their mobile counterparts are much easier, allowing a larger number of users to connect. The International Digital Media and Association hosted the first serious talks and assessments of social networking communities in 2006. According to the researcher, a quantifiable study of social networking societies of college graduate educators and academic staff, as well as professionals in their field, has been investigated in this research. The Face book is most of the time, and their share identity information has been disclosed in the social sides. Relationship status, cell phone number, full name, geographical information, and political beliefs are just a few instances of identification information shared on social media.

Professionals may use social networking sites like LinkedIn to share their successes, best attributes, achievements, and talents, which can help them, gain attention from future employers, colleges, and like-minded people. As the number of people using social media grows, so has the likelihood of fraudulent activity [1]. The researchers highlight the substantial interaction between young adults and social media, as well as the obstacles and options for preserving personal information obtained by criminals on social media. The researchers provide a detailed study of information concerning the government's challenges in this area, using the identity theft legislation enacted by Congress [2].

Natural convection in enclosures, according to the experts, is one of the most active topics in heat transfer study nowadays of identity implementing innovative on social media and internet sites [3]. The researchers discovered that attackers can hijack routes and delete data for un-trusted apps; as a result of their examination, the researchers recommended that some solutions be developed to permanently protect data from hackers.

Authors give an in-depth research of identity theft on social media, demonstrating that many victims who have been directly harmed by cyber identity theft are ignorant of the criminal who has obtained their information from social network IDs [4].

However, they were aware that they had been a victim of social networking site scam. According to past surveys, the majority of identity theft victims had no idea how their personal information was stolen; similarly, in 2001, the American Federal Trade Commission stated that only 1% of cyber-crime cases could be linked to the internet.

Theft of an individual's legal identity is a severe problem. He claims that much of this information is currently available on social media platforms, which customers post on web sites to suit their everyday banking needs, as well as information on social media channels [4]. He believes that while taking someone's data from web pages and then use it for personal gain is considered a crime in advanced countries, much work remains to be done in developing countries to tackle the issues of fraudulent activity for their online audiences with the goal of ensuring their online safety.

According to the author, the most studies have verified that alleged crimes in cyber-attacks have an impact on finance, institutional, and economic sectors in Pakistan, resulting in the development of an illegal network for the sale, purchase, and provision of illegal services to criminals to support their criminal activities [5].

According to previous studies, "the hatred and substance" of crime makes it more difficult to eradicate crime in our culture. According to several researches, successful public-private partnerships and worldwide transnational corporations are the transfer nations of cybercrime. There is a substantial body of published research on cyber-crime that describes the use of psychological tricks and deceptive behavior by cyber-criminal activities.

According to the researchers' findings, cyber theft happens as a result of the grey areas that exist in the online system, which allows hackers to exploit it for their own gain. They say that the second reason is that generic consumer information is available to online communities or users in the form of data posted on web forums. Customers submit their content on the web, such as their user name, basic information, and other data according to the website's nature, and they use it for their own purposes [6].

They may also promote their businesses and startups, as well as look for employment related to their sectors. As more individuals become aware of this technology and choose to utilize these platforms to communicate and interact with others, numerous firms see this as an opportunity and are concurrently developing more of these platforms [7]. Insurance cybercrime, criminal information theft, driver's license identity theft, and the most recent identity theft of social media profiles are all examples of identity fraud [6].

Apart from increasing prominent for socializing, these services have also grown famous for numerous crimes that occur within these virtual worlds and have terrible consequences for an individual's life. The dramatic surge in identity theft in recent decades has been ascribed by researchers to the rise in social media usage. Although social media has made communication easier and more accessible, it also appears to have rendered users' personal information vulnerable to identity theft [8].

Cybercrime, such as identity theft, cost UAE customers over than five billion dirhams' in late 2015 [9]. In 2014, identity theft affected 17.6 million people in the United States, costing the country $15 billion [9].

Identity thieves appear to be mostly targeting young people aged 15 to 33, sometimes known as millennial, throughout the world [9]. In 2015, 770,000 Australians were victims of identity theft, which cost them $15 billion yearly. Identity theft results in more than simply financial damage. The effects include an average loss of 30 hours of emotional, psychological, and social time coping with the implications [9].

The most popular social networking websites in the world, as per June 2017 statistic with 1.94 billion active users, Face book once again takes the lead. This may have followed by YouTube with 1 billion users. This data demonstrates Face book's popularity, as there is a significant difference between Face book and its competitors [10].

According to academics, there has been an increase in the amount of literature on the issue of fraud and victims throughout the globe in recent years. Even though the researchers have discussed the four major categories of fraud and victims in their dynamic research with a focus on the UK based research along with the United States of America and Australia, the writers investigate that frauds from other countries cannot be directly applied to the United Kingdom without changing the situation [11]. They analyzed the existing literature on theft and victims in order to provide a quick summary of the fraud steers and explain the strategies utilized.

The amount to which people share information on social networking sites is governed by their choices, which are impacted by a variety of factors. On most social networking sites, the control method is often the user privacy options, which allow a person to manage the appearance of their identity to others. Most users keep these at the default settings specified by the social networking provider, which may be less than ideal in terms of privacy for the final user of these services. A social networking profile is how online social users represent themselves online, and it supports their visibility and distributes information about them [12].

Crimes relating to the holding of identity information and offences relating to the interaction with it distinguish these criminal kinds in Australia. While both are offenses, those that cause financial damage to the victim are more serious. Even if there is no financial loss, it is still conceivable to be harmed by this act. Users who want to communicate their feelings, sentiments, and experiences online find the ability to exchange information appealing. The reciprocal nature of such information sharing is one of the attractions of social networking sites, although many sites strive to strike a balance between user security and convenience of use [13].

According to previous study, sensitive data includes personal photos, names, and gender, all of which are vulnerable to leaking on social media websites. According to previous studies, Users of face book in especially are much more willing to share personal information (including their real names) and include email addresses in their accounts. While the papers required to implement identity differ, most governments recognize a variety of identification documents, and under global standards, names, nationality, date of birth, and citizenship are completely unique attributes that are taken into account together to meet identity criteria [14].

The leaked information might be utilized by the identity thief to fulfill their goal of establishing identity. They could use it to commit identity cybercrimes after it is established. In the United States, for example, credentials have been snatched and then used to commit a variety of crimes, with the victims being falsely implicated [14].

The European Cyber Security Convention aims to standardize international legislation of cybercrime by providing domestic criminal law agencies with cooperation procedures for investigating and prosecuting computer offences. The word "cybercrime" is a term used by the European Convention to define crimes that use a computer or a computer network. Whereas a computer is used to conduct the crime, it is distinguished from traditional crimes. As a result, scams involving the use of a computer as a tool are included. Similarly, when identity theft occurs through the use of a computer, it is potentially covered by the European Convention. The European Convention, on the other hand, does not address identity theft explicitly [15].

According to Fogel and Nehmad's research, certain social media users are more willing to participate in risky activities than others. Furthermore, Face book users show a higher level of confidence in the service than MySpace users. In this context, it has been discovered that males are more willing than women to accept friendship requests on social networking sites. Men are also more willing to reveal personal information such as phone address and telephone number than women. There appears to be a divide here between individuals who share information on the Internet and those who are victims. Anyone who uses social networking sites, however, is at danger, and it appears so the more data transmitted, the higher the risk [16].

According to some academics, advancements in the technology to support the architecture employed by social networking sites are required to improve privacy. Because technology is what stimulates sharing of information in the first instance, it may also be used to reduce future crime. To decrease identity theft, business interests might have to set their interests aside for the greater good. While there are still numerous methods to abuse people, social networking has given rise to new ways to do it, and the platform itself can help to reduce it. The availability of social networking facilitates identity theft, and the low cost underlying identity theft contributes to the crime's persistence [16].

According to Wang et al., IT is normally accomplished in one of two ways. The first involves low-tech methods such as "grabbing wallets, dumpster diving, paying staff for client data, or physically taking files or computer hard drives." Skimming (using a computer to gather information from either the bar code of an ATM or prepaid card), impersonation attack (sending message to people from such a site posing as a trusted source), and phishing (sending email messages posing as a trusted source and requesting secret identification information) are examples of the second [17].

According to RSA Security Incorporations "2017 Global Fraud and Cybercrime Forecast," social media identity theft began in 2011, when e-commerce accounts and credit cards began to be published on social media. Because they were usually simple, free, and had a global reach, these sites became a breeding ground for scammers. With the exception of crimes such as bullying, stalking, and harassment that occur on social media sites, Identity Theft is another one, but it has a greater impact on the victim than the others. Social networking sites such as Face book, Twitter, and LinkedIn have infiltrated the lives of even those with only rudimentary knowledge of how to use the Internet. Criminals, particularly identity thieves, have taken use of these venues [18].

Using Babaei et al's strategy, removing endpoints from the network after seeing the first dubious activity minimizes the number of available nodes and, in many cases, the frequency of standard network operations. While these issues are expected to arise over time, if they are resolved, the node may become a companion node in the hereafter. The number of steps, remaining energy, and cooperation history are all considered inputs to the fuzzy logic procedure, with the level of cooperation between nodes being the output of the present scheme. To transfer data, each node is chosen with a sufficient trust factor [19].

One of the key challenges with social media, according to K. Krombholz et al., is information accuracy. Many phony Face book profiles exist because many individuals establish profiles with incorrect information. To avoid phony identities, social media guidelines said that users should supply accurate information. According to the author, Face book's first focus should be user safety in order to prevent users from diverting [20].

F. Stutzman and J. Kramer-Duffield offer suggestions for improving user privacy in Social networks. They recommend making users' accounts hidden for colleagues only to minimize identity theft, which will lessen the danger of information theft on social media [21].

A. Verma et al. discovered that the infrastructure of centralized social networks, such as the ones in use today, does not protect users' privacy and security. As a result, they presented a decentralized and distributed design for online networks, particularly social networks, that protects users' privacy and security. The "Freedom box" serves as a private server in this architecture, which is built on decentralized media. Diaspora is used as a social network. As a result, each user gets their own Freedom box in which they may keep their personal information. They used a cryptographic technology called (Random Sequential Algorithm) RSA and electronic signatures to improve privacy and security [22].

L. Bilge et al. investigated how simple it is for a hacker to launch systematic crawlers and identity fraud attacks on famous websites in order to gain access to user personal data. They demonstrated two types of assaults against victims with online profiles, one for active members, and the other for non-registered users. Automated identity theft is the first type of assault. The designers of this assault established cloned victims' accounts and then sent friend requests to those victims' connections. An autonomous cross-site profile cloning assault is conducted as the second attack. In this technique, the attacker can establish a fake profile where the target is not yet enrolled and contact the target's friends who are joined on both networks. The findings of the experiments reveal that automation attacks are both successful and practical in practice [23].

Identity theft is a sort of crime, according to C. Marcum et al., and the rapid advancement of technology has created new techniques for stealing the personally identifiable information of hundreds of victims simultaneously. Indeed, the problem has been compounded by the rising number of users on social media sites, as well as the relatively inadequate security and authentication mechanisms. The study also found that users may be unaware of the dangers of providing details or the possibility of using it to forecast highly sensitive data such as social security numbers [24].

R. Demyati also discussed Face book privacy concerns. She talked about how Face book handled concerns about privacy and how it impacted its users. The primary worries were whether Face book distributes user information with marketers and who may see users' images when they are tagged by their friends. In terms of Face book's response, they consented to amend certain aspects of their privacy statement while refusing to change others. She continued with some suggestions, such as advising readers to view the privacy policies, keep important information secret, refuse friend requests from strangers, create a new Face book personal email, and report any difficulties to the social media team [25].

If a criminal is seeking for consumer information, he or she would most likely go to the resource that can be obtained for the least amount of money. Private details are at the disposal of the least cautious possessor, comparable to the "weakest link" concept articulated by Hirshleifer, from the consumer's perspective (1983). The privacy of a consumer's information will be determined by whoever has the least incentives to keep the information safe. As a result, attempts to increase data protection in certain areas may be ineffective if security in these other areas remains weak [26].

People put a lot of faith in Face book because it is the most popular social network. This is frequently because they believe their Face book 'Friends' are individuals they can trust. According to reports, one-third of social media users supply at least three essential elements that might lead to identity theft. Names, dates of birth, pet identifiers, contact information, and the mother's middle name are examples of data. The many methods a thief uses to steal someone's identity on Face book. The most prevalent way was to send a private message instructing the recipient to visit a Scam Website. In 2012, Face book acknowledged to a data breach in which 6 million users' phone numbers and email addresses were accessible to unauthorized users.

2.2 Identity theft cases review

On Twitter, an assailant impersonated Saudi Arabian football club Al-president, Ahli's Prince Fahad bin Khalid. He spread false information about the club quitting sports networks, causing concern among Saudi athletes until it was revealed that the individual behind the profile was impersonating the prince [27].

Another phony Twitter account arose, this time imitating Prince Abdulaziz bin Fahad Al Saud, who already has a verified page with the same photo and identity as Prince Abdulaziz bin Fahad Al Saud, but with one more letter. The counterfeiter's profile imitated the prince's style and persona, luring many people into contacting and trusting the assailant.

A hacking gang known as "Cyber of Emotion" hacked Prince Sultan bin Salman's official Twitter page, which is the chairman of the Saudi Commission for Tourism and Antiquities. The cyber collective sent out many tweets that were disparaging of the tourism industry. In a statement, the commission confirmed that the account had been hacked and stated that they embrace constructive feedback and recommendations from the public. The commission affirmed that they will respond to all of the hacking group's queries with transparency and clarity. With the aid of Twitter's technical support, they were able to retrieve the profile [27].

Unsatisfied customers attacked PayPal's Twitter page, using the site to complaint about their service. PayPal issued a swift statement to reassure consumers and clarify that the assault was limited to the Twitter profile [27].

The profile of Burger King was hacked, exposing that the firm had been purchased by McDonalds. Nevertheless, the mishap turned into a funny and successful commercial for them, and they received over 60,000 new supporters as a result.

Anonymous apparently hacked the Fox News Twitter account. "Fox News was picked because we anticipated their encryption would be as much of a parody as their reporting," a member of the Script kiddies told Think Magazine [28].

The Guardian's Twitter account was hijacked by the Syrian Electronic Army [28].

It's important noting that "hackers" have gained access to some of the official government identities. Abuse and conveying messages in insensitive and contentious ways are the targets of these attacks.

We can see from all of the instances above that data theft is a significant issue with serious repercussions. To solve this problem, more efforts must be made to track down violators and impose harsh financial penalties on them.

Celebrity news and formal government institutions now use social media as a source of information. As a result, their social media profiles must be vetted in order to avoid the spread of misinformation and misunderstanding. The goal of this study was to investigate through fuzzy logics link among identity crime and social networking, which has received little attention in the literature, and to provide the groundwork for future research. To further understand the dimensions of this interaction, more empirical study is required. It is intended that by expanding knowledge of this association, greater interest in this study would be developed.

2.3 State of the Art Work And Features Comparison

Due to the importance of protecting user accounts, numerous researchers have been carried out, resulting in several recommendations for increasing user account security. For instance, organizations such as OWASP [29] give several tips for making user accounts secure. These include length, complexity, and password topology. There is also an article which sheds light on the role of fake identities in advanced persistent threats and covers the mentioned approaches of detecting fake social media accounts [31]. On the other hand, a paper [32] also describes a series of rules for users and developers that allow fighting deception in social networks.

Paper	[30]	[33]	[34] [31]	[35]	[36] [29]
Foundation of the Paper	Detection based on abrupt changes in limited User behavior features	Based on Mobile Social Network only, both online and offline	Detecting if your identity is prone to theft	Identity fraud in a Korean Online Game	To Identity theft preventive models and frameworks

		activity.			
Use of multiple Platform	NO	Yes	Yes	Only Mobile Game	Ecommerce only
Use of Fuzzy Logic	Yes	No	No	Yes	No
Known /Unknown Cases	Both cases	Only Known Cases	Not Defined	Both cases	Both cases
Posts	Based on Content of Posts	No	Not Defined	Not Defined	Not Defined
Messages	Based on Content of Messages	No	No	No	No
Anomaly in Recipients	No	No	No	No	No
Spam Link Detection	Yes	No	No	No	No
Location	Average Latitude and Longitude Points	Check-ins based	No	No	No

Login Device Type Changes	No	No	No	Yes	No
Change of Display Picture	Yes	No	No	No	No
Change of Display Name	Yes	No	No	No	No
Mobile Application Based	No	Yes	No	No	No

Table-1

Because social media is used by millions around the world as a messaging platform, a lot of data is connected with user postings, such as locations, interests, and social ties. These websites are the quickest and easiest way to obtain personal information from visitors. Users of Twitter or Face book, in particular, should be wary of what personally identifiable information they post on their accounts and how others could exploit it. As a result, when users join up for social media sites, they should consider how these sites safeguard them and whether they can trust them with their personal information.

Fake accounts often include false information, photographs, and other data, which tarnishes the victim's public image. As a result, in order to limit the risk of fraudulent activity, adequate identification mechanisms must be employed to verify the user's identity. Biometric technologies, such as fingerprint and iris recognition, are examples of approaches that can be applied. These systems communicate biometric data from input terminals to a computer system, which uses it to verify the users' identities. We can ensure a person's authenticity and keep the public safe before they occur by employing this method.

Identity theft on social media as a result of the lack of use of identity verification techniques such as fingerprint and eye print. Furthermore, it will not be a wise decision and would endanger users. As a result, the need for additional verification mechanisms arose in order to ensure that users' identities are genuine and not forged.

There is a distinction between affirmation and verification on social media. To secure the account, the site will send a message to the user's email or phone. Some sites utilized contact information or emails for identity verification, while others combined the two to offer greater security for consumers. Personality verification remains a challenge, as does the danger of identity fraud.

With today's widespread usage of social media networks, these crimes have transferred to new platforms that provide thieves easy access to massive amounts of data. Social engineering is a relatively widespread crime that takes place on social media sites like Face book. When nations see that specific social media sites are becoming big contributors to crime, they frequently impose bans as a form of retaliation. Several social media programs such as Face book, WhatsApp, Twitter, YouTube, Instagram, and Skype were banned in multiple countries in 2016, as well as the number of countries where users were arrested for their engagement in social media crimes.

PROPOSED METHODOLOGY

3.1 Type And Nature Of Research

Both a virtual and a physical body are affected by cybercrime, although the impacts on each are different. Identity theft is the clearest example of this problem. Individuals in the United States, for example, do not have an official identity card but instead have a Social Security number, which has long been used as a de facto identifying number. Each citizen's Social Security number is used to collect taxes, and many private organizations use it to maintain track of their employees, pupils, and patients. With access to a person's Social Security number, it is possible to collect all of the documentation pertaining to that person's citizenship, i.e., to steal his identity.

When fraudsters steal a company's credit card According to statistics, over 600,000 Face book profiles are hacked every day, with one in every six users reporting that their identity has been stolen. On social network, four out of ten individuals have been victims of cybercrime, and one out of ten users has been a victim of a bogus link. Three out of four people believe that cyber thieves are concentrating their efforts on social network platforms. When it comes to potentially risky activity on social networking sites, data show that one out of every three users does not check out after each session. One out of every five users does not double-check received links before sharing them. One out of every six users has no idea if their privacy settings are public or private information, they have two potential outcomes. First, they steal digital information on individuals, which may be beneficial in a variety of ways. They may, for example, use credit card information to rack up massive bills, causing credit card companies to incur significant losses, or they could give the data to others who could use it in a similar way. Second, they may utilize the names and numbers of individual credit cards to construct new identities for additional offenders. A thief may, for instance, approach the bank directly of a stolen card and request that the account's postal address be changed. The perpetrator may then obtain a passport or driver's license bearing his own photo but bearing the identity of the victim.

3.2 Research Design

The increasing use of social networks in society is a fact. Proof of this are the millions of users who have created an account on a social network such as Facebook, Twitter or LinkedIn among many others. The purpose of these platforms is to create connections between people, allowing them to share with their contacts aspects of daily life, thoughts, etc. In this way, many cybercriminals have seen these platforms as an easy way to extort people, stealing their identities, taking advantage of their data, and even impersonating the victims themselves to share phishing campaigns, malware or other types of viruses. Therefore, it is necessary to have a system that allows a fast detection of identity theft. This paper proposes the use of a novel system that can be implemented on any of these social networks, allowing to detect cases of identity theft.

In this proposed work, Fuzzy logic with Mamdani FIS model is used to achieve detection of identity theft in social media because Fuzzy logic, first introduced by L. A. Zadeh in 1965. Following three conventional steps of Fuzzy process are performed in our methodology as shown in figure-1 and figure-2 portraying the Basic Fuzzy Models:

- Fuzzification of input variables
- Inference (Mamdani FIS)
- Defuzzification

The method has been frequently utilized to portray the unpredictability of real-world events. The quantity and diversity of applications based on fuzzy set theory have increased dramatically during the last decade. Fuzzy set theory was utilized to examine the quality performance of social media security system in the realm of computer security. The study focuses on the complicated and evolving nature of the different elements that are taken into account when evaluating social media security. They were satisfied that the fuzzy logic model is a useful tool for examining and evaluating the security and theft quantity. In dispersed intrusion forecasting and crucial method, fuzzy set theory is frequently used to analyses online risk.

The following phases were used to determine the empirical methodology and design execution for evaluating users' perceptions of privacy on online social networking sites using fuzzy set theory:

3.3 Linguistic Variables Design Model

Transparency, consistency, and scalability were the system's inputs. These criteria, or linguistic factors, are believed to be of equal weight, and each is assigned a numerical value based on responses to questions regarding this particular social network. The frequently appears (that is, secrecy, integrity, and availability) must be represented by fuzzy sets while designing the fuzzy system. A feature set is used to represent the fuzzy sets.

Fuzzy Basic Model-1

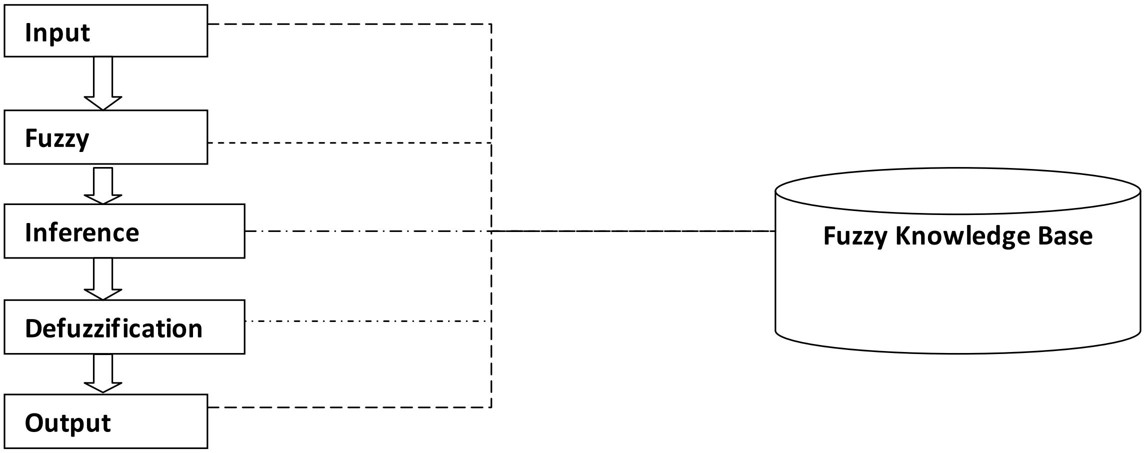


Figure 1

Fuzzy Basic Model-2

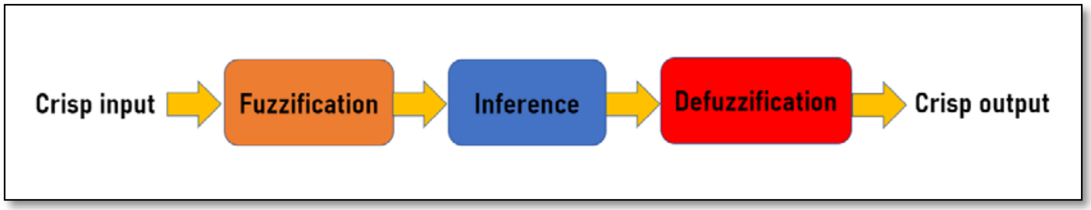


Figure 2

3.4 Fuzzy Sets

Fuzzy sets created by triangle membership functions were used to uphold the views of linguistic variables. The triangle membership function was chosen primarily for its simplicity and suitability for this project. Each linguistic variable is given five values, which is regarded an optimal number because more than five values complicate the design. Not private, slightly top secret information, classified information, very confidential, and highly confidential were used to characterize the degree of anonymity as a language variable. The degree of integrity was likewise described on a scale of low, high, and extremely high, while the level of unavailability was defined on a scale of never, rarely, frequently, very often, and always available.

To detect identity theft, our research relies on behavioral features of an individual. To do so, the natural language parameters and numerical values of those were considered.

In our work data of more than 200 individuals' behaviour was taken for the year 2021. The data consisted of parameters like Posts uploaded by an individual, text messages sent by an individual, text messages sent to new recipients, display picture changes done by an individual, average location: City and Country, device used to perform the activities and display name of an individual.

Our work was based upon 9 Different Parameters based on user's behaviour mostly, taken from given dataset. To clarify in more convenient way, we segregated these variables into 2 categories:

- Linguistic Variables.
- Non-Linguistic Variables.

As shown in below Table 2, where F.P represents Changes in Frequency of Posts, F.M represents Changes in Frequency of Messages, N.R represents Presence of New Recipients in Messages List, F.N.R represents Changes in Frequency of New Recipients in Messages List, D.P represents Changes in Display Picture, F.P represents Changes in Display Name, D.T represents Changes in Type of device used.

3.5 Linguistic Variables

In our case, Linguistic Variables of a social media profile were taken into consideration like Posts, Messages Display Name etc.

3.6 Non-Linguistic Variables

In case of Non-Linguistic variables, discrete valued membership functions were taken into considerations i.e. Yes and No or Changed or Unchanged.

Linguistic Variables	
F.P	In our case unlike previously proposed idea of using text-mining and Machine Learning technique [3] being applied to each post done by user, an average of post per hour for each month was taken and then analyzed if this parameter is pointing towards an identity compromise.
F.M	Here an average of messages sent by user to different recipients per hour for each month was taken and comparison was done to see deviation and then analyzed if this parameter is pointing towards an identity compromise.
D.N	Normally it is not usual thing to see a change in display name. When there are abnormalities found in other considered parameters along with change in display name, this puts a huge impact on the credibility of genuineness of a user account.
Non-Linguistic Variables	
N.R	Normally it is expected from a regular user to text someone who has already texted before, despite that presence of a new recipients is not quite uncommon. However, when there are abnormalities found in count or frequency of texts sent

	along with presence of new recipients as well, this puts a huge impact on the credibility of genuineness of a user account.
F.N.R	In the presence of new recipients, we tracked the average count of new recipients contacted in an hour during that whole month, then compared it with data of whole year, this greatly enghance our ability to accurately decide if the user's identity is compromised.
D.P	Normally the change is profile picture is considered a usual behavior. However, when there are abnormalities found in other considered parameters along with change in display picture, this puts a huge impact on the credibility of genuineness of a user account.
Location	In our datasets, average geolocation based upon threat intelligence applied on the IPs of user were considered, the result of these were in the form of user's presence in a city and country averagely in a month. Users normally travel from one city to another, which is quite close to normal human behavior, however, change of city, country in combination of abnormalities found in other parameters raises concerns.
D.T	Normally it is not usual thing to see a change in type of device used by an individual to perform his regular social media activities, despite that this change is considered benign. However, when there are abnormalities found in other considered parameters along with change in device type, this puts a huge impact on the credibility of genuineness of a user account.

Table 2

Since there were 9 parameters and 24 membership functions combined shown in below complete flow chart (figure 3), these contributed to huge number of rules and were causing excessive usage of resources and computation, therefore, these 9 parameters were distributed into 4 different groups L1, L2 and L3 and Geo-Location.

L1 – gave the fuzzified output of Average Posts per hour and Average messages per hour, in the form of L1 Member ship function (figure 4)

Complete Flow Chart

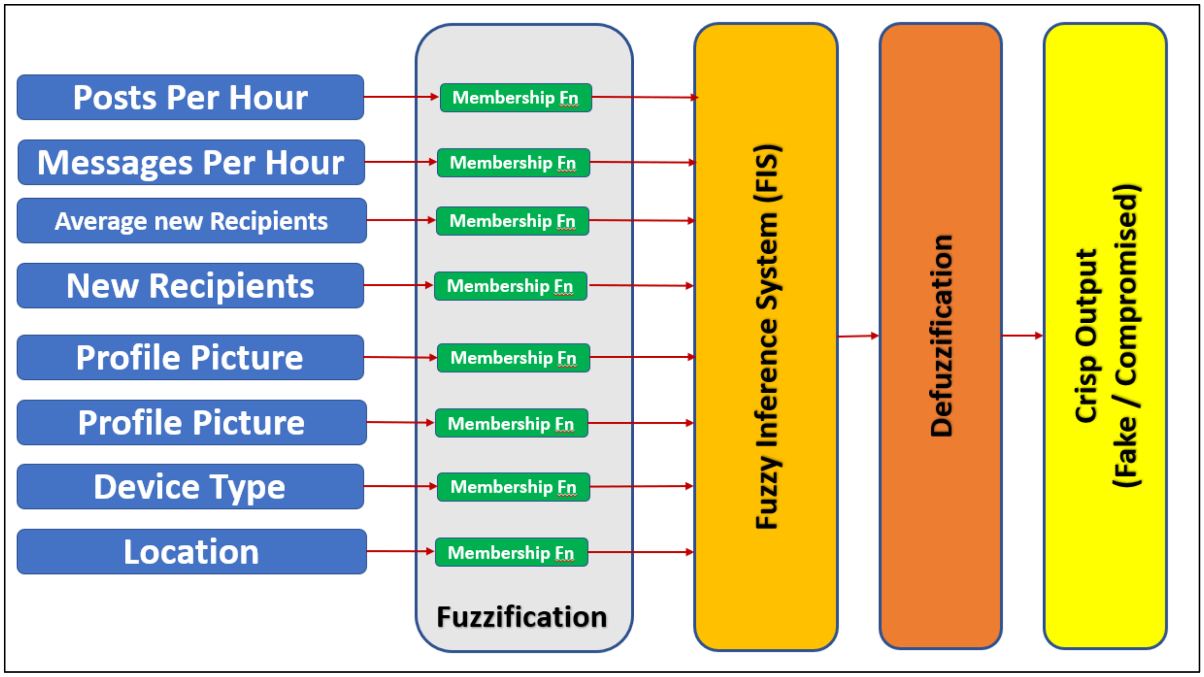


Figure 3

L1 Member ship function

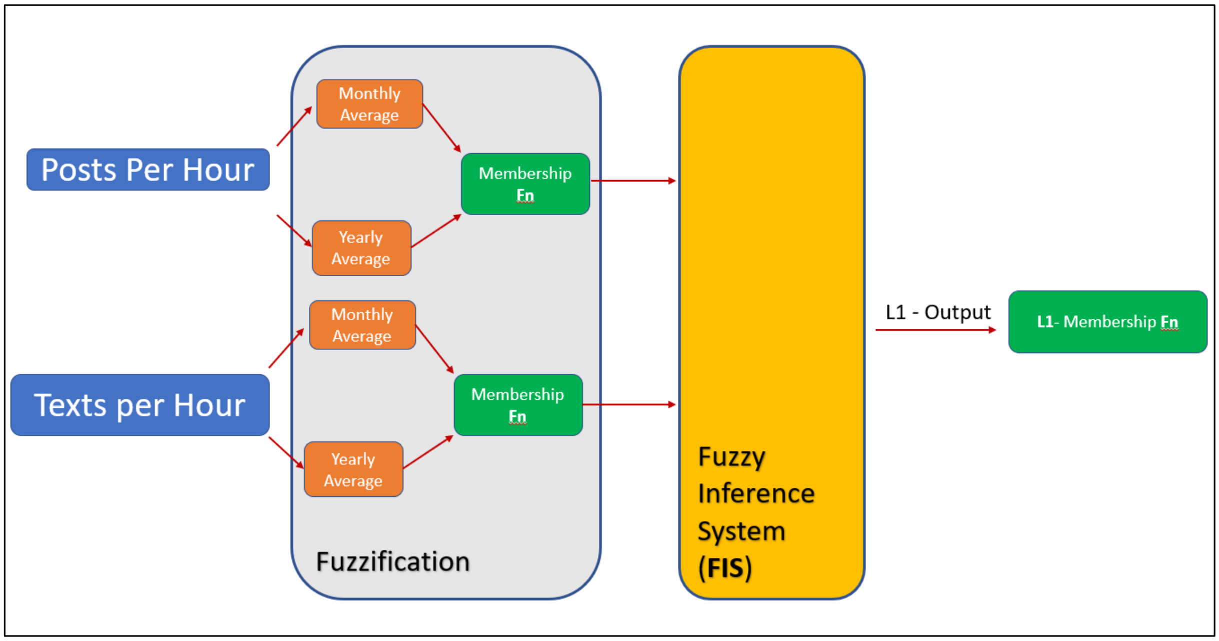


Figure 4

L2 – gave the fuzzified output of Average New Recipients per hour and normality of New Recipients' presence in comparison to past behavior, this result was given out in the form of L2 Member ship function (figure 5).

L2 Member ship function

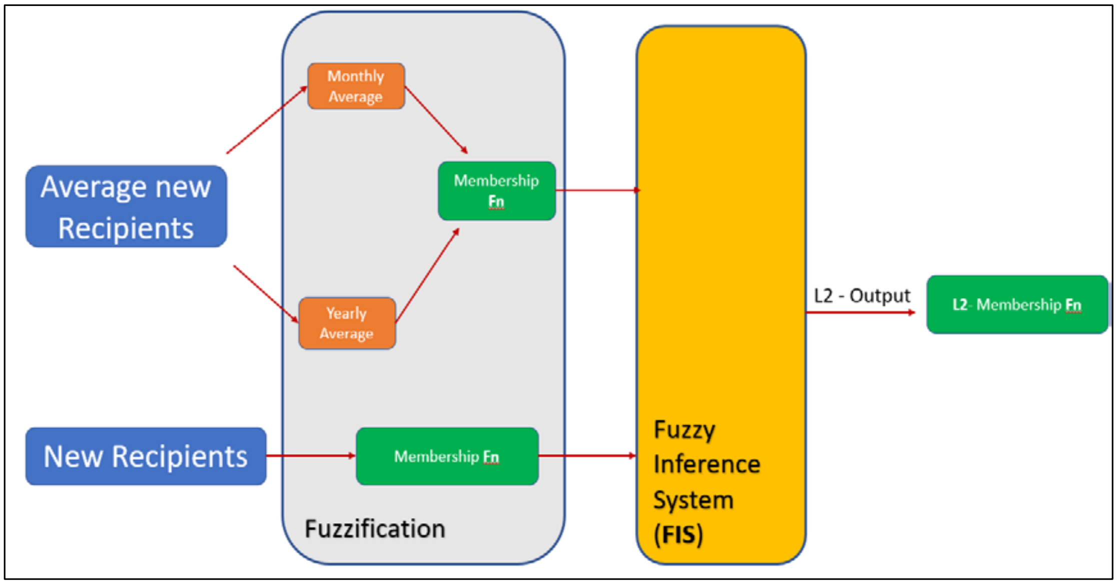


Figure 5

L3 – gave the fuzzified output of any change in Display name, picture, and normality of Device type under the use, in comparison to past behavior, this result was given out in the form of L3 Member ship function (figure 6).

L3 Member ship function

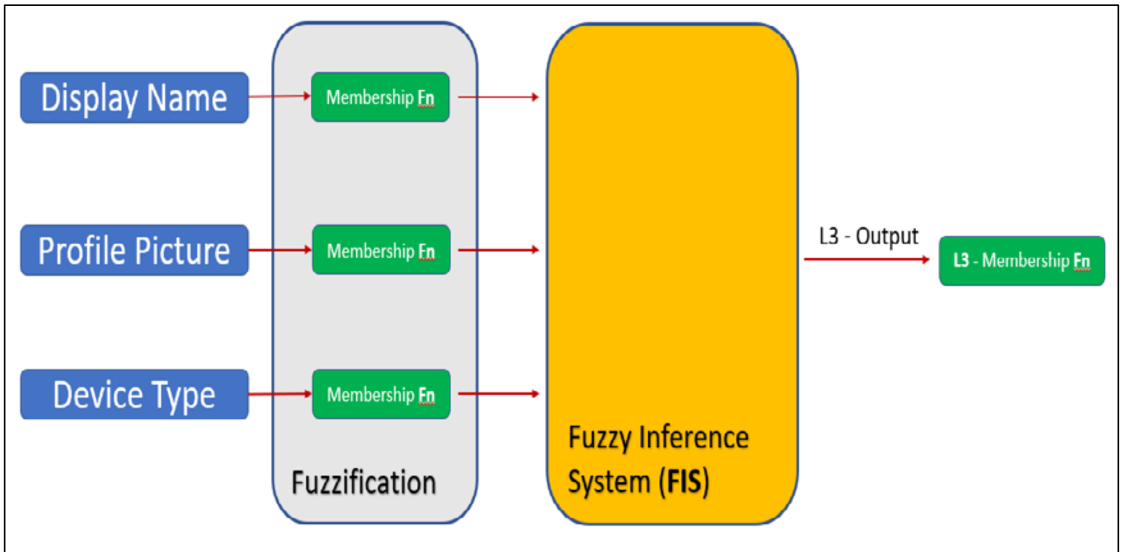


Figure 6

Geo-Location Member ship function

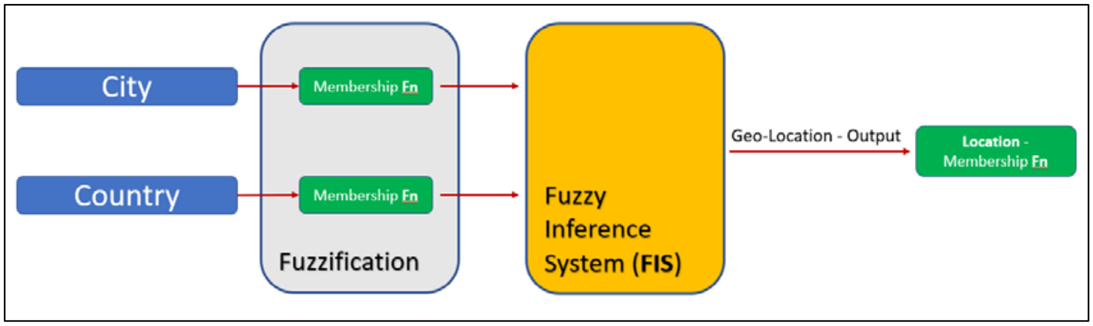


Figure 7

All the above L1, L2 and L3 membership functions were fed into another Fuzzy inference system give out a Linguistic Variables which was later used in combination with Geo-Location. Membership Fn (figure 7) in our final Mamdani Fuzzy inference system.

Sample Size And Sampling Technique

Teenagers have the highest level of trust in news and information posted on social media platforms. In comparison to the rest of the participants, who do not trust any information save news provided by government-verified accounts, they do not trust any information. However, in many circumstances, these accounts are vulnerable to theft and may disclose a large amount of false information before being restored.

The most common source of confusion about how social networks handle user information is that consumers never read the privacy policies. Many users are unaware of how social network companies safeguard and manage their personal information.

As a result, the vast majority of participants have little faith in social network providers.

Survey Result

Twitter, Instagram, YouTube, Face book, and Google+ are the most popular social media platforms. Instagram is the most popular social media platform among primary students. Twitter was the place to be for high school, college, and graduate education. The most common uses of social networks were to enjoy and pass the time, obtain important information, and, of course, interact with friends / relatives. When it comes to choosing a false identity, teens have the best success rate; participants prefer to write their whole genuine names. The degree of privacy varied depending on the amount of schooling. School kids and undergraduates, on the other hand, were more accessible in their privacy since they chose to make their profiles public. Some of them seem to be completely unaware of it.

As a result, the majority of individuals who wished to verify their accounts chose to make their profiles public. Individuals who made their accounts private were a small percentage of those who wished to authenticate their profiles. Emails, city, contact information, hobbies, photos, local geography, education, and marital status are the top information that users disclose on social networks. 15.5 percent of participants had no information on their profiles, indicating that they do not care about validating their profiles because they do not see the need to secure their information.

Experimental Setup

4.1 Lab Environment

In our experimental setup, the tests were performed on Core-i7 10th Generation System with 16GB ROM. The tool used to perform the experiment was MATLAB 2021b version. Total 9 parameters were taken into consideration.

4.2 Defining The Membership Functions

Unfocused singleton

A fuzzy set having a membership function that is zero at all other points and unity at one specific point.

Single output method

An output function that, rather of following a continuous curve, is represented as a spike at a single integer. It is only supported in the Fuzzy Logic Toolbox when used as a component of a zero-order Sugeno model.

Functions in tool box

The mapping of each point in the input space to a membership value (or network function) between 0 and 1 is defined by a curve known as a membership function (MF). The universe of discourse is another name for the input space.

Really, the sole prerequisite for a membership function is that it must range between 0 and 1.

There are 11 different membership function types included in the Fuzzy Logic Toolbox. These 11 functions are constructed from a number of fundamental functions in turn:

- Linear functions in pieces
- Function of the Gaussian distribution

- The sigmoid curve
- Polynomial quadratic curves
- Polynomial cubic curves

Straight lines can be used to provide the most basic membership functions. The benefit of these straight line membership functions is their simplicity.

- Trimf is a triangular membership function
- Function of trapezoidal membership: trapmf
- On the Gaussian distribution curve, two model parameters are constructed: a straightforward Gaussian curve and a two-sided composite of two separate Gaussian curves. Gaussmf and Gauss2mf are the two functions.
- Three parameters define the generalized bell membership function, which is known as gbellmf, Sigmf, or the sigmoid membership function.

Curves based on polynomials The Z, S, and Pi curves, all called for their shapes, and are three related membership functions (The functions zmf, smf, and pimf).

The nonlinear transformation of a data set to a scalar outcome data is known as a fuzzy logic system. Four components make up a Fuzzy logic system:

- A fuzzifier (Fuzzification)
- Rules
- Engine of inference
- Fuzz remover (Defuzzification)

Fuzzy linguistic parameters, linguistic variables words, and membership functions are used to transform a crisp set of input data into a fuzzy set in the course of fuzzy logic. Fuzzification is the process that follows.

Consisting of a set of rules, an inference is drawn. In the defuzzification process, the fuzzy output is translated into a crisp value using the membership functions.

Fuzzification

The method of employing membership functions to produce similarity score for a fuzzy variable.

Defuzzification

The process of turning a fuzzy inference system's output into a clear output. Defuzzification is necessary to get a clear output since the outcome is fuzzy. The defuzzifier part of a FLS serves this function. According to the output variable's membership function, defuzzification is carried out.

Numerous methods are feasible for this defuzzification, which is not a component of "mathematical fuzzy logic." The most popular defuzzification algorithms are presented.

1. Determining the gravitational centre
2. Determining the singletons centre of gravity
3. Calculating the mean average
4. Locating the maximum on the left
5. Deciding on the ideal maximum

Choose between two appropriate membership function types. It is highly beneficial from the service providers' point of view to understand the user's opinion in a relevant way. We already have a comprehensive understanding of the relationship between the quality component (responsiveness) and the shifting trends in user experience thanks to the fuzzy sets. From each of the current fuzzy sets, one may notice a symmetrical bell-shaped curve that depicts the dispersion of fuzzified opinion scores throughout the reaction periods. We investigated what sort of membership function may be a suitable option to represent the distribution of FOS as a result of this visual observation.

The traditional bell-shaped membership function is one potential contender. The first option is logically rejected because of the significant approximation error that is brought on by the irregularity in the data points. Another option is to use the approximation technique known as the "truncated function" in order to reduce approximation error. The issue of a significant approximation error has been resolved by this strategy, which involves building a chain of bell-shaped functions. It is important to note that this approach must modify numerous bell-shaped functions in order to achieve high accuracy. In other words, accuracy will grow as the number of functions increases. We direct the readers to for further information on this strategy.

Utilizing a Gaussian membership function is the third choice. For three major purposes, we used this membership function:

- Owing to the data's randomness: The Central Limit Theorem establishes that a large number of random variables added together have a normal distribution.
- The Gaussian function makes it easy to describe the resulting fuzzy sets because of their bell-curve form.
- Due to its straightforward formulation and favorable statistical characteristics, the Gaussian membership function is fairly used in practice.

The associated 3D visualization would be challenging to see here due to the enormous number of data points in the final partition matrix. Figure 4.4 shows a few of the original partition matrix's elements that were chosen in order to clarify the observation. We have formatted the graph such that it matches the partition matrix. The strength of the relationship between the collection of response times and the clusters is therefore represented by each individual bar. The clustered bars also contrast values on the horizontal axis.

Figure 6 shows the agent's first path after leaving a specific location in the environment. With Bug World, this setting is included. After completing the upper-left corner target, the agent moves on by adhering to bounds. The charging station is in the lower left corner where it stops. The upper right target and lower right target are then consecutively reached by continuing to follow boundaries. The agent then returns to the charging station and the upper left objective. At this point, the agent is able to position itself in its topological graph, and decides that it is time to start exploring other areas of the room.

The membership functions were formulated as per their characteristic’s outcome. In case of majority linguistic variables, they were broken down into following 3 categories.

- Normal Behavior: Normal Behavior contributes to regular behavior which is expected from a normal non compromised user.
- Little Deviated Behavior: Little Deviated behavior contributes to slightly suspicious behavior and might raise some concern.
- Highly Deviated Behavior: Highly Deviated Behavior contributes to very high chances of identity compromise and is a big concern

Membership Fn of Linguistic Variables

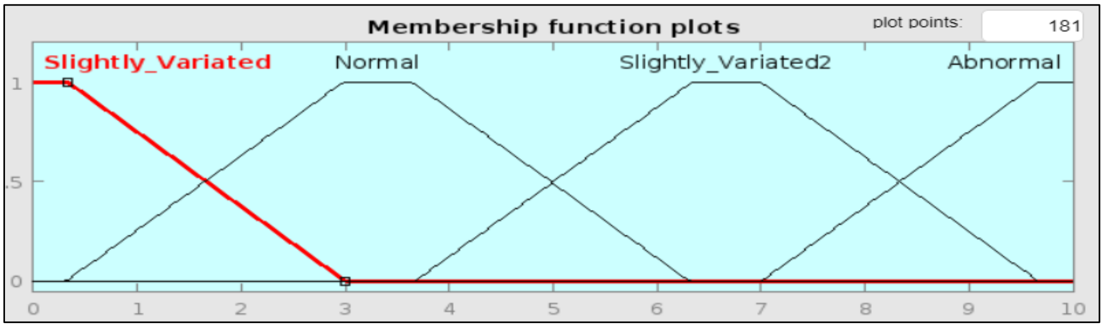


Figure-8

Apart from above categorization in figure 8, some of the variables had discrete membership functions, like Changed-Same or Normal-Abnormal values as shown below in figure 9.

Membership Fn of Non-Linguistic Variables

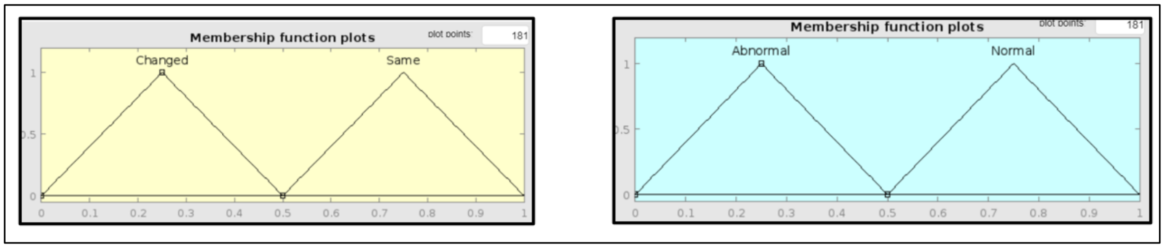
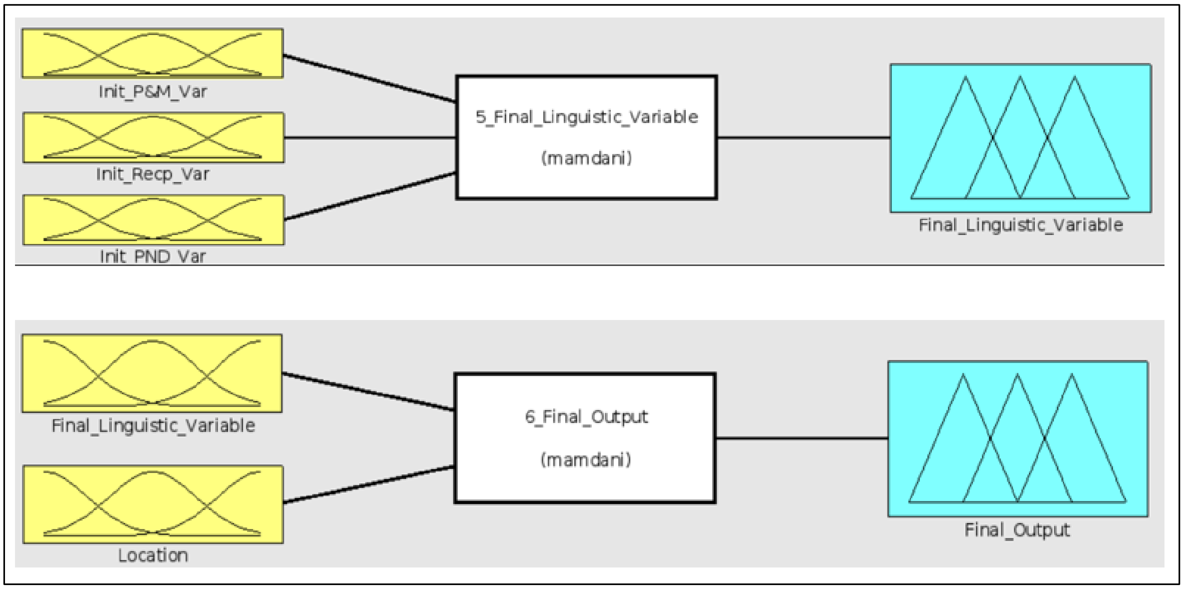


Figure 9

4.3 Expected Outcome

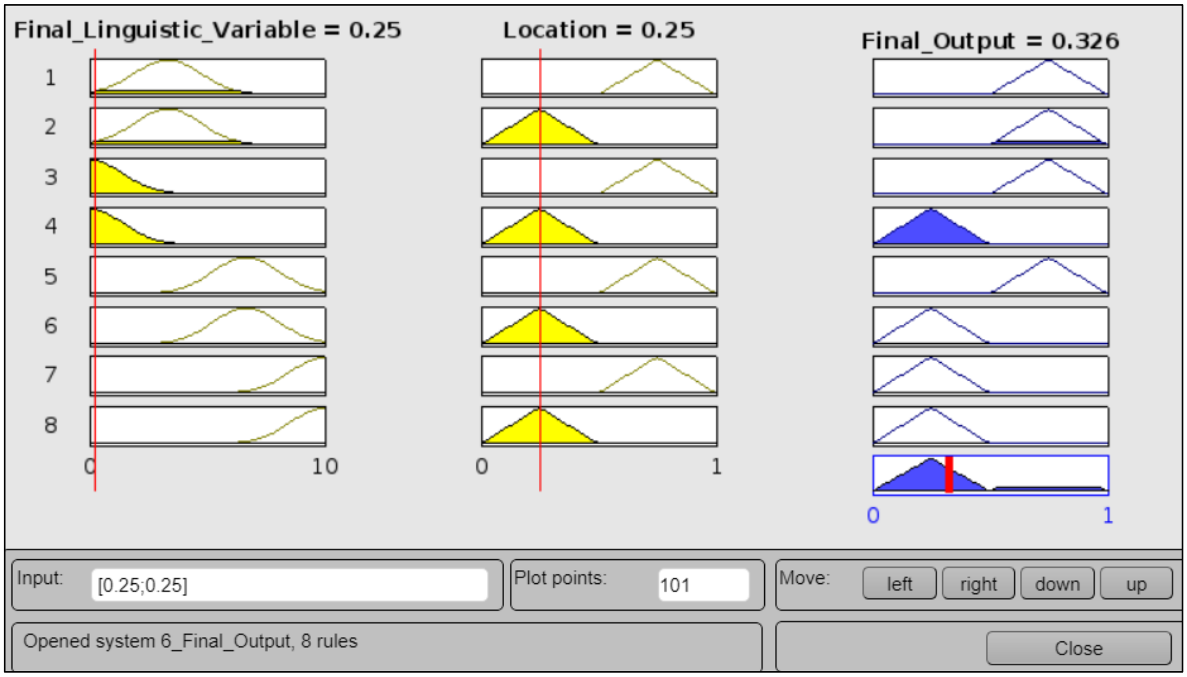
The fuzzified non-crisp outcome which then had solid grounds to give verdict about genuineness of a user’s account was then fed into Defuzzification step which showed final Output if the user’s identity is real or fake/compromised as can be seen in figure 10-12

Final Outcome -1



Figuree-10

Final Outcome - 2


Final_Linguistic_Variable = 0.25

Location = 0.25

Final_Output = 0.326

Input: [0.25;0.25]

Plot points: 101

Opened system 6_Final_Output, 8 rules

Figure - 11

Graphical Representation of Final Outcome

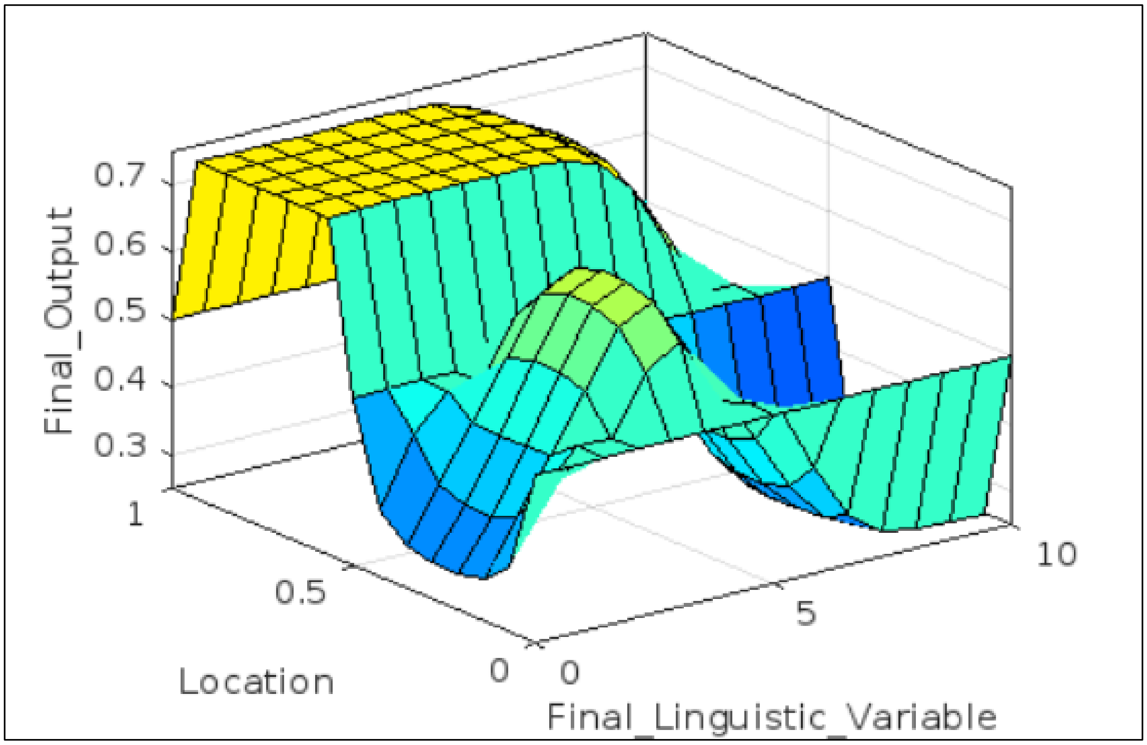


Figure-12

There are several membership functions, including:

- Triangle
- Trapezoid
- Linearly piecewise
- Gaussian
- Singleton

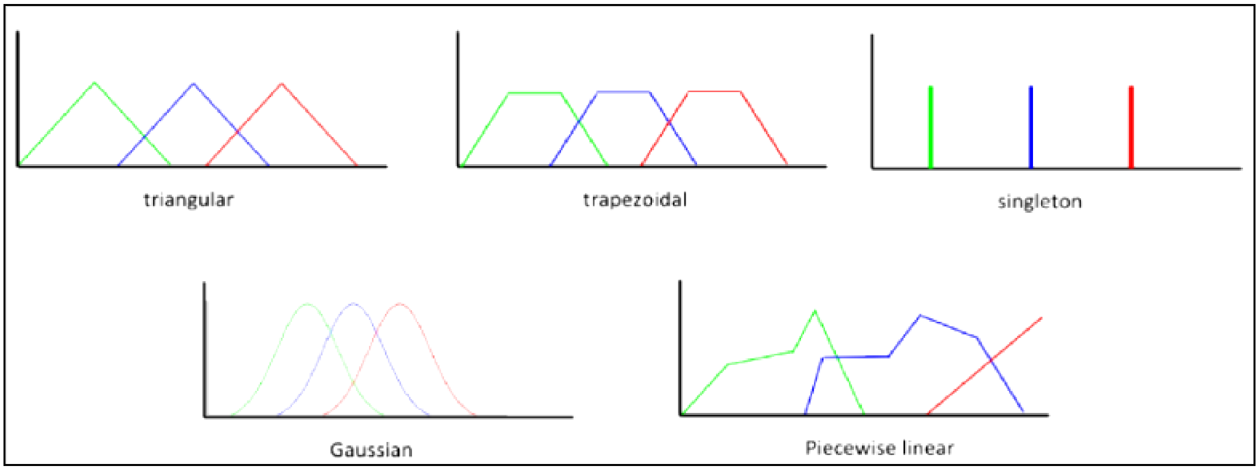


Figure-13

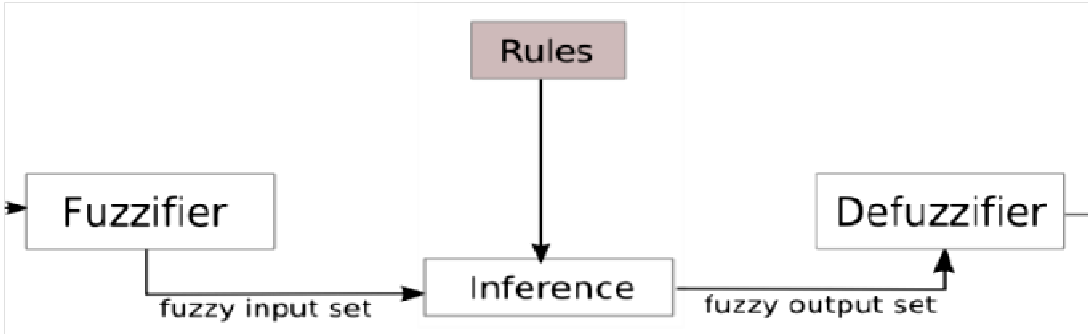


Figure-14

Results and Analysis

5.1 Testing Phase

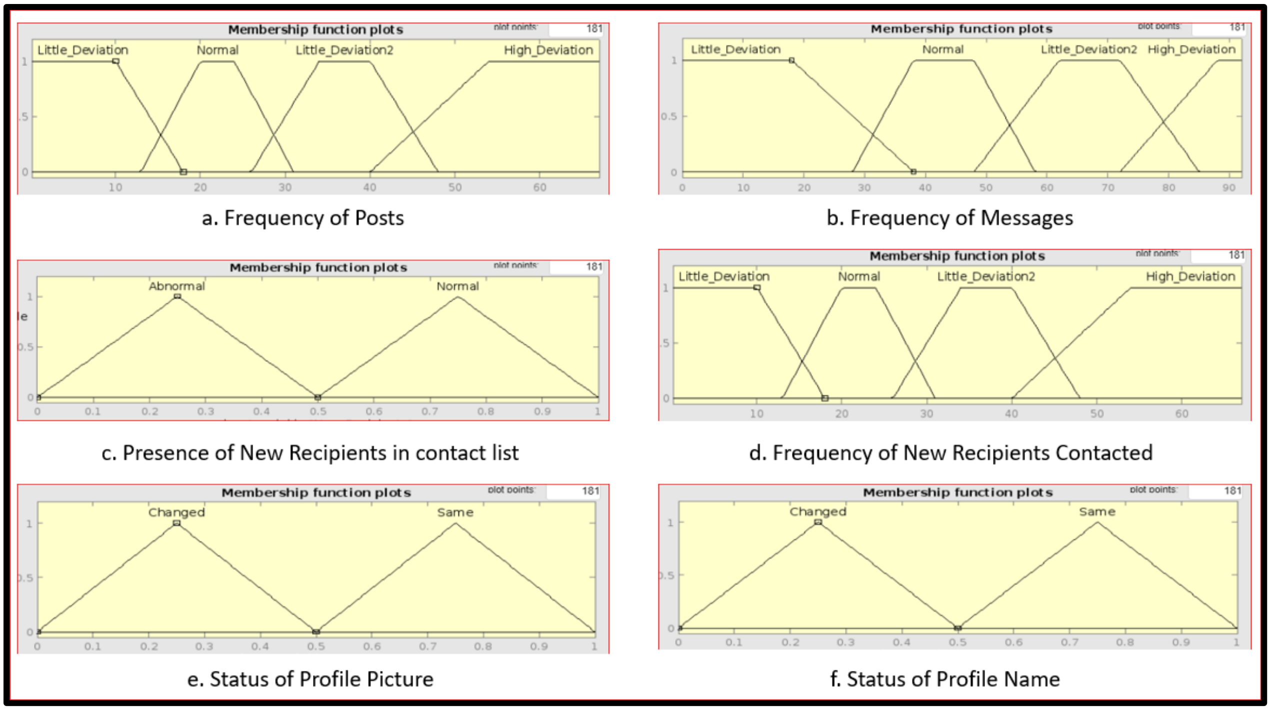
Utilizing a Gaussian membership function is the third choice. For three major purposes, we used this membership function:

- Owing to the data's randomness: The Central Limit Theorem establishes that a large number of random variables added together have a normal distribution.
- The Gaussian function makes it easy to describe the resulting fuzzy sets because of their bell-curve form.
- Due to its straightforward formulation and favorable statistical characteristics, the Gaussian membership function is fairly used in practice.

The associated 3D visualization would be challenging to see here due to the enormous number of data points in the final partition matrix. A few of the original partition matrix's elements that were chosen in order to clarify the observation. We have formatted the graph such that it matches the partition matrix. The strength of the relationship between the collection of response times and the clusters is therefore represented by each individual bar. The clustered bars also contrast values on the horizontal axis.

5.2 Membership Functions

Our framework achieves adaptive and practical approach to detect identity frauds based on user behavior as shown in figure 15 and figure 16 without depending on traditional machine learning techniques which are computationally expensive and still able to give better accuracy. Our research uses Mamdani FIS model.



a. Frequency of Posts

b. Frequency of Messages

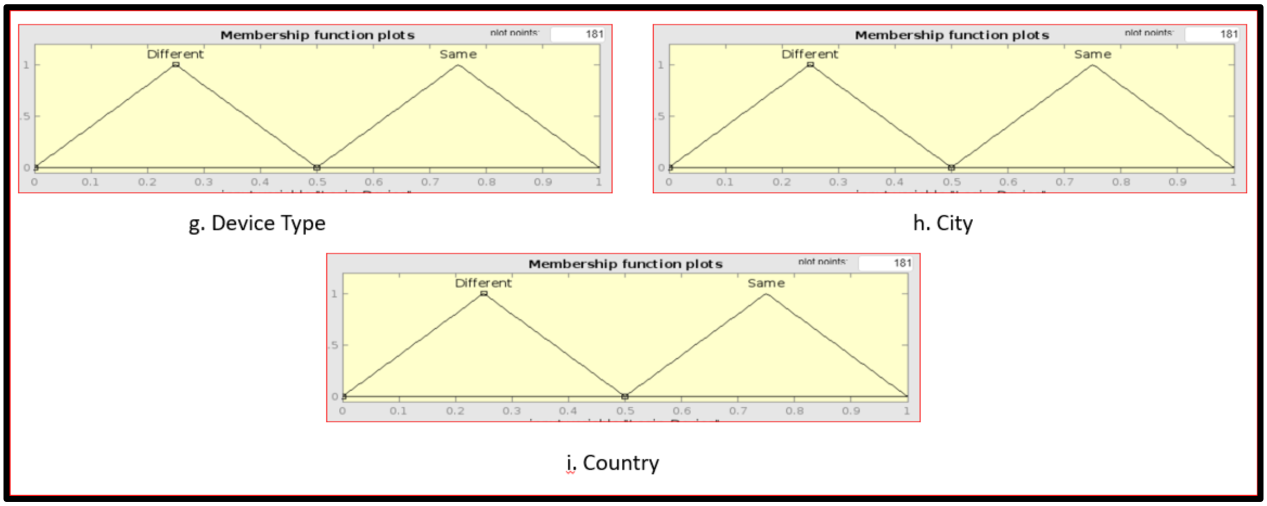
c. Presence of New Recipients in contact list

d. Frequency of New Recipients Contacted

e. Status of Profile Picture

f. Status of Profile Name

Figure 15



g. Device Type

h. City

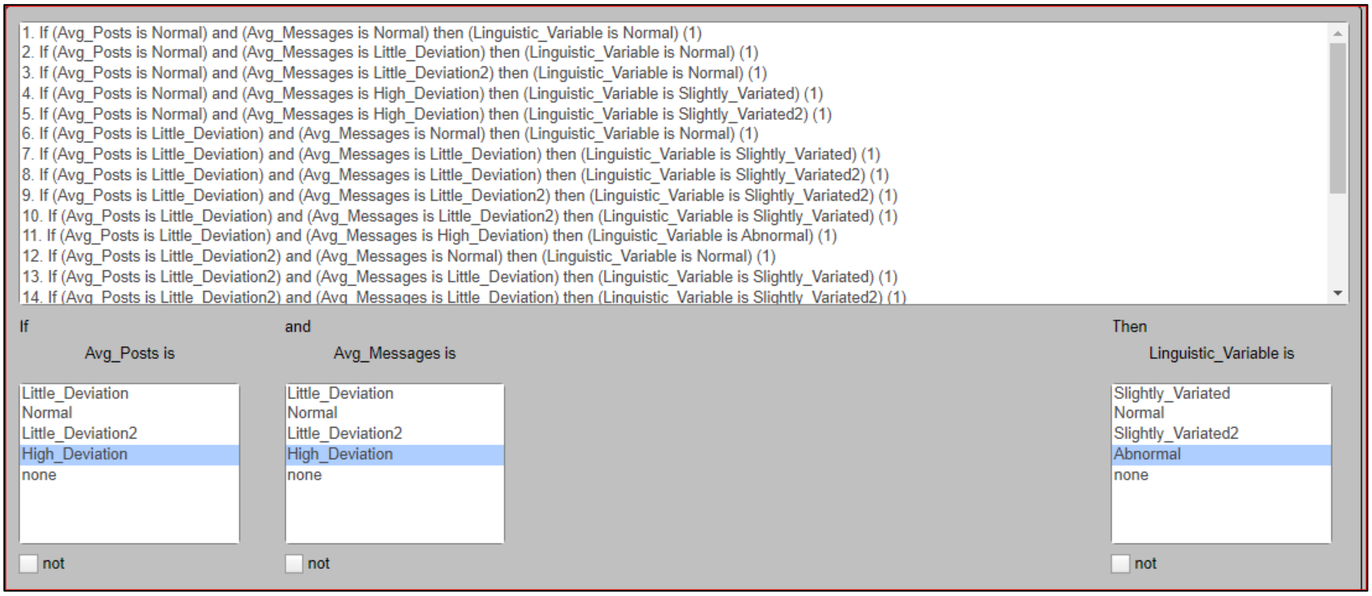
i. Country

Figure 16

5.3 Defining Rules

Since there were 9 parameters and 24 membership functions combined, these contributed to huge number of rules and were causing excessive usage of resources and computation, therefore, these 9 parameters were distributed into 4 different groups L1(figure-17), L2(figure-18) and L3(figure-20) and Geo-Location(figure-19).

Rules for average on frequency of posts and messages (L1)



1. If (Avg_Posts is Normal) and (Avg_Messages is Normal) then (Linguistic_Variable is Normal) (1)
2. If (Avg_Posts is Normal) and (Avg_Messages is Little_Deviation) then (Linguistic_Variable is Normal) (1)
3. If (Avg_Posts is Normal) and (Avg_Messages is Little_Deviation2) then (Linguistic_Variable is Normal) (1)
4. If (Avg_Posts is Normal) and (Avg_Messages is High_Deviation) then (Linguistic_Variable is Slightly_Variated) (1)
5. If (Avg_Posts is Normal) and (Avg_Messages is High_Deviation) then (Linguistic_Variable is Slightly_Variated2) (1)
6. If (Avg_Posts is Little_Deviation) and (Avg_Messages is Normal) then (Linguistic_Variable is Normal) (1)
7. If (Avg_Posts is Little_Deviation) and (Avg_Messages is Little_Deviation) then (Linguistic_Variable is Slightly_Variated) (1)
8. If (Avg_Posts is Little_Deviation) and (Avg_Messages is Little_Deviation) then (Linguistic_Variable is Slightly_Variated2) (1)
9. If (Avg_Posts is Little_Deviation) and (Avg_Messages is Little_Deviation2) then (Linguistic_Variable is Slightly_Variated2) (1)
10. If (Avg_Posts is Little_Deviation) and (Avg_Messages is Little_Deviation2) then (Linguistic_Variable is Slightly_Variated) (1)
11. If (Avg_Posts is Little_Deviation) and (Avg_Messages is High_Deviation) then (Linguistic_Variable is Abnormal) (1)
12. If (Avg_Posts is Little_Deviation2) and (Avg_Messages is Normal) then (Linguistic_Variable is Normal) (1)
13. If (Avg_Posts is Little_Deviation2) and (Avg_Messages is Little_Deviation) then (Linguistic_Variable is Slightly_Variated) (1)
14. If (Avg_Posts is Little_Deviation2) and (Avg_Messages is Little_Deviation) then (Linguistic_Variable is Slightly_Variated2) (1)

If

Avg_Posts is: Little_Deviation, Normal, Little_Deviation2, High_Deviation, none

□ not

and

Avg_Messages is: Little_Deviation, Normal, Little_Deviation2, High_Deviation, none

□ not

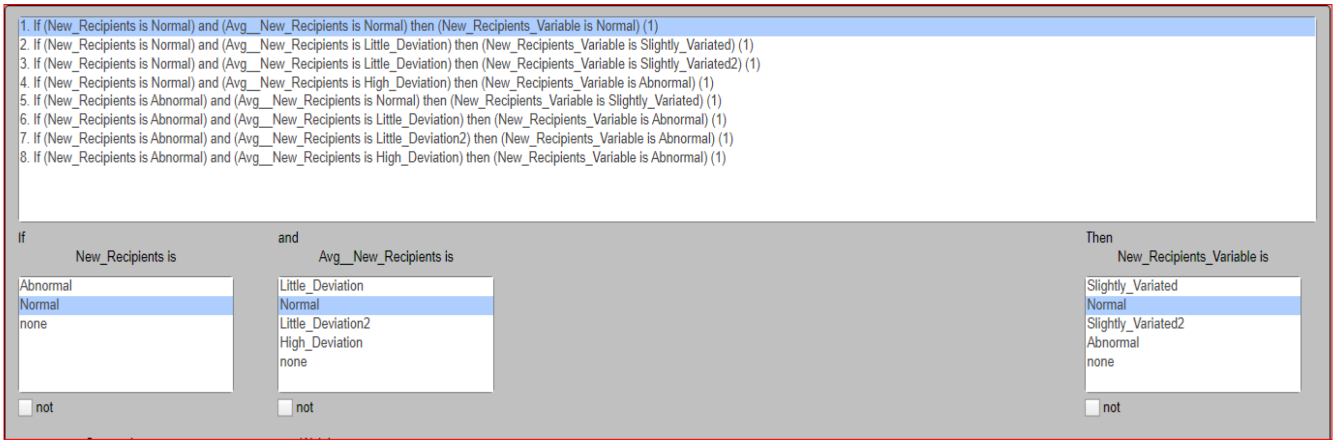
Then

Linguistic_Variable is: Slightly_Variated, Normal, Slightly_Variated2, Abnormal, none

□ not

Figure – 17

Rules for presence of new recipients and average of new recipients contacted (L2)



1. If (New_Recipients is Normal) and (Avg__New_Recipients is Normal) then (New_Recipients_Variable is Normal) (1)
2. If (New_Recipients is Normal) and (Avg__New_Recipients is Little_Deviation) then (New_Recipients_Variable is Slightly_Variated) (1)
3. If (New_Recipients is Normal) and (Avg__New_Recipients is Little_Deviation) then (New_Recipients_Variable is Slightly_Variated2) (1)
4. If (New_Recipients is Normal) and (Avg__New_Recipients is High_Deviation) then (New_Recipients_Variable is Abnormal) (1)
5. If (New_Recipients is Abnormal) and (Avg__New_Recipients is Normal) then (New_Recipients_Variable is Slightly_Variated) (1)
6. If (New_Recipients is Abnormal) and (Avg__New_Recipients is Little_Deviation) then (New_Recipients_Variable is Abnormal) (1)
7. If (New_Recipients is Abnormal) and (Avg__New_Recipients is Little_Deviation2) then (New_Recipients_Variable is Abnormal) (1)
8. If (New_Recipients is Abnormal) and (Avg__New_Recipients is High_Deviation) then (New_Recipients_Variable is Abnormal) (1)

If

New_Recipients is: Abnormal, Normal, none

□ not

and

Avg__New_Recipients is: Little_Deviation, Normal, Little_Deviation2, High_Deviation, none

□ not

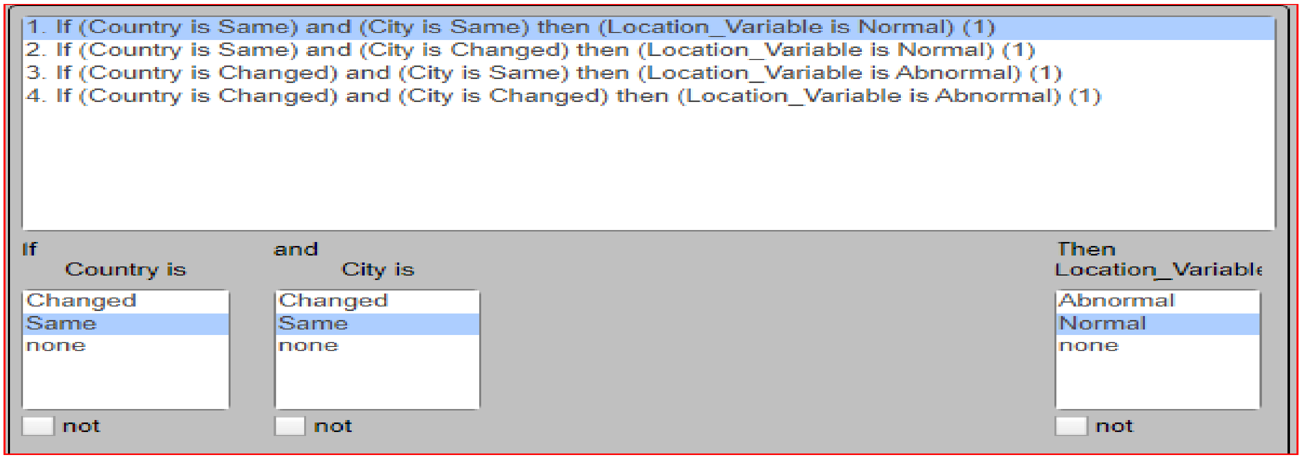
Then

New_Recipients_Variable is: Slightly_Variated, Normal, Slightly_Variated2, Abnormal, none

□ not

Figure - 18

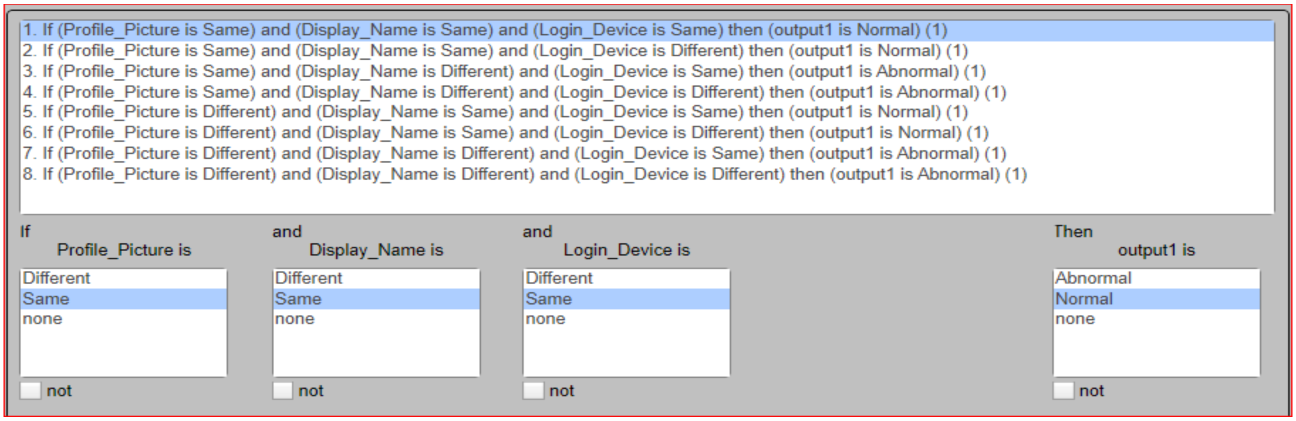
Rules for change in location



```
1. If (Country is Same) and (City is Same) then (Location_Variable is Normal) (1)
2. If (Country is Same) and (City is Changed) then (Location_Variable is Normal) (1)
3. If (Country is Changed) and (City is Same) then (Location_Variable is Abnormal) (1)
4. If (Country is Changed) and (City is Changed) then (Location_Variable is Abnormal) (1)
```

Figure - 19

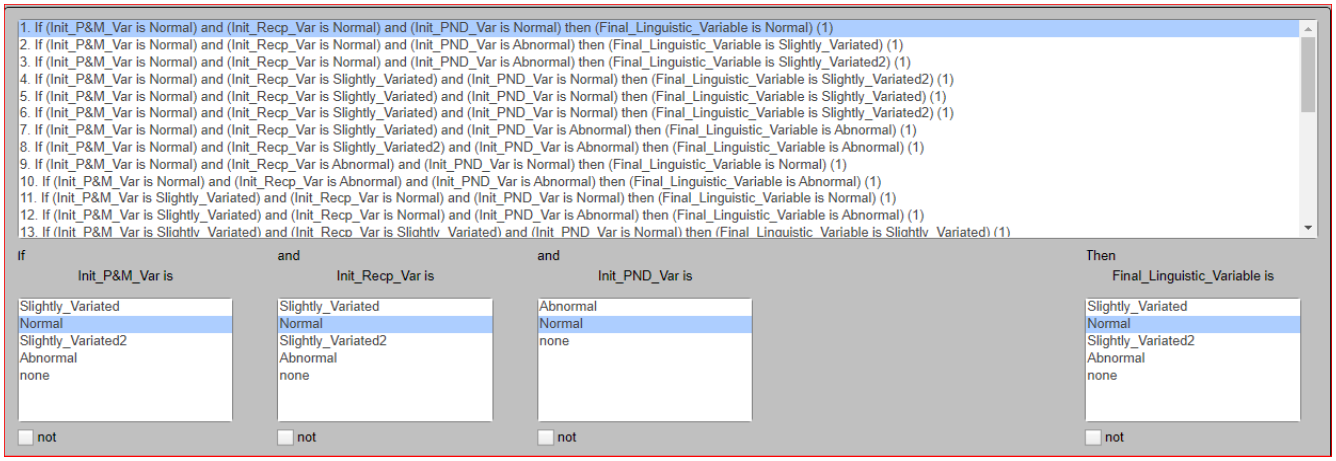
Rules for change in non-linguistic variables (L3)



```
1. If (Profile_Picture is Same) and (Display_Name is Same) and (Login_Device is Same) then (output1 is Normal) (1)
2. If (Profile_Picture is Same) and (Display_Name is Same) and (Login_Device is Different) then (output1 is Normal) (1)
3. If (Profile_Picture is Same) and (Display_Name is Different) and (Login_Device is Same) then (output1 is Abnormal) (1)
4. If (Profile_Picture is Same) and (Display_Name is Different) and (Login_Device is Different) then (output1 is Abnormal) (1)
5. If (Profile_Picture is Different) and (Display_Name is Same) and (Login_Device is Same) then (output1 is Normal) (1)
6. If (Profile_Picture is Different) and (Display_Name is Same) and (Login_Device is Different) then (output1 is Normal) (1)
7. If (Profile_Picture is Different) and (Display_Name is Different) and (Login_Device is Same) then (output1 is Abnormal) (1)
8. If (Profile_Picture is Different) and (Display_Name is Different) and (Login_Device is Different) then (output1 is Abnormal) (1)
```

Figure – 20

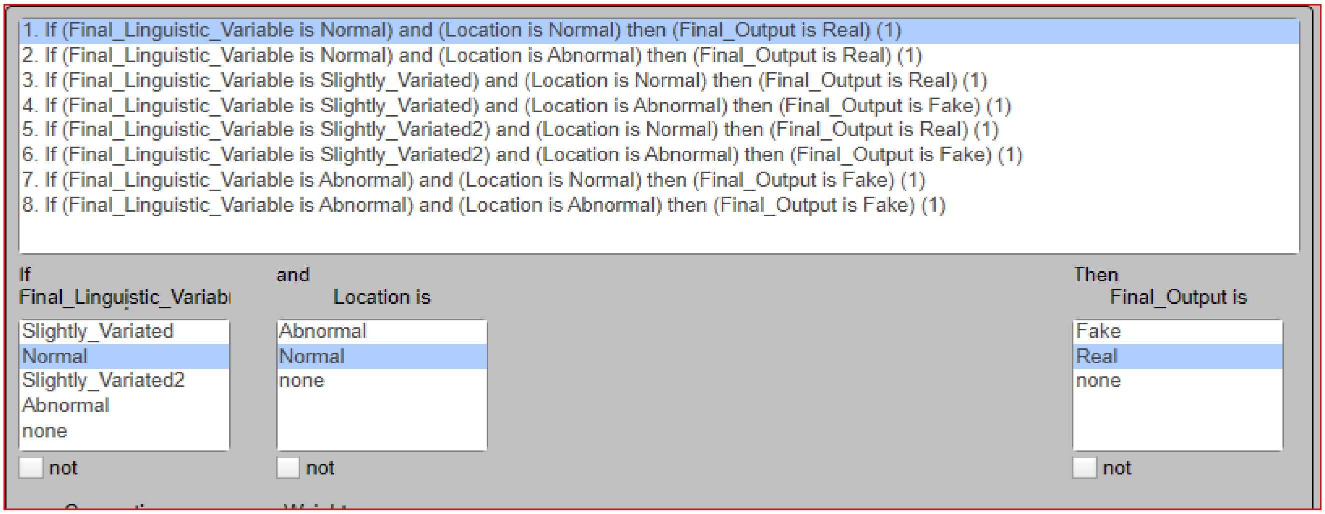
Rules for calculating Final linguistic variable



```
1. If (Init_P&M_Var is Normal) and (Init_Recp_Var is Normal) and (Init_PND_Var is Normal) then (Final_Linguistic_Variable is Normal) (1)
2. If (Init_P&M_Var is Normal) and (Init_Recp_Var is Normal) and (Init_PND_Var is Abnormal) then (Final_Linguistic_Variable is Slightly_Variated) (1)
3. If (Init_P&M_Var is Normal) and (Init_Recp_Var is Normal) and (Init_PND_Var is Abnormal) then (Final_Linguistic_Variable is Slightly_Variated2) (1)
4. If (Init_P&M_Var is Normal) and (Init_Recp_Var is Slightly_Variated) and (Init_PND_Var is Normal) then (Final_Linguistic_Variable is Slightly_Variated2) (1)
5. If (Init_P&M_Var is Normal) and (Init_Recp_Var is Slightly_Variated) and (Init_PND_Var is Normal) then (Final_Linguistic_Variable is Slightly_Variated) (1)
6. If (Init_P&M_Var is Normal) and (Init_Recp_Var is Slightly_Variated) and (Init_PND_Var is Normal) then (Final_Linguistic_Variable is Slightly_Variated2) (1)
7. If (Init_P&M_Var is Normal) and (Init_Recp_Var is Slightly_Variated) and (Init_PND_Var is Abnormal) then (Final_Linguistic_Variable is Abnormal) (1)
8. If (Init_P&M_Var is Normal) and (Init_Recp_Var is Slightly_Variated2) and (Init_PND_Var is Abnormal) then (Final_Linguistic_Variable is Abnormal) (1)
9. If (Init_P&M_Var is Normal) and (Init_Recp_Var is Abnormal) and (Init_PND_Var is Normal) then (Final_Linguistic_Variable is Normal) (1)
10. If (Init_P&M_Var is Normal) and (Init_Recp_Var is Abnormal) and (Init_PND_Var is Abnormal) then (Final_Linguistic_Variable is Abnormal) (1)
11. If (Init_P&M_Var is Slightly_Variated) and (Init_Recp_Var is Normal) and (Init_PND_Var is Normal) then (Final_Linguistic_Variable is Normal) (1)
12. If (Init_P&M_Var is Slightly_Variated) and (Init_Recp_Var is Normal) and (Init_PND_Var is Abnormal) then (Final_Linguistic_Variable is Abnormal) (1)
13. If (Init_P&M_Var is Slightly_Variated) and (Init_Recp_Var is Slightly_Variated) and (Init_PND_Var is Normal) then (Final_Linguistic_Variable is Slightly_Variated) (1)
```

Figure - 21

Final Outcome



1. If (Final_Linguistic_Variable is Normal) and (Location is Normal) then (Final_Output is Real) (1)
2. If (Final_Linguistic_Variable is Normal) and (Location is Abnormal) then (Final_Output is Real) (1)
3. If (Final_Linguistic_Variable is Slightly_Variated) and (Location is Normal) then (Final_Output is Real) (1)
4. If (Final_Linguistic_Variable is Slightly_Variated) and (Location is Abnormal) then (Final_Output is Fake) (1)
5. If (Final_Linguistic_Variable is Slightly_Variated2) and (Location is Normal) then (Final_Output is Real) (1)
6. If (Final_Linguistic_Variable is Slightly_Variated2) and (Location is Abnormal) then (Final_Output is Fake) (1)
7. If (Final_Linguistic_Variable is Abnormal) and (Location is Normal) then (Final_Output is Fake) (1)
8. If (Final_Linguistic_Variable is Abnormal) and (Location is Abnormal) then (Final_Output is Fake) (1)

Figure - 22

5.4 Analyzing The Data

In our case total 32 cases from the given dataset were tested as can be seen in Table 3 below where F.P represents Changes in Frequency of Posts, F.M represents Changes in Frequency of Messages, N.R represents Presence of New Recipients in Messages List, F.N.R represents Changes in Frequency of New Recipients in Messages List, D.P represents Changes in Display Picture, F.P represents Changes in Display Name, D.T represents Changes in Type of device used, F.O represents Final Outcome, N.B represents Normal Behavior, S.D represents Slightly Deviated behavior, H.D represents Highly Deviated behavior,

Out of those 32 users 12 were found to be fake or compromised while the rest of the tested users showed normal results as per their expected past behavior.

In our testing, it was found that in case of real identities there were slightly deviations in frequency of posts or messages but in major contributing factors like display name or major changes in geo-locations like country and device type the behaviors remained consistent.

In case of fake or compromised identities it was found that there were normal behaviors or slightly deviations in frequency of posts or messages but in major contributing factors like display name or major changes in geo-locations like country and device type the behaviors were quite deviated and unusual from expected ones.

Users	F.P	F.M	N.R	F.N.R	City	Country	D.P	D.N	D.T	F.O
Brian Collins	S.D	S.D	N.B	N.B	Changed	Unchanged	Changed	Unchanged	Unchanged	Real
Adriel Bullock	H.D	S.D	N.B	S.D	Changed	Unchanged	Changed	Unchanged	Unchanged	Real
Gabriela	N.B	S.D	N.B	S.D	Unchanged	Unchanged	Changed	Unchanged	Unchanged	Real
Zach	S.D	S.D	N.B	H.D	Changed	Changed	Changed	Changed	Changed	Fake
Alexandra Gorn	H.D	S.D	N.B	S.D	Changed	Unchanged	Changed	Unchanged	Unchanged	Real
Aayla Dianne	N.B	S.D	N.B	S.D	Unchanged	Unchanged	Changed	Unchanged	Unchanged	Real
Abmel	S.D	S.D	N.B	H.D	Changed	Changed	Changed	Changed	Changed	Fake
Adnan Sulehri	S.D	N.B	N.B	N.B	Changed	Unchanged	Unchanged	Unchanged	Unchanged	Real
Aaron Jefs	N.B	S.D	N.B	H.D	Changed	Changed	Changed	Changed	Changed	Fake
Ali Haider	S.D	N.B	N.B	N.B	Changed	Unchanged	Unchanged	Unchanged	Unchanged	Real
Bexley	N.B	S.D	N.B	S.D	Unchanged	Unchanged	Changed	Unchanged	Unchanged	Real
Giuseppe	S.D	S.D	N.B	N.B	Changed	Unchanged	Changed	Unchanged	Unchanged	Real
Ishfaq Chatha	S.D	N.B	N.B	N.B	Changed	Unchanged	Unchanged	Unchanged	Unchanged	Real
Rufus	S.D	S.D	N.B	N.B	Changed	Unchanged	Changed	Unchanged	Unchanged	Real
Kashif Saleem	S.D	N.B	N.B	N.B	Changed	Unchanged	Unchanged	Unchanged	Unchanged	Real
Quinn	N.B	S.D	N.B	H.D	Changed	Changed	Changed	Changed	Changed	Fake
Vidor	H.D	H.D	N.B	H.D	Changed	Changed	Unchanged	Changed	Changed	Fake
Fergus	H.D	S.D	N.B	S.D	Changed	Unchanged	Changed	Unchanged	Unchanged	Real
Simon	S.D	S.D	N.B	N.B	Changed	Unchanged	Changed	Unchanged	Unchanged	Real
walter Ulrick	S.D	S.D	N.B	H.D	Changed	Changed	Changed	Changed	Changed	Fake
Thea	N.B	S.D	N.B	S.D	Unchanged	Unchanged	Changed	Unchanged	Unchanged	Real
Yvette	S.D	S.D	N.B	H.D	Changed	Changed	Changed	Changed	Changed	Fake
Walter	H.D	S.D	N.B	S.D	Changed	Unchanged	Changed	Unchanged	Unchanged	Real
Delaney	N.B	S.D	N.B	H.D	Changed	Changed	Changed	Changed	Changed	Fake
Akarsh	N.B	N.B	Abnorma	S.D	Unchange	Unchange	Unchange	Unchange	Unchange	Real

			l		d	d	d	d	d	
Dilip	N.B	N.B	Abnormal	S.D	Unchanged	Unchanged	Unchanged	Unchanged	Unchanged	Real
Feivel Fergus	H.D	H.D	N.B	H.D	Changed	Changed	Unchanged	Changed	Changed	Fake
Darack wang	N.B	S.D	N.B	H.D	Changed	Changed	Changed	Changed	Changed	Fake
Dhruv	N.B	N.B	Abnormal	S.D	Unchanged	Unchanged	Unchanged	Unchanged	Unchanged	Real
Ehsaan Mian	N.B	N.B	Abnormal	S.D	Unchanged	Unchanged	Unchanged	Unchanged	Unchanged	Real
Kees Grunden	H.D	H.D	N.B	H.D	Changed	Changed	Unchanged	Changed	Changed	Fake
Irvin	H.D	H.D	N.B	H.D	Changed	Changed	Unchanged	Changed	Changed	Fake

Table-3

5.5 Comparative Analysis

As previously mentioned, there was a proposal idea in a paper published [3] in which it was suggested to use Fuzzy logic based on behavioral analysis but there was no practical implementation with very limited parameters considered.

DITOS by Fuzzy gives 98% accuracy, which gives it edge over the previously work done in this domain. Table 4 shows comparative analysis of our framework with existing solutions.

Paper	**Social Media Platform**	**Fuzzy Logic Implementation**	**Identity Theft**	**Use of Behavioral Parameters**	**Use of Feature Parameters**	**Use of Geo-Locations**
[1][2][18][5]	✓	×	✓	✓	×	×
[3]	✓	✓	✓	✓	×	✓
[19][20]	×	×	✓	✓	×	×
[11][7][12]	✓	×	✓	✓	×	✓
[21]	✓	×	✓	✓	✓	×
D.I.T.O.S BY Fuzzy	✓	✓	✓	✓	✓	✓

Table 4

In the fuzzification and defuzzification phases of a fuzzy logic system, membership functions are employed to convert non-fuzzy input data into fuzzy linguistic words and vice versa.

5.6 Discussion

It displays the overall number of categories created, the total number of data sets used, and the outcomes obtained using each method. The data sets are taken from newsgroups, publications, corpus pages of various text documents, websites, Reuters, and archives. In the beginning of the training phase, many categories are formed. These methods have been applied to documents from small to huge corpora and have taken into account a few too many categories. The outcomes of the experiments demonstrate how the corresponding approach is superior to others. Results include improved performance and accuracy, increased speed, decreased storage, and many other beneficial aspects.

In order to handle situations where decision-makers must properly examine and analyze information that is imperfect in nature, fuzzy set theory was established. In situations when precise information and facts are lacking, fuzzy sets offer a theoretical foundation and an analytical tool. Fuzzy sets and probabilistic reasoning have been used to solve a variety of scientific and technical issues since its inception. Fuzzy models, for instance, have been utilized to create microchips that execute fuzzy algorithms [37]. Fuzzy industrial controllers for improved functioning of today's trains, cars, and elevators, military equipment and operations, and robots all employ fuzzy logic at the moment. Fuzzy logic chips are also used in a range of consumer items, including television sets, cars, camcorders, and cameras [38]. In order to evaluate a fuzzy logic query for a hepatic technology, Hamam et al. created taxonomy. The taxonomy was then modeled with a fuzzy logic system and, in the end, was assessed by a Madman fuzzy system. The influence of several perception measures characteristics, such as representation quality, physiological and psychological, was investigated in the aforementioned study by establishing some assumptions, such as rule selection and Gaussian membership selection. In this case, a fuzzy logic system was used to measure the parameters based on fuzzy logic objectively [39].

To determine the impact of various parameters, Yu DU et al. built a hierarchy model. A fuzzy judgment matrix was set up based on a survey and the subsequent statistical analysis. A total weight vector was generated using this matrix and its counterpart matrix. In reality, this overall vector was employed as an indication to determine which aspects have the greatest influence on consumer satisfaction. Gomes et al. used a fuzzy logic approach in Wireless Mesh Networks to improve user happiness and maximize multimedia distribution performance [40].

We came to the conclusion that the Gaussian membership functions may be better appropriate to model the fuzzy logic model for the reasons listed below. Although it would be feasible to enhance the accuracy of the -MF by incorporating additional bell-shaped functions, this is not preferred in practice by service providers due to the associated difficult formula of the fitted function. A Gaussian MF is specified by just two parameters, in contrast to the -membership function. In this work, we have put forth an approach that will allow service providers to choose how to allocate RT intervals to the appropriate quality scales based on the Gaussian MFs they have acquired. In general, appropriate types of membership functions may be used to simulate the fuzzy operating core based on the quality factor and the observed distribution [41].

Recent studies have shown that fuzzy logic and the challenges it addresses serve as an effective foundation for problems like text classification, dimension reduction, extraction of features and retrieval, and similarity analyzers. A kind of logic known as fuzzy logic is said to have been specifically created for the purpose of capturing knowledge and human reasoning in a way that makes them computer process able. Fuzzy sets, linguistic variables, possibility distributions, and fuzzy if-then rules are some of the key ideas in fuzzy logic. A system's degree of uncertainty, or "fuzziness," refers to the reality that nothing can be anticipated with absolute accuracy, according to the Published in journal of Information Retrieval & Knowledge Management Process. Practically speaking, the values of variables are not always exact; rather, it is more likely that approximate values are known [42].

Right side, corridor, internal corner of 90-degree turns, etc. are some examples of typical nodes utilized in this graph. This topological representation is constructed only based on the presence or absence of a barrier in the agent's immediate vicinity and the timing of this identification process; it is unaffected by any GPS locations. This work shows that various sorts of information may be incorporated in the control architecture suggested for advocating activities, rather than describing the properties of this representation method.

These modules can utilize fuzzy logic to suggest actions. The fuzzy rules employed by the External Situation module. An undesirable conduct in these rules is a result that is followed by NOT. Similar methods are used by the Needs module to choose behaviors, but in addition to employing fuzzy states produced from the motivations, it also takes into account external situations. The Cognition module is distinct from the other two recommendation modules in that it does not develop its suggestion using fuzzy logic. The Cognition module is used to create a topological map of the environment in the experiment carried out using the simulated world for mobile robots.

Conclusion

6.1 Conclusion

Over a period, the user’s behavior must remain consistent and a user using features like Display Name and Picture is most likely to use same with same location and device. Thus, this behavior must fall within a specific range. This research normalizes user’s behavior through FIS using behavioral analysis to achieve better trust in individual’s identity by utilizing fuzziness and removing inaccuracies in detection.

With the underlying methods, processes, and procedures, several studies show positive outcomes. The experimental findings offer very accurate fuzzy-based text categorization that is good. These models concentrate on novel difficulties and methods for categorization. The knowledge regarding advanced fuzzy classification, associated models, and approaches is therefore provided by these research papers and their survey. The analytical review combines summary and synthesis to provide a straightforward description of the sources in an organized pattern and offers a fresh perspective of previously published content. As a result, it tries to cover the key elements of the most recent research, including both substantive findings and theoretical and methodological contributions. Additionally, both their parametric data and experimental outcomes are adequately discussed and contrasted separately.

Our framework successfully detects abnormal behavior and thwarts various types of threats against identity in social media.

6.2 Future Work

In future, we intend to expand our capability of detection by including more parameters and expanding this framework to other parts of network for foolproof detection.

Such comparative research and technical analysis charts offer a solid foundation for comprehending the usage of fuzzy and the issues it raises. The models and approaches have been validated by a number of experimental outcomes. Text mining and text categorization benefit greatly from fuzzy logic's application and application domains. As a result, fuzzy similarity is employed for classification across a wide range of application areas and fields worldwide.

6.3 Precautionary Measures

As a useful tool to categories these behaviors, such as changes in operating conditions or process failures, the fuzzy systems developed in this work enabled the securing of surface graphs that facilitate the interpretation of the behavior of the variables and the effects they are having on the process.

Never put personal or financial information on display: Too many, this may seem like a no-brainer, but there are some individuals who get overly thrilled when they acquire a new credit card or driver's license and want the entire world to know about it. As a result, people publish images of themselves exhibiting their accomplishments. This is what most Identity Theft thieves are looking for when stealing identities. If you still want to upload photos of your documentation with your private details on them, make sure to blur out any names or numbers that are printed on them.

Turn off the features that allow you to log in automatically: Allowing social media applications to auto log you in there and allowing WebPages to retain your log in data are both bad ideas. If someone gets their hands on your device, they won't be able to access your account directly. As a result, they won't be able to see your personal information.

Publishing Location Updates Should Be Avoided: When people write about their travels location information on the internet, it gives criminals solid knowledge that they will be out of respective houses at the time of the update, allowing them to break inside and steal goods or, more significantly, identifying credentials that can be used to impersonate them.

Setting Strict Privacy Preferences: Because users should be aware that their personal information, such as their name, photograph, birth date, address, and place of employment, is sensitive data, they should configure their privacy settings so people they trust may see it. This can be done by altering the choices for your personal data in your Face book, Instagram, LinkedIn, and Twitter accounts.

Use of Passwords That Are both Strong and Unique: When establishing an account on a social media site, that was the first and most important step to take. Building positive, safe, and unique passwords, such as alphanumeric with extra characters, will help keep fraudsters at away.

With all these findings we can easily incur that how much identity frauds have spread and almost everybody is prone to this. For this a few precautionary steps can be taken like:

Always make connections with genuine people: While this may not appear to be a problem because, for example, when someone makes you a request on Face book, you can verify that person by looking at his name, profile photo, or other visible attribute, this is not the case. When crooks imitate someone, they do so in this manner. They impersonate other users by using their genuine identities and publicly available photos to try to catch them off guard.

Two factor authentication: This is a feature that some social media platforms have adopted to make it more difficult for criminals to obtain access to a user's account. When users log in for the first time via Twitter, for example, they can enable a setting that requires them to enter a one-time code delivered to their mobile phones.

Use Different Passwords for Different Accounts: People have a habit of using the same password for all of their accounts, primarily because it is easier and easier to remember. When Mark Zuckerberg's LinkedIn password was revealed, he used the same password for all three sites, which resulted in the hijacking of his Pinterest and Twitter accounts. Although it may appear to be a bother to keep separate passwords for each account, it will save you a lot of time and worry in the long run.

Never store credit card information on the internet: Under no circumstances should you store credit card information on social networking sites. Your identity will almost certainly be exploited if information falls into the wrong hands.

Avoid tagging your photos with their location: You don't have to include a location tag with every photo you share on social media sites. It is usually preferable not to divulge the location of a particular photograph because this reveals where you have been. This is critical information for identity thieves because it reveals a pattern of where they have been.

Use of Protection Services: Identity Guard and Life Lock are two services that can help you protect your social media accounts from identity theft. These services are the way to go if one feels they require expert services to protect their accounts.

Activating Unusual Activity Alerts: When someone else tries to log into your account from a different device or maybe from a different location, Face book and other applications provide you the choice to set alerts. To prevent unauthorized users from entering into your accounts, it's always a good idea to enable these notifications.

References

1. Hoar, S. B. (2001). Identity theft: The crime of the new millennium. *Or. L. Rev.*, *80*, 1423.
2. Newman, G. R., & McNally, M. M. (2005). Identity theft literature review.
3. Ganesh, L., & Zhao, B. Y. (2005, November). Identity theft protection in structured overlays. In *1st IEEE ICNP Workshop on Secure Network Protocols, 2005.(NPSec)*. (pp. 49-54). IEEE.
4. Marshall, A. M., & Tompsett, B. C. (2005). Identity theft in an online world. *Computer Law & Security Review*, *21*(2), 128-137.
5. Bilge, L., Strufe, T., (2009,). All your contacts are belong to us: automated identity theft attacks on social networks. In *Proceedings of the 18th international conference on World wide web* (pp. 551-560).
6. Roberts, L. D., Indermaur, D., & Spiranovic, C. (2013) Fear of cyber-identity theft and related fraudulent activity. *Psychiatry, Psychology and Law*, *20*(3), 315-328.
7. M. Fita, November 2012. Available: https://www.brandignity.com/2012/11/6-reasons-whysocial-networking-is-so-popular-these-days/.
8. Lewis, K. (2016). How social media networks facilitate identity theft and fraud. *Entrepreneurs' Organization, accessed on*, *15*.
9. Khan, Z. R., Rakhman, S., & Bangera, A. (2017). Who Stole Me? Identity Theft on Social Media in the UAE. *Journal of Management and Marketing Review (JMMR)*, *2*(1), 79-86.
10. Irshad, S., & Soomro, T. R. (2018). Identity theft and social media. *International Journal of Computer Science and Network Security*, *18*(1), 43-55.
11. Mitts, J., & Talley, E. (2019) Informed trading and cybersecurity breaches. *Harv. Bus. L. Rev.*, *9*, 1.
12. P. Kallas, June 2017. https://www.dreamgrow.com/top-15-most-popular-socialnetworking-sites/.

13. Kolaczek, G. (2009, April). An approach to identity theft detection using social network analysis. In *2009 First Asian Conference on Intelligent Information and Database Systems* (pp. 78-81). IEEE.
14. Holm, E. (2014). Social networking and identity theft in the digital society. *The International Journal on Advances in Life Sciences*, *6*(3&4), 157-166.
15. Bendiek, A., & Porter, A. L. (2013). European cyber security policy within a global multistakeholder structure. *European Foreign Affairs Review*, *18*(2).
16. Fogel, J., & Nehmad, E. (2009). Internet social network communities: Risk taking, trust, and privacy concerns. *Computers in human behavior*, *25*(1), 153-160.
17. Khandpur, R. P., Ji, T., Jan, S., Wang, G., Lu, C. T., & Ramakrishnan, N. (2017, November). Crowdsourcing cybersecurity: Cyber attack detection using social media. In *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management* (pp. 1049-1057).
18. Bleau, H. (2016). 2017 Global Fraud and Cybercrime Forecast. *https://www. rsa. com/en-us/blog/2016-12/2017-global-fraud-cybercrime-forecast, last visit*, *9*, 2017.
19. Ribeiro, F. N., Henrique, L., Benevenuto, F., Chakraborty, A., Kulshrestha, J., Babaei, M., & Gummadi, K. P. (2018, June). Media bias monitor: Quantifying biases of social media news outlets at large-scale. In *Twelfth international AAAI conference on web and social media*.
20. Krombholz, K., Merkl, D., & Weippl, E. (2012). Fake identities in social media: A case study on the sustainability of the Facebook business model. *Journal of Service Science Research*, *4*(2), 175-212.
21. Stutzman, F., & Kramer-Duffield, J. (2010, April). Friends only: examining a privacy-enhancing behavior in facebook. In *Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 1553-1562).
22. Saini, N., Sangwan, G., Verma, M., Kohli, A., Kaur, M., & Lakshmi, P. V. M. (2020). Effect of social networking sites on the quality of life of college students: a cross-sectional study from a city in north India. *The Scientific World Journal, 2020*.
23. Gao, H., Hu, J., Huang, T., Wang, J., & Chen, Y. (2011). Security issues in online social networks. *IEEE Internet Computing*, *15*(4), 56-63.

24. Marcum, C. D., & Higgins, G. E. (2019). Cybercrime. In *Handbook on crime and deviance* (pp. 459-475). Springer, Cham.
25. Al-Daraiseh, A. A., Al-Joudi, A. S., Al-Gahtani, H. B., & Al-Qahtani, M. S. (2014). Social networks' benefits, privacy, and identity theft: KSA case study. *Social Networks*, *5*(12), 129-143.
26. Bali, T. G., Hirshleifer, D., Peng, L., & Tang, Y. (2021). *Attention, social interaction, and investor attraction to lottery stocks* (No. w29543). National Bureau of Economic Research.
27. Allison, S. F. (2003). A case study of identity theft.
28. Shah, M. H., Ahmed, J., & Soomro, Z. A. (2016). Investigating the Identity Theft Prevention Strategies in M-Commerce. *International Association for Development of the Information Society*.
29. OWASP, https://www.owasp.org/index.php/Main Page, (Accessed on 20, February 2018).
30. Jose´ A´. Concepcio´n-Sa´nchez Et al. "Fuzzy Logic System for Identity Theft Detection in Social Networks" in 2018 4th International Conference on Big Data Innovations and Applications.
31. Aleksei Romanov, Alexander Semenov, Oleksiy Mazhelis and Jari Veijalainen "Detection of Fake Profiles in Social Media".
32. M. Tsikerdekis and S. Zeadally (May 22, 2015). Detecting and Preventing Online Identity Deception in Social Networking Services in Internet Computing, vol. 19, no. 3, pp. 41-49.
33. "B. Z. He, C. M. Chen, Y. P. Su and H. M. Sun (2014). A defence scheme against identity theft attack based on multiple social networks. Expert Systems with Applications."
34. Esma et al. - The ultimate invasion of privacy: Identity theft.
35. Jiyoung Woo "An automatic and proactive identity theft detection model in MMORPGs"
36. Zoran et al. 2018 - Analyzing of e-commerce user behavior to detect identity theft
37. Beheshti, H. M., & Lollar, J. G. (2008). Fuzzy logic and performance evaluation: discussion and application. *International Journal of Productivity and Performance Management.*

38. KR, M. S., & Raorane, M. S. Performance Evaluation of Management Faculties from ‘Students Feedback Performa’using Fuzzy Logic Model.

39. Hamam, A., Eid, M., Saddik, A. E., & Georganas, N. D. (2008, June). A fuzzy logic system for evaluating quality of experience of haptic-based applications. In *International Conference on Human Haptic Sensing and Touch Enabled Computer Applications* (pp. 129-138).

40. Escobar, L. M., Aguilar, J., Garcés-Jiménez, A., De Mesa, J. A. G., & Gomez-Pulido, J. M. (2020). Advanced fuzzy-logic-based context-driven control for HVAC management systems in buildings. *IEEE Access*, *8*, 16111-16126.

41. Jang, J. S., & Sun, C. T. (2009). Neuro-fuzzy modeling and control. *Proceedings of the IEEE*, *83*(3), 378-406.