applied sciences

MDPI

Article

A Mamdani Type Fuzzy Inference System to Calculate Employee Susceptibility to Phishing Attacks

Yahya Lambat [1], Nick Ayres [1], Leandros Maglaras [1,*] and Mohamed Amine Ferrag [2]

[1] School of Computer Science and Informatics, De Montfort University, Leicester LE1 9BH, UK; P17218692@my365.dmu.ac.uk (Y.L.); nick.ayres@dmu.ac.uk (N.A.)
[2] Department of Computer Science, Guelma University, Guelma 24000, Algeria; ferrag.mohamedamine@univ-guelma.dz
* Correspondence: leandros.maglaras@dmu.ac.uk

Abstract: It is a well known fact that the weakest link in a cyber secure system is the people who configure, manage or use it. Security breaches are persistently being attributed to human error. Social engineered based attacks are becoming more sophisticated to such an extent where they are becoming increasingly more difficult to detect. Companies implement strong security policies as well as provide specific training for employees to minimise phishing attacks, however these practices rely on the individual adhering to them. This paper explores fuzzy logic and in particular a Mamdani type fuzzy inference system to determine an employees susceptibility to phishing attacks. To negate and identify the susceptibility levels of employees to social engineering attacks a Fuzzy Inference System FIS was created through the use of fuzzy logic. The utilisation of fuzzy logic is a novel way in determining susceptibility due to its ability to resemble human reasoning in order to solve complex inputs, or its Interpretability and simplicity to be able to compute with words. This proposed fuzzy inference system is based on a number of criteria which focuses on attributes relating to the individual employee as well as a companies practices and procedures and through this an extensive rule base was designed. The proposed scoring mechanism is a first attempt towards a holistic solution. To accurately predict an employees susceptibility to phishing attacks will in any future system require a more robust and relatable set of human characteristics in relation to the employee and the employer.

Keywords: fuzzy logic; FIS; mamdani; social engineering; rule set

Citation: Lambat Y.; Ayres, N.; Maglaras, L.; Ferrag, M.A. A Mamdani Type Fuzzy Inference System to Calculate Employee Susceptibility to Phishing Attacks. *Appl. Sci.* **2021**, *11*, 9083. https://doi.org/10.3390/app11199083

Academic Editor: Luis Javier Garcia Villalba

Received: 1 September 2021
Accepted: 28 September 2021
Published: 29 September 2021

1. Introduction

It is well established that humans are the weakest link in any secure system and where there is human/computer interaction it can undoubtedly lead to security breaches [1–3]. There are several factors that contribute to weak cyber security awareness in employees [4,5]. Businesses can find themselves vulnerable to attacks which result in security or system breaches, this often leads to the release of personal and sensitive data that ultimately results in damage to reputation and reduction in consumer confidence as well as monetary fines [6,7]. Moreover, during the COVID-19 crisis, employees are encouraged to increase teleworking while at the same time most products and services have become available over the web. The COVID-19 epidemic has resulted in a significant increase of fraudulent mails. Phishing emails, social engineering attacks, malware, ransomware, spyware, are only a few examples of the cyber crime incidents reported during the COVID-19 crisis period. Phishing mails take advantage of misinformation, isolation, and a lack of awareness to make people a vulnerable target for those types of attacks and persuade victims to hand over money, personal information, and even authentication credentials.

Employers mitigate potential cyber risks through various system protections, for example, automated software that detects potential phishing attacks and automatically resolves issues. In terms of employees, i.e., the human element; they are put through training courses which can be ineffective and "boring" [8]. Training courses are often

several hours in duration, and sometimes lack user involvement which subsequently results in a high probability of the attendees forgetting the training content within a short period of time. The proposed Fuzzy Inference System (FIS) constructed aims to solve the problem of employees who possess weak phishing awareness qualities by identifying their level of susceptibility to phishing attacks and thus more tailored and specific training can be given based on the susceptibility of the individual.

This paper investigates how fuzzy logic can be used to identify employees who are susceptible to phishing type attacks to varying degrees. This research demonstrates how a Mamdani type FIS can combine various human states with numerical variables, such as 'procrastination' by assigning a gradient value to produce a crisp output which is assessed by the most logical action. To test the proposed system a set of broad generic inputs were required. The inputs selected relate to the individual as well as the company the individual is employed. The proposed Fuzzy system, as initially designed, considered some basic human characteristics, security levels as well as company policies. A scoring system with various weightings for each input was defined for each input and an extensive rule-base was designed which was applied to the fuzzified inputs. Depending on the range of each input variable the FIS calculated the degree to which the employee is susceptible to phishing type attacks. To achieve this the proposed Fuzzy system designed considered different variables which included employees career experience, personal attributes and through this the FIS calculates the range of how much an employee is susceptible to phishing type attacks. The contributions and novelty of the article are:

- Presents the design and implementation of an effective Mamdani FIS
- Proposes the use of fuzzy logic to calculate the susceptibility of phishing attacks in employees
- Sets out the basic foundations of future advancement for a system that could mimic employee attributes

The rest of this research is organized as follows: In Section 2, we discuss related work and in Section 3 we give a system overview. Section 4 presents a Generic Fuzzy Inference System. In Section 5, we present The proposed Fuzzy System. In Section 6, we detail the individual components of the FIS while in Section 7 we evaluate its efficiency. Finally, Section 8 includes the conclusions that we draw from this research and Section 9 further work that we plan to conduct.

2. Related Work

In a recent work [9], the authors tried to categorize the existing phishing detection mechanisms into three groups; Content-Based Approaches, Heuristic-Based Approaches and Fuzzy Rule-Based Approaches. In the first category lie all methods that exploit visual similarities between a legitimate and a fake web site in order to detect the malicious ones. The second category of detection mechanisms are using several features that are extracted by phishing web pages in order to train models that can detect malicious pages. Finally in the third category fuzzy logic is applied in order to create rules and membership function of fuzzy sets for detecting phishing web sites.

Conducting a phishing simulation in health care institutions the authors in [10] showed that, during this simulation, a quite large number of employees clicked on phishing email with the click rates ranging from 13% to 49%. This research proved that for US health care organizations phishing emails indicate one of the major cybersecurity risks. One of the mitigation techniques that are proposed by the authors is to raise awareness about this risk by running simulated phishing campaigns or other similar initiatives. The authors in [11] conducted a real phishing study that involved 191 employees. One interesting finding was that when there was a sense of urgency in a phishing mail users tended to be more vulnerable. The authors also concluded that targeted training could could reduce employees susceptibility to phishing.

In [12], a theoretical model of factors that influence users on clicking phishing e-mails was developed. The model helped the researchers reveal the positive effects of

habit and proactive measures while on the same time showcased that only procedural https://www.overleaf.com/project/61407bb8ef6f86cc9a3484ef countermeasures were not affecting the users' behavior.

In an interesting work [13] that focused on phishing mails, it was proved that by clearly defining the main differences between legitimate and phishing emails, the risk of a phishing attacks reduces. By applying fuzzy logic the authors developed a novel AI detection system for classifying legitimate against malicious emails. Following a similar approach the authors in [14] developed a phishing detection mechanism using fuzzy logic.

This proposed method builds upon the previous mentioned work. This research investigates a framework where specific rules have been designed that include a number of aspects relating to particular employee characteristics, security policies and phishing sophistication levels. To fully understand an individuals susceptibility to a potential phishing attack, employee characteristics include individual experience and security level as well as personal behaviours: job satisfaction and procrastination. These employee traits are fed into a Mamdani FIS that outputs a overall phishing susceptibility rating for that particular employee.

3. System Overview

The motivation for fuzzy is to enable a computational paradigm that is based on how humans think, or as described by [15] "more often than not, the classes of objects encountered in the real physical world do not have precisely defined criteria of membership". This equates to the notion that an attribute, whatever that is, is inherently abstract. For example, tall, short, warm and cold are subjective terms, as one person's understanding of short can vastly differ to another person's understanding. This does not imply one is right nor the other is wrong, it is the fact that both persons understand each other when they communicate without using precise measurements, but their individual understandings are quite different. This state infers that observations have different meanings to different individuals. Fuzzy logic can be utilised to bridge the gap between the vagueness of difference of understanding.

Fuzzy logic in contrast to set theory can be used to model humanistic traits where the human brain reasons with uncertainty and vagueness. Computers are limited to evaluations of precise values and this is where fuzzy logic attempts to conjoin this uncertainty and vagueness. Fuzzy logic differs from classical logic in which statements are no longer absolutes such as yes/no, black/white, true/false, 1 or 0, but rather statements that find the degree of truth between the value of 0 and 1 (e.g., to what extent is a statement true).

For example, through fuzzy set theory Bob is 185 cm tall, is he average or is he tall? Traditional set theory shown in Figure 1 dictates that Bob is of average height however a fuzzy logic system would suggest that Bob is partly average and partly tall as seen in Figure 2 because he falls within the range of [0,1], Bob has a truth degree to average of 0.6 and a degree of 0.4 to tall. Thus, fuzzy logic aims to eliminate vagueness by assigning precise values to these graduations.

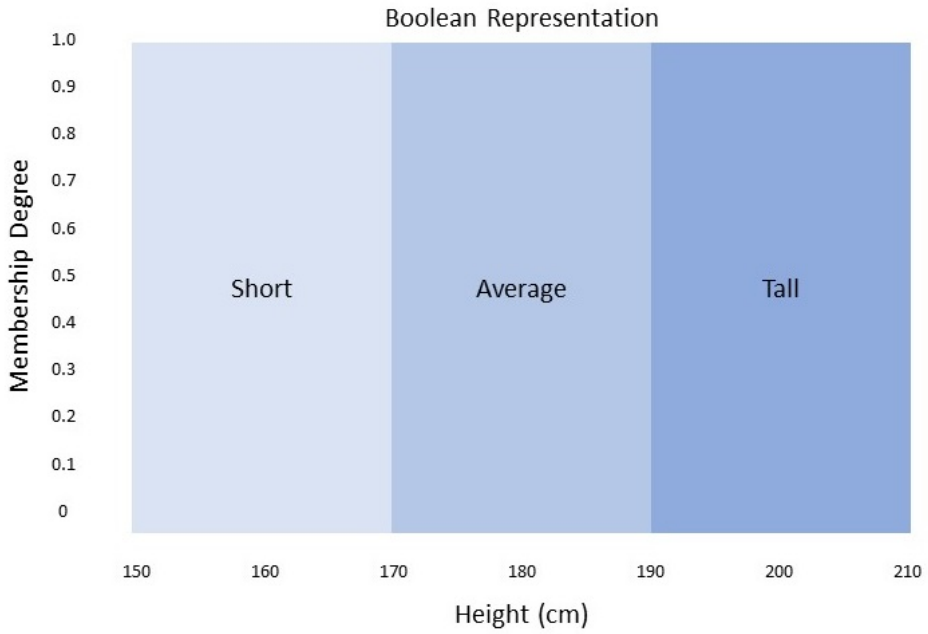

Figure 1. Boolean Representation.

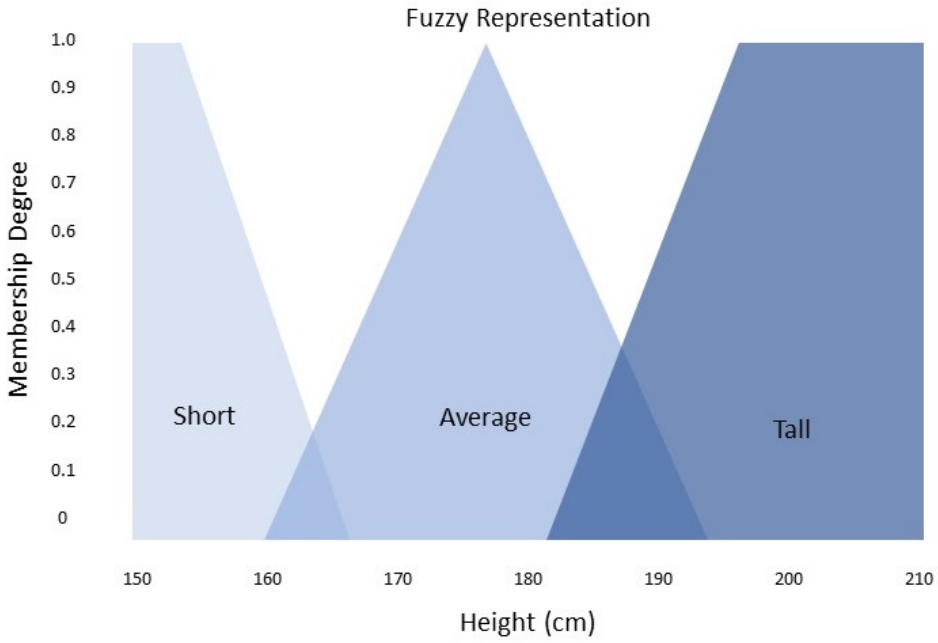

Figure 2. Fuzzy Representation.

4. Generic Fuzzy Inference System

To aid in the clarity and understanding of the Susceptibility Fuzzy Inference System, a generic Mamdani type FIS is described and detailed below. Each step of the inference process is outlined to enable a full comprehension of how a simple FIS functions. Figure 3 is an example of calculating the optimal climate rating and is used to portray the process of the fuzzy system. The system takes in, 2 input variables (Temperature, Humidity) and are fuzzified and processed by 3 IF-then rules, lastly the results are defuzzified into crisp values to represent the optimal climate rating.

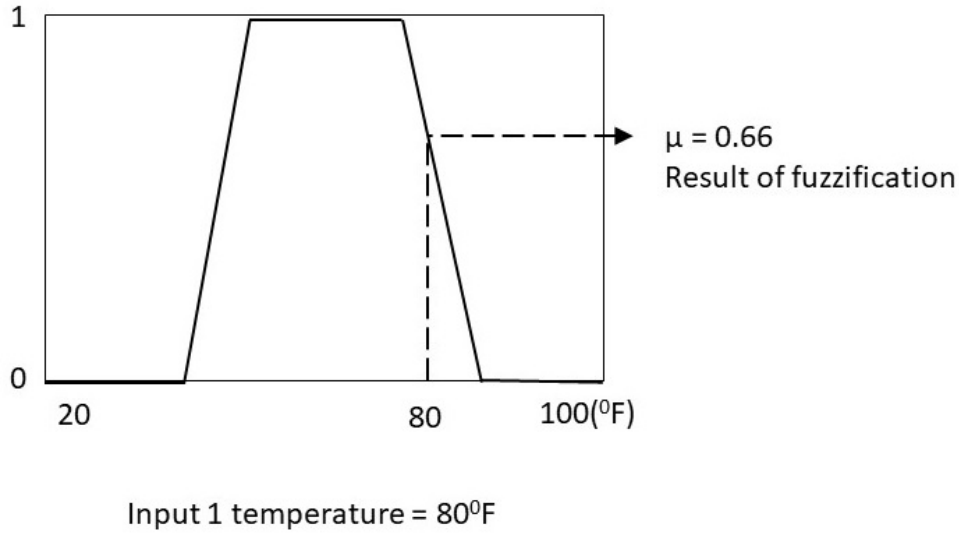

Figure 3. Fuzzification of Input variable (Temperature).

4.1. Fuzzification of Input Variables

The first step is to convert the crisp numerical values of the input variable "into the equivalent membership values of the appropriate fuzzy sets via membership functions" [16]. Following on from the example, the 3 If-then rules showcase 4 linguistic sets such as: temperature is (cold, warm) and humidity is (low, high). The 2 input variables (temp + humidity) must be fuzzified in accordance to their membership functions of the linguistic sets.

For instance, in Figure 3 a temperature of 80 degrees Fahrenheit is shown to have a degree of membership of 0.66μ to the linguistic set variable 'temperature is warm'.

4.2. Application of Rule List

The IF-then rules are defined and configured by the user to correspond with a logical fuzzy set to allow the system to calculate the degree of membership. The rules follow a forward chaining link wherein everything left of the THEN is named the antecedent and everything right of THEN is the consequent, indicative of forward chaining that uses existing facts and its rule base to derive new acts. Rules are comprised of subjective understanding and are defined to relay a logical statement, for example, the simple rule list below contains 3 If- then rules:

- Rule 1: 'If temperature is warm AND humidity is high THEN climate is ok'
- Rule 2: If temperature is warm OR Humidity is low THEN climate is good'
- Rule 3: 'If temperature is not warm THEN climate is undesirable'

In the above rule list, rule 1 contains 2 fuzzy linguistic sets, 'temperature is warm' and 'humidity is high', the rule then follows on to produce a fuzzy output of 'temperature is ok'. The fuzzy operators commonly used are the AND operation + OR operation, this is to allow a logical configuration of rule (other popular operators include min + max but they are irrelevant to this implementation). Rule 1 showcases the AND operator implying when the temperature is warm AND when the humidity is high then the climate is OK. Both statements must be true in order for the consequent to be true. Rule 2 shows the OR operator suggesting when the temp is warm or the humidity is low then climate is good, the OR operator allows more flexibility in the rule as it does not need a positive result (i.e., both antecedents do not need to be true) for the first set rather it allows both sets to not be dependant upon each other, as opposed to rule 1 wherein the climate is OK only if the temp is warm AND humidity is high. Rule 3 portrays a statement in which only 1 set is being used and still allowing a result.

4.3. Aggregation of Rule Outputs

As all decisions of the FIS are established from the rule-list, the rule outputs have to be amalgamated in which the aggregation process allows each fuzzy set to be combined into a single output value which is the representation of all inputs.

4.4. Defuzzification

The input for the last process of the FIS is defuzzification; this takes the aggregate fuzzy sets to produce an output of a single number (value). Defuzzification is the opposite process of fuzzification, as it takes out the precise quantity of the range of the fuzzy set to the output variable. There are different defuzzification methods however the centroid method (centre of area) is the most appropriate and common as it returns the centre of the area under the aggregate fuzzy set.

5. The Proposed Fuzzy System

The system (see Figure 4) designed for a susceptibility application uses a "Mamdani fuzzy inference" [Mamdani, 1975] aspect which enables six inputs to be taken into the system through the process of fuzzification and produce a single output. A Mamdani Fuzzy Inference method was chosen as they are easier to manipulate, have understandable rule bases and most importantly are appropriate to work with human input, as it takes into consideration the state which cannot be precisely defined by computing standards.

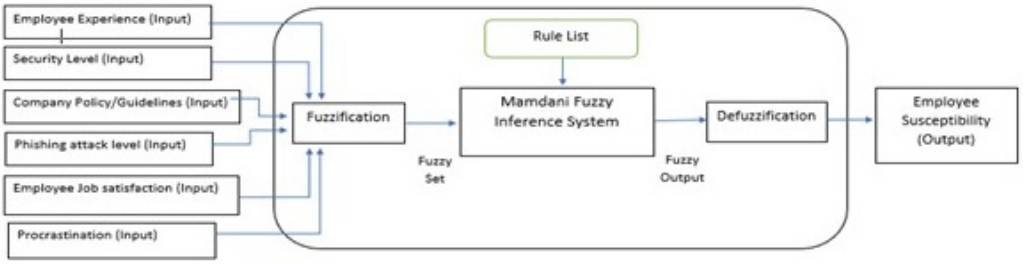

Figure 4. Susceptibility system design.

5.1. Stage 1: Fuzzification of Input Variables

In the beginning stages input variables are inserted into the system to be fuzzified, this is the process of the system to determine to which membership degree they (inputs) belong to each fuzzy set through membership functions. For example, Figure 5 represents the input 'employee experience' with the defined values between '0–100%'. These are also known as crisp numerical inputs. During the process of fuzzification the crisp input is fuzzified by assigning a corresponding linguistic variable via is membership degree. The fuzzy sets are then defined by the members it holds.

Membership degree	Linguistic Variable
00 - 30%	Non-Technical
20 - 70%	Technical
55 - 100%	Cyber Security Analyst

Figure 5. Input 1 Memberships.

The example of input 1 in Figure 4 'Employee Experience' details the membership degree and linguistic variable assigned to each value through the process of fuzzification.

5.2. Stage 2: Inference System (Rule List/Base)

The function of the inference system is to apply predefined rules to the fuzzy input to allow the configuration of the generation of a fuzzy output. The rule list is used to evaluate the linguistic values (e.g., Non-technical) and assign them to the corresponding fuzzy set.

Mamdani designed systems allow the rule list to be configured by the user to be catered to their logical needs, 'IF Then' statements are conditional as Figure 6 details the rule that states that 'IF Employee experience is….' alternatively, this equates to 'If Employee experience = 0–30 and security level = 0–10 and phishing attack level is 60–100 then employee susceptibility' = 70–100, this is also represented as:

IF (X is x1) AND (Y is y2) AND (Z is z3) THEN (W is w2)

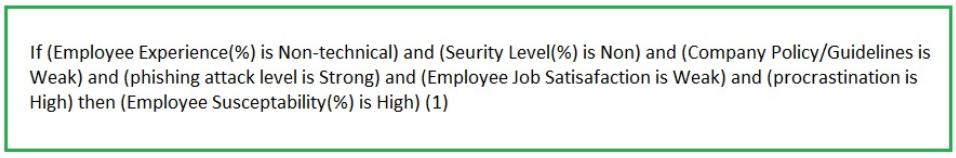

Figure 6. Example Rule.

Once the rule list is created it contains logical rules catered to the system. The rule list is then applied to every variation of the fuzzified inputs and assigned the correct linguist terms. Logic is used to describe the reality of the rule, for example, if an employee is a security analyst then it would be logical for them to have lower susceptibility, whereas an illogical statement would be if the phishing attack level is extremely high and company guidelines is weak then employee susceptibility is low.

5.3. Stage 3: Defuzzification

The process of defuzzification is the opposite of the beginning process of fuzzification, by this stage the input will have had its membership degree calculated and be in the form of a fully linguistic statement such as Figure 6 illustrates, in order to get a crisp numerical value as an output it must be reverted back through defuzzification. The Mamdani type FIS allows for alternative methods to find the defuzzification value.

Below are the described defuzzification methods what can be defined by the user to gain the desired output.

- Centroid—The centroid method returns its value from the centre of area of the fuzzy set as shown in Figure 7. It uses the formula (1) below where $\mu(x_i)$ is the membership value for point x_i.

$$xCentroid = \frac{\sum_i \mu(x_i)x_i}{\sum_i \mu(x_i)x_i} \quad (1)$$

- Bisector—This defuzzification method finds the line that bisects the fuzzy set into 2 sub regions of the same equal area, it usually concurrent with the centroid line (Figures 8 and 9).
- Middle, Largest and smallest of maximum (MOM), (LOM), (SOM) These 3 defuzzification methods take the maximum value of the plateau, as shown in Figure 9 SOM takes the smallest value, MOM the middle and lastly LOM takes the largest value.
- Membership Functions (MF) Membership functions are viewed as the graphical representation of a fuzzy set, it is a curve that defines how each plot in the input space is mapped to the degree of membership. The software used MATLAB creates visual representations for the MFs that are used for analysis. MFs for fuzzy sets are defined as μA:X → [0,1], wherein each element of X is shown to be a value between 0 and 1, this value is referred to the "membership value or the value with a degree of membership" [Emaths,2020] as it signifies the grade or degree of membership of the element to the fuzzy set.

The most common forms of membership functions are triangular, trapezoidal, Gaussian. In the susceptibility FIS, only the Trapezoidal MF and the Triangular MF were used as this was to allow for an efficient use of the systems output. The triangular membership function takes on the representation of a triangle as observed in Figure 10 whereby the 'trimf' function returns the fuzzy membership values using the following function (2):

$$f(x;a,b,c) = max\left(min\left(\frac{x-a}{b-a}, \frac{c-x}{c-b}\right), 0\right) \quad (2)$$

This is also defined by a lower limit (a), upper limit (b) and a value (m) wherein a $\leq$ m $\leq$ b as portrayed in Figure 11.

The trapezoidal MF uses the 'trapmf' function to compute wherein the output takes on a trapezoidal shape, Figure 12, and returns the fuzzy membership values using function (3):

$$f(x;a,b,c,d) = max\left(min\left(\frac{x-a}{b-a}, 1, \frac{d-x}{d-c}\right), 0\right) \quad (3)$$

Figure 13 illustrates the plots in which the lower limit (a) upper limit (d) lower support limit (b) and upper support limit (c) wherein a $\leq$ b $\leq$ c $\leq$ d.

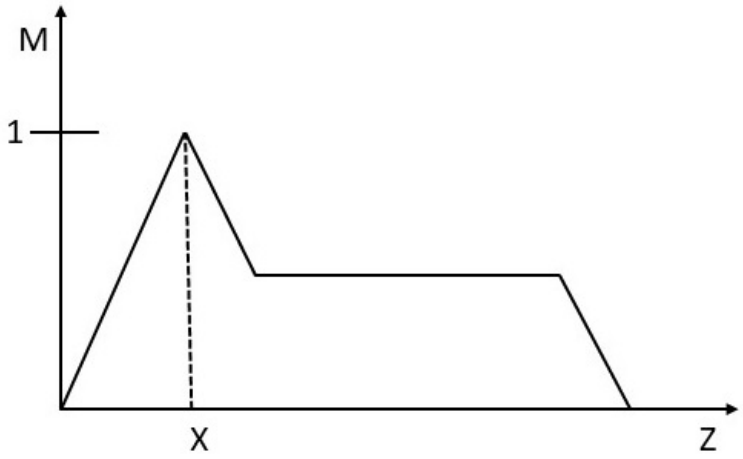

Figure 7. Centroid Visualisation.

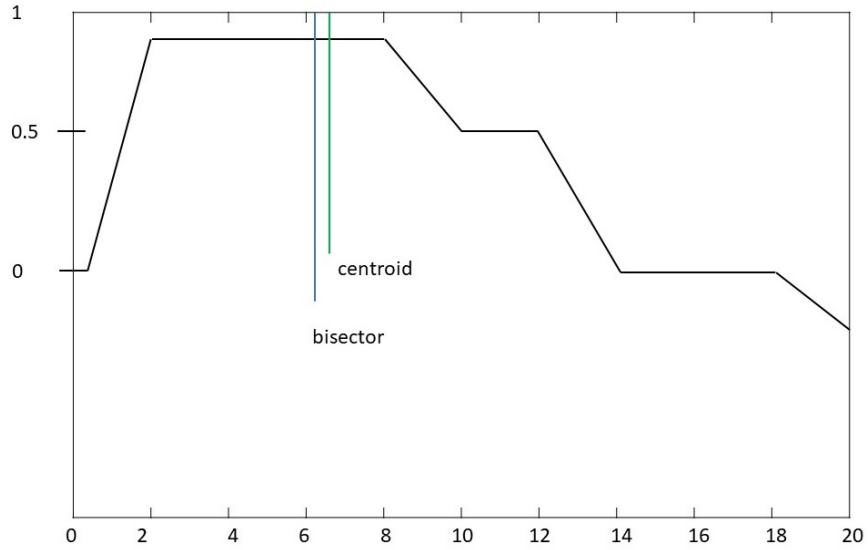

Figure 8. Visualisation of Bisector Defuzzification.

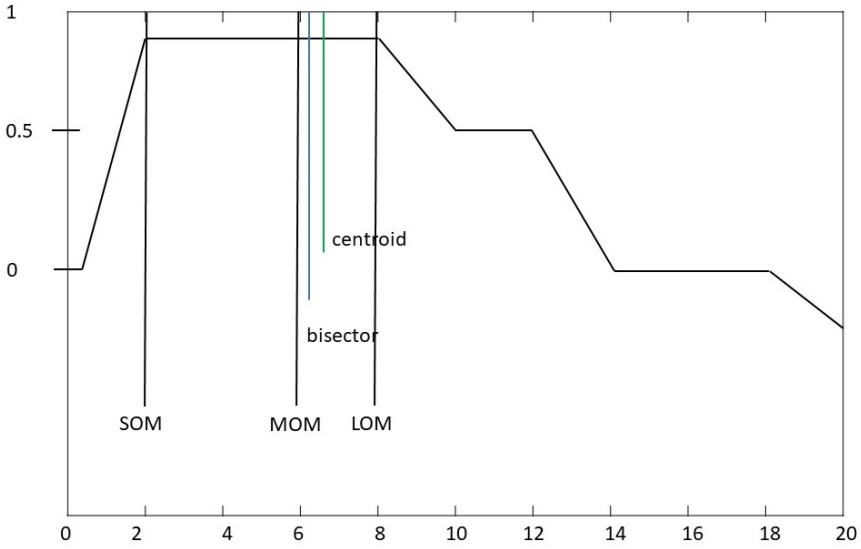

Figure 9. Visualisation of SOM, MOM and LOM.

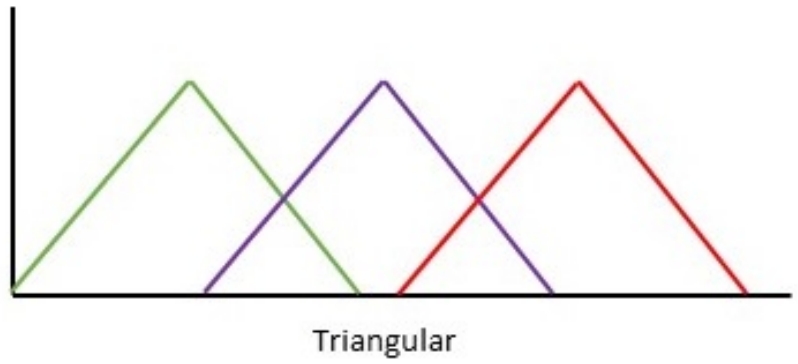

Figure 10. Tri MF Representation.

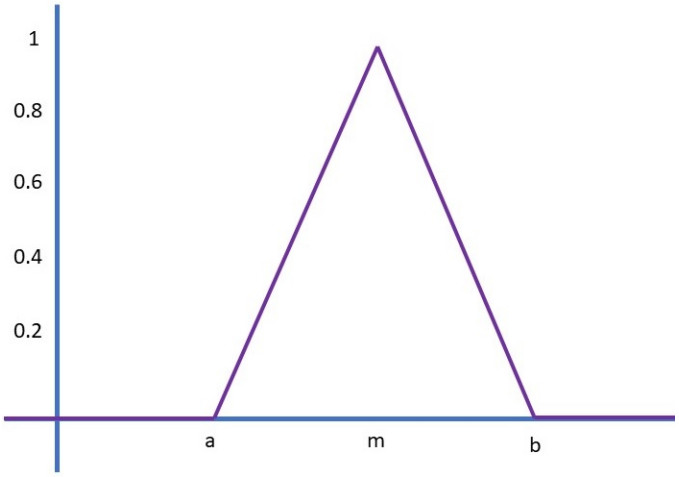

Figure 11. Tri MF Plot Points.

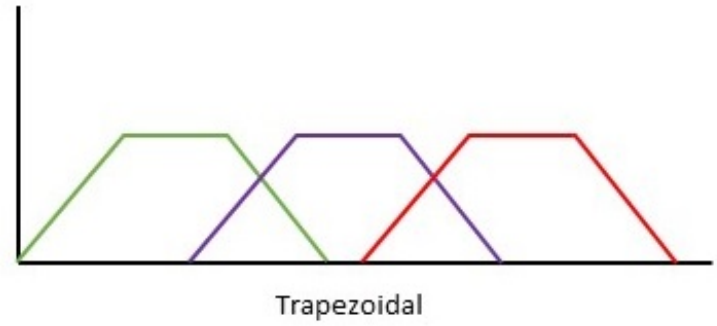

Figure 12. Visual Representation of Trapmf.

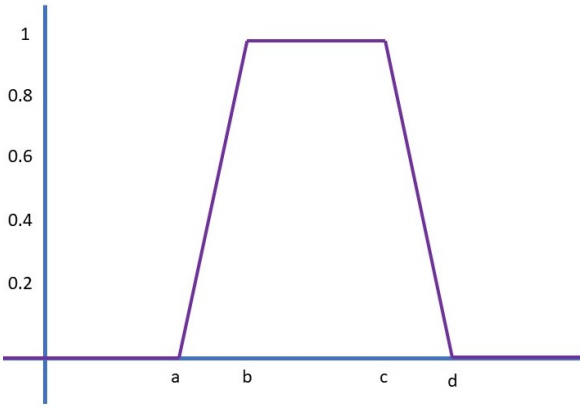

Figure 13. TrapMF Plot Points.

6. Variables and Fuzzy Set Justification

As shown in Figure 4 the system is relatively simple, however, the overall objective of the system and what it aims to achieve is quite complex. The system considers employee skill levels combined with human attitudes (e.g., procrastination and job satisfaction) to calculate how susceptible an employee is to phishing attacks. We believe that job satisfaction and procrastination can be causes of a security breach by an unintentional insider threat. To fully understand an employee's level of susceptibility to phishing attacks we wanted to include some basic personal character traits as well as how a particular individual performed within their job role. Job satisfaction and procrastination were

selected on the grounds that if an employee is bored or unhappy within their job role or function this could have a detrimental effect on their performance where corners may be cut or work left unfinished. A simple rule for this would be when a non-technical person is targeted by a high level phishing attack and has a high procrastination rating then the employee is considered as extremely susceptible.

With regard to employee experience we have found in our own practice as well as within academic work that individuals with "a little bit of knowledge" can often be the cause of a security breach. In taking this system forward both employee experience and background inputs will be fully reviewed and the inputs based on relevant research or subdivided into multiple inputs to fine tune these broad terms. Originally the system considered an additional two inputs; education level and phishing awareness. The idea behind education level was to take into consideration that a person with higher levels of education would be less susceptible to a phishing attack. For example, it would not be completely logical to have an employee with a HE degree in a non technical discipline being less susceptible then an employee with no HE degree but is technically trained. Alternatively, the phishing awareness variable initially used was a simple yes/no function and was removed due to its impact, for example, if an employee was aware of phishing attacks then this would imply that there was no requirement for the system to calculate their susceptibility and this had a negative impact on the entire system. The decision to remove these two variables was due to their ineffectiveness to the overall system outcome.

The following sections detail the individual components of the FIS with justification and reasoning behind why they were selected and configured for this particular research.

6.1. First Input: Employee Experience

The employee experience variable holds the greatest importance as it encompasses the career background of every employee. For the purpose of this research this input is broad and assumes that individuals with a medium to high level of technical experience will therefore be more inclined to better security practices. However, in reality this is often not the case. Risk compensation and risk homeostasis relate to increased levels of risky behaviour when increased safety measures are installed [17]. The idea was to put employees into as few categorisations as possible where employee experience could be categorised as non-technical and technical. The non-technical input allowed for a variety of non-technical employees such as those who did not study a science, technology, engineering, or mathematics (STEM) HE degree and did not have any technical skills regarding computer science or cyber security. Alternatively, the second set contained the employees who were STEM degree educated and do possess technical skills and were in a technical specific role. These two sets completely encompass all types of employees with no anomalous factors. However, to showcase the difference between the two, a third set was introduced as a controlled wildcard. A set named 'security analyst' was selected to allow for a vast difference between technical and non-technical employees.

The rules created for non-technical employees were principally not on a higher susceptibility rating then the technical employees, however, there were several included rules in the rule list where non-technical employees had outputted a low susceptibility rating, this was achieved through variable manipulation.

Employee Experience (%) [Range: 0–100]

- Non-Technical [0 0 25 30]
- Technical [20 30 60 70]
- Security Analyst [55 70 100 100]

The range (0–100%) was selected as an appropriate scale. It allowed a classification of the three inputs where the non-technical set was at the lowest end of the spectrum (0–30%) and security analyst was at the higher end (55–100%) this follows a logical order where employees included in the non-technical set were classed lower then the cyber security analyst. Evaluating the system for further changes in hindsight the system would have been more efficient if this variable was configured with the intent of the end result being

parallel with the ranges, meaning the most susceptible employees (non-technical) would be in the higher bracket (55–100%) and the security analyst would be in the lower bracket (0–30%). However, this would be a purely aesthetic modification as it is not the range that dictates the output rather the rule list (nonetheless the system would look congruent with the low to high susceptibility).

The range values of the non-technical set [0 0 25 30] begins with overlapping values and the security analyst set [50 70 100 100] ends similarly. This was to allow for rule strength to fire and hit the extreme left value for non-technical employees and for the rules to fire on the extreme right for the security analyst. This was modified from the original version of [0 5 15 25 30] for the non-technical set as the rules being fired were not hitting the extreme values due to the defined ranges. Repeating the first and last values, allowed for the membership function to begin and end at the extreme points. Comparing the difference of rules hitting the extreme values, shows that changing the range to a duplicate value allows for a more accurate response.

The input itself contributes a significant weighting to the system as 95% of all the rules configured contain the employee experience input. As it is indicative of and plays a big part in the output, the set shapes the susceptibility of an employee from the very early stages.

The initial design of input 1 considered additional inputs such as manager and expert. These were aimed to further expand the employee skill levels. For example, the assumption that a manager would have some prior knowledge of social engineering considering their role, additionally the same would apply to the expert. However, these sets were taken out system inefficiency. This was due to the variables leaving far too many circumstances open including questions arising to 'what if a manager has no technical experience whatsoever' and 'expert is too subjective', there could be SQL experts or marketing experts neither of which would further clarify the susceptibility of the employee, therefore they were removed to allow for a simpler input one. The Trapezoidal membership function was chosen for this variable due to the advantages of the polygonal MFs, these advantages included:

- A small scale of data was required to define the MF
- TrapMF permits an easy modification of the input values
- The TrapMF allowed for a definition of values in four membership ranges (e.g., 20 30 60 70)
- Output visualisation graphs generated are easy to read and interpret

6.2. Second Input: Security Level

An employee's security level was chosen as input two with the logical assumption that the higher security clearance an employee has, the more aware they will be of social engineering. Thus, following on from this the sets that was chosen was sectioned out into four memberships. This differs from the UK's official security vetting structure [18,19], however, the generalisation of the levels is portrayed ranging from an employee having a low clearance (Counter terrorism check) to high clearance (developed vetting). The range similar to input one takes the form of 0–100%, however the draft system initially had a rating of 1–10 but this was changed to match the other inputs. Similar to input one overlapping of values was included to maximise the rule firing efficiency.

Input 2: Security Level (%) [Range 0–100]

- Non [0 0 20]
- Low [15 25 35]
- Medium [26 50 80]
- High [70 100 100]

For input two the Triangular MF function was selected as it allowed a smooth transition and overlapping between the security clearance sets. The Gaussian MF was initially chosen but was discarded due to inefficient overlapping, this meant that the security levels are absolute and not interchangeable.

- Trimf was selected as it is defined by three values, e.g., Low 15 25 35
- The graphical output is also simple and easy to read
- The overlap of degree memberships is easy to compute
- TriMF functions have less complexity when splitting the values

6.3. Third Input: Company Policy

Input three is a generic input regarding the target company and their cyber security policies or ISO's. A company that has strong policies and safeguards in place to negate a cyber security issues they are less likely to succumb to an attack. This is indicative of a company that educates their employees though specific training and the assumption that they follow ISO guidelines. However, if a company has a weak security policy and ISO's then the employees are more likely to have a higher susceptibility rating. This input relates to an overall cyber security policy the company has in place or an aggregate of many cyber security related policies. In reality this input would be based on a specific cyber related activity, for example, accessing the Internet from a company computer or the use of company email.

Input 3: Company Policy (%) [Range: 0–100]

- Weak [0 0 40]
- Medium [25 50 75]
- Strong [60 100 100]

Input three system weighting is used in 90% of the rules. It dictates an employee's susceptibility due to the reasoning that if a company takes an active role in ensuring their workforce is cyber aware and can identify social engineering tactics then they have a workforce that is not so susceptible to phishing attacks.

6.4. Fourth Input: Phishing Attack Level

Input four was included as it details the potential level of attack an employee might be presented with. The set starts with a weak membership where the range begins at nine. This was selected due to the fact that even a weak phishing attack can still produce a certain level of damage if successful. This input has a medium weighting and was configured so the other inputs contribute more to the overall output. When a phishing attack is strong, in most cases the susceptibility result will always be high, this ultimately negates all the other variables and defeats the purpose of the system.

Input 4: Phishing attack level (%) [Range: 0–100]

- Weak [9 9 30]
- Medium [20 50 75]
- Strong [57 100 100]

The phishing attack level uses a Tri MF which allows for perfect overlapping of the phishing attack severity levels. This allows for a smooth transition between the sets as there is not an exact value if a phishing attack is medium or strong, that is why the medium range is 20% to 75% with the strong set starting at 57% to allow for a generous overlapping of memberships. An improvement for further research on this input is to subjugate it to extreme scrutiny wherein there will not be 3 simple categorisations but rather a deeper intricate classification of phishing attacks and their severity. This would include sets of Email phishing, spear phishing, whaling etc., along side their severity level (an additional input). This will not only create a more optimised system but increase its efficiency exponentially as it relates real phishing attacks rather than simplifying the attacks.

6.5. Fifth Input: Employee Job satisfaction

The initial aim of the FIS was to incorporate human personality traits into a technical system. The variable job satisfaction input was chosen to aid in the calculation of susceptibility. Job satisfaction was selected as a human trait variable that effects the output significantly [20].

Input 5: Employee Job satisfaction (%) [Range: 0–100]

- Weak [0 0 10 20]
- Medium [15 30 45 60]
- Strong [50 65 100 100]

Job satisfaction is directly linked to employees who lack motivation within their role, an unhappy employee is less likely to pay attention in training courses, progress their career development (i.e., learn new skills) or be a loyal employee [21]. Though the latter is an extreme suggesting sabotage of a kind, the system takes into account that if an employee has a low job satisfaction then they are more likely to fall prey to phishing attacks, this could mean by accident or intentionally.

6.6. Sixth Input: Procrastination

The second form of human trait included into the system is procrastination as it relates to two different types of employees, one productive and the other not. The assumption taken is that an unproductive employee may spend their time browsing the internet [22], watching online videos or doing performing actions other than their job thus they would be more prone to security breaches [23]. This is taking into account the visiting of non-work related websites which is indicative of visiting compromised websites and allowing potential phishing opportunities.

Input 6: Procrastination (%) [Range: 0–100]

- Weak [0 0 20 35]
- Medium [25 35 55 60]
- Strong [50 70 100 110]

The defuzzification method used for input six was the TrapMF, this was to allow for a wide range of values to be defined. Input six is more of an experimental variable and though does play a part within the system, it does not have a large weighting which relates to the output when compared with input one or input two. The third set for membership 'high' has an anomaly as the right most extreme value is 110 (right point of the plateau), this was done purposely as the rules firing were not hitting the extremes efficiently when the value was 100.

6.7. Output: Employee Susceptibility

All six inputs were configured through the specialised user defined rule base enabling the value of susceptibility as the principle output. The end result was developed to ensure the susceptibility value was suited to identify the employees who are weak or more susceptible to a phishing attack. The systems main objective was to determine how much an employee is susceptible and in doing so three members were chosen. For clarity a simple susceptibility range of low to high was chosen. This was done to allow to a more efficient understanding through 3 members rather then separating out the classifications even further which would become overly complex. Originally the research factored six member outputs. These outputs included extremely low, low, neutral, medium, high and extremely high. This original design was disregarded due to the overly complex nature of the classifications. It was found that this introduced more illogical irregularities in the process of explaining why such an employee is susceptible to such an extent.

Output 1: Employee Susceptibility (%) [Range: 0–100]

- Low [0 0 15 30]
- Medium [20 30 55 65]
- High [55 75 100 110]

The range follows the same precedent as all the inputs and outputs a value from 0 to 100%, however, the far right extreme value was set at 110 to ensure the system functioned correctly with no errors. The output is decided by the values of all six input variables, this is done by the FIS which calculates the output value. The defuzzification method chosen was the trapezoidal MF that allowed for a more efficient output value.

7. Testing, Evaluation and Results

The system incorporated 69 rules that were specifically designed to create the logic outcome. The rules that were optimised catered for all employees depending on the six input variables. The rule base followed a logical process such as:

Example Rule 1:

- **IF** Employee Experience is Non-technical **AND**
- Security Level is Non **AND**
- Company Policy is Medium **AND**
- Phishing Attack Level is Medium **AND**
- Employee Job Satisfaction is Medium **AND**
- Procrastination is High **THEN**
- Susceptibility is **High**

The above rule highlights an employee with a high susceptibility rating, the main contributing factor is that the employee is non-technical and the procrastination level is high.

The next rule indicates that with strong company guidelines and a technical employee who is productive at work results in a low susceptibility rating.

Example Rule 2:

- **IF** Employee Experience is Technical **AND**
- Security Level is Medium **AND**
- Company Policy is Strong **AND**
- Phishing Attack Level is Weak **AND**
- Employee Job Satisfaction is Medium **AND**
- Procrastination is Low **THEN**
- Susceptibility is **Low**

The next rule results in a low output, however, it only considers 4 variables, it showcases that when a technical employee with a medium security level clearance is faced with a weak attack the susceptibility is considerably low.

Example Rule 3:

- **IF** Employee Experience is Technical **AND**
- Security Level is Medium **AND**
- Company Policy is Strong **AND**
- Phishing Attack Level is Weak **THEN**
- Susceptibility is **Low**

The rule below considers the various inputs that all contribute to the susceptibility rating of the employee being weak. This rule is different to the others as it takes into consideration an employee who is a security analyst, however, due to the company having weak guidelines, phishing attacks rating string and the employee at best is below average (job satisfaction + procrastination) thus resulting in a high susceptibility rating.

Example Rule 4:

- **IF** Employee Experience is Security-Analyst **AND**
- Security Level is High **AND**
- Company Policy is Weak **AND**
- Phishing Attack Level is Strong **AND**
- Employee Satisfaction is Medium **AND**
- Procrastination is High **THEN**
- Susceptibility is **High**

The rule base is built on logic, for instance if a non-technical employee is faced with a strong phishing attack level and the company has weak security policies then the most logical and oblivious outcome would be that the employee is susceptible to phishing attacks. To mimic a workplace environment of having anomalous variables such as, a non-technical employee who is developing their skills by studying for a security based

certification or similar were ultimately left out of the FIS as the aim of this research was to build a simple system that showcased the susceptibility of the average employee. However, if this research is taken forward then a more sophisticated rule base should be configured whereby employees who fall into an anomalies category could then be accommodated.

The initial method taken to create a rule was the utilisation of a random number generator. Random numbers were used to generate more then 150 rules that attempted to hit every single variable in the defined sets. The idea was to produce the most efficient and logical output where, if every value is hit then the output has to be correct. In doing so the random number generator was used to allow for every possibly that the defined combinations could make. However, when examining the quality of the output, the susceptibility of rating for variables were highly illogical and inefficient. For example, a rule that stated a non-technical employee had a low susceptibility rating also contained the variables of a high phishing attack level and was an extremely unproductive worker. Indicative of this the need to change the rule base entirely and address the problem was required. If the rule base did not work then the entire system is unfit. The rules were altered to cater for the most logical outcome where only logical rules were included enabling a rule base that was efficient and well structured.

The entirety of the rule base contained the AND operator. It was found that when the OR operator was used to define rule input variables, there were a number of illogical errors as well as abnormal decision of events observed. For example, if the employee experience is 'security analyst OR phishing attack level is weak then employee susceptibility is High'. This result is deemed inaccurate, however, a very small number of OR rules were included to satisfy the outcome and not negate the earlier defined rules. This is further proven through Figure 14 where the fifth column (SOM) shows values that are of no use, e.g., a high susceptibility rating is only 60. The AND operator was shown to have the optimal result when configuring the rulebase (Centroid).

High Suseptibility	78.97674	80	100	60	54.41595
Low Susceptibility	13.03863	11	22	0	13.03863
Analyst High S	13.53571	12.5	25	0	11.53961
Analyst Low S	80.59916	83.5	100	67	82.28169
Weak employee	78.92807	80	100	60	54.41595
Strong Employee	14.10959	13.5	27	0	71.24356
Average Employee	50	50	50	50	12.25729
Loaths Job	81.103	85	100	70	54.41595
Max Susceptibility	82.28169	87.5	100	75	54.41595
Min Susceptibility	11.4058	7.5	15	0	54.41595
Defuzz Methods:	Centroid	MOM	LOM	SOM	Bisector

Figure 14. Defuzzification Output Values.

A series of tests were developed to push the bounds of the system to ensure full robustness and that it meets the requirements and expectations. A spreadsheet was created as a secondary input process. The spreadsheet acted as a alternative method of inputting values into the system without having to reconfigure the system. This external file contained values for each defined variable, this was linked to the system using C code. The spreadsheet allowed for multiple to test cases (employees) to be ran simultaneously.

The input values in the external spreadsheet was executed through the system evaluated through the predefined rules. This resulted in a value being written to the spreadsheet for analysis (Figure 15). The analysis took the form of cross checking all the defuzzification values returned and calculating the most accurate susceptibility rating. Each row of the sheet mimicked a employee test case that was fed into the system. For example, row 1 was predicted to be a high susceptible employee and the variables inputted catered for this. There are various defuzzification methods (Centroid, LOM, MOM, SOM), however, after evaluation of the system outputs, it was found that the MOM method produced the most effective output. Figure 15 shows the MOM values result in accurate and efficient

output values that are congruent with the system aims. The centroid method produces results closely matching the MOM method (Figure 14). The MOM and centroid values were analysed to have the correct and most logical results for the susceptibility rating. However, comparing the two methods, MOM was found to be a more suitable. The LOM, SOM and Bisector methods were unsuitable and were not utilised for the system where inaccurate values returned within the defuzzification method as illustrated in columns 4, 5 and 6 in Figure 14.

15	20	90	10	70	Hig Sus	78.97674	80	100	60	79
50	60	20	10	15	Low Sus	13.03863	11	22	0	13
80	80	10	10	15	SEC High	13.53571	12.5	25	0	13
50	60	60	10	90	SEC low	80.59916	83.5	100	67	81
5	90	25	5	95	weak employee	78.92807	80	100	60	79
25	50	25	80	10	Strong	14.10959	13.5	27	0	14
28	60	15	45	25	middle	50	50	50	50	50
0	100	15	5	100	Hates Job	81.103	85	100	70	81
0	0	100	0	100	Max	82.28169	87.5	100	75	83
100	100	9	100	0	Min	11.4058	7.5	15	0	11
Input 2	Input 3	Input 4	Input 5	Input 6		Centroid	MOM	LOM	SOM	Bisector

Figure 15. Defuzzification Values from an Excel File.

8. Conclusions

A key issue concerning a human workforce is their vulnerability to social engineering attacks. Millions of pounds are spent by companies securing their systems in an attempt to evade cyber related attacks, however this protection can often be easily undone by a single employee that is either unaware or has not been trained to identify social engineering tactics. To minimise the success of social engineered attacks a system similar to the one proposed here could be utilised to find and target individual employees that may be more susceptible or vulnerable to phishing attacks. Employee training is often generic, however the results of this type of system could be incorporated into the induction process of new employees to tailor specific training of individuals who have a higher susceptibility score. In summary, the FIS designed and implemented dictates an effective Mamdani controller that utilises fuzzy logic to calculate the susceptibility of phishing attacks in employees. The research conducted has generally focused on the application of using Fuzzy logic to identify and negate phishing attacks from a technical level and considers the human element which is the main factor for social engineering. Due to the advancement of technology the system successfully bridges the gap between technology and human state (traits) which was the aim of the susceptibility FIS. This research sets out the foundations of future advancement for a system that mimics employees attributes to gain maximum effectiveness in the targeting and the identification of those employees who are considered to have weak security skills and would therefore pose a risk to the company.

9. Further Work/Study

The system designed here is an experimental framework and proof of concept, subsequently there are several advancements that can be made to develop the system further. Firstly, to improve the accuracy of the system output requires the addition of more responsive variables. To accurately predict an employees susceptibility to phishing attacks will in any future system require a more robust and relatable set of human characteristics in relation to the employee and the employer. The field of Organisational Psychology has some interesting topics which can be applied to this FIS. Topics such as psychometric tests, motivation, occupational stress and organisational culture could be included. Other more personal topics could be incorporated, for example, employee satisfaction level, employee IQ and career motivation.

This proposed FIS considers multiple different systems. The primary system comprised of three subsystems, for example, system 1 takes only the career experience into account. System 2 relates to the the business/company (policies, size of workforce) and

the last involves the employees humanistic traits such as the ones described above. These three different systems produce their own crisp output values which are then combined and used to create the susceptibility rating. In doing so certain variables will require different weightings to the result of the output. Any future iteration of this FIS will require more finely tuned inputs that relate to the individual as well as the company. Finally, human thinking methods are not always predictable, considering this the configuration of rules and variables to cater to anomalous circumstances will undoubtedly increase system accuracy, efficiency and enable companies to identify the weak links within their workforce.

Author Contributions: Conceptualization, Y.L., N.A. and L.M.; Methodology, Y.L., N.A. and L.M.; Software, M.A.F., Y.L. and N.A.; Validation, N.A. and L.M.; formal analysis, M.A.F. and N.A.; investigation, Y.L., N.A. and L.M.; resources, M.A.F., Y.L. and N.A.; data curation, Y.L., N.A., L.M. and M.A.F.; writing—original draft preparation, Y.L., N.A. and M.A.F.; writing—review and editing, N.A. and L.M.; visualization, M.A.F., N.A. and L.M.; supervision, N.A. and L.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Conflicts of Interest: All authors declare no conflict of interest.

References

1. Bhusal, C.S. Systematic Review on Social Engineering: Hacking by Manipulating Humans. *J. Inf. Secur.* **2021**, *12*, 104–114.
2. Aldawood, H.; Skinner, G. Educating and Raising Awareness on Cyber Security Social Engineering: A Literature Review. In Proceedings of the IEEE International Conference on Teaching, Assessment, and Learning for Engineering (TALE), Wollongong, Australia, 4–7 December 2018; pp. 62–68
3. Frauenstein, E.D.; Solms, R.V. *Combating Phishing: A Holistic Human Approach*; Information Security for South Africa: Pretoria, South Africa, 2014; pp. 1–10.
4. Fielding, J. The people problem: How cyber security's weakest link can become a formidable asset. *Comput. Fraud Secur.* **2020**, *2020*, 6–9. [CrossRef]
5. Darwish, A.; Zarka, A.E.; Aloul, F. Towards understanding phishing victims' profile. In Proceedings of the International Conference on Computer Systems and Industrial Infomatics, Sharjah, United Arab Emirates, 18–20 December 2012; pp. 1–5.
6. Parthy, P.P.; Rajendran, G. Identification and prevention of social engineering attacks on an enterprise. In Proceedings of the International Carnahan Conference on Security Technology (ICCST), Chennai, India, 1–3 October 2019; pp. 1–5.
7. Knight, R.; Nurse, R.C.J. A framework for effective corporate communication after cyber security incidents. *Comput. Secur.* **2020**, *99*, 102036. [CrossRef]
8. Ferrara, J. Why Most Cyber Security Training Does Not Work. 2012. Available online: https://www.proofpoint.com/us/security-awareness/post/why-most-cyber-security-training-doesnt-work (accessed on 15 September 2021).
9. Zuraiq, A.A.; Alkasassbeh, M. Phishing detection approaches. In Proceedings of the 2019 2nd International Conference on new Trends in Computing Sciences (ICTCS), Amman, Jordan, 9–11 October 2019; pp. 1–6.
10. Gordon, W.J.; Wright, A.; Aiyagari, R.; Corbo, L.; Glynn, R.J.; Kadakia, J.; L.; man, A.B. Assessment of employee susceptibility to phishing attacks at US health care institutions. *JAMA Netw. Open* **2019**, *2*, e190393. [CrossRef] [PubMed]
11. De Bona, M.; Paci, F. A real world study on employees' susceptibility to phishing attacks. In Proceedings of the 15th International Conference on Availability, Reliability and Security, Virtual Event, Ireland, 25–28 August 2020; pp. 1–10.
12. Shahbaznezhad, H.; Kolini, F.; Rashidirad, M. Employees' Behavior in Phishing Attacks: What Individual, Organizational, and Technological Factors Matter? *J. Comput. Inf. Syst.* **2020**, 1–12. [CrossRef]
13. Salem, O.; Hossain, A.; Kamala, M.. Awareness program and ai based tool to reduce risk of phishing attacks. In Proceedings of the 2010 10th IEEE International Conference on Computer and Information Technology, Bradford, UK, 29 June–1 July 2010; pp. 1418–1423.
14. Shirsat, S.D. Demonstrating different phishing attacks using fuzzy logic. In Proceedings of the 2018 Second International Conference on Inventive Communication and Computational Technologies (ICICCT), REC kal, India, 20–21 April 2018; pp. 57–61.
15. Zadeh, L.A. Fuzzy sets. *Inf. Control* **1965**, *8*, 338–353. Available online: https://www.sciencedirect.com/science/article/pii/S0019995865902410X (accessed on 10 September 2021). [CrossRef]
16. Wang, C. *A Study of Membership Functions on Mamdani-Type Fuzzy Inference System for Industrial Decision-Making*; Lehigh University: Bethlehem, PA, USA, 2015.

17. Hedlund, J. Risky business: Safety regulations, risk compensation, and individual behavior. *Inj. Prev.* **2000**, *6*, 82–89. [CrossRef] [PubMed]
18. Scott, P.F. The contemporary security vetting landscape. *Intell. Natl. Secur.* **2020**, *35*, 54–71. [CrossRef]
19. GOV.UK. 2020. National Security Vetting: Clearance Levels. Available online: https://www.gov.uk/government/publications/united-kingdom-security-vetting-clearance-levels/national-security-vetting-clearance-levels (accessed on 28 June 2021).
20. Sainju, B.; Hartwell, C.; Edwards, J. Job satisfaction and employee turnover determinants in Fortune 50 companies: Insights from employee reviews from Indeed.com. *Decis. Support Syst.* **2021**, *148*, 113582. [CrossRef]
21. Roadbert, M. Why organisational readiness is vital in the fight against insider threats. *Netw. Secur.* **2020**, *8*, 7–9. [CrossRef]
22. Askew, K.L. *The Relationship Between Cyberloafing and Task Performance and an Examination of the Theory of Planned Behavior as a Model of Cyberloafing*; University of South Florida, Department of Psychology: Tampa, FL, USA, 2012.
23. Hadlinton, L.; Parsons, K. Can cyberloafing and Internet addiction affect organizational information security? *Cyberpsychol. Behav. Soc. Netw.* **2017**, *20*, 567–571. [CrossRef] [PubMed]