PAR - Assignment 3 - Report

Bruno Sánchez & Jean Dié

15th December 2024

Contents

1 Introduction **3**

2 Task 1: Definition of Linguistic Variables **3**

2.1 Input Linguistic Variables . . . 3

2.1.1 Ratio of Null Hyperlinks . . . 3

2.1.2 Number of External CSS Files . . . 4

2.1.3 Domain Registration Length . . . 4

2.1.4 Web Traffic . . . 5

2.1.5 Page Rank . . . 5

2.2 Output Linguistic Variable . . . 6

3 Task 2: Definition of the Rule Base **7**

3.1 Rule Development . . . 7

3.2 Complete Rule Set . . . 8

3.3 Rule Set Design Rationale . . . 8

4 Task 3: Implementation Using MATLAB Fuzzy Toolkit **9**

4.1 System Configuration . . . 9

4.2 Validation . . . 11

5 Task 4: Testing the System **11**

5.1 Test Cases . . . 12

5.2 Results and Discussion . . . 13

6 Task 5: Design of an Enhanced Fuzzy Expert System **14**

7 Conclusion **15**

1 Introduction

2 Task 1: Definition of Linguistic Variables

In this task, we define the input and output linguistic variables for our fuzzy expert system designed to detect phishing on websites. Based on the features identified in [1], we have selected a subset of five features, each with different reference scales and units, to serve as input variables. For each variable, we have determined the number of terms, labels, and corresponding fuzzy sets using trapezoidal membership function in the (a, b, c, d) format, ensuring they satisfy the property of Fuzzy Partition.

2.1 Input Linguistic Variables

For our fuzzy phishing detection system, we have carefully selected five input linguistic variables that capture key characteristics of websites commonly exploited in phishing attacks. Each variable represents a different aspect of website legitimacy, from structural elements to reputation indicators. The variables have been chosen to provide complementary signals while maintaining interpretability. We specifically selected non-binary features because binary features cannot be effectively represented using triangular or trapezoidal membership functions.

2.1.1 Ratio of Null Hyperlinks

The variable f60 measures the percentage of non-functional or empty links within a webpage. As identified in [1], this feature is particularly significant for phishing detection as it reflects both the structural integrity and maintenance quality of a website. Legitimate websites typically maintain their hyperlinks carefully, while phishing sites often contain numerous broken or placeholder links due to hasty creation or poor maintenance. The range spans from 0 to 100 percent and is characterized by three linguistic terms with trapezoidal membership functions:

Linguistic Term	**Range**	**Argumentation**
Low	[0, 0, 15, 25]	Typical of legitimate websites with properly maintained links. A low ratio of null hyperlinks suggests that the website is well-maintained and regularly updated, reducing the likelihood of it being a phishing site.
Moderate	[15, 25, 35, 45]	Indicates potential suspicious activity. A moderate ratio of null hyperlinks may suggest that the website has some issues with link maintenance, which could be a sign of a less reputable site or one that is not regularly updated.
High	[35, 45, 100, 100]	Strong indicator of a phishing webpage. A high ratio of null hyperlinks is a red flag, as it suggests that the website may have been hastily created with little regard for link validity, a common trait of phishing sites.

Table 1: Linguistic Terms for Ratio of Null Hyperlinks

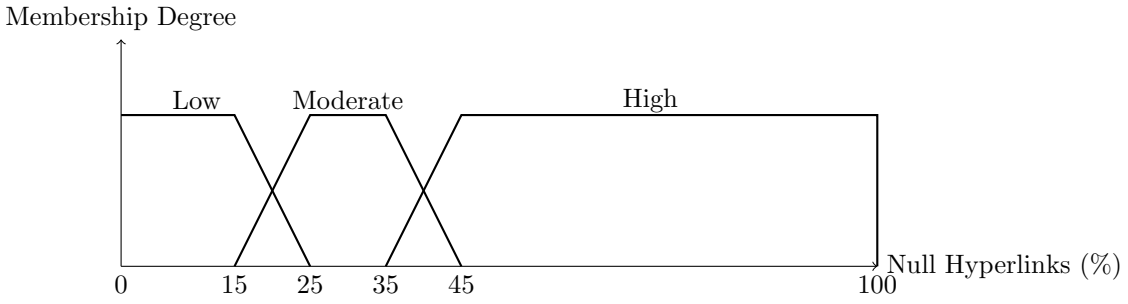


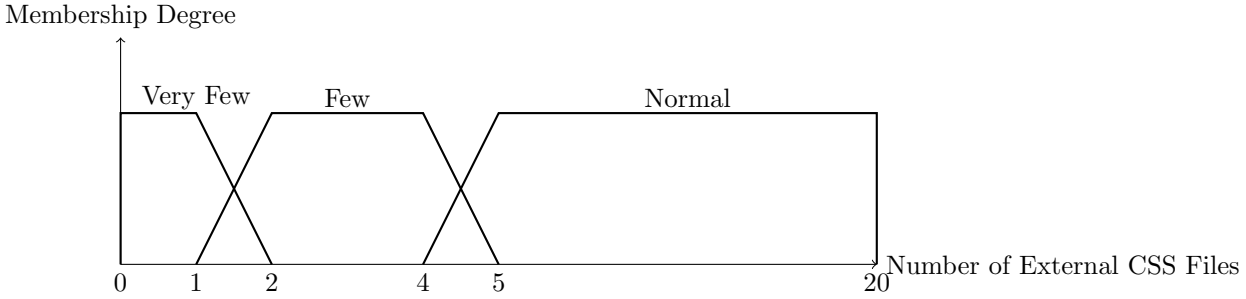
Figure 1: Membership Functions for Ratio of Null Hyperlinks

2.1.2 Number of External CSS Files

The variable f61 quantifies external stylesheet references in a webpage. This feature is particularly relevant for phishing detection as legitimate websites typically maintain consistent styling through multiple CSS files, following modern web development practices. Phishing websites, in contrast, often minimize their styling implementation to quickly replicate legitimate pages with minimal effort. Additionally, as noted in [2], the number of external stylesheets can indicate the level of sophistication and development effort invested in the website, where an unusually low number might signal a hastily created phishing page. The range spans from 0 to 20 files, with three linguistic terms:

Linguistic Term	**Range**	**Argumentation**
Very Few	[0, 0, 1, 2]	Common in phishing sites that employ minimal styling to create basic replicas of legitimate pages. A single CSS file often indicates oversimplified implementation that warrants scrutiny.
Few	[1, 2, 4, 5]	Represents balanced styling implementation. This range aligns with modern web development practices where CSS is modular but optimized for performance.
Normal	[4, 5, 20, 20]	Characteristic of properly developed sites with comprehensive styling needs. The upper limit reflects current best practices in CSS organization while acknowledging that excessive CSS files may indicate poor optimization.

Table 2: Linguistic Terms for Number of External CSS Files

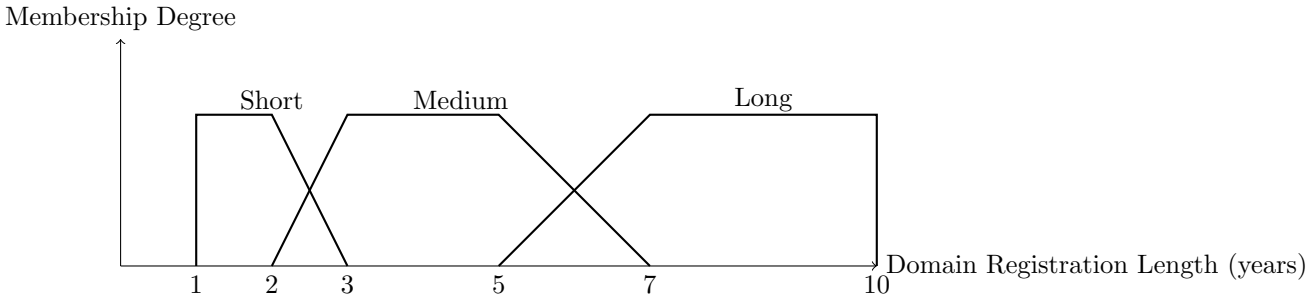


2.1.3 Domain Registration Length

The variable f82 reflects the duration of domain registration in years. According to [3], domain registration length is a crucial temporal feature for phishing detection, as legitimate websites tend to be registered for longer periods while phishing domains are typically short-lived. As noted in [2], attackers often register domains for the minimum required period to minimize costs and avoid detection, making registration length a strong indicator of potentially malicious intent. This pattern emerges because phishing campaigns are typically brief, targeted operations that don't require long-term domain maintenance. The range spans from 1 to 10 years and is characterized by the following:

Linguistic Term	**Range**	**Argumentation**
Short	[1, 1, 2, 3]	Common for phishing domains. Short registration lengths are often used by phishing sites to avoid detection and minimize costs.
Medium	[2, 3, 5, 7]	Moderate commitment to domain maintenance. Medium registration lengths may indicate a more established site, but not necessarily a highly reputable one.
Long	[5, 7, 10, 10]	Typical of established legitimate websites. Long registration lengths suggest a strong commitment to maintaining the domain, which is characteristic of legitimate websites.

Table 3: Linguistic Terms for Domain Registration Length



2.1.4 Web Traffic

The variable f84 measures the percentile rank of a website's traffic. Web traffic is a significant indicator of website legitimacy as it reflects the site's established presence and user trust. As noted in [2], phishing websites typically exhibit very low traffic patterns due to their short lifespan and targeted nature, while legitimate websites maintain consistent traffic levels over time. This metric is particularly valuable because it's difficult for attackers to artificially inflate genuine user traffic, making it a reliable indicator of legitimacy. The range spans from 0 to 100 percentiles and is characterized by four linguistic terms with trapezoidal membership functions:

Linguistic Term	**Range**	**Argumentation**
Very Low	[0, 0, 10, 20]	Indicative of low visibility and potential risk. Websites with very low traffic are less likely to be well-known and may be newly created or less reputable.
Low	[10, 20, 60, 70]	Suggests limited reach and moderate risk. Low traffic websites may be legitimate but are not widely recognized, which can be a characteristic of less established sites.
Moderate	[60, 70, 80, 90]	Represents average visibility and lower risk. Moderate traffic websites are more likely to be established and have a consistent user base, indicating a higher likelihood of legitimacy.
High	[80, 90, 100, 100]	Characteristic of popular and reputable websites. High traffic websites are well-known and widely visited, reducing the likelihood of them being phishing sites.

Table 4: Linguistic Terms for Web Traffic

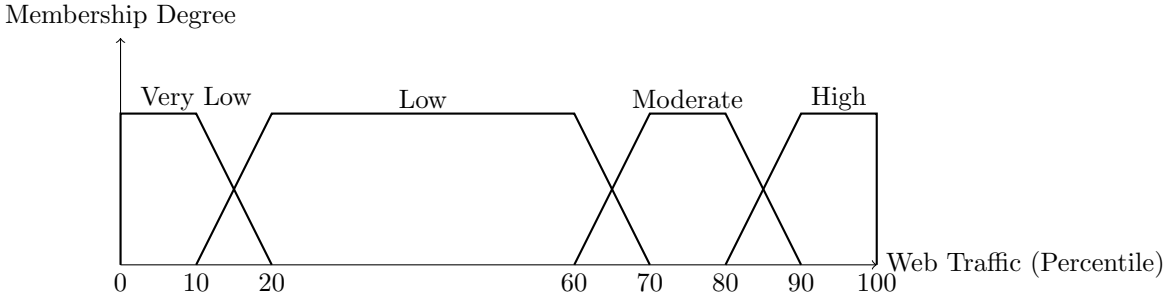


Figure 2: Membership Functions for Web Traffic

2.1.5 Page Rank

The variable f87 measures the popularity score of a website, ranging from 0 to 10. This feature is meaninglful for phishing detection as it quantifies a website's reputation based on its relationships with other sites across the internet. While legitimate websites naturally accumulate higher page ranks through genuine backlinks and long-term presence, phishing sites typically have very low page ranks due to their recent creation and lack of legitimate connections. Unlike many other metrics, page rank is particularly difficult to manipulate as it requires building a genuine network of trusted referring sites. The range spans from 0 to 10 and is characterized by three linguistic terms with trapezoidal membership functions:

Linguistic Term	Range	Argumentation
Popular	[0, 0, 2, 3]	Indicative of highly reputable websites. A high page rank suggests that the website is well-known and widely trusted, reducing the likelihood of it being a phishing site.
Moderately Known	[2, 3, 5, 7]	Represents websites with moderate popularity. These sites are somewhat known and trusted, but not as widely recognized as those with higher page ranks.
Not Very Known	[5, 7, 10, 10]	Characteristic of less reputable websites. A low page rank indicates that the website is not widely known or trusted, which could be a sign of a phishing site.

Table 5: Linguistic Terms for Page Rank

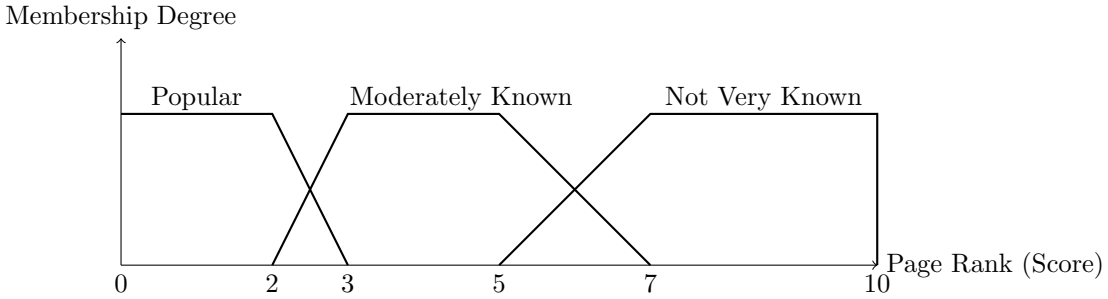


Figure 3: Membership Functions for Page Rank

2.2 Output Linguistic Variable

The output variable measures the phishing risk of a website, ranging from 0 to 100 percent. We define four fuzzy categories: safe, weakly suspicious, strongly suspicious, and phishing. Each category is represented by a trapezoidal membership function.

Fuzzy Category	Range	Argumentation
Safe	[0, 0, 15, 25]	Indicates a very low risk of phishing. Websites in this category are considered safe and trustworthy.
Weakly Suspicious	[15, 25, 35, 45]	Suggests a low risk of phishing. These websites may have some minor issues but are generally considered safe.
Strongly Suspicious	[45, 55, 65, 75]	Represents a moderate risk of phishing. Websites in this category exhibit several suspicious characteristics and warrant caution.
Phishing	[65, 75, 100, 100]	Indicates a high risk of phishing. Websites in this category are highly likely to be phishing sites and should be avoided.

Table 6: Fuzzy Categories for Phishing Risk

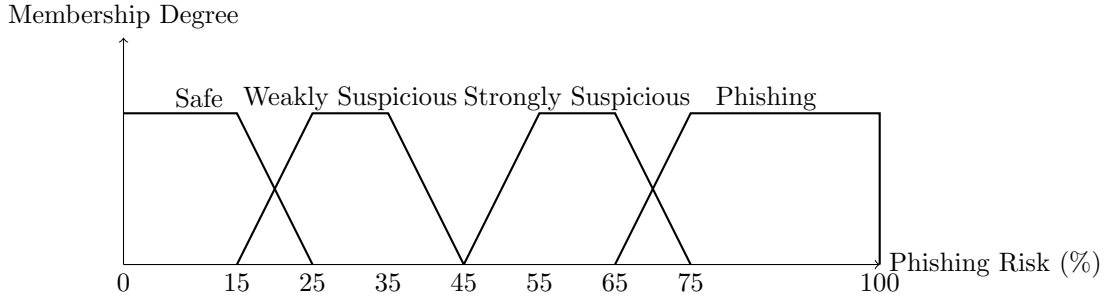


Figure 4: Membership Functions for Phishing Risk

3 Task 2: Definition of the Rule Base

3.1 Rule Development

The rule base for the fuzzy expert system was developed using a systematic methodology to accurately assess phishing risk while maintaining manageability. Five critical input variables were identified: Ratio of Null Hyperlinks, Number of External CSS Files, Domain Registration Length, Web Traffic, and Page Rank. Each variable was defined with appropriate linguistic terms and membership functions to handle uncertainty. Conjunctive rules were formulated by logically combining these variables' linguistic terms to reflect how different conditions contribute to phishing risk; for example, a high ratio of null hyperlinks and a short domain registration length indicate a potential phishing site. The output variable, Phishing Risk, was specified using linguistic terms ranging from Very Low to Very High based on the expected influence of input conditions. To cover a wide range of scenarios, including both typical cases and edge conditions, the rule set includes rules of varying lengths and complexities, enhancing the system's flexibility and responsiveness. Redundant or conflicting rules were eliminated to maintain the integrity of the rule base. The total number of rules was limited to 35 for manageability.

3.2 Complete Rule Set

#	Null Links	Ext. CSS	Domain Reg.	Traffic	Page Rank	Weight	Risk
1	High	*	*	*	*	0.50	**Phishing**
2	Moderate	Few	Short	Very Low	Not Very Known	0.85	**Phishing**
3	Moderate	Few	Short	Very Low	Moderately Known	0.85	**Phishing**
4	Moderate	Few	Medium	Very Low	Not Very Known	0.80	**Phishing**
5	Moderate	Few	Medium	Very Low	Moderately Known	0.80	**Phishing**
6	Moderate	Normal	Short	Very Low	Not Very Known	0.75	**Phishing**
7	Moderate	Normal	Short	Very Low	Moderately Known	0.75	**Phishing**
8	Moderate	Normal	Medium	Very Low	Not Very Known	0.70	**Phishing**
9	Moderate	Normal	Medium	Very Low	Moderately Known	0.70	**Phishing**
10	*	Very Few	*	*	*	0.50	**Strongly Suspicious**
11	Moderate	Few	Short	Low	Not Very Known	0.85	**Strongly Suspicious**
12	Moderate	Few	Short	Low	Moderately Known	0.85	**Strongly Suspicious**
13	Moderate	Few	Medium	Low	Not Very Known	0.80	**Strongly Suspicious**
14	Moderate	Few	Medium	Low	Moderately Known	0.80	**Strongly Suspicious**
15	Moderate	Normal	Short	Low	Not Very Known	0.75	**Strongly Suspicious**
16	Moderate	Normal	Short	Low	Moderately Known	0.75	**Strongly Suspicious**
17	Moderate	Normal	Medium	Low	Not Very Known	0.70	**Strongly Suspicious**
18	Moderate	Normal	Medium	Low	Moderately Known	0.70	**Strongly Suspicious**
19	Moderate	Few	Short	Moderate	Not Very Known	0.85	**Weakly Suspicious**
20	Moderate	Few	Short	Moderate	Moderately Known	0.85	**Weakly Suspicious**
21	Moderate	Few	Medium	Moderate	Not Very Known	0.80	**Weakly Suspicious**
22	Moderate	Few	Medium	Moderate	Moderately Known	0.80	**Weakly Suspicious**
23	Moderate	Normal	Short	Moderate	Not Very Known	0.75	**Weakly Suspicious**
24	Moderate	Normal	Short	Moderate	Moderately Known	0.75	**Weakly Suspicious**
25	Moderate	Normal	Medium	Moderate	Not Very Known	0.70	**Weakly Suspicious**
26	Moderate	Normal	Medium	Moderate	Moderately Known	0.70	**Weakly Suspicious**
27	*	Few	*	High	*	0.50	**Safe**
28	Low	*	*	*	*	0.50	**Safe**
29	*	*	Long	*	*	0.50	**Safe**
30	*	*	*	*	Popular	0.50	**Safe**
31	Moderate	Normal	Short	High	Not Very Known	0.85	**Safe**
32	Moderate	Normal	Short	High	Moderately Known	0.85	**Safe**
33	Moderate	Normal	Medium	High	Not Very Known	0.80	**Safe**
34	Moderate	Normal	Medium	High	Moderately Known	0.80	**Safe**
35	*	Few	*	High	*	0.50	**Safe**

Table 7: Complete Rule Set for Phishing Detection with Weights

3.3 Rule Set Design Rationale

The development of our fuzzy rule set was guided by a comprehensive analysis of the practical constraints and resources required for attackers to manipulate various website features. By understanding the cost-benefit relationships from an attacker's perspective, we were able to establish a hierarchy of feature importance and create rules that reflect real-world attack patterns.

Our approach began by analyzing each feature through the lens of attacker constraints, considering the financial costs, technical complexity, and time investments required to manipulate them. This analysis directly influenced

how we weighted and combined features in our rule set.

Domain registration length emerged as a significant indicator due to its direct financial implications. Attackers typically aim to minimize operational costs, making long-term domain registrations financially unattractive. Legitimate businesses often register domains for extended periods, while phishing operations tend to use short-term registrations to reduce expenses. However, we recognize that this feature alone is insufficient for classification, as demonstrated by our rules where short registration periods are offset by other positive indicators.

Web traffic presents a more complex scenario for classification. While it can be artificially inflated through various means, including advertising campaigns, bot networks, and embedding techniques, these methods require financial investment. This understanding influenced our rule design, where high traffic alone does not guarantee a safe classification without corroborating positive indicators from other features. Attackers can potentially manipulate traffic through paid advertising or automated tools, making it a valuable but not definitive indicator.

Page rank represents a more resilient metric due to its multifaceted nature. While traffic manipulation can influence page rank, achieving high rankings in search engine results requires sophisticated SEO strategies and sustained effort. The technical expertise and time investment required for SEO manipulation makes this feature particularly valuable in our classification system, especially when combined with other positive indicators.

The ratio of null hyperlinks serves as a crucial technical indicator. Modern web scraping tools like HTTrack [4] and GoClone [5] can easily copy website structures, but often result in broken or placeholder links (*href="#"*). This pattern emerges when attackers hastily replicate legitimate sites without properly maintaining link functionality. Our rules heavily weight this feature, as high ratios of null hyperlinks strongly suggest automated copying typical of phishing attempts.

External CSS file count provides insight into website development practices. While tools can copy entire CSS files, maintaining multiple external stylesheets represents a level of development sophistication typically associated with legitimate websites. Phishing sites often consolidate CSS to simplify deployment, making the number of external stylesheets a useful discriminator between legitimate and malicious sites.

These understandings directly shaped our rule set structure. High-risk classifications (Phishing) typically require multiple high-cost negative indicators, such as a high ratio of null hyperlinks combined with short domain registration and very few CSS files. Moderate-risk classifications (Strongly Suspicious) often involve mixed signals, where some high-cost features appear legitimate while others show suspicious patterns. Low-risk classifications (Safe) generally require multiple positive indicators that would be costly or time-consuming for attackers to fake simultaneously.

4 Task 3: Implementation Using MATLAB Fuzzy Toolkit

Following the approach defined throughout the previous sections, the Fuzzy Expert System was implemented using *MATLAB's Fuzzy Logic Designer Toolkit*. We defined five input variables and one output variable, each with their respective membership functions, modeled using the trapezoidal functions detailed in Section 2. Additionally, we constructed the full set of 35 rules described in Section 3, ensuring comprehensive coverage of the system's logic.

Our implementation employs the *minimum* (min) operator as the t-norm for conjunctions and the *maximum* (max) operator as the t-conorm for disjunctions. These operators are selected to ensure consistency with standard fuzzy logic principles while maintaining computational simplicity.

For defuzzification, we utilized the *Center of Area* (CoA) method, as it provides a precise and interpretable output. This method calculates the centroid of the aggregated fuzzy set to generate the crisp output value, ensuring a balanced representation of all contributing rules.

We validated the system using 3D surface plots generated by MATLAB's visualization tools. These plots depict the relationships between input variables and the output, offering a clear representation of how the defined rules influence the system's behavior. This step verified the correctness of the rule base and the effectiveness of the membership functions.

4.1 System Configuration

The fuzzy inference system was set up using MATLAB's *Fuzzy Logic Designer Toolkit* with the following configuration:

- **Type**: Mamdani Type-1

- **Input Variables**: 5
  - Ratio of Null Hyperlinks
  - Number of External CSS
  - Domain Registration Length
  - Web Traffic
  - Page Rank
- **Output Variable**: Phishing Risk
- **Membership Functions**:
  - Each input variable is modeled with **3 or 4 membership functions** using trapezoidal shapes.
  - The output variable is modeled with **4 membership functions**.
- **Rule Base**: A total of **35 rules** were defined.
- **Operators**:
  - *AND Method*: Minimum (t-norm)
  - *OR Method*: Maximum (t-conorm)
  - *Implication Method*: Minimum
  - *Aggregation Method*: Maximum
- **Defuzzification Method**: Center of Area (centroid)

The configuration and connections between the variables are illustrated in Figure 5.

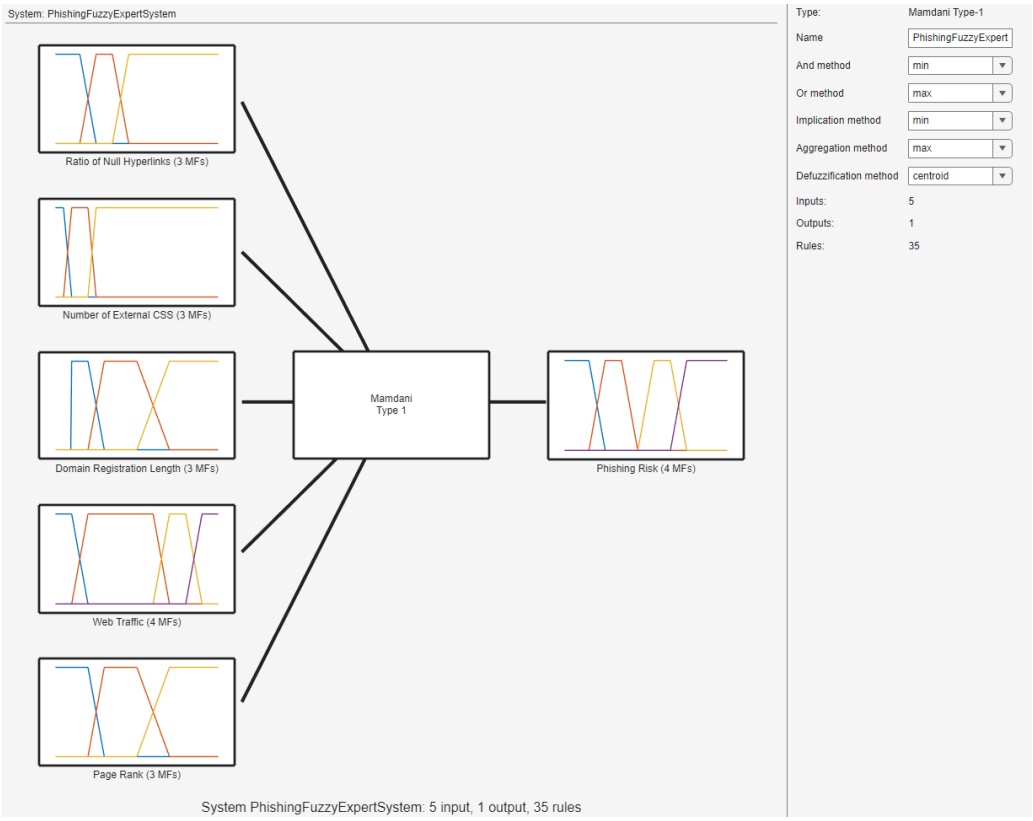


Figure 5: System Configuration in MATLAB's Fuzzy Logic Designer Toolkit

4.2 Validation

The validation of the fuzzy expert system was performed using the 3D Control Surface plots generated by the Fuzzy Logic Designer Toolkit, which depict the output variable (*Phishing Risk*) as a function of two input variables at a time. These plots are generated based on the defined set of 35 rules.

Several 3D plots were analyzed, each illustrating the relationship between different pairs of input variables and the output. An important argument validating the system's behavior is that the resulting surfaces exhibit **monotonicity**. This indicates that the output (*Phishing Risk*) consistently increases or decreases as the input variables change, aligning with expectations based on the rule base.

Four of these 3D plots are presented in Figure 6 for visualization.

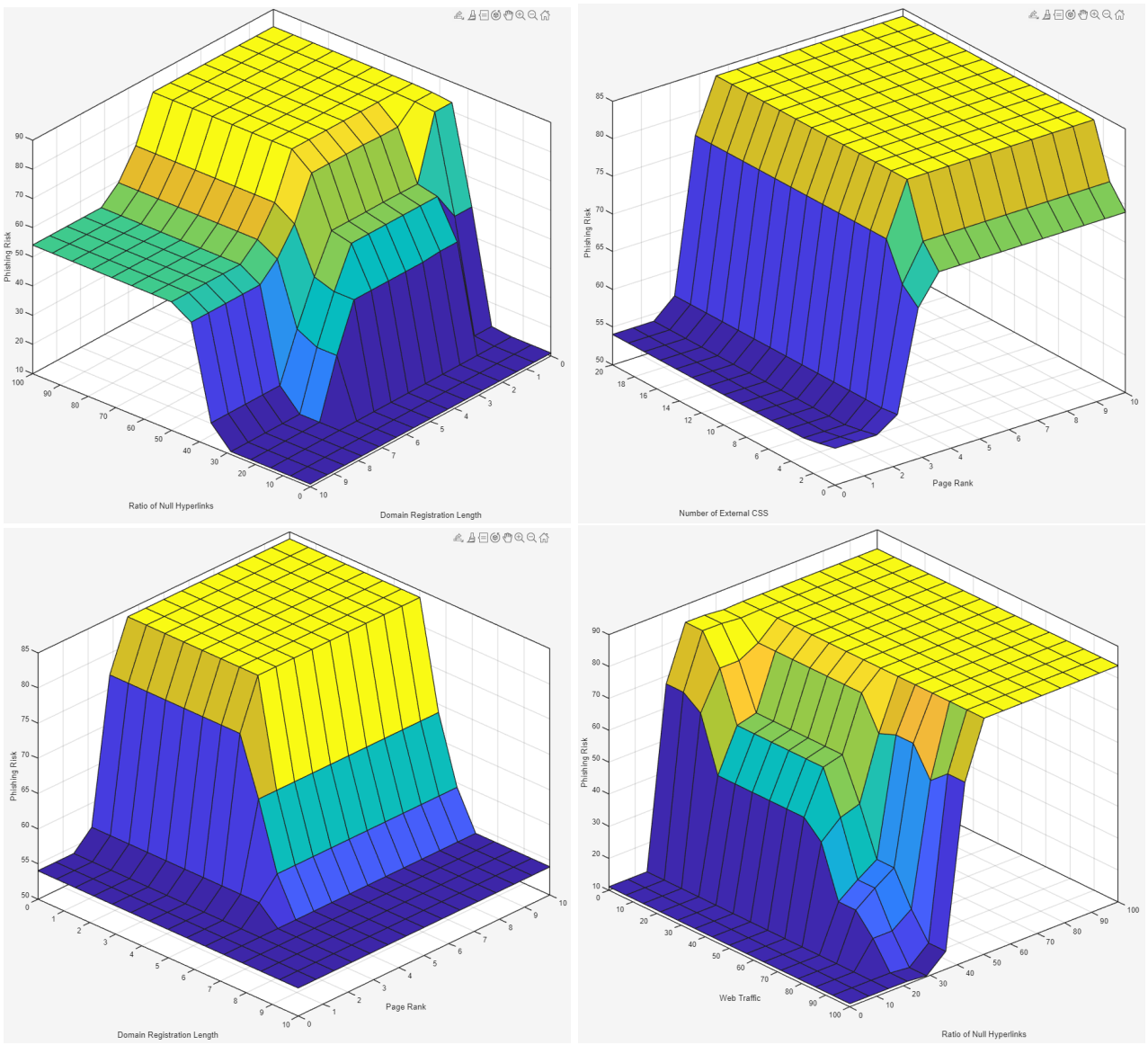


Figure 6: 3D surface plots of the output variable as a function of input variable pairs.

5 Task 4: Testing the System

To evaluate the performance and reliability of the fuzzy expert system, we conducted tests using four different websites. These test cases were carefully created to represent various scenarios, ensuring a diverse range of inputs and outputs. The objective was to analyze the system's behavior under different conditions and validate that the output aligns with expectations.

Each test case includes specific input values for the five input variables, with some cases intentionally designed to activate more than one membership label for the same variable. This approach ensures the system utilizes

multiple rules to compute the output (*Phishing Risk*), demonstrating the effectiveness of the rule base and inference mechanism.

5.1 Test Cases

The following test cases describe the input values and expected outcomes for the fuzzy expert system:

- **Test Case 1: Legitimate Website**

  This case represents a legitimate website with benign characteristics. The input values were selected to reflect a low likelihood of phishing activity:

  - *Ratio of Null Hyperlinks*: 5%
  - *Number of External CSS*: 15 files
  - *Domain Registration Length*: 8 years
  - *Web Traffic*: 85-th percentile
  - *Page Rank*: 2.5 popularity score

  Expected Outcome: Low *Phishing Risk*.

- **Test Case 2: Suspicious Website**

  This case simulates a borderline suspicious website with mixed indicators:

  - *Ratio of Null Hyperlinks*: 40%
  - *Number of External CSS*: 3 files
  - *Domain Registration Length*: 4 years
  - *Web Traffic*: 65-th percentile
  - *Page Rank*: 4.8 popularity score

  Expected Outcome: Medium *Phishing Risk*.

- **Test Case 3: Phishing Website**

  This case reflects a highly suspicious phishing website with characteristics strongly indicative of phishing activity:

  - *Ratio of Null Hyperlinks*: 85%
  - *Number of External CSS*: 1 file
  - *Domain Registration Length*: 1 year
  - *Web Traffic*: 18-th percentile
  - *Page Rank*: 8.2 popularity score

  Expected Outcome: High *Phishing Risk*.

- **Test Case 4: Ambiguous Website**

  This case represents an ambiguous website with conflicting indicators that activate multiple rules:

  - *Ratio of Null Hyperlinks*: 50%
  - *Number of External CSS*: 2 files
  - *Domain Registration Length*: 6 years
  - *Web Traffic*: 80-th percentile
  - *Page Rank*: 2.1 popularity score

  Expected Outcome: Medium *Phishing Risk*.

5.2 Results and Discussion

The following results demonstrate the system's behavior for each of the test cases. Rules activated during the inference process and the corresponding *Phishing Risk* outputs are presented for each case.

- **Test Case 1: Legitimate Website**
  - *Rules activated:* 28, 29, 30
  - *Phishing risk:* 11.0 (Safe)

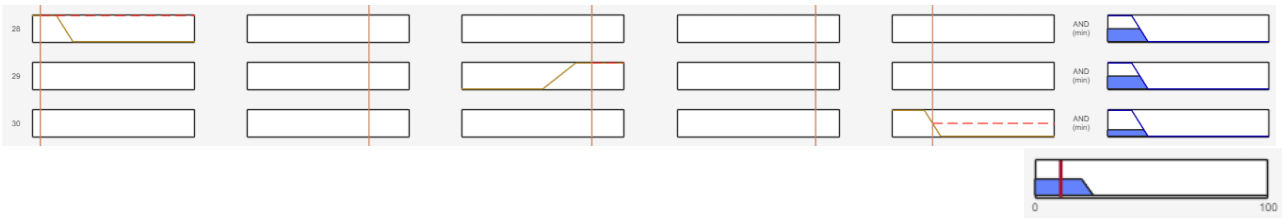


Figure 7: Rules activation and output for Test Case 1.

For this case, the system activated rules 28, 29, and 30, which primarily correspond to input variables indicating a low risk of phishing. The resulting *Phishing Risk* value is **11.0**, categorizing the website as *Safe*. The surface plot confirms the expected behavior, as all inputs align with low-risk labels.

- **Test Case 2: Suspicious Website**
  - *Rules activated:* 1, 14, 22
  - *Phishing risk:* 55.3 (Strongly suspicious)

Figure 8: Rules activation and output for Test Case 2.

Here, rules 1, 14, and 22 were activated, reflecting medium input values for various variables. The output *Phishing Risk* is **55.3**, categorizing the website as *Strongly Suspicious*. This result is consistent with the rule base, where medium risk variables combine to produce a moderately high output.

- **Test Case 3: Phishing Website**
  - *Rules activated:* 1, 10
  - *Phishing risk:* 74.0 (0.1 Strongly suspicious, 0.9 Phishing)

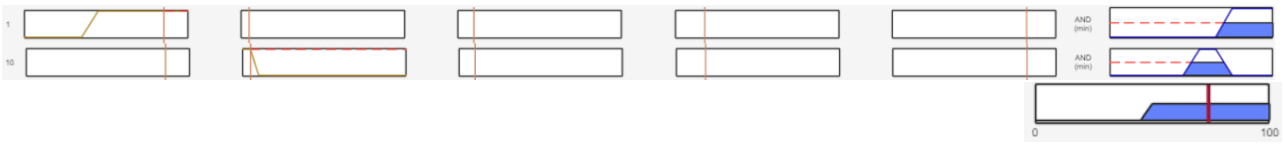


Figure 9: Rules activation and output for Test Case 3.

In this case, rules 1 and 10 were triggered, primarily due to high-risk indicators across all input variables. The system computed a *Phishing Risk* of **74.0**, mostly categorizing the website as *Phishing*. The activation of high-risk membership functions and rules aligns with expectations, as the inputs strongly suggest phishing activity.

- **Test Case 4: Ambiguous Website**

- *Rules activated:* 1, 29, 30
- *Phishing risk:* 55.7 (Strongly suspicious)

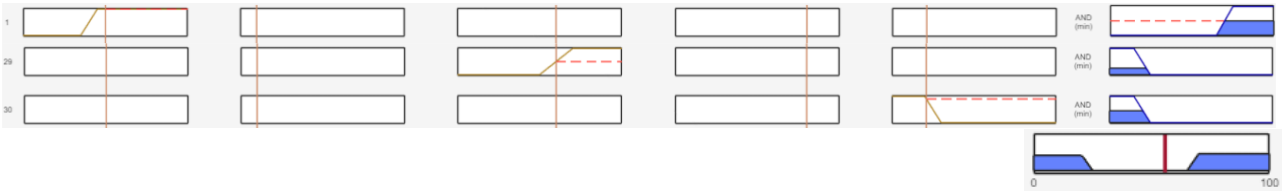


Figure 10: Rules activation and output for Test Case 4.

For this ambiguous case, rules 1, 29, and 30 were activated. The inputs include both high and low-risk indicators, leading to a *Phishing Risk* of **55.7**. The output categorizes the website as *Strongly Suspicious*. This result highlights the system's ability to handle conflicting inputs, as multiple rules contribute to the final output.

6 Task 5: Design of an Enhanced Fuzzy Expert System

To enhance the phishing detection capabilities, we propose a hierarchical fuzzy expert system that processes website features through multiple specialized analysis layers. Figure 11 illustrates this enhanced architecture:

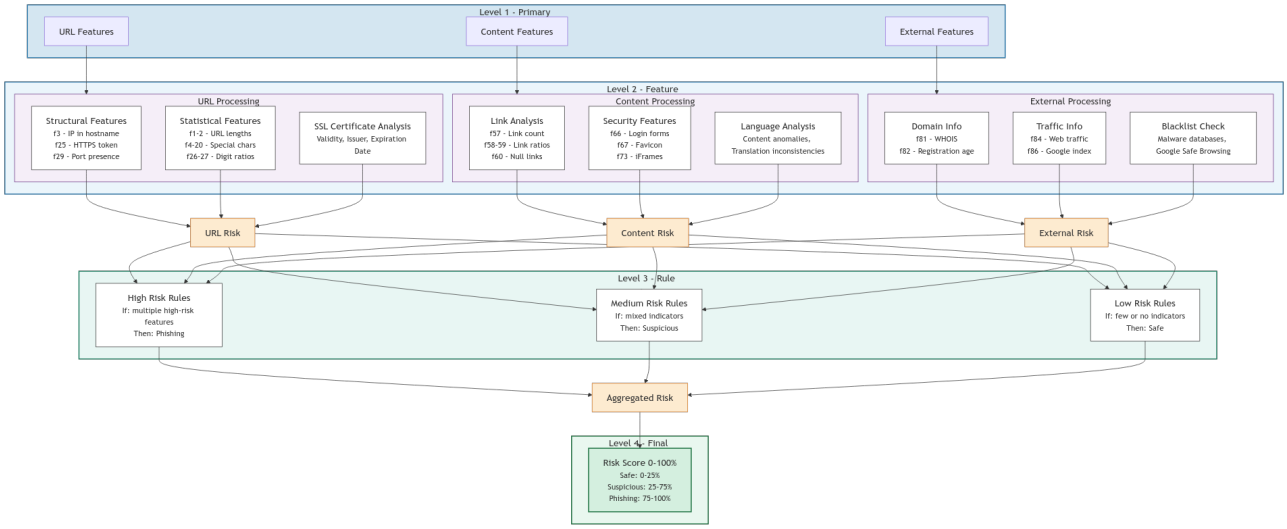


Figure 11: Enhanced Hierarchical Fuzzy Expert System Architecture

For a more detailed view of the system architecture, the reader may refer to the rotated version of the diagram on the **last page of this report** (Figure 12).

The proposed architecture follows a four-level hierarchical structure, systematically refining the analysis of website characteristics to improve phishing detection. This design ensures that information flows progressively, with each level contributing to a deeper and more nuanced understanding of potential threats.

At the **first level**, the system processes raw inputs by focusing on three complementary perspectives: *URL analysis*, *content evaluation*, and *external verification*. Each channel addresses distinct dimensions of website legitimacy, capturing structural details, embedded resources, and external reputation metrics. By combining these perspectives, the system lays a comprehensive foundation for detecting phishing indicators.

In the **second level**, feature analysis becomes more granular through dedicated processing modules. The *URL component* examines structural properties, such as IP usage, port presence, and SSL certificate validity, alongside statistical patterns like character distributions and word anomalies. *Content analysis* dives into hyperlink behavior, security indicators (e.g., login forms and iFrames), and language consistency to identify suspicious textual anomalies. Finally, *external verification* evaluates domain registration history, web traffic data, and blacklist status using trusted sources. This modular breakdown enhances feature extraction while enabling specialized analysis that goes beyond traditional single-layer systems.

The **third level** introduces a stratified rule evaluation mechanism, which assigns risk profiles to detected patterns. High-risk rules flag explicit phishing indicators, while medium-risk and low-risk rules identify subtle or ambiguous feature combinations. This layered approach not only improves detection granularity but also reduces false positives by carefully balancing strict rules with adaptive thresholds.

At the **final level**, the system integrates the outputs of prior layers into a unified risk assessment. By aggregating risk scores and analyzing interactions across the channels, the system delivers both a *numerical risk score* and a *categorical classification* (e.g., Safe, Suspicious, Phishing). This dual-output format ensures flexibility for integration into various security workflows.

The hierarchical design represents a significant improvement over our previous flat architectures. By progressively refining the analysis, the system achieves greater accuracy and robustness without introducing excessive complexity. Intermediate aggregation nodes streamline the rule evaluation process, while the clear separation of concerns enhances maintainability. In addition, the modular structure accommodates the integration of emerging detection techniques, such as SSL certificate analysis, advanced language processing, or blacklist verification, ensuring adaptability to evolving phishing tactics.

7 Conclusion

In this project, a fuzzy expert system was developed to classify websites based on their risk of phishing using a combination of linguistic variables, rule-based reasoning, and fuzzy logic techniques. The system successfully integrates structural, content-based and external features to provide a nuanced evaluation of the legitimacy of the website.

The proposed architecture incorporates a four-level hierarchical design that progressively refines the feature analysis and risk assessment. Using specialized processing at each level, the system overcomes the limitations of flat, single-layered approaches, enabling improved detection accuracy and flexibility. The rule base, constructed with 35 conjunctive rules, effectively captures combinations of characteristics to differentiate between safe, weakly suspicious, strongly suspicious, and phishing websites.

Testing across diverse scenarios demonstrated the system's robustness in handling conflicting inputs and ambiguous cases. The system outputs both a numerical risk score and a categorical classification, ensuring flexibility for integration into various security applications. The use of the MATLAB Fuzzy Logic Toolkit further validated the system's behavior through visualized rule activations and 3D surface plots.

The enhanced design presented in Task 5 highlights the potential for future improvements by incorporating additional features such as SSL certificate analysis, language consistency, and blacklist verification. This modular and adaptive design ensures the system's ability to evolve alongside emerging phishing tactics.

In general, the fuzzy expert system provides a solid foundation for phishing detection, combining interpretability, accuracy, and extensibility. Future work could explore hybrid approaches with machine learning techniques and real-time feature extraction to further enhance the system's performance in dynamic environments.

References

[1] A. Hannousse and S. Yahiouche, "Towards benchmark datasets for machine learning based website phishing detection: An experimental study," *arXiv [cs.CR]*, 2020.

[2] R. Zieni, L. Massari, and M. C. Calzarossa, "Phishing or not phishing? a survey on the detection of phishing websites," *IEEE Access*, vol. 11, pp. 18 499–18 519, 2023. DOI: 10.1109/ACCESS.2023.3247135.

[3] Z. Dou, I. Khalil, A. Khreishah, A. Al-Fuqaha, and M. Guizani, "Systematization of knowledge (sok): A systematic review of software-based web phishing detection," *IEEE Communications Surveys & Tutorials*, vol. 19, no. 4, pp. 2797–2819, 2017. DOI: 10.1109/COMST.2017.2752087.

[4] X. Roche, *Httrack website copier*, https://www.httrack.com/, Accessed: 2024-12-01.

[5] Imthaghost, *Goclone*, https://github.com/imthaghost/goclone, Accessed: 2023-12-01.

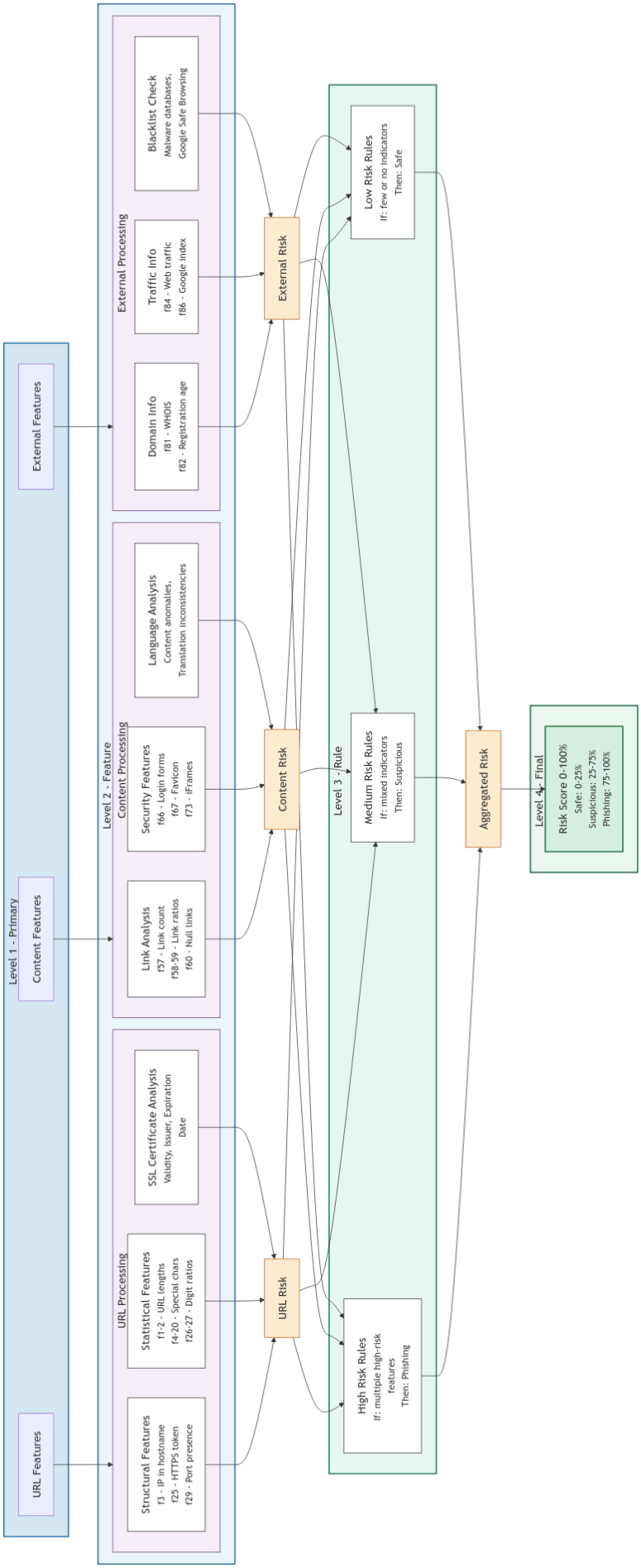


Figure 12: Rotated Enhanced Hierarchical Fuzzy Expert System Architecture for Better View