# Methods and Means of Identifying Fraudulent Websites

Ivanna Sabadash [1][0000-0002-7805-8726] and Nestor Dumanskyi [1][0000-0001-6908-2751]

Igor Korobiichuk [2][0000-0002-5865-7668]

[1]LvivPolytechnic National University, Lviv, Ukraine
[2]University Professor of Warsaw University of Technology, Poland

ivanna.t.sabadash@lpnu.ua , nestor.o.dumanskyi@lpnu.ua

**Abstract.** The present article deals with the basic methods and means of identifying fraudulent websites. The activity of scammers was analyzed: methods and means of attracting users to visit illegal websites and provide their personal, compromising or financial information. The statistics of the negative activity of fraudsters on the use of personal information of users and the total cost of the damage caused have been investigated. Possible harmful consequences of providing personal information to fraudulent sites are described. The methods of self-identification of sites that can engage in fraudulent or illegal activity are presented. A variety of online resources to help identify fraudulent or suspicious sites are considered. The article shows the effectiveness of avoiding fraud on the Internet using methods of check the databases of fraudulent websites and online fraudulent website verification services. The efficiency of using the methods of visual identification and analysis of the site and the efficiency of combining these methods are also shown.

**Keywords:** Fraudulent Websites, Identification of Fraudsters, illegal site, Identifying Methods.

## 1 Introduction

One of the major changes that the business world is experiencing now (especially during the quarantine period) is the progressive development and introduction of e-commerce. Given the rapid development of web and internet technology along with e-commerce, they are increasing the volume fraudulent online services.

The overwhelming majority of Ukrainian consumers are just starting to get acquainted with the features of e-commerce and it is not always successful. The younger generation is making online purchases unconditionally since they are used to «living there». The older generation is too cautious and often afraid to freely use all the benefits of the Internet. There are, however, a fraction of users who are fast paced, reasonably and practically fit for financial activities online. The purpose of this article is to increase the number of the latter by protecting consumers of Internet services from social engineering.

Social engineering, or the luring of user data by criminal means, based on the basic human weaknesses, trust, fear and haste. [5]

## 2 Related works

Asha S. Manek, P. Deepa Shenoy, M. Chandra Mohan, K. R. Venugopal in their work analyze the problem of online shopping platform. Users on the Web sell and buy goods in e-store, active use online banking sphere and often give review about their online shopping experience. Often people deliberately give false feedback with malicious intent to promote fraudulent schemes and distribute fraudulent online business. It is not always necessary to rely on feedback from people on WEB, although for many it is an important component of decision-making. Scientists in their work propose a new method Bayesian logistic regression classifier (BLRFier) that detects fraudulent and fishing websites by analysing user reviews for online shopping websites. They have built dataset by crawling reviews of truthful and fraudulent e-shopping websites to apply supervised learning techniques. Experimental evaluation of BLRFier model reach 100% accuracy meaning the effectiveness of this approach for real-life use. [1] Daisuke Miyamoto, Hiroaki Hazeyama, Youki Kadobayashi in their research introduce HumanBoost, special an approach that aims at improving the correctness of detecting phishing sites by utilizing users' past trust decisions (PTDs). When a text forms of website asks for fill personal information, the user must make a decision about trusting to this site. The researchers in their article suppose that a database of user PTDs would be transformed into a binary vector, representing phishing or not-phishing, and the binary vector can be used for detecting phishing websites, like the existing heuristics For pilot investigate, in November 2007, scientists invited 10 members and conducted a topic experiment. The members of experiment browsed fourteen simulated fraudulent sites and six rightful sites and judged whether each of the site appeared to be a phishing one. Participants' trust decisions used as a new heuristic and let AdaBoost incorporate it into eight existing heuristics. The results show that the average error rate for HumanBoost was 13.4%, whereas for participants it was 19.0% and for AdaBoost 20.0%. Scientists also managed two follow-ups investigate in 2010, observed that the average error rate for HumanBoost was below the others. In the end, they conclude that PTDs are available as new heuristics, and HumanBoost has the potential to improve detection accuracy fraudulent websites for user of Internet. [2]

Daisuke Miyamoto, Toshiyuki Miyachi, Yuzo Taenaka, Hiroaki Hazeyama in their scientific works developed PhishCage, an experimental infrastructure for phishing identifying systems. Because of its short-term existence of fraudulent sites is difficult doing comparison of effectiveness the detection systems. Basic idea of scientific work is developing a tested in which phishing sites can be browsed as if they existed realistically. According to researcher's inspection for phishing detection systems, article defines the requirements for the phishing identifying systems, and designs

PhishCage to according to these requirements. The experiment of PhishCage demonstrates designing algorithm for 121 fraudulent sites into the emulated Internet topology. Researchers confirm that phishing detection systems can obtain the realistic IP address and autonomous system number of the phishing sites in PhishCage, and few modifications enable the websites to work as if they are in the real Internet. Also, they analyzed the limitation of PhishCage, and discuss the expedience of emulation technique. [3]

## 3 Identification of fraudulent sites

According to the Ukrainian Unified Interbank Association (UIA) of all fraudulent activities, 65% account for the use of the Internet and social engineering (see Fig. 1).

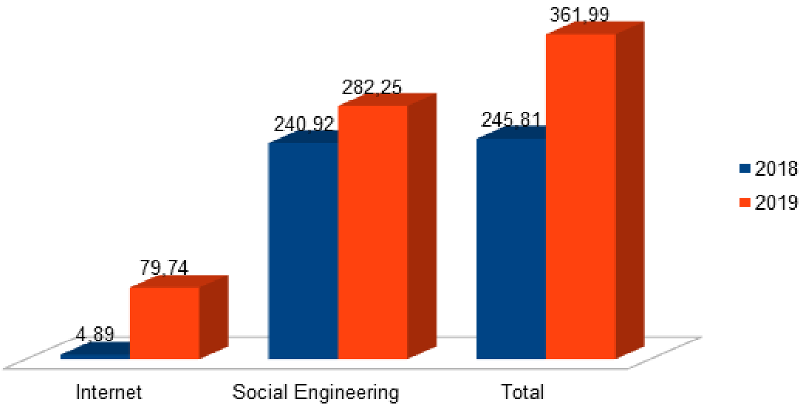

**Fig. 1.** Total fraud revenue (UAH million)

In the case of ATM fraud, banks sometimes compensate for losses, if they can be confirmed by facts. But in social engineering there are no grounds for compensation, because the client voluntarily discloses personal information or transfers money. Some initiative banks may block or require additional confirmation of funds transfer to accounts with suspicious status. But if the client has confirmed and completed the transaction, it will be difficult to return the money. At this time, the return is possible only if the fact of the crime has been acknowledged by the court. [6, 7]

That is why it is advisable to correctly identify whether a fraudulent transaction is taking place. The basic schemes of work of fraudulent sites are divided into copying / cloning of official sites and creation of sites with short term of existence (see Fig. 2).

Also, it is worth mentioning the methods of attracting customers to fraudulent sites. For this the fraudsters use all possible means. They include a variety of mailings (email, SMS, social networks, messengers ...) and advertising on various systems and sites, as well as printed flyers and even advertising on billboards. It is worth remembering how you got on a site, because clicking on a link — the easiest way, but not always the safest.

Global informatization has also made possible to make fast payments via the Internet, make purchases, study and work online. About one million websites are created every day in the world. Among them, about 25% intended to deceive users to get their money or personal information (bank card numbers, usernames and passwords of accounts, compromising personal information for future blackmail).

The online environment is filled with sites that offer easy money, cheap online courses, incredible winnings, and other easy-to-enrich methods. According to statistics, about a third of such sites are fraudulent. Every half a minute on the Internet there is a new website which is designed for phishing money from its users. Fraud is the unlawful enrichment of site owners or the dishonest use of bank, passport or other personal information.

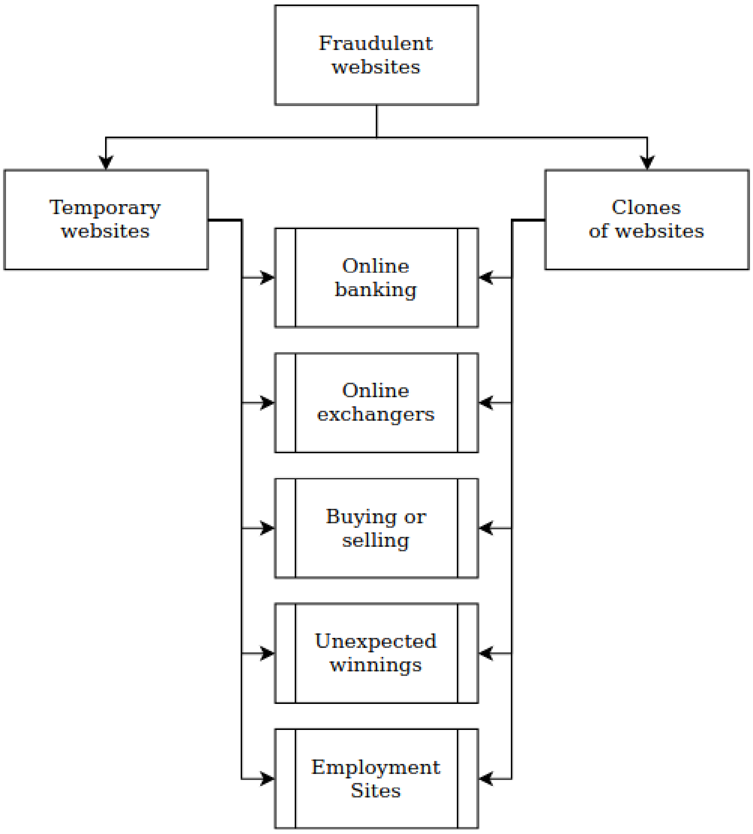

**Fig. 2.** The main directions of work of fraudulent sites

The main method of fraudulent schemes is to deceive users on the websites of giveaways or lotteries and online job offers. According to statistics compiled by Internet users, banks, their customers and the cyber police department, they account for most fraud cases. The fake online stores and study material suggestions are next in the ranking. A large share of fraudulent websites is those that offer to invest money in a business that should generate high profits (cryptocurrency, stocks and bonds, valuable materials, alternative fuel, etc.). Websites that offer poll money and have paid registration in most cases are also deceptive. In the post-Soviet space, another widely spread scheme for fraud is the offer of gambling. Quick loans on the Internet offer to pay a fee for the provision of a loan, but in the end, they do not provide any funds to their customers, receiving their money and bank data. Some sites promise users to pay compensation for medical expenses, subsidies, social benefits or tax payments. Such actions are aimed at obtaining passport and bank data for further enrichment of fraudsters. One type of online fraud is phishing. Under false pretenses invented by cybercriminals, individuals or entire organizations are forced to disclose incriminating and personal information. Often, this method of fraud is called "spearfishing" or phishing attacks. Data Breach and Incident Response (DBIR) in its report published the results of research that phishing was the main vector of attacks in 32% of all data leaks. Most often, this is in the form of emails with offers to confirm the registration of the account and a link to the site, which completely replicates the design of the original resources. This forces users to disclose personal information on their own.

One type of phishing is “farming”. This is an attack by a malicious user to retrieve personal information by slightly changing the Domain Name System (DNS) address, but the site's interface remains the same as the original. This is to ensure that the user accepts the fake site for real and no doubt enters personal information.

Thorough checking of websites will help to circumvent the schemes of fraud on the Internet.

There are several methods to identify fraudulent websites:

- self-checking;
- check in databases of fraudulent sites;
- online verification services.

Fraudulent sites can be defined according to typical criteria. Users do not need to connect their browser extensions or check sites using special services. It is possible to determine the authenticity of the site yourself if you know some indicators.[4]

*Attendance.* By the number of visitors to a website over a period, you can conclude that it is safe for the user. Many sites install meters on their pages. Also, you can check attendance yourself using plugins such as Liveinternet or Similarweb.

*Website reputation and reviews.* You can check a fraudulent site by searching the Internet (search engines) for information about it. If there are no reviews about the web resource, or they are negative, or there is no information at all — this is a sign of a fraudulent site. To check the reputation of websites, there is a popular extension of "The Web of Trust". This is a rating tool that helps you make informed decisions about whether to trust a site when searching, purchasing products, or browsing online. "The Web of Trust" is based on the approach of the so-called "crowdsourcing". Many users evaluate the site by sharing their experiences. This approach helps to prevent the risky and threatening occurrence that exist in real time online.

*Feedback.* A form on a website where users can write to site administrators and specify email. In most cases, fraudulent sites do not install such forms on their pages or respond to such hits.

*Date of creation.* The longer operates the site, the less likely fraudulent acts on it. After committing a scam, sites cease to exist, or moving to another domain. This is to ensure that users do not have time to complain. Some websites independently indicate the date of creation (placing information in the footer). You can check the date the web resource was created by entering in the search bar: whois.com/whois/ site name.

*Inaccurate, illiterate content.* Text content plays an important role in the functionality of the site. Professional websites hire content managers, copywriters, and they competently and structurally present information on the resource. Sites that are intended to deceive users are filled by amateur or scammers themselves. Therefore, the content on them contains a lot of inaccuracies and errors. Image and video content on fraudulent sites are not original but copied (for example, commercials, logos of well-known brands, etc.). Often, scammers use the logos of popular banks and payment systems to trick people and making money for fictitious services. But non-professionals do this carelessly and the user may notice the erroneous location of the graphics and video content and understand that this site is fraudulent. Frequent transitions to other domains and creation of new sites require a transfer of content, which also leads to inaccuracies and illiteracy of the content.

*The presence of analogues in social networks.* Most successful sites are featured on well-known social networks (Facebook, Instagram, Twitter). The official pages of the social networking profiles indicate the link to the corresponding site. Developers of fraudulent sites usually do not create analogue pages on social networks.

*Domain.* Unreliable sites rarely use national domains. For owner fraudulent websites It is difficult and unprofitable to pay for the short-term existence of a national domain.

*Contact Information.* In order to identify the authenticity of the site it is necessary to check the actual address and telephone numbers indicated on the contact information.

*Payment systems.* It is necessary checking the payment system options accepted by the customer. Traditionally, fraudulent sites do not have a payment card acceptance system. The most commonly used option is the payment in virtual currency, the registration of which may allow fraudulent actions, because it does not verify the authenticity of the data of the payee.

*Website address.* On fraudulent sites that aim to steal user login and password data, the layout of the interface is fully consistent with the original site. If such a site is visited by an uninitiated user, it may accept the original site and log in. Accordingly, leave in text form username and password, which can be used by fraudsters. This type of websites can be distinguished from the original at the site address where there must be a mismatch with the real site.

Specialized databases provide instant web site verification. They are filled with lists of dangerous sites and information about them (see Table 1).

*Ukrainian Interbank Payment Systems Member Association "EMA".* The fraudulent site database is the result of daily monitoring conducted by the Association's experts, as well as information from bank security services, payment services, cyber police and verified information left by users through the "Report Incident" functionality. If users encounter payment fraud, then this information can be left here, which will be reviewed, verified and used to prevent future fraud incidents. User fills out a simple form. Indicates whether there is a suspicion of misconduct, whether it has already been harmed, the type of fraud and the address of the suspected site.

**Table 1.** Databases of fraudulent websites.

| Database Name | Address (URL) |
|---|---|
| Ukrainian Interbank Payment Systems Member Association "EMA" | https://www.ema.com.ua/ [8] |
| Malware Domain List | https://www.malwaredomainlist.com/ [9] |
| Cyber police blacklist | https://cyberpolice.gov.ua/stopfraud/ [10] |
| Phishtank | https://phishtank.com/ [11] |

There are two ways to check if a site is dangerous using the EMA database: a user can view the blacklist or enter the site address in a text form and immediately retrieve the result of the check. In addition to the BlackList EMA, the site has a WhiteList — a list of tried and trusted payment services.

*Malware Domain List.* — list of fraudulent websites with the specified information: Date (UTC), Domain, IP, Reverse Lookup, Description, ASN and the country of creation of the site. There is a function of sorting by these indicators.

*Cyber police blacklist.* The project was created for the implementation of the state policy in the field of combating cybercrime, informing the population about the emergence of new cybercriminals, introducing software for the systematization of cyber accidents. On the site you can check the information by parameters: bank card number, phone or sitelink.

*Phishtank.* A free website where anyone can send, verify, track and share phishing data. Users must register on a site to report a suspected phish. A feature of the database is that it is possible to check not only sitelinks but also suspected email addresses. You can also find a lot of useful information about online phishing on the Phishtank website.

If required to enter your personal information — such as credit card number, social security number, account number or password on unfamiliar sites, they should immediately be checked for fraud. One or more popular and most effective online services can be used for verification. [13]

Online fraudulent website verification services.

- Check website status with *Google Safe Browsing* — a service developed by the Google security team to detect unsafe websites on the web and alert users and webmasters of potential harm. The service is installed and effectively protects about three million devices (computers, phones) worldwide.
- *Webmoney Advisor* — to check online exchangers and payment sites. The developers of the largest payment system have developed their own algorithm for checking

websites and pages. Real feedback is analyzed to determine if it is a reliable or fraudulent resource. To check online for fraudulent websites, you must enter a URL on the Webmoney Advisor page and wait for the review to complete. In the end, the service will provide traffic statistics, the ability to use the currency of the Webmoney system, the rating of positive and negative user estimation. You can leave your own review and evaluation of the website. This increases the likelihood of preventing unauthorized access to user data worldwide. You can connect a browser plugin that allows you to get the reputation of the site you visit, whether it accepts WebMoney and other information in one click.

- *URLVoid* — free service that allows users to scan a website to help detect malware, phishing, and fraud threats. The URLVoid has collected data on fraudulent websites over the last ten years. In total service have analyzed more than sixty million unique websites.
- *Unmask Parasites* is an easy-to-use website security service that helps expose hidden_illicit (parasites) content that hackers insert into high-quality web pages using various security gaps.
- *Virustotal* is an online service that analyze suspicious files and facilitates the rapid detection of viruses, worms, Trojans and all types of malware in real time. Allows scanning sites URLs or download files (up to 550 Mb in size) from your computer for verification. It uses more than seventy antivirus systems and automatically updates its virus databases.
- *Dr.Web LinkChecker* is a free plugin for Chrome, Firefox and Opera browsers. It automatically checks the links navigated by the user and the download files. Instantly detects modified links and checks links for scripts and frames. Also blocks ads on web pages.
- *Avast Online Security* is a browser extension for phishing protection on the Internet. Check pages that a user visits and warns of potential danger. In addition, the program issues a notification if on the Internet is tracking user activity.
- *Is It Hacked?* Performs several website checks and monitor the blacklists in real time. Check websites built in all programming languages (PHP, .NET, Java, Ruby, Scala, Erlang, GO, C#, Python etc).
- *WHOIS.* This is a resource that provides domain registration information and who is the hosting provider. This service will be useful for those users who have already been the victims of fraudulent activity on the Internet and who want to find the perpetrator. When you find out who exactly supports the site, you can write to the hosting provider and get information about the fraudster. Data will be provided subject to proof of fraud.
- *Trustorg.com* or *"Trust in the Web"* is a service designed to identify fraudulent websites and determine the level of Internet users' trust in websites. It is very easy to use the service; it is enough to enter the domain of the site in the form of verification and review the results. "Trust on the Web" shows time of existence of the domain, location of IP-address, unique rating of trust, reputation of the site. The analysis is based on the integrated use of reliable verification technologies. In particular, the Yandex Directory, Web Of Trust, Safe Browsing. Users will also be able to see the site owner's registered phone number and office address. In addi-

tion, you can register on the site of the service, which will allow you to leave comments on sites and change their level of trust.

- *Web of Trust* is an extension for web browsers, an international site verification service. Protects users from phishing, dangerous software, viruses. Issues a security warning message if a user navigates to a suspicious page. You can also view the reputation of the site yourself through the extension icon. The WoT database contains over one million sites from popular providers.

## 4 Results

After conducting the research on the identifying of fraudulent sites (Fig. 3), we can be concluded that the progress of fraudulent methods is proportional to the level of its detection.

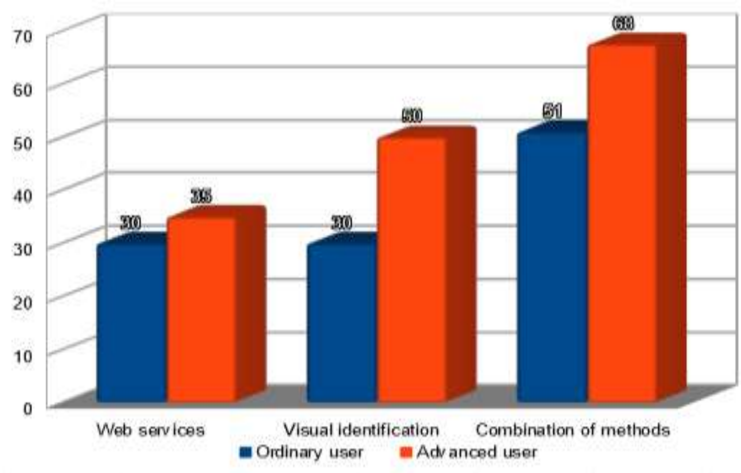

**Fig. 3.** Percentage of recognized fraudulent sites

The illegal site gets into the databases of fraudulent websites and online fraudulent website verification services very late, since the services themselves are not quick to find them, and users rarely report such cases. More often, an appeal gets to banking institutions (if an unlawful financial transaction has been committed) and to mobile operators (if a phone number has been "stolen"), which, in turn, report to the relevant authorities only after a repeated number of such appeals.

Accordingly, the method of using databases of fraudulent websites reduces the opportunity to get trapped by scammers by 30-35%. If you use the methods of visual identification and analysis of the site – this reduces the probability of 30-50%.

Such a large gap is related to that fact that not all users of the global network are able to independently evaluate the truth of the website only through the analysis of visual information and content. In general, if you use all the methods proposed, you can reduce the number of people affected by fraudulent sites by 51-67%.

## 5 Conclusion

The development of Internet technologies and the informatization of society opens great opportunities and accessibility of information, but on the other hand requires caution and protection of personal data. Every year, the opportunities of fraudsters increase in proportion to the methods of combating them. Therefore, the use of various means of site identification does not always produce 100% results. Internet users must be reasoned and guide of logic before making payment or providing personal information. Users need to learn not to make emotional and quick decisions. This is very often used by scammers "organizing promotions", which will end in a few minutes, leaving no time for reflection.

## References

1. Manek, A., Shenoy, P., Mohan, M., Venugopal, K.: Detection of fraudulent and malicious websites by analysing user reviews for online shopping websites. I. J. Knowledge and Web Intelligence 5(3): 171-189 (2016).
2. Miyamoto, D., Hazeyama, H., Kadobayashi, Y.: HumanBoost: Utilization of Users' Past Trust Decision for Identifying Fraudulent Websites. JILSA 2(4): 190-199 (2010).
3. Miyamoto, D., Miyachi, T., Taenaka, Y., Hazeyama, H.: PhishCage: reproduction of fraudulent websites in the emulated internet. SimuTools 2013: 242-247.
4. Christou, O., Pitropakis, N., Papadopoulos, P., McKeown, S., Buchanan, W.: Phishing URL detection through top-level domain analysis: A descriptive approach. ICISSP. Proceedings of the 6th International Conference on Information Systems Security and Privacy: 289-298 (2020).
5. Shao, J., Zhang, Q., Ren, Y., Li, X., Lin, T.: Why are older adults victims of fraud? Current knowledge and prospects regarding older adults' vulnerability to fraud. Journal of Elder Abuse and Neglect 31(3): 225-243 (2019).
6. Mostard, W., Zijlema, B., Wiering, M.: Combining visual and contextual information for fraudulent online store classification. IEEE/WIC/ACM International Conference on Web Intelligence: 84-90 (2019).
7. Marchal, S., Szyller, S.: Detecting organized eCommerce fraud using scalable categorical clustering. ACM International Conference Proceeding Series: 215-228 (2019).
8. Ukrainian Interbank Payment Systems Member Association "EMA" Homepage, https://www.ema.com.ua/, last accessed 2020/04/24.
9. Malware Domain List Homepage, https://www.malwaredomainlist.com/, last accessed 2020/04/24.
10. Cyber police blacklist Homepage, https://cyberpolice.gov.ua/stopfraud/, last accessed 2020/04/24.
11. Phishtank Homepage, https://phishtank.com/, last accessed 2020/04/24.
12. Fedushko, S., Ustyianovych, T., Gregus, M. Real-time high-load infrastructure transaction status output prediction using operational intelligence and big data technologies. Electronics (Switzerland). 9(4),668 (2020).
13. Tiwari, M., Gepp, A., Kumar, K.: The future of raising finance - a new opportunity to commit fraud: a review of initial coin offering (ICOs) scams. Crime, Law and Social Change 73(4): 417-441 (2020).