IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received March 7, 2022, accepted April 2, 2022, date of publication April 11, 2022, date of current version April 25, 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3166474

Phishing Classification Techniques: A Systematic Literature Review

RAHMAD ABDILLAH, ZARINA SHUKUR, MASNIZAH MOHD, AND TS. MOHD ZAMRI MURAH

Center for Cyber Security (CYBER), Faculty of Information Science and Technology, Universiti Kebangsaan Malaysia, Bangi 43600, Malaysia

Corresponding author: Rahmad Abdillah (rahmad.abdillah@uin-suska.ac.id)

This work was supported by the Ministry of Higher Education of Malaysia under Grant FRGS/1/2019/ICT01/UKM/01/2 and Grant KKP/2020/UKM/4/3.

ABSTRACT Phishing has become a serious and concerning problem within the past 10 years, with many reviews describing attack patterns and anticipating different method utilizations. This indicates that the results are still not comprehensive, subsequently leaving a critical gap in phishing reports. Therefore, this study aims to conduct a systematic review, to show a more crucial issue in phishing attacks, namely classification techniques. These issues were categorized into techniques, datasets, performance evaluation, and phishing types. The obtained results are expected to help developers prevent future phishing attacks more effectively, especially in selectively and carefully determining the techniques and evaluations to address specific types of phishing.

INDEX TERMS Classification, phishing, systematic literature review.

I. INTRODUCTION

Phishing attacks are among the most common cyberattack threats on the internet [1], due to being a technique used in obtaining sensitive data, such as bank account numbers or accessing larger computerized systems through fraudulent email or website requests (National Institute of Standards and Technology). This indicates that the attackers often perform actions similar to an entity, to steal information from members or users [2]. These attacks often focus on the requests to change identity, passwords, and other important information, using email, social media, and others. In the industrial sector, the Anti-Phishing Working Group stated that the main targets of these attacks were presently webmail, financial institutions, payments, social media, and e-commerce [3]. Phishing attacks also involves the utilization of the world's top internet services, such as Namecheap (24%), Google (16%), Public Domain Registry (PDR) (19%), NameSilo (6%), Tucows (7%), and other channels (28%). Moreover, the SSL on phishing websites is one of the attackers' mainstays to deceive their victims. One of the most effective prevention techniques for detecting these attacks is classification, which has been widely used to detect fraudulent activities on websites and emails. To improve the accuracy of detecting phishing attacks, various techniques have reportedly been developed by study researchers, such as feature selection [4]–[9] and ensemble learning [8], [10]–[14]. This led to the performance of classification technique reviews, to prevent phishing attacks. These reviews are expected to obtain more insight into the attack techniques.

Although various classification techniques are continuously emerging, they are found to still affect performance accuracy when using big and recent data [15]. For example, the classification of phishing has reportedly become a trend in previous years, although produces different performance results based on the objectives and dataset used. Therefore, a mechanism is needed to conduct a systematic analysis on the performance and variety of the techniques presently available in detecting phishing, especially the classification method. This study complements existing reports, such as [2], [16]–[18], into a systematic literature review (SLR), to focus on phishing classification techniques. According to Qabajeh *et al.* [16], the phishing prevention techniques were analyzed based on education and legal aspects, which were computerized using human-crafted and intelligent machine learning methods, respectively. This focused on the comparison of conventional and intelligent prevention techniques. To conduct SLR, a phishing development is

The associate editor coordinating the review of this manuscript and approving it for publication was Zhan Bu.

being evaluated in this present report, accompanied by the description of classification technique usages over the last 10 years. Based on the study of Akinyelu [17], phishing websites and email detection prevention techniques were analyzed, as well as the utilization of datasets as performance benchmarks. Furthermore, Gangavarapu *et al.* [18] focused on email phishing prevention techniques, by evaluating feature selection (extraction), applicability, learnability, and generalizability of several state-of-the-art machine learning. The review of this present study is also based on [2], [17], [18] with an SLR on phishing classification techniques, including emails, financial data, short messaging services (SMSs), tweets, uniform resource locators (URLs), web pages, and websites. Subsequently, more insights are provided on the use of feature selection, with the SLR answering the following questions, (1) What phishing types mostly occur for classification techniques?, (2) What data sources do phishing classification techniques mostly use?, (3) What methods are often used for phishing classification techniques?, and (4) What performance evaluations are often used for phishing classification techniques?

This study aims to provide a more comprehensive SLR while focusing on the classification techniques for phishing attacks. It is also used as a guide for developing the prevention of phishing attacks, through more accurate classification techniques. The following contains the contributions of this research:

1) The identification of more comprehensive future development opportunities, such as determining the limit value of performance evaluation, expert collaboration, as well as the exchange of data and information, for phishing detection of different languages.
2) This review focuses on performance evaluation, data sources, phishing attack types, as well as the explanation of parameter settings and validation techniques.
3) The investigation of popular phishing attacks, such as emails, financial data, SMS, tweets, URLs, web pages, and websites.

This study is subsequently organized into the following sections, (1) Section II, where the SLR is compared with previous related results, (2) Section III, where the obtained literature related to this review are thoroughly evaluated, (3) Section IV, where the basics of phishing attacks are explained from various sources, (4) Section V, where phishing is being assessed, including the technical aspects, datasets, types, accuracy performance, recommendations, and subsequent future insights, and (5) Section VI, where the conclusions are provided.

II. RELATED WORKS

Many reviews are found to comprehensively describe phishing attacks in the last 10 years, starting from the environment to technical and non-technical preventive techniques, respectively. However, not all these reviews focused on the classification techniques. This indicated that several previous studies specifically carried out a more comprehensive assessment on the utilized classification methods, performance evaluation, datasets, and phishing types, within the last ten years (2010-2020). Based on this condition, many related reviews were mainly divided into several groups, namely Twitter, SMS, Email, Website, URL, and Financial data. To add more in-depth insight, these were subsequently divided into more variables, namely Dataset, Classifier, Parameter Settings, Features, Validation method, and Evaluation metrics. The summary of the reports related to phishing reviews is shown in Table 1.

Based on Das *et al.* [2], a review of phishing URLs, websites, emails, and user studies was conducted, indicating the subsequent utilization of various parameters, namely feature, detection method, dataset, and evaluation metrics. This showed that the diversity of the dataset was evaluated for each phishing review. It also provided recommendations for dealing with phishing email issues, although did not mention the detection technique parameters used by the researchers URLs, websites, and electronic mail. This was because the parameters were indispensable for the researchers willing to perform comparisons with others. Furthermore, the study of Khonji *et al.* [19] involving a survey based on human factors, blacklists, heuristics, visual similarity, and data mining, which only focused on a variety of phishing detection techniques and solutions. Unlike Varshney *et al.* [20], the review only surveyed phishing detection techniques without a preventive solution. This indicated the sole utilization of the search engines, heuristics and machine learning, phishing black and white lists, visual similarity, DNS, proactive URLs, and mobile websites. Meanwhile, Khonji *et al.* [19] and Varshney *et al.* [20] did not describe the detection and performance evaluation techniques mostly used against these phishing attacks, leading to the development of other methods by study researchers.

According to Qabajeh *et al.* [16], a brief detection technique survey was conducted for website phishing attacks, with the analytical categories being grouped into traditional and computerized methods. This only provided limited information on existing techniques, as well as the description of the methods found in phishing detection. Therefore, several information such as performance evaluation and parameters were not stated in the analysis. This was in line with [21], which conducted a systematic review on phishing websites. Subsequently, the criteria used were datasets, features, techniques, and evaluation metrics. This indicated that the results obtained by [21] were more comprehensive than [16], although some explanations were still needed, such as the phishing detection technique parameters. Based on Gangavarapu *et al.* [18] and Almomani *et al.* [22], phishing emails were found to be the point of focus, indicating the production of methods, datasets, and evaluation metrics. This review only surveyed the techniques used in phishing emails, with the differences observed between Almomani *et al.* [22] and Gangavarapu *et al.* [18] being the provided solution. The prevention technique categories were also network-level protection, authentication,

TABLE 1. Comparison of the phishing review articles with SLRs within the last ten years.

Survey	Year's Range	Phishing type						Data set	Classifier	Parameter Setting	Feature	Validation method	Evaluation metrics
		Twitter	SMS	Email	Website	URL	Financial data						
Das *et al.* [2]	2004-2018			√	√	√		√			√		√
Gangavarapu *et al.* [18]	2005-2017			√					√				√
Akinyelu [17]	2007-2019			√	√			√	√		√		√
Qabajeh *et al.* [16]	2005-2017				√						√		
Dou *et al.* [19]	2005-2016				√				√		√		√
Aleroud and Zhou [20]	2005-2015			√	√						√		
Varshney *et al.* [21]	2004-2016				√						√		
Almomani *et al.* [22]	2004-2012			√	√						√		√
Khonji *et al.* [15]	2006-2011			√	√	√							√
This Survey	2010-2020	√	√	√	√	√	√	√	√	√	√	√	√

client-side tools, user education, server-side filters, and classifiers. This indicated that the results obtained by Almomani *et al.* [22] were very comprehensive than Gangavarapu *et al.* [18], due to the use of methods and solutions in preventing phishing attacks. However, the methods by which the parameters were utilized were not comprehensively explained for the detection techniques.

Akinyelu *et al.* [17] also conducted a review on phishing websites and emails, with the detection techniques and datasets being the only utilized categories. This indicated the performance of a brief categorical survey, where the influence of the obtained literature was described with the subsequent provision of feedback. Meanwhile, the detection technique parameters were not explained, indicating less understanding of the influence of the variables. According to Aleroud and Zhou [23], a review was conducted on phishing systems, namely environment, techniques, and countermeasures, through a very comprehensive survey. The prevention techniques also included machine learning, text mining, human users, profile matching, and other methods (ontology, honeypots, search engines, and client-server authentication). The results showed only seven phishing attack classification techniques, compared to this present study. Additionally, performance evaluations were only conducted on anti-phishing tools. Various related studies also described the techniques used to detect phishing attacks, although had several undisclosed issues, such as (1) the methods by which the dataset was distributed against the phishing attacks, (2) the use of popular techniques for phishing types, (3) the use of phishing type evaluations, and (3) the use of parameters for phishing classification techniques. For example, the utilization of many parameters was not described by the researchers, leading to unbalanced comparisons with other studies. Therefore, the information on parameters is crucial to researchers during comparative analyses.

III. METHODOLOGY

This describes the SLR method, questions (Qs), search strategy, study selection, data extraction, quality assessment, and data synthesis.

A. SLR METHOD

Based on Fig.1, the SLR had nine stages, namely (i) identifying SLR needs, (ii) building a code of conduct, (iii) evaluating the code of conduct, (iv) searching for related reviews, (v) selecting related reviews, (vi) obtaining the information related to SLR, (vii) evaluating the quality of related studies, (viii) combining the results, and (ix) describing the SLR results.

B. STUDY QUESTIONS

The questions raised in this study are as follows,

Q1, What phishing types mostly occur for phishing classification techniques?

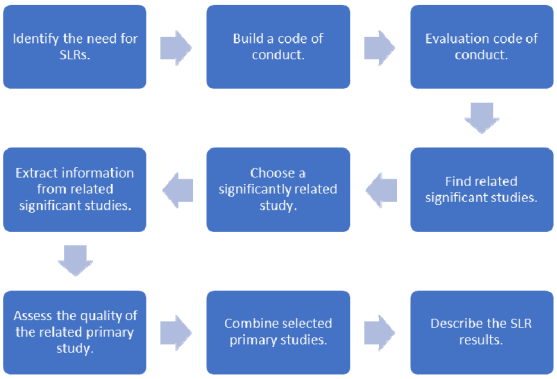

FIGURE 1. Systematic literature review step.

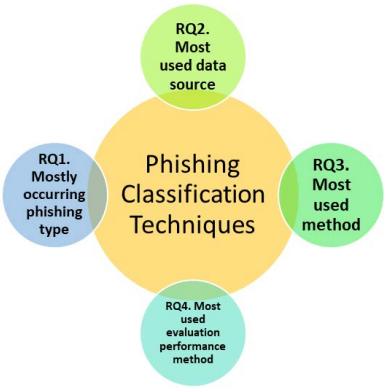

FIGURE 2. SLR on phishing classification techniques.

Q2, What data sources are mostly used for phishing classification techniques?

Q3, What methods are often used for phishing classification techniques?

Q4, What performance evaluations are often used for phishing classification techniques?

C. SEARCH STRATEGY

After identifying the study questions (Fig. 2), several queries and a journal database were selected and evaluated. This utilized database was obtained from many quality publishers, such as IEEE, Springer, ACM, Wiley, and Emerald. Moreover, the queries were defined based on the predefined questions, where phishing and classification descriptions were utilized. These queries were subsequently applied to the researchers titles, abstracts, and keywords. The parameters for the publication year, document type, and article category were also added in this process. This indicated that the required document type should be reviewed in the computer science category. For the results to be closer to the completion of the study questions, the search should focus on the documents conducted between January 2010 to December 2020.

D. STUDY SELECTION

Based on Table 2, the inclusion criteria for determining related studies included phishing topics, as well as the comparison and application of classification techniques. The articles outside the inclusion criteria were also separated from the priority SLR journals.

TABLE 2. Criteria of inclusion and exclusion.

Criteria of Inclusion	Criteria of Exclusion
Comparative research of phishing classification techniques.	Studies unrelated to phishing
Phishing classification techniques on various datasets.	Studies unrelated to the classification
It covers a wide variety of phishing topics.	Studies that unfulfilled any inclusion criteria
Priority articles from journals.	

TABLE 3. Mapping information that has been extracted to RQ.

Extracted information	RQ
Data Source	RQ2
Classification Techniques	RQ3
Performance evaluation	RQ4
Phishing type	RQ1

Based on the search process, "phishing" provided 1,669 articles in TOPIC, leading to the rearrangement of the keywords. To subsequently produce phishing articles, the keyword was also used in the TITLE, leading to the production of 225 reviews. Moreover, the keywords were continuously rearranged for more specifications in the TITLE (phishing) AND TOPIC (classifi*), leading to the production of 86 articles. The reviews irrelevant to computer science were then discarded, leading to the selection of 68 studies of phishing attack classification techniques.

E. DATA EXTRACTION

After the search and selection processes (Fig. 3), all related reviews were extracted to obtain information on the completion of the study questions. These data were obtained and used to map the selected articles, with adjustments conducted to the extracted information against the questions, as shown in Table 3.

F. STUDY QUALITY ASSESSMENT AND DATA SYNTHESIS

Based on this stage, the transformed articles were interpreted to answer the SLRs, where several graphic models were used to facilitate translation. In addition, the produced interpretations were found to be quantitative and qualitative.

IV. BACKGROUND

Phishing is one of the most dangerous cyberattacks capitalizing on human weaknesses, through the leveraging of social engineering and technology collaborations [24]. This indicates the loss of confidential information to hackers, through emails, SMS, social media including Twitter, Facebook, and Google+, or a web browser pop-up [25]–[27].

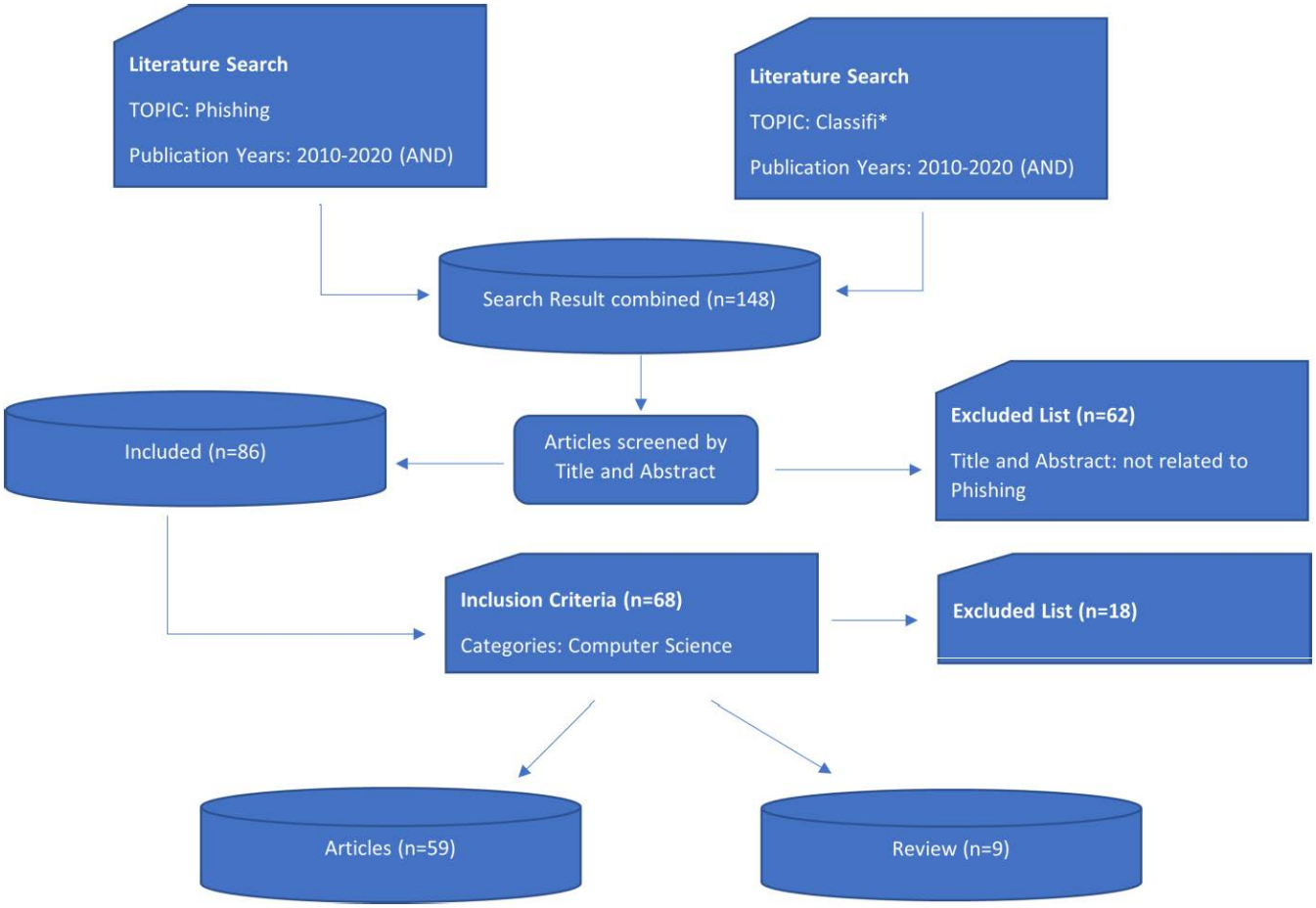

FIGURE 3. The search and selection of primary studies.

It is also found in almost all web pages, including auction and online payment websites [24]. Based on this condition, are known to capitalize on human weaknesses by sending emails or text messages containing gifts/security alerts from an organization [26]. This is often performed to direct the action of users according to the desired wishes. Therefore, a phishing attack aims to capitalize on the trust of users interacting with an institution believed to be safe and legitimate [24]. This indicates that phishing is encountered when a user obtains an illegal link in their email, with the subsequent response being often influenced by following the pop-up directions. These techniques are commonly found on computers and mobile platforms [7].

The technique is also divided into three groups based on the attack target [26]: namely general, spear, and whale phishing. General phishing is often massively carried out with hackers just throwing baits without using maximum effort, indicating that the chances of success are meagre [26]. This type of attack is found to only trap careless users. Furthermore, the spear-phishing attack targets a specific group of people with an essential organizational role, as hackers only need minimal effort to locate victims, such as using social scams. This indicates that the hackers constantly change methods, although use similar objectives with the failure of their attacks. The chance of success in this attack is found to be better, compared to general phishing. Meanwhile, the whale phishing attack has a target on organizational CEOs or political party leaders. This indicates that hackers do their best to profile victims and modify emails, as well as engage in various exploits to expose specific vulnerabilities. The difference between this attack and spear phishing is only based on the impacts achieved and performance efforts.

The disadvantages of phishing attacks are found to be very fatal, including the losses in financial institutions are observed as a reputation failure, due to the customers becoming insecure with the safety of transactions [24]. Meanwhile, finances are often disturbed based on users being unable to reaccess their financial accounts, such as through the illegal use of credit cards by hackers [28]. According to this condition, a phishing attack survey was conducted regarding the occurrences observed over the past 10 years. This indicated that phishing attack variations always occur yearly. Based on Fig. 4, the variation of these attacks was observed to increase yearly, with most of the occurrences found in 2019 and 2020, through websites, webpages, emails,

URLs, SMS, and tweets. Subsequently, the phishing attack types that occurred over the past 10 years were described.

A. WEBSITE

Hackers often create a replica of the original website, for the full interaction of victims [29]. This is then accompanied by the transfer of the website through various media, according to the phishing target, such as emails, SMS, social media, and browser pop-ups [30]. Subsequently, hackers capitalize on human weaknesses by leading unknowing victims to the replica website, with the instructions to complete a validation request file/credential renewal and financial information [29]. When the user follows these instructions, accessing financial accounts become impossible with the immediate loss of money [31].

According to Abbasi *et al.* [32], two types of phishing websites were found to presently exist, namely concocted and spoof sites, respectively. This indicated that a concocted site replicates a legitimate website for commercial purposes, to conduct sale/purchase transactions or fraud. These sites are for buying and selling transactions between hackers and victims, based on the acceptance of money without providing the purchased product. This phishing technique often uses social engineering to obtain its victims. For example, a hacker creates a shop or an account at an e-commerce service provider, although performing transactions at the concocted sites. In this condition, hackers always use various excuses for victims to believe in carrying out transactions on the concocted site. This is not in line with the spoof sites, which only creates similar replica websites as the original, including web domains and content. When consumers select a replica website to log in, the user credentials are stolen by hackers and then used to assess the original platform for financial gain [33].

B. WEBPAGES

Phishing webpages manipulate textual forms or the appearance of legitimate websites, including the URL structure [5], [34]. This indicates that the ability of hackers to imitate legitimate websites is likely to deceive victims due to a lack of phishing knowledge. Lost personal information is also likely to lead to identity theft and loss of large amounts of money. Moreover, other forms of webpage phishing are found to exist, such as exploiting vulnerabilities to enter a legitimate website. This leads to the hacked website automatically having similar capabilities, such as the domain and appearance of legitimate sites [34].

Some users often place their trust in a website's safety from phishing attacks, based on the green padlock icon in the browser address bar. In recent years, almost all legitimate or phishing websites have reportedly used HTTPS, leading to the inability to become the standard against prevention [30]. Using a small amount of JavaScript technology, hackers also create green padlock icons and fake HTTPS in the browser address bar [35]. According to the Anti-Phishing Working Group [3], hackers used domain-validated to enable the SSL feature on phishing sites, due to being the weakest form of certificate validation.

C. EMAIL

Email phishing is a present problem so difficult to solve [36], as spam are limited to legitimate marketing emails with the occurrence of other types of phishing emails [37]. For example, a hacker often sends a phishing email to the victim as an essential or influential person in an organization, to obtain important information. This problem is increasingly developing in the big data era, with hackers assessing the profiles of victims, such as name, gender, contact information, and daily activities [26]. This indicates that hackers use fake emails to trick victims into providing confidential information. For example, a phishing victim is often directed to access a replicated bank website, with the instructions to provide an account number or credit card [37].

Besides causing the loss of important information, email phishing is also a means of accelerating the spread of malware [38]. When the email contains links leading to dangerous websites, malware is often being unknowingly clicked, leading to a fatal impact on the user [39]. Based on Sur [12], the message categories in phishing emails are as follows,

- Authority
  An email from a law enforcement agency or authoritative institution, to regulate social life.
- Commitment
  An email from an organization or community group acting on humanitarian concerns, such as raising funds for natural disaster victims.
- Liking
  An email from people sharing similar fates, leading to users leaving the safe zone. Hackers often use this model to obtain organizational information through their members.
- Perceptual Contrast
  This email capitalizes on the victim's lack of knowledge in profit utilization information.
- Reciprocation
  This email capitalizes on the concept of reciprocity among humans (give and take).
- Scarcity
  These emails capitalize on human feelings, such as being provided with a short-term profit, which accordingly causes anxiety or loss when not immediately performed.
- Social proof
  This email originates from a trusted partner or neighbour. When users receive this message, they often become surer of the email's information due to being sent by a colleague or neighbour.

D. URL

Hackers are becoming constantly innovative through the creation of phishing URLs and various methods of obtaining

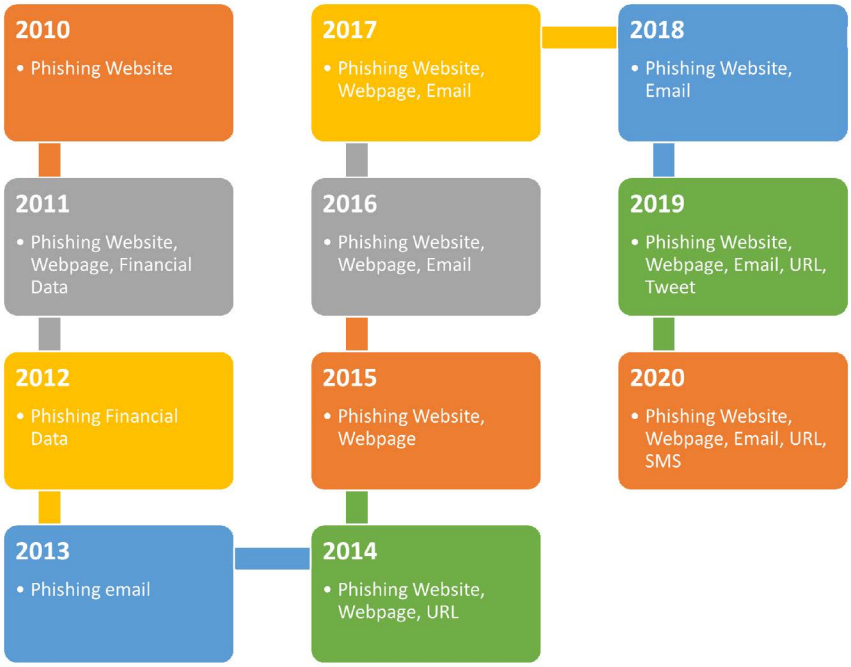

FIGURE 4. Different types of phishing within the years.

TABLE 4. Number of papers.

Source	Latest Quartile	Number of Papers
IEEE COMMUNICATIONS SURVEYS AND TUTORIALS	Q1	4
COMPUTER SCIENCE REVIEW	Q1	1
IEEE TRANSACTIONS ON DEPENDABLE AND SECURE COMPUTING	Q1	1
FUTURE GENERATION COMPUTER SYSTEMS-THE INTERNATIONAL JOURNAL OF ESCIENCE	Q1	1
INFORMATION SCIENCES	Q1	1
ARTIFICIAL INTELLIGENCE REVIEW	Q1	1
JOURNAL OF NETWORK AND COMPUTER APPLICATIONS	Q1	1
APPLIED SOFT COMPUTING	Q1	3
EXPERT SYSTEMS WITH APPLICATIONS	Q1	4
INFORMATION & MANAGEMENT	Q1	1
NEURAL COMPUTING & APPLICATIONS	Q1	2
DECISION SUPPORT SYSTEMS	Q1	2
JOURNAL OF AMBIENT INTELLIGENCE AND HUMANIZED COMPUTING	Q1	3
JOURNAL OF MANAGEMENT INFORMATION SYSTEMS	Q1	1
IEEE TRANSACTIONS ON NEURAL NETWORKS	Q1	1
COMPUTERS & SECURITY	Q1	7
IEEE TRANSACTIONS ON NETWORK AND SERVICE MANAGEMENT	Q2	1
IEEE ACCESS	Q2	10
CLUSTER COMPUTING-THE JOURNAL OF NETWORKS SOFTWARE TOOLS AND APPLICATIONS	Q2	1
SOFT COMPUTING	Q2	1
WORLD WIDE WEB-INTERNET AND WEB INFORMATION SYSTEMS	Q2	1
ACM TRANSACTIONS ON INFORMATION AND SYSTEM SECURITY	Q2	2
COMPUTER COMMUNICATIONS	Q2	1
PERVASIVE AND MOBILE COMPUTING	Q2	1
INFORMATICA	Q3	1
JOURNAL OF EXPERIMENTAL & THEORETICAL ARTIFICIAL INTELLIGENCE	Q3	1
KYBERNETES	Q3	1
CONCURRENCY AND COMPUTATION-PRACTICE & EXPERIENCE	Q3	1
SECURITY AND COMMUNICATION NETWORKS	Q4	3
IET INFORMATION SECURITY	Q4	2
DATA TECHNOLOGIES AND APPLICATIONS	Q4	1
INTERNATIONAL JOURNAL ON ARTIFICIAL INTELLIGENCE TOOLS	Q4	2
INTERNATIONAL ARAB JOURNAL OF INFORMATION TECHNOLOGY	Q4	2
IEICE TRANSACTIONS ON INFORMATION AND SYSTEMS	Q4	2

TABLE 5. Phishing type in years.

Year	N	Phishing type						
		Web page	Website	URL	Email	SMS	Tweet	Financial data
2020	16	3	8	2	2	1	0	0
2019	15	2	7	4	1	0	1	0
2018	5	0	1	0	4	0	0	0
2017	3	1	1	0	1	0	0	0
2016	4	2	1	0	1	0	0	0
2015	2	1	1	0	0	0	0	0
2014	6	3	2	1	0	0	0	0
2013	3	0	0	0	3	0	0	0
2012	1	0	0	0	0	0	0	1
2011	3	1	1	0	0	0	0	1
2010	1	0	1	0	0	0	0	0

TABLE 6. Evaluate the performance of the classification technique based on the type of phishing.

Performance evaluation	Phishing type						
	Email	Financial data	SMS	Tweet	URL	Web page	Website
Accuracy	√	√	√	√	√	√	√
AUC	√					√	
CCR						√	√
F1-Macro							√
FAR						√	
F-Measure	√		√		√	√	√
FNR	√				√	√	√
FPR	√				√	√	√
GM							√
MCC						√	
Precision	√		√	√	√	√	√
PER							√
RT							√
ROC	√					√	
SW-Test							√
TNR	√				√	√	√
TPR	√		√	√	√	√	√
W T-Test							√

AUC = Area Under the Curve, CCR = Correct Classification Rate, FAR = False Alarm Ratio, FNR = False Negative Rate, FPR = False Positive Rate, GM = Geometric Mean, MCC = Matthew's Correlation Coefficient, PER = Prediction error rate, RT = Ranking Techniques, ROC = Receiver Operating Characteristics, SW Test = Shapiro–Wilk Test, W T-Test = Welch's T-Test.

confidential information [40]. The characteristic of this URL is based on the replication of a legitimate address, which then redirects to a website being modified for phishing [10]. The chances of being exposed to phishing URLs are high for those not comprehensively aware. This indicates that the awareness of URL phishing should be increased by identifying similar website appearance and the actual address [41]. However, present URL attacks often lead to other phishing websites, avoiding the detection of users.

According to Volkamer *et al.* [42], there were several reasons people always fall victim to URL phishing, such as,

- The awareness of phishing URLs is lacking, leading to inappropriate decisions.
- Reliable URLs are often unknown when written in the email, as well as the browser address and status bars, respectively.
- The final destination of the URL is unknown, due to being redirected or using tiny addresses.
- URLs are not carefully checked before or accidentally being clicked, due to not knowing a phishing address.

E. FINANCIAL DATA

Financial data is known to be of interest to some phishing researchers, although not all have the desire to explore this specific data. This indicates that the data has reportedly attracted the attention of several researchers, such as [43] and [44], due to being associated with phishing attacks. In this data, phishing attacks often determines a correlation between textual descriptions, as organizational problems reportedly have financially significant impacts. Based on Nishanth *et al.* [43] and Chen *et al.* [44] phishing attacks were detected on the financial department through fiscal data. This indicated that the Decision Trees (DT), support vector machines (SVM), and multilayer perceptron (MLP) classification techniques were used to obtain phishing risk scores [43]. Furthermore, Nishanth *et al.* [43] developed the study of Chen *et al.* [44], which focused on the use of data imputation techniques. The results showed that this technique replaced the value of missing financial data, based on soft computing [43]. Nishanth *et al.* [43] and Chen *et al.* [44] also used financial statements and textual data to derive three levels of phishing attacks.

F. TWEET

Twitter is one of the media intermediaries for hackers to quickly and undetectably spread phishing across networks [45], due to using short content and tiny URLs. This medium does not need technique, as conveyed by [12], with the simple concept of Twitter, Follower, and Following indicating that users are especially interested in hackers. This indicates that hackers have always used these concepts to launch phishing attacks. Based on this feature, very little study has been observed, as Liew *et al.* [45] detected an attack on Twitter using the Random forest classification technique. This technique was used on the datasets obtained through a crawl of Twitter, indicating the classification of phishing attacks with 94.64% precision and 95.49% recall. However, the proposed warning mechanism against these attacks was approximately detected at 97.50% security alerts, for real-time phishing detection.

TABLE 7. The list of primary studies in phishing classification.

Publication year	Studies	Publication	Dataset	Method	Phishing type
2020	[35]	COMPUTERS & SECURITY	private and public	C4.5	Web page
2020	[70]	JOURNAL OF AMBIENT INTELLIGENCE AND HUMANIZED COMPUTING	private and public	Random Forest	website
2020	[71]	COMPUTERS & SECURITY	public	SVM	Web page and email
2020	[13]	INFORMATICA	public	Multilayer Perceptron	Website
2020	[9]	APPLIED SOFT COMPUTING	public	NN	website
2020	[72]	COMPUTER COMMUNICATIONS	private	SVM	URL
2020	[59]	INTERNATIONAL ARAB JOURNAL OF INFORMATION TECHNOLOGY	public	Intelligent association classification	website
2020	[62]	NEURAL COMPUTING & APPLICATIONS	public	Twin Support Vector Machine (TWSVM)	website
2020	[73]	IEEE ACCESS	public	Extreme Gradient Boosting (XGBoost)	email
2020	[29]	IEEE ACCESS	public	Random Forest	website
2020	[64]	IEEE ACCESS	public	LogitBoost-Extra Tree (LBET)	website
2020	[10]	IEEE ACCESS	public	SVM	URL
2020	[34]	IEEE ACCESS	private and public	spwalk	Web page
2020	[25]	DATA TECHNOLOGIES AND APPLICATIONS	public	Deep Belief Neural Networks	Email
2020	[74]	IEEE ACCESS	private and public	Logistic Regression	website and email
2020	[46]	FUTURE GENERATION COMPUTER SYSTEMS-THE INTERNATIONAL JOURNAL OF ESCIENCE	public	Naïve Bayes	SMS
2019	[61]	JOURNAL OF AMBIENT INTELLIGENCE AND HUMANIZED COMPUTING	public	Logistic Regression	Web page
2019	[30]	NEURAL COMPUTING & APPLICATIONS	public	Random Forest	website
2019	[75]	PERVASIVE AND MOBILE COMPUTING	public	SVM	URL
2019	[8]	INFORMATION SCIENCES	private and public	Random Forest	website
2019	[41]	EXPERT SYSTEMS WITH APPLICATIONS	public	Random Forest	URL
2019	[40]	JOURNAL OF AMBIENT INTELLIGENCE AND HUMANIZED COMPUTING	public	kNN	URL
2019	[60]	SOFT COMPUTING	public	Nonlinear Regression	website
2019	[37]	IEEE ACCESS	public	RCNN	email
2019	[45]	COMPUTERS & SECURITY	private and public	Random Forest	tweet
2019	[69]	IEEE ACCESS	public	CNN-LSTM	website
2019	[65]	INTERNATIONAL JOURNAL ON ARTIFICIAL INTELLIGENCE TOOLS	public	Firefly Algorithm	website
2019	[26]	COMPUTERS & SECURITY	public	Logistic Regression	Web page
2019	[6]	IET INFORMATION SECURITY	public	Deep Neural Network	website
2019	[7]	IEEE ACCESS	public	NN	website
2019	[14]	IEICE TRANSACTIONS ON INFORMATION AND SYSTEMS	private and public	C4.5	URL
2018	[68]	IEEE TRANSACTIONS ON DEPENDABLE AND SECURE COMPUTING	private	RUSBoost	email
2018	[76]	DECISION SUPPORT SYSTEMS	public	NN	email
2018	[12]	JOURNAL OF EXPERIMENTAL & THEORETICAL ARTIFICIAL INTELLIGENCE	private	Deep Neural Network	email
2018	[77]	INTERNATIONAL ARAB JOURNAL OF INFORMATION TECHNOLOGY	public	Random Forest	website
2018	[78]	IEEE ACCESS	public	NN	email
2017	[4]	SECURITY AND COMMUNICATION NETWORKS	private and public	C4.5	website
2017	[11]	WORLD WIDE WEB-INTERNET AND WEB INFORMATION SYSTEMS	private and public	Extreme Learning Machine (ELM)	Web page
2017	[67]	CONCURRENCY AND COMPUTATION-PRACTICE & EXPERIENCE	private	Random Forest	email
2016	[24]	APPLIED SOFT COMPUTING	public	Fast Associative Classification Association Algorithm (FACA)	Website
2016	[5]	IEICE TRANSACTIONS ON INFORMATION AND SYSTEMS	public	Logistic Regression	Web page
2016	[79]	EXPERT SYSTEMS WITH APPLICATIONS	public	SVM	Web

TABLE 7. *(Continued.)* **The list of primary studies in phishing classification.**

2016	[39]	KYBERNETES	public	Firefly Algorithm (FFA) + SVM	page Email
2015	[32]	JOURNAL OF MANAGEMENT INFORMATION SYSTEMS	private and public	Genre Tree Kernel	website
2015	[80]	SECURITY AND COMMUNICATION NETWORKS	private and public	Bagging	Web page
2014	[28]	CLUSTER COMPUTING-THE JOURNAL OF NETWORKS SOFTWARE TOOLS AND APPLICATIONS	private and public	SVM	Web page
2014	[81]	COMPUTERS & SECURITY	private and public	SVM	Web page
2014	[82]	IEEE TRANSACTIONS ON NETWORK AND SERVICE MANAGEMENT	public	Random Forest	URL
2014	[58]	ACM TRANSACTIONS ON INFORMATION AND SYSTEM SECURITY	private and public	Logistic regression	Web page
2014	[31]	IET INFORMATION SECURITY	private and public	C4.5	website
2014	[33]	INFORMATION & MANAGEMENT	public	SVM	website
2013	[66]	COMPUTERS & SECURITY	public	AdaBoost	email
2013	[36]	JOURNAL OF NETWORK AND COMPUTER APPLICATIONS	public	SVM+AdaBoost+Naïve Bayes	Email
2013	[27]	APPLIED SOFT COMPUTING	private	SVM	email
2012	[43]	EXPERT SYSTEMS WITH APPLICATIONS	private	Decision trees (DT)	Financial data
2011	[44]	DECISION SUPPORT SYSTEMS	private	Decision trees (DT)	Financial data
2011	[83]	ACM TRANSACTIONS ON INFORMATION AND SYSTEM SECURITY	private and public	Bayesian Network	Website
2011	[84]	IEEE TRANSACTIONS ON NEURAL NETWORKS	private	Bayesian Network	Web page
2010	[57]	EXPERT SYSTEMS WITH APPLICATIONS	private and public	P.A.R.T.	Website

G. SMS

Phishing attacks on telecom areas, namely SMS (Short Message Service), has reportedly reached users through the emergence of technology, with the provided convenience leading to the fall of victims. This indicates that phishing SMS (Smishing) is almost similar to email attacks, which involves stealing the victim's credentials. In this process, hackers often send messages to victims, containing a phone number for subsequent transactions or a URL with the directives to a malicious website [46]. These websites are similarly designed to the original platform, by imitating all the actual source code. When using Smishing, hackers often disguise themselves as trustworthy people, companies, or government organizations [47]. This indicates that more utilization of SMS leads to higher smishing [25]. Furthermore, many organizations use Blacklisting techniques against suspicious URLs, although the method is effortless for anticipation through a shortener service. This indicates the toughness in determining the safety and danger of a URL when hackers use the shortener service [47]. A similar modification of URLs to the original website address is also being performed by the hackers, such as misplaced or misspelt characters, e.g., gooogle.com or facebok.com.

The Smishing trend in the last five years (2016-2020) was found to increase yearly, with the highest observed at 241342 reports in 2020 [48]. It also contributed to the spread of the Banking Trojan [49], due to being used to avoid the screening process mechanism carried out by Google Play. This indicated that McAfee Mobile Security found a significant increase of 141% in Q3 and Q4 within 2020 [49]. Therefore, bank customers are still the target of phishing attacks through telecommunications media [50]. Smishing also retrieves the information stored on smartphones when the user clicks on a malware-based URL [47], [51]. This information includes contacts, notes, financial information, pictures, etc. This trend subsequently steals personal information such as security cards, photo identification, or accreditation certificates for other crimes [52]. Various forms of text messages are also found to be compromised by malware, such as the provision of coupons to wedding invitations. This indicates that mobile phishing is very difficult to detect than that of the email, due to the tiny message size [50].

The opportunity for smishing attacks on mobile devices is reportedly enormous, due to the higher interaction of people with mobile devices, compared to laptops or computers [53]. For personal and official purposes, SMS is commonly utilized, with many popular organizations adopting it as a source of communication with their customers. Most organizations also use SMS to send information, promotions,

and surveys [54], with study researchers utilizing various methods to prevent smishing attacks, such as the Blacklist technique. This method prohibits SMS from the phone numbers included in the Blacklist category, although still has a weakness with hackers performing Smishing from another mobile contact [53]. Using machine learning collaboration and feature extraction, the concept of smishing detection often identify various sources of telephone numbers in SMS [53]. Smishing messages have also become a part of Spam texts, although have a similar goal in obtaining users' personal and financial information [46], [50]. Based on this condition, various researchers had difficulties in obtaining the Smishing dataset to be used in comparing their proposed detection technique [50]. This indicated that the use of language translation techniques was an option to obtain new datasets for study researchers in specific countries. According to Rastenis *et al.*, an English phishing email dataset was used and subsequently translated into Russian and Lithuanian [55]. This indicated that the Google Translate service was used to interpret the phishing email dataset into Russian and Lithuanian. Based on Wu *et al.* [56], mobile device users were exposed to phishing attacks due to the following,

- Limitations of mobile device capabilities, such as the screen size and the ability to perform computations.
- User habits, such as leisure activities, enables more convenience to click on a URL than typing to check. This is because the use of a virtual keyboard is not as comfortable as using a physical device.

V. RESULT AND DISCUSSION

Phishing attacks have reportedly led to catastrophic losses, with various detection and prevention attempts being carried out by several researchers. Based on this condition, SLRs were conducted against phishing attacks as the classification technique for the last 10 years. Several reviews were also obtained from the best journal ranking in the computer science category. According to Table 4, the literature obtained were 50, 26, 6, and 18% of Q1, Q2, Q3, and Q4 Web of Science (WoS) articles, respectively. The most dominant source journals were also the IEEE ACCESS (15%), COMPUTERS & SECURITY (10%), EXPERT SYSTEMS WITH APPLICATIONS (6%), as well as IEEE COMMUNICATIONS SURVEYS AND TUTORIALS (6%).

Phishing study was found to be increasing yearly, especially for classification techniques. Based on Fig. 5, the highest increase was observed in 2019 (23%) and 2020 (26%), with an annual rate of 22%. This indicated that there was a significant increase in the yearly study of phishing.

A. MOST OCCUR PHISHING TYPE

The distribution of phishing type is shown in Fig. 6, where the higher dominance was found in the website (39%), webpage (22%), email (20%), URL (12%), and others (7%). This indicated the relevance to present phishing incidents, where a collaboration between emails or other social media increased the types of attacks, such as URLs, webpages, and websites.

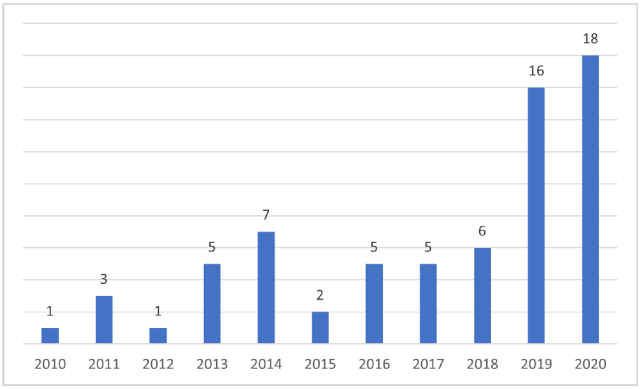

FIGURE 5. Average numbers of relevant publications in ten years.

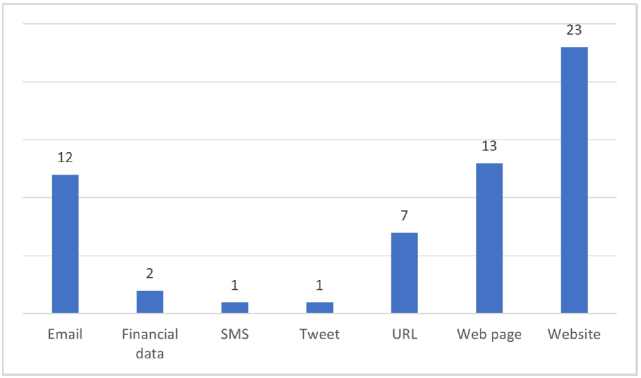

FIGURE 6. Phishing type.

The yearly distribution of each type of phishing is shown in Table 5, where most studies were observed in 2020 (16 articles). During this period, the distribution of articles was based on phishing types, namely webpage (3), website (8), URL (2), email (2), and SMS (1). However, the distributions focused on the webpage (2), website (7), URL (4), email (1), and tweet (1). This subsequently indicated that only a few studies were conducted in 2010 and 2012. According to Table 5, phishing studies were increasing, especially in 2019 and 2020, where dominance was found within websites at 7 and 8, respectively. This indicated that there were yearly studies on websites, webpages, and emails, showing that researchers were highly focused on solving phishing problems.

B. MOST USED DATA SOURCE

In this study, the dataset was used as a medium to test the performance of the researchers proposed method. Based on Fig. 7, the most widely used dataset was the PhishTank (34 authors), Alexa (15 authors), and UCI Machine learning (12 authors). This was because the PhishTank dataset provided publicly suspicious phishing URL information [57], with Alexa being a Web analytic obtaining data from clients through an installed toolbar [58]. Meanwhile, the UCI Machine learning was mostly used to conduct benchmarks, especially the phishing studies [8]. A total of 13 authors also used the datasets obtained from their institutions, such

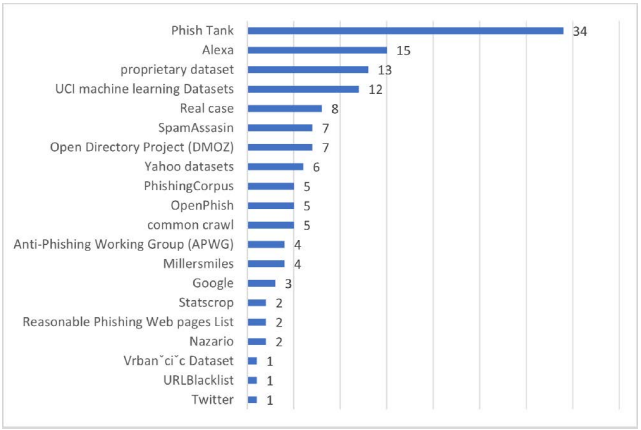

FIGURE 7. The distribution of data sources is based on the most widely used.

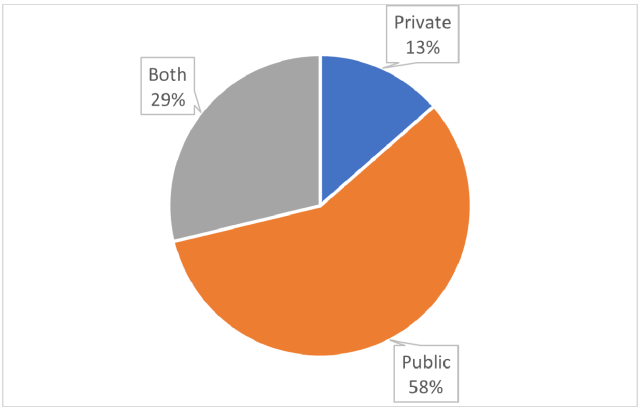

FIGURE 8. Total distribution across various data source types.

as email servers, spam filters, honeypots, financial data, and common-crawl websites. Besides that, other researchers utilized the datasets according to their problems, such as 3Sharp, DeepPhish, Enron, Phishload, starting point directory, and Stuffgate Free Online Website Analyzer.

Based on phishing, the use of datasets was divided into public, private, and blended (public and private) groups, respectively. According to Fig. 8, the uses of public, private, and blended datasets were 58, 13, and 29%, respectively. This indicated that the researchers simultaneously used the public and private datasets, to ensure that the proposed method's performance remained superior. However, the use of private datasets was constrained when compared with the proposed methods.

Based on Fig. 9, the use of public datasets began to increase from 2016 to 2020 (30 articles), indicating that the utilization of this information remained a reference for phishing studies to benchmark the proposed methods. The uses of both private and public datasets were also found to be increasing, subsequently providing new phishing insights. However, it remained an obstacle when compared to the proposed method's performance, due to the collision with a private dataset.

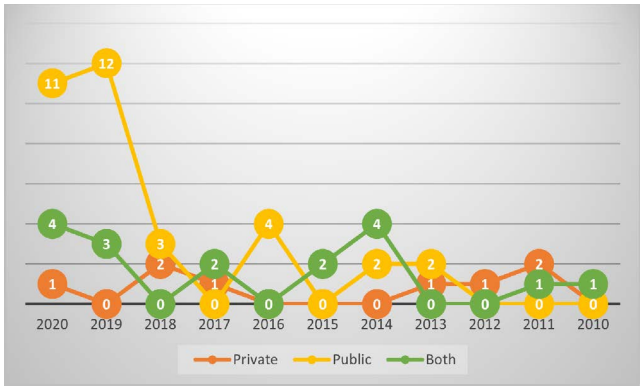

FIGURE 9. Distribution of data sources in years.

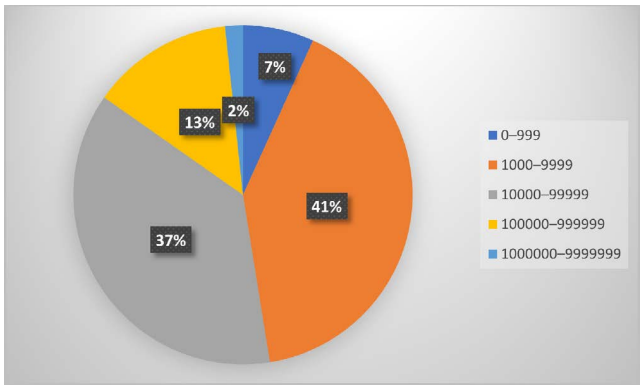

FIGURE 10. Data source size distribution.

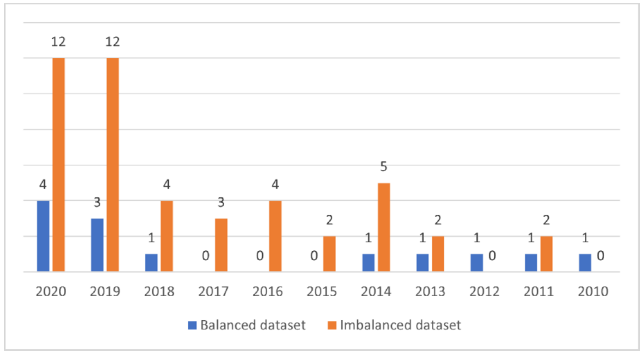

FIGURE 11. Data source position.

According to Fig. 10, the most widely used dataset size ranged between 1000–9999 (41%), accompanied by 10000–99999 (37%) and 100000–999999 (13%). This indicated that using the number of datasets was a consideration to test the method's performance.

The use of balanced and imbalanced datasets was also surveyed in this report. This indicated that an imbalanced dataset occurred due to the uneven distribution of data classes. Based on Fig. 11, 13 and 46 researchers (23 and 78%) used both balanced and imbalanced datasets, respectively. This showed that the increase in the use of imbalanced datasets began in 2011 and peaked within 2019 and 2020, although did not significantly increase.

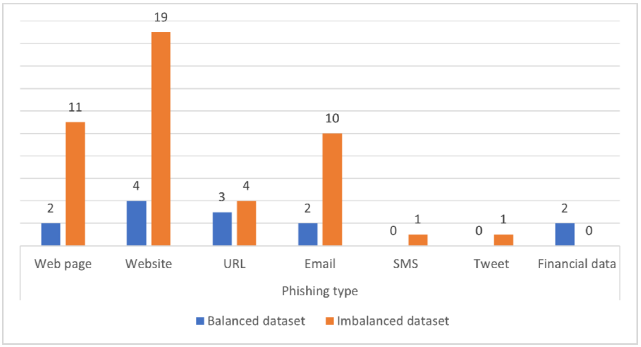

FIGURE 12. Distribution of balanced and imbalanced data sources used on phishing types.

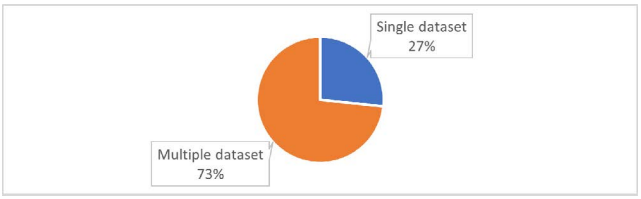

FIGURE 13. Use of data sources.

In Fig. 12, the use of imbalanced datasets was more dominant than the balanced datasets, due to being commonly used on phishing websites (41%), webpages (24%), and emails (22%). However, the use of balanced datasets was only used on websites (31%), webpages (15%), and URLs (23%).

Several researchers also used one dataset, with others using various types such as Alexa and PhishTank, to achieve the required standards. For these researchers, the use of a dataset depended on the standards being set. From Fig. 13, 73% and 27% of researchers used multiple and one dataset collaborations, respectively.

The utilization of dataset features was also surveyed, due to being an essential element in detecting phishing attacks [5]. For researchers, the evaluation of feature use affected the classification performance [12], [13], [59]. This indicated the elimination of ineffective features on attack detection [9], subsequently reducing the training data processing time [60]. Based on Fig. 14, the feature evaluation was carried out by 58 articles, containing 33%, 2%, and 65% ordinary, FVV index, and cross-validation assessments, respectively. Some researchers also added numerous mechanisms to assess the feature evaluation models, such as the FVV index and cross-validation. This indicated that the FVV index was created by Zhu *et al.* [7] to evaluate sensitive features' impact, subsequently showing that the variable conquered overfitting, especially in the neural network classification. Meanwhile, cross-validation was the most widely used technique in machine learning, especially phishing attack detection techniques. The variable also overcame overfitting in machine learning [4], [24]. In addition, performance evaluation was conducted using majority cross-validation in accuracy, TPR, precision, and f-measure.

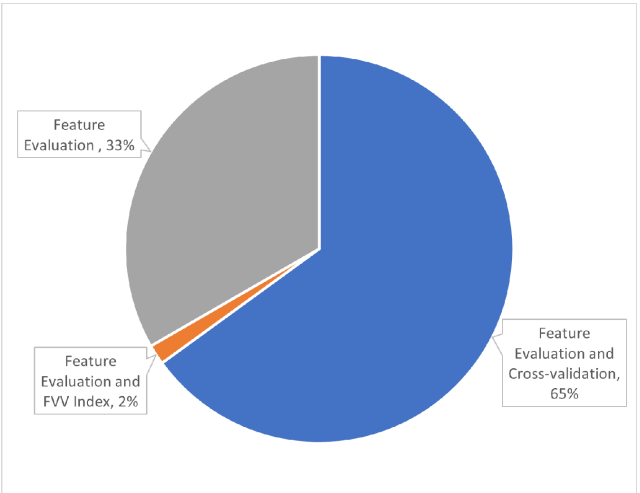

FIGURE 14. Distribution of use cross-validation and feature evaluation.

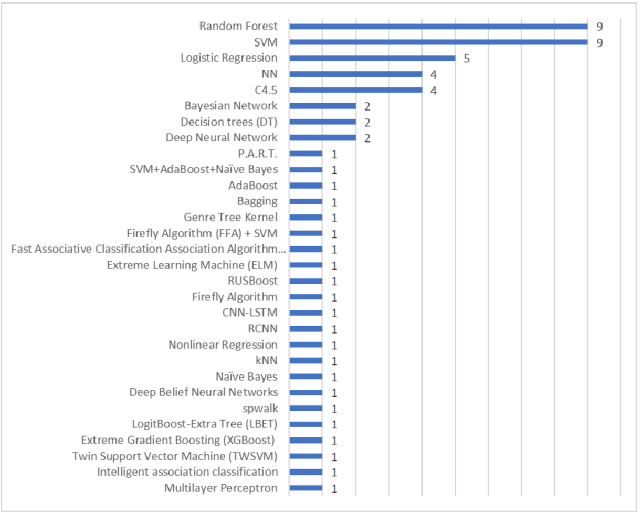

FIGURE 15. The method used in phishing classification techniques.

C. MOST USED METHODS

From 2010 to 2020, 30 methods were used to classify phishing attacks, as shown in Fig. 15, where the most widely used techniques were Random Forest, Support Vector Machine (SVM), Logistic Regression, Neural and Bayesian Networks, C4.5, Decision Tree (DT), and DNN (Deep Neural Network). This showed that Random Forest, SVM, and Logistic Regression were the three most used methods in this study (39%).

Each expert reported that their results were better than the techniques used by other studies. According to Chen *et al.* [58], the Logistic Regression performance was better than that of SVM and C4.5 on the phishing webpage. However, the DNN performance was better than that of SVM and C4.5 on the websites, according to Ali and Ahmed [6]. The study of Zhang *et al.* [33] consequently showed that the SVM performance was better than that of the Random Forest and Logistic Regression on a phishing website. Besides that, the Random Forest performed

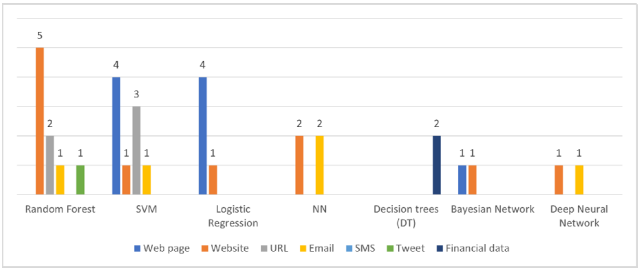

FIGURE 16. Most used methods in phishing type.

better than DT and SVM on phishing URLs, according to Sahingoz *et al.* [41].

This indicated that the specific methods performed better with the different phishing types. Although these methods had similar phishing types, they were still not necessarily able to produce the same performance. Moreover, differences were obtained from the collection, processing, and testing of data. Zhang *et al.* [33] also collected specific data on phishing and legitimate e-Business websites, namely http://www.315online.com.cn and http://www.anquan.org. The number of websites obtained was also 1,416 and 1,462 for phishing and 1,462 legitimate platforms. Subsequently, the study of Ali and Ahmed [6] used UCI Machine learning with 1353 websites, which contained 548, 702, and 103 legitimate, phishing, and suspicious platforms, respectively. This indicated that cross-validation was used to validate the developed model. However, Zhang *et al.* [33] used the precision, recall, and F-measure values to evaluate classification techniques.

Based on Fig. 16, the most explained phishing type was the website (13 articles), accompanied by the webpage (10 articles), URL (6 articles), and email (5 articles). Random Forest (5 articles) was also the most analyzed phishing technique on websites, accompanied by the SVM and Logistic Regression, each with four articles on webpages, respectively. Meanwhile, the phishing type with little analysis were financial data, tweets, and SMS.

D. MOST USED PERFORMANCE EVALUATION METHOD

Various methods of performance evaluation were used also used on the proposed classifications, with variations caused by the researchers efforts to obtain the best results compared with similar studies. This indicated that accuracy was the commonly used evaluation method [61], although was unable to be used as a benchmark for measuring all types of classification abilities. Therefore, the more the performance evaluation methods used, the better the opportunities for effective model developments.

Based on Fig. 17, general performance evaluation, namely accuracy, was mostly utilized. This indicated that most performance evaluation was accuracy (45 articles), accompanied by the True Positive Rate (TPR) (30 articles), F-measure (22 articles), and Precision (21 articles). The following are the nine most used performance evaluation techniques for phishing classification,

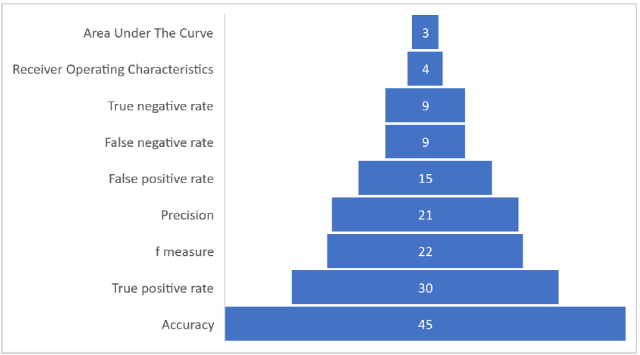

FIGURE 17. Top nine performance evaluations.

- *Accuracy*
  The percentage of correct phishing predictions and legitimate websites to the total number of platforms (websites) [62].
- *TPR/Recall/Sensitivity*
  The percentage of successfully and accurately predicted phishing websites from the total number of platforms (websites) [62].
- *Precision*
  The percentage of successfully and accurately predicted phishing websites from the total number of expected platforms (websites) [62].
- *F-Measure*
  Harmonic values indicated precision and recall [62].
- *False Positive Rate (FPR)*
  The percentage of legitimate websites incorrectly predicted from the total number of original platforms [62].
- *False Negative Rate*
  The percentage of phishing websites incorrectly predicted from the total number of platforms [62].
- *True Negative Rate/Specificity*
  The percentage of successfully and accurately predicted legitimate websites from the total number of original platforms [62].
- *Receiver Operating Characteristics*
  The plot values of TPR against FPR, using various threshold settings [64].
- *Area Under the Curve*
  The probability that the classification technique performed a higher ranking of randomly selected positive instances than the negative conditions [63].

Based on Table 6, the most widely used performance evaluations were accuracy, precision, and TPR/Recall/Sensitivity. However, the least used methods were Welch's T-Test, Shapiro-Wilk Test, Ranking Techniques, Prediction error rate, Geometric Mean, F1-Macro, and Matthew's Correlation Coefficient. According to the review, the classification technique parameter settings were compared with related studies, with some researchers exhibiting specific parameters. From Fig. 18, the information disclosure on the use of parameters increased yearly, especially in 2019 and 2020. Meanwhile, the researchers that did not mention specific

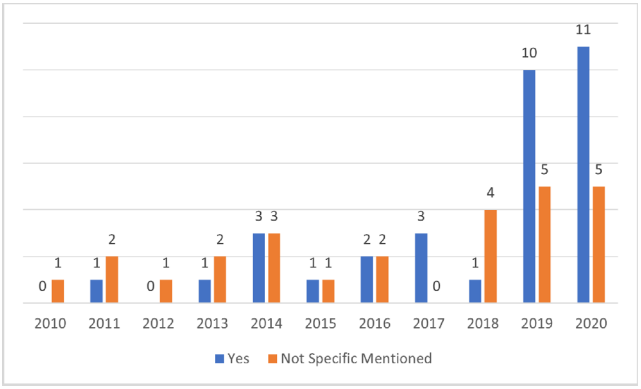

FIGURE 18. Parameter setting Classification technique based on research year.

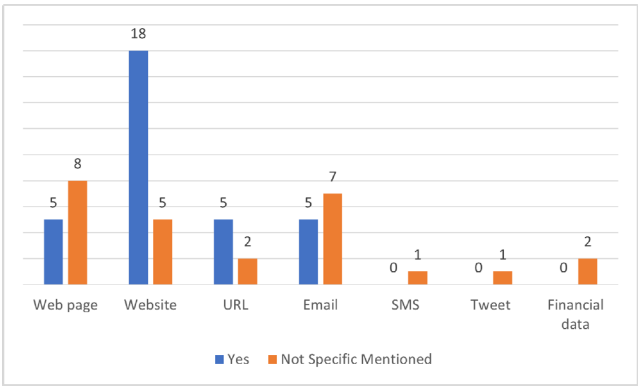

FIGURE 19. Parameter setting Classification techniques based on the type of phishing attack.

parameters experienced an increase of almost half, compared to others.

Based on Figs. 6 and 19, the phishing types containing websites, webpages, emails, and URLs mostly correlated with information disclosure, using the classification technique parameters. This indicated that the website was most significant for phishing webpages, emails, and URLs, based on the parameters. Approximately 54.55% of website phishing researchers also comprehensively conveyed the use of the classification parameters. Meanwhile, the phishing webpages (30.77%), emails (26.92%), and websites (19.23%) researchers did not provide information on the use of specific classification technique parameters.

According to the parameter setting classification techniques, Al-Fayoumi *et al.* [60] used the confidence values of (0.2, 0.5), (0.1, 0.4) and (0.05, 0.3) to produce the best performance. This was in line with Alsariera *et al.* [64], although only the number of iterations = 100 was able to produce a significant performance comparison. Therefore, disclosing information on the use of parameters produced phishing development continuity.

These results were also in line with several studies, as the parameter setting was used to ensure a fair comparison of related reports. Subsequent comparisons were also carried out in this study, such as observing changes in the performance

FIGURE 20. Different types of phishing attacks and their classification techniques.

of classification techniques to the parameters. However, some studies only utilized the parameters recommended by the related reports. This setting is subsequently a big challenge [65], as there were no general rules or standards to be followed towards obtaining the best results. These indicated that many researchers used various modifications to the size of the parameters used in improving the performance of classification techniques. The settings used by the researchers were also the number of folds in cross-validation [66], layer [67], hidden nodes [14], [11], learning rate [68], threshold [7], activation function [6], epoch [6], [64], [69], minimum support and confidence values [59], related studies parameter utilizations [39], and automation [60]. Therefore, the parameter settings improved the performance of classification techniques to maximum accuracy [6].

Several studies also conducted experimental classification techniques on dataset changes to parameter settings. This indicated that the best classification technique performance was achieved by changing the dataset [66], Therefore, the role of parameter setting was very important to the performance of the techniques [89]. In Table 7, the list of the primary studies with six attributes were also presented, namely year, main study, publication, dataset, method, and phishing type. This primary study contained 68 articles (January 2010-December 2020), and was ordered by the most recent year of publication. Fig. 20 shows the various phishing attacks with the classification techniques.

E. INSIGHTS AND FUTURE STUDY DIRECTIONS

Based on the SLR, the following are some future classification technique contributions for phishing attacks,

- *No studies used a different language dataset.*

Most researchers used the datasets in English, such as phishing emails [36], [39], [66], [73] and SMS [46], indicating that the chances of increasing the prevention of attacks are small. Therefore, the phishing attacks in various languages were needed to measure the classification techniques.

- *There are no expert-based feature recommendations.*

Many researchers only depended on the preprocessing features, especially emails, subsequently confusing the message's classification [68]. For example, an email containing a URL from a colleague was received after attending a webinar or meeting, categorized as a regular mail. However, it is categorized as a suspicious email when the mail received

contained a short message as a URL. The complex behavior can lead to various innovations in crime [86], such as using persuasion techniques. Many suspicious emails use persuasion techniques to deceive their victims [85], [87], [88]. Therefore, the expert validation of the email is required, especially when related to the phishing classification features.

- *There is no standard value or cut-off range for performance evaluation.*

No categories were determined for the performance assessment of the classification technique. This indicated that the researchers used a value close to 1, indicating the best performance [28]. Besides accuracy, many researchers also used alternative measurements, based on the observation of higher TPR or lower FPR values. Another critical issue is the difficulty in providing comparative evaluations among different phishing detection techniques. This was mainly due to the restriction of sharing, as well as the lack of standard benchmarks and reference datasets, based on the attackers' dynamic nature and potential data sensitivity [21].

F. LIMITATIONS

This study had limitations in the search for articles, as only journal-based publications with an impact factor were used. Clarivate analytics-WoS was also used to obtain the articles, as journals with emerging index citation sources in the WoS core collection were ignored. Additionally, only the phish-based articles were selected and classified in the computer science category.

VI. CONCLUSION

In this study, a more in-depth evaluation was observed at the phishing classification techniques, using systematic literature review. This was the first systematic review in the past 10 years, with a comprehensive focus on the classification techniques. Several recommendations were also provided, based on helping study researchers obtain more insight into the development of phishing. The results showed that many researchers performed comparisons without describing the parameter setting of the utilized classification technique. Other issues found were also the incorrect evaluation and validation of the classification technique performances, as well as the diversity of the dataset's utilization. In addition, the proposed systematic literature review thoroughly described the gaps in the classification techniques, for the development of phishing studies to be highly focused on more efficient solutions.

ACKNOWLEDGMENT

This research was funded by the Malaysia Ministry of Education, Universiti Kebangsaan Malaysia under grants, FRGS/1/2019/ICT01/UKM/01/2 and KKP/2020/UKM/4/3.

REFERENCES

[1] Verizon, A. J. Nathan, and A. Scobell. (2020). *2020 Data Breach Investigations Report*. [Online]. Available: https://enterprise.verizon.com/resources/reports/2020-data-breach-investigations-report.pdf%0Ahttp://bfy.tw/HJvH

[2] A. Das, S. Baki, A. El Aassal, R. Verma, and A. Dunbar, "SoK: A comprehensive reexamination of phishing research from the security perspective," *IEEE Commun. Surveys Tuts.*, vol. 22, no. 1, pp. 671–708, 1st Quart., 2020.

[3] Anti-Phishing Working Group. (2020). *Phishing Activity Trends Report 3 Quarter 2020*. [Online]. Available: https://docs.apwg.org/reports/apwg_trends_report_q3_2020.pdf

[4] K. D. Rajab, "New hybrid features selection method: A case study on websites phishing," *Secur. Commun. Netw.*, vol. 2017, pp. 1–10, 2017.

[5] W. Zhang, H. Ren, and Q. Jiang, "Application of feature engineering for phishing detection," *IEICE Trans. Inf. Syst.*, vol. 99, no. 4, pp. 1062–1070, 2016.

[6] W. Ali and A. A. Ahmed, "Hybrid intelligent phishing website prediction using deep neural networks with genetic algorithm-based feature selection and weighting," *IET Inf. Secur.*, vol. 13, no. 6, pp. 659–669, Nov. 2019.

[7] E. Zhu, Y. Chen, C. Ye, X. Li, and F. Liu, "OFS-NN: An effective phishing websites detection model based on optimal feature selection and neural network," *IEEE Access*, vol. 7, pp. 73271–73284, 2019.

[8] K. L. Chiew, C. L. Tan, K. Wong, K. S. C. Yong, and W. K. Tiong, "A new hybrid ensemble feature selection framework for machine learning-based phishing detection system," *Inf. Sci.*, vol. 484, pp. 153–166, May 2019.

[9] E. Zhu, Y. Ju, Z. Chen, F. Liu, and X. Fang, "DTOF-ANN: An artificial neural network phishing detection model based on decision tree and optimal features," *Appl. Soft Comput.*, vol. 95, Oct. 2020, Art. no. 106505.

[10] M. Sameen, K. Han, and S. O. Hwang, "PhishHaven—An efficient real-time AI phishing URLs detection system," *IEEE Access*, vol. 8, pp. 83425–83443, 2020.

[11] W. Zhang, Q. Jiang, L. Chen, and C. Li, "Two-stage ELM for phishing web pages detection using hybrid features," *World Wide Web*, vol. 20, no. 4, pp. 797–813, Jul. 2017.

[12] C. Sur, "Ensemble one-vs-all learning technique with emphatic & rehearsal training for phishing email classification using psychology," *J. Exp. Theor. Artif. Intell.*, vol. 30, no. 6, pp. 733–762, Nov. 2018.

[13] P. Vaitkevicius and V. Marcinkevicius, "Comparison of classification algorithms for detection of phishing websites," *Informatica*, vol. 31, no. 1, pp. 143–160, Mar. 2020.

[14] Y.-H. Chen and J.-L. Chen, "Ai@ntiphish—Machine learning mechanisms for cyber-phishing attack," *IEICE Trans. Inf. Syst.*, vol. 102, no. 5, pp. 878–887, May 2019.

[15] M. Khonji, A. Jones, and Y. Iraqi, "An empirical evaluation for feature selection methods in phishing email classification," *Int. J. Comput. Syst. Sci. Eng.*, vol. 28, no. 1, pp. 37–51, 2013.

[16] I. Qabajeh, F. Thabtah, and F. Chiclana, "A recent review of conventional vs. automated cybersecurity anti-phishing techniques," *Comput. Sci. Rev.*, vol. 29, pp. 44–55, Aug. 2018.

[17] A. A. Akinyelu, "Machine learning and nature inspired based phishing detection: A literature survey," *Int. J. Artif. Intell. Tools*, vol. 28, no. 5, Aug. 2019, Art. no. 1930002.

[18] T. Gangavarapu, C. D. Jaidhar, and B. Chanduka, "Applicability of machine learning in spam and phishing email filtering: Review and approaches," *Artif. Intell. Rev.*, vol. 53, no. 7, pp. 5019–5081, Oct. 2020.

[19] M. Khonji, Y. Iraqi, and A. Jones, "Phishing detection: A literature survey," *IEEE Commun. Surveys Tuts.*, vol. 15, no. 4, pp. 2091–2121, 4th Quart, 2013.

[20] G. Varshney, M. Misra, and P. K. Atrey, "A survey and classification of web phishing detection schemes," *Secur. Commun. Netw.*, vol. 9, no. 18, pp. 6266–6284, Dec. 2016.

[21] Z. Dou, I. Khalil, A. Khreishah, A. Al-Fuqaha, and M. Guizani, "Systematization of knowledge (SoK): A systematic review of software-based web phishing detection," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 4, pp. 2797–2819, 4th Quart., 2017.

[22] A. Almomani, B. B. Gupta, S. Atawneh, A. Meulenberg, and E. Almomani, "A survey of phishing email filtering techniques," *IEEE Commun. Surveys Tuts.*, vol. 15, no. 4, pp. 2070–2090, 4th Quart., 2013.

[23] A. Aleroud and L. Zhou, "Phishing environments, techniques, and countermeasures: A survey," *Comput. Secur.*, vol. 68, pp. 160–196, Jul. 2017.

[24] W. Hadi, F. Aburub, and S. Alhawari, "A new fast associative classification algorithm for detecting phishing websites," *Appl. Soft Comput.*, vol. 48, pp. 729–734, Nov. 2016.

[25] M. Arshey and K. S. A. Viji, "An optimization-based deep belief network for the detection of phishing E-mails," *Data Technol. Appl.*, vol. 54, no. 4, pp. 529–549, Jul. 2020.

[26] Y. Ding, N. Luktarhan, K. Li, and W. Slamu, "A keyword-based combination approach for detecting phishing webpages," *Comput. Secur.*, vol. 84, pp. 256–275, Jul. 2019.

[27] C. K. Olivo, A. O. Santin, and L. S. Oliveira, "Obtaining the threat model for e-mail phishing," *Appl. Soft Comput.*, vol. 13, no. 12, pp. 4841–4848, Dec. 2013.
[28] R. Gowtham and I. Krishnamurthi, "PhishTackle—A web services architecture for anti-phishing," *Cluster Comput.*, vol. 17, no. 3, pp. 1051–1068, Sep. 2014.
[29] W. Ali and S. Malebary, "Particle swarm optimization-based feature weighting for improving intelligent phishing website detection," *IEEE Access*, vol. 8, pp. 116766–116780, 2020.
[30] R. S. Rao and A. R. Pais, "Detection of phishing websites using an efficient feature-based machine learning framework," *Neural Comput. Appl.*, vol. 31, no. 8, pp. 3851–3873, Aug. 2019.
[31] R. M. Mohammad, F. Thabtah, and L. McCluskey, "Intelligent rule-based phishing websites classification," *Inf. Security, IET*, vol. 8, no. 3, pp. 153–160, May 2014.
[32] A. Abbasi, F. Zahedi, D. Zeng, Y. Chen, H. Chen, and J. F. Nunamaker, "Enhancing predictive analytics for anti-phishing by exploiting website genre information," *J. Manage. Inf. Syst.*, vol. 31, no. 4, pp. 109–157, Jan. 2015.
[33] D. Zhang, Z. Yan, H. Jiang, and T. Kim, "A domain-feature enhanced classification model for the detection of Chinese phishing e-business websites," *Inf. Manage.*, vol. 51, no. 7, pp. 845–853, Nov. 2014.
[34] X. Liu and J. Fu, "SPWalk: Similar property oriented feature learning for phishing detection," *IEEE Access*, vol. 8, pp. 87031–87045, 2020.
[35] C. L. Tan, K. L. Chiew, K. S. C. Yong, S. N. Sze, J. Abdullah, and Y. Sebastian, "A graph-theoretic approach for the detection of phishing webpages," *Comput. Secur.*, vol. 95, Aug. 2020, Art. no. 101793.
[36] R. Islam and J. Abawajy, "A multi-tier phishing detection and filtering approach," *J. Netw. Comput. Appl.*, vol. 36, no. 1, pp. 324–335, Jan. 2013.
[37] Y. Fang, C. Zhang, C. Huang, L. Liu, and Y. Yang, "Phishing email detection using improved RCNN model with multilevel vectors and attention mechanism," *IEEE Access*, vol. 7, pp. 56329–56340, 2019.
[38] A. K. Jain and B. B. Gupta, "A novel approach to protect against phishing attacks at client side using auto-updated white-list," *EURASIP J. Inf. Secur.*, vol. 2016, no. 1, p. 9, Dec. 2016.
[39] O. A. Adewumi and A. A. Akinyelu, "A hybrid firefly and support vector machine classifier for phishing email detection," *Kybernetes*, vol. 45, no. 6, pp. 977–994, Jun. 2016.
[40] H. Abutair, A. Belghith, and S. Alahmadi, "CBR-PDS: A case-based reasoning phishing detection system," *J. Ambient Intell. Humanized Comput.*, vol. 10, no. 7, pp. 2593–2606, Jul. 2019.
[41] O. K. Sahingoz, E. Buber, O. Demir, and B. Diri, "Machine learning based phishing detection from URLs," *Expert Syst. Appl.*, vol. 117, pp. 345–357, Mar. 2019.
[42] M. Volkamer, K. Renaud, B. Reinheimer, and A. Kunz, "User experiences of TORPEDO: TOoltip-poweRed phishing email detection," *Comput. Secur.*, vol. 71, pp. 100–113, Nov. 2017.
[43] K. J. Nishanth, V. Ravi, N. Ankaiah, and I. Bose, "Soft computing based imputation and hybrid data and text mining: The case of predicting the severity of phishing alerts," *Expert Syst. Appl.*, vol. 39, no. 12, pp. 10583–10589, Sep. 2012.
[44] X. Chen, I. Bose, A. C. M. Leung, and C. Guo, "Assessing the severity of phishing attacks: A hybrid data mining approach," *Decis. Support Syst.*, vol. 50, no. 4, pp. 662–672, Mar. 2011.
[45] S. W. Liew, N. F. M. Sani, M. T. Abdullah, R. Yaakob, and M. Y. Sharum, "An effective security alert mechanism for real-time phishing tweet detection on Twitter," *Comput. Secur.*, vol. 83, pp. 201–207, Jun. 2019.
[46] S. Mishra and D. Soni, "Smishing detector: A security model to detect smishing through SMS content analysis and URL behavior analysis," *Future Gener. Comput. Syst.*, vol. 108, pp. 803–815, Jul. 2020.
[47] J. W. Joo, S. Y. Moon, S. Singh, and J. H. Park, "S-detector: An enhanced security model for detecting smishing attack for mobile computing," *Telecommun. Syst.*, vol. 66, no. 1, pp. 29–38, Sep. 2017.
[48] FBI's IC3. (2020). *2020 Internet Crime Report*. [Online]. Available: https://www.ic3.gov/Media/PDF/AnnualReport/2020_IC3Report.pdf
[49] G. Davis and R. Samani. (2021). *McAfee Mobile Threat Report*. [Online]. Available: https://www.mcafee.com/enterprise/en-us/assets/reports/rp-mobile-threat-report-2019.pdf
[50] A. K. Jain and B. B. Gupta, "Feature based approach for detection of smishing messages in the mobile environment," *J. Inf. Technol. Res.*, vol. 12, no. 2, pp. 17–35, Apr. 2019.
[51] S.-H. Moon and D.-W. Park, "Forensic analysis of MERS smishing hacking attacks and prevention," *Int. J. Secur. Appl.*, vol. 10, no. 6, pp. 181–192, Jun. 2016.
[52] A. Lee, K. Kim, H. Lee, and M. Jun, "A study on realtime detecting smishing on cloud computing environments," in *Advanced Multimedia and Ubiquitous Engineering* (Lecture Notes in Electrical Engineering), vol. 354. Berlin, Germany: Springer, 2016, pp. 495–501.
[53] A. K. Jain and B. B. Gupta, "Rule-based framework for detection of smishing messages in mobile environment," *Proc. Comput. Sci.*, vol. 125, pp. 617–623, Jan. 2018.
[54] G. Sonowal and K. S. Kuppusamy, "SmiDCA: An anti-smishing model with machine learning approach," *Comput. J.*, vol. 61, no. 8, pp. 1143–1157, Aug. 2018.
[55] J. Rastenis, S. Ramanauskaitė, I. Suzdalev, K. Tunaitytė, J. Janulevičius, and A. Čenys, "Multi-language spam/phishing classification by email body text: Toward automated security incident investigation," *Electronics*, vol. 10, no. 6, p. 668, 2021.
[56] L. Wu, X. Du, and J. Wu, "Effective defense schemes for phishing attacks on mobile computing platforms," *IEEE Trans. Veh. Technol.*, vol. 65, no. 8, pp. 6678–6691, Aug. 2016.
[57] M. Aburrous, M. A. Hossain, K. Dahal, and F. Thabtah, "Intelligent phishing detection system for e-banking using fuzzy data mining," *Expert Syst. Appl.*, vol. 37, no. 12, pp. 7913–7921, Dec. 2010.
[58] T.-C. Chen, T. Stepan, S. Dick, and J. Miller, "An anti-phishing system employing diffused information," *ACM Trans. Inf. Syst. Secur.*, vol. 16, no. 4, pp. 1–31, Apr. 2014.
[59] M. Al-Fayoumi, J. Alwidian, and M. Abusaif, "Intelligent association classification technique for phishing website detection," *Int. Arab J. Inf. Technol.*, vol. 17, no. 4, pp. 488–496, Jul. 2020.
[60] M. Babagoli, M. P. Aghababa, and V. Solouk, "Heuristic nonlinear regression strategy for detecting phishing websites," *Soft Comput.*, vol. 23, no. 12, pp. 4315–4327, Jun. 2019.
[61] A. K. Jain and B. B. Gupta, "A machine learning based approach for phishing detection using hyperlinks information," *J. Ambient Intell. Humanized Comput.*, vol. 10, no. 5, pp. 2015–2028, May 2019.
[62] R. S. Rao, A. R. Pais, and P. Anand, "A heuristic technique to detect phishing websites using TWSVM classifier," *Neural Comput. Appl.*, vol. 33, no. 11, pp. 5733–5752, Sep. 2020.
[63] T. Fawcett, "An introduction to ROC analysis," *Pattern Recognit. Lett.*, vol. 27, no. 8, pp. 861–874, Jun. 2006.
[64] Y. A. Alsariera, V. E. Adeyemo, A. O. Balogun, and A. K. Alazzawi, "AI meta-learners and extra-trees algorithm for the detection of phishing websites," *IEEE Access*, vol. 8, pp. 142532–142542, 2020.
[65] G. Vrbančič, I. Fister, and V. Podgorelec, "Parameter setting for deep neural networks using swarm intelligence on phishing websites classification," *Int. J. Artif. Intell. Tools*, vol. 28, no. 6, Sep. 2019, Art. no. 1960008.
[66] V. Ramanathan and H. Wechsler, "Phishing detection and impersonated entity discovery using conditional random field and latent Dirichlet allocation," *Comput. Secur.*, vol. 34, pp. 123–139, May 2013.
[67] M. U. Chowdhury, J. H. Abawajy, A. V. Kelarev, and T. Hochin, "Multilayer hybrid strategy for phishing email zero-day filtering," *Concurrency Comput., Pract. Exper.*, vol. 29, no. 23, p. e3929, Dec. 2017.
[68] C. N. Gutierrez, T. Kim, and R. D. Corte, "Learning from the ones that got away: Detecting new forms of phishing attacks," *IEEE Trans. Dependable Secure Comput.*, vol. 15, no. 6, pp. 988–1001, Dec. 2018.
[69] P. Yang, G. Zhao, and P. Zeng, "Phishing website detection based on multidimensional features driven by deep learning," *IEEE Access*, vol. 7, pp. 15196–15209, 2019.
[70] R. S. Rao, T. Vaishnavi, and A. R. Pais, "CatchPhish: Detection of phishing websites by inspecting URLs," *J. Ambient Intell. Humanized Comput.*, vol. 11, no. 2, pp. 813–825, Feb. 2020.
[71] A. S. Bozkir and M. Aydos, "LogoSENSE: A companion HOG based logo detection scheme for phishing web page and E-mail brand recognition," *Comput. Secur.*, vol. 95, Aug. 2020, Art. no. 101855.
[72] R. S. Edwin and R. Ravi, "A performance analysis of software defined network based prevention on phishing attack in cyberspace using a deep machine learning with CANTINA approach (DMLCA)," *Comput. Commun.*, vol. 153, pp. 375–381, Mar. 2020.
[73] E. S. Gualberto, R. T. De Sousa, T. P. De B. Vieira, J. P. C. L. Da Costa, and C. G. Duque, "From feature engineering and topics models to enhanced prediction rates in phishing detection," *IEEE Access*, vol. 8, pp. 76368–76385, 2020.
[74] A. El Aassal, S. Baki, A. Das, and R. M. Verma, "An in-depth benchmarking and evaluation of phishing detection research for security needs," *IEEE Access*, vol. 8, pp. 22170–22192, 2020.

[75] R. S. Rao, T. Vaishnavi, and A. R. Pais, "PhishDump: A multi-model ensemble based technique for the detection of phishing sites in mobile devices," *Pervas. Mobile Comput.*, vol. 60, Nov. 2019, Art. no. 101084.
[76] S. Smadi, N. Aslam, and L. Zhang, "Detection of online phishing email using dynamic evolving neural network based on reinforcement learning," *Decis. Support Syst.*, vol. 107, pp. 88–102, Mar. 2018.
[77] V. Muppavarapu, A. Rajendran, and S. K. Vasudevan, "Phishing detection using RDF and random forests," *Int. Arab J. Inf. Technol.*, vol. 15, no. 5, pp. 817–824, 2018.
[78] T. Chin, K. Xiong, and C. Hu, "Phishlimiter: A phishing detection and mitigation approach using software-defined networking," *IEEE Access*, vol. 6, pp. 42516–42531, 2018.
[79] M. Moghimi and A. Y. Varjani, "New rule-based phishing detection method," *Expert Syst. Appl.*, vol. 53, pp. 231–242, Jul. 2016.
[80] G. Geng, X. Lee, and Y. Zhang, "Combating phishing attacks via brand identity and authorization features," *Secur. Commun. Netw.*, vol. 8, no. 6, pp. 888–898, Apr. 2015.
[81] R. Gowtham and I. Krishnamurthi, "A comprehensive and efficacious architecture for detecting phishing webpages," *Comput. Secur.*, vol. 40, pp. 23–37, Feb. 2014.
[82] S. Marchal, J. Francois, R. State, and T. Engel, "PhishStorm: Detecting phishing with streaming analytics," *IEEE Trans. Netw. Service Manage.*, vol. 11, no. 4, pp. 458–471, Dec. 2014.
[83] G. Xiang, J. Hong, C. P. Rose, and L. Cranor, "CANTINA+: A feature-rich machine learning framework for detecting phishing web sites," *ACM Trans. Inf. Syst. Secur.*, vol. 14, no. 2, pp. 1–28, 2011.
[84] H. Zhang, G. Liu, T. W. S. Chow, and W. Liu, "Textual and visual content-based anti-phishing: A Bayesian approach," *IEEE Trans. Neural Netw.*, vol. 22, no. 10, pp. 1532–1546, Oct. 2011.
[85] C. Naksawat, S. Akkakoson, and C. K. Loi, "Persuasion strategies: Use of negative forces in scam E-mails," *GEMA Online J. Lang. Stud.*, vol. 16, no. 1, pp. 1–17, 2016.
[86] N. Farhana, N. M. Tahira, M. Alic, and N. D. K. Ashard, "Mini-review of street crime prediction and classification methods," *Jurnal Kejuruteraan*, vol. 33, no. 3, pp. 391–401, 2021.
[87] M. R. Abd Rahman, "Online scammers and their mules in Malaysia," *Jurnal Undang-Undang Dan Masyarakat*, vol. 26, no. 2020, pp. 65–72, Sep. 2020.
[88] A. H. Shaari, M. R. Kamaluddin, W. F. Paizi-Fauzi, and M. Mohd, "Online-dating romance scam in Malaysia: An analysis of online conversations between scammers and victims," *GEMA Online J. Lang. Stud.*, vol. 19, no. 1, pp. 97–115, Feb. 2019.
[89] Y. M. Mohmad Hassim and R. Ghazali, "An approach to improve functional link neural network training using modified artificial bee colony for classification task," *Asia–Pacific J. Inf. Technol. Multimedia*, vol. 2, no. 2, pp. 63–71, Dec. 2013.

RAHMAD ABDILLAH received the S.T. degree from the Faculty of Science and Technology, Universitas Islam Negeri Sultan Syarif Kasim Riau, Indonesia, in 2010, and the M.T. degree in informatics from the Institut Teknologi Bandung (ITB), Indonesia, in 2013. He is currently pursuing the Ph.D. degree with the Faculty of Information Science and Technology, Universiti Kebangsaan Malaysia (UKM). He has been working as a Lecturer with the Department of Informatics, Universitas Islam Negeri Sultan Syarif Kasim Riau, since 2010. His research interests include cybersecurity, artificial intelligence, and social engineering.

ZARINA SHUKUR received the Ph.D. degree from the University of Nottingham, in 1999. She is currently a Professor with the Center for Cyber Security Studies, Universiti Kebangsaan Malaysia. Her research interests include formal methods and cybersecurity.

MASNIZAH MOHD received the B.I.T. and M.I.T. degrees in information science from Universiti Kebangsaan Malaysia (UKM), in 1999 and 2002, respectively, and the Ph.D. degree in computer and information sciences from the University of Strathclyde, in 2010. She is currently an Associate Professor with UKM. Her current research interests include information retrieval, natural language processing, and cyber intelligence.

TS. MOHD ZAMRI MURAH received the B.Sc. and M.Sc. degrees in statistics from the University of Iowa, Iowa, USA, in 1987 and 1989, respectively. He is currently a Senior Lecturer with the Center for Cyber-security, Universiti Kebangsaan Malaysia. His current research interests include the development of deep learning models for cybersecurity, automated penetration testing, and cyber range.