IJMECIE

International Journal of Mechanical, Electrical and Civil Engineering

E-ISSN: 3047-4531
P-ISSN: 3047-4523

(Research/Review) Article

Formulating Tracing and Linking Algorithm For Identifying Cyber Terrorism Website

Wan Ahmad Ramzi W.Y [1*], Aznizah Ab.Karim [2], Mohamad Shaufi Kambaruddin [3], Afzanizam Alias [4], Mohd Faizal Yahaya [5]

[1] Politeknik Sultan Abdul Halim Mu'adzam Shah, Kementerian Pendidikan Tinggi Malaysia, Bandar Darul Aman, 0600 Jitra, Kedah Darul Aman, Malaysia
Email : ramzi@polimas.edu.my

[2,3,4,5] Departemen Teknologi Informasi dan Komunikasi, Politeknik Sultan Abdul Halim Mu'adzam Shah Malaysia. Syarikat Air Melaka Berhad, Malaysia
Email : ramzi@polimas.edu.my

* Corresponding Author : Wan Ahmad Ramzi W.Y

Abstract: Nowadays, the number of cyber terrorists that using Internet as a medium keep increasing around the world are really worrying. Even though the cyber terrorist is publishing their post openly and public, there are quite difficult to recognize their main communication media. This happen because the cyber terrorist might hide their agenda through any sympathetic base website. Therefore, the aim of this project is to formulate tracing and linking algorithm using data mining technique in order to identify the relationship between each cyber terrorism components that may result to cyber terrorism website.

Keywords: cyber-attack, cyber terrorist, data mining, Fraud, Internet, Intrusion, technology, website.

1. Introduction

In this new era technology, terrorist around the world have been implementing Internet as one of communication medium to fulfil their purpose and fully utilized the Internet in term of spreading their propaganda, connected to each other's around the world, planning and recruiting. Even though government has taken several actions including blocking the website, yet the cyber terrorists are smart as they are hiding through website with sympathetic motive within the situation. This might confuse the government to take action as it is difficult to identify the real cyber terrorism website. Based on prior study, number of frameworks regarding on cyber terrorism have been developed. One that has been focused is on the cyber terrorism component. There are six (6) main components of cyber terrorism components that have been identified in order to recognize the cyber terrorism website involving of actor, method, motive, target, tools and method. [1] A potential website can identify manually based on the cyber terrorism components identified. To enhance and assist the identification process for a potential website, a few numbers of technique from data mining have been identified. The techniques are used in the process of tracing and linking thus combining both to formulate an algorithm that will help to identify the potential website.

Related Work

In this section, a study related to the topic is conducted to support the importance of the research. The definition of Cyber Terrorism, their issues, framework and data mining techniques approach will be discussed in this section..

Cyber Terrorism

Cyber terrorism has already emerged, yet a clear definition remains to be established. Various interpretations of the term "cyber terrorism" exist. These differing definitions are summarized in Table 2.1.

Received: 08 March , 2025
Revised: 28 March , 2025
Accepted: 18 April, 2025
Online Available : 23 April , 2025
Curr. Ver.: 23 April , 2025

Table 1: Definition of Cyber Terrorism

Definition	Author
Cyber terrorism is the convergence of terrorism and cyberspace. Also known as unlawful attacks against computers, networks to intimidate or force a government for political or social objectives.	[2]Denning, 2000
Cyber terrorism is referred to as electronic terrorism or information war.	[3]Rouse, 2010
Cyber terrorism can be referred as the use of cyber tools that being used by terrorist organization to attack the infrastructure of economy or civilian systems as a target.	[4]Schweitzer et al., 2011
Cyber terrorism is the premeditated, politically motivated attack against information, computer systems, computer programs and data which result in violence against non-combatant targets by sub-national groups or clandestine agents.	[5]Curran, 2011
"The use of computer network tools to shut down critical national infrastructures (such as energy, transportation, government operations) or to coerce or intimidate a government or civilian population".	[6] A. Lewis, 2013
"Exploitation of electronic vulnerabilities by terrorist groups in pursuit of their political aims."	[7] Cruz, 2013
This cyber terrorism is best define as "transnational terrorist organization" whereby the insurgent and jihadist that have used the internet as tools for planning attack, radicalization, recruitment, a method of propaganda distribution and a mean for communication and for disruptive purpose.	[8] Theohary et al., 2015

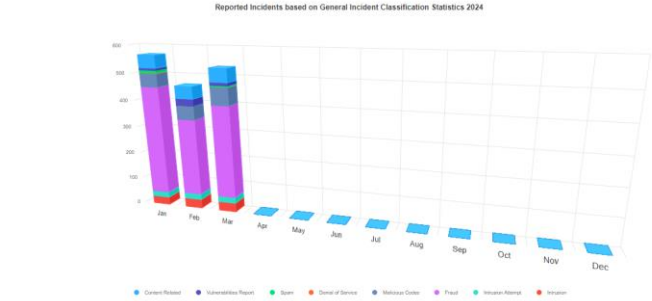
Despite the various definitions provided, they all centre around the concept of cyber terrorism. In this research, cyber terrorism is defined as the use of the Internet as a medium and communication tool by cyber terrorists to carry out their activities and plan attacks aimed at causing damage to their targets and the cyberspace infrastructure. This scope includes both internal and external networks. Issue Cyber Terrorism

The rise in cybercrime has amplified issues related to cyber terrorism. The Internet enables global communication, providing cyber terrorists a platform to plan attacks and spread propaganda. They exploit cyberspace to incite fear and panic among the public, often to achieve political goals [9]. Activities like recruitment, fundraising, acquiring advanced knowledge, and disseminating propaganda are also conducted in this space, as digital platforms facilitate spreading and disposing of messages easily.

Cyber terrorists commonly use tactics such as distributed denial-of-service (DoS) attacks, hate websites, emails, and breaches of sensitive networks [10]. They also engage in hacking, spreading extremist propaganda, and targeting ICT infrastructures, sometimes

leading to the failure or destruction of national assets. Despite the increasing frequency of such cases, research on identifying websites used for these activities remains insufficient, allowing cyber terrorism to persist and grow [11].This has been proved by reported by Rapid7 in National Exposure Index [12]with Malaysia placed at 31 from top 50 countries most vulnerable attacks. Besides, as reported by MyCert, Cyber security Malaysia in Incidents based on General Incident Classification Statistic 2024 shows the vulnerabilities rank 2 from January to Mac 2024 as shown in Table 2.2

Table 2: Incidents based on General Incident Classification Statistic 2024



#	JAN	FEB	MAC	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	TOTAL
Content Related	47	45	50	0	0	0	0	0	0	0	0	0	142
Vulnerabilities Report	8	25	11	0	0	0	0	0	0	0	0	0	44
Spam	9	1	5	0	0	0	0	0	0	0	0	0	15
Denial of Service	2	0	0	0	0	0	0	0	0	0	0	0	2

Framework of Cyber Terrorism

There are several of frameworks of cyber terrorism proposed since today. However, one of the most familiar are the frameworks Actor-Target-Effect as shown in figure 2.1 proposed by [13], The framework consists of six components: actors, motives, resources, targets, activities, and effects. **Actors** refer to organizational structures, risk tolerance, and ethical or moral constraints. **Motives** for cyber terrorism include social, psychological, economic, and political factors. **Resources** encompass weapons, finances, methods, logistics, or personnel. **Targets** are categorized as infrastructure, realization, disorganization, and revision. **Effects** of cyber terrorism can be physical, syntactic, or semantic, with both immediate and indirect impacts.

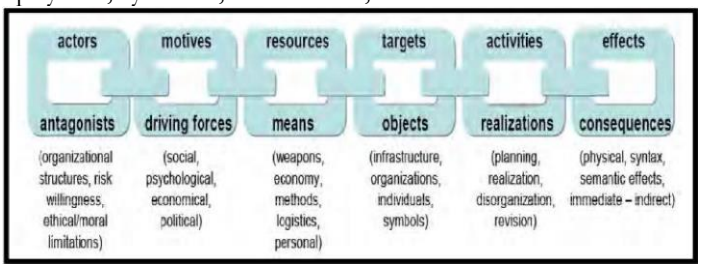


Figure 1: Heickero (Actor-Target-Effect chain) Framework[13]

While Gordon and Ford introduced a framework to study cyber terrorism behavior, as illustrated in Figure 2.2. This framework comprises seven components: perpetrator, place, action, tool, target, affiliation, and motivation.**Perpetrator** refers to either an individual or a group. **Place** signifies that cyber terrorism activities occur globally, unrestricted by physical locations, as the Internet fosters a globalized environment. **Action** includes threats, violence, education, recruitment, and strategic planning. **Tool** encompasses methods such as kidnapping, harassment, propaganda, and education, with computers often utilized to execute attacks like hacking and deploying computer viruses.

Targets typically include government officials or corporations. **Affiliation** refers to either actual or claimed associations. Finally, **motivation** is often driven by a desire for social or political change.

Components	LTTE (Example)	Description
Perpetrator	Group/Individual	In cyber context, virtual interactions can lead to anonymity.
Place	Worldwide	The event does not have to occur in a particular location. The Internet has introduced globalization of the environment.
Action	Threats/Violence/ Recruitment/ Education/Strategies	Terrorist scenarios typically are violent or involve threats of violence. Violence in virtual environment includes psychological effects, possible behavior modification and physical trauma.
Tool	Kidnapping/ Harassment/ Propaganda/Education	Terrorist use the computer as tool. Facilitating identity theft, computer viruses, hacking are examples fall under this category.
Target	Government Officials/Corporations	Potential targets are corporations and government computer systems.
Affiliation	Actual/Claimed	Affiliation refers to recruitment in carrying out given instructions. Affiliation can result in strengthening of the individual organizations as they can immediately acquire access to the information resources of their allies.
Motivation	Social/Political Change	Political, social and economic are the motivations present in the real-world terrorism.

LTTE = Liberation Tigers of Tamil Eelam (Sri Langka)

Figure 2: Gordon & Ford Cyber Terrorism Framework [14]

Based on Figure 2.2, [15] has proposed their framework in order to identify the cyber terrorism behaviour. The framework consists of six components that are targets, impact, method of action, domain, tool of attack and motivation. The framework for data collection and data analysis involved two stages. The first stage is conducted on data analysis on the context of cyber terrorism framework by using grounded theory approach by questionnaires and semi-structured interview. Meanwhile for the second stage is conducted based on statistical analysis on the cyber terrorism framework by using the questionnaires and structured survey.

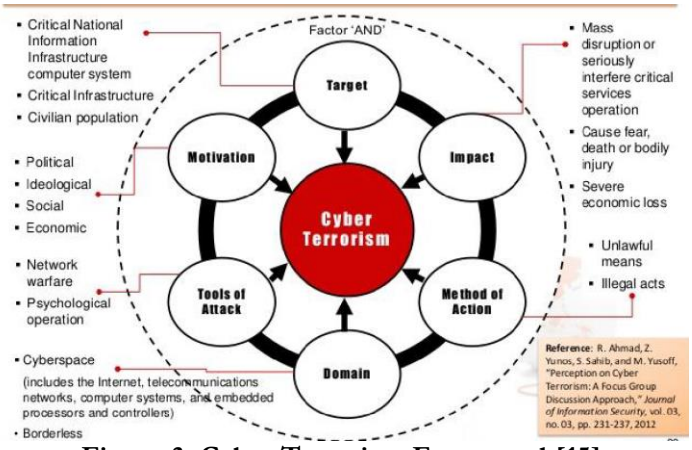


Figure 3: Cyber Terrorism Framework[15]

Salleh et. al has proposed a framework to know the cyber terrorism pattern and behaviour, there are six components involved in this framework which are actor, motivation, tool, impact, target and method. The actor can be either organization or person. Motivation can be explained as concept, ideology or economic change. Next component, tool can be either weapon or network warfare. Method should be operation or action. Next component that is target can be either person, place, victims or organization. Meanwhile impact can be described as violence or threat. as shown in Figure 2.4.

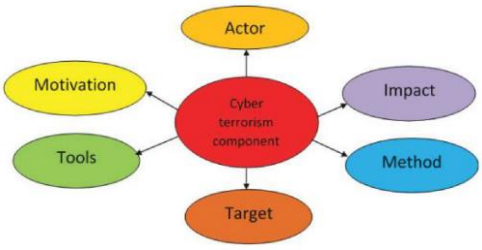


Figure 4: Cyber Terrorism Components [16]

Tracing Pattern

Trace pattern can be define as a regular way of process to discover the origin or starting point of scenario that has been occurred [17]. Trace pattern plays important roles in cyber terrorism by identify the behaviour of cyber terrorism. With the trace pattern identified, it can assist to overcome the cyber terrorism that occur in our country. It is an essential element in helping investigator in a crime scene to find the evidence which may lead to the effective strategy to counter cyber terrorist. Trace pattern plays important role in tracing any activities of cyber terrorism. The evidence of this crime can be in term of data a record that consist of

user activities such as login, logout, and computer shutdown, file execution and network packet.

Besides, trace pattern will assist the forensic investigator to find any evidence about the cyber terrorism because any activities of cyber terrorist or attacker can be identified based on the traces data found in the attack pattern that represent in the form of trace pattern. In this situation, trace pattern will help to determine how cyber terrorism could be happened [17]. The conceptual framework will assist to develop more efficient trace pattern. It is because of all elements in conceptual framework will give details information and could be use as a necessary tool to trace cyber terrorism activities in a cyber space.

In this research, trace pattern is defined as a way to discover the origin or starting point of a scenario that has happened. The main role of trace pattern is to identify the potential cyber terrorism website. This will help to understand context of website in conducting investigation.

Linking Technique

Linking technique is mainly process of linking data that can be done by comparing attributes value of instances in source class with target class [18]. Besides it can be used to find relationship between different database either two or more relational databases [19]. It is also can a powerful method to quickly understand and observe on the situation. This will help the forensic investigator in identifying the potential website by linking and get the relationship between related components of cyber terrorism.

Link analysis technique is one of the data mining technique categories. This technique has been implemented in many fields including healthcare, medical and cybercrime. [20]Lun Yen stated that this technique can be implemented in order to predict future action by studying the historical data. It can be functioned to analyse the relationship between different relational databases [19]. Besides, data mining technique will used data or information that are extracted from the web base system or website in order to achieve the knowledge of those system itself [21].

This analysis technique basically relates with record linkage. It can identify the corresponding database that have same entity in quick and accurately [22]. Basically, it will assist in matching the same entity between source data and training data ad relate to other components. This would help to make the decision-making process become easier and faster. There are two top of record linkage that are known as deterministic and probabilistic. It is need to study the scenario before choosing the suitable algorithm. Deterministic linkage technique is normally determined the agreement between record pairs [23]. Meanwhile probabilistic linkage has been determined to assess the discriminatory power of each identifier and the likelihood between record pairs [23]. Probabilistic record linkage can merge and mapped with database with the absence of unique identifier [24].

Based on Yannik, there are anount of records that cannot be linked automatically to match data in every pair records. In order to minimize the issue, record linkage technique has been combined with the machine learning technique [25]. With aid of supervised machine learning, this would minimize the interference of human by providing training data to map with the source data.

A heuristic technique is being used in order to detect and scan the specific website through a signature database that being built and contained with important information. If any signature match with heuristic pattern provided in database, it will then link to the machine learning technique [26]. The study of detection will involve on checking the URL, features and counting the values of heuristic data. The output will be analysed based on threshold value and resulting either as true positive, true negative, false positive or false negative.

Pooja implementing data mining in order to identify on online radicalization through textual analysis on message that have been post in the group where Social Network Analysis as a basic. Clustering of data is being used where it is referred as a group of objects or situation are being putting together, respectively reassemble based on category or member [27].

The process flow of tracking the extremist is by using data mining to capture the data, filter and cauterize the data respectively and then create a dynamic social network analysis.

Proposed Solution

Based on the prior study and literature review, it is proposed to use a trace pattern to identifying the potential cyber terrorism website. A proposed trace pattern would be the six

main components of cyber terrorism that involving of actor, method, motive, target, tool and impact. The analysis will be done by finding the relationship of each components for a suspected website. To identify the relationship, it is proposed to developing on the heuristic method. This heuristic method is fusing on the web searching technique and entity matching technique for tracing the related keywords meanwhile hierarchical clustering technique in order to link the relationship for each attributes and cyber terrorism components found. Both trace and linking technique will be combined in a developed algorithm where it then will be test and run on controlled environment. For this research, it is proposed to have a training database that containing keywords for each cyber terrorism components for the developed algorithm to be run in order to identify the potential cyber terrorism website based on the threshold value.

Web searching technique is being proposed in order to identify the URL link or websites that need to be crawled or extracted the data. It is proposed to find on the website or blogs to trace the specific website. Next proposed solution will relate on entity matching technique in supervised machine learning as this technique will be used to trace the keywords found from potential website and map into the training database. Hierarchical clustering technique is being proposed as a solution as it will cluster all keywords that successfully traced in the training database and then counted to achieve the threshold value and create a relationship between others attributes and cyber terrorism components respectively. This will lead to retrieve a result on identifying status of approached website either it is a potential cyber terrorism website or a normal website.

2. Methodology

Data Collection

For this research, quantitative data collection is used that rely on structured data to predetermine the potential output. The strategy used in this project involving the experiments based on the prior study. The categorization of keywords of each component are identified based on cyber terrorism components identified in prior study.

There are also approached of mixed methods design that combining the qualitative and quantitative method for data collection [28]. Even though only few find it useful, but it may help to analyst and verify the data collected. 33

In this research, a training data is being used in a controlled environment. Training database will contain with a number of general keywords or vocabulary that normally used by cyber terrorism in their communication through the website that have been extracted and classified. The data have been trained so that algorithm used will analyses based on the existing keywords.

Framework

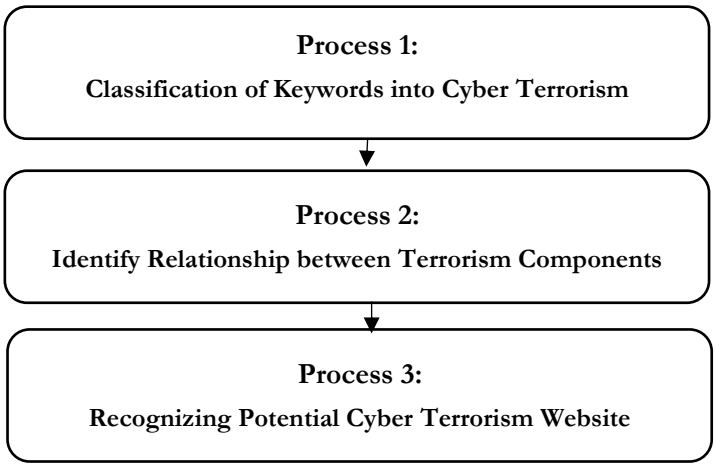


Figure 5 Framework Methodology

Figure 3.1, it shows the framework methodology used for this project. The elaboration for each process is shown below.

Process 1: Classification of Keywords into Cyber Terrorism Components

In this process, a list of keywords that have been extracted will be used and classified referred to cyber terrorism components. There are six main components of cyber terrorism that consist of actor, method, motivation, tools, target and impact. The identified keywords are collected before using observation base. In this process, the classified keywords then will be matched with synonym term of keywords to be store together into the components to provide varies of keywords scheme. In the meantime, the classification is also running on the keywords that have been extracted from website, it then will be traced into the training database. If it matches with any existed keywords, it will be classified and counted.

Process 2: Identify Relationship between Cyber Terrorism Components

In this process, each cyber terrorism components will represent a dataset for collected keywords. To find out the relationship between components, first will provide the frequency or threshold value for each component being appeared in the potential website forum. With the specific threshold value for each component, then it will be linked and generated the output. In this process, a clustering linking technique will be used where each keyword that match with respective attributes and cyber terrorism components will be calculated. The process of identifying the value each cyber terrorism components will be counted based on the attributes and keywords that are traced between the training database containing with keywords and the potential approached website. If any keywords extracted and traced with any attributes for respective cyber terrorism components, the attributes will be counted and clustered thus will create an alive status for that cyber terrorism components.

Process 3: Recognizing Potential Cyber Terrorism Website

In this process, the output will be either the website status is either a potential cyber terrorism website. The process of recognizing the potential cyber terrorism website will depend on the six main cyber terrorism components with their respective attributes and keywords involving of actors, method, motive, target, tools and impacts. As mentioned in process 2, each cyber terrorism components will holding a status alive if their respective attributes contributing on a threshold value that greater than zero where counted on the keywords traced. Each cyber terrorism components are compulsory to have alive status to comply the process in identifying the potential cyber terrorism website. The recognition process will be depend on the formulation of trace and link technique.

3. Results and Discussion

The research conducted on identifying cyber terrorism websites using the tracing and linking algorithm has yielded significant findings. This section outlines the outcomes of testing and analysing the data sets with the developed algorithm.

Data Sets Tested

The algorithm was applied to four data sets to evaluate its performance:

Table 4.1: Dataset tested

Key Findings

Data Set	Keywords Extracted	Components Activated	Conclusion
DS1	**49 keywords, including "mujahid," "qaeda," "perang," "Palestin"**	**Actor, Method, Motive, Tools, Target, Impact**	**Potential cyber terrorism website**
DS2	**Keywords on extremist ideology (e.g.,**	**Actor, Motive, Method**	**Potential cyber terrorism website**

	"jihad," "syahid")		
DS3	**Terms like "umat Islam," "kaum penjajah"**	**Actor, Method, Target**	**Borderline; leaning potential**
DS4	**Limited matches, generic terms**	**None**	**Normal website**

Key Findings

DS1	**49 keywords, including "mujahid," "qaeda," "perang," "Palestin"**	**Actor, Method, Motive, Tools, Target, Impact**	**Potential cyber terrorism website**
DS2	**Keywords on extremist ideology (e.g., "jihad," "syahid")**	**Actor, Motive, Method**	**Potential cyber terrorism website**
DS3	**Terms like "umat Islam," "kaum penjajah"**	**Actor, Method, Target**	**Borderline; leaning potential**
DS4	**Limited matches, generic terms**	**None**	**Normal website**

Table 4.2: Key Findings

Summary of Results

The algorithm effectively identified potential cyber terrorism websites by analysing the relationship between the six components. DS1 and DS2 showed clear activation across all components, while DS3 presented partial activation. DS4 served as a control and was correctly identified as a normal website. This validates the algorithm's robustness in detecting cyber terrorism indicators while avoiding false positives.

Conclusion and Future Work

The findings from this research demonstrate the potential of using tracing and linking algorithms to identify cyber terrorism websites effectively. The results validate the approach's ability to differentiate between malicious and benign websites by analysing keywords and their relationships to cyber terrorism components. Future work will focus on enhancing the algorithm's scalability and accuracy by incorporating more extensive datasets, refining the keyword database, and integrating advanced machine learning techniques to improve detection precision and minimize false positives.

References

[1] N. Mohd Salleh, S. R. Selamat, R. Yusof, and S. Sahib, "Discovering cyber terrorism using trace pattern," *Int. J. Netw. Secur.*, vol. 18, no. 6, pp. 1034–1040, 2016.

[2] D. E. Denning, *Activism, Hacktivism, and Cyberterrorism: The Internet as a Tool for Influencing Foreign Policy*. Arbor House Publishing Company, 2000.

[3] M. Rouse, "Electronic terrorism and information warfare," *TechTarget*, 2010. [Online]. Available: https://www.techtarget.com

[4] Y. Schweitzer and S. Shay, *The Globalization of Terror: The Challenge of Al-Qaida and the Response of the International Community*. Sussex Academic Press, 2011.

[5] J. R. Curran, "Cyber terrorism: Its place in global security," *Int. Secur. Stud. J.*, vol. 5, no. 3, pp. 29–42, 2011.

[6] A. Lewis, "Critical infrastructures: Securing the networks of the 21st century," *Cybersecurity J.*, vol. 7, no. 2, pp. 14–28, 2013.

[7] F. J. Cruz, "The exploitation of vulnerabilities: Analyzing cyber terrorism," *J. Inf. Warfare*, vol. 12, no. 4, pp. 34–46, 2013.

[8] C. A. Theohary and J. Rollins, *Cyberwarfare and Cyber Terrorism: Infrastructural Attacks in the Digital Era*. Congressional Research Service, 2015.

[9] S. Dombe and A. Golandsky, "The evolution of cyber terrorism: Strategies and countermeasures," *J. Cybersecurity Stud.*, vol. 5, no. 3, pp. 87–100, 2016.

[10] M. Alisha, "Cyber terrorism: The technological dimension of modern terrorism," *Int. J. Secur. Netw.*, vol. 6, no. 2, pp. 33–40, 2010.

[11] R. Noor, "Emerging threats: Analyzing the impact of cyber terrorism on global security," *J. Digit. Crime*, vol. 3, no. 4, pp. 45–56, 2011.

[12] R. James, "National exposure index: A study on global cyber vulnerabilities," *Rapid7 Cybersecurity Reports*, 2016. [Online]. Available: https://www.rapid7.com

[13] R. Heickero, "Terrorism online: The new generation," *J. Inf. Warfare*, vol. 6, no. 1, pp. 25–34, 2007.

[14] S. Gordon and R. Ford, "Cyberterrorism? Computers as weapons of terror," *Comput. Secur.*, vol. 21, no. 7, pp. 636–647, 2002, doi: 10.1016/S0167-4048(02)01007-X.

[15] Z. Yunos and R. Ahmad, "Evaluating cyber terrorism components in Malaysia," in *Proc. 5th Int. Conf. Inf. Commun. Technol. Muslim World (ICT4M)*, 2014, doi: 10.1109/ict4m.2014.7020582.

[16] N. Mohd Salleh, S. R. Selamat, R. Yusof, and S. Sahib, "Discovering cyber terrorism using trace pattern," *Int. J. Netw. Secur.*, vol. 18, no. 6, pp. 1034–1040, 2016.

[17] S. R. Selamat et al., "Traceability in digital forensic investigation process," in *2011 IEEE Conf. Open Syst.*, pp. 101–106, 2011, doi: 10.1109/icos.2011.6079259.

[18] E. Simperl, "Collaborative ontology engineering: A survey," *Knowl. Eng. Rev.*, vol. 27, no. 1, pp. 1–29, 2012, doi: 10.1017/S0269888912000058.

[19] A. Smita and G. Amol, "Techniques for efficient linking of data in relational databases," *Int. J. Database Manag.*, vol. 5, no. 2, pp. 12–20, 2014.

[20] L. Yen, "Predictive analytics using link analysis: Applications in cybercrime and medical fields," *J. Data Min. Anal.*, vol. 8, no. 3, pp. 45–60, n.d.

[21] P. Minky, "Mining web-based systems for knowledge acquisition," *J. Inf. Syst.*, vol. 6, no. 4, pp. 19–26, 2014.

[22] L. Lifang, X. Wang, and Y. Zhang, "Advances in record linkage techniques for data integration," *J. Database Manag.*, vol. 10, no. 2, pp. 33–40, 2003.

[23] S. B. Dusetzina et al., "Linking data to support comparative effectiveness research," *Med. Care*, vol. 52, no. 10 Suppl 3, pp. S41–S47, 2014, doi: 10.1097/MLR.0000000000000077.

[24] C. Jared, "Advanced probabilistic record linkage for relational databases," *Database Technol. J.*, vol. 12, no. 5, pp. 67–75, 2016.

[25] S. Yannik, "Machine learning-assisted record linkage: A hybrid approach," *J. Mach. Learn. Appl.*, vol. 9, no. 4, pp. 56–70, 2014.

[26] L. A. Nguyen, B. L. To, H. K. Nguyen, and M. H. Nguyen, "Detecting phishing websites: A heuristic URL-based approach," in *2013 Int. Conf. Adv. Technol. Commun. (ATC)*, pp. 187–192, 2013, doi: 10.1109/atc.2013.6698185.

[27] R. Pooja, "Textual analysis for detecting radicalization in online communities," *J. Soc. Netw. Anal.*, vol. 4, no. 1, pp. 22–30, 2013.

[28] National Science Foundation, *The Use of Mixed Methods in the Social Sciences: Evaluating Qualitative and Quantitative Research*. Washington, D.C.: NSF, 2002.