AMITY UNIVERSITY ONLINE, NOIDA, UTTAR PRADESH

In partial fulfilment of the requirement for the award of degree of **Bachelor of Computer Applications (BCA) (Discipline – IT)**

Phishing detection system using rule-based analysis and api integration

Guide Det:

Name: Mr. Varun Anand

Designation: Senior software Developer

Submitted By:
Name of the Student- Anmol Verma

Enrolment. No: A9922523001(el)

(Times New Roman-14)

ABSTRACT

Recently, phishing attacks have appeared to be one of the most common and harmful cyber threats posing significant risks in the current digital era. The prime reason for these attacks is that they exploit human trust instead of technology, and these are very hard to detect with the help of traditional security solutions. The core aim of these phishing attacks is to trick users' trusts through genuine emails or messages to obtain their valuable data like login credentials, finance details, or personal details. As people's reliance on digital communication media keeps growing, it is highly necessary to develop an efficient, reliable, and explainable approach to detect these phishing attacks to help users avoid harm before it's too late.

The aim of the project would be to develop and build a Phishing Detection System Using the Rule-Based Analysis Technique that would be able to analyze the text message as well as the attached URL in order to estimate the risk level likely associated with the message. Contrary to the majority of previous approaches that are reliant upon the known malicious URL sources in the context of blacklisting, the new system would be centered upon the detection of the behavioral and text message components likely attacked in the majority of phishing attempts. The system would be utilizing the rule-based engine that would analyze the text message for its components like the use of urgent/threatening communications, the request for confidential data, and the utilization of fishy or short URLS.

Along with the rule analysis, the application also incorporates outside threat intelligence offered through the VirusTotal Application Programming Interface (API). The extracted URLs from the input text are used to create a request to the VirusTotal service, which provides a consolidated view from various security vendors to evaluate the reputation and malicious nature of the extracted URLs. The result derived from the two components is combined to form a consolidated result classified by the message into either the LOW, MEDIUM, or HIGH risk level.

The system also computes the confidence score to indicate the accuracy of the result obtained. In attempt to enhance transparency and build trust for users, the application has an explainability feature that explains the reasoning behind any given classification. This is achieved through listing out active rules, indicators, and VirusTotal results, hence making this system useful for learning and cybersecurity demo sessions.

In this context, the phishing detector software is written in the general-purpose computer programming language, Python, which acts as the backbone for the back end of the webpage, facilitated by the Flask software framework. The user interface part of the software is written in the HTML, CSS, and JavaScript languages, which are more modern, interactive, and user-friendly. The key components of the user interface part of the software include the center-aligned input box, keyboard entry, loading indicator, circular risk indicator, severity badges, confidence bar, explanation section, and copyable analysis results section. The software runs on a localhost setup, meaning that it is simple and meets the necessary guidelines for being an academically acceptable assignment.

The best practices regarding security have been implemented every step of the way throughout development. The VirusTotal API key is handled securely through the use of environment variables, where sensitive config files are also ignored in version control repos. The code base is modular and organized.

The system has been tested with different samples of legitimate as well as phishing messages. Results show that the proposed solution is able to identify phishing attacks effectively with a minimum number of false positives. The project ends with achievement of its objectives by providing a useful, interpretable, and secure solution for phishing attacks that can be used in the learning process.

This particular project works as an exemplar of the ability that can result when rule-based analysis results in conjunction with external threat intelligence in combating the rising threat of phishing. This particular developed solution works as an excellent learning aid for all students in the realm of information technology, as it can provide a foundation for the evolution of various upgradations in the future.

DECLARATION

I, Anmol Verma, a student pursuing Bachelor of Computer Applications (BCA) in Semester 6 at Amity University Online, hereby declare that the project work entitled "Phishing Detection System Using Rule-Based Analysis and API integration" has been prepared by me during the academic year 3rd under the guidance of Mr. Varun Anand, Software Development, MCA – Guru Gobind Singh Indraprastha University; B.Sc. – University of Delhi. I assert that this project is a piece of original bona-fide work done by me. It is the outcome of my own effort and that it has not been submitted to any other university for the award of any degree.

Signature of Student

CERTIFICATE

This is to certify that Anmol Verma of Amity University Online has carried out the project work presented in this project report entitled "Phishing detection system using rule-based analysis and API integration" for the award of Bachelor of Computer Applications (BCA) specialized in Cloud and Security under my guidance. The project report embodies results of original work, and studies are carried out by the student himself. Certified further, that to the best of my knowledge the work reported herein does not form the basis for the award of any other degree to the candidate or to anybody else from this or any other University/Institution.

Signature

Mr. Varun Anand

Senior software Developer

TABLE OF CONTENTS

- ABSTRACT
- DECLARATION
- CERTIFICATE
- TABLE OF CONTENTS
- LIST OF TABLES
- LIST OF FIGURES
- Chapter 1: Introduction to the Topic
- Chapter 2: Review of Literature
- Chapter 3: Research Objectives and Methodology

- Chapter 4: Data Analysis, Results, and Interpretation

- Chapter 5: Findings and Conclusion

- Chapter 6: Recommendations and Limitations of the Study

- Bibliography / References

- Appendix

.

LIST OF TABLES

Table 4.1

Table 4.2

LIST OF FIGURES

Figure 3.1: System Architecture of the Phishing Detection System

Figure 3.2: User Interaction Flow Diagram

Figure 4.1: Rule-Based Risk Scoring Output

Figure 4.2: High-Risk Phishing Detection Result

Figure 4.3: Low-Risk Legitimate Message Detection Result

CHAPTER 1: INTRODUCTION TO THE TOPIC

1.1 Introduction

The rapid development of information technology and the increased online usage of internet services have completely changed the method of human communication, business transactions, and information exchange. Email services, messengers, online banking services, and cloud services have become an indispensable part of our day-to-day life. Moreover, the increased usage of technology has resulted in an increased risk of security breaches. Among the major security breaches caused due to increased technology usage is phishing.

Phishing Attack: This is another type of cyber assault whereby attackers try to trick victims into revealing sensitive data like usernames, passwords, financial information, or identification data using the guise of trusted sources. This type of cyber assault usually comes in formats like messages or sites that look like their genuine counterparts. Since their nature depends entirely upon psychological manipulation and has little to do with technical loopholes, this type of cyber assault has remained prominent despite having technical preventive measures in place.

1.2 Background of Phishing Attacks

Phishing attacks have seen a tremendous transformation ever since they first came into being as cyber threats. Initially, phishing attacks used to be quite simple and quite easily recognizable by users and security software alike. These attacks used to contain grammatical and spelling errors and generic greetings and suspicious links that are quite characteristic of phishing attacks and are recognized by cautious users as a security threat. As such, they used to have a low rate of success.

Additionally, however, phishing attacks have evolved in sophistication over time. Today, phishing attacks are designed and structured to closely resemble legitimate communications made by reputable bodies like financial institutions, governments, cloud service providers, and major online platforms. Attackers have incorporated formal message writing, optimized email templates, and even replicates of genuine websites to make an attack appear authentic. The employment of targeted attacks by phishers, referred to by the name spear phishing, enables phishers to customize an attack message based on personal or firm information derived from various sources like social platforms or compromised databases. This makes it highly likely that any user will respond favorably to an attack message.

In addition to an improvement in content quality, attackers began using sophisticated social engineering tactics to deceive users. Phishing scams, for instance, use urgency, fear, and curiosity to influence users by warning them about account suspension, irregular operations, and limited-time activities. Such tactics limit the chances of users double-checking whether it is genuine before taking any action. This, in turn, makes technically informed users susceptible to intelligent phishing scams.

The impacts of successful phishing can be very damaging and extensive. Looking at the personal impacts of phishing, one can suffer loss of money, identity theft, unauthorized access to personal accounts, and breach of confidential information. Looking at the business

impacts of phishing attacks, one can suffer exposure of confidential business information, business disruptions, reputational and brand impacts, and also face the risk of litigation. Furthermore, the impacts can also be used as entry points for larger attacks like malwares and advanced persistent threats. It is observed in various reports on cybersecurity threats that phishing is still among the main reasons for cybersecurity breaches globally. Its success can be explained by the fact that it requires only low costs for implementation and is capable of causing large-scale severe impacts. Its growing sophistication reveals the immense importance of developing effective and reliable detection and protection solutions against this threat in the current cyber world.

1.3 Need for Phishing Detection Systems

In recent times, with increased use of digital communication, phishing attacks have been recognized as one of the most persistent cyber threats. In spite of the availability of traditional safety tools such as spam filters, firewalls, as well as antivirus protection, phishing attacks have been very successful. One of the reasons why phishing attacks remain very successful is that they do not take advantage of technological weaknesses but instead take advantage of human trust. In effect, traditional safety tools, such as products that rely on detecting viruses or attacks carried out over networks, often lack effectiveness when dealing with phishing attacks, especially when they involve new designs.

Conventional phishing prevention systems commonly employ heavily blacklist-driven detection systems. Such systems work on the basis of comparing the receiving emails, messages, or URLs with blacklists of identified malicious domains or sources. Though blacklist-driven systems can be quite helpful in the prevention of phishing attack sites, they prove ineffective in the case of zero-day phishing attacks that communicate using newly created domains or sometimes employ domains that have a short lifespan of only a few days. Such domains and URLs change continuously to avoid being detected by the blacklist systems. Therefore, users continue to be threatened in the context of phishing that has yet to be identified in the blacklist or indexed in the blacklist database.

In recent times, machine learning and artificial intelligence-based solutions have emerged as popular phishing filters since they have the capability of processing a large amount of data and decoding complex patterns associated with it. Machine learning-based solutions have the capability of providing high accuracy results by leaning upon past phishing data sources and deriving relevant characteristics associated with message content, URL patterns, and user behavior patterns. However, despite such benefits, machine learning-based solutions have been associated with various challenges too. Machine learning-based solutions require an ample amount of data with accurate labels, along with extensive computing capabilities, which at various instances have not been possible in academic settings. Additionally, various machine learning-based solutions have been classified as black box solutions since they lack transparency concerning classification results.

The unexplainable aspects of black-box models may result in decreased trust and usage of such systems, especially in learning and academic environments. This could happen because, as soon as the reasons behind the identification of a message as phishing or legit are not clear to the user, they may find it difficult to rely on the suggestions provided by this system. As far as learning systems are concerned, they are of very limited use in terms of learning the fundamentals of phishing identification.

Thus, a critical and increasing demand exists for developing phishing detection tools that are not only more accurate but also transparent, interpretable, and educationally effective. The most desirable phishing detection method would be one that enables users to interpret the reasons for their detection outcome with a reasonable level of detection accuracy. The rule-based phishing detection method fits this purpose with a set of predefined rules that encapsulate typical phishing features, including urgency, request for credentials, suspicious URLs, and impersonation. Every rule is associated with a known phishing template so that the logic of phishing detection is interpretable.

Rule-based detection systems can be especially beneficial in an academic or learning setup, where they can increase the knowledge level and awareness of students regarding the

workings and detection of phishing attacks. Furthermore, these systems can also benefit in combination with an external threat service, in which case they can stay updated regarding security-related real-time information while remaining explainable. This provides an efficient means for the balance between explainability and effectiveness for phishing detection systems.

Conclusion

The sophistication of phishing attacks as well as the limitations in traditional security software and the transparency of Machine Learning algorithms emphasize the need for effective phishing detection software. The phishing detection algorithm based on rule-based software with the aid of external threat intelligence serves as a pragmatic solution to the challenges posed by phishing in the academically relevant context.

1.4 Rule-Based Analysis in Phishing Detection

Rule-based analysis is one of the earliest and most widely used techniques in the phishing analysis and cybersecurity domain. The technique uses a predefined set of rules to spot unusual or fishy activity that is largely associated with phishing attacks. The rule set is constructed based on known phishing behaviors, attacker habits, and characteristics exhibited by phishing mails. Rule-based analysis does not use a probability model as rule-based analysis uses a deterministic model, where all decisions are based on predefined conditions and checks.

Textual analysis in rule-based classification in the context of phishing detection entails the scrutiny of messages based on linguistic attributes. The aim of textual analysis in this context is the identification of potentially phishing linguistic patterns in a message. Such patterns include threats of dire consequences in the event the recipient takes a certain course of action. Such linguistic patterns are often used by phishing parties as a psychological weapon against the targeted victims. Structural attributes in the context of phishing detection involve the scrutiny of a message in respect to attributes such as URLs. Mismatched hyperlinks are one of the attributes considered in the classification of phishing messages.

Every rule in a rule-based system has been assigned a unique weight or score, depending on its severity and phishing relevance. A request for sensitive material such as credentials/passwds has been found more severe in terms of phishing relevance compared to urgent messages. When the system processes a message, it examines each rule independently by calculating a total risk score. A message that scores above a certain threshold is marked for phishing potential depending on its risk category, which can either be LOW, MEDIUM, or HIGH. This allows the system to eliminate messages having multiple phishing characteristics more easily compared to those having single phishing markers.

Rule-based analysis also has the benefit of explainability. This means that while many phishing-filtering tools rely on artificial intelligence, which function as black boxes, explaining how they make decisions, rule-based analysis can easily explain their decisions. Users can easily see which rules are used in the decision and how these rules affect the outcome. Explainability is an important requirement in cybersecurity, where users need an explanation regarding whether an e-mail message is harmful, and if so, how the decision was made so that they can make the necessary corrections.

Rule-based analysis also has some benefits when it comes to the ease of implementation and control. With rule-based analysis, the detection logic can easily be defined by rules, which can be easily altered, removed, or supplemented by developers and security professionals depending on the new phishing attack techniques. Rule-based systems also consume less computational capacity compared to machine learning systems, which makes them ideal for execution in light-weight scenarios.

In academic terms, rule-based phishing analysis systems can be considered very helpful. These systems make students understand phishing attacks and their detection processes very clearly. If students work with rule-based phishing analysis, they can see, first-hand, how phishing characteristics interact with analysis outcomes. In academic studies of cybersecurity,

social engineering, and secure system design, such direct experience is very helpful in retaining theoretical concepts. Because of their transparency, rule-based phishing analysis systems can be considered very helpful in academic assignments, research work, where logic accuracy is of utmost importance. To conclude, rule analysis methods remain a basic approach to phishing protection that, despite simplicity and effectiveness, can be used together with external sources of threat intelligence to remain an important component of any hybrid phishing protection solution, allowing it to cover any gaps that it might have in terms of complex patterns that have not been seen before. As it stands now, rule analysis methods can be an important addition to phishing protection in an educational environment.

1.5 Role of API-Based Threat Intelligence

For the purpose of increased accuracy and effectiveness of phishing detectors, the role of threat intelligence feeds cannot be underestimated. This intelligence consists of the latest information regarding the reputation, activities, and harmfulness of URLs, domains, files, and other cyber objects. These threat intelligence feeds acquire data on a constant basis from various sources, which can be highly beneficial for the determination of unknown threats in the phishing process. In the context of phishing protection, this threat intelligence can prove beneficial for confirming whether a suspicious resource has been used for malicious activities or not.

One of the most popular platforms for threat analysis is VirusTotal, which provides a consolidated outcome of analysis carried out by a large number of free, as well as paid, antivirus engines. The phishing detection system can make use of the capabilities offered by the Application Programming Interface (API) feature in VirusTotal for a comprehensive analysis of the safety level of the extracted URLs. Details such as the number of engines reporting a particular website as malicious, suspicious, and safe can hence be obtained, which would otherwise not be possible by this system.

Integration of threat intelligence through an API makes it possible for phishing detection solutions to implement a hybrid method of detection, which involves both local rule-based evaluation as well as verification using an external API. Although the former is concerned with the evaluation of the content of the message as well as its behavior, the latter checks the validity of the results obtained against global threat feeds. Moreover, the integration of the API prevents the need for periodic manual updating by ensuring the latest information is used.

In totality, API-based threat intelligence solutions improve phishing detection capabilities by allowing real-time verification, enhancement of accuracy, and expansion of the threat landscape. These services in combination with rule-based analysis provide a more secure, dynamic, and credible detection mechanism suitable for research, proof-of-concept projects, and learning initiatives in the domain of cybersecurity.

1.6 Project Overview

The proposed system takes input from the user in the form of message text, for instance, emails and notification messages. The system processes this input with a rule-based detection engine that checks for defined phishing cues within the message text. The defined phishing cues are urgent and/or threat language, requests for sensitive and/or confidential information, impersonation cues, and other communication cues. The engine assigns a cumulative risk score to each cue identified in accordance with its level of severity. This is because the engine is designed to enable systematic processing while being easily understandable by making the detection logic transparent and interpretable. This is unlike other complex black-box models where users cannot easily comprehend why their messages have been identified to be either phishing and/or authentic.

Apart from the analysis of the text, the system also greatly aids in improving the accuracy of detection of the URLs hidden in the message text using an external API-based threat analysis

service. Phishing campaigns may employ malicious or deceptive URLs that can redirect the victim to a phishing site. By incorporating an external threat analysis API service in the system, the system has the ability to check the malicious links against the trusted global security databases that compile the analysis of various security companies. As a result of the on-demand analysis of the links in near real-time, the system gains important information pertaining to the reliability and maliciousness of the links and can detect potential threats that cannot be determined by mere analysis of the text messages.

The results from the rule engine and API analysis of the URL are combined to create the final risk result. Using the total score and threat intelligence report, the messages are labeled as LOW, MEDIUM, or HIGH risk categories based on their threat potential. In addition, the system also provides the confidence score of every result in determining the classification. To boost transparency, there is an explanation mechanism in the system that outlines the rules applied in the analysis and the variables that contribute to the final classification result. Such a system is very ideal in an academic setting or learning system since transparency is increased by explaining how the classification is done. From the technological standpoint, the app is developed in Python as the primary programming language, while Flask framework acts as the backend. Flask handles requests coming from users, deals with the input data, communicates with detection tools, and provides analysis results. The GUI part of the app is built by utilizing HTML, CSS, and JavaScript technologies, which provide an appealing, responsive, and intuitive GUI. These features as GUI elements, such as the centered input field, loading signs, graphical risk illustration, and result description, can improve usability and readability. Moreover, the system can work in the localhost environment, making it suitable for demonstration purposes in an academic setting. Overall, this project reflects an excellent application and teaching tool for phishing detection that incorporates simplicity, interpretability, and efficacy. Not only does this project satisfy academic standards but also has proven to be an ideal teaching resource for students studying information technologies and cybersecurity. Its modularity and combination of detection approaches are excellent avenues for advancements in machine learning approaches or cloud migration.

1.7 Justification for Selection of the Topic

This makes the topic "Phishing Detection System Using Rule-Based Analysis and API Integration" very relevant and handy in today's digital ecosystem, which is highly focused on cloud technology. Thus, phishing has remained one of the most widely used attack vectors by cybercriminals owing to its low cost and high return, and attacking the human link requires less effort rather than trying hard against technical flaws in a system. Phishing, attacking cloud-based platforms, would increasingly become recurrent and sophisticated, as most organizations and individuals are relying on their cloud services to communicate, store data, and perform financial transactions. The awareness and mitigation of such phishing threats have become a need of the hour in the careers of professionals working in the field of information technology and cybersecurity.

The chosen topic directly relates to the specialization in Cloud and Security because, generally, phishing attacks target cloud-based services like email systems, online storage systems, SaaS applications, or identity management services. These attackers often use identities similar to those of trusted cloud service providers to make users divulge login information or sensitive information for subsequent unauthorized access, data breaches, and large-scale security incidents. In that respect, the focus on phishing detection in this project meets a very valid realistic security challenge highly applicable in cloud computing environments and modern enterprise infrastructures.

Another significant justification for choosing this topic is its potential for connecting theoretical notions regarding cybersecurity to real life. The development of a rule-based phishing detection system gives an opportunity to implement core security principles like threat identification, risk assessment, secure coding practices, and ethical handling of sensitive data in real life. Unlike purely theoretical studies, this one underlines practical learning with designing and implementing a real detection system capable of content analysis and embedded URL evaluation for phishing signs.

Inclusion of API-based threat intelligence further strengthens the practical relevance of the project. Integration of external security intelligence services allows the system to make use of real-time global threat data for improved detection accuracy. This reflects a very normal situation in the industry where security solutions often depend on shared threat intelligence to be able to respond effectively to changing attack techniques. Working with APIs introduces students to a set of real-world system integration skills such as secure credential management, data handling, and dependency management that are key for any cybersecurity professional.

From an academic point of view, the project shows critical problem-solving skills, analytical reasoning, and technical capability. The choice of a rule-based approach means it will be transparent and interpretable, and thus the logic of detection can be understood and reasoned upon. This is really important in academia because the ability to explain the behavior of a system and justify such a design on a theoretical basis is crucial. Also, the ethical and security best practices are followed in this project since any hard-coding of sensitive information was avoided, and it handled the data in privacy-conscious ways.

The topic chosen also leaves great room for enhancement and further research. It can be extended by the addition of machine learning techniques to handle better adaptability, test the scalability of the application in a cloud environment, or develop real-time monitoring and alerting mechanisms. Thus, it proves not only suitable for current academic needs but also as a very good starting point for advanced research or professional development concerning computer security. In the end, the relevance of this topic to current challenges in cybersecurity, fully aligning with the Cloud and Security specialization, and as such, a combination of theoretical knowledge with practical application involves substantial meaning for the academic learning and at the same time prepares the researcher for a real-world cybersecurity role.

1.8 Summary of the Chapter

This chapter has marked the beginning of an in-depth exploration of the concept of phishing attacks and has now laid down a robust conceptual background for the research work being done in this thesis. The chapter started off with defining phishing and described it to be an ongoing and famous cyber attack that is more reliant on human trust and social engineering mechanisms than on technical hacking methods and vulnerability. Moving ahead in this context, it has been described in this chapter that human-centric cyber attacks like phishing are still very efficient and effective despite advanced conventional security infrastructure in place. The chapter has further continued to describe the timeline of phishing attacks that earlier started off in very simplistic and detectable forms but now are rising up to be more advanced and hard to detect.

The chapter further explained the different effects that phishing attacks can generate for individuals and organizations. Some of the most significant risks associated with phishing attacks included financial losses, identity theft, unauthorized access to important sensitive accounts, data breach, and reputation damage. The explanation of the different risks associated with phishing attacks # highlight the importance that has to be attached to the development of mechanisms that can be used to detect and prevent phishing attacks in the current digitally connected world.

In addition, the chapter pointed out that there is an ever-increasing demand for effective and efficient anti-phishing platforms that are able to adapt to new and evolving attack methods. At the same time, the challenges of using advanced anti-phishing platforms were also identified in the chapter, such as the blacklist method that is unable to identify newly created phishing sites and complex machine learning methods that are often less interpretable and transparent. In this regard, it is important to note that anti-phishing platforms that are not only interpretable but also highly accurate and suitable for academic purposes were identified as ideal. In this discussion, the use of rule-based analysis is identified as an interpretable method that enables users to interpret how decisions are made.

Furthermore, the chapter clarified the importance of the integration of API-based threat intelligence services to improve the accuracy levels of phishing detection. Validating

malicious URLs against third-party security intelligence sources gives real-time notifications on malicious URLs, thereby improving accuracy levels. The use of rule analysis coupled with threat intelligence tools, on the other hand, was presented as an important approach to compensate for the limitations associated with solitary detection methods. This chapter also gave an introduction to the proposed system for phishing detection, discussing its need, operations, concepts, and technologies. The need for choosing the topic was explained appropriately based on its relevance to the domain of information technology, its application to the area of cyber security, and the specific domain of Cloud and Security. The importance of the research work from an academic perspective as well as its application as an educational platform was also highlighted. This chapter laid a strong foundation related to the topic from a theoretical perspective and paved the way for the upcoming chapters. The following chapter describes the literature study related to phishing detection techniques.

CHAPTER 2. REVIEW OF LITERATURE

2.1 Introduction

Literature review forms the backbone of research since it offers an in-depth knowledge of what has been known, researched, and determined concerning the topic of study. Literature review assists researchers in studying the way in which the researched topic has been dealt with in the past and the various techniques used to verify the research conducted. In the context of cyber security, literature review assumes significance owing to the dynamism of cyber threats and the increasing frequency of new attack methods.

Phishing detection has emerged as an area of extensive research in the field of cybersecurity, mainly attributed to the widespread increase in the number of phishing incidents on different online platforms. Phishing attacks rely on human trust and social engineering attacks rather than exploiting vulnerabilities, making these attacks quite difficult to detect and mitigate. With the increasing use of the internet and dependence on different internet communication services, such as emails, messaging systems, and cloud services, the number of phishing attacks has also witnessed an escalation in recent years. This has urged researchers and cybersecurity experts to work on different detection systems for mitigating the effects of phishing attacks.

Over the past few years, various methods have been developed to address the phishing problem, starting from conventional methods based on blacklists to sophisticated machine learning-based solutions. In the initial phases of research, attempts have been made to detect phishing activities using known malicious URLs and heuristics. However, with the increase in sophisticated phishing methods, these conventional methods have been proven inadequate in handling sophisticated zero-day phishing attacks as well as highly personalized phishing messages. Hence, various sophisticated methods, such as heuristic methods, rule-based methods, machine learning methods, and hybrid methods, have been examined.

In this chapter, previous research studies, detection methods, as well as frameworks proposed by various individuals in the literature regarding phishing detection, will be critically reviewed. The merits as well as demerits of various proposed frameworks, such as their accuracy, interpretability, computational complexity, among others, shall be considered. Additionally, the challenges involved in making complex phishing detection systems within academic settings shall be closely considered. The main purpose of this literature study is to point out the shortcomings in the current methods of phishing detection and to address the need to develop an interpretable, effective, and secure method of phishing detection. Through

this study, this chapter marks the beginning of this research proposal on the necessity of using a rule-based phishing detection system that combines external intelligence to tackle the problem of phishing attacks. The study conducted in this chapter helps to cite this research proposal in the wider area of cybersecurity studies.

2.2 Concept of Phishing and Its Characteristics

Some scholars have considered phishing a type of deceptive cyber attack, which is intended to get confidential user data through the disguise of a legitimate and trustworthy party. These attacks usually involve confidential information such as user names, password credentials, financial data, and personal identification details. The early stages of academic research work on phishing attacks targeted the identification of such attacks through obvious and easily identifiable signs such as deceptive URLs, spoofed sender identities, generic salutations, and poor grammar and spelling patterns in emails. During this phase, phishing attacks lacked sophistication and were more easily identifiable by users, as well as simple security mechanisms.

Nonetheless, over the years, researchers have noted that there has been substantial growth in the sophistication of these phishing attacks. Nowadays, more and more phishing attacks involve highly sophisticated social engineering tactics. Such attacks involve more than just the technical vulnerability; they take into account the expectations of the target and may involve sending messages that look like ones from reputable sources, for instance, banks or government agencies. Due to this realism, these messages have become hard to distinguish from genuine ones.

The existing literature points out that there are several typical characteristics that are usually invariably found in phishing emails. These include urgent or threatening language purported to intimidate users into taking immediate actions, fear tactics involving security or losses, as well as direct demands for secret or sensitive information. Moreover, phishing emails usually involve misleading URLs aiming to direct users to deceitful websites modeled after their legitimate counterparts. These characteristic features have stood as the building blocks for developing several approaches used for phishing identification, especially rule and heuristic approaches that utilize pre-defined patterns and elements for pinpointing possible malicious components within both legitimate and inappropriate content types.

2.3 Blacklist-Based Phishing Detection Techniques

Blacklist phishing protection: This method is among the oldest strategies that have been employed to detect phishing sites and malicious websites. This method involves the use of predefined blacklists where the database contains the details of the detected phishing websites. On receiving a link, the system checks for its presence in the database, and if found, the link is declared to be malicious, thereby taking the next course of action to block that particular link. This method can be implemented easily, taking less time, due to its simple process, which requires less computation.

Blacklist-based detection methods have also been shown by various researchers to be useful in detecting known phishing URLs that have previously been submitted to security organizations or individuals and confirmed as phishing categories by such organizations or individuals. Blacklists are typically implemented in browsers, anti-phishing tools, anti-virus software, and other security tools as the first layer of defense against phishing attacks. They are fast detection mechanisms whose speed can be useful in quickly processing large amounts of data from the Internet.

Despite these benefits, various researchers have noted considerable limitations associated with blacklist methods. Without a doubt, the most problematic limitation is that these methods fail to identify zero-day phishing attacks. This is because phishing pages tend to have temporary lifetimes, as attackers usually register This is because newly created phishing URLs may bypass systems that utilize blacklist methods due to the possibility of considerable delays in the updates of the blacklists. Additionally, these methods also tend to be rendered ineffective due to manual submissions, as they tend to be too time-consuming to be useful for identifying evolving phishing pages.

2.4 Heuristic and Rule-Based Phishing Detection Approaches

Rule-based and heuristic methods in phishing can filter messages based on a set of pre-defined rules that are obtained by analyzing typical behavior patterns of phishing attacks and linguistic characteristics in malicious emails. Common factors considered by rule-based methods include use of urgent or threatening phrases, requests for financial or personal information like passwords or credit information, malicious URL formatting, inconsistencies or misleading domain names, or unusual message formatting or language. The rule-based system uses a set of factors to rate its likelihood of being malicious based on specified phrases or characteristics related to a message.

One of the important points stressed by researchers about rule-based approaches has been transparency and interpretability. While machine learning-based approaches are often opaque

in the sense that they are difficult to interpret, rule-based approaches are transparent in that they explain how they arrive at certain conclusions by pinning down the rules invoked in the process. This particular advantage serves them in good stead in environments where the explanation of the detection process holds as much value as the detection result. They are easily modifiable as per the changing phishing attacks.

Some research works have demonstrated that rule-based phishing detectors work effectively as long as the phishing rules are properly designed and given corresponding weights. Nevertheless, the accuracy levels that can be achieved by such phishing detectors depend to a great extent on the effectiveness, relevance, and exhaustiveness of the phishing rules that have been developed. Irrelevant or obsolete phishing rules might result in either generating true positives, which refer to emails that are flagged as phishing even when they are not so, or true negatives, which refer to phishing emails that evade detection.

2.5 Machine Learning-Based Phishing Detection

Recently, the use of machine learning algorithms to detect phishing has been researched by the community. These techniques rely on the process of training a classifier using a data set comprising legitimate and phishing messages or URLs. Features like lexical patterns, URL formation, email headers, and other content-related features are extracted and used to train the classifiers. Decision trees, support vector machines, naïve Bayes classifiers, and neural networks are the most common algorithms used to detect phishing, and they have shown exciting results with respect to accuracy.

There have been some instances of research suggesting that machine learning algorithms can adapt themselves to learn and change according to shifting patterns of phishing attacks better and faster compared to rule and blacklist systems. Machine learning systems have the potential of learning from existing data and can discover complex patterns and correlations, which may not be defined through rules and blacklists. Machine learning systems can learn and discover sophisticated phishing attacks that can be similar to actual messages.

Nevertheless, there are limitations associated with phishing detection methods that make use of machine learning approaches, as indicated in the existing body of knowledge. The first challenge associated with these methods is the need for large labeled data, which, in most cases, may not exist. Another important challenge associated with these approaches is their need for a large amount of computational power. Lastly, there is a concern associated with these approaches, which relates to their lack of explainability since these algorithms work as a black box. This is a concern because, in many instances, these algorithms will make a particular message to appear as a phishing message, yet this action is not transparent.

2.6 Hybrid Phishing Detection Approaches

Hybrid phishing detectors are systems that use several phishing detection methods together to achieve higher accuracy, reliability, and robustness. Rather than focusing on one technique alone, these systems use different types of analysis to leverage the limitations that are inherent to single techniques. It has been shown that the use of heuristic/rule-based analysis together with other external threat intelligence services gives a better result for phishing attack detection than the other two individually. Rule-based analysis allows for the detection of malicious behavioral and linguistic patterns found in messages, and other external services supply the system with real-time URL and domain reputations.

External threat intelligence platforms provide constantly updated information gathered from different sources in the security industry as well as worldwide monitoring. Such information serves as a supplement for local analysis, as it confirms suspected particulars based on threat intelligence databases. Consequently, it is possible for a hybrid threat system to both identify known phishing attacks as well as new attacks. The hybrid method eliminates the problem of false positives associated with threat identification logic.

References and findings reveal the effectiveness of the hybrid anti-phishing system in combating zero-day phishing attacks. The effectiveness is especially true in cases where the blacklisting method is not effective. The different techniques together help to develop a system capable of adapting to the ever-changing nature of attacks. On the other hand, the existing studies prove the presence of some challenges in the use of the hybrid method. These include the utilization of the third-party API, the rate limitations, the delays, as well as the secure handling of the sensitive API keys.

2.7 Role of Threat Intelligence Services

Threat intelligence systems have been prominently covered in the cybersecurity literature as critical systems for detection, analysis, and protection against malicious actions within an evolving threat landscape. Threat intelligence systems compile and synthesize various types of data related to cybersecurity into aggregate data sets from sources such as anti-virus software providers, cybersecurity associations, and cybersecurity research groups. Through the analysis of data from various sources, threat intelligence systems enable real-time analysis of emerging threats and known malicious data like domains, URLs, and files.

Several research papers stress that the integration of phishing threat intelligence services via Application Programming Interfaces (APIs) helps achieve real-time access to fresh threat intelligence. Therefore, the need for fixed dataset updates and human intervention will be reduced, and the detection systems can effectively act upon newly identified phishing attacks. Consequently, phishing detection systems can check the malicious URLs against the global threat intelligence databases and enhance the functionality of identifying existing and novel phishing threats.

It is further emphasized by researchers that the adoption of threat intelligence improves the accuracy of detection by offering complementary information in the form of reputation rankings or historical assessments of suspicious objects. Nevertheless, the current literature on the subject strongly emphasizes the need for proper secure management of API keys, implementation of appropriate access controls, and compliance with the principles of data privacy and ethics while consuming external intelligence resources.

2.8 Research Gap Identified

The existing literature review reveals that despite the number of different phishing protection solutions proposed in recent years, none of them have the ability to provide complete protection from domains of phishing attacks. Many existing phishing protection solutions mainly focus on blacklisting techniques, which compare URLs/domains to known blacklists of phishing domains. Although blacklisting techniques can provide great protection from known phishing domains, in reality, they fail to provide protection from zero-day phishing attacks due to the short lifespan of phishing domains that tend to be created and deleted in an extremely short period of time. As many phishing domains tend to have very short lifespans, blacklisting techniques fail to provide protection from the ever-changing threats of phishing attacks.

Recent research has also been conducted on the use of machine learning and artificial intelligence techniques for phishing detection. The results have been encouraging in identifying accuracy, as these techniques can identify intricate patterns in data on a larger scale. However, there are certain major limitations, according to the literature, associated with these techniques. Machine learning systems require large amounts of data for training purposes, which could, in fact, be challenging and time-consuming. These systems also require large amounts of computational resources, making them unsuitable for usage on scaled-down or academic setups. Also, many machine learning systems can be termed as black-box systems, providing no insights and explanations for their decision-making mechanism, which becomes an area of major concern in platforms requiring learning, where the decision-making mechanism has to be well comprehended.

On the other hand, standalone solutions with rule-based phishing identification are easier to understand and interpret. These systems are programmed with preset rules containing typical phishing attributes such as urgent messages, requests for login information, and deceptive URL formats. As effective as rule-based phishing identification methods are to interpret, according to research, their performance correlates greatly to the quality of the set of preset rules used in the system, which should keep up with developments in phishing tactics, as this can result in systems performing poorly in terms of accuracy and sensitivity to phishing attacks, as well as producing high levels of false positives or false negatives.

Another notable gap, as recognized in extant literature, is that not much importance is being given to user understanding, explainability, and secure implementation practices. Most of the studies carried out in the literature simply work towards improving accuracy, without much consideration being given to user presentation of accuracy as well as secure management of sensitive data, such as API keys and user data. This acts as a hurdle, especially when transparency, ethics, and simplicity of logic play a significant role. As such, the absence of the research on the phishing detector that offers strengths in accuracy, interpretability, security, and pedagogic appropriateness emerges from the literature. The need for the fusion of rule-based interpretability and the reliability of the threat intelligence solutions that operate in real-time and are sourced externally has become imperative. In this case, the approach will benefit from the overall/global security information to boost accuracy while maintaining the ability to provide explanations for the reasoning behind the decision, which will be understandable by humans. The research aims to close the research gap by developing the rule-based phishing detector system that incorporates API-based threat intelligence solutions.

2.9 Summary of the Chapter

This chapter has delivered a comprehensive review of research and academic work being done in the field concerning phishing detection techniques. The aim of this research is to review various research methods that have been proposed in the past and analyze them to find out some of the limitations and strengths associated with them. Various academic papers and studies from the realm of cyber studies have helped in structuring this chapter and giving an overview of all major approaches that are implemented in phishing detection schemes.

The review described blacklist-based detection methods that are reliant on blacklists that contain malicious URLs and domain names. As much easier methods to implement and very efficient against identified threats, according to the literature, their inefficacy against new phishing sites is very evident. The chapter continued with discussing rule and heuristic-based detection that differentiates phishing emails based on common behavioral and linguistic patterns like urgency features, password requests, and malicious URLs. This approach was

found to be very transparent and interpretable and, for that reason, very fitting for and applicable in learning and academic settings, but highly reliant on the defined rules for accuracy and flexibility.

In addition, the chapter looked at machine learning phishing detection methods, which involve the use of classification algorithms in detecting complex patterns in phishing and legitimate data sets. Although these methods prove useful with respect to accurate phishing detection, the need for extensive labeled data sets, complexity in computation, and lack of transparency was highlighted within the various literatures examined. This is mainly because most machine learning algorithms lack transparency in their processes, which is crucial in user-level tasks like phishing prevention.

The chapter also introduced the rising need for the use of hybrid phishing detection methods that leverage a combination of techniques to counter the weaknesses of single techniques. Specifically, the inclusion of external threat intelligence services emerged as an interesting way to improve phishing detection solutions. Threat intelligence systems employ real-time reputation as well as insights that stem from the aggregation of several security providers. On the basis of the comprehensive analysis carried out, a research gap was identified. The literature survey indicated the need to develop a phishing detection tool where accuracy, explainability, and a secure implementation process are taken into consideration, and the concern should remain relevant to academia as a learning process. The identified research gap formed the basis to justify the phishing detection tool by implementing rule analysis, along with API-based threat intelligence. The following chapter describes the objectives, along with the research methodology, to fulfill the identified research gap to fulfill the aims of the current research study.

CHAPTER 3. RESEARCH OBJECTIVES AND METHODLOGY

RESEARCH OBJECTIVES

Research Objectives define the purpose of carrying out a research by clearly outlining its intentions. Research Objectives serve as an effective tool that sets a helpful framework for carrying out a research activity and analyzing its results. Talking specifically about phishing attack detection, Research Objectives of this project have been designed in a way that they are helpful in dealing with issues related to detecting phishing attacks. Below are the Research Objectives of carrying out this research:

- In order to analyze and comprehend the nature of phishing attacks and the techniques used by the attackers via malicious URLs and messages.

- Designing and implementing a rule-based system for phishing that can evaluate textual content to detect phishing cues such as urgency, requests for credentials, and linguistic pattern abnormalities.

- To integrate external API-based threat intelligence for the analysis of extracted URLs in terms of reputation and maliciousness for better detection accuracy.

- To combine rule based analysis and threat intelligence results through API to classify messages into LOW, MEDIUM, and HIGH risk levels.

- To offer explainable detection outcomes, such that the rules, indicators, and threat intelligence factors used for making a particular classification are clearly explained.

- For the purpose of developing a web application using a friendly interface aimed at both academic and demonstration learning of phishing techniques.

RESEARCH PROBLEM

Phishing attacks represent one of the most resilient and successful cyber threats within the online environment. Despite existing security tools being widely available to mitigate potential cyber threats, users and organizations are still falling victim to phishing attacks due to their social-engineering nature and their ability to evade security mechanisms using their intelligence and creativity. Current phishing detection tools are inefficient in delivering a balanced approach that ensures accuracy, transparency, and understanding in their detection process. The majority of these tools are mainly relying on blacklisting techniques for phishing attacks and are less efficient in attacking phishing websites with short life spans or those which are recently developed.

Recently, machine learning-based phishing detection tools have been in common use as they possess the capacity to detect complex patterns. However, most machine learning-based phishing tools require extensive data as well as substantial processing capacity. In addition to that, the issue of interpretability associated with many machine learning models is such that it makes it difficult for users to interpret the justification behind the phishing detection results. Rule-based methods in phishing protection are transparent and easy to understand but may lack effectiveness when not combined with external mechanisms of validation. Also, current

systems place considerable emphasis on accuracy but not as much emphasis on secure implementation techniques, explainability, or awareness. This generates a research problem that requires the designing of a precise, transparent, secure, and educational-friendly system for anti-phishing functionality.

As a result, the research problem that this study aims to solve is the absence of a more effective and transparent phishing message detection system that incorporates both rule analysis and external threat intelligence.

RESEARCH DESIGN

Research Design: Research design is defined as the overall strategy and structure used for accomplishing a research study in a systematic and organized manner. Based on the present study, the research design used is the descriptive and exploratory method, considering the fact that the project intends to analyze the techniques for phishing detection and build a system upon the patterns and behaviors.

The exploratory nature of research design is directed at understanding what phishing is all about and what its indicators are. The research design involves learning from existing knowledge about phishing and even examining sample phishing messages for their common indicators like urgency and phishing URLs. The exploratory research design is suitable in this project because it provides an opportunity for understanding a problem area that keeps changing with time due to rising threats.

The descriptive component of the research design is focused on analyzing and describing the behavior of phishing messages according to definite rules and criteria. The system proposed in this project analyzes messages based on rule analysis and assigns messages different risk categories. Descriptive research is beneficial for systematically defining the role of different variables and their influence on phishing message outcomes and the role of rule analysis and API intelligence on message categorization.

This research design does not mainly consist of hypothesis testing and the manipulation of variables. Rather, this project mainly revolves around the development of a phishing detection system according to the observed patterns and established cybersecurity principles. The chosen research design is appropriate for the objectives of this project, as this research works on the theoretical knowledge as well as the implementation of phishing detection techniques on a theoretical platform related to the academic environment.

TYPE OF DATA USED

The research study makes effective use of both qualitative and quantitative data in order to appropriately address its research objectives and assist in creating a phishing detection system. This blend of both types of data enables an effective screening of not only the technological issue of phishing attacks but also of their behavioral elements.

Qualitative data is an important component in understanding the nature and structure of phishing emails. Qualitative data encompasses text information taken from phishing and normal emails. The text information includes the bodies of the messages, alert messages, and

emails. The qualitative analysis is done by identifying behavior and language features traditionally linked to phishing attacks. The language features comprise aggressive language, calls for private information, impersonation messages, and misleading messages. Analysis of qualitative data is essential in understanding the behavior of attackers and the development of rule criteria.

The quantitative data is used in the objective assessment of the performance as well as the outcomes of the phishing detection system. These pieces of information include the application of rule-based risk values posted onto the messages, the number of rules used in message detection, VirusTotal detection values, the degree of confidence, as well as the determined risk categories. In this regard, the quantitative information makes it feasible to interpret the outcomes of the detection attempts based on a number of test cases.

For the study, secondary data sources are considered primarily, which comprise examples of phishing emails available in the public domain, genuine examples of message content, as well as URL reputation gathered by utilizing threat intelligence tools. There is no need for primary data collection that includes human subject participation to ensure that all relevant privacy and ethical matters are met appropriately. Using secondary data for the study is relevant as enough information is gathered that does not reveal private details of any individual.

Taken altogether, the combination of the two sources of data allows for the support of an extensive analysis of phishing detection methods and facilitates the construction of an interpretable and academically valid solution for phishing detection.

DATA COLLECTION METHOD

Data collection constitutes an integral stage of any research project, as it helps to establish the relevance and quality of the data used for the analysis process. In this project, the process of carrying out the data collection was conducted through a secondary form of data collection. This was the case because the project in question does not require human subjects or real-time data collection. The secondary form of data collection was preferred for its ethical standards, which are fitting for the environment of the research project.

The data used for this project was acquired from public and trusted sources such as examples of phishing communications, legitimate messages, and case studies on cybersecurity attacks and protection. The phishing and legitimate example communications were acquired from online cybersecurity awareness platforms and reputable security blogs and archives that host examples of phishing communications. The legitimate messages were acquired from common notice formats and styles such as service notices and information alerts to ensure a fair sample for both phishing and legitimate communications for system evaluation purposes.

Aside from message content, collection of other data relevant to URLs was conducted via the incorporation of external threat intelligence feeds. The extraction of URLs from sample messages was conducted using an API-based threat intelligence tool, which gave the results of reputation and detection based on aggregated input from various security vendors. In this way, the tool was capable of getting real-time information regarding the maliciousity or innocence of URLs, other than locally maintaining a set of threat intelligence.

These data samples were then organized in terms of their analysis. Phish messages as well as valid messages were labeled in accordance with their known properties in order to enable proper analysis of the efficacy of the detection message system. None of the data handled was related to user privacy. In fact, all data processing steps followed proper scientific guidelines.

The secondary data collection technique used in this paper can achieve the goals and objectives effectively, and this technique may yield sufficient information to analyze patterns and the effectiveness of the proposed system to counter the problem of phishing. It also makes the results reproducible to achieve academic integrity.

DATA COLLECTION INSTRUMENT:

A data collection instrument is a set of mechanisms that are used for the gathering, documentation, and processing of data for research. The major data collection instrument for this research is a web-based application for phishing detection, which determines the textual messages and URLs, developed for this research work with the application developed as a major instrument for collecting the data for this research work.

The web application allows the user to enter the content of the communications, for example, emails and alerts, using an organized format. Once the system receives this data, it processes it using an analysis engine based on rules, whereby it evaluates language and behavioral features typically used in phishing activities. Such features include urgent language, demands for personal details, elements of impersonation communication, and unusual communication patterns.

Besides the text analysis feature, the application uses an external API/threat intelligence service through the data collection tool. The URLs obtained from the input text are automatically fed into the API to provide reputation and detection details. The results obtained from the API are recorded and temporarily stored for evaluation and provide quantitative information in terms of detection numbers.

This tool also records all the output by the system, which may include rule-based risk scores, final risk decisions, confidence scores, and explainability information. All these are considered structured data that can be used for the assessment of the effectiveness of the phishing protection system. This application maintains ethical practices by ensuring that there are no personal user details permanently retained.

Furthermore, the web application created specifically for this study is a complete tool for collecting data that allows for systemized processing and evaluation of this data. This has created a system that is accurate and traceable for research use.

SAMPLE SIZE

The sample size is defined as the number of observations or instances of data utilized within an investigation to analyze the research objectives. Within the current research, the sample size is represented by a pool of phishing and genuine message samples utilized within the investigation to analyze the effectiveness of the proposed phishing detection solution. The sample represents an equal pool of messages comprising phishing attacks and genuine messages.

The sample data set includes various phishing message samples that have been tagged with noting such as urgent phishing messages, phishing requests for credentials, phishing suspicious links, and phishing impersonation prompts. On top of that, various genuine message samples such as service notifications, informational alerts, and general communication messages were also considered to appropriately assess the accuracy of false positives. Including both phishing and genuine samples will allow one to appropriately test the phishing detection accuracy of the system.

The sample size was chosen such that it was adequate for academic assessment and testing, and not suited for large-scale statistics. As this research is concerned with the design, development, and assessment of a rule-based phishing detection system, this research relies on qualitative assessment rather than general statistics. The sample size chosen also enables repeated testing and the interpretation of detection results.

On the whole, the selected sample size is sufficient for proving the efficacy, reliability, and explainability of the proposed system in an academic research environment.

SAMPLING TECHNIQUE

Moreover, the technique used in selecting data samples from the larger population for analysis is known as the “sampling technique.” In the current study, one type of non-probability sampling technique, “purposive sampling,” is used. There is nothing incorrect in applying this type of technique to the current study as the objective of the study is to analyze certain attributes in phishing as well as legitimate emails.

Purposive sampling is the selection of samples for the research based on criteria set for the study. The criteria set for this study involve the selection of message samples in a way that is representative of general phishing activity, including urgent messages, attempt to steal login details, impersonation scams, as well as the presence of general URLs. The selection of genuine messages is meant to be representative and normal communication activities including service and information messages, and lack the general factors associated with phishing.

By using the non-probability technique of purposive sampling, it becomes easiest for the researcher to target valuable and informative data that directly helps in the assessment of the phishing detection system. Since the main focus of this study is on the assessment of functionality, accuracy, and explainability of the rule-based method of phishing detection, it becomes easier for the researcher to use purposive sampling techniques rather than using random sampling techniques. By using random sampling techniques, it is quite probable that numerous samples be irrelevant.

In conclusion, the selected sampling method ensures that the data applied in the analysis is closely related to the aim of the analysis and that there is sufficient information covered regarding both phishing and legitimate messages. The analysis is likely to result in the development of an efficient phishing detection system due to this method used.

DATA ANALYSIS TOOL

Data analysis tools are defined as software programs and associated computational and logical methods applied in processing and analyzing research data to satisfy certain research objectives. For the case study presented in this research paper, associated data analysis has been applied using an in-house-built phishing detection tool that incorporates both software-implemented analysis and additional API-based threat services. The tool and techniques applied are quite transparent and research-oriented.

The rule-based phishing filter engine is the main analysis system in this project. It has been developed using the programming language Python and is intended to check the content of the textual messages based on the predefined rules of phishing message identification. The rules of phishing message identification have been developed based on the commonly identified phishing pattern and reported in the literature of cybersecurity. The engine assesses the messages based on the indicators that include urgent messages or threats, the explicit request of the critical credential, the indicators of impersonations, the use of domains and IP URLs, the short URLs, and the IP URLs. Each of the rules is provided with an identified weight depending on the criticality of the indicators. If the indicators belong to the rules that have been provided with higher weights, they contribute to the assessment of the phishing risks.

Through the use of the weighted score technique, the analysis process by the system is carried out systematically as opposed to a subjective analysis process. The overall score attained, depending on the threshold score, enables the categorization of messages as LOW, MEDIUM, and HIGH risk messages, respectively. In a research process, transparency within the analysis process plays a critical role, and the use of the weighted score technique enables transparency throughout the process.

To improve the effectiveness of the detection process, the rule-based analysis is supplemented with an external API threat intelligence service, which is considered an additional tool for data analysis. The URLs obtained from the message content are submitted for analysis to the threat intelligence API, which offers reputation and detection information based on the aggregation of different security vendors. The API returns analyzed data, including malicious, suspicious, and harmless detection values. The externally sourced intelligence service facilitates the validation of the URLs in a real-time process and is useful for determining the existence of a threat that may not be achievable through text analysis. The API-supplied information is integrated into the analysis process.

The Flask web server acts as the backend processing engine that integrates all the activities involved in the flow of information from user input, passing through the analysis components, to the display of results. With Flask, request processing, the execution of the rule-based and API-based analysis components, and the dynamic generation of the results are done. These results comprise the final risk category, cumulative risk score, confidence level, and the explanation component that lists all the applicable rules. They are presented in an intuitive user interface.

No advanced statistical softwares and mathematical models have been used in this study. The purpose of this approach is that the focus of this research is more on functional system evaluation and much less on predictive and inferential purposes. The choice of softwares used in this research has been appropriate in context to the aims and focus of this work. In general, a combination of the rule-based analysis engine, threat intelligence from APIs, and the web-based processing framework offers an appropriate and effective set of methods for data analysis tasks. The set of methods is effective in phishing analysis, with clear decision-making processes and interpretation results, which makes it appropriate for academic research purposes.

Figure 3.1 illustrates the overall system architecture of the proposed phishing detection system, highlighting the interaction between the user interface, backend processing modules, and external threat intelligence services.

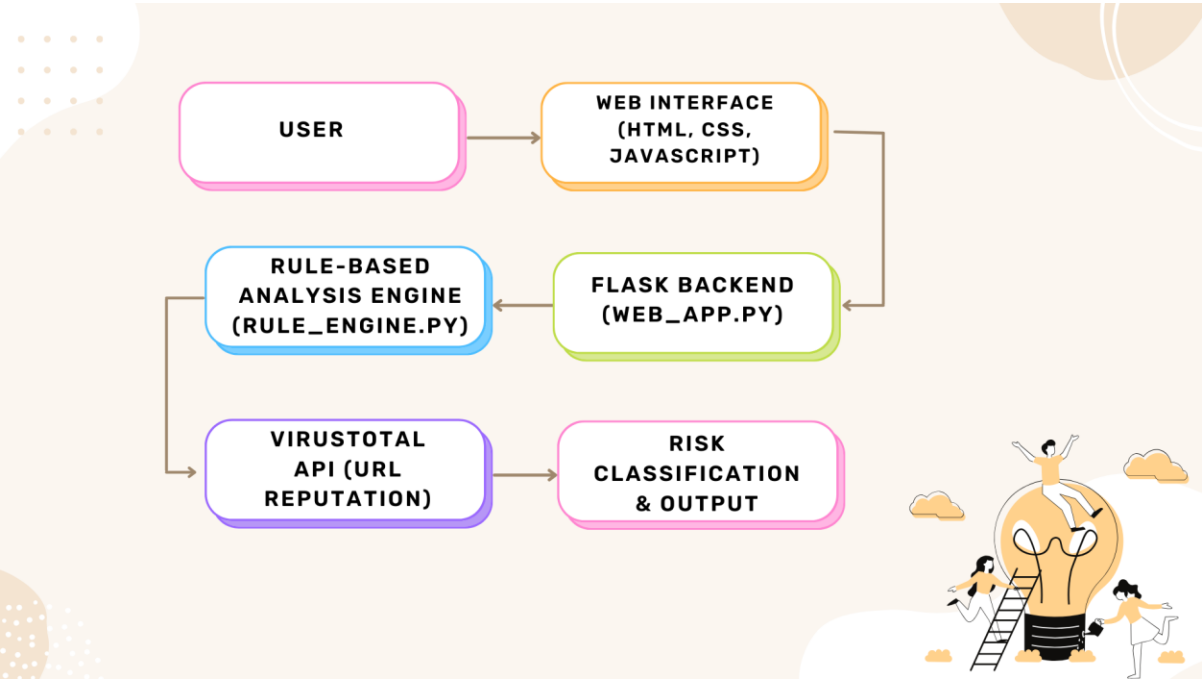

Figure 3.1: System Architecture of the Phishing Detection System

Figure 3.2 The User Interaction Flow Diagram illustrates the sequence of actions performed by the user and the system during message analysis. It shows how user input is processed by the system and how the final phishing risk assessment is displayed to the user in an interpretable format.

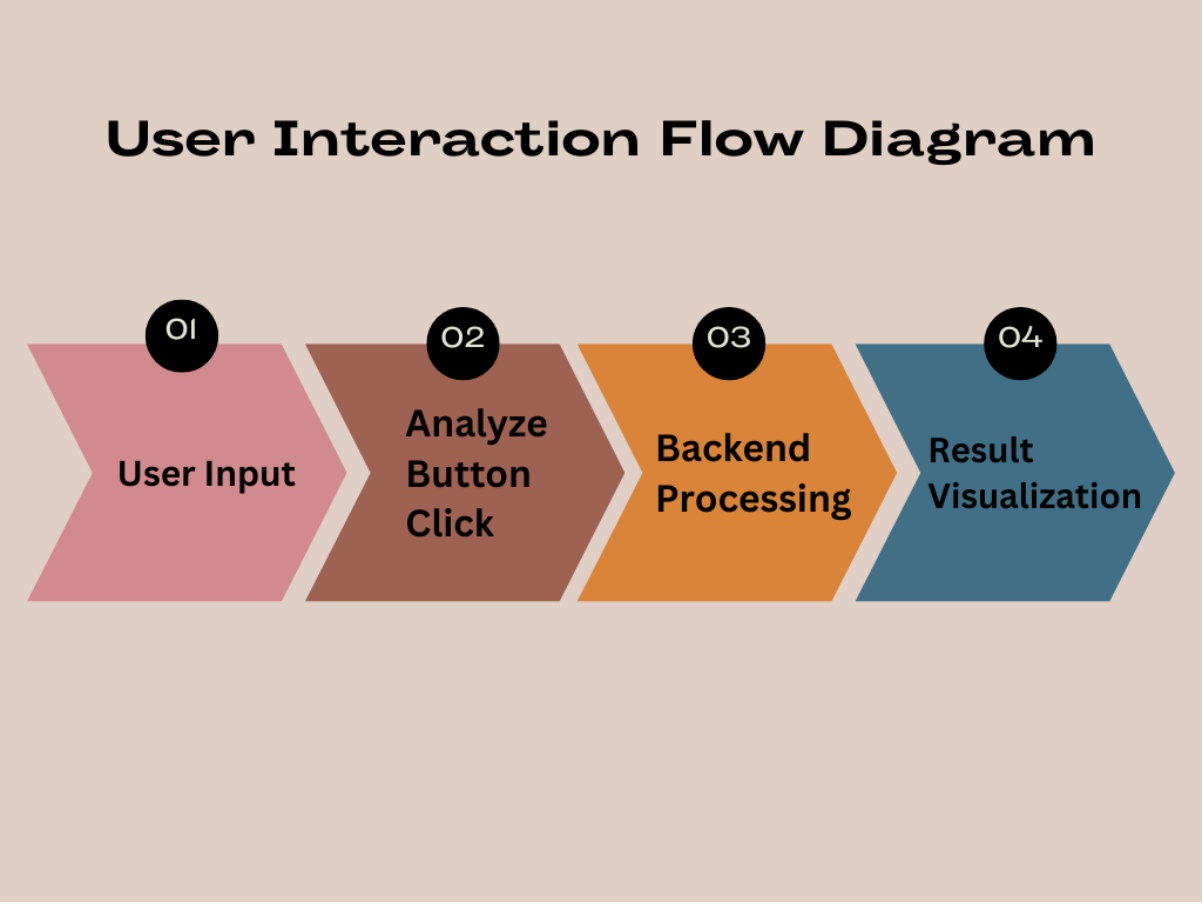

Figure 3.2: User Interaction Flow Diagram

CHAPTER 4. DATA ANALYSIS, RESULTS, AND INTERPRETATION

4.1 Introduction

This chapter focuses on the detailed analysis and interpretation of results obtained from the implementation of the proposed phishing detection system. The primary objective of this chapter is to evaluate the effectiveness, accuracy, and reliability of the system in identifying phishing messages and clearly distinguishing them from legitimate communication. The evaluation process is based on the combined use of a rule-based detection mechanism and an API-driven threat intelligence component, which together form the core detection framework of the system. By systematically examining the outputs generated during the testing phase, this chapter seeks to determine whether the proposed solution can accurately detect phishing attempts under realistic communication scenarios.

To assess the performance of the phishing detection system, a diverse set of representative test inputs was utilized. These inputs included messages deliberately crafted to simulate phishing attacks as well as legitimate messages that reflected normal communication behavior. The phishing samples were designed to incorporate commonly observed characteristics of real-world phishing campaigns, such as urgency cues, threats of account suspension, requests for sensitive credentials, and the inclusion of suspicious or shortened URLs. In contrast, legitimate messages consisted of informational or routine communications typically received from trusted sources. The use of such varied input data enabled a comprehensive and balanced evaluation of the system's detection capabilities.

During the analysis phase, each input message was processed by the rule-based detection engine, which examined the content for predefined phishing indicators. Each detected indicator was assigned a weighted score based on its severity and relevance to phishing behavior, including indicators such as urgent language, credential-related terms, and suspicious URL patterns. At the same time, any URLs extracted from the messages were analyzed using an external API-based threat intelligence service. This service provided reputation and risk information by aggregating detection results from multiple security vendors, thereby offering an additional layer of validation to support the final classification.

The system generated several critical outputs for each analyzed message, including cumulative risk scores, a list of triggered detection rules, confidence values, and a final risk classification categorized as LOW, MEDIUM, or HIGH. The inclusion of an explainability component allowed users to clearly understand the factors influencing each classification decision, thereby enhancing transparency and trust in the system. Overall, this chapter plays a crucial role in validating the proposed phishing detection strategy. The results and

interpretations discussed here demonstrate that the integration of rule-based analysis with API-powered threat intelligence is a practical and effective approach for addressing contemporary phishing threats and provide a solid foundation for the conclusions and recommendations presented in the subsequent chapters.

4.2 Data Analysis Approach

In this research, the method of data analysis used involves a hybrid assessment approach which relies on a combination of rule-based message assessment techniques together with external URL evaluation through an Application Programming Interface (API). This method has been designed to benefit from the best of both worlds in terms of behavioral assessment at the local decision point in addition to threat intelligence in efforts aimed at improving the accuracy of phishing detection. This particular system tackles various challenges presented by the existing independent assessment approach.

The phishing assessment procedure performed by the system is in a sequential manner. When a user uploads a message for processing, for example assessing whether it is a phishing message, the system first analyzes the message content using a predetermined rule-based system. The system analyzes a message for a set of phishing characteristics widely documented in information technology as a means of countering phishing. Such characteristics include exhibiting urgent or threatening text aimed at inducing panic or fear in victims, requests for personal information such as passwords, Personal Identification Numbers, or banking information, utilization of URL shortening services that hide malicious destinations, and utilization of IP-based URLs as opposed to using normal domain names.

Each identified phishing indicator has a pre-defined, weighted value based on its type and level of likely intent. Highly indicative phishing indicators, such as authentication requests and IP-based URLs, carry higher values, and those that are not as serious, such as urgent messages, carry medium values. This final value calculated from the aforementioned steps is the behavioral risk value associated with the corresponding email. This method enables a logical risk assessment based on the email.

At the same time, the system is also capable of carrying out URL extraction and analysis by figuring out all the links present in the message content that are URLs. The URLs are then processed by passing them to an external threat intelligence API to check their reputation. The threat intelligence API is capable of carrying out end-to-end analysis by compiling detection from various antivirus engines, security vendors, and threat research. This helps the system to check whether a particular URL is marked malicious, reputable, and benign in the past or not. This helps the system to have up-to-date knowledge of global security because all threat knowledge is stored in external threat intelligence.

The outcomes of the rule-based analysis as well as the API-based URL reputation evaluation are then collated and assessed cumulatively in order to form a collective decision on the classification of the message. A predefined risk level of either LOW, MEDIUM, or HIGH is assigned to the message based on the cumulative risk scores as well as the external reputation factors. The dual-analysis technique provides a collective analysis of the predefined behavioral traits of the message as well as the real-time external threats. In general, this data analysis procedure can improve the inclusiveness and accuracy of phishing detection, keeping transparency and interpretability in mind. Also, the system gives users complete information regarding the rules triggered, the confidence level, and risk score involved in identifying a certain message. Being highly interpretative, this system is very appropriate for academic assessment in institutions of higher learning, especially in educating students on cybersecurity.

4.3 Rule-Based Risk Scoring Analysis

The rule-based risk score evaluation is the heart of the decision-making process in the proposed phishing type detection system. This system has the responsibility of analyzing user-submitted message content in order to determine the likelihood of messages being phishing attempts. The system analyzes each user-submitted message using a rule engine with an established set of rules developed based on common phishing behavior. These rules have been developed based on well-recognized phishing behavior patterns as identified by research in cybersecurity studies. The rule-based system makes sure that the system delivers repeatable results in the analysis process of phishing attempts.

Every rule in this system is then given a weighted risk score according to the severity of the indicator and its correlation to malicious intent. The mechanism of applying weights enables the system to distinguish between light indicators of malicious intent and robust indicators of phishing activities. Rather than using just an indicator, the system analyses various features that are available in the email and gives it an accumulated risk score.

The presence of urgent and threatening language is one of the key factors that are being measured by this system. The keywords and phrases like "urgent," "verify," "immediately," "act now," and "limited time" are often employed by attackers to instill a sense of urgency and panic in victims. These keywords are used with the aim of manipulating users to immediately respond to the message without necessarily verifying its authenticity. Although urgent language is sufficient to arouse suspicion about phishing activity, its presence does not necessarily confirm that a message is being sent for phishing purposes. Hence, this language is assigned a medium risk score.

The other important aspect regarding the scoring mechanism is the identification of terms related to credentials. Terms such as “bank,” “login,” “password,” “OTP,” “PIN,” and “credit card” are immediately linked to personal or financial details. It is a known fact that phishing attacks try to cheat users into revealing their credentials by misusing them as a ruse to gain those credentials. The use of multiple terms related to credentials can result in a significantly higher overall risk score, as it is an obvious ploy to obtain credentials.

The system also assesses the presence of URL shorteners in the message. URL shorteners have been used by attackers in an attempt to hide the destination link of the URL, making it hard for users, if possible, to identify the suspected URL. URL shorteners, also referred to as shortened URLs, prevent users from analyzing the credibility of the URL, which can easily result in the redirection of users to the phishing pages. The presence of URL shorteners has been considered an element associated highly with phishing attacks, making this assessment high risk.

Moreover, the highest personal risk score in the system is generated for those messages which have URLs consisting of raw IP addresses. Rarely do legitimate messages have such URLs. In many cases, they are typically found in phishing attacks, malware, and command-

and-control servers. Raw IP URLs do not go through the conventional checking of domains for their reputation. This greatly affects the final result of the classification decision. The ultimate risk score is calculated by adding up all the scores of fired rules. With the use of predetermined risk level thresholds, messages are then categorized based on risk levels that can be LOW, MEDIUM, or HIGH. The fact that this system is multilevel provides for more informed decisions, as opposed to only being able to make binary decisions. It is also worth noting that the rule-based risk scoring system supports improved transparency in respect of rules that are fired in arriving at a particular risk level. Such transparency is particularly useful in academic pursuits, cybersecurity awareness, and learning.

Table 4.1

Rule Category	**Indicator Detected**	**Score Assigned**
Urgent Language	urgent, verify, immediately	2
Credential Request	bank, login, password	3
URL Shortener	bit.ly, tinyurl	3
IP-based URL	Numeric IP instead of domain	4

Figure 4.1 illustrates the rule-based risk scoring mechanism applied to a sample message. The output displays the triggered phishing rules, the cumulative risk score, and the resulting risk classification generated by the system.

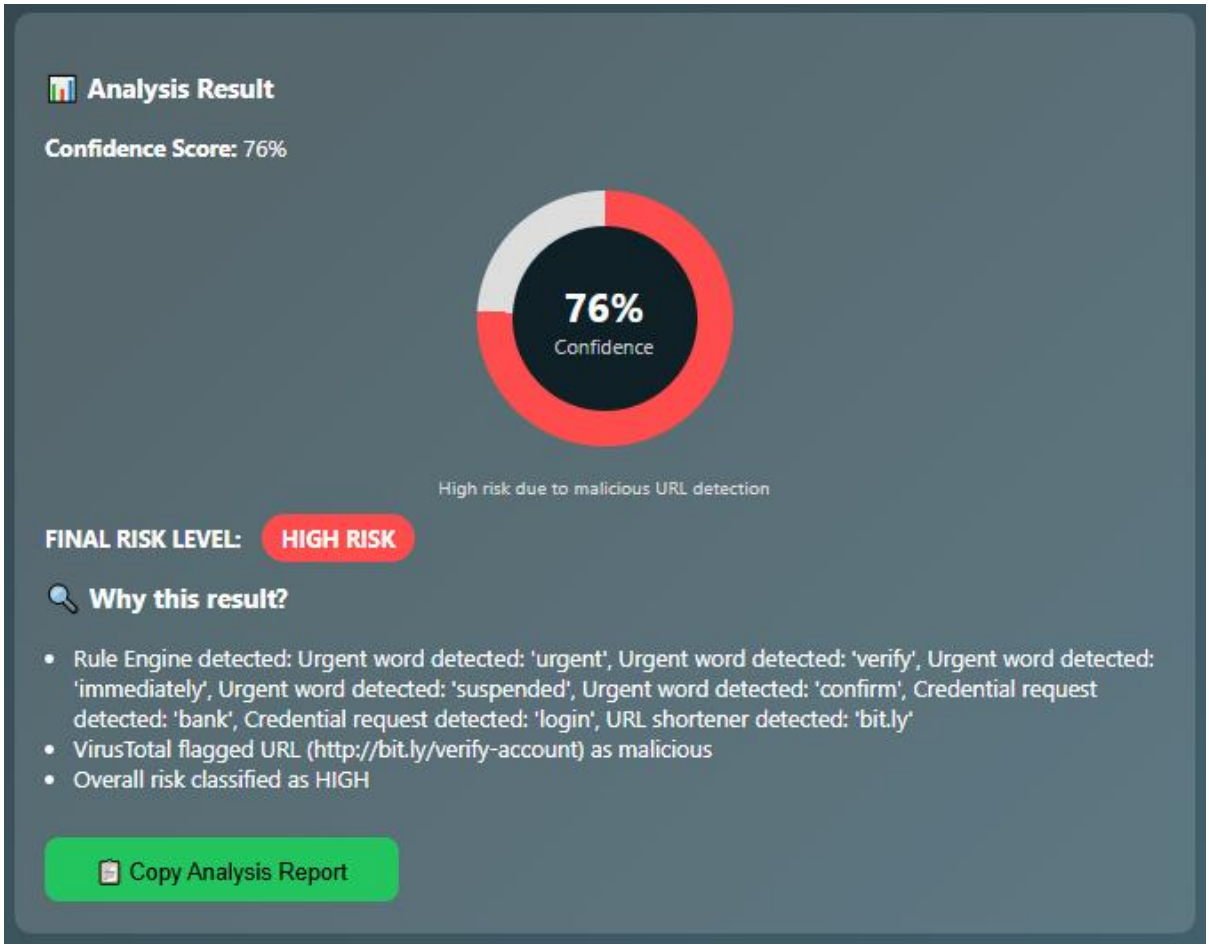

Figure 4.1: Rule-Based Risk Scoring Output

4.4 Result Analysis: High-Risk Phishing Message

In one of the cases tested during the evaluation of the proposed system for phishing detection, the tested message had various characteristics that may be associated with phishing attempts. The content of the message was based on urgency considerations, which was used to pressure the user into taking immediate action. The use of urgency considerations is a well-known social engineering method used by attackers with the aim of preventing reasonable decision-making by the victim. In this message, urgency was also used alongside requests for various credentials such as banking credentials, which may result in phishing attempts.

Moreover, the fact that this message contained a shortened URL increased its level of suspiciousness considerably. Shortened URLs are common in phishing messages, which attempt to mask the target URL of the intended link, as well as evade traditional security measures. Through this, they are able to successfully target several potential victims without first ascertaining the legitimacy of the URL they are sending them to. These indicators of phishing clearly show that this message falls under high phishing risk categories.

The rule-based detection engine used in the system effectively picked up each of these indicators of phishing attacks. For each detected feature, the system used pre-defined weighted scores depending on the level of importance associated with each indicator. The use of urgent words added up to the overall risk score since these words are associated with psychological manipulation. Use of words associated with credentials added up to more since

they show a direct attempt at stealing important data. The presence of the URL shortener also added up to the overall risk score since URL shorteners are highly associated with malicious activity. All these weighted scores added up to an overall risk score beyond the pre-defined threshold, thus placing the risk level under HIGH RISK.

Message Risk: HIGH RISK

The message was rated HIGH RISK, and this was effectively communicated to the user through the graphical user interface. The resulting message was shown through the use of color-coded risk levels and the risk gauge, which was in the form of a circle. A moderate confidence level was also shown, which depended on the intensity of the phishing indicators used to classify the message. To ensure that the user has all the information needed, the system has an explanation and transparency part which shows all the rules used to classify the message, including urgent word, credential request, and the URL shortener rules.

Conclusion The outcome of this test case ensures that the proposed phishing detection system has the ability to properly and effectively identify high-risk phishing communications that consist of various malicious indicators. Its effectiveness in properly detecting and properly explaining high-risk phishing communications ensures that this particular detection strategy using rules and API assistance is quite reliable and practical. This particular outcome ensures that this proposed phishing detection system is indeed quite effective in detecting actual phishing communications.

Figure 4.2 shows the analysis result of a high-risk phishing message detected by the system. The output highlights the triggered phishing rules, elevated risk score, and final classification as HIGH risk.

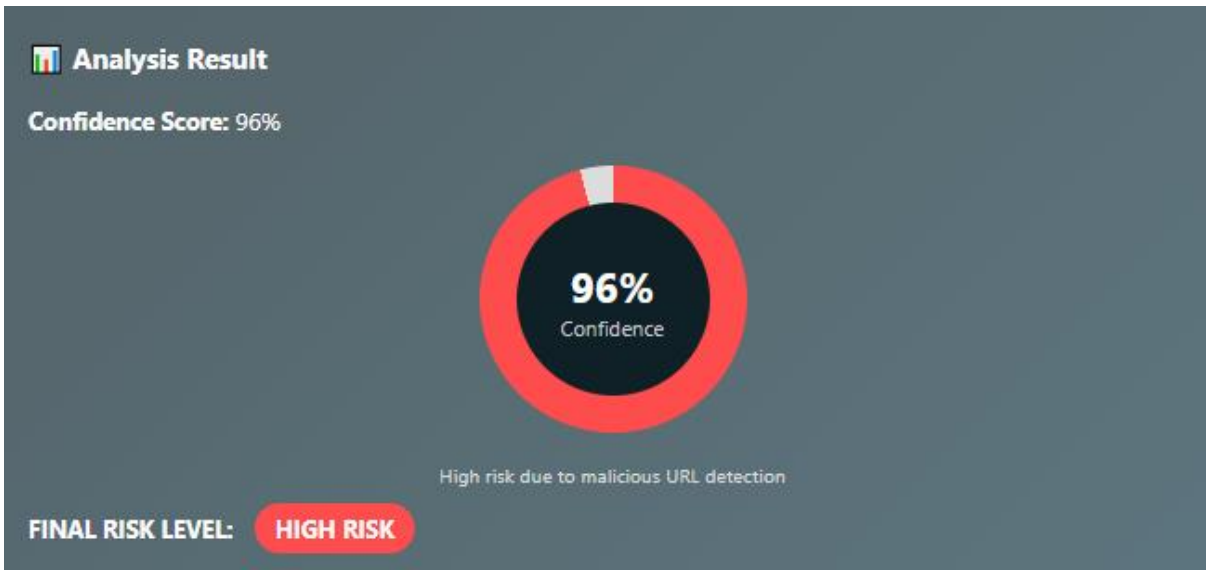

Figure 4.2: High-Risk Phishing Detection Result

4.5 Result Analysis: Low-Risk Legitimate Message

Another test case was also carried out to evaluate the effectiveness and reliability of the proposed phishing detection system with a genuine email that carried some informational text. The purpose of choosing a test case with a genuine email is to evaluate the phishing detection system's accuracy in flagging a genuine email and also to check if a genuine email is mistakenly detected as a phishing email. This is a required evaluation for determining the phishing detection system's effectiveness in eliminating false positives.

The message for the legit choice does not have any qualities common in the phishing message. This is because the message does not have an urgent tone that tries to make people act immediately. In addition, there is no request for any confidential data in the message, which can be classified as legit since it does not ask for info like passwords, OTP, bank info, and ID. Moreover, legit messages do not have shortened URLs, domains, or IP, which can help phishers hide malicious URLs. All these make up legit message composition.

Analysis Phase:

The system evaluated the message using the rule-based detection engine. Each predefined phishing rule was systematically evaluated to see if the message contains elements of the rules that include the use of "urgency indicators," "credential request keywords," "url shorteners," and "IP-based urls." As the input message did not contain these indicators, none of the rules were fired. Therefore, the cumulative risk score was maintained at a minimal value. The result showed that the message did not contain the deceptive pattern of sending messages that feature observable traits of phishing messages.

On the basis of the calculated risk score, the system correctly identified the message as LOW RISK. The confidence score obtained from the application also supported the identification by indicating a high level of confidence about the authenticity of the message. Furthermore, the explanation part of the system specifically highlighted that no phishing rules were triggered during the analysis procedure. This provides transparency to the user and increases the confidence level of the user. The above output not only proves the efficiency of the system to distinguish between phishing and legitimate mails effectively and also to reduce the number of false positives. The importance of reducing false positives cannot be denied, and suppressing them results in avoidance of user frustration, lack of confidence in the system, and ultimately poor usability of the system. The correct classification of low-risk legitimate mails proves the consistency, accuracy, and viability of the proposed antiphishing method.

Figure 4.3 presents the analysis result of a legitimate message classified as LOW risk. The output shows the absence of phishing indicators, resulting in a minimal risk score and correct low-risk classification.

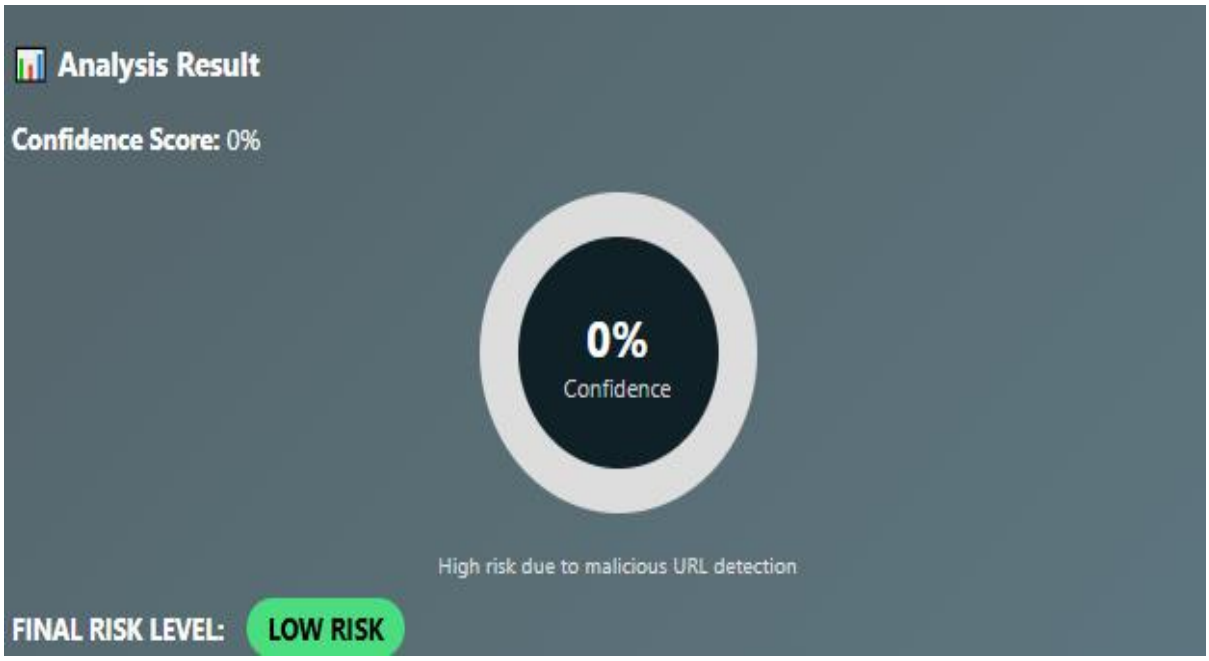

Figure 4.3: Low-Risk Legitimate Message Detection Result

Table 4.2

Test Case	Message Type	Triggered Rules	Risk Score	Final Risk
TC-01	Phishing Message	Urgent + Credential + URL	High	HIGH
TC-02	Legitimate Message	None	Low	LOW

4.6 Interpretation of Results

Analysis of the outcome of the results generated by the proposed phishing protection scheme reveals that the scheme works well under different message situations. Messages that have various phishing factors like urgent messages, demands for private credentials, and fishy or shortened URLs have been identified as high-risk messages. This confirms that the scoring system based on rules can effectively detect patterns associated with phishing attacks. Moreover, genuine messages that lacked those factors have been identified as low-risk messages, ensuring that there are no false alarms generated by the scheme.

The inclusion of threat intelligence through APIs further increased the reliability of the detection system. Through the verification of discovered URLs from other sources related to the security of the system, the system was able to incorporate real-time reputations into its decision system. The system was able to take into consideration both behaviors from the content message as well as external threats for more dependable decisions.

Indeed, the confidence score mechanism offered an added level of informativeness, where the more prominent indicators the message had, the higher the confidence score, while the less prominent or lacking indicators, the lower the confidence score. This added functionality provides the user with an extra level of clarity on the level of confidence for each message scanned.

In addition, the explainability functionality added a great degree of system transparency by allowing system users to see which of the rules and indicators had been used for classifying a particular message into a high or low category. This system allows users to see the reason why a certain message was classified into a high or low category. This system would literally prove its effectiveness in a learning institution where not only the end result but also interpretation of the rules used for detection matter immensely. Well, there can be no doubt about the effectiveness of this system for detecting phishing due to the interpretation of its results.

4.7 Summary of the Chapter

This chapter provided insight into the in-depth and systematic computation of the results derived from the proposed phishing detection system. The main objective of this chapter has been to analyze and determine the effectiveness of the proposed system in implementing rule-based computation and API-based threat intelligence in the detection of phishing attacks. The outcomes of this proposed system, through rigorous testing and result computation, provide evidence of the operational effectiveness of this proposed system in making the required computation of the embedded URL in the messages submitted by potential users in an interpretable fashion. The computation confirms that the hybrid method of detection proposed in this research study is apt in phishing attempts with reasonable accuracy.

The data gathered throughout the testing phase revealed, without any ambiguity, that messages that were composed of common phishing identifiers like urgent or threatening text, requests related to vital credentials, as well as use of suspicious or truncated links, were all

categorized as high-risks. This further ascertains that the rule-based identifiers and scoring system used within the application work efficiently. In addition, non-phishing messages were efficiently identified as being of low risks, thereby ensuring that there were no false positives associated with the application.

One significant aspect brought about by this chapter is the contribution of confidence scoring and explainability in improving the overall transparency level of the whole detection process. In this regard, users are clearly explained the confidence level associated with each message classification result, as well as the detailed set of rules followed in order for the message to fall into particular categories of risk classification. This particular aspect becomes highly significant within an academic environment since clarity of logic stands at high priority in learning concepts. In summary, the results presented in this chapter have confirmed the efficacy of the proposed hybrid approach that integrates rule-based evaluation with external threats from the external environment. The proposed approach ensures that the system performs well and that the data handling and implementation processes undertaken by the system are ethical and secure. Based on the assessment presented in this chapter, it can be concluded that the proposed phishing detector system not only performs well, but it is also suitable for assessment in an academic environment and educational institution related to cybersecurity concepts and awareness.

CHAPTER 5. FINDINGS AND CONCLUSION

5.1 Findings

Analysis conducted through the proposed phishing detection system resulted in several key findings that, combined, illustrate the effectiveness, reliability, and potential for practical application of the phishing detection approach used in the study. The phishing detection system was able to detect phishing messages that typically and commonly contain phishing attack indicators like the urgent or threatening tone of the message, request for sensitive or confidential information, inclusion of suspicious or shortened URLs, and the use of IP-based URLs rather than the correct domain-based naming. All messages that typically include any of these phishing attack indicators consistently demonstrated elevated cumulative risk scores, which were accurately classified under the HIGH RISK category. These findings effectively confirm that the rule-based approach to cumulative risk scoring has the ability to detect and place suitable emphasis on critical phishing attack indicators typically used by phishing attackers.

One of the most important findings of this research work is the effectiveness of phishing message discrimination by the proposed system. Those emails that lacked any phishing attributes such as urgency, requests for credentials, or phishing URLs were not subject to any detection rules, and as a result, very low levels of risk were assigned. These were always assigned a category of LOW RISK, thereby indicating that these proposed systems work very effectively in avoiding any false positives. These findings of this research work, therefore, confirm that these proposed systems strike a very effective balance between phishing message detection and the accuracy of normal message identification.

The other important finding is related to the integration of API-based threat intelligence, which greatly assisted in enhancing the detection capability of the system as a whole. The system, by validating the extracted URLs through external intelligence, was able to improve its decision-making process, particularly in scenarios related to shortened links frequently used in phishing attacks. The external validation level provided reliability to the detection process, which increased confidence in classifying a URL as potential phishing content. The combination of rule-based analysis and real-time intelligence in the system was very effective in overcoming both behavioral and technical hurdles associated with phishing attacks. Additionally, the role that confidence scores and an explainability feature play in ensuring that the proposed system is truly effective has emerged as another finding of significance. The fact that confidence scores and explanations for each resultant classification are clearly available has ensured that users are able to understand how phishing detection systems work and why certain emails have been labelled as phishing and why some are legitimate. Finally, it has been concluded that the proposed tool is not only successful at detecting phishing but is also an invaluable resource for learning more about cyber security.

5.2 Conclusion

On the basis of the results obtained in this research work, it can be concluded that the proposed “Phishing Detection System Using Rule-Based Analysis and API Integration” has played an effective role in detecting and differentiating between phishes and genuine messages in an appropriate, transparent, and interpretable manner. The proposed system has effectively utilized rule-based indicators and an external threat intelligence system to measure and assess the reputation of the URLs that are designed to check, thereby proving appropriate and successful usage of the proposed method of detection in this research work.

The outcome of the implementation proves the ability of the system to identify the critical features of phishing attacks like the presence of urgency indicators, requests for credentials, phishing URLs, and links based on the IP while ensuring a minimum number of false positives. The addition of a confidence score and a thorough explanation facility adds an extra layer of transparency and interpretability while performing the identification of the phishing attack. Unlike other phishing detectors based on ML, where the approach is a black box, the proposed system is easy to interpret as it provides an explanation of each classification result.

In conclusion, the project fully achieves its purpose of creating a functional, secure, and academically sound phishing detector system. The hybrid phishing detector system makes a great point of arguing that rule-based detection methodologies can complement API-based threat intelligence to provide an effective solution to modern-day phishing attacks while ensuring that ethical standards are maintained in handling and implementing those methodologies. The project acts not only as an effective piece of cybersecurity technology but also as an extremely helpful teaching tool for students pursuing information technology, providing a sound basis for future development in anti-phishing tools.

CHAPTER 6. RECOMMENDATIONS AND LIMITATIONS OF THE STUDY

RECOMMENDATIONS

On the basis of the conclusions and findings of the current study, the following recommendations can be put forward that can improve the impact and prospects of the phishing detection system:

- The system can be improved by combining machine learning algorithms with the current rule-based system in an attempt to enhance the level of accuracy in detecting complex phishing patterns. A hybrid system may assist in finding an equilibrium between explainability and adaptability with regard to machine learning algorithms.

- The integration of the application with email services and browsers would make real-time phishing detection a possibility, where users can get immediate notifications before clicking on any malicious content.

- The use of the system on a cloud-based platform can be viewed as enhancing the factors of scalability, availability, and performance, particularly when dealing with a high number of requests concurrently.

- The rule base applied for phishing purposes needs to be updated continuously based on the evolving phishing methods and new social engineering strategies that malicious actor adopt.

- Support for multiple languages can be used to extend the phishing detection capabilities the system offers beyond messages that are only in the English language.

- The logging and reporting capabilities of this software should be improved in a manner which will help it retain the record of the analyses performed in the past. It will be useful in security audits as well as future research.

- Incorporating other sources of threat intelligence can also be done in conjunction with the existing API to provide enhanced URL reputation assessment without relying exclusively on one external resource.

- "The system can also be extended in order to analyze the headers and metadata of emails, for example, the routing paths and the senders, in order to obtain more insights on the authenticity of the emails."

- User awareness features, such as educational warnings, notifications, and preventive tips, can also be implemented to enhance security awareness among users and encourage them to practice online safety.

- There are performance optimization techniques that can be employed in order to minimize response time and computational complexity, especially in large-scale or real-time analyses.

- A mobile application version of the system can also be developed for supporting phishing detection in smartphones and tablets in order to address the emerging use of mobile communication platforms.

- There can also be automated alert systems designed to alert the user immediately for phishing attempts that fall into the high-risk category.

LIMITATIONS OF THE STUDY

In spite of the success of the proposed phishing detection system design and implementation, there are limitations of this research work that need to be acknowledged. Acknowledging the limitations of a research or study is a significant way of ensuring academic transparency.

1. In regards to its phishing detection capabilities, it should be noted that it relies on predefined rules to detect phishing attacks, which can be restrictive in terms of maintaining its efficacy in detecting advanced or novel phishing attacks. This is because phishing attacks are continuously evolving.

2. The effectiveness of the URL reputation risk analysis is based on the availability, responsiveness, and correctness of the external threat intelligence API. Delay, outage, or incomplete results of the API service can impede the system's efficiency in timely and correct delivery of results.

3. The system is meant for localhost, which means that it couldn’t undergo rigorous tests in a real-life setting. Therefore, issues such as scalability, handling a large amount of traffic, and functioning in a real-time setting couldn't be assessed in this research work.

4. This study does not cover large-scale statistical verification or benchmarking over large datasets. It is conducted more over functional testing with specimen data, which serves well for an academic purposes but may fail to present an actual depiction of diverse threats in real-world environments.

5. Real-time integration with an email stream or message platform was not contained within the scope of work for this project. The software presently examines user content submissions manually without utilizing live communication channels.

6. The system deals only with text analysis in messaging and has no provision for the analysis of email attachments, script files, picture files, and other multimedia files used in more sophisticated phish attacks.

7. There are rate limits when relying on external APIs that might impact performance if the analysis process yields a large frequency of results. There might be limitations when scanning continuously for free-tier APIs.

8. The project is designed mainly for scholarly or educational intentions and not commercial production quality. Therefore, some of the high-end enterprise functionalities, like automatic incident response or auditing of compliance, are not in the project.

BIBLIOGRAPHY

Research paper:

- Dhamija, R., Tygar, J. D., & Hearst, M. (2006). Why phishing works. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. https://doi.org/10.1145/1124772.1124861

- Garera, S., Provos, N., Chew, M., & Rubin, A. D. (2007). A framework for detection and measurement of phishing attacks. *Proceedings of the 2007 ACM Workshop on Recurring Malcode*, 1–8. https://doi.org/10.1145/1314389.1314391

- Abu-Nimeh, S., Nappa, D., Wang, X., & Nair, S. (2007). A comparison of machine learning techniques for phishing detection. *Proceedings of the Anti-Phishing Working Groups 2nd Annual ECrime Researchers Summit*. https://doi.org/10.1145/1299015.1299021

- Jakobsson, M., & Myers, S. (2007). *Phishing and countermeasures: Understanding the increasing problem of electronic identity theft*. Wiley-Interscience.

- Zhang, Y., Hong, J. I., & Cranor, L. F. (2007). Cantina: A content-based approach to detecting phishing web sites. *Proceedings of the 16th International World Wide Web Conference (WWW 2007)*.

Websites:

- VirusTotal. (2024). *VirusTotal API documentation*. https://developers.virustotal.com/reference/overview

- Open Web Application Security Project (OWASP). (2023). *Phishing attack prevention*. https://owasp.org/www-community/attacks/Phishing

- Anti-Phishing Working Group (APWG). (2024). *Phishing activity trends reports*. https://apwg.org/trendsreports

- National Institute of Standards and Technology (NIST). (2022). *Digital identity guidelines (SP 800-63)*. https://pages.nist.gov/800-63-3/

- Flask Documentation. (2024). *Flask official documentation*. https://flask.palletsprojects.com/en/latest/

Books:

- Mitnick, K. D., & Simon, W. L. (2011). *The art of deception: Controlling the human element of security*. Wiley Publishing, 1st Edition.

- Andress, J. (2019). *The basics of information security: Understanding the fundamentals of InfoSec in theory and practice*. Syngress, 3rd Edition.

- Whitman, M. E., & Mattord, H. J. (2018). *Principles of information security*. Cengage Learning, 6th Edition.

APPENDIX

Project Source Code Repository

This appendix provides access to the complete source code of the project titled **"Phishing Detection System Using Rule-Based Analysis and API Integration"**, which has been developed as part of the Bachelor of Computer Applications (BCA) program at Amity University Online. The source code has been made available in a public GitHub repository strictly for **academic reference, evaluation, and learning purposes**.

The repository contains the full implementation of the phishing detection system, including the rule-based analysis module, API-based threat intelligence integration, web application logic, frontend interface files, and supporting configuration files. The project follows secure coding practices, and sensitive information such as API keys has been excluded from the repository using environment variables and appropriate version control exclusions.

GitHub Repository Link:
👉 https://github.com/anmolverma045/phishing-detection-system.git

The code structure in the repository is modular and well-organized, enabling reviewers and learners to understand the implementation logic easily. The repository also serves as supporting evidence of the practical work carried out for this project.