World Academics Journal of
Engineering Sciences
Vol.8, Issue.2, pp.18-23, June (2021)

Research Paper

E-ISSN: 2348-635X

Identifying Internet Crime Using Mamdani Fuzzy Rules and Fuzzy Classification Algorithm

Musa Mojarad[1*], Afsaneh Haghanin Nia[2], Neda Galehdari[3], Khayronsa Bazarganzadeh[4]

[1]Dept. of Computer Engineering, Firoozabad Branch, Islamic Azad University, Firoozabad, Iran
[2,3,4]Dept. of Computer Engineering, Liyan Institute of Education, Bushehr, Iran

[*]Corresponding Author: m.mojarad@iauf.ac.ir, Tel.: +98-91788-63397

Available online at: www.isroset.org

Received: 30/Apr/2021, Accepted: 20/Jun/2021, Online: 30/Jun/2021

***Abstract*—**Crime prevention has always been one of the basic and important issues in human life that has been applied in different ways throughout history. In this research, fuzzy decision tree classification model, Mamdani fuzzy rules are used to identify cybercrime. The aim of this study is to identify the important features of specific crimes and classify crimes into three different categories. We use only a few features for classification work, which reduces the size of the collection. Reducing the complexity of the tree reduces the fuzzy decision for classifying mass data sets. In order to improve the results of fuzzy decision tree classification, the optimal number of language semesters in fuzzy sets for each feature is identified. Experiments have been performed to evaluate the proposed model on the actual data set of the crime and the results of the proposed model in the average mode show an accuracy of 98,13% in detecting the type crime. The results show the superiority of the proposed method compared to other methods.

Keywords*—Internet crime detection, Fuzzy decision tree, Mamdani fuzzy rules, Feature selection.*

I. INTRODUCTION

Crime prevention is one of the basic and important issues in human life that has been studied in different ways throughout history. Due to the development of information and communication technologies and the launch of comprehensive information systems in police and judicial organizations, as well as law enforcement and registration of criminals in databases in order to identify crime and detect crimes is one of the necessities. Of the Iranian police and the judiciary. Undoubtedly, human social conditions make confronting the phenomenon of crime inevitable [1].
Today, the expansion of databases and their processing plays an important role in detecting crime patterns for security police organizations. Meanwhile, data mining methods and artificial intelligence processing are powerful and at the same time low-cost tools for pattern discovery and knowledge extraction from the database in line with the decision-making process for crime prevention and control. Efforts made by developed countries towards data mining applications in the field of crime analysis confirm that there are many capabilities in discovering the rules of crime and preventing its occurrence [2, 3]. This study uses documentary methods (libraries) to answer questions about the role of data mining methods and artificial intelligence in the prevention of violent crime, what are the most widely used algorithms and methods to discover the rules of crime and analyse how this process works. Methods are.
Cybercrime encompasses a wide range of offenses and includes a variety of catastrophic crime harassment that occurs in a virtual environment. Cybercrime is short-lived and has only become popular in the last 20 years. Third-generation crime is computer-dependent and Internet-dependent, and can be committed in cyberspace or cyberspace. Meanwhile, scientific and intelligent crime detection solutions have attracted the attention of many criminologists due to their scientific background and mathematical knowledge. One of these solutions is data mining. Data mining is a process that uses intelligent methods to extract knowledge from a set of data [4].

In the continuation of this research, the problem statement is discussed in section II and then we review some of the latest work done in section III. In Section IV, the proposed model based on Mamdani fuzzy rules and fuzzy classification algorithm for detecting cybercrime is presented and the necessary parameters are presented. The results of evaluating the proposed method and discussing it are given in Section V, and finally the conclusions and suggestions are listed in Section VI.

II. STATEMENT OF PROBLEM

With the prevalence of the use of computers, the Internet and electronic devices such as mobile phones, payment devices, etc. in personal life and office relationships, delinquency and abuse in the use of these tools are also inevitable events. What are now referred to as cybercrime and cybercrime is a collection of the same offenses and crimes that occur through computers and electronic or computer-influenced devices, and there are numerous examples of this. It has a role in our minds [1].

In this research, to achieve the set goals, we use real and organized data in order to create a cybercrime detection model. The data set used is valid and is from the UCI repository. Our goal is to use data mining techniques to create a diagnostic model based on existing data and to use it to predict cybercrime in new data. In general, data mining is the process of discovering patterns and regular and hidden processes in large and distributed data, using a wide range of algorithms based on mathematical sciences and statistics. Data mining techniques use sciences such as artificial intelligence, machine learning, statistics, operational research, and database management to build models and answer questions. Extracting and analyzing organizational information from data available to employees is a process that has been done for many years and is not a new task in organizations [5]. But in recent years, with the growing computing power of computers and the possibility of achieving results from complex computing in a short time, has led to the development of more advanced algorithms. These algorithms refine and analyze the data by considering different dimensions of the data and extract and present complex and unrecognizable patterns by old methods.

In this paper, we use classification techniques to create a crime detection model. Classification is one of the learning methods with the observer for predicting data categories, which identifies new data classes based on default and predefined classes. Some data classification algorithms include decision tree, backup vector machine, regression logistics, NEOBIZ, and so on. Since detecting the behaviour of criminals in the Internet system is associated with uncertainty and the records of transactions can be helpful in understanding these movements, so in this study we use fuzzy logic to deduce the rules and detect different behaviour's.

III. Related Work

Crime analysis in a sense refers to the implementation of a particular discipline in the police community. This analysis includes more than one type of crime, which is why some authors refer to it as a public safety analysis. In general, data mining methods and algorithms are divided into three general categories: crime detection, crime prediction, crime analysis and detection [2]. Figure 1 shows the methods used in crime data mining.

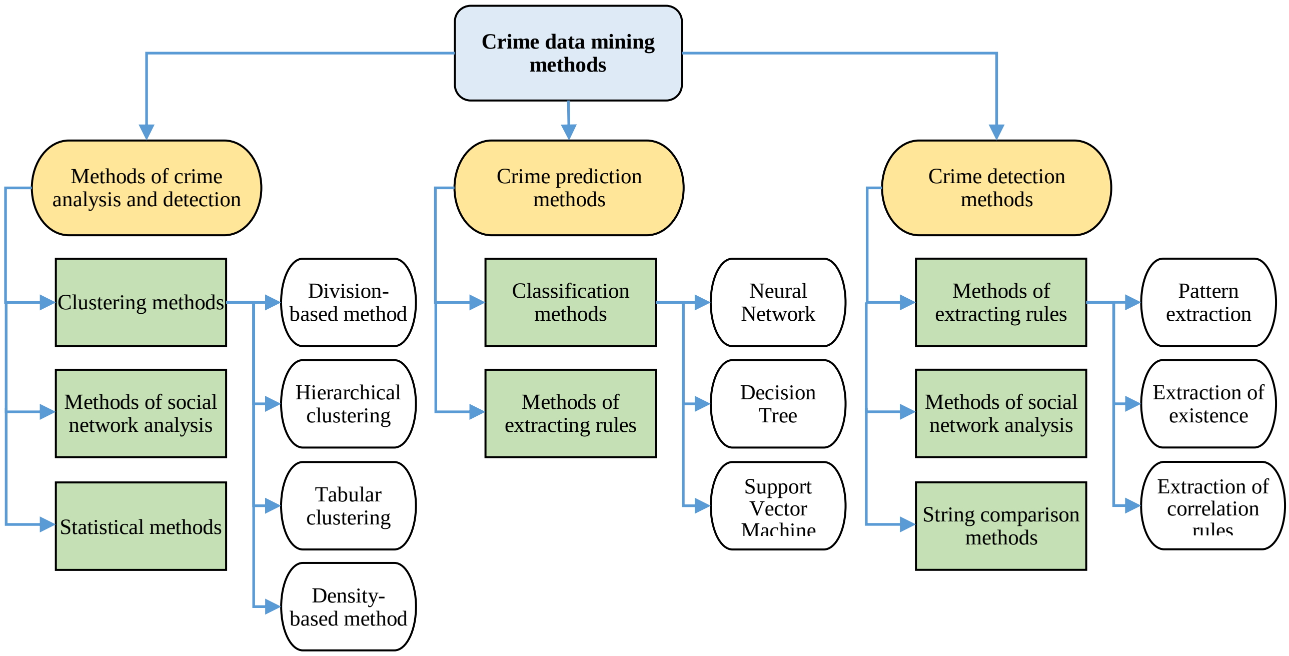

Figure 1. Crime data mining methods

Here are some examples of research into the use of different data mining techniques in crime modeling. Inouye et al. (2017) presented a model based on clustering technique to identify and group different types of crimes [3]. Moon et al. (2013) also used regression to predict cybercrime. Based on the results, the number of hours of computer use and membership in groups and Internet networks increased the rate of cybercrime and were introduced as the main variables predicting the rate of crime [4]. McMahon et al. (2015) reviewed the applications of data mining techniques in cybercrime [5]. Li et al. (2010) presented a decision support model based on fuzzy construction technique to identify and analyze patterns and trends in crime occurrence. This model is implemented in the data of the Taiwan International Police [6]. Gerber et al. (2014) presented a model for predicting crime scene next week based on current week data that could be very effective in crime prevention [7]. In this regard, Tuominen et al. (2017) conducted a study on the moral demographic characteristics of recidivists in the city of Izmit [8].

In another study, Akers (2017) presented data mining techniques to predict and prevent social media crimes in the Internet environment [9]. One of the joint studies conducted by the British Police and the Department of Psychology at the University of Sunderland in the field of application of data mining techniques in the police force is a study conducted by Prabakaran and Mitra (2018) [10]. Chastain (2016) used time series to predict the date of occurrence of thefts that are repeated in one day [11]. Li et al. (2010) presented a decision support model based on the SOM fuzzy technique for identifying and analyzing patterns and trends in crime occurrence [12]. Annabathula (2007) study states that after the 9/11 terrorist attacks, the CIA, the FBI, and other federal agencies decided to gather internal and external security information to prevent terrorist attacks [13].

IV. METHODOLOGY

This section presents a hybrid model based on Mamdani fuzzy rules and fuzzy decision tree classification algorithm to identify cybercrime. Through this system, the police can make a difference between the two parties and their votes and votes in each of them.

The classical fuzzy decision tree fuzzies the values of each property with respect to a fuzzy set with a different number of terms, and then creates a rule base based on the segmentation performed. Often fuzzy sets and membership grade functions are provided by an expert based on the importance of the features. Membership degree functions can include one or more fuzzy sets with different decision values. For example, a two-state membership degree function has two fuzzy sets, Low and High. The greater the number of cases (number of fuzzy subsets), the more cases the decision process covers, and this ultimately leads to increased accuracy in the classification model. In general, increasing the number of language terms in fuzzy sets increases the complexity and production of large-scale fuzzy decision trees, so it is necessary to determine the optimal number of language terms for each feature.

In this research, the number of terms of fuzzy sets is optimally searched for each feature by trial and error. For example, for the first attribute we consider a membership degree function with three values A_1, A_2 and A_3 and for the second attribute we consider a membership degree function with two values B_1 and B_2. Therefore, in line with the objectives set for each feature, fuzzy sets with a different number of A_n (linguistic terms with a different number n) are assigned. A_n is searched according to the separation values and the importance of each property in the sample classification (class recognition).

In the proposed method, before creating a fuzzy decision tree classification model in order to statistically reduce the size of data related to cybercrime, the data pre-processing step and selection of effective features is performed. For the selection of effective features, we use the features used in other articles on the data set used in this research. In order to build the whole model, the data set samples are divided into two sets: training (E^T) and experimental (E^P). E^T is used to design the model and E^P is used to evaluate and test the classification model. Finally, to create a model for classifying cybercrime data, we use the fuzzy decision tree based on the number of optimal fuzzy sets, and from the fuzzy tree, the fuzzy rules database to predict the input data by the method of Mamdani We use. In the fuzzy decision tree, the number of fuzzy subsets as well as threshold values are calculated by error attempt and stored as the BestGlobal variable for each attribute.

A. Data pre-processing

Pre-processing is one of the most important parts of data mining models because bad input leads to bad output. Pre-processing is used to improve the quality of real data in processing models such as classification. In this study, data processing includes "preparation" and "normalization". In the preparation step, we remove the fields from the database that have no role in classification. Also, the information of some properties is missing for some samples, so we remove samples with missing values from the database. Missing values are often marked with a "?" Have been specified.

Normalization; we normalize the properties of the properties when they are in different ranges. This causes the values of all attributes to be in the same range. In this research, we use the Max-Min method to normalize the properties in the range [0-1] according to Eq. (1).

$$X_{i,f}^{new} = \frac{X_{i,f} - \min(X_f)}{\max(X_f) - \min(X_f)} \qquad (1)$$

Where, $X_{i,f}$ and $X_{i,f}^{new}$ are the actual and normalized values of the i-th example of the feature f, respectively. $\min(X_f)$ and $\max(X_f)$ are the smallest and largest values of all instances in the f feature, respectively.

B. Classify data with fuzzy decision tree

The purpose of this section is to create a fuzzy rule database based on the fuzzy decision tree. The division of attribute values in the fuzzy decision tree is determined by trial and error, and the fuzzy rules extracted from the fuzzy decision tree are argued based on the Mamdani method for identifying the target class.

Here we consider a fuzzy set with different divisions (different language terms) for each feature. For example, Figure 2 defines a trapezoidal fuzzy set with 5 modes VeryLow, Low, Medium, High and VeryHigh with 10 decision values (T_1 to T_{10}) for an feature. Figure 2 (a) shows a trapezoidal fuzzy set with 5 states and 10 decision values, which assigns the value of the input property to one of the 5 sets that belongs more according to the membership degree function. Figure 2 (b) shows that it is assigned to the input property of a fuzzy subset with $n = 5$ (A_i, $\forall\, i = 1,2,3,4,5$). Finding n in this research is searched for each feature in a specific range by trial and error.

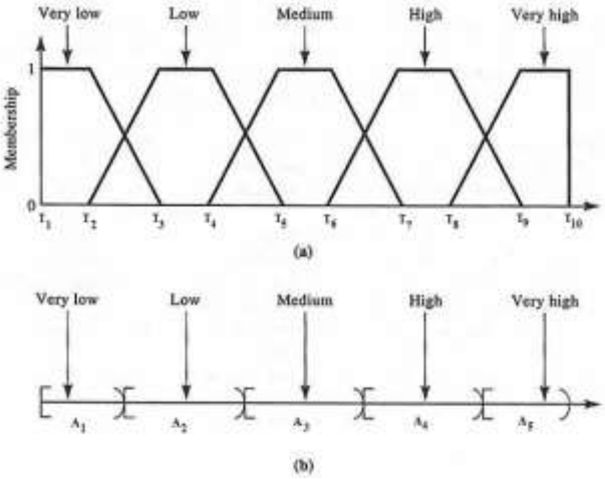

Figure 2. Trapezoidal fuzzy set with 5 modes and 10 decision values

In order to draw a decision tree, here we try to reduce the amount of impurities in the sets created in each property to use it to create a tree with a minimum height at the end. The entropy of a set is the amount of impurity in that set of data, also called the disorder of a set. The entropy of a data set is calculated by the following equation, in which P(c) represents the proportions of the data belonging to class c.

$$I(S) = \sum_{c} p(c)log_2 p(c) \tag{2}$$

In this method, first for all properties, the irregularity of the entropy relation is calculated and then using this value for all properties, the usefulness of the information of that property is defined as Eq. (3).

$$Gain(A) = I(S) - I_{res}(A) \tag{3}$$

Where, $I_{res}(A)$ indicates the amount of irregularity remaining in the data due to the selection of feature A. Using the sum of each of the features for all cases, an $I_{res}(A)$ feature can be computed as Eq. (4).

$$I_{res}(A) = -\sum_{A} p(a) \sum_{c} p(c|a)log_2 p(C|a) \tag{4}$$

Where, a will be equal to the resulting subsets (sample values for feature A) if feature A is selected. Each feature that has the highest information gain is placed at the root of the tree, then this feature is removed from the data set and the steps are repeated for the other features to complete the tree.

C. Extraction of fuzzy rules by Mamdani method

In this research, we use Mamdani fuzzy inference system to create a database of fuzzy rules. In this system, both the first part of the rules and the last part are the result (fuzzy rules). Here, the form of fuzzy rules is used as Eq. (5).

$$R_i : if\ (x_1\ \ is\ \ \hat{A}_{i1})\ and(or)(x_2\ \ is\ \ \hat{A}_{i2})\ and(or) \ldots \\ (X_m\ \ is\ \ \hat{A}_{im})\ then\ \ y_i = \ \hat{B}_i \tag{5}$$

Where, x_j is the input variable defined in the domain $\hat{A}_j$ and similarly y_i is the output variable defined in the domain $\hat{B}$. Parameter j refers to the attribute number and parameter i refers to the rule number. Each $\hat{a}_j$ is a language term from the fuzzy set in the corresponding $\hat{A}_j$. Each $\hat{a}_j \in \hat{A}_j$ is a function of the degree of membership in the form $\mu_{\hat{A}_j,a_j}$, which indicates the degree of compatibility $\hat{a}_j$ with $\hat{A}_j$. Similarly, $\hat{b}$. Is a language term from the fuzzy set $\hat{B}$, which is defined for each $\hat{b} \in \hat{B}$ membership degree function $\mu_{\hat{B},\hat{b}}$. According to the set of rules and based on the fuzzy decision tree created in the previous step, the rules are extracted from the fuzzy tree. Each path from root to leaf is a rule in a fuzzy rule database.

For a specific process, the purpose of building fuzzy inference systems is to determine the fuzzy rules governing that process. Inference in a fuzzy inference system means how the output value is determined for an input. To do this, we first fuzzy the values based on the values of the input properties for a sample. In the next step, the weight of each rule is calculated according to the inputs by the method of Mamdani. The weight of each rule is to determine the degree of compliance or degree of fire of each rule. The degree of compatibility of each training sample $X_i = \{x_1, x_2, \ldots, X_n\}$ with the introductory part of the rule $A_j = (\hat{a}_1, \hat{a}_2, \ldots, \hat{a}_n)$ is calculated using the multiplication operator (Mamdani method).

$$\mu(X_i) = \prod_{k=1}^{n} \mu_{\hat{a}_k(x_i)} \tag{6}$$

Eventually the fuzzy output becomes a non-fuzzy number. This output actually specifies the target class of the sample for mass detection.

V. Results and Discussion

In this section, to compare the superiority of the proposed method, we compare the results with a number of rule-based classification models as well as a number of the latest crime detection methods. To perform the simulation, MATLAB software version 2017 was used to simulate and analyze the proposed method. The simulations and all the tests were performed using a 5-core Intel processor with a frequency of 2.2 GHz, 4 GB of memory.

In this study, we use the Communities and Crime Unnormalized dataset to identify crimes that is obtained through UCI machine learning and is accessible from reference [17]. The information in this dataset focuses on communities in the United States. This dataset contains social and economic data from 90 law enforcement censuses. Law enforcement data from the 1990 Law Enforcement Oversight and Geographic Information and Research Data Management Statistics from the 1995 FBI UCR. This data set contains 147 features and 2215 samples in 18 different types of crimes. In this research, the type of crime (Assaults) has been investigated.

In this study, we use various criteria such as Precision, Recall, F-Measure, and Time to evaluate the results of the proposed classification model. These criteria are one of the most important indicators for evaluating the performance of a classification model and are calculated based on the irregularity matrix. At each stage of model validation, the data set is divided into two parts: training data (E^T) and experimental data (E^P). In this segmentation, 70% of the data is used for E^T and another 30% for E^P.

Table 1 shows the effective characteristics for the type of rape crime that have been selected based on the results of previous research [14]. This table also reports the results of the final segmentation for each feature.

Table 1. Selected features from the crime dataset and details of their classification

Features	Number of divisions	Symbol	Division threshold
F6	3	A	{0, 0.33, 0.66, 1}
F11	4	B	{0, 0.25, 0.5, 0.75, 1}
F37	2	C	{0, 0.5, 1}
F119	2	D	{0, 0.5, 1}

The results of this experiment show that the best classification accuracy is calculated with the number of features 4. The accuracy attributed to this feature is 95.11. Based on the selected features, the final fuzzy decision tree is shown in Figure 3. This tree is drawn in two classes: Crime (is guilty) and Non-Crime (is not guilty).

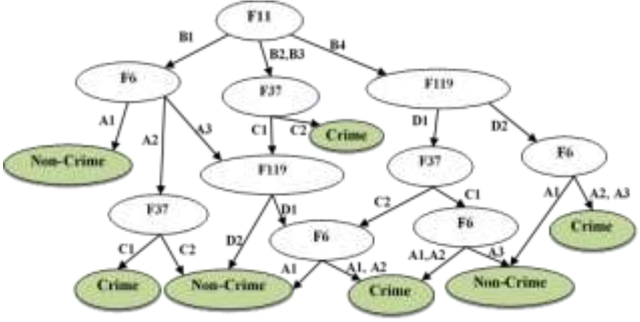

Figure 3. The final fuzzy decision tree for mass detection

In another experiment, in order to further evaluate the performance of the proposed method, a comparison is made against the three recently proposed innovative algorithms. Here, three algorithms, PSO-ANN, EA-ANN and GA-ANN, as well as the classical artificial neural network (ANN), are used to compare and evaluate the results of the proposed method. Table 2 shows the results of this comparison.

Table 2. Comparison of the performance of the proposed method with ANN, PSO-ANN, EA-ANN and GA-ANN methods

Algorithms	Precision	Recall	F-Measure	Time(s)
ANN	90.60	90.20	91.40	1480
PSO-ANN	90.20	95.00	92.50	26
EA-ANN	90.20	95.00	92.50	43
GA-ANN	90.20	94.90	92.50	108
Proposed Method	92.05	95.38	93.69	365

The results show that the proposed method is superior to other crime detection algorithms in many evaluation criteria and in other criteria has provided good results. In all experiments, the best result was obtained in the criterion F with a value of 93.69%. The execution time of the proposed algorithm is higher than other methods and it converges in 365 seconds at best. The reason for this is the two-part method proposed in search of the best features and the creation of a fuzzy decision tree classification model.

VI. Conclusion and Future Scope

In this study, a hybrid model based on Mamdani fuzzy rules and fuzzy decision tree classification algorithm to identify cybercrime was presented. Through this system, the police can make a difference between the two parties and their votes and votes in each of them. The results show that the proposed method is superior to other crime detection algorithms in many evaluation criteria and in other criteria has provided good results. In all experiments, the best result was obtained in the criterion F with a value of 93.69%. For future work, it is suggested that an effective feature selection algorithm be applied before the proposed classification model in order to reduce the number of rules and reduce the complexity of the problem.

References

[1] Soleimanian Gharehchopogh, F., & Haggi, S. (2020). An Optimization K-modes clustering algorithm with elephant herding optimization algorithm for crime clustering. *Journal of Advances in Computer Engineering and Technology*, *6*(2), 79-90.

[2] Farsi, M., Daneshkhah, A., Far, A. H., Chatrabgoun, O., & Montasari, R. (2018). Crime data mining, threat analysis and prediction. In *Cyber Criminology* (pp. 183-202). Springer, Cham.

[3] Inouye, D. I., Yang, E., Allen, G. I., & Ravikumar, P. (2017). A review of multivariate distributions for count data derived from the Poisson distribution. Wiley Interdisciplinary Reviews: Computational Statistics, 9(3), e1398.

[4] Moon, B., McCluskey, J. D., McCluskey, C. P., & Lee, S. (2013). Gender, general theory of crime and computer crime: An empirical test. International journal of offender therapy and comparative criminology, 57(4), 460-478.

[5] McMahon, R., Serrato, D., Bressler, L., & Bressler, M. (2015). Fighting cybercrime calls for developing effective strategy. Journal of Technology Research, 6, 1.

[6] Li, S. T., Kuo, S. C., & Tsai, F. C. (2010). An intelligent decision-support model using FSOM and rule extraction for crime prevention. Expert Systems with Applications, 37(10), 7108-7119.

[7] [Gerber, M. S. (2014). Predicting crime using Twitter and kernel density estimation. Decision Support Systems, 61, 115-125.

[8] Tuominen, T., Korhonen, T., Hämäläinen, H., Katajisto, J., Vartiainen, H., Joukamaa, M., ... & Lauerma, H. (2017). The factors associated with criminal recidivism in Finnish male offenders: importance of neurocognitive deficits and substance dependence. Journal of Scandinavian Studies in Criminology and Crime Prevention, 18(1), 52-67.

[9] Akers, R. (2017). Social learning and social structure: A general theory of crime and deviance. Routledge.

[10] Prabakaran, S., & Mitra, S. (2018, April). Survey of Analysis of Crime Detection Techniques Using Data Mining and Machine Learning. In Journal of Physics: Conference Series (Vol. 1000, No. 1, p. 012046). IOP Publishing.

[11] Chastain, B., Qiu, F., & Piquero, A. R. (2016). Crime theory evaluation using simulation models of residential burglary. American Journal of Criminal Justice, 41(4), 814-833.

[12] Li, S. T., Kuo, S. C., & Tsai, F. C. (2010). An intelligent decision-support model using FSOM and rule extraction for crime prevention. Expert Systems with Applications, 37(10), 7108-7119.

[13] Annabathula, R. (2007). A Web-based tool for analysis of crime laboratory data. West Virginia University.

AUTHORS PROFILE

Mr. Mousa Mojarad received his PhD in Computer-Software Engineering in 2020. He is currently a lecturer and faculty member of the Islamic Azad University, Firoozabad Branch. His hobbies are big data, cognitive computing, clustering, software engineering, classification models, and cloud computing. He has more than 8 years of teaching experience and 6 years of research experience.

Ms. Afsaneh Haghanin Nia received the B.Sc. Degree in computer engineering from University of Applied Sciences, Bushehr Province, in 2016, and the M.Sc. degree in computer software engineering from Liyan Institute of Education, Bushehr, in 2021. His primary research interest is in using meta-heuristic algorithms for routing, although he has concurrent research in robotics, machine learning, optimization algorithms, and artificial intelligence.

Ms. Neda Galehdari received the B.Sc. Degree in computer engineering from Imam Javad University, in 2012, and the M.Sc. degree in computer software engineering from Liyan Institute of Education, Bushehr, in 2021. His current research interests include Evolutionary Computation, Optimization Methods, Optimization Algorithms, and Swarm Intelligence.

Ms. Khayronsa Bazarganzadeh received the B.Sc and M.Sc. degree in Computer Engineering from Borazjan Branch, Islamic Azad University, in 2016 and Liyan Institute of Education, Bushehr, in 2021, respectively. Khayronsa’s current research interests include the affective computing, virtual reality, human-computer interaction, and computer networks.