
Presenting a Model for Recognizing Phishing Sites and Privacy Violations in the Tourism Industry

Fereidoon Rezaei¹, Mohammad Ali Afshar Kazemi^{2*}, Mohammad Ali Keramati³

1. PhD Student, Department of Information Technology Management, Central Tehran Branch, Islamic Azad University, Tehran, Iran.
2. Department of Industrial Management, Central Tehran Branch, Islamic Azad University, Tehran, Iran. (Corresponding Author)
3. Department of Industrial Management, Central Tehran Branch, Islamic Azad University, Tehran, Iran.

Article history:

Received date: 2022/11/12

Review date: 2023/02/02

Accepted date: 2023/02/15

Keywords:

Phishing, Privacy, Data Mining, Multilayer Perceptron.

Purpose: Electronic Tourism is one of the important components of expanding Tourism by synchronizing this industry with information technology.

It has not been long since its emergence.

Methodology: this field is a combination of tourism and information technology that is one of the most common types of income-generating businesses which is producing job opportunities in the modern world. The advancement of science alongside communication and information technologies presented many opportunities and threats to this field due to tech such as smartphones and sensors, virtual and augmented reality tools, NFC, RFID, etc.

Findings: The disclosure of the tourists' information and the possible abuse of it is one such threat. Therefore privacy and non-disclosure of information should be important factors. Recognition of reputable sites is an important factor in solving this problem. In this study, we have presented a model for recognizing fake and phishing sites which use the CFS+PSO and a combination of Info+Ranger alongside their results to reduce the test dataset features so that it could present a model for categorizing and higher accuracy in recognizing phishing sites by using the Multilayer Perceptron method. The proposed model was successful in recognizing 95.5% of phishing sites.

Conclusion: The effect of information technology on the tourism industry and the usage of internet websites for selling and providing tourism services to tourists have created new security challenges. Protecting the privacy and personal information of people and tourists is one of these challenges and the disclosure of such information could lead to abuse by unqualified people and dissatisfaction and distrust of such systems.

Please cite this article as: Rezaei F, Afshar Kazemi MA, Keramati MA. (2023). Presenting a Model for Recognizing Phishing Sites and Privacy Violations in the Tourism Industry, *Iranian Journal of Educational Sociology*. 6(1): 222-232.

* Corresponding Author: M_afsharkazemi@iauec.ac.ir

1. Introduction

Nowadays tourism is considered one of the most important service industries in the world. Governments recognize tourism as a tool for developing and protecting culture and tradition with the least possible negative effect (Buhalis, D & Law, R. 2008). Tourism is a series of activities that the tourist performs away from their place of living and working due to personal or work reasons (Morvati, S & Ali and Asadian, A & Faezeh. 2014).

Today internet has turned into an integral part of people's socio-economical infrastructure of daily activities. Huge amounts of private and secret information on the internet creates a plethora of threats and attacks which might cause financial loss, identity theft, loss of private information, defamation of company name and loss of clients' trust in electronic commerce (Buhalis, D & Law, R. 2008). Expansion of the scope of e-commerce has resulted in increase in energy consumption, energy management becoming more complex, growth in the volume of data, the emergence of need for high bandwidth for sending data, and high performance process systems. Two of the most important ones include privacy protection and information security (Atafar, A & Khazaei Pool, Javad & Pour Mostafa Khoshkroudi, M 2012), (Büyükoçkan, G. & B. Ergün 2011), (Khosravi, A; Ganjoo, Maziar and Mazarei, Hassan 2017).

Elements which lead to threat and attack to an internet network include limitless access to internet, anonymity of people, high speed release, no face to face contact, free access to valuable services and contents and also lack of appropriate laws. Therefore, internet must be functional as a channel for commercial trade.

Intruder detection systems which work with data mining can be very effective. These systems can use training in real networks and real and unrestricted data sets to detect new and hybrid methods. In fact, detection system acts like a gateway by receiving a package as input. The system starts collecting information and deleting their unused features by preprocessing data. Then, data are sent for processing. After data is analyzed, it is classified. Natural records then are sent to the main network and unnatural ones are sent to next levels of processing (Awazu, Y. & Desouza, K. C. 2004).

On the other hand, the advancements of communication and information technology have had a considerable effect on different fields which have prepared the context for the transformation, acceleration, and facilitation of affairs by creating an integrated foundation. The world of business is not separate from these advancements and the tourism industry as a profitable and important business in the world has faced many new experiences and developments during the past few years and tried to take a step in matching itself with the world of modern technologies (Atafar, A & Khazaei Pool, Javad & Pour Mostafa Khoshkroudi, M 2012). Since the development of this industry is one of the most important money-making parts of an economy, the countries were successful that uses new aspects of communication and information technology to find newer ways of using the capabilities of this field.

Technology in the tourism industry has created the foundation of creation and evolution by affecting operational processes. Nowadays, the customer would like to do the selection of their tourism product similar to other services on a platform using information and communication technologies; and this is one of the features of the smart citizen of the third millennium.

Using the current communication and information technology foundation, tourism organizations can present themselves and services to potential customers without any time and space constraints; and the tourists can receive the information they want with high speed and accuracy and also pay for these services (Büyükoçkan, G. & B. Ergün 2011).

On the other hand, personal information and corporate secrets are important assets that tempt others to perform business with them. Andrew Grove, one of the execute managers at Intel, expresses his privacy concerns and says: "Privacy is one of the biggest problems in the new technological age. If some companies don't pay enough attention to the privacy of their stakeholders (customers), employees, business partners, providers, and other real and legal persons, they will lose their market share." (Khosravi title, A, Ganjoo, M and Mazarei, H, 2017)(Awazu, Y. & Desouza, K. C. 2004)(Bamberger, K.A. & Mulligan, D.K.

2011)(Bamberger, K.A. & Mulligan, D.K. 2013) In reality, modern technologies have made it possible for companies to put their intended information into the privacy of others or use illegal tools or even seemingly legal ones to disclose this information to gain monetary value, more profits, eliminating the competition or monopolize their business (Shafi'i, S. 2010).

The concerns regarding information privacy are related to the understanding of people over their right to control their own private data so disclosing private data is considered a dangerous act because it compromises people against the opportunistic actions of a company (Milne, GR., Gordon, ME. 2010)(Laufer, RS., Wolfe, M. 1977).

Factors that can create a threat and attack in internet networks are limitless access to the internet, people's anonymity, high release rate, lack of face-to-face communication, free access to valuable services and content, and also lack of proper rules and agreements (Ma'ouni, M. 2015). Therefore, the appropriateness of the Internet as a channel for performing business transactions is expressed.

In the early 1990s, we saw the rise of the internet due to its popularity which gave way to a new method of cybercrime: Phishing (Mohammad, R. M., Thabtah, F., McCluskey, L. 2015).

In Phishing unlike other hack and penetration methods, there are no penetrations and no weaknesses are used, but it's the user itself that is tricked using different methods to share information such as user name, password, and banking information with the attacker (A.K.A Phisher) (Chaudhry, J. A., Rittenhouse, R. G., 2015). Based on a study by the "Anti Phishing Working Group", phishing attacks have decreased during the final quarter of 2019 and gotten closer to the average. Of course, this number has had a 232% increase in Brazil in the same year. Phishing attacks that target web, e-mail, and software service users are still the largest type of phishing attacks. Around three-fourths of all phishing sites use SSL protection, this is the highest value since early 2015 which shows that users cannot just rely on SSL and require an understanding of more features (Anti Phishing Working Group. 2019).

There have been many definitions of web phishing, all of which can be summarized into one sentence: "Web Phishing is the process of creating a copy of a legal site and using social skills to trick the victim into sharing their personal information" (Mohammad, R. M., Thabtah, F., McCluskey, L. 2015).

Also with the growth of electronic tourism and the important actions taken by centers and agencies providing tourism services, especially in the field of providing new tourism services in the country, It has become important to analyze the effects of privacy worries of the user on their interaction with tourism services and also airplane and hotel ticket selling websites. Therefore, this paper presents a smart and effective algorithm for recognizing phishing sites that have created concerns for the customers and made them lose their trust.

The data used in this study are from a series of real and fake websites extracted from the UCI mining website (Mohammad, R. M., Thabtah, F., McCluskey, L. 2015).

Research Background

In 1967, Westin published "Privacy and Freedom" to complete the definition of privacy and present a fourfold criterion for people's privacy. Westin defined the revolution and evolution of information privacy in four steps by following the information technology revolution (Jensen, C., Potts, C. & Jensen, C. 2005)(Jutla, D.N., Bodonk, P. & Zhang, Y. 2006).

Ibaris et al. (2010) presented a fuzzy rule-based classification algorithm for recognizing electrical banking phishing websites (commerce contexts on the semantic web. 2006).

Pandi and Ravi (2012) used text mining and data mining for recognizing phishing emails. They reached 12 important features using the t-statistic feature selection alongside SVM classification, logistic regression, decision-making table, MLP, MDH, PNN, and genetic programming. These algorithms were once analyzed while selecting the features and once without selecting them (Pandey, M., Ravi, V. 2012). They presented a better method in another study and tested it using a cross-validation method with 10 iterations (Pandey, M., Ravi, V. 2013).

Langari and AbdalRazaghNezad (2016) used features extracted from the Inclined Planes System Optimization (IPO) algorithm to divide websites into the three categories of legal, suspicious, and phishing (Langari, Nafiseh, Abdolrazaqnejad, Majid, 1394).

In the paper (Rezaei F, Afshar Kazemi M A, Keramati M A. 2021), neural network is used as a way to deep training and detecting attacks and anomalies in e-commerce system. Error of “overfitting” is very common in multi-layered neural networks. In the aforementioned paper, the number of features have been cut by using firefly algorithm in order to prevent the effect. Results of simulation show that neural network system with the help of decreasing features have higher precision.

Using models and mechanisms of machine learning is one of the methods for countering threats and attacks (Doshi, Rohan, Noah Apthorpe, and Nick Feamster. 2018) (Syed, Naeem Firdous, et al. 2020). Machine learning is a subset of artificial intelligence and mostly relies on machine’s own experience and forecasting based on this experience. Machine learning algorithms use data sets called training data sets to learn and make required models. When new data are introduced to machine learning algorithm, system can forecast based on created model. Artificial Neural Network is one the most used and famous, and also one of the strongest, algorithms and models of machine learning (Manimurugan, S., et al. 2020) (Latif, Shahid, et al. 2020).

In 2015, Sink et al. used four preprocess algorithms – CFS, IG, consistency subset and PCA – for decreasing the dimensions. They also used five classifying algorithms – J48, naïve Bayes, SVM, random forest and AdaBoost – for evaluating feature-decreasing algorithms in terms of precision and AUC (Singh, P., Jain, N., Maini, A., 2015).

In 1396, Alireza Bahraami et al. used features-selection methods and decision tree J48 to decrease features and detect intrusion (Alizadeh Bahrami, Karimi, Abdullahi Fard, 2016). In 1399, Baharlou et al. used the two-step method called Wrapper+PCA and random forest algorithm to detect and recognize fishing websites (Baharlo, Yari, 2019).

2. Methodology

Figure 1 shows the general flowchart and structure of the proposed method. As you can see, the proposed system's data must be gathered in the first step. This paper uses the Phishing Dataset of the UCI website (Mohammad, R. M., Thabtah, F., McCluskey, L. 2015). Preprocessing operations such as removing duplicate data, normalization, feature engineering, and vectoring is done during the second step. Afterward, we use different methods to select the important features and then implement different algorithms such as the J48, BayesNet, and Multilayer Perceptron on the data. Finally, the results are compared.

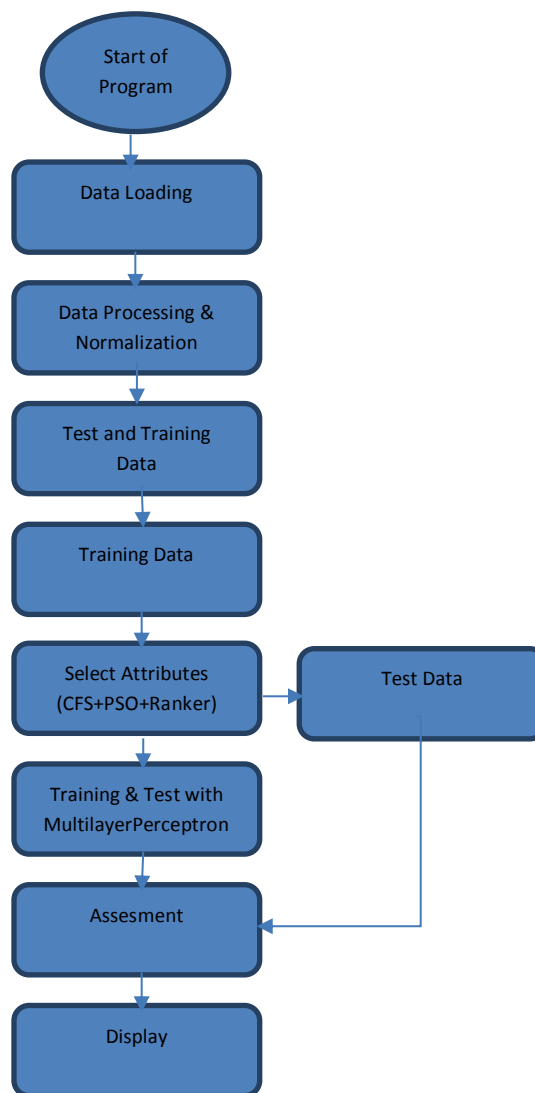


Figure 1: Phishing recognition algorithm

Test Tools

The Weka 3.9.2 was used for the implementation of this study; Weka is an open-source software developed by the Waikato university (Pandey, M., Ravi, V., 2012) used in the Windows 7 environment with an Intel Core i5 processor and a 4GB ram.

3. Findings

Every website has its own unique features which makes it stand out among other websites and even websites that closely resemble them. As such, fishers always try to come up with a trick to make a fake website and change some features of legal websites (Langari, Nafish, Abdolrazaqnejad, Majid, 1394). Data set of UCI, a data mining website, was examined. This set of data contains 30 features based on four following criteria (Table 1).

Table 1: Phishing sites' features

Row	name
1.1.	Address Bar based Features
1.1.1	Using the IP Address

1.1.2	Long URL to Hide the Suspicious Part
1.1.3	Using URL Shortening Services “TinyURL”
1.1.4	URL’s having “@” Symbol
1.1.5	Redirecting Using “//”
1.1.6	Adding Prefix or Suffix Separated by (-) to the Domain
1.1.7	Sub Domain and Multi Sub Domains
1.1.8	HTTPS (Hyper Text Transfer Protocol with Secure Sockets Layer)
1.1.9	Domain Registration Length
1.1.10	Favicon
1.1.11	Using Non-Standard Port
1.1.12	The Existence of “HTTPS” Token in the Domain Part of the URL
1.2. Abnormal Based Features	
1.2.1	Request URL
1.2.2	URL of Anchor
1.2.3	Links in <Meta>, <Script> and <Link> tags
1.2.4	Server Form Handler (SFH)
1.2.5	Submitting Information to Email
1.2.6	Abnormal URL
1.3. HTML and JavaScript based Features	
1.3.1	Website Forwarding
1.3.2	Status Bar Customization
1.3.3	Disabling Right Click
1.3.4	Using Pop-up Window
1.3.5	Iframe Redirection
1.4. Domain based Features	
1.4.1	Age of Domain
1.4.2	DNS Record
1.4.3	Website Traffic
1.4.4	PageRank
1.4.5	Google Index
1.4.6	Number of Links Pointing to page
1.4.7	Statistical-Reports Based Feature

The data used in this research are a series of fake and real websites extracted from the UCI data mining site (Mohammad, R. M., Thabtah, F., McCluskey, L. 2015). The data for this research are batch data. Each website in this dataset has a class or goal feature that shows their batch. The value of one shows a fake website, zero shows suspicion and a negative one shows a real website. This dataset includes 11055 phishing and legal websites from 2015, 4898 of which are legal websites and 6157 are phishing sites (Baharloo, Yari, 1399).

Feature Reduction Methods

Data dimensionality reduction methods are divided into feature extraction and feature selection methods (Ismaili, Mehdi, 2013). Feature selection methods try to reduce the data dimensions by selecting a subset of the initial features. On the other hand, feature extracting methods maintain the main semantic of the features after the reduction.

In this paper, we have used a combination of the Cfs+PSO and Info+Ranker methods for feature reduction. Then implemented a combination of these methods after iterative experiments and selected the important features.

Test Results

As table 2 shows, the number of features decreased after applying feature-decreasing methods. According to table 2, the number of features in the normal situation was 30 which was reduced to 8 after using CFS+PSO hybrid method. After applying CFS+BestFirst hybrid method, the number of features reached 8. And in Info+Ranker mode, 10 first features were selected. Finally, combining CFS+PSO and InfoRanker put out 8 important features.

Table 2: The number of features after the reduction

Method Name	Number of Features
Normal Data	30
CFS+PSO	8
CFS+BestFirst	8
Info+Ranker	First 10
CFS+PSO & Info Ranker	8 Important feature (a combination of both methods)

The extracted features in different situations are:

Table 3 shows the extracted features in the CFS+PSO and CFS+BestFirst methods.

Table 3: 8 Features

CFS+PSO & CFS+BestFirst	
1.1.6	Adding Prefix or Suffix Separated by (-) to the Domain
1.1.8	HTTPS (Hyper Text Transfer Protocol with Secure Sockets Layer)
1.2.1	Request URL
1.2.2	URL of Anchor
1.2.3	Links in <Meta>, <Script> and <Link> tags
1.4.2	DNS Record
1.4.3	Website Traffic
1.4.5	Google Index

Table 4 shows the first 10 features selected using the Info+Ranker method in order of their priorities.

Table 4: Features selected using the Info+Ranker method

Info+Ranker	
1.1.8	HTTPS (Hyper Text Transfer Protocol with Secure Sockets Layer)
1.2.2	URL of Anchor
1.1.6	Adding Prefix or Suffix Separated by (-) to the Domain
1.4.3	Website Traffic
1.4.4	PageRank
1.1.7	Sub Domain and Multi Sub Domains
1.4.1	Age of Domain
1.1.9	Domain Registration Length
1.2.1	Request URL
1.2.3	Links in <Meta>, <Script> and <Link> tags

Table 5 includes the 8 important features that were extracted using the combination of the CFS+PSO and Info+Ranker methods.

Table 5: Results combining the CFS+PSO and Info+Ranker methods

CFS+PSO & Info+Ranker	
1.1.8	HTTPS (Hyper Text Transfer Protocol with Secure Sockets Layer)
1.2.2	URL of Anchor

1.1.6	Adding Prefix or Suffix Separated by (-) to the Domain
1.4.3	Website Traffic
1.4.4	PageRank
1.1.7	Sub Domain and Multi Sub Domains
1.2.1	Request URL
1.2.3	Links in <Meta>, <Script> and <Link> tags

After the feature reduction, the J48, BayesNet, and Multiplayer Perceptron classification methods and algorithms were executed in the Weka software; they were tested by allocating 80% of the evaluation criteria teaching and 20% of it for testing to find the best possible feature reduction and best possible classification methods. The results are shown in Table 6.

Table 6: Classification methods' results

Feature Reduction Method	Modeling Methods		
	J48	BayesNet	Multilayer Perceptron
Normal Data	93.5	93.5	94.7
CFS+PSO	92.3	92.7	91.0
Info+Ranker	93.3	93.3	95.1
CFS+PSO & Info+Ranker	93.7	93.3	95.5

In tables (2) and (6), we compared the accuracy of three data mining algorithms to find the best classification algorithm and feature reduction method.

Table (6) shows that the Multilayer Perception algorithm has the highest accuracy with 30 features in Table 1, 10 features in Table 4, and 8 features in Table 5.

4. Discussion

Personal and financial information is one of the upcoming threats on the path to the development of information technology in virtual space. This threat is called fishing. Examining and analyzing present methods shows increased flexibility when selecting influential features in detection process of fishing websites, making classifier algorithm for development behavior of targeted websites dynamic. And also the possibility of analyzing and controlling great volume of websites have been overlooked. So in the paper (Abdolrazaq-Nejad, Majid, 2015), three aforementioned goals were followed. First, a mechanism has been defined based on design of a change threshold for flexible decreasing of features included in detecting fishing websites. Then, memory was assigned to algorithms which optimize slope pages. The effect of memory on algorithm's performance in high iterations was decreased. And 12 fuzzy rules were defined in a system of fuzzy deduction. These steps were taken to give smart dynamic nature to the algorithm in order to classifying evaluation society websites into three categories: legal, suspicious, and fishing. Our paper has used simpler and more agile methods, has had better performance in terms of decreasing features, and has shown better performance in detecting fishing websites.

UCI database, which includes 30 features belonging to web pages such as URL address and IP address, was classified into two evolutionary algorithms: standard cuckoo and comparative cuckoo. This method reduced the number of features to 9. Also, fishing websites were detected with the precision of 93.33% by using Bagging algorithm as classifier. We used data with lowered dimensions. In present paper, first for the sake of simplification we decreased the number of features by using methods and algorithms of CFS, PSO, and Ranker. Then, we did the classifications based on data mining algorithms such as J48, BayesNet, and MultilayerPerceptron. At the end, analysis showed that MultilayerPerceptron with the precision of 95.5% has the highest precision rate among detection algorithms. This Algorithm worked better that previous ones.

One of the limitations of the study was lack of real data for detecting e-commerce attacks and fishing. We used data mining data set of UCI for decreasing the effect of this problem. This data set contains 30 features based on 4 criteria (Table 1).

Nowadays, we are passing from the age of industrialism into post-industrialism; a rapid passage that is creating a challenging age, in a way that innovations such as information technology have become vital for continued survival and stability. It is clear that using technology requires specific foundations, studies, and tools. Since tourism can have many positive and negative effects on the economy of a country and provide the necessities for sustainable development; we should not forget that planning in this industry should be done using correct and logical principles and rules considering the available facilities of today and tomorrow. During the last few decades, information and communication technologies have changed our societies in unpredictable ways. Travel and tourism are some of the sections that underwent the most change. The effect of information technology on the tourism industry and the usage of internet websites for selling and providing tourism services to tourists have created new security challenges. Protecting the privacy and personal information of people and tourists is one of these challenges and the disclosure of such information could lead to abuse by unqualified people and dissatisfaction and distrust of such systems.

Considering the many problems and complexities and the other challenges in recognizing phishing pages, it is essential to analyze and present a smart method of recognizing phishing in websites. This study used the CFS, PSO, and Ranker algorithms to reduce the features for simplification and then used data mining algorithms such as J48, BayesNet, and Multilayer Perceptron for data classification. Finally, the results showed that the Multilayer Perceptron method had the highest accuracy of 95.5% in classification recognition.

The advantage of this proposed method over other similar systems is finding the best possible features after performing feature reduction which leads to the simplification and complexity reduction of the model.

Acknowledgments

In this research, the ethical standards including obtaining informed consent, guaranteeing privacy, confidentiality, etc. are observed, and the participants are hereby thanked.

References

- Alizadeh Bahrami K, Abdullahi Fard. (2016) "J48 Decision Tree in Intelligent Intrusion Detection Systems", National Conference on New Researches in Electrical, Computer and Medical Engineering, Islamic Azad University, Kazeroun Branch. [Persian]
- Atafar A, Khazaei Pool J, Pour Mostafa Khoshkroudi M. (2012), Factors Affecting the Adoption of Information Technology in Tourism Industry, *Tourism Management Studies*. 7(18): 131-156. [Persian]
- Awazu Y, Desouza KC. (2004). The Knowledge Chiefs: CKOs, CLOs and CPOs. *European Management Journal*. 22(3): 339-344.
- Baharloo Y. (2020). "Improving the method of identifying phishing websites using data mining on web pages", *Iranian Journal of Information and Communication Technology*, Iranian Information and Communication Technology Association. 12(44): 27-38. [Persian]
- Bamberger KA, Mulligan DK. (2013). Business perceptions and satisfaction with e-government Information Quarterly. 30(1): 1-9.
- Buhalis D, Law R. (2008). Progress in information technology and tourism management: 20 years on and 10 years after the internet- the state of etourism research. *Tourism management*. 29(4): 609:623.
- Büyüközkan G, Ergün B. (2011). Intelligent System Applications in Electronic Tourism, *Expert Systems with Applications*. 38(6): 6586- 6598.
- Chaudhry JA, Rittenhouse RG. (2015). "Phishing: Classification and Countermeasures", 7th International Conference on Multimedia, Computer Graphics and Broadcasting. 28-31.
- Doshi R, Noah A, Nick F. (2018). "Machine learning ddos detection for onsumer internet of things devices." 2018 IEEE Security and Privacy Workshops (SPW).
- Ismaili M. (2013). *Concepts and Techniques of Data Mining*, Kashan: Sura. [Persian]
- Jensen C, Potts C, Jensen C. (2005). Privacy practices of Internet users: Self-reports versus observed behavior. *International Journal of Human-Computer Studies*. 63(1-2): 203-227.
- Jutla DN, Bodonk P, Zhang Y. (2006). PeCAN: An architecture for users privacy-aware electronic commerce contexts on the semantic web. *Information Systems*. 31(4): 295-320.
- Khosravi title A, Ganjoo M, Mazarei H. (2017), Presenting a Model for Privacy Concerns in Electronic Banking, Third International Conference, Web Research.
- Langari N, Abdolrazaqnejad M. (2015). "Identification of Phishing Website in Internet Banking Using Sloping Page Optimization Algorithm", *Journal of Electronic and Cyber Defense*. 1: 40-29. [Persian]
- Latif Sh. (2020). "A Novel Attack Detection Scheme for the Industrial Internet of Things Using a Lightweight Random Neural Network." (8): 89337-89350. [Persian]
- Laufer RS, Wolfe M. (1977). Privacy as a Concept and a Social Issue: Amultidimensional Developmentl Developmental Theory. *Journal of Social ISSues*, 33(3): 22-42.
- Ma'ouni M. (2015). "Detection of attacks in electronic banking using fuzzy-rough combination system (Fuzzy_rough)" Department of Computer, Imam Reza University (AS). [Persian]
- Manimurugan S. (2020). "Effective Attack Detection in Internet of Medical Things Smart Environment Using a Deep Belief Neural Network." (8): 77396-77404.
- Milne GR, Gordon ME. Direct mail privacy-efficiency trade-offs within an implied social contract framework. *Journal of Pulicy and Marketing*. 12(2): 206-215.
- Mohammad RM, Thabtah F, McCluskey L. (2015). "Tutorial and critical analysis of phishing websites methods", *Computer Science Review*. 17: 1-24.
- Morvati Sharifabadi A, Asadian Ardakani F. (2014), Presenting a health tourism development model with an integrated approach of fuzzy TOPSIS and interpretive structural modeling in Yazd province, *Health Management*. 17(55): 73-8. [Persian]
- Pandey M, Ravi V. (2012). "Detecting phishing e-mails using Text and Data mining", IEEE International Conference on Computational Intelligence and Computing Research(ICCIC). 2012.
- Rezaei F, Afshar Kazemi MA, Keramati MA. (2021). Detection of E-commerce Attacks and Anomalies using Adaptive Neuro-Fuzzy Inference System and Firefly Optimization Algorithm . 13(1) :32-39. [Persian]
- Shafi'i S. (2010). *Civil Liability for Violation of Privacy*, Master Thesis in Private Law, Kashan University. [Persian]

- Singh P, Jain N, Maini A.(2015). "Investigating the Effect Of Feature Selection and Dimensionality Reduction On Phishing Website Classification Problem", 1st International Conference on Next Generation Computing Technologies (NGCT) Dehradun, India. 388-393.
- Syed Naeem F. (2020). "Denial of service attack detection through machine learning for the IoT." Journal of Information and Telecommunication. 1-22.