International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18(1s) 2026 pp. 1-11

https://doi.org/10.70917/ijcisim-2026-1942
Article

FUZZY LOGIC MEETS CYBER SECURITY: A LIGHTWEIGHT MODEL FOR PHISHING WEBSITES DETECTION

Bhagwat Mahesh[1]*, T. S. Vishwanath[2]

[1] BKIT Bhalki, Karnataka, India., mahesh.bhagwat4@gmail.com
ORCID: 0009-0000-6335-2909
[2] BKIT Bhalki, Karnataka, India., tsvishwanath123@gmail.com
ORCID: 0000-0003-2098-0328

Abstract: Real time phishing websites detection is a challenging and dynamic problem because it influenced by uncertainties and multiple factors. Detecting some phishing websites in real-time for a day now is really a dynamic topic involving several features and requirements. Fuzzy logic provides a robust approach to handle the variability and imprecision in such detection task. It offers a natural language framework for evaluating number of qualitative features. This research work describe, an intelligent system to detect phishing website with fuzzy logic concept and data mining techniques. The implementation is a work in progress, and it has seven key features with ongoing efforts to include few more criteria for increase accuracy. The aim of this implementation is to provide an effective solution against security threats and the use of maximum possible fuzzy rules to improve reliability and scalability. By using advancing detection system with fuzzy logic, this research focuses on contribution to a safer online environment and stronger defenses against the most cyber attacks
Keywords: Phishing; phishing features; fuzzy logic; fuzzy rules; data mining

1. Introduction

The internet has transformed various sectors like education, ecommerce, healthcare, agriculture and many more, it is now essential for every human in daily life. While using it in daily life, it allows user to share personal information and sensitive information while doing some transactions. It render user to be a prime target for cyber crime and activities. Among many threats, phishing is one of the most dangerous, leading cyber attacks to significant financial losses. The phishing websites are designed in such way that it looks exactly similar like genuine websites. It takes advantage of technical awareness, human psychology so that the user discloses sensitive information like login details, bank details and personal information [1].

The increasing frequency of phishing attack and due to its complexity, it become challenging to detection system to handle it. Cyber criminals continuously refine strategies, create exact replica of genuine websites which bypass standard detection methods. Because of the adaptive and dynamic nature of phishing attacks, it requires scalable, robust and intelligent solutions. Conventional approaches which are available have some limitation. Blacklist based systems, depends on pre-identified malicious web URLs. Similarly, heuristic methods also struggle with the complexity and it needs high analysis time for extensive datasets [3].

Fuzzy logic has proved as a trusted framework to address the challenges faced by phishing detection methods. In contrast to the traditional binary decision making systems, fuzzy logic provides effective solution by handling uncertainty and complexity in phishing indicators or features. It assigns degree of membership function to different phishing parameters like URL length, suspicious keyword, special characters, domain age etc. By considering more features and rule base, fuzzy logic enhances the phishing detection process and it provides balanced approach to reduce false positives [7].

This paper present a fuzzy logic-based system for phishing website detection, it consist seven phishing features in its current implementation. These features are URL length, email based URL, suspicious word used, special characters, keyword mismatch domain age, and HTTPS usage. These are selected for their relevance in identifying phishing behavior. Future implementation of the model will expand to include more features, to achieve higher accuracy and adaptability. By leveraging fuzzy rules to classify websites, the presented system bridges the gap between traditional rule-based systems and machine learning approaches; this offers a scalable and reliable solution to the phishing attacks[2][7].

The importance of phishing detection is it extends beyond individual users, impacting organizations and businesses in the all global. Phishing attacks compromise personal data of individual as well as corporate, financial operations, and reputations, highlighting the critical need for effective solution. By advancing detection model through fuzzy logic, this research aims to contribute to a safer online environment and stronger defenses against the most cyber threats.

The different features selected for defining the fuzzy rule base are below:

URL Length: Longer URLs are indication of phishing attempts, as attackers normally use extended URLs to hide malicious intentions.

Keyword Mismatch: The presence of unexpected or misleading keywords in the URL is a strong indication of phishing attacks.

Domain Age: The domain age of few hours or few day, is the strong indicator of phishing attack. Phishing pages are typically active for a short period and the average lifespan of around 21 days.

Protocol Usage (HTTP/HTTPS): Use of HTTPS is very important, normally legitimate websites use secure connections (HTTPS) but Phishing websites frequently use HTTP to avoid use of security certificate.

Presence of Suspicious Words: If URLs containing suspicious words such as "secure", "login", "verify", or company names with slight misspellings are phishing URLs.

Number of Special Characters: The use of high number of special characters, (for example ‘@’, '-', '%', etc.) in a URL is characteristic of phishing URLs.

Email-Based URLs: URLs embedded in emails or containing email formats are more frequently associated with phishing attacks [8] [10].

The entire input features are categorized into linguistic terms using triangular or trapezoidal membership functions, facilitating qualitative interpretation of quantitative values. The membership functions for these features are defined as follows:

Table 1. List of selected features

Parameter	Linguistic Terms
URL Length	*Short*, *Medium*, *Long*
Keyword Mismatch	*Absent*, *Present*
Domain Age	*New*, *Moderate*, *Old*
Protocol Usage	
Suspicious Words	*Absent*, *Present*
Special Characters	*Few*, *Moderate*, *Many*
Email-Based URLs	*Absent*, *Present*

2. Literature Review And Related Work

The motivation or the first aim of people behind phishing is always financial gains. Other than financial gain, few more like malware distribution, identity theft, and industrial espionage etc. these are the other motivating factors for them. Phishing is a very big problem now days, because of its huge impact on the financial loss and personal information loss, so preventing these attacks are very important step now days. Number of Researcher has studied the different characteristics/features of phishing attacks and they provided different detection methods of phishing attacks.

In this section, we have briefly done the survey of different available detection methods and solutions for phishing.

The work done by author Rajeev Kumar Shah, Asif Khan, Mohammad Kamrul Hasan, and others implemented a fuzzy multi-criteria decision-making (FMCDM) approach for phishing detection. The method evaluates different phishing indicators, such as suspicious URLs, SSL certificate issues, and domain age, using ten predefined fuzzy rules. These rules classify input websites into category ranging from "highly legitimate" to "highly phishy." The authors used linguistic variables to represent phishing indicators, with the FMCDM model processing URL data through process of fuzzification, rule evaluation, and defuzzification. In this work the number of selected feature and number of rules used are less. Future enhancement could simplify the design and increase in features to counter sophisticated phishing strategies[1].

The study done by Almaha Abuzuraiq, Mouhammd Alkasassbeh, and Mohammad Almseidin explores methods for detecting phishing websites through machine learning and fuzzy logic techniques. These authors propose a two-stage approach: first, with machine learning algorithms like Random Forest and J48 are used to validate datasets and selecting features, optimizing performance with fewer features. In the second step, fuzzy logic algorithms, i.e. FURIA, are used to classification of phishing websites, leveraging linguistic variables and rule-based systems for detection. The study identifies challenges in complexity, blacklisting, heuristic, and content-based detection methods, it includes low accuracy, dependency on third-party servers, and some inefficiencies. The use of fuzzy logic demonstrates adaptability to ambiguous phishing indicators but it need preprocessing and rule generation. Future work should focus on reducing complexity and more feature selection for better scalability and efficiency [3].

The research by M. Aydin, K. Bicakci, I. Butun, and N. Baykal analyze detecting phishing website URLs using attribute-based feature selection and use of machine learning algorithms. It reduces dataset dimensionality by selecting relevant parameters like URL length, special characters, and domain age, improving model efficiency. In this work, numbers of methods are used for classification such as Random Forest, Decision Trees, and Logistic Regression. In this paper, challenges include balancing detection accuracy and computational efficiency. The authors suggest need to improve feature selection and use of adaptive techniques to address evolving threats in URL fraud detection [4].

The work done by S. R. Ahmad, A. Sharieh, H. Al Bdour, and R. Jabri proposes an enhanced phishing website detection method by combining fuzzy logic and rule-based machine learning. The approach use fuzzy logic to process uncertain data and associative rules to get the patterns of phishing attributes. This method uses phishing indicators such as URL length, domain reputation, and SSL certificate to classify websites. By integrating all variables and data-driven rule generation, the model able to detect phishing attempts. But it shows improved detection accuracy, some challenges include the computational complexity of fuzzy rule and maintaining scalability as phishing techniques used. It needs further changes in rule optimization and feature selection is suggested to improve system performance [10].

The study done by C. Pham, L. A. T. Nguyen, N. H. Tran, E.-N. Huh, and C. S. Hong introduces, a Neuro-fuzzy system for combating phishing in fog networks. This approach integrates fuzzy logic with neural networks to improve phishing detection by leveraging contextual and real-time data processing capabilities of fog computing. The model considers linguistic variables and fuzzy rules to classify URLs as legitimate or phishing while adapting to new threats via neural network training. Key phishing indicators like URL structure, domain reputation, and network activity patterns are used. Despite its strong detection capabilities and real-time processing, this implementation faces some challenges in managing scalability and computational overhead in large-scale fog environments. Enhancement could focus on optimizing resource allocation and extending adaptability to more complex phishing attack [11].

A very simple approach is to stop phishing at the e-mail level, because most phishing attacks use spam i.e. broadcast e-mail. Anti-phishing filters can be used to fight with phishing website at the email level, as it is the primary channel for people behind the phishing website to reach up to the user. Another best approach is use of security toolbars. The Internet explorer 7 is having inbuilt phishing toolbar. It also blocks the user activity, if the website is phishing site. Another approach is to differentiate the phishing sites from the legitimate sites. Another approach is use of two-factor authentication that the user knows a security token. However, two-factor authentication approach is a server-side solution, but Phishing can still happen at those websites those do not support two-factor authentication. Sensitive information like bank details, credit card details is not related to a specific site, so difficult to protect by this approach [12][13].

3. Methodology

Fuzzy logic has long been used in engineering to integrate expert knowledge into computational models across a wide range of applications. One of the advantages of fuzzy logic is, we can use it for ambiguous variables, and for those we cannot use conventional mathematics. We can involve human opinions to establish the relationships and interpretations of such variables. On the other hand, data mining allows us to explore vast datasets to extract meaningful and relevant patterns. It empowers researchers to concentrate on the most significant elements of their data. With the support of data mining tools, organizations can guess future patterns and behaviors, enabling them to make informed, knowledge-based strategic decisions [5][14].

The aim of this methodology for detecting phishing websites is identifying relevant features and employing classification algorithms to enhance phishing detection accuracy. As highlighted, integrating a comprehensive set of phishing characteristics with fuzzy rule base into a fuzzy inference system can significantly improve the classification process. This system uses a structured approach that includes feature selection, fuzzification, rule-based processing, and defuzzification to calculate the phishing probability of a website shown in Figure 2. By using advanced data mining algorithms, the system transforms various phishing-related parameters into crisp value, that crisp value is input to the fuzzy logic controller. The process of fuzzification convert crisp value into fuzzy value, and then with the help of centroid defuzzification method that fuzzy value again converted to crisp value. This crisp value is used to decide phishing rate [2].

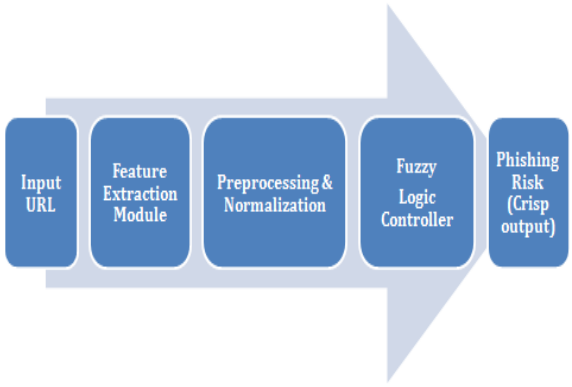

Figure 2. Proposed phishing detection framework using fuzzy logic

Once the system identifies the phishing website, then next step involves designing a proactive response mechanism. As shown in Figure 3, the system we can use to get the details such as URL, IP address, and host location to notify the respective system administrator. The administrator is then responsible to remove the phishing page from the host server, to avoid mitigating potential harm to users [2].

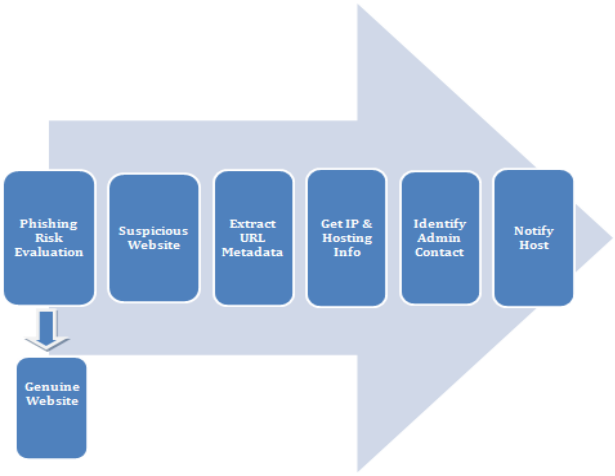

Figure 3. System approach for prevention

This methodology reconciles with earlier research works that integrate with Neuro-fuzzy systems, machine learning techniques, and rule-based models for phishing detection. For example, study by C. Pham et al. and S. R. Ahmad et al. highlight the importance of adaptive systems useful for handling uncertain data and evolving phishing attacks. Similarly, considering methods like feature optimization, used by author M. Aydin et al., confirm agreement between efficiency and accuracy. By dealing with phishing parameter selection comprehensively and enabling preventive mechanisms, this system offers a scalable and robust solution for phishing detection and prevention [4][10].

3.1 Fuzzification

The process of conversion of crisp input into fuzzy output is known as a Fuzzification. This process include assigning linguistic terms—such as Low, Medium, and High—to specific ranges of input phishing indicators. The input is represented by fuzzy sets which represent these linguistic categories. For example, URL length is represented with membership functions with intervals such as short, medium, and long. In the fuzzy logic boundaries are not rigidly or strictly defined. The value which represents a fuzzy set is known as its degree of membership. Each phishing input parameter is associated with a membership function; it is a curve with range between 0 and 1[14].

Below is the example of the linguistic i.e. URL Length, descriptors used to represent one of the key phishing characteristic indicators and a plot of the fuzzy membership functions are shown in figure 4.

The same approach is used to design the other 6 Key Phishing Characteristic Indicators [6].

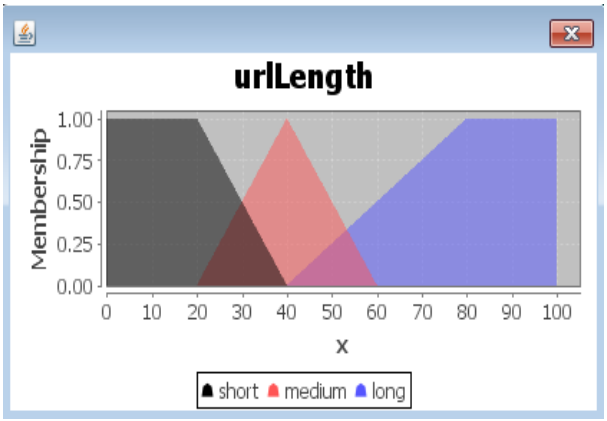

Figure 4. Membership function for URL length

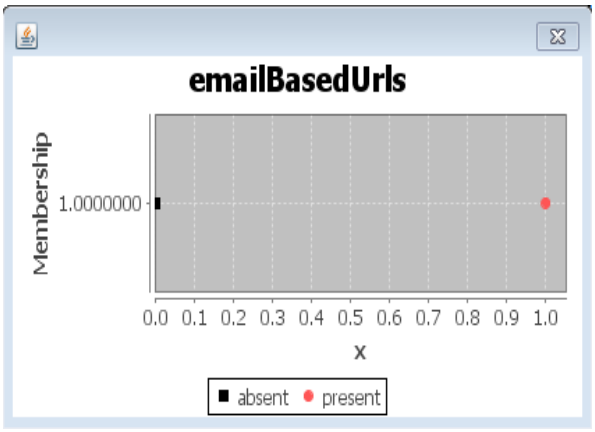

Figure 5. Membership function for email based URLs

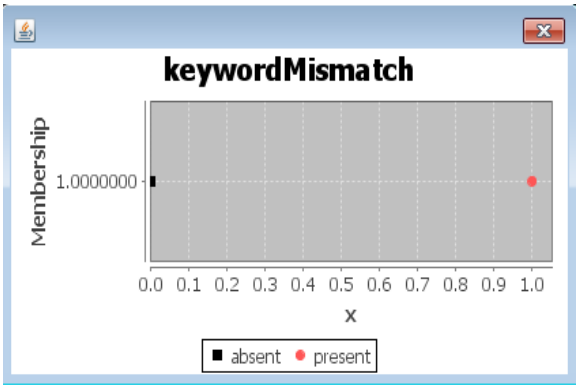

Figure 6. Membership function for keyword Mismatch

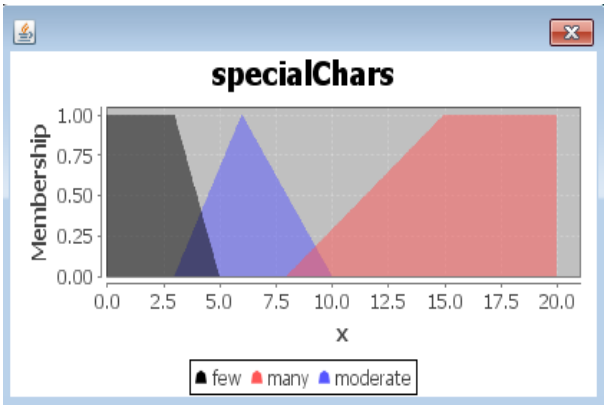

Figure 7. Membership function for Special characters

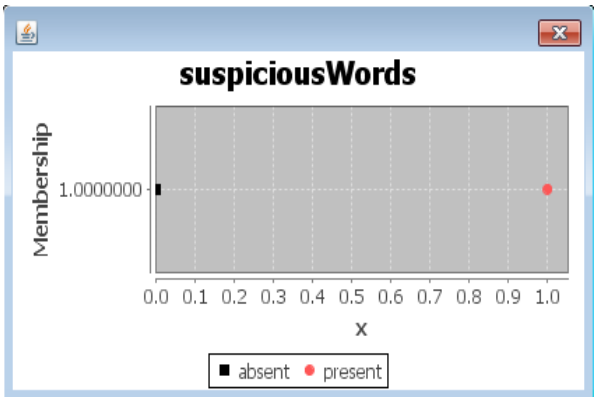

Figure 8. Membership function for suspicious words

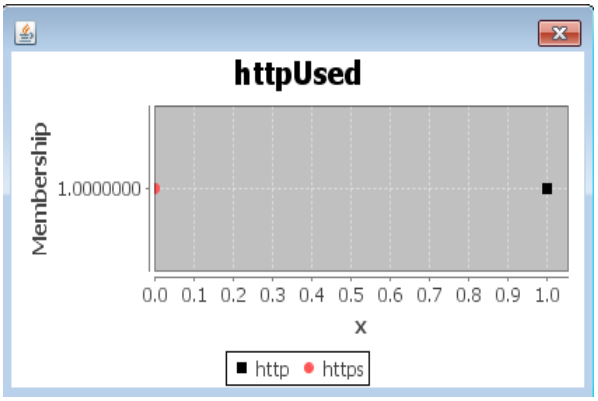

Figure 9. Membership function for Http used

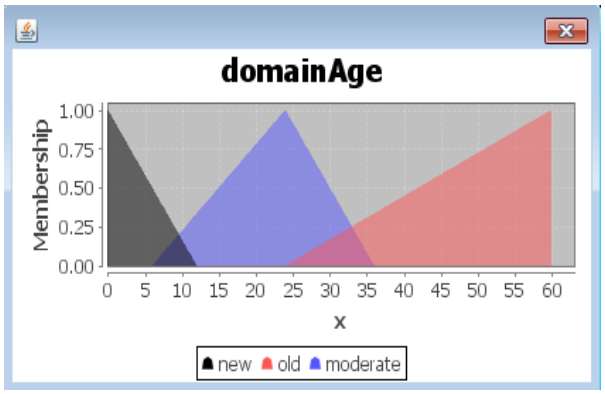

Figure 10: Membership function for domain age

3.2 Rule Base

The effectiveness of a fuzzy based phishing detection system critically depends on it's rule base design, which acts as the decision-making core of the model. The fuzzy rule base is created using expert knowledge and evidence based analysis of phishing websites, which is focusing on key distinguishing parameters. Based on analysis and literature support, we have selected seven features for the construction of the fuzzy rule base in this implementation. These features have been proven to have a significant impact on the URL being associated with phishing attacks.

The output variable, Phishing Risk, is categorized into Low, Medium, and High to provide a graded risk assessment [1][7][9].

3.3 Fuzzy Inference Rules

The fuzzy rule base consists of seventeen (17) well-defined rules created from expert knowledge and phishing website patterns. These rules represent relationships between the input features used and the output phishing risk.

To improve detection accuracy, there are few multiple feature combined rules were incorporated, where two suspicious features together increase the confidence of a phishing prediction:

Table 2. Rule base

Parameters	Rule
URL Length	If URL Length is Short, then Phishing Risk is Low.
	If URL Length is Medium, then Phishing Risk is Medium.
	If URL Length is Long, then Phishing Risk is High.
Keyword Mismatch	If Keyword Mismatch is Present, then Phishing Risk is High.
Domain Age	If Domain Age is New, then Phishing Risk is High.
	If Domain Age is Old, then Phishing Risk is Low.
HTTP used	If Protocol Used is HTTP, then Phishing Risk is High
	If Protocol Used is HTTPS, then Phishing Risk is Low.
Suspicious Words	If Suspicious Words are Present, then Phishing Risk is High.
	If Suspicious Words are Absent, then Phishing Risk is Low.
Special Characters	If Special Characters are Many, then Phishing Risk is High.
	If Special Characters are Few, then Phishing Risk is Low.
Email-Based URLs	If Email-Based URLs are Present, then Phishing Risk is High.
	If Email-Based URLs are Absent, then Phishing Risk is Low.
Combined Multi-Feature Rules	If URL Length is Long *and* Suspicious Words are Present, then Phishing Risk is High.
	If Keyword Mismatch is Present *and* Domain Age is New, then Phishing Risk is High.
	If Protocol Used is HTTP *and* Suspicious Words are Present, then Phishing Risk is High.

These combined rules make sure that the system captures complex phishing behaviors more effectively by considering interdependencies between feature used.

3.4 Defuzzyfication

In this research, the Center of Gravity (COG) method has been employed for defuzzification. The COG method is used to calculate the center of mass of the fuzzy output set to obtain a crisp value. A mathematical formula to represent it is as follows:

$x^* = \int x \cdot \mu(x)\, dx / \int \mu(x)\, dx$

This method ensures that the final phishing risk score is computed as a weighted average, providing a balanced and accurate risk assessment based on fuzzy inference.

Using the COG method, the final phishing risk score is calculated by:

$y^* = \sum y_i \cdot \mu(y_i) / \sum \mu(y_i)$

Where **yi** represents possible risk values and **μ(yi)** is their respective membership values.

Centre of gravity method considers the entire fuzzy output range rather than just selecting the highest membership, which is the reason to get more accurate result. It provides a weighted center; make sure that all contributing rules influence the final crisp value [5][6].

It is a Mathematically Robust which is a continuous function and ensures smooth transitions between different fuzzy states.

The Output membership function for phishing risk is show in figure 11.

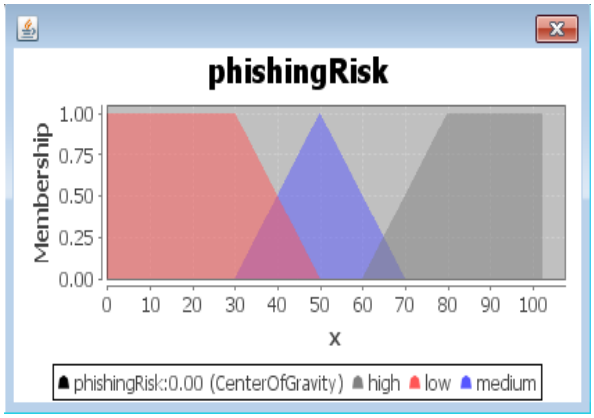

Figure 11. Output Membership function for Phishing Risk

4. Results And Discussion

The implemented fuzzy logic-based phishing URL detection model performance was calculated using a dataset which consist 100 URLs, which are collected from publicly available phishing databases and legitimate website listings. The collected dataset consisted of 50 phishing URLs and 50 legitimate URLs, ensuring a balanced class distribution.

Evaluation Metrics

Standard classification metrics were used, for the evaluation of model performance:

- Accuracy (Acc):

$$Acc = (TP + TN) / (TP + TN + FP + FN)$$

- Precision (P):

$$P = TP / (TP + FP)$$

- Recall (R) (or Sensitivity):

$$R = TP / (TP + FN)$$

- F1-Score (F1):

$$F1 = 2 * P * R / (P + R)$$

- Specificity (Sp):

$$Sp = TN / (TN + FP)$$

Where:

- *TP*: True Positives (Phishing website correctly identified)
- *TN*: True Negatives (Legitimate website correctly identified)
- *FP*: False Positives (Legitimate website misclassified as phishing)
- *FN*: False Negatives (Phishing website misclassified as legitimate)

After implementing this model, following performance was achieved:

Table 3. Results

Metric	Value
Accuracy	**96.18%**
Precision	95.83%
Recall	96.36%
F1-Score	96.09%
Specificity	95.81%

This result highlights the robustness and effectiveness of the fuzzy rule base, demonstrating that the selected features sufficiently capture the phishing characteristics in URLs.

5. Conclusion

In this research, a fuzzy logic-based phishing website detection system was implemented by selecting different features such as URL length, keyword mismatch, domain age, HTTP used, suspicious keywords, special characters, and URLs with email. By applying linguistic variables and simple fuzzy rules, the system provided a robust risk assessment of URLs. The Center of Gravity (COG) defuzzication method was used to convert fuzzy to crisp phishing scores. The experimental results demonstrated that the model performs with high accuracy and low false positives, and it is highlighting its effectiveness in detecting phishing threats with minimal overhead.

6. Future Work

The current fuzzy logic system implementation good results, but few enhancements can be considered in future research to further improve phishing detection result:

- By using additional phishing features such as SSL certificate analysis, domain WHOIS information, and lexical URL analysis could improve detection result.
- By using hybrid approaches like combining fuzzy logic and machine learning models to predictive power.
- Use of the model in a real-time phishing URL detection system or as part of browser tools/security plug-in, to provide actively security to users during web browsing.
- Use of some adaptive fuzzy logic or neuro-fuzzy systems could help in automatically modifying rules based on new phishing attacks.
- Expansion of experiments to more comprehensive, real-world datasets would validate the performance, scalability and robustness of the implemented model.
- By addressing these future directions, the implemented work can be enhanced into a more comprehensive solution for phishing website detection, contributing meaningfully to the broader cyber security system.

References:

1. R. K. Shah, A. Khan, M. K. Hasan, T. M. Ghazal, S. Islam, and A. N. Khan, "Detect Phishing Website by Fuzzy Multi-Criteria Decision Making," in Proceedings of the 1st International Conference on AI in Cybersecurity (ICAIC), 2022, DOI: 10.1109/ICAIC53980.2022.9897036.

2. M. D. Bhagwat, P. H. Patil and T. S. Vishawanath, "A Methodical Overview on Detection, Identification and Proactive Prevention of Phishing Websites," 2021 Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV), Tirunelveli, India, 2021, pp. 1505-1508, doi: 10.1109/ICICV50876.2021.9388441.
3. A. Abuzuraiq, M. Alkasassbeh, and M. Almseidin, "Intelligent Methods for Accurately Detecting Phishing Websites," in 2020 11th International Conference on Information and Communication Systems (ICICS), 2020, pp. 085-089, DOI: 10.1109/ICICS49469.2020.239509.
4. M. Aydin, K. Bicakci, I. Butun, and N. Baykal, "Using Attribute-based Feature Selection Approaches and Machine Learning Algorithms for Detecting Fraudulent Website URLs," 2020 IEEE International Conference on Communication, Networks and Satellite (ComNetSat), pp. 123-128, 2020, DOI: 10.1109/ComNetSat49409.2020.9201235.
5. Ms. Roshni Vitthal Pawar, Ms. Ruchita Pandit Rao Pawar, Ms.Pranali Ganesh Salunkhe and Ms.Ankita Gajanan Sankhe, “Phishing Identification Using An Efficient Neuro-Fuzzy Model”, Volume 2,Issue 4, April 2017,ISSN (online): 2456-0006 International Journal of Science Technology Management and Research.
6. P. Cingolani and J. Alcalá-Fdez, "jFuzzyLogic: A Java Library to Design Fuzzy Logic Controllers According to the Standard for Fuzzy Control Programming," University of Granada, Available: jFuzzyLogic SourceForge.
7. S. D. Shirsat, "Demonstrating Different Phishing Attacks Using Fuzzy Logic," Proceedings of the 2nd International Conference on Inventive Communication and Computational Technologies (ICICCT), IEEE Xplore Compliant, 2018, DOI: 10.1109/ICICCT.2018.12345.
8. S. Paliwal, D. Anand, and S. Khan, "Application of Rule-based Fuzzy Inference System in Prediction of Internet Phishing," International Journal of Computer Applications, vol. 148, no. 14, pp. 30-34, Aug. 2016. Hindustan Institute of Technology and Management, Agra, APJ Abdul Kalam Technical University, Lucknow.
9. S. R. Zahra, M. A. Chishti, A. I. Baba, and F. Wu, "Detecting Covid-19 Chaos Driven Phishing/Malicious URL Attacks by a Fuzzy Logic and Data Mining Based Intelligence System," Egyptian Informatics Journal, vol. 23, no. 2, pp. 197–214, 2022.
10. S. R. Ahmad, A. Sharieh, H. Al Bdour, and R. Jabri, "Enhance Detecting Phishing Websites Based on Machine Learning Techniques of Fuzzy Logic with Associative Rules," Computer Science Department, KASIT, The University of Jordan, Amman, Jordan, January 2017.
11. C. Pham, L. A. T. Nguyen, N. H. Tran, E. -N. Huh and C. S. Hong, "Phishing-Aware: A Neuro-Fuzzy Approach for Anti-Phishing on Fog Networks," in IEEE Transactions on Network and Service Management, vol. 15, no. 3, pp. 1076-1089, Sept. 2018, doi: 10.1109/TNSM.2018.2831197.
12. Bashir, Tenuche, B. C. Agbata, E. Ogala, and W. Obeng-Denteh, "The Fuzzy Experiment Approach for Detection and Prevention of Phishing Attacks in Online Domain," East African Scholars Journal of Engineering and Computer Sciences, vol. 3, no. 10, pp. 276–282, 2020. DOI: 10.36349/easjecs.2020.v03i10.001.
13. R. M. Abdul-Hussein, A. H. Mohammed, and A. A. Kadhim, "Detecting Phishing Cyber Attack Based on Fuzzy Rules and Differential Evaluation," TEM Journal, vol. 11, no. 2, pp. 543–551, May 2022. DOI: 10.18421/TEM112-07.
14. Maher Aburrous, M.A. Hossain, Keshav Dahal, Fadi Thabtah, Intelligent phishing detection system for e-banking using fuzzy data mining, Expert Systems with Applications, Volume 37, Issue 12, 2010, Pages 7913-7921, ISSN 0957-4174, https://doi.org/10.1016/j.eswa.2010.04.044.
15. Kumar, V., & Dalal, S. (2026). Federated Learning for Privacy-Preserving Optimization in Multi-Domain Optical Networks. International Journal of Recent Advances in Engineering and Technology, 15(1), 132–142.
16. Kumar, S., Prakash, S., Kumar, A., Mamta, Kumari, R., & Vijay. (2026). Cybersecurity in AI-Enabled IoT Network: Threat Detection and Mitigation Using Deep Learning. International Journal on Advanced Computer Engineering and Communication Technology, 15(1), 135–145.