Hierarchical expert system for security evaluation and its implementation on an Android smartphone

Ishika Prasad, Tejas Raval, Lipisha Chaudhary
Department of Computer Science
Golisano College of Computing and Information Sciences
Rochester Institute of Technology
Rochester, NY 14623
(ip1262, tr7550, lc2919)@rit.edu

***Abstract*— This paper talks about building of an expert system for evaluating security of a Windows platform. It contains study about various paper describing a similar type of problem to be implemented. This paper also contains ideas very specific to the teams' requirement with rather in depth elucidation of how the flow of the project is done. A couple of new ideas are proposed which makes this project uniquely define itself among others.**

Keywords—expert system; security evaluation; OS security; knowledge base; machine learning; neural networks.

I. Introduction

Security evaluation means determine how secure a platform is based on some parameters defined. Initially this process was conducted manually but now with the help of AI and some machine learning algorithm we can automate this process. In this paper we have specifically focused on OS security, i.e. we will be evaluating security of a certain Windows platform. This can be using expert systems. An expert system is a computer program that uses AI technologies to simulate the judgment and behavior of a human or an organization that has expert knowledge and experience in a particular field. Additionally, there is knowledge base contained trained dataset which can be used as a basis to conclude desired results. The final goal is to make the system work in a similar way as a human works.

II. LITERATURE SURVEY

A. Paper 1[1]

On daily basis, all kinds of business activities use information technologies. This is the reason information becomes the most valuable resource of the organization. The main part of any organization related security operation is to examine the information security. The intent of auditing is to know the information security management and recommend few solutions for the developing expert system. This paper is based on international standards on information security and protection tools which discusses the development of an application.

With the arrival of recent intrusive and virus attacks, the need of effective computer security measures has become essential. Post facto security auditing has come into view to handle the abnormal and anomalous events that cannot be addressed in real time. The use of special tools such as questionnaires and checklists are one of the efforts taken in reducing expenses and facilitating audit. This is used to identify gaps between certain security standards and existing organization's security practices. A process of asking questions and making conclusions from answers can be sighted as auditing process.

According to this paper which includes one more effective tool for the audit is to develop a knowledge base which will provide information for Chief Information Security Officers and it will help them to find the right management decisions on the information security policy. The key components of knowledge base are Asset, Source (standard), Vulnerability, Step (a refinement of the part of Guideline in special section of the standard) and others. Every Step introduced to the protected Object, to the cross-references to other stored Guidelines as well as to the type of Vulnerability it is against. This instrument supply analysis of the components, standards, search-based knowledge management directives, and issuing recommendation which can be established.

The process of auditing the Information security can be extremely expensive. For any organization, the priority is the reduction of the cost of the audit process. For this problem, Automating the audit process by creating expert system can significantly reduce the cost, as the main work on decision making is executed automatically, which is based on computer analysis of the situation and issuing guidelines and recommendations.

In case of Information Security audit automation, expert systems have many things to offer. Firstly, the expert systems approach fits question answer format of auditing. Secondly, expert system functions on the basis of meta model that reflects knowledge in target field. The way expert thinks in a particular field, mirroring and implementing common human logic can give a system, ability to estimate the situation and make decisions.

The development of expert system is based on the platform, which keeps the knowledge base and automated gathers information of the objects, analyzes using fuzzy logic and makes recommendations based on the results. By using an expert system, apart from saving time, we can eliminate the possibility of errors in the calculation of results.

The architecture of expert system contains five parts: Database, interface for experts, interface for risk managers, interface for analytics and interface for information security officers. The database is the main component of expert system as different components directly interacts with it. It contains questions, the list of users, answer, question weights, risk levels, recommendations, analysis results and tools. For the second part, Interface for Information Security is considered. After passing the authorization phase, they determine the range of questions as set of linguistic variables like LOW, MEDIUM and HIGH which is relevant to set of numeric values. The authorization and evaluation of risk level of questions which is an interface for risk managers that covers the third part. For the fourth part, the interface for analytics covers authorization and own interface, in which analytics can run different calculations of results and take output results. Interface for providing output based on the output is the last part.

In conclusion, there are several software designed to help person decision making according to requirements and need of the organization. This process is ease of decision making processes which may bring telling practical value in information security and fuzzy expert system development. There are still lot of remaining areas for reducing issues and developing the expert systems.

B. Paper 2[2]

As the technology is getting advance day by day, various components of the computer networks are getting connected and hence they are prone to cyber-attack. This paper gives a Fuzzy Rule based expert system which can secure and provide robustness to a network from
such cyber-attacks.

Basically this system uses Fuzzy logic and Triangular Fuzzy Number for computation. The expert system is based on rules which are made by defining input and output variables. For this, initially domain experts are consulted to make a dataset for the rules. This step is done to understand the attacks which can be done on the network and how should an expert system react in such situations. Various attacks which can be done on the network can be Denial of Service attack, Logic bombs etc. Also, the expert system is expected provide explanation similar to human expert and hence the input from the domain experts in vital.

Furthermore, this expert system is designed to forward chained, i.e a rule is formulated in this system as if A, then B. The 3 basic components of this User interface, Decision making inference engine and Database. The database is used to store the fuzzy rules and other datasets. The user of this system interacts with the system via the user interface. The inference engine takes input from the user and gives a decision/ suggestion based on the fuzzy rules stored in the database.

Vital part of this expert system is defining the Fuzzy Rule based model. Initially, a fuzzification module is used to convert the crisps input to a grade by a fuzzy set. This is done by the Triangular Membership Function. Later fuzzy rules are set which from the base of the inference.

There are various examples of the rules given in the paper by the author which are based on factors like Cyber Intruders aim or target. Later and defuzzification module is used as bridge between the fuzzy logic control and inference system as the output given to the user should be crisps.

C. Paper 3[3]

As the world is advancing technologically day by day, information security is the utmost priority. Audit of the information security in organizations is and option but its time consuming and costly. And hence, less costly way of automation of security evaluation and implementations can is necessary. Basically, Expert System(ES) tries to replicate the way a domain expert would think and behave in a certain situation.

There are 3 stages in developing the Expert System for Information Security.

a) Building a high-level structure of the knowledge base for the IS.
b) Development of the system workflow.
c) Developing a methodology for population of the knowledge base.

In order to create a Knowledge Base and Ontology of the IS domain is created. This includes terms like Threat, Vulnerability, Control and Asset. These terms are very well explained in the paper. These terms form the knowledge model on top which the knowledge base is made.

Now the next step is to collect the data to store in knowledge base and this is done by asking a set of questions of various forms to the organization's employees. After these data points are collected by the system, it now tries to find the threats relevant to the assets of the company. Next, several questions related to the control measures of the organization are asked. The control measures can be the frequency of the data backup.

Furthermore, a classification of the threats and their effects is calculated by analysis of the logical chains. These logical chains may be constructed by using security policy and analyzing the possible risks. This classification is done as subjective and objective by the author. In the paper, author have written in details about such classification of threat.

Later author talks about the classification of the vulnerabilities. Classification of vulnerabilities is necessary as threats can exploit vulnerabilities to break safety information to obtain illegal benefits. Author has explained this classification in details in the paper.

Moreover, the author has explained a real life security state scenario of a university and have made detailed analysis and classification of the threats, vulnerabilities and controls. This analysis of author helps us to think about any other scenario and write down the threats, vulnerabilities and controls for it. Also, based on these 3 factors, author has made if then rules. This helps us to define appropriate rules according to ISO standards. And according to controls expert system can generate recommendations for any particular situation.

Thus this paper explained to us the importance of the identification of the assets and need to set the initial level of security that meets in the Information Systems. The analysis of this paper helped us to understand the ontological scheme of the subject area with in-depth branches and the necessity of the classification of the vulnerability, threat and the asset.

As further in the project, we plan to implement our own expert system, building a knowledge for our system will be benefited by this study. Also, in one of the papers we have studied the use of designing a Fuzzy Rule Based Expert System. The study of current paper will help us to make stronger knowledge base and if else rules which can be effectively used by the Fuzzy Rule Based Expert System.

D. Paper 4[4]

In the new environment of trading where development of e-commerce especially the type B2C(business-to-consumer) has increased exceptionally. The critical fact for success of e-commerce is considered as Trust. This paper presents an application of expert system on trust in e-commerce. The challenges and new problems with online trading to online buyers is uncertainty about quality of products or services and sellers can be anonymous which lead in high level of risk in online transaction environments, virtual communities and online auctions. The purpose is to build a trust management system which can help to reduce risk and make it easier for buyer and seller to interact with each other in low risk environment.

Analytical Hierarchy Process (AHP) is a mathematical technique which is used for multi-criteria decision-making. The use of AHP leads to more transparency of the quality of management decisions. Based on decision maker's understanding of the problem, the hierarchy can be designed and pair wise comparisons can be made of the decision elements. AHP uses redundant judgements for checking consistency and it can exponentially increase the number of judgements to be taken out from decision makers.

The information can be uncertain. The human thinking process can handle inexact, uncertain and vague concepts in appropriate manner but cannot be expressed or measured in mathematical way. To overcome this, Fuzzy logic provides a framework to model uncertainty, the human way of thinking, reasoning and the perception process. A Fuzzy expert system is an expert system that uses a collection of fuzzy membership functions and rules, instead of Boolean logic, to reason about data. For the extraction of models from numerical data representing the behavior of a system, we can use Neuro-fuzzy modeling. Without loss of information, learning capability of feedforward neural networks supports the model extraction if the architecture of the network, once properly trained, may be translated into rules.

The basic architecture of a fuzzy expert system includes fuzzification interface, a fuzzy rule base (knowledge base), an interface engine (decision making logic) and a defuzzification interface. Neuro-fuzzy is mixture of fuzzy logic and neural networks to give a system of postulates, data and interfaces to describe an object or process. One example of Neuro-fuzzy systems is Adaptive Network Fuzzy Interface System (ANFIS), which has good software support. For modeling a nonlinear function, a dynamic system identification and a chaotic time series prediction, we can use AFNIS. The AFNIS architecture chosen as given its potential in building fuzzy models with good prediction capabilities.

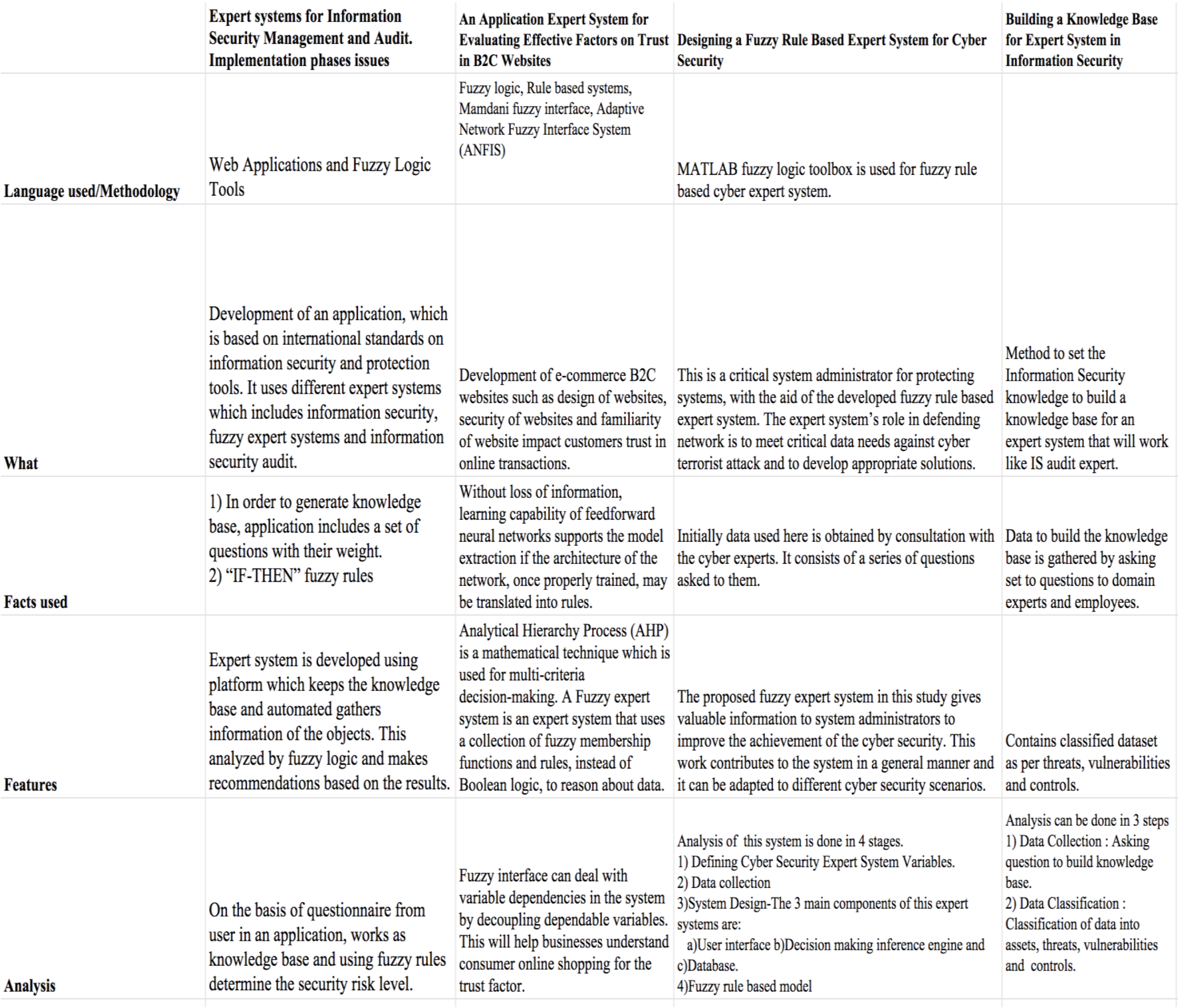
The general trust model is composed of two modules. First model will be used to quantify the trust measure on the basis of three factors which can be identified in trust model; security, design and familiarity based on Mamdani fuzzy interface system. The second module will be same model based on AFNIS.

In conclusion, due to uncertainties involved, the trust relationships among customers and vendors are hard to assess. Advantages of using fuzzy logic to quantify trust in e-commerce applications are: Fuzzy interface is capable of quantifying imprecise data and uncertainty in measuring the trust index of the vendors. Also, Fuzzy interface can deal with variable dependencies in the system by decoupling dependable variables. The result of this paper study will help businesses understand consumer online shopping for the trust factor.

E. Conclusions

The comparative study given below shows the comparative study for four papers: Expert systems for Information Security Management and Audit. Implementation phase issues, An Application Expert System for Evaluating Effective Factors on Trust in B2C Websites, Designing a Fuzzy Rule Based Expert System for Cyber Security and Building a Knowledge Base for Expert System in Information Security. Based on the factors of comparison i.e., Language used/Methodology, What does paper infer about, Facts used in the paper, Features described, Analysis of the expert system in individual paper. The comparative study based on language used/methodology is Web Application and Fuzzy Logic Tools, Fuzzy logic, Rules based systems, Mamdani fuzzy interface, Adaptive Network Fuzzy Interface System(ANFIS) and MATLAB fuzzy logic toolbox is used for fuzzy rule based cyber expert system. The comparative study for all four papers are useful for determining the security evaluation in different platform and applications. These comparison will be useful when modeling the expert system. While designing the expert systems, considering these facts and use of the expert system will let the application to be more secure and fits for the type of model to be used according to the situation. These comparison study will give the information about what type of expert system is useful for which kind of application and which kind of application is more secure or less secure based on the factors. These papers were useful in determining various metrics and algorithms which will be used in our project.

F. Comparative Study



	Expert systems for Information Security Management and Audit. Implementation phases issues	**An Application Expert System for Evaluating Effective Factors on Trust in B2C Websites**	**Designing a Fuzzy Rule Based Expert System for Cyber Security**	**Building a Knowledge Base for Expert System in Information Security**
Language used/Methodology	Web Applications and Fuzzy Logic Tools	Fuzzy logic, Rule based systems, Mamdani fuzzy interface, Adaptive Network Fuzzy Interface System (ANFIS)	MATLAB fuzzy logic toolbox is used for fuzzy rule based cyber expert system.	
What	Development of an application, which is based on international standards on information security and protection tools. It uses different expert systems which includes information security, fuzzy expert systems and information security audit.	Development of e-commerce B2C websites such as design of websites, security of websites and familiarity of website impact customers trust in online transactions.	This is a critical system administrator for protecting systems, with the aid of the developed fuzzy rule based expert system. The expert system's role in defending network is to meet critical data needs against cyber terrorist attack and to develop appropriate solutions.	Method to set the Information Security knowledge to build a knowledge base for an expert system that will work like IS audit expert.
Facts used	1) In order to generate knowledge base, application includes a set of questions with their weight. 2) "IF-THEN" fuzzy rules	Without loss of information, learning capability of feedforward neural networks supports the model extraction if the architecture of the network, once properly trained, may be translated into rules.	Initially data used here is obtained by consultation with the cyber experts. It consists of a series of questions asked to them.	Data to build the knowledge base is gathered by asking set to questions to domain experts and employees.
Features	Expert system is developed using platform which keeps the knowledge base and automated gathers information of the objects. This analyzed by fuzzy logic and makes recommendations based on the results.	Analytical Hierarchy Process (AHP) is a mathematical technique which is used for multi-criteria decision-making. A Fuzzy expert system is an expert system that uses a collection of fuzzy membership functions and rules, instead of Boolean logic, to reason about data.	The proposed fuzzy expert system in this study gives valuable information to system administrators to improve the achievement of the cyber security. This work contributes to the system in a general manner and it can be adapted to different cyber security scenarios.	Contains classified dataset as per threats, vulnerabilities and controls.
Analysis	On the basis of questionnaire from user in an application, works as knowledge base and using fuzzy rules determine the security risk level.	Fuzzy interface can deal with variable dependencies in the system by decoupling dependable variables. This will help businesses understand consumer online shopping for the trust factor.	Analysis of this system is done in 4 stages. 1) Defining Cyber Security Expert System Variables. 2) Data collection 3)System Design-The 3 main components of this expert systems are: a)User interface b)Decision making inference engine and c)Database. 4)Fuzzy rule based model	Analysis can be done in 3 steps 1) Data Collection : Asking question to build knowledge base. 2) Data Classification : Classification of data into assets, threats, vulnerabilities and controls.

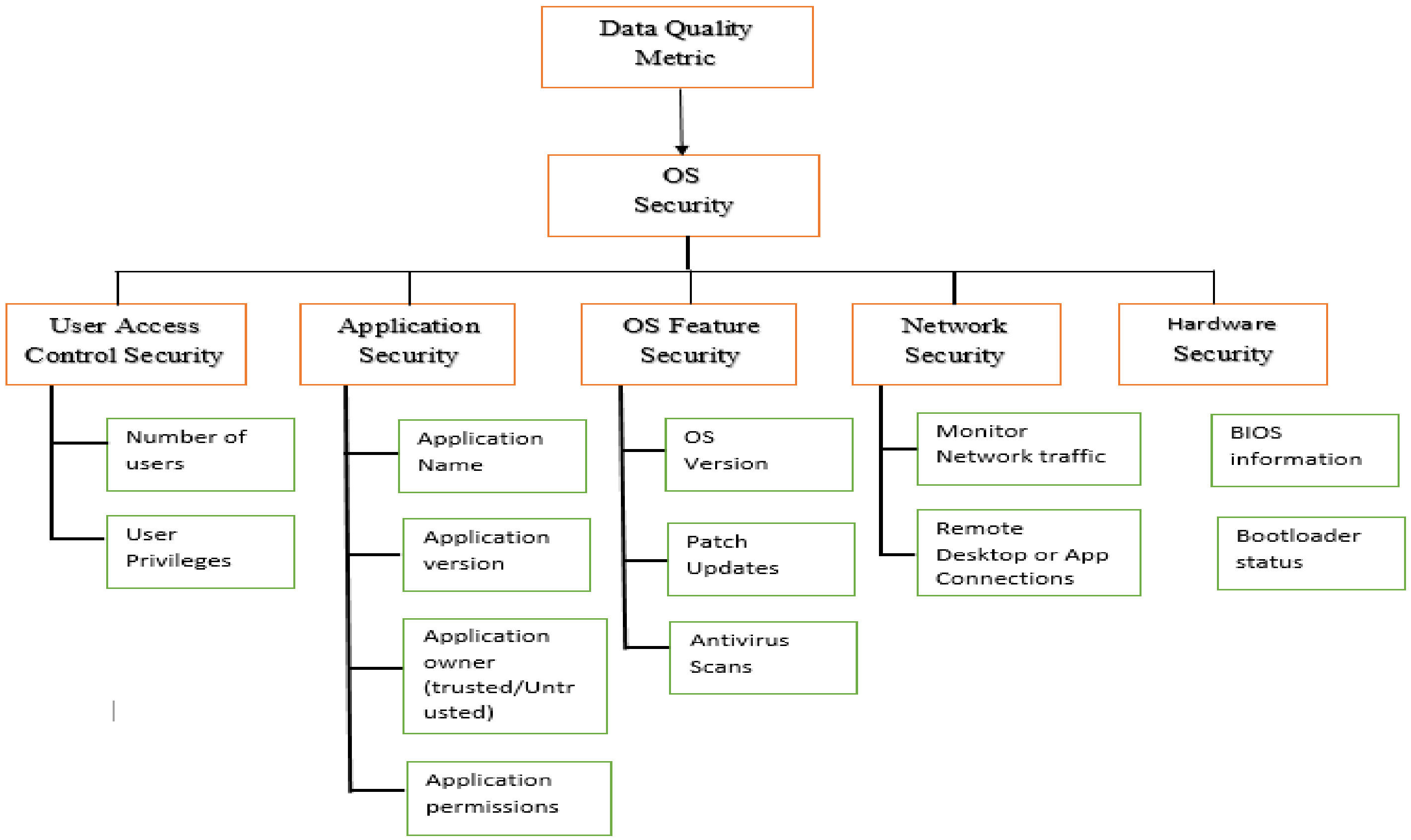


Figure 1: Metric Tree for security evaluation for Windows platform

III. Project 2 Specification

This expert system evaluates the security of any OS based platform, taking in the information obtained from the system itself or from the user, it generates an inference based on this information and gives out the security measure of that particular system.

Security now-a-days is playing an important role for assessing the vulnerability of the system and securing it from any possible threats. We have chosen Windows as our OS platform to access its security using our expert system. The reason for choosing Windows was, after conducting some research we concluded that Windows is the least secured OS platform among all other platforms.

A. Approach And Basic Flow of the Project

Our approach to solving this problem was very simple, choose few pragmatic metrics, obtain relevant information from the system, compare it with our knowledge base using predefined rule set, run our trained machine learning algorithm on this data and finally give the conclusion in the form of rating, which will

flowchart gives a fundamental flow of how all the process works:

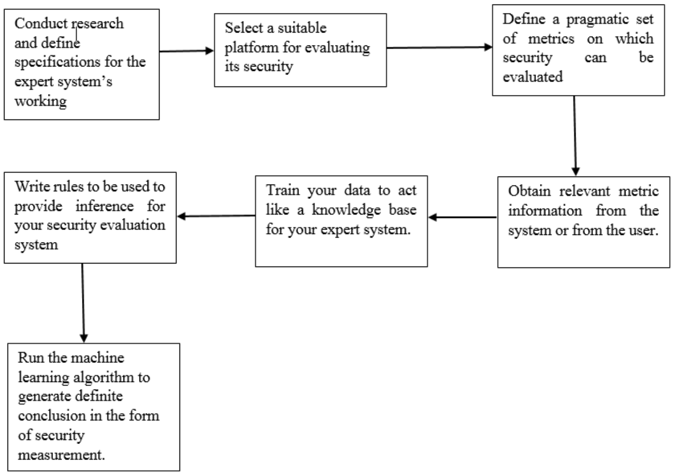

Conduct research and define specifications for the expert system's working
Select a suitable platform for evaluating its security
Define a pragmatic set of metrics on which security can be evaluated
Obtain relevant metric information from the system or from the user.
Train your data to act like a knowledge base for your expert system.
Write rules to be used to provide inference for your security evaluation system
Run the machine learning algorithm to generate definite conclusion in the form of security measurement.

Figure 2: Basic flow of the project

B. Metrics

determine how secure the system in question is. The following

The metrics in this project provide a skeleton on which the security would be evaluated. We have constructed a metric tree which classifies different parameters on which the security of the selected platform can be evaluated. We have started by categorizing it by labeling them as Data Quality metrics with OS security as the major focus of classification. The security of the platform can be evaluated based on these major parameters/metrics, user access control security, application security, OS feature security, Network security,

Antivirus security etc. Figure 2 shows the metric tree.

For our project we will not be using all the metrics specified, the reason been some of the information which is required to be obtained for these metrics can be challenging to acquire.

While implementing our expert system, taking into consideration all the metrics can be very difficult hence we have narrowed down our metrics on which we will be building our expert system. Figure 3 shows the condensed metrics used to evaluate the security of the said platform.

C. Implementation

The implementation of this expert system requires metrics, which we have defined in the previous section, and a knowledge base.

Metrics to be used: Implementing all the metrics specified in the Figure 000 can be a bit of an arduous task, hence we have condensed it to a few metrics which be pragmatic to be implemented plus can cover all the elements required to conclude the security of the system.

Metric	Values
Patch Updates	Yes – latest date, No – Others
Antivirus Scan	Yes – Performed, No – Otherwise
Application Status	No – Potentially Harmful, Yes - Otherwise
Application version	Yes – Latest, No – Not latest
OS version	Yes – Latest, No - Otherwise
Windows Defender	Yes – If active, No – Not active

Facts Gathering: The method which we will use for gathering our facts will be both directly from the system (automatically) or in the form of user inputs (manually).

Knowledge Base: The knowledge base in our expert system will act as a database of rule sets which after receiving the relevant data from the system or the user, will provide us a standard data on which comparison will be done and inference model will be built.

This knowledge base will be a rule of IF-THEN statements, where IF will be a condition which will be checked against the data received and THEN statement will act as an input to the inference model to be built. The following gives an example as to how the rules will be implemented:

1) ***IF*** *patchUpdateTime = latestDate*
   THEN *patchUpdateTime = Yes;*

2) ***IF*** *anitvirusScan = performed*
   THEN *anitvirusScan = Yes*;

Inference Model: The inference model for this expert system takes in the answers obtained from the THEN statement and creates a data set. This data set after obtaining all the required data is trained, this is done to achieve maximum accuracy of results after the ML algorithm is run over this trained data.

patchUpdateTime	*anitvirusScan*	*OSVersion*	*TargetSolution*
Yes	Yes	Yes	High
Yes	No	Yes	Medium
No	No	No	Low

Figure 3: Example of dataset

The above table gives an example of how the data after training will give out conclusions based on the values of the metrics. We have considered Yes and No values for the attributes to understand a particular metric holds true or not. For example, in the table we have taken few metrics which will decide the security of the system, so if all the metrics have an attribute value of *Yes* then it can easily be concluded that the security of that particular Windows system is High, i.e. the system is highly secured.

The training of the data can be done using Boolean operations where Yes is the truth value and No is the false value, additionally using an and operation to obtain a definite value which will help our ML model to decide the *TargetSolution* value. This value is nothing but the conclusion of our expert system.

User Interface: The user interface of our expert system will have a user window which will take inputs form the user, if required, and a window which displays their security level. Additionally, there will be window which will display the reasons why their security level was Low, and possibly a solution on how they can improve their security.

Hardware Requirements: Any Windows based platform which will have support for Python.

Software Requirements: Python IDE for compiling the program, preferably python version for after 2.7

IV. PROJECT 3 SPECIFICATION

In our project, we are using Machine Learning algorithm which is based on supervised learning.

Supervised learning is defined as an algorithm which has a definite output solution, i.e. a target attribute which gives a meaningful inference. The attributes used for this algorithm will be the metrics specified in the project 2 specification, Figure 000. The inference obtained from this algorithm will determine the final security level of the system in question. From the raw data obtained from the user or system directly, we will build a dataset containing all the metric attributes plus a TargetSolution attribute which will be our final conclusion to the security decision to be made. This dataset will be trained by generating combinations of various attribute values, and the similar value data cluster will have the same target attribute value. Based on this trained dataset and the information obtained, a final verdict will be presented to the user.

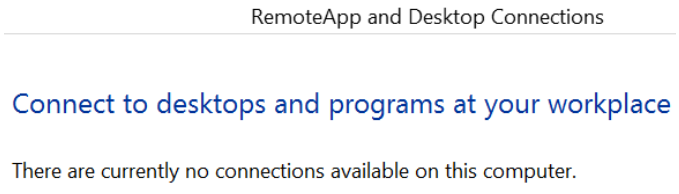
For instance, if the system has a torrent software installed then the security will without a doubt have a target attribute value as Low security. In this way if attribute value having an apparent value, does not need to take into consideration any other attribute values for deciding the security of the systems.

V. NOVEL AND INTERESTING IDEAS

Initially our expert system idea was limited to an input-output based model. Few ideas which can be possibly implemented and which will enhance the working of our ES:

1) Going ahead with it we thought of adding a new component to our interfence model, i.e. we decided to display the reason why the security of the said sytem is Low, this will be only true of there is a valid reason for the security to be Low/Medium. We cannot have any specific reasons for systems which are highly secured. This can be implemented during the fact gathering phase, while taking in the inouts from the user or directly from the system we will be storiing this information into our knowledge base so after we get the final result we can check the values from the trained datset, we can collect the attributes with values as No and the get the corresponding values from our knowledge base, which can be then displayed on th euser interface. The advantage of this will give reassurance to the user that there is a chance to repair the damage that can be caused by having Low security for the system. Additionally, the system will be secured from any vulnerabilities or threats that can cause serious harm to the system and its crucial data.

2) We have also considered a different type of metric that will check any remote connection going from or to the system. This metric can be considered a classification of the network security. The implemetntaion of this can be a little bit complex because for the facts gathering we need to obtain the incoming and outgoing network connections, this can be obtained from RemoteApp and Desktop connections.



RemoteApp and Desktop Connections

Connect to desktops and programs at your workplace

There are currently no connections available on this computer.

Figure 4: Screenshot showing any remote or desktop connections in Windows

We can check from which IP the connection is established and whether the IP is private or public. If the IP came out ot be public we can check if it is a malicious IP or not. Additionally, we can also check the source from which the connection has been established. But again this can be a far-fetched notion and will probably be a bit complex to implement, with the auxiliary requirement of user intervention.

VI. REFERENCES

[1] Maksat Kanatov, Lyazzat Atymtayeva, Bagdat Yagaliyeva, “Expert systems for Information Security Management and Audit. Implementation phase issues.”

[2] Mehrbakhsh Nilashi, Karamollah Bagherifard, Othman Ibrahim, Nasim Janahmadi, Mousa Barisami, “An Application Expert System for Evaluating Effective Factors on Trust in B2C Websites”.

[3] Kerim Göztepe, ”Designing a Fuzzy Rule Based Expert System for Cyber Security”.

[4] L. Atymtayeva, K. Kozhakhmet, G. Bortsova, “Building a Knowledge Base for Expert System in Information Security”