

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/382446662>

A review of a website phishing detection taxonomy

Conference Paper · June 2024

DOI: 10.6084/m9.figshare.26345473

CITATIONS

0

READS

35

2 authors, including:



[Damian Frąszczak](#)

Military University of Technology

31 PUBLICATIONS 65 CITATIONS

SEE PROFILE

A Review of A Website Phishing Detection Taxonomy

Edyta FRĄSZCZAK,
Military University of Technology, Warsaw, Poland
edyta.fraszczak97@gmail.com

Damian FRĄSZCZAK
Military University of Technology, Warsaw, Poland
damian.fraszczak@wat.edu.pl

Abstract

Phishing is a common cybercrime that uses social engineering to trick people into revealing sensitive information such as personal identity data and financial credentials. This paper thoroughly analyzes various techniques used to detect website phishing, categorizing them based on the primary features they analyze and the methods they use. By examining the strengths and weaknesses of each approach, this review aims to emphasize the challenges and advancements in phishing detection, highlighting the need for ongoing research and adaptation to counter evolving phishing tactics effectively. The detailed comparison and discussion of these techniques offer valuable insights for developing more robust and accurate phishing detection systems.

Keywords: phishing, website analysis, phishing detection, artificial intelligence

Introduction And Research Motivation

Phishing is a type of cyber attack involving social engineering techniques to deceive individuals into providing sensitive information, such as login credentials, credit card numbers, and personal identification details. Attackers pretend to be trustworthy entities through emails, messages, or fake websites to trick victims into revealing their confidential information. This stolen information is then used for malicious purposes, such as financial fraud, identity theft, and unauthorized access to personal and corporate data (Alkhalil et al., 2021; Grimes, 2024; Peltier, 2006).

Nowadays, there are over 1 billion registered websites (Gupta, 2024), closely linked to the nearly 5.5 billion Internet users ("Internet use in 2024," 2024). Phishing attacks present a significant threat in today's digital world, with both the number and sophistication of attacks increasing. According to the Anti-Phishing Working Group, December 2021 witnessed the highest monthly total of phishing attacks in history. In 2023 alone, CERT Polska received 95,696 phishing reports in Polish networks (*Raport roczny z działalności CERT Polska w 2023 roku*, 2024). These reports involved 51,374 URLs from 49,039 domains, which resolved to 3,097 IP addresses. These attacks can lead to substantial financial losses, compromised user data, and damage the reputation of organizations (Grimes, 2024).

Additionally, phishing attacks rank as the second most expensive cause of data breaches, costing businesses an average of \$4.65 million per breach. These trends reflect the growing prevalence of phishing websites over the years, highlighting the dual nature of opportunities and threats brought by the development of the Internet. It is important to note that phishing can be used for other adversarial activities like disinformation campaigns (Saxena, 2024; "Top Cybersecurity Statistics for 2024," n.d.). Disinformation, often spread through deceptive websites and fraudulent content, shares similarities with phishing in terms of tactics used to mislead users (Frąszczak, 2021a, 2021b). By leveraging phishing detection techniques, it is possible to identify and mitigate the spread of false information online, thus enhancing the overall integrity and security of the digital ecosystem.

Website phishing detection has become a popular research topic due to its critical importance in cybersecurity. Numerous solutions have been proposed, ranging from rule-based methods to advanced machine-learning techniques (Arshad et al., 2021; Divakaran and Oest, 2022; "Phishing Detection Leveraging Machine Learning and Deep Learning," n.d.). Despite significant progress, there is still room for improvement. Attackers now employ more intelligent tools to create phishing websites, such as generative artificial intelligence (e.g., GAN networks). Existing

methods often struggle with new and evolving phishing tactics, highlighting the need for continual advancements in this field.

The main aim of this paper is to provide a comprehensive overview of phishing, its techniques, taxonomy, and the methods used for detecting phishing websites. The paper is divided into four main parts. The first part introduces the problem and research motivation. The second part provides a taxonomy of website phishing detection. The third part reviews state-of-the-art methods. Finally, the fourth part presents the conclusion of the paper, summarizing the findings and insights.

Phishing Detection Taxonomy

Formally, the detection of phishing websites can be presented as a binary classification problem. In its simplest form, this involves classifying a website as either being a "phishing" (positive class, $y = 1$) or a "legitimate" one (negative class, $y = 0$) based on a set of features (x) extracted from the website.

Given a set of features $x = [x_1, x_2, \dots, x_n]$, where each x_i represents a specific attribute of the website (e.g., URL length, presence of IP address, use of URL shortening services), the task is to determine the class label $y \in \{0, 1\}$. The classification function f can be expressed then as:

$$y = f(x) \quad (1)$$

where function f maps the feature vector x into the output y , where:

$$\begin{cases} 1 & \text{if the website is classified as phishing} \\ 0 & \text{if the website is classified as legitimate} \end{cases} \quad (2)$$

In the context of machine learning, the classification function f is often represented by a model, such as a logistic regression, decision tree, support vector machine or neural network. The model is trained on a labeled dataset $D = \{(x_i, y_i)\}, i \in [1, m]$ where m is the number of training example.

Phishing detection depends on a wide range of features extracted from different parts of websites. These features fall into four main categories: URL and HTTP features, graphical features, network features, and content-based features. Each category contains specific attributes that aid in identifying phishing websites. In Table 1, you can find a summary of the most frequently used features. (Alkhalil et al., 2021; Divakaran and Oest, 2022; Mohammad et al., 2012; "Phishing Detection Leveraging Machine Learning and Deep Learning," n.d.; "Phishing Detection Leveraging Machine Learning and Deep Learning," n.d.; Sönmez et al., 2018).

Table 1 A review of the most common website features used for phishing detection

Feature Type	Feature	Description	Sample rule
URL and HTTP Features	URL Length	Longer URLs are often used to conceal malicious parts.	$URLs \geq 54$ characters are likely phishing.
	IP Address in URL	Using an IP address instead of a domain name.	IP address in URL indicates phishing.
	HTTPS Token in Domain	Adding 'HTTPS' in the domain part to mislead users.	Presence of 'HTTPS' token in domain suggests phishing.
	Redirection (//) in URL Path	Double slashes in the path often indicate redirection.	If '/' appears beyond the seventh position, it is likely phishing.
	TinyURL Usage	URL shortening services used to mask phishing URLs.	Use of TinyURL is indicative of phishing.
	URLs with '@' Symbol	'@' symbol in URL leads browsers to ignore preceding parts.	Presence of '@' in URL indicates phishing.
Graphical Features	Favicon	Favicon loaded from an external domain.	External favicon indicates phishing.
	Images from Different Domains	Loading images from domains other than the main domain.	High percentage of external images suggests phishing.
	Pop-up Windows	Use of pop-up windows to gather personal information.	Pop-up windows with forms indicate phishing.
	HTML Frames	Use of invisible iFrames to display additional content.	Presence of iFrames suggests phishing.
	Status Bar Customization	JavaScript used to fake URLs in the status bar.	Manipulation of status bar indicates phishing.
	Disabling Right Click	JavaScript to disable right-click functions.	Disabled right-click is a phishing indicator.
Network Features	Domain Registration Length	Short registration periods are common in phishing domains.	Registration length ≤ 1 year suggests phishing.
	DNS Record	Absence of DNS records for the domain.	No DNS record indicates phishing.
	Website Traffic	Low traffic and poor ranking on web analytics tools.	Low Alexa rank or unrecognized domains indicate phishing.
	PageRank	Low PageRank values.	PageRank < 0.2 suggests phishing.
	Google Index	Whether the site is indexed by Google.	Absence from Google's index indicates phishing.

	Number of Links Pointing to the Page	Few or no backlinks.	0 links suggest phishing; 1-2 links are suspicious.
Content-Based Features	Textual Content Analysis	Analysis of specific words and phrases indicative of fraud.	Presence of certain keywords can indicate phishing.
	HTML and JavaScript Features	Use of specific HTML tags and JavaScript functions.	Features like 'onMouseOver' and disabling right-click suggest phishing.
	Form Handlers	Presence of suspicious form handlers, such as those pointing to different domains.	Form handlers pointing to different domains or using 'mailto' suggest phishing.
	Meta, Script, and Link Tags	Percentage of external links in meta, script, and link tags.	High percentage of external links in these tags suggests phishing.
	Request URL	External objects loaded from different domains.	High percentage of external requests suggests phishing.

The required properties of web pages can be processed and scraped independently or retrieved from publicly available datasets. The most commonly used datasets in the area of website phishing detection have been summarized in **Table 2**. (“Alexa Top Websites | Last Save before it was closed,” n.d.; “APWG | The APWG eCrime Exchange (eCX),” n.d.; “Home - UCI Machine Learning Repository,” n.d.; “OpenPhish - Phishing Intelligence,” n.d.; “PhishTank | Join the fight against phishing,” n.d.)

Table 2 A review of the most common available datasets used for phishing detection

Name	Description
PhishTank	Collaborative clearing house for phishing data. Provides an API for submitting and querying phishing URLs.
UCI Machine Learning Repository	Online available high-quality datasets of different domains including phishing detection websites.
Alexa Top Sites	Dataset of the top-ranked websites globally, used to compare against suspected phishing sites.
OpenPhish	Provides a continuously updated dataset of phishing URLs. Available as a paid service with API access
APWG eCrime Exchange (eCX)	Repository of phishing and cybercrime data contributed by APWG members, including detailed information about phishing attacks.

Furthermore, platforms like PhishTank and OpenPhish work in real-time, allowing users to submit URLs they suspect are phishing attempts. Each submission is verified, and users can check if a URL has been reported as a phishing attempt. For instance, the PhishTank API allows developers to incorporate URL checks into their applications, responding to JSON or XML formats. For instance, a sample API response may look like presented in **Listing 1**. This response indicates that the URL "https://www.example.org/" is in the PhishTank database, has been verified as phishing, and provides details about the report.

```

<url0>
  <url>https://www.example.org/</url>
  <in_database>true</in_database>
  <phish_id>11728</phish_id>
  <phish_detail_page>
    http://www.phishtank.com/phish_detail.php?phish_id=11728
  </phish_detail_page>
  <verified>true</verified>
  <verified_at>2006-10-01T02:32:23+00:00</verified_at>
  <valid>true</valid>
  <submitted_at>2006-10-01T02:28:46+00:00</submitted_at>
</url0>

```

Listing 1 Sample PhishTank JSON API response

In cases where the available data is insufficient or a new aspect is being considered, researchers have to use available tools and build a dedicated ETL process for that task. The sample architectures used for crawlers have been presented in many publications (Frąszczak and Magdziarz, 2023; Sunil Kumar and Neelima, 2011) . It is worth mentioning that the process often relies on open-source tools for gathering website data and extracting critical features. A summary of these tools is presented in **Table 3** (Almaqali et al., 2020; Frąszczak, 2022a; Frąszczak and Frąszczak, 2024; Haddaway, 2015; Morina and Sejdiu, 2022). Building that process sometimes involves the UI component. In such situations, using UI-generated views is recommended, as they can reduce development time and achieve great results.(Frąszczak, 2022b)

Table 3 Review of open-source tools that can be used for extracting website parameters

Purpose	Tool name	Description
Downloading website data	HTTrack	HTTrack is a free and open-source web crawler and offline browser utility.
	Scrapy	Scrapy is an open-source web-crawling framework for Python used to extract data from websites.
	Selenium	Selenium is a portable framework for testing web applications, useful for automating interactions.
	StormCrawler	StormCrawler is an open-source SDK for building low-latency, scalable web crawlers based on Apache Storm.
Extracting content-related features	Beautiful Soup	Beautiful Soup is a Python library for parsing HTML and XML documents.
	lxml	lxml is a library that provides a convenient API for parsing XML and HTML.
	HTMLParser	HTMLParser is a simple and fast HTML and XHTML parser library in Python.
	Tesseract	Tesseract is an OCR tool for recognizing text in images.
	Nltk	Nltk is a platform for building Python programs to work with human language data.
Extracting network-related features	NetworkX	NetworkX is a package for the creation, manipulation, and study of complex networks.
	Gephi	Gephi is an open-source network analysis and visualization software.
	RPaSDT	A sophisticated GUI-based tool for network analysis dedicated to disinformation propagation and source detection.
	NetCenLib	The networkX-based library is used for identifying crucial nodes in networks.

	Graph-tool	Graph-tool is an efficient module for the manipulation and statistical analysis of graphs.
Extracting and analyzing page response	Scrapy	Scrapy is an open-source framework for extracting data and analyzing HTTP responses.
	Fiddler	Fiddler is a web debugging proxy tool that logs HTTP(S) traffic.
	Charles Proxy	Charles Proxy is a web debugging proxy application for monitoring HTTP and HTTPS traffic.
Gathering website structural data	Selenium	Selenium is used for testing and automating interactions with HTML/CSS structures.
	Puppeteer	Puppeteer provides a high-level API to control headless Chrome or Chromium.
	PhantomJS	PhantomJS is a headless WebKit scriptable with JavaScript API for automated browsing.
	Playwright	Playwright is a Node library to automate Chromium, Firefox, and WebKit.

After collecting data, models are tested and validated. This paper also presents a review of current techniques for phishing detection, treating it as a classification problem and using classical evaluation metrics. Typically, these techniques are evaluated using key metrics based on the confusion matrix, including True Positives (TP), True Negatives (TN), False Positives (FP), and False Negatives (FN). A summary of them is presented in **Table 4**. Evaluating these metrics helps understand the performance of phishing detection models, optimizing their effectiveness and minimizing false alarms (Arshad et al., 2021; Divakaran and Oest, 2022; Tharwat, 2021).

Table 4 Metrics used for the evaluation of phishing website detection

Name	Formula	Analysis
Accuracy	$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$	The percentage of all correct classifications (both phishing and non-phishing) out of the total number of examples.
Precision	$Precision = \frac{TP}{TP + FP}$	The percentage of correctly identified phishing sites among all sites predicted as phishing.
Recall	$Recall = \frac{TP}{TP + FN}$	Recall is the ratio of the number of correctly identified outbreaks over the real true outbreaks.
F-measure	$F\text{-measure} = \frac{2 \times Precision \times Recall}{Precision + Recall}$	The harmonic mean of precision and recall, provides a balance between the two.

Review of website phishing detection methods

Phishing detection involves using different techniques depending on the available data and the aspects being tested. This paper categorizes the available techniques into four main groups: machine learning-based methods, content-based analysis, list-based approaches, and heuristic strategies (Arshad et al., 2021; Divakaran and Oest, 2022; Kara et al., 2022; “Phishing Detection Leveraging Machine Learning and Deep Learning,” n.d.; Tang and Mahmoud, 2021). The division is presented in.

Figure 1.

Machine learning techniques are widely utilized for phishing detection due to their effectiveness and adaptability (Ludl et al., 2007) utilized a decision tree algorithm, relying on 18 HTML and URL features, achieving an accuracy of 83.09%. (Kulkarni and Iii, 2019) employed multiple classifiers, including decision trees, Naive Bayesian classifiers, SVM, and neural networks, reaching 90% accuracy on a dataset of 1,353 safe URLs. In (Fette et al., n.d.) PILFER has been introduced, utilizing SVM and ten specific features to achieve 92% accuracy, despite concerns about the dataset's size. (Chiew et al., 2019) used the Random Forest algorithm with a dataset of 5,000 URLs, showing an accuracy of 94.6%



Figure 1 Overview of Phishing Detection Techniques

List-based approaches involve whitelists and blacklists. (Jain and Gupta, 2016) proposed an auto-updated whitelist approach, achieving 86.02% accuracy with a false-positive rate of 1.48%, suitable for real-time environments. Blacklists, such as those used by Google Safe Browsing, identify phishing sites based on manually reported URLs. These methods provide quick and straightforward protection but struggle with the rapid creation and short lifespan of phishing sites.

Heuristic strategies identify phishing webpages based on extracted features compared to legitimate ones. CANTINA presented in (Zhang et al., 2007) uses the TF-Df method for content-based detection, effective for text-heavy pages but less so for image-based content. (“PhishShield: A Desktop Application to Detect Phishing Webpages through Heuristic Approach - ScienceDirect,” n.d.) extracts features like footer links and website logos, offering faster and more comprehensive detection than blacklist methods. However, heuristic methods can be bypassed when attackers understand and counter the heuristics.

Visual similarity techniques detect phishing by comparing the visual content of suspect websites with trusted domains. (Hadi et al., 2016) used spatial features, while (Mao et al., 2017) employed the EMD algorithm to analyze visual similarity. Recent deep learning approaches, like VisualPhishNet (Abdelnabi et al., 2020) and Phishpedia (Lin et al., 2021), use screenshot analysis to detect phishing, achieving high accuracy by learning similarities from reference lists of known websites. However, these methods require the full loading of webpages, leading to higher latency and vulnerability to adversarial attacks.

Each phishing detection method has its strengths and limitations. Machine learning-based methods offer high accuracy but require continuous updates to adapt to evolving threats. Content-based models provide more detailed analysis but at the cost of higher computational resources. List-based approaches offer quick, real-time detection but are less effective against new phishing sites. Heuristic methods and visual similarity techniques add valuable layers of detection, though they also face challenges in adapting to sophisticated evasion techniques (Arshad et al., 2021; Divakaran and Oest, 2022; “Phishing Detection Leveraging Machine Learning and Deep Learning,” n.d.).

Conclusions

Phishing is a common cybercrime that uses social engineering to trick people into revealing sensitive information such as personal identity data and financial credentials. This paper thoroughly analyzes various techniques used to detect website phishing, categorizing them based on the primary features they analyze and the methods they use. By examining the strengths and weaknesses of each approach, this review aims to emphasize the challenges and advancements in phishing detection, highlighting the need for ongoing research and adaptation to counter evolving phishing tactics effectively. The detailed comparison and discussion of these techniques offer valuable insights for developing more robust and accurate phishing detection systems.

References

- Alexa Top Websites | Last Save before it was closed [WWW Document], n.d. URL <https://www.expireddomains.net/alexa-top-websites/> (accessed 6.16.24).
- Alkhalil, Z., Hewage, C., Nawaf, L., Khan, I., 2021. Phishing Attacks: A Recent Comprehensive Study and a New Anatomy. *Front. Comput. Sci.* 3, 563060. <https://doi.org/10.3389/fcomp.2021.563060>
- Almaqbali, I.S.H., Al Khufairi, F.M.A., Khan, M.S., Bhat, A.Z., Ahmed, I., 2020. Web Scrapping: Data Extraction from Websites. *J. Stud. Res.* <https://doi.org/10.47611/jsr.vi.942>
- APWG | The APWG eCrime Exchange (eCX), n.d. URL <https://apwg.org/ecx/> (accessed 6.16.24).
- Arshad, A., Rehman, A.U., Javaid, S., Ali, T.M., Sheikh, J.A., Azeem, M., 2021. A Systematic Literature Review on Phishing and Anti-Phishing Techniques. <https://doi.org/10.48550/ARXIV.2104.01255>
- Chiew, K.L., Tan, C.L., Wong, K.S., Yong, K.S.C., Tiong, W.K., 2019. A new hybrid ensemble feature selection framework for machine learning-based phishing detection system. *Inf. Sci.* 484, 153–166. <https://doi.org/10.1016/j.ins.2019.01.064>
- Divakaran, D.M., Oest, A., 2022. Phishing Detection Leveraging Machine Learning and Deep Learning: A Review. <https://doi.org/10.48550/ARXIV.2205.07411>
- Fette, I., Sadeh, N., Tomasic, A., n.d. Learning to Detect Phishing Emails.
- Frąszczak, D., 2022a. RPaSDT—Rumor Propagation and Source Detection Toolkit. *SoftwareX* 17, 100988. <https://doi.org/10.1016/j.softx.2022.100988>
- Frąszczak, D., 2022b. NEFBDA — .NET Environment for Building Dynamic Angular Applications. *SoftwareX* 19, 101163. <https://doi.org/10.1016/j.softx.2022.101163>
- Frąszczak, D., 2021a. Fake News Source Detection – The State of The Art Survey For Current Problems And Research, in: *Proceedings of the 37th International Business Information Management Association (IBIMA)*.

Presented at the Innovation management and information technology impact on global economy in the era of pandemic., International Business Information Management. <https://doi.org/10.6084/M9.FIGSHARE.16545675>

- Frąszczak, D., 2021b. Information Propagation In Online Social Networks - A Simulation Case Study, in: Proceedings of the 38th International Business Information Management Association (IBIMA). Presented at the Innovation management and information technology impact on global economy in the era of pandemic. Proceedings of the 38th International Business Information Management Association Conference (IBIMA), International Business Information Management, Cordoba, Spain. <https://doi.org/10.6084/M9.FIGSHARE.18974987.V1>
- Frąszczak, D., Frąszczak, E., 2024. NetCenLib: A comprehensive python library for network centrality analysis and evaluation. *SoftwareX* 26, 101699. <https://doi.org/10.1016/j.softx.2024.101699>
- Frąszczak, D., Magdziarz, K., 2023. The Architecture Concepts for Building Highly Scalable Crawling Cluster For Data-Driven On-Page Optimization. <https://doi.org/10.6084/M9.FIGSHARE.21909273.V1>
- Grimes, R.A., 2024. Fighting Phishing: everything you can do to fight social engineering and phishing, 1st ed. John Wiley and Sons, Indianapolis.
- Gupta, R., 2024. How Many Websites Are There? The Web in 2024. Themeisle Blog. URL <https://themeisle.com/blog/how-many-websites-are-there/> (accessed 6.16.24).
- Haddaway, N., 2015. The Use of Web-scraping Software in Searching for Grey Literature. *Grey J.* 11, 186–190.
- Home - UCI Machine Learning Repository [WWW Document], n.d. URL <https://archive.ics.uci.edu/> (accessed 6.16.24).
- Internet use in 2024 [WWW Document], 2024. . DataReportal – Glob. Digit. Insights. URL <https://datareportal.com/reports/digital-2024-deep-dive-the-state-of-internet-adoption> (accessed 6.16.24).
- Jain, A., Gupta, B.B., 2016. Comparative analysis of features based machine learning approaches for phishing detection. 2016 3rd Int. Conf. Comput. Sustain. Glob. Dev. INDIACom.
- Kara, I., Ok, M., Ozaday, A., 2022. Characteristics of Understanding URLs and Domain Names Features: The Detection of Phishing Websites With Machine Learning Methods. *IEEE Access* 10, 124420–124428. <https://doi.org/10.1109/ACCESS.2022.3223111>
- Kulkarni, A., Iii, L.L.B., 2019. Phishing Websites Detection using Machine Learning. *Int. J. Adv. Comput. Sci. Appl. IJACSA* 10. <https://doi.org/10.14569/IJACSA.2019.0100702>
- Ludl, C., McAllister, S., Kirda, E., Krügel, C., 2007. On the Effectiveness of Techniques to Detect Phishing Sites. Presented at the Proceedings of the Detection of Intrusions and Malware and Vulnerability Assessment Conference (DIMVA), pp. 1–20.
- Mohammad, R., Thabtah, F., McCluskey, T., 2012. An assessment of features related to phishing websites using an automated technique.
- Morina, V., Sejdiu, S., 2022. Evaluating and comparing web scraping tools and techniques for data collection.
- OpenPhish - Phishing Intelligence [WWW Document], n.d. URL <https://openphish.com/> (accessed 6.16.24).
- Peltier, T.R., 2006. Social Engineering: Concepts and Solutions. *Inf. Syst. Secur.* 15, 13–21. <https://doi.org/10.1201/1086.1065898X/46353.15.4.20060901/95427.3>
- Phishing Detection Leveraging Machine Learning and Deep Learning: A Review [WWW Document], n.d. . ar5iv. URL <https://ar5iv.labs.arxiv.org/html/2205.07411> (accessed 5.24.24).
- PhishShield: A Desktop Application to Detect Phishing Webpages through Heuristic Approach - ScienceDirect [WWW Document], n.d. URL <https://www.sciencedirect.com/science/article/pii/S1877050915013411> (accessed 6.17.24).
- PhishTank | Join the fight against phishing [WWW Document], n.d. URL <https://phishtank.org/> (accessed 6.12.24).
- Raport roczny z działalności CERT Polska w 2023 roku, 2024. . CERT Polska.
- Saxena, A., 2024. 100+ Phishing Attack Statistics You Should Know in 2024. Sprinto. URL <https://sprinto.com/blog/phishing-statistics/> (accessed 6.16.24).
- Sönmez, Y., Tuncer, T., Gökal, H., Avcı, E., 2018. Phishing web sites features classification based on extreme learning machine, in: 2018 6th International Symposium on Digital Forensic and Security (ISDFS). Presented at the 2018 6th International Symposium on Digital Forensic and Security (ISDFS), IEEE, Antalya, Turkey, pp. 1–5. <https://doi.org/10.1109/ISDFS.2018.8355342>