A Data-Deriven Model for Forensic Policy Making in Electronic Banking Using Agent-Based Simulation

Afshin Khodamoradi [1], Alireza Pourebrahimi[2]*, Mohammad Ali Afshar Kazemi[3]

1.Department of Industrial Management, Qeshm Branch, Islamic Azad University, Qeshm, Iran.
2.Department of Industrial Management, Karaj Branch, Islamic Azad University, Karaj, Iran (Corresponding author).
3.Department of Industrial Management, Central Tehran Branch, Islamic Azad University, Tehran, Iran

*** Corresponding author email address**: Poorebrahimi@gmail.com

Received: 2024-10-20 **Reviewed:** 2024-11-10 **Revised:** 2024-12-01 **Accepted:** 2024-12-23 **Published:** 2025-03-30

Abstract

The rapid growth and expansion of information and communication technology has brought about a revolution in various aspects of human life and organizational performance. Every new technology, in order to become widespread and develop, requires legal acceptance before its full potential can be utilized. In a world where we are increasingly dependent on the virtual space and store our sensitive information on cloud and online platforms, the existence of criminals aiming to obtain this information is an undeniable and self-evident reality. With the expansion of the internet, an old topic that existed even before the advent of the internet entered the digital world. This topic is referred to as forensic science. Therefore, the main objective of this research is to propose a grounded data model for criminology policy-making in electronic banking using agent-based simulation. After presenting the model, simulators, especially Matlab, will be used according to the project's requirements, and the results will be evaluated based on execution speed. The study will continue by creating several simulations to analyze the system, with detailed results and related graphs provided. After extracting features and applying algorithms, the proposed method will be evaluated through accuracy measures.

Keywords: *Electronic Banking, Criminology, Forensics, Virtual Space.*

How to cite this article:

Khodamoradi A , Pourebrahimi A, Afshar Kazemi M. (2025). A Data-Deriven Model for Forensic Policy Making in Electronic Banking Using Agent-Based Simulation. Management Strategies and Engineering Sciences, 7(1), 127-148.

1. Introduction

Today, significant transformations have occurred in the banking industry, making its nature increasingly challenging and complex. In the modern world, banking issues and problems not only spread within national borders to other banks and similar organizations but also transcend national borders, affecting financial institutions in other countries [1].

Banks and credit institutions play a central role in the growth and prosperity of national economies. The banking industry in our country has undergone significant transformations in recent years. Some of the most important developments in banking include extensive interactions between domestic banks and international banks and organizations, efforts to combat money laundering, the increasing number of banks and credit institutions, intense competition among them, the use of new banking tools, and the expansion of electronic banking activities [2, 3]. Naturally, as the banking industry evolves, its supervision also undergoes profound changes, with increased importance attached to it. Given the significant impact that banks and credit institutions have on economic growth and their influence on macroeconomic variables such as liquidity growth, inflation, and unemployment, the need for transformation in banking supervision has become more pressing [4, 5]. This transformation aims to align banks and credit institutions with monetary and banking policies, prevent violations of laws and regulatory provisions, ensure the stability and soundness of the banking system, and protect the interests of depositors [6].

In today's banking world, with the complexity of banking tools, the diversity of banking activities, and internal communications, maintaining the health and stability of the banking system is one of the key reasons for monitoring banks and credit institutions. This is because, in the event of a problem or disruption in a financial institution, the entire financial system will be impacted due to extensive interconnections [7].

On the other hand, cybercriminals can inflict severe damage on the computers of companies or even individuals without leaving the slightest trace (no fingerprints, no bullet holes, not even cigarette butts). The internet has the potential to help or harm leaders of groups, senior executives of financial institutions, lawmakers, high-level security officers, risk managers, and insurance brokers or intermediaries. Internet or other data transmission protocols have facilitated the flow of information processing. In the new information economy, information is the primary, fluid capital (like blood in arteries) of every company, and companies carefully protect it. Even well-managed and controlled companies are interested in discussing and addressing information security and issues such as information technology, which requires informational solutions [8].

Ensuring the security of information and organizational information structures is a matter that is discussed at the corporate level, with responsibility lying with top management (board members) and other important individuals within the organization. Any security issues in the organization's information network can have widespread and detrimental effects throughout the organization. In general, service disruption or failure can significantly impact the operational level of the organization, to the extent that if information and news are compromised or disrupted, the organization's brand and reputation will be harmed [9].

The banking network, like many other organizations and companies, is not immune to internet-related problems and crimes. Today, or in the near future, most financial transactions in the economic market will be carried out through computers. As a result, the development of computer data in banks could lead to a fear of imminent execution of criminal activities by cybercriminals [10].

With the advent of computers, a new type of crime has emerged that confronts criminologists and security forces. Cybercrime has diversified over time, and now includes known forms such as sabotage, espionage, theft, and illegal use and manipulation of computers, among others. In the 1970s, the first scientific and criminological studies were conducted in this area, leading to the identification of various forms of cybercrime and the significant "dark figure" associated with such crimes [1].

For example, the list of criminal activities in the realm of informatics and computers, introduced by the Council of Europe, includes: computer fraud, information technology fraud, causing damage to data through informatics programs, sabotage and illegal manipulation of computers, unauthorized access to informatics systems and data, unauthorized interception of computer communications, unauthorized production of protected computer and informatics programs, illegal duplication of a topology, computer espionage, and unauthorized use of a computer. These are internet-related crimes that have caused significant damage to their victims [11].

In the age of communication, despite the widespread use of advanced communication tools and the utilization of communication achievements of this century, we are still confronted with some unresolved issues. Terms such as "cybercrime" are familiar to many of us, and from time to time, reports or news about the commission of such crimes by criminals are published. Therefore, given the aforementioned points and the fact that banks have suffered significant losses due to internet-related crimes, identifying these crimes and focusing on methods for preventing them seems essential. This highlights the importance of the present research [12, 13].

In summary, the rapid growth and expansion of information and communication technology have caused a revolution in various aspects of human life and organizational performance. Every new technology requires legal acceptance before it can become widespread and fully developed. In a world that is increasingly dependent on virtual space, where we store our sensitive information in the cloud, the existence of criminals aiming to obtain such information is an undeniable and self-evident reality. With the expansion of the internet, an old topic that existed even before the internet came into existence has entered the digital world: forensic science. Therefore, the primary aim of this research is to propose a grounded data model for criminology policy-making in electronic banking using agent-based simulation.

2. Methodology

This research is a descriptive-quantitative study that employs both in-depth thinking and survey methods. Various tools of this technique, including interviews, observation, questionnaires, and document review, were used for data collection.

Regarding the collection of information related to the components of the research and its background, library-based methods and the use of articles from reputable databases (such as Science Direct, Springer, IEEE, etc.) related to the subject matter are employed.

To collect the data for this study, log databases from banks were used.

The statistical population of this study consists of cybersecurity event logs from banks over the past year, and no specific sampling method was applied.

After presenting the model, based on the project's requirements, common simulators, especially Matlab, will be used, and the results will be evaluated based on execution speed. In the following, the research will conduct several simulations in the system under investigation, and the results will be explained in detail, along with relevant charts.

After extracting features and applying algorithms, the proposed method will be evaluated using accuracy metrics.

2.1. Combination of Fuzzy Sets and Rough Sets

The theory of rough sets introduces the concept of indiscernibility to reduce dimensionality. Two objects in a dataset are said to be equivalent if their feature values are the same, meaning they belong to the same equivalence class. In the theory of deterministic rough sets, objects belong to an equivalence class with degree one, and their membership value in other equivalence classes is zero. In practice, it is preferable for an object to belong to each class with a membership value in the range [0, 1]. This concept leads to the idea of fuzzy equivalence classes, which is the core idea in hybrid approaches combining fuzzy set theory and rough sets [14].

A fuzzy rough set is an extension of rough sets that is derived by approximating a fuzzy set in the space of deterministic approximation. Fuzzy rough sets can further be generalized into fuzzy rough sets, where all equivalence classes are fuzzy. In the context of feature reduction for rough sets, this situation corresponds to when decision and condition values are fuzzy [15].

Grounded theory is a qualitative research method where a theory is developed from data. In fact, whenever a researcher wishes to explore the experiences and perspectives of individuals to formulate and generate a theory, grounded theory is an appropriate approach.

The goal of grounded theory is to create a theory that is based on and adheres to evidence. This method is used to discover new theories. In this approach, the researcher compares dissimilar phenomena to understand their similarities. The researcher sees micro-level events as the foundation for macro-level explanations [16].

It should also be noted that the questions answered by the grounded theory approach focus on "how" and "why." The primary focus of this research is to explain how human resource measurement is performed and why companies enter the human resource measurement process. Therefore, it is expected that the grounded theory approach will assist the researcher in answering the main research questions in an optimal manner [17].

The data analysis process consists of three main stages: open coding, axial coding, and selective coding. In open coding, open sampling is performed, meaning that participants are selected who offer the greatest opportunity to gather the most relevant data regarding the phenomenon under investigation. In open sampling, as a researcher, we are still unsure which concepts are theoretically appropriate, so we take an open approach, disregarding individual differences, and increase the number of interviews.

In axial coding, preliminary hypotheses about each category and the relationships between them are made, and new questions are posed to validate these hypotheses. New comparisons are also made. Participants are chosen who provide the best opportunity to gather relevant data. In selective coding, the goal is to maximize the opportunity to confirm the narrative flow and relationships between categories and fill gaps in weak and incomplete categories. This may involve returning to previous participants or selecting new participants who have relevant information [18].

2.2. Agent-Based Simulation

2.2.1. Distributed Data Mining

Distributed data mining refers to techniques for discovering important patterns in separate databases, examining patterns from a unified perspective, and uncovering specific relationships between different datasets. Common distributed data mining algorithms perform local data analysis, which is then integrated using knowledge integration methods to extract general knowledge from them.

2.2.2. Agent-Based Data Mining

Agent-based data mining refers to versions of multi-agent systems developed to improve the data mining process. The development of multi-agent systems can help solve data mining problems. For example, agent-based data mining architecture, agent-based interactive exploration, agent-based user interaction, automatic pattern discovery, distributed agent-based data mining, dynamic agent-based mining, multi-source agent-based mining, peer-to-peer agent-based mining, and web-based agent mining. The roles of agents in supporting distributed data mining will be described in detail below.

Agent technology can contribute to overcoming these challenges with its capabilities for autonomy, interaction, dynamic selection, scalability, multiple strategies, and collaboration. Other reasons include encapsulation, mobility, time constraints (such as data flow, the time-consuming extraction process, and exploration), and computational costs and efficiency requirements. In fact, multi-agent technology complements distributed data mining in many ways.

2.3. Association Rule Discovery Algorithms

2.3.1. AIS Algorithm

This algorithm was one of the first developed for extracting all frequent items from a database in 1993 by Agrawal, Imielinski, and Swami. The name of this algorithm is derived from the initials of the developers. The algorithm performs several passes over the database and in each pass, scans all transactions.

In the first step, the support for each item is calculated, and those with support greater than the minimum threshold are recorded in L1. In the second step, for each item in L1, the algorithm returns to the main database and creates all pairs of items, then calculates their support. The output of this step is stored in 2C. In the third step, a similar process is performed to calculate 3C. These steps continue until no new frequent sets are added.

2.3.2. SETM Algorithm

This algorithm was proposed by Houtsma in 1995, and in 1996, a second version was introduced by Srikant for calculating frequent items in SQL. In this algorithm, each item in the set is in the form <TID, Itemset>.

The support for each item is calculated separately, and the largest supports are selected. Candidate items (CK) are created by extending each of these frequent items with other items in each transaction. Furthermore, during this stage, TIDs related to each of the CKs are stored in an ordered structure called 2C, and the support for each CK is calculated by summing its repetitions from the previous step, producing 3C. These steps continue until no new frequent sets are found. The main disadvantage of this algorithm arises from the large number of CKs, and since the TID for each CK is stored, more memory space is required.

2.3.3. Apriori Algorithm

This algorithm leverages the fact that all subsets of frequent items are also frequent and that items should be sorted based on alphabetical order. The key difference between this algorithm and other algorithms is in the calculation of Ck items and their selection for subsequent steps. In other algorithms, frequent items were created by extending each of the frequent items to other individual items (which may not themselves be frequent) in each transaction, thus generating many CKs, which would then be pruned in later steps, and the database would be scanned multiple times. In contrast, this algorithm only scans the database once to find frequent items.

The Apriori algorithm takes this important factor into account by generating Ck items by joining frequent items from the previous phase and eliminating those that were already part of the previous phase, regardless of individual transactions. This significantly reduces the number of unnecessary Ck items.

2.3.4. AprioriTid Algorithm

The Apriori algorithm scans the entire database in each pass to calculate supports, but scanning the entire database may not always be necessary in every phase. Based on this issue, another algorithm called AprioriTid was developed. This algorithm uses a method similar to the Apriori algorithm to compute Ck items in each phase. The key difference between this algorithm and the Apriori algorithm is that it does not scan the entire database after the first phase and instead uses Ck sets to compute supports. Similar to the SETM algorithm, the items in this algorithm are also stored in the form <TID, Xk>.

3. Findings and Results

3.1. Characteristics Affecting the Detection of Cybercrime Events

In cybercrime attacks, perpetrators attempt to design fake websites in such a way that users do not notice the difference between them and the original site, making it easy for users to unknowingly disclose their confidential information. Despite the efforts of cybercrime perpetrators, there are indicators and features in fake websites that help us detect their authenticity. Clearly, in order to design a system capable of identifying any cybercrime-related events and alerting users, the first step is to identify the features of cybercrime-related websites. Therefore, the initial step involved reviewing articles related to cybercrime detection and examining real-world examples of cybercrime websites available on the FishTank platform. Based on this, an initial list of all the characteristics of cybercrime attacks was extracted.

Table 1. Characteristics of Cybercrime Attacks

No.	Cybercrime Indicators	No.	Cybercrime Indicators
1	Use of IP address in URL	21	Use of similar characters in URL
2	Unusual request URL	22	Adding prefixes and suffixes
3	Unusual anchor URL	23	Use of @ symbol in URL
4	Unusual DNS record	24	Use of hexadecimal codes
5	Unusual URL	25	Use of switching gateways
6	Use of SSL certificate (lock indicator or https:// in the address bar)	26	Excessive emphasis on security in the website
7	Certificate Authority (CA)	27	Public email addresses
8	Unusual cookies	28	Manipulation of time for account access
9	Details in the digital certificate	29	Request for confidential information in emails
10	Web page redirection	30	HTML files as email attachments that open locally on the client
11	Cross-site scripting (XSS)	31	SSL error messages
12	Pharming attack	32	"Invalid information" error message after login
13	Hidden page link upon mouse hover	33	Presence of valid domain names in the URL
14	Unusual SFH	34	Websites providing financial services
15	Spelling and grammatical errors in emails or websites	35	<iframe> tag in website code
16	Copying website	36	<script> tag in website code
17	Presence of forms with submit buttons	37	Use of email code on the website
18	Pop-up windows	38	Presence of https in website code
19	Right-click disabled	39	Illegal pop-up code
20	Long URL		

3.2. Identifying Input Variables

At this stage of the research, it is necessary to identify the features that are deemed most significant and influential according to experts. As mentioned earlier, the initial list included 39 cybercrime indicators. To remove redundant indicators, the list was narrowed to 28 important characteristics, which covered the other indicators, and the list was formatted into a questionnaire that was distributed among the experts. This process aimed to identify the most important and effective indicators according to their perspectives.

The experts involved in this research were all network and information security specialists working in government and private companies, including Hamkaran System, Fanap, Mamut IT, Bank Tejarat, Bank Mellat, Post Bank of Iran, Saba System Sadra, Telecommunication Research Center, and also faculty members from Sharif University of Technology and Amirkabir University of Technology. Questionnaires were also sent to researchers in the United States and Australia whose previous research in the field of cybercrime had been published in reputable international journals.

According to the general principle, when accuracy considerations are met, the larger the sample size, the lower the error of generalization. However, due to financial, human, and time constraints in most research, controlling some of the variables that affect the results is not always possible. In such situations, the results are reliable when the sample size is large enough.

In this questionnaire, experts were asked to assign a number between 1 and 10 to indicate the importance and impact of each of the indicators on detecting cybercrime events.

In the first phase (validity assessment), 10 experts provided feedback on the accuracy of the items and definitions. Their suggestions were incorporated into the questionnaire, and Cronbach's alpha was calculated to assess reliability. Since Cronbach's alpha was calculated as 0.849, the questionnaire was found to be valid and reliable.

The questionnaire was sent to 100 individuals, of whom 43 responded. It should be noted that each completed questionnaire was carefully reviewed, and if any discrepancies were found in the responses, the respective questionnaire was discarded and not used. For the analysis of the questionnaires, the average of each indicator was calculated (Table 2). These averages ranged between 5 and 8, which correspond to the "moderate to high" range. Given the relative importance of all the indicators in the initial list, this result was somewhat predictable. Since calculating the mean did not highlight the relative superiority of certain indicators, a different method was needed for ranking the indicators based on their importance and influence.

Generally, there are three methods for extracting and determining factors in factor analysis: principal component method, maximum likelihood method, and principal axis factoring.

Since simultaneous analysis of 28 variables is practically difficult and even impossible, factor analysis was employed in this research, with the principal component method used to determine the factors.

In the principal component method, the goal is to estimate factor loadings in such a way that the sum of the common variance is maximized. This allows the factors to explain a higher percentage of the variance of the observable variables.

Table 2. Mean of Indicators and Calculation of Loadings for Two Factors

No.	Cybercrime Indicators	Experts' Mean Opinion	Factor 1	Factor 2
1	Use of IP address	5.46	-0.08401774	0.199054046
2	Unusual request URL	7.12	-0.22826583	-0.104964362
3	Unusual anchor URL	6.12	-0.14394513	-0.244070268
4	Unusual DNS record	6.23	-0.14437648	-0.106664333
5	Unusual URL	7.39	-0.20734723	0.068678146
6	Use of SSL certificate	6.77	-0.07838130	-0.084525211
7	Certificate Authority	7.07	-0.13576264	-0.371317275
8	Unusual cookies	5.30	-0.09332888	-0.238611891
9	Details in the certificate	6.12	-0.14091565	-0.279801187
10	Website redirection	6.63	-0.18054822	-0.138555086
11	Code injection attack	6.86	-0.21122304	-0.187975143
12	Pharming attack	6.16	-0.26348457	-0.025571380
13	Hidden link on mouse hover	5.46	-0.22675047	-0.216272437
14	Unusual SFH	6.07	-0.13630730	-0.334370265
15	Syntax and spelling errors	6.81	-0.18546897	-0.003587774
16	Website copying	6.88	-0.02585877	-0.014771853
17	Presence of information forms	6.91	-0.20168410	-0.076945395
18	Use of pop-up windows	6.46	-0.22799395	-0.069789606
19	Right-click disabled	5.25	-0.20285741	-0.215584942
20	Long URL	5.14	-0.18344816	-0.029323293
21	Use of similar characters	6.35	-0.21359875	-0.230520356

22	Adding prefixes and suffixes	7.18	-0.27041966	-0.061858127
23	Use of "@" symbol	5.88	-0.18614920	-0.099840976
24	Hexadecimal codes	6.23	-0.15518396	-0.250053487
25	Use of switching gateway	5.69	-0.17025524	-0.320978427
26	Overemphasis on security	5.05	-0.13984636	-0.136110076
27	Public email addresses	6.44	-0.20882279	-0.114166808
28	Manipulating time	5.93	-0.22962968	-0.239332106

The R software divides the variables into six categories based on their impact directions. As explained earlier, these categories are introduced in Table 3. The variables correspond to the cybercrime indicators listed in Table 2 in order:

Table 3. Categorization of Variables Based on Impact Direction

No.	Category	Variables	No.	Category	Variables
1	Category 1	Variables (3, 7, 8, 9, 14)	4	Category 4	Variables (5, 12, 15, 16, 18, 20, 22, 23, 27)
2	Category 2	Variables (4, 6, 10, 11, 13)	5	Category 5	Variables (26, 19, 21, 28)
3	Category 3	Variables (2, 17)	6	Category 6	Variables (1, 24, 25)

The greater the vector corresponding to each variable, the greater the influence of that variable in the model presented. Given that the scale of the factors in the factor loading diagram is not uniform, this diagram cannot be directly used to determine the impact of variables. To address this, the loadings for each variable were extracted and arranged in descending order, as shown in Table 4.

Table 4. Ranking of Cybercrime Indicators in Descending Order

No.	Cybercrime Indicators	Loading	No.	Cybercrime Indicators	Loading
1	Trustworthiness of the certificate authority	0.07815401	15	Website copying	0.03157417
2	Use of switching gateway	0.06600700	16	Unusual request URL	0.03156140
3	Unusual SFH	0.06519158	17	Use of pop-up windows	0.02842592
4	Manipulating time	0.05500482	18	Public email addresses	0.02832051
5	Use of similar characters in URL	0.04938203	19	Website redirection	0.02589759
6	Hidden link on mouse hover	0.04909477	20	Unusual URL	0.02385478
7	Details in the certificate	0.04907296	21	Use of IP address in URL	0.02334075
8	Right-click disabled	0.04381400	22	Presence of information forms	0.02309735
9	Hexadecimal codes	0.04330440	23	Use of "@" symbol in URL	0.02230987
10	Unusual anchor URL	0.04014525	24	Overemphasis on security	0.01891602
11	Code injection attack	0.03997491	25	Long URL	0.01725654
12	Adding prefixes and suffixes	0.03847661	26	Syntax and spelling errors	0.01720581
13	Pharming attack	0.03503901	27	Unusual DNS record	0.01611092
14	Unusual cookies	0.03282296	28	Use of SSL certificate	0.00664407

As observed in Table 4, the variable "Trustworthiness of the certificate authority" has the highest impact, while "Use of SSL certificate" has the least impact on detecting cybercrime events.

To determine the most influential category, the total loadings associated with each category were used, and these loadings are arranged in descending order in Table 5.

Table 5. Loadings Associated with Categories

No.	Category	Loading	No.	Category	Loading
1	Category 1	0.2653868	4	Category 2	0.1377223
2	Category 4	0.2424632	5	Category 6	0.1326521
3	Category 5	0.1671169	6	Category 3	0.05465875

Thus, Category 1 holds the greatest importance, followed by Categories 2 (Blue), 4 (Turquoise Blue), 5 (Red), 6

(Pink), and 3 (Green) in descending order. As a result, 8 variables out of the initial 28 were removed at this stage. The remaining 20 significant variables are listed in Table 6.

Table 6. Effective Indicators for Detecting Cybercrime Events in Iranian Banks

No.	Cybercrime Indicators	No.	Cybercrime Indicators
1	CA trustworthiness	11	Unusual request URL
2	Unusual anchor URL	12	Unusual URL
3	Unusual cookies	13	Use of pop-up windows
4	Details in the certificate	14	Long URL
5	Unusual SFH	15	Adding prefixes and suffixes
6	Unusual DNS record	16	Use of "@" symbol for obfuscation
7	Use of SSL certificate	17	Use of similar characters in URL
8	Website redirection	18	Use of IP address
9	Code injection attack	19	Use of hexadecimal codes
10	Hidden link with mouse hover	20	Use of switching port

3.3. Determining Output Variables

The goal of designing the expert system is to identify cybercrime attacks with the highest agility and accuracy. The output variable of the fuzzy inference engine is the "website's cybercrime risk level," which is assigned the linguistic terms: "legal," "slightly suspicious," "suspicious," "highly suspicious," and "fraudulent." In other words, the system classifies a website using one of the following terms:

- **Legal**: The website is sufficiently secure, and its credibility can be trusted.
- **Slightly Suspicious**: The website is not entirely trustworthy; before entering any information, one should ensure its validity and legality.
- **Suspicious**: The website has characteristics that violate its legality.
- **Highly Suspicious**: The website is highly likely to be fraudulent. No information should be entered.
- **Fraudulent**: The website is entirely fraudulent.

3.4. Evaluation of the Fuzzy Expert System for Cybercrime Detection

At this stage, the fuzzy expert system was tested on several real samples of websites from Iranian banks and also on samples of cybercrime attacks registered on the PhishTank website. It is worth mentioning that the criterion for comparing and validating the results of the expert system is the experts' opinions. This is because an expert system aims to simulate the closest possible model to that of a human expert.

3.5. Improving the Fuzzy Expert System Using the Rough Set Theory

3.5.1. Reducing Input Variables Using Rough Set Theory

At this stage, the goal is to identify and eliminate ineffective and redundant indicators from Table 2 using the fuzzy-rough feature selection algorithm for real-world cybercrime events. This will reduce the number of input variables to the fuzzy system and, consequently, minimize the number of rules, thus reducing the time required to determine the website's validity, making the system more agile. To achieve this, the fuzzy-rough set algorithm was used to identify which input variables have the most significant influence on the output of the fuzzy cybercrime detection system, and the system's rules were defined based on these indicators.

For implementing the fuzzy-rough algorithm, 60 real-world websites from the electronic banking domain were extracted, of which more than 50% were related to Iranian banks, while the rest were associated with cybercrime attacks on other banking websites worldwide. These websites were then processed using a special version of the WEKA data mining software. It should be noted that, in this stage, the values for all 28 initial indicators listed in Table 2 were considered for each website and fed into the software. The six most influential indicators derived from the fuzzy-rough algorithm are as follows: URL length, certificate authority trustworthiness, certificate details, unusual anchor URL, unusual SFH, and adding prefixes and suffixes (subdomain existence).

3.5.2. Evaluation of the Fuzzy-Rough Expert System

At this stage, the fuzzy-rough expert system was tested on 50 electronic banking websites, and the results were compared with those obtained from the fuzzy expert system. The comparison shows that the fuzzy-rough expert system is capable of detecting the website's credibility with the same accuracy as the fuzzy expert system.

The results showed that the cybercrime risk for this test website was calculated to be 43.7%, and the website was classified as "suspicious." The identical output with the previous fuzzy system output clearly demonstrates that the indicators removed from the fuzzy system's inputs did not significantly impact the final result, and they only increased the number of calculations. The role of the indicators in the fuzzy-rough rule base remained unchanged compared to the fuzzy system rules. Therefore, when the input values are identical for the common indicators, the final detection result from the fuzzy-rough system will be the same as that from the fuzzy system. The advantage of using the fuzzy-rough system is its speed in detecting cybercrime, which is discussed further below.

Table 7. Results of Implementing the Fuzzy-Rough Expert System

No.	URL Length	Certificate Trustworthiness	Certificate Details	Unusual Anchor URL	Unusual SFH	Adding Prefixes and Suffixes	Output (Percentage)	Detection	Result Accuracy
1	64	0	0	0	0	0	4.91	Fraudulent	Correct
2	28	0	0	0	0	1	4.91	Fraudulent	Correct
3	60	3.8	7	0	0	1	3.19	Slightly Suspicious	Incorrect (Legal)
4	45	9	6.9	0	0	1	78.4	Legal	Correct
5	25	9.8	5.9	0	0	1	76.5	Legal	Correct
6	37	9.8	5.9	1	0	1	78.4	Legal	Correct
7	55	0	0	0	0	1	4.91	Fraudulent	Correct
8	38	9.8	5.9	0	0	1	78.4	Legal	Correct
9	30	9.1	6.7	1.2	0	0	4.91	Fraudulent	Incorrect (Legal)
10	20	8	5.3	2.3	1	1	7.43	Suspicious	Correct

The results of the fuzzy-rough expert system applied to 50 banking websites show that the system's detection accuracy is 88%, with an error rate of 12%. It was noted that the main advantage of using rough set theory is the reduction in computation time. As mentioned, the calculation time is directly related to the number of variables and fuzzy rules. The fuzzy-rough system, which has six input variables and 40 rules, was implemented and tested on a computer with 4 GB of RAM and an Intel Core i7 processor at 3.2 GHz. The system was able to compute and report the result within one second. In contrast, the fuzzy expert system, which has 20 input variables and 159 rules, took at least 16 seconds to compute the cybercrime risk and report the website's status. Therefore, it is evident that the fuzzy-rough expert system is a much more suitable option for real-time use, as only a few seconds are needed to uncover sensitive and confidential user information in cyberspace, and timely identification of fraudulent websites is critical.

3.6. Agent-Based Simulation

In this section, we aim to present a proposed architecture for data mining based on multi-agent systems and briefly describe distributed data mining.

There are numerous research topics in the field of agent-based data mining. In establishing an organizational data mining infrastructure based on agents, one may focus on techniques for designing and analyzing community-oriented and organizational systems for large-scale agent-based systems. Similarly, solutions for integrating agent-based service applications, distributed data preparation, distributed agent collaboration, and parallel agent computation must be considered. In many cases of data mining, researchers must investigate algorithms capable of adapting to dynamic changes in data and user requests. Adaptive and automatic data mining algorithms should be explored.

Figure 1 illustrates the framework that includes the research components in this area, which are: the foundation of agent exploration, agent-based data processing, agent-based knowledge discovery, agent-based multi-agent systems, agent-based information processing, common topics in agent exploration, agent exploration systems, agent exploration applications, agent-based knowledge management, and agent exploration performance evaluation.

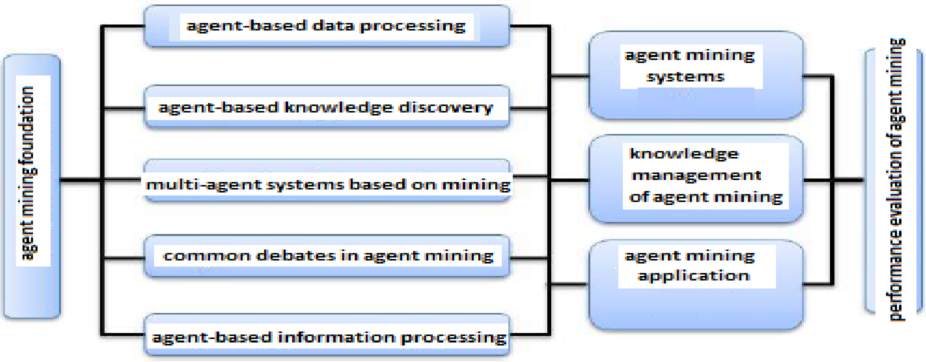

agent mining foundation
agent-based data processing
agent-based knowledge discovery
multi-agent systems based on mining
common debates in agent mining
agent-based information processing
agent mining systems
knowledge management of agent mining
agent mining application
performance evaluation of agent mining

Figure 1. Agent-Based Exploration Research Framework

This research falls under the domain of agent-based knowledge discovery. The overall operation of the proposed architecture is as follows: after the user logs in and successfully registers, a user agent is created to represent the user. This agent is deployed within the system, establishes communication with the user, queries the user for the necessary knowledge, and receives its input parameters. Then, a group of active agents is triggered, and the data mining operation is initiated. After the mining results are provided to the user, they are asked whether they wish to search for additional knowledge. If a new search request exists, the system waits for the user to enter the necessary parameters; otherwise, the user agent is removed. Figure 2 presents a workflow diagram based on a distributed multi-agent data mining system.

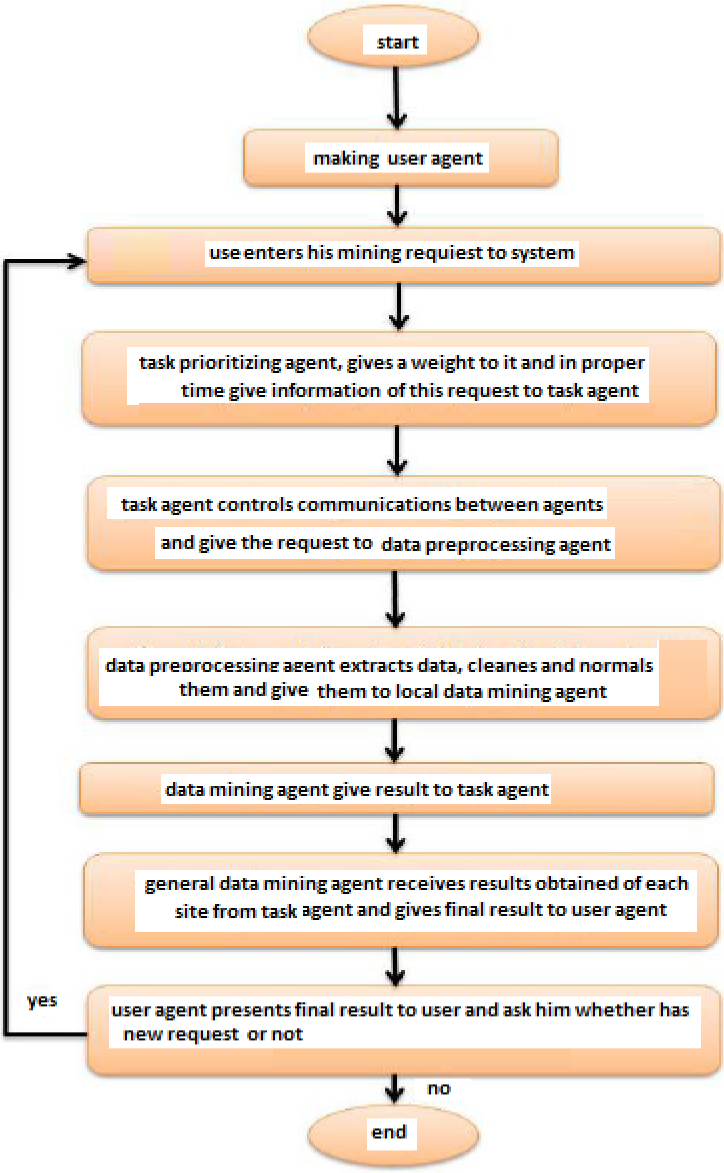

start
making user agent
use enters his mining requiest to system
task prioritizing agent, gives a weight to it and in proper time give information of this request to task agent
task agent controls communications between agents and give the request to data preprocessing agent
data preprocessing agent extracts data, cleanes and normals them and give them to local data mining agent
data mining agent give result to task agent
general data mining agent receives results obtained of each site from task agent and gives final result to user agent
user agent presents final result to user and ask him whether has new request or not
yes
no
end

Figure 2. Workflow of the Distributed Data Mining System Based on a Multi-Agent System

3.7. *Proposed Architecture for Distributed Data Mining Based on Multi-Agent Systems*

Given the tasks that a distributed data mining system must perform, the system's functionality is divided into two parts: the user interaction functions and the data preprocessing and mining functions. Each function is associated with an agent, so the system mainly consists of two types of agents: the user agent, the data preprocessing agent, and the data mining agent. To explore data from various websites, the data preprocessing agent and the data mining agent are considered as a pair and distributed across the sites that will be mined. The data preprocessing agent prepares the data for the data mining agent. The data mining agent completes the mining task and submits the results to the central data mining site for integration and obtaining the correct results.

3.7.1. *Proposed Four-Layer Architecture*

As will be detailed below, the proposed architecture is four-layered. The description of each layer is as follows:

- **First Layer: User Layer**

As shown in Figure 3, the first layer of this architecture consists of the users, the user agent, and the user information database.

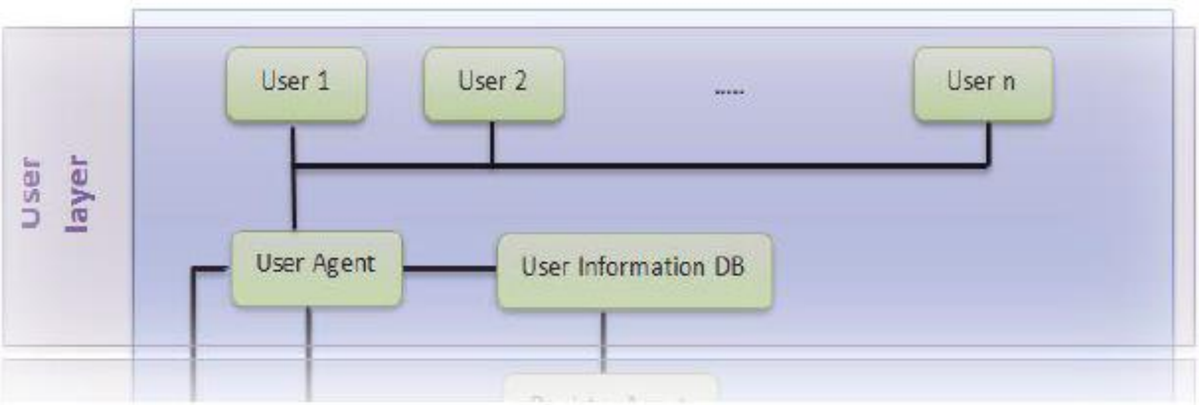


Figure 3. User Layer

User Agent: The primary function of this agent is to complete the interactions between the system and the user. For this purpose, the user agent provides an interactive interface through which the user submits requests, and the data mining results are shown. The intelligence of the user agent manifests in the long term, as during the data mining process, the exploration requests of the user are stored in the task information database, which can represent the user’s interests.

User Information Database: This database stores the records and information of the user.

- **Second Layer: Management Layer**

As seen in Figure 4, the components of this layer include the registration agent, the task prioritization agent, the task agent, the general data mining agent, and the task information database.

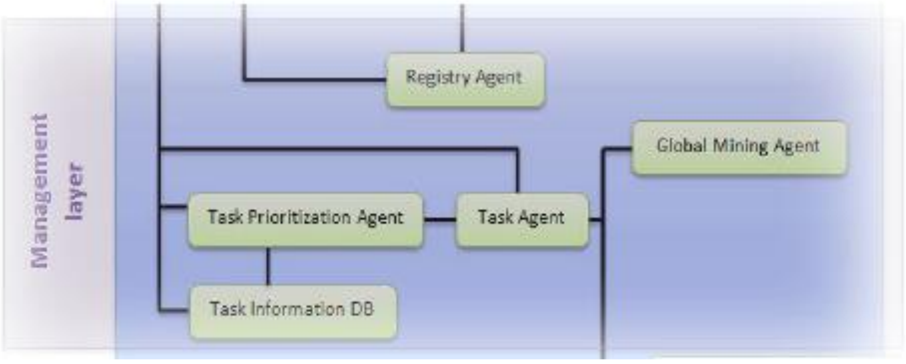


Figure 4. Management Layer

Task Prioritization Agent: These agents register the number of requests in a given time interval and their associated information. Then, they assign a weight to each request based on the recorded information, which is the main function of this agent. When the system and network are idle, this agent can find requests with weights higher than a critical threshold. Based on these weights, the required exploration information for the requests is sent to the task agent.

Task Agent: These agents are temporary and are automatically created by the task prioritization agent to handle data mining requests. They exist until the corresponding request is fully processed.

Task Information Database: This database is used to store the request information received from the user via the task prioritization agent.

Registration Agent: This agent is composed of functions responsible for controlling various access types, changes, logs, etc., for different users, including admins, regular users, developers, and others.

- **Third Layer: Processing Layer**

As shown in Figure 5, this layer consists of the data mining agent and the data preprocessing agent.

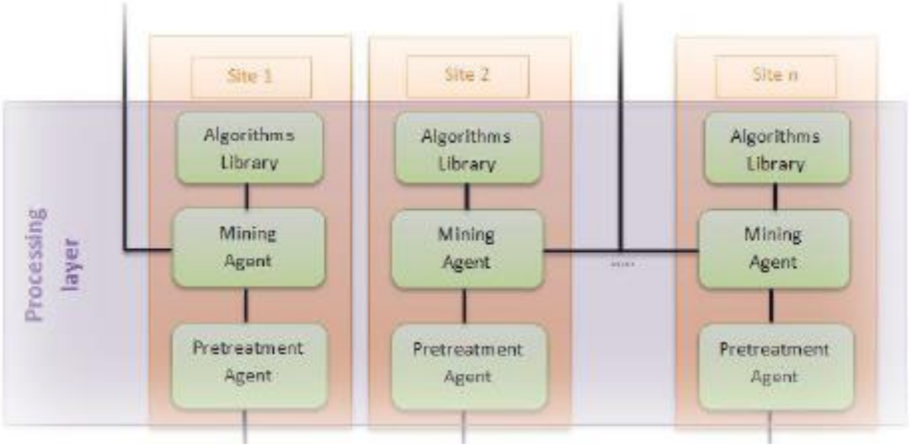


Figure 5. Processing Layer

Data Preprocessing Agent: This agent is located at the local site and handles the data sources of that site. Its main task is to perform preprocessing operations and provide normalized data. For this purpose, it extracts the relevant data from the data sources, performs preprocessing tasks such as normalization, and finally sends the data to the data mining agent.

Data Mining Agent: Each of the data mining agents implements a specific data mining algorithm or technique. Data mining agents contain methods for initialization, performing exploration, and presenting results to the corresponding task agent. Depending on the function of the data mining agent, it can be divided into two components: the general data mining agent and the local data mining agent. The general data mining agent is located at the central site and works on the data generated from the results of the local data mining agents across various sites. The local data mining agent performs the exploration task on the local site and uses the data provided by the data preprocessing agent at that site.

Algorithm Library: In this library, the exploration algorithms and the results obtained from using each algorithm are stored. These algorithms can be deleted, updated, and developed.

- **Fourth Layer: Resource Layer**

This layer consists of the data resources (Figure 6).

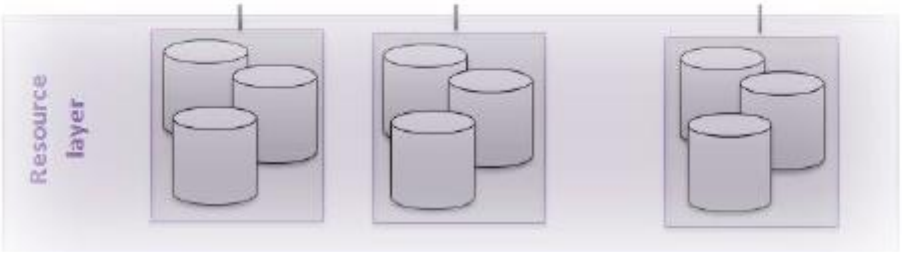


Figure 6. Resource Layer

As mentioned earlier, each layer of this architecture consists of components. To implement system communication between different agents, a communication server is used, and each agent transfers information to the next step and agent through this server. An overview of this architecture is shown in Figure 7.

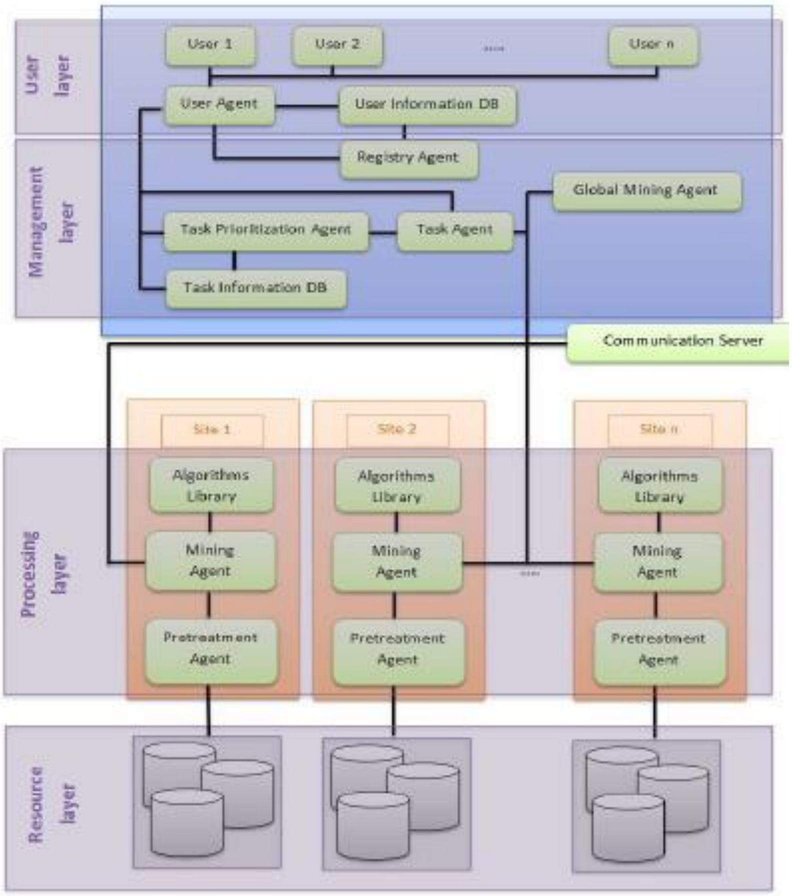


Figure 7. Proposed Architecture for Distributed Data Mining Based on Multi-Agent Systems

In this system, the user agent must establish reciprocal communication with other agents and perform the necessary interactions to complete the user's exploration request. The task assignment method to each agent is described as follows:

1. The user gives the task to the user agent, which creates a data mining task.
2. Through the task agent, the exploration task is assigned to the data preprocessing agent located at each site.
3. After the data preprocessing agent completes its task, it sends the results to the data mining agent at that site to perform the data mining.
4. After the results from the data mining agents of each site are collected by the task agent, they are submitted to the general data mining agent at the central site to be integrated.
5. The general data mining agent prepares the final results and presents them to the user agent.
6. The user agent presents the results to the user.

3.8. *Case Study*

3.8.1. *Performance Analysis of Cyberattack Discovery Algorithms*

The primary difference between the algorithms lies in the method of generating frequent itemsets (L). The performance of the algorithms has been compared using two types of data: experimental data and real data. The parameters used for comparing these algorithms are as follows:

T5.I2.D100k $\Rightarrow$ T=5, I=2, D=100,000
T10.I2.D100k
T10.I4.D100k
T20.I2.D100k
T20.I4.D100k
T20.I6.D100k

- D: Number of transactions
- T: Average transaction size
- I: Average size of frequent items
- L: Number of frequent items
- N: Number of items

- K: 1000, number of items = 1000

The symbol above each of the charts indicates the transactions, average frequent item size, and average transaction size. For example, K100D2I5T indicates 10,000 = D, 2 = I, and 5 = T, meaning the experiment was conducted for 10,000 transactions with an average frequent item size of 2 and an average transaction size of 5. The horizontal axis represents the minimum support. Various experiments were conducted for different sample sets, and the results are shown in the charts below. The execution times of the SETM algorithm were so high that they could not be included in the following charts (Figure 8).

Upon careful observation of these charts, it is apparent that: the Apriori algorithm consistently outperforms the AIS algorithm, and Apriori performs better than AprioriTid in large-scale data sets. In the AprioriTid algorithm, values are considered in place of the database. If the database can fit in memory, this algorithm operates faster than Apriori. However, when the database becomes too large, it cannot fit in memory, and as a result, the calculation time increases significantly. Therefore, Apriori performs faster than AprioriTid.

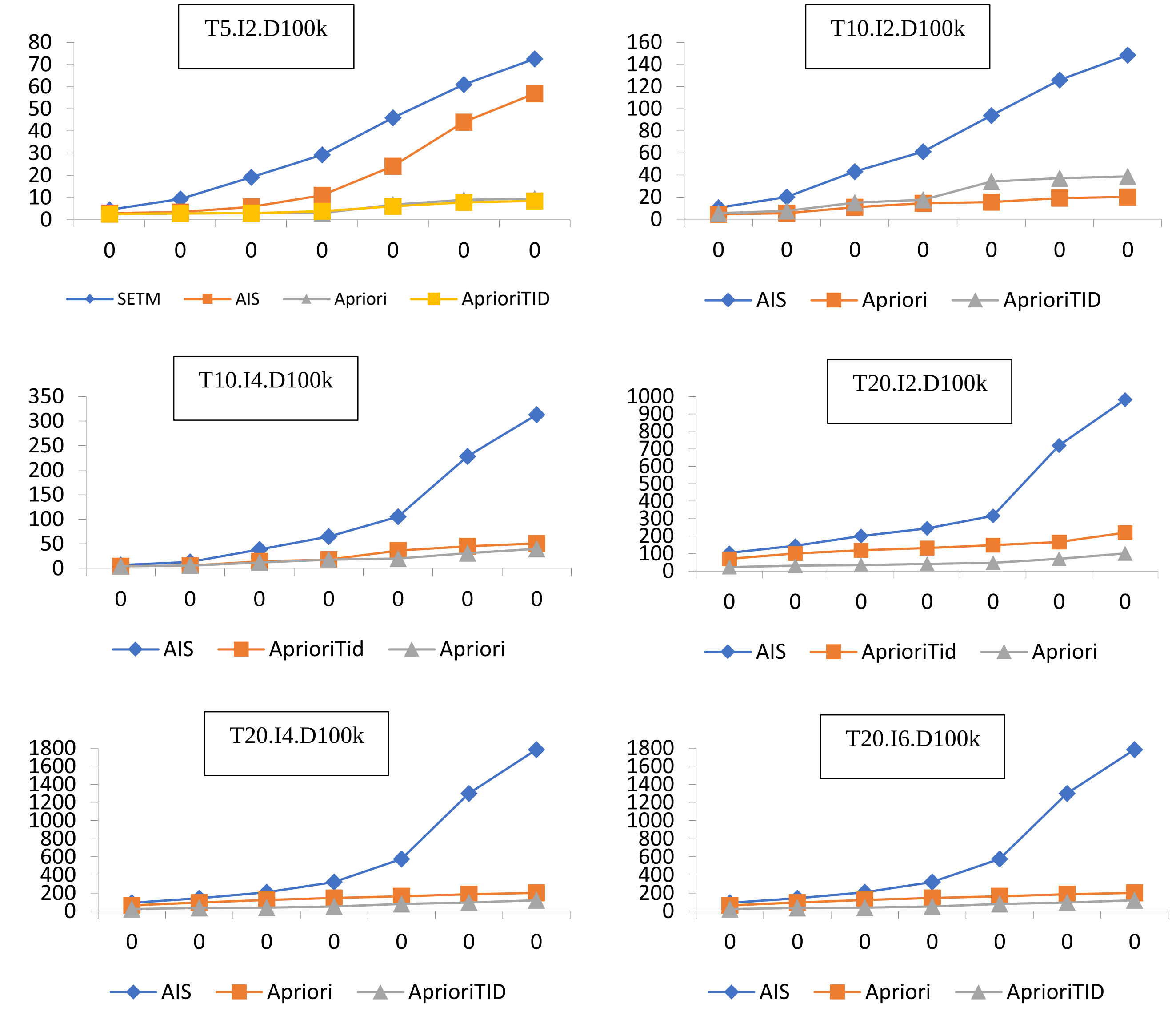


Figure 8. Performance Behavior of Various Algorithms

3.8.2. Real Data

The bank includes:

- 63 branches
- 46,873 transactions (with an average size of 47.2)

As seen, the size of the database is small, so there are no memory issues. Therefore, the AprioriTid algorithm executes faster than the Apriori algorithm. So, which one is better: AprioriTid or Apriori? To answer this question, a comparison between these two algorithms was conducted across different phases, and the results are shown in the following chart (Figure 9):

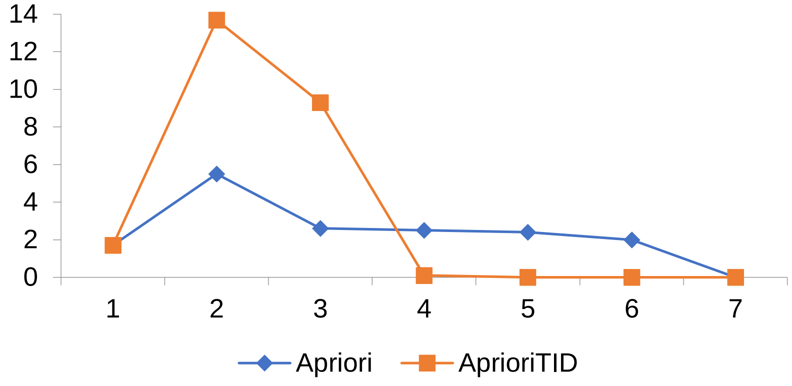
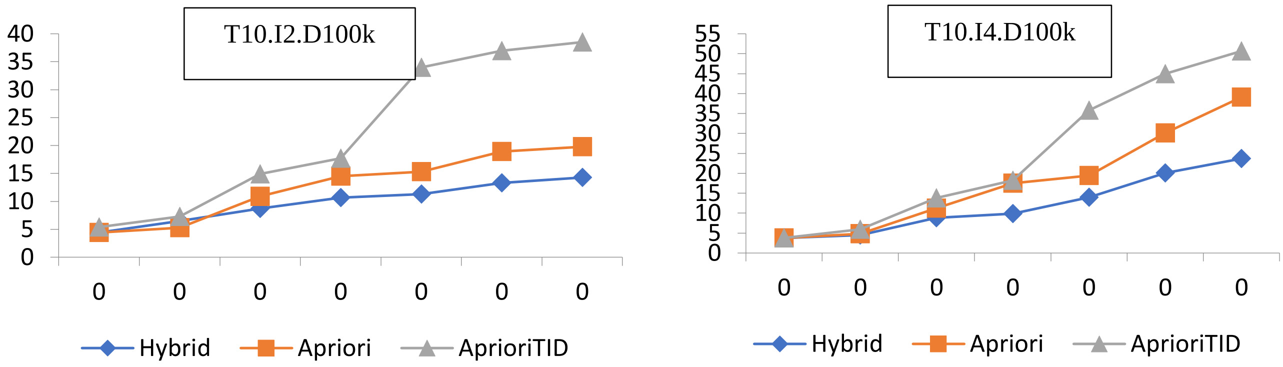


Figure 9. Comparison of the Performance of AprioriTid and Apriori Algorithms

At later stages, the size of the data becomes sufficiently small, and memory consumption is reduced. Therefore, from Phase 4 onward, the execution time of the AprioriTid algorithm decreases significantly, nearly reaching zero. To optimize the use of these two algorithms, a new algorithm called AprioriHybrid was developed. The characteristics of this algorithm are as follows:

- In the initial phases, it operates in the same way as the Apriori algorithm.
- The estimated size is calculated as follows:
- Total transactions + sum of support of all items = estimated size
- When the sizes become small enough and memory consumption is reduced, the algorithm switches to AprioriTid and follows the procedure of this algorithm.
- Although the switch from Apriori to AprioriTid is time-consuming, it often results in positive outcomes. In the charts below (Figure 10), the performance of the three algorithms is compared. In all of these charts, it is shown that the hybrid algorithm has a shorter execution time compared to both Apriori and AprioriTid.



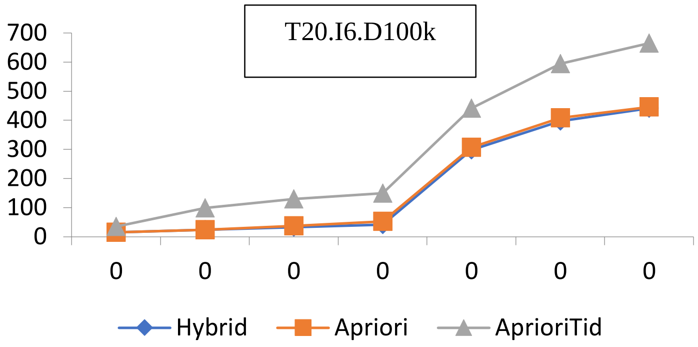


Figure 10. Comparison of the Performance of AprioriTid, Apriori Hybrid, and Apriori Algorithms in Various Experiments

3.8.3. Multi-Source Data Mining Based on Agents

Organizational data mining applications often use multiple, distributed, and heterogeneous data sources, and integrating them can be costly. Data mining agents, each designed for specific algorithms, focus on exploring local patterns in each individual data source, while other coordinating agents work together to organize the distributed pattern mining process. Pattern integration agents control the aggregation of local patterns and the creation of global patterns.

For example, in the diagram below, the basic principles of multi-source hybrid mining within the MSCM-AKD framework are shown. The system can be implemented by agents. As seen in Figure 10, this system may include the following agents:

- **Data Resource Management Agents**: Data controllers and coordinators responsible for managing communication between data sets, partitioning data, and efficiently dispatching data mining tasks to the data sets. In adaptive learning, change controllers are developed to monitor significant data changes.
- **Local Data Mining Agents**: Data miners m1 through mN, designed to perform pattern mining on data sets DB1 through DBN.
- **Data Mining Coordinating Agents**: These agents are responsible for coordinating the scheduling of local pattern mining by data miners on each data set. They carry out this task according to a specific protocol, which could either be a predefined instruction or a decision made dynamically. Another responsibility of these agents is to control the execution state of a data miner and notify others when to begin the pattern mining process.
- **Local Pattern Set Management Agents**: These agents assist in managing the local patterns generated by the local data mining agents.
- **Pattern Integration Agents**: These agents are responsible for merging local patterns and creating a set of global patterns. These agents are designed based on pattern integration methods, such as clustering methods.
- **Knowledge Management Agents**: These agents are responsible for managing the knowledge collected from various sources.

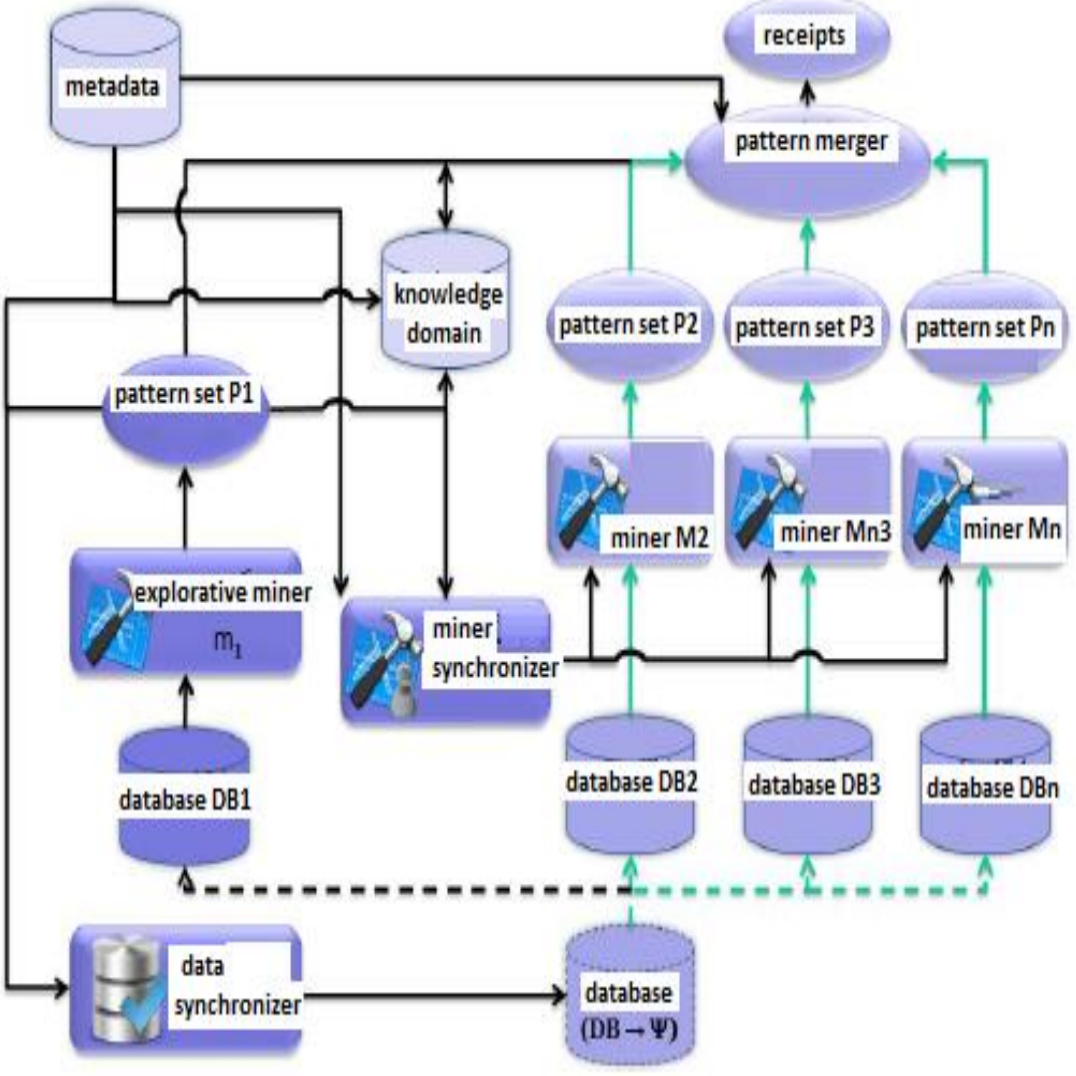


Figure 11. MSCM-AKD Framework

In practice, the agents mentioned above are introduced as application-specific agents to perform specific tasks. For example, a data mining agent may be customized as a cyberattack pattern miner, a repeated sequence miner, etc. The data mining coordinator may be customized to coordinate the execution of data mining, focusing on tasks such as scheduling local dataset exploration, sending messages from one data miner to another, etc.

Many other types of agents may also be employed. For instance, user interface agents are created to interact with domain experts in order to gather domain knowledge from them and send it to knowledge management agents. Negotiation agents may also be used. For example, in an agent-based business system, as described, commercial agents with multiple strategies may negotiate with each other to increase their individual profits and benefits, as well as the overall profit for the business agent managed by the global coordinator.

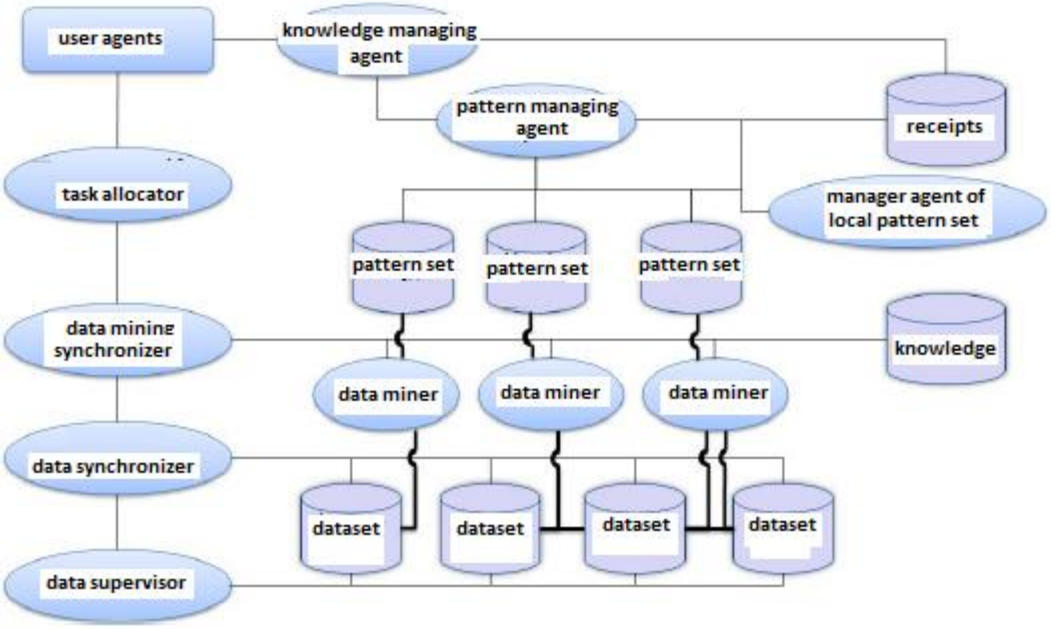


Figure 12. Hybrid Multi-Source Mining Based on Agents

Another example of distributed agent-based data mining is illustrated in Figure 12, which shows the mechanism for the participation of distributed agents in data mining tasks and in the collection of global patterns.

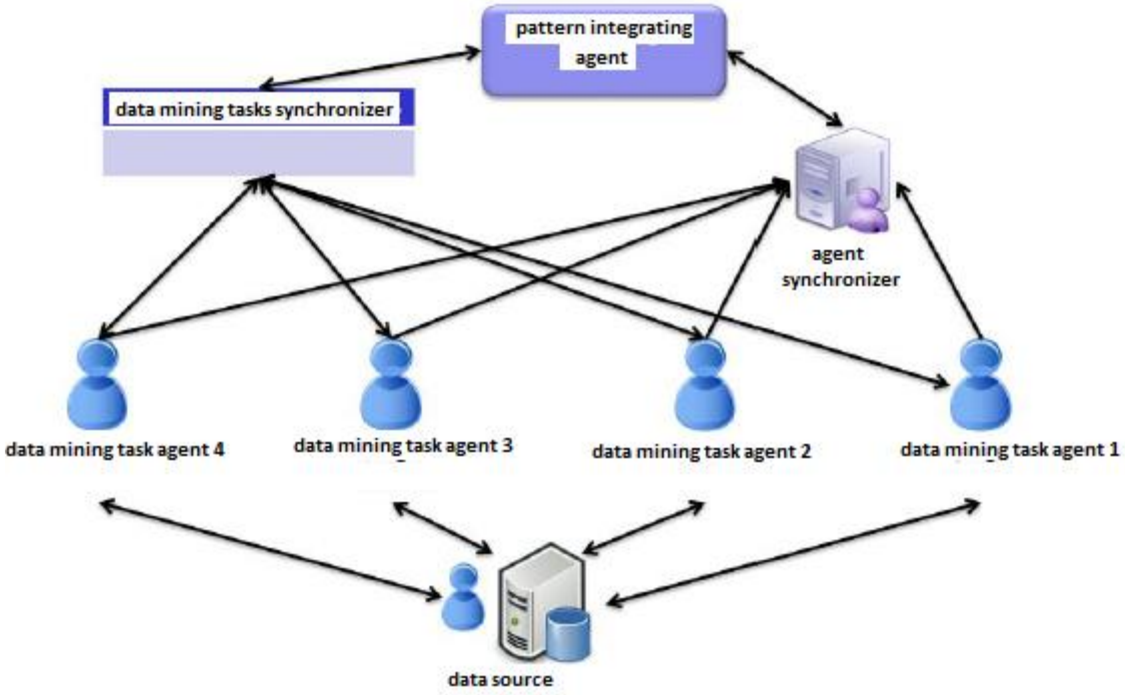
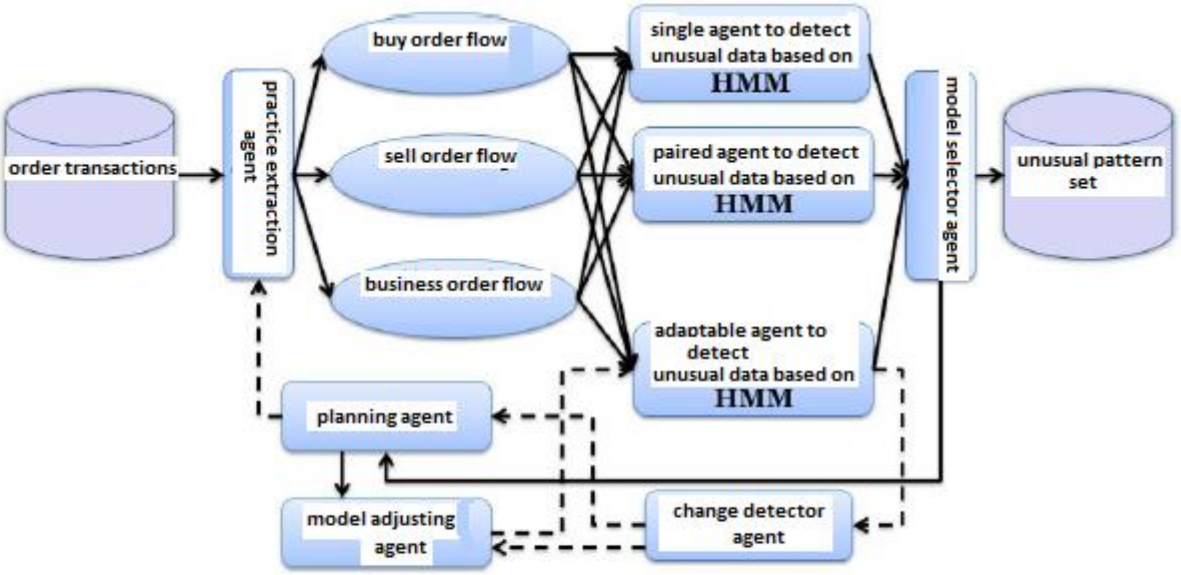


Figure 13. Multi-Source Mining with Agent-Based Networking

3.8.4. Agent-Based Adaptive Behavior Pattern Mining Using HMM

In the discovery of abnormal business behavior in market data, agents are used to identify anomalous business behavior in dynamic activity streams. The framework of this system, designed using agent technology, is shown below.



order transactions
practice extraction agent
buy order flow
sell order flow
business order flow
single agent to detect unusual data based on HMM
paired agent to detect unusual data based on HMM
adaptable agent to detect unusual data based on HMM
model selector agent
unusual pattern set
planning agent
model adjusting agent
change detector agent

Figure 14. Agent-Based Anomaly Pattern Discovery Framework

This system consists of several agents, including: activity extraction agent, anomaly detection agent based on HMM, change detection agent, model tuning agent, and planning agent. These agents collaborate to find the best training model and then use it to develop the activity pattern discovery model.

Each agent has specific objectives and roles and follows particular communication rules to cooperate with other agents and perform data mining tasks. The fundamental objectives and roles of all agents are outlined below.

- **Agent: Activity Extraction Agent**
- **Objective**: Extract activity sequences from data sources
- **Role**: Activity extractor
  - Understand types of activities in data sources
  - Partition data sources according to activity types
  - Extract activity sequences
- **Agent: HMM-Based Anomaly Detection Agent**
- **Objective**: Identify anomalous data in activity streams using HMM-based models
- **Role**: Anomaly detector
  - Train HMM-based models
  - Test HMM-based models
  - Present identified anomalous patterns
- **Agent: Model Selection Agent**
- **Objective**: Select the best models from a list of given models
- **Role**: Model selector
  - Collect performance data according to given criteria for each model
  - Compare model performances
  - Determine the best model
- **Agent: Planning Agent**
- **Objective**: Coordinate the scheduling of agent activities based on inputs
- **Role**: Planner
  - Receive inputs from the change detection agent and model selection agent
  - Send response requests to relevant agents
  - Activate related agent activities

Each agent uses protocols to perform its roles and achieve its objectives. For example, one of the communication protocols used by the planning agent is outlined below.

PROTOCOL: RetrainModelRequest
Requester: PlanningAgent
Responder: HMM-based AnomalyDetectionAgent
Input: The best model with model_id
Rule: model_id. Training. Fulfilled()
Output: retrain.Successful() = True

4. Discussion and Conclusion

Information security is one of the critical issues in e-commerce. Additionally, software development and increased security in banking systems are crucial factors for the acceptance and expansion of electronic banking processes. If the necessary conditions for meeting these two requirements are established, the general use of electronic systems will be expanded and facilitated. Furthermore, the risks associated with using such systems, while maintaining high security, will decrease, and customer trust and satisfaction will increase. Research indicates that banks must first establish trust with their customers, ensuring that electronic banking and the operations conducted within it are sufficiently secure. Significant advancements in technology have greatly contributed to improving the security of transmitted data in electronic banking. However, challenges in the expansion and development of electronic banking still persist. Electronic banking security is most often discussed in the context of online banking. The internet, as a public network, faces issues of confidentiality and information security. Generally, online and real-time banking can pose significant risks to financial institutions and businesses. Customer information and financial transactions are highly sensitive and confidential, and conducting these transactions over the internet introduces challenges in transaction security and trust. Without security, electronic banking would not only be useless but would also cause significant losses. While absolute security does not exist, at least, a non-vulnerable status must be ensured through investments in security measures.

This paper aims to provide an effective solution for detecting and controlling one of the most critical types of cybercrime attacks, which poses the greatest damage to customer trust in the field of e-commerce, particularly electronic banking. To enhance the accuracy of detecting cybercrime attacks arising from cyber events, this study utilized fuzzy theory and, for the first time, improved the fuzzy expert system for detecting cybercrime attacks arising from cyber events using rough set theory.

Distributed multi-agent data mining systems represent a relatively new area of research. They align well with emerging artificial intelligence technologies, offering specific flexibility and intelligence, improving resource utilization, and enhancing system stability. Agent and multi-agent technology is rapidly growing and offers a strong development mindset. Various parts of the system can be studied as independent research topics to improve system functionality by leveraging human thinking, thus reducing the complexity of system design and development. The starting point for agent and multi-agent technology coincides with the development of computer technology, and many discussions and concepts on development and software design ideas exist. Some researchers believe that agent technology could be the next generation of software systems after object-oriented technology. However, this goal is still far from realization, and developers can only conduct extensive research on agent technology in software design to develop a software system with excellent specifications. It is noteworthy that, with the development of this technology, the complexity of developing multi-agent data mining systems based on software will decrease, while the system's intelligence, flexibility, and stability will increase significantly.

In this study, a four-layer architecture based on multi-agent systems for data mining in distributed environments is proposed. The PADMA model was used in the design of this architecture, with modifications such as the addition of user agent, general data mining agent, user information database, task information database, task prioritization agent, and task agent. These modifications offer advantages such as increased scalability, intelligence, and security of the system, better resource utilization, reduced data redundancy, and decreased communication costs between agents.

The results of this research are as follows:

a) The output of the designed fuzzy expert system depends on the design of the fuzzy knowledge base rules, which were created based on expert opinions in the field of online banking.

b) The proposed model was simulated in the MATLAB software environment, making it easy to calculate flow values for any input data and obtain the corresponding output. The fuzzy expert system can automatically retrieve input data by connecting to a local database and also automatically record the output results. Therefore, processing a large volume of data in a short time is possible.

c) The system was tested using real data from cybercrime attacks arising from cyber events on bank websites.

d) The factors influencing the detection of cybercrime attacks arising from cyber events in electronic banking in Iran were extracted.

e) Although the system designed in this paper focuses on the banking sector, it is highly flexible for detecting various types of cybercrime attacks arising from cyber events on the internet. With minor adjustments, it can be applied to other e-commerce websites.

Authors’ Contributions

Authors equally contributed to this article.

Acknowledgments

Authors thank all participants who participate in this study.

Declaration of Interest

The authors report no conflict of interest.

Funding

According to the authors, this article has no financial support.

Ethical Considerations

All procedures performed in this study were under the ethical standards.

References

[1] O. Osho, U. L. Mohammed, N. N. Nimzing, A. A. Uduimoh, and S. Misra, "Forensic Analysis of Mobile Banking Apps," in *International Conference on Computational Science and Its Applications*, 2019: Springer, Cham, pp. 613-626, doi: 10.1007/978-3-030-24308-1_49.

[2] B. Kaulu, G. Kaulu, and P. Chilongo, "Factors influencing customers' intention to adopt e-banking: a TAM and cybercrime perspective using structural equation modelling," *Journal of Money and Business,* vol. 4, no. 1, pp. 38-53, 2024, doi: 10.1108/JMB-01-2024-0007.

[3] A. Pambudi, R. Widayanti, and P. Edastama, "Trust and Acceptance of E-Banking Technology Effect of Mediation on Customer Relationship Management Performance," *ADI Journal on Recent Innovation,* vol. 3, no. 1, pp. 87-96, 09/30 2021, doi: 10.34306/ajri.v3i1.538.

[4] J. Jose, "The Influence of Gamification on Customer Experience in Digital Banking Practices," *SSRN Electronic Journal*, 2024, doi: 10.2139/ssrn.4715254.

[5] P. Kumar, A. K. Mokha, and S. C. Pattnaik, "Electronic customer relationship management (E-CRM), customer experience and customer satisfaction: evidence from the banking industry," *Benchmarking: An International Journal,* vol. 29, no. 2, pp. 551-572, 2022, doi: 10.1108/BIJ-10-2020-0528.

[6] K. D. Kolawole, R. T. Salman, S. E. Durodola, D. Babatunde, and E. O. Igbekoyi, "Determinants of forensic accounting and its effects on alleviation of fraud practices in Nigerian Deposit Money Banks," 2018.

[7] A. W. Henry and A. B. Ganiyu, "Effect of forensic accounting services on fraud reduction in the Nigerian banking industry," *Advances in Social Sciences Research Journal,* vol. 4, no. 12, 2017, doi: 10.14738/assrj.412.3342.

[8] D. K. Gupta, "Growing Needs of Forensic Audit in Corporate and Banking Frauds in India," 2020, doi: 10.2139/ssrn.3624001.

[9] R. Chanajitt, W. Viriyasitavat, and K. K. R. Choo, "Forensic analysis and security assessment of Android m-banking apps," *Australian Journal of Forensic Sciences,* vol. 50, no. 1, pp. 3-19, 2018, doi: 10.1080/00450618.2016.1182589.

[10] A. A. Uduimoh, O. Osho, I. Ismaila, and M. A. Shafi'i, "Forensic Analysis of Mobile Banking Applications in Nigeria," *i-manager's Journal on Mobile Applications and Technologies,* vol. 6, no. 1, p. 9, 2019, doi: 10.26634/jmt.6.1.15704.

[11] O. Aigienohuwa, E. I. Okoye, and E. O. Uniamikogbo, "Forensic accounting and fraud mitigation in the Nigerian banking industry," *Accounting and Taxation Review,* vol. 1, no. 1, pp. 177-195, 2017.

[12] M. H. A. Abdulrahman, M. S. Ab Yajid, A. Khatibi, and S. F. Azam, "THE IMPACT OF FORENSIC ACCOUNTING ON FRAUD DETECTION IN THE UAE BANKING SECTOR: AN EMPIRICAL STUDY," *European Journal of Social Sciences Studies,* 2020.

[13] S. Abdulrahman, "Forensic accounting and fraud prevention in Nigerian public sector: A conceptual paper," *International Journal of Accounting & Finance Review,* vol. 4, no. 2, pp. 13-21, 2019, doi: 10.46281/ijafr.v4i2.389.

[14] N. Scudder, R. Daniel, J. Raymond, and A. Sears, "Operationalising forensic genetic genealogy in an Australian context," *Forensic Science International,* vol. 316, p. 110543, 2020, doi: 10.1016/j.forsciint.2020.110543.

[15] H. K. Khanuja and D. Adane, "To Monitor and Detect Suspicious Transactions in a Financial Transaction System Through Database Forensic Audit and Rule-Based Outlier Detection Model," in *Organizational Auditing and Assurance in the Digital Age*: IGI Global, 2019, pp. 224-255.

[16] N. Phumkaew and V. Visoottiviseth, "Android Forensic and Security Assessment for Hospital and Stock-and-Trade Applications in Thailand," in *2018 15th International Joint Conference on Computer Science and Software Engineering (JCSSE)*, 2018: IEEE, pp. 1-6, doi: 10.1109/JCSSE.2018.8457347.

[17] Á. MacDermott, T. Baker, P. Buck, F. Iqbal, and Q. Shi, "The Internet of Things: Challenges and considerations for cybercrime investigations and digital forensics," *International Journal of Digital Crime and Forensics (IJDCF),* vol. 12, no. 1, pp. 1-13, 2020, doi: 10.4018/IJDCF.2020010101.

[18] A. Akintunde, "Contemporary issues in Forensics Accounting and Forensics Audit," *JOURNAL OF ACCOUNTING, FINANCE AND DEVELOPMENT (JAFID),* p. 147, 2019.