S & M 3667

Development of Disinformation Verification System with Criminal Record Based on Previous Systems

Mu-Chuan Chen, I-Long Lin,* and Hung-Cheng Yang

Department of Computer Science and Engineering Tatung University, Zhongshan N. Rd., Taipei City 104, Taiwan

(Received January 3, 2024; accepted May 22, 2024)

Keywords: policy, misinformation, crime database, verification platform

It has become more difficult to distinguish disinformation due to the increasing amount of information on the Internet daily. Disinformation is disseminated through various communication technologies and endangers the quality of life by harming a sound information society. To ensure effective and safe information exchange, public-led information transparency is demanded to screen systems and databases that produce and spread disinformation. We compared false verification platforms on the Internet, namely, the Thailand Verification Center, Taiwan FactCheck Center, MyGoPen, and Cofacts, and constructed a new system to screen disinformation. This system based on the cybercrime theory can be used to verify disinformation and prevent fraud more efficiently than the previous false verification systems.

1. Introduction

The "New Generation Strategy and Action Program to Combat Fraud" was formulated by the Taiwanese government in 2022 to involve groups and tools to prevent fraud and block the illegal transfer of money.(1) Since 2017, fraud has been the most serious crime in Taiwan. More than 20000 cases of fraud have been reported yearly for the past six years (Fig. 1). Resulted financial losses exceeded USD 0.8 billion in 2021 (Fig. 2). However, the amount recovered was only about USD 83 million. The patterns of fraud have been constantly evolving. Given that the prosecution rate was only 6% out of 1629 cases in 2020, it has become necessary to enhance the detection rate for disinformation-related fraud and improve the efficiency of evidence collection.

Sensor technology plays a crucial role in verifying disinformation by providing objective data and enhancing detection methods. Sensor technology is used for fact-checking and linguistic analysis to monitor online content, including news articles, social media posts, and videos. For instance, for news articles, sensor technology can be used for cross-referencing satellite imagery and weather reports to verify their accuracy.(2) If disinformation aims to manipulate public sentiment, sentiment analysis can be used to find suspicious narratives.(3) When integrated with AI and other tools, sensor technology can contribute to effective disinformation detection and the prevention of the spread of false information.

*Corresponding author: e-mail: cyberpaul@gm.ttu.edu.tw

https://doi.org/10.18494/SAM4866

ISSN 0914-4935 © MYU K.K.
https://myukk.org/

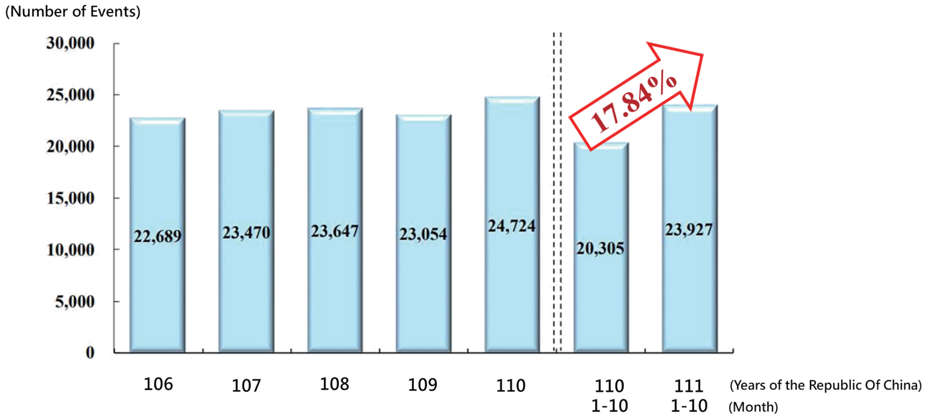

Fig. 1. (Color online) Number of fraud cases from 2017 to 2022.

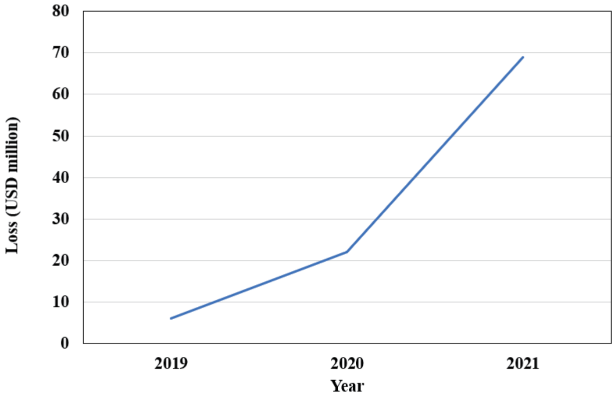

Fig. 2. (Color online) Amount of loss in fraud cases from 2019 to 2021.

We investigated previous systems that detect fraud and disinformation efficiently by comparing the functions of the Thailand Verification Center, Taiwan FactCheck Center, MyGoPen, and Cofacts. On the basis of the results, we constructed a new system integrating a criminal record database to verify disinformation effectively. The constructed system can be integrated with sensor technology in the future for better disinformation screening and preventing related cybercrimes.

2. Technologies for Screening Disinformation

2.1 Support vector machine (SVM)

An SVM model is used to estimate a hyperplane (or decision boundary) with defined attributes and classify data on the hyperplane. There are many options for hyperplanes in n-dimensional space, depending on the number of attributes. In the SVM model, two types of data point are defined to detect disinformation. An extreme learning machine (ELM) is often used with SVM to detect and screen disinformation.[4]

2.2 ELM

ELM is used to classify queries based on lexical attributes. This requires a large amount of processing power because complex data structures used as attributes are highly dimensional. Therefore, the dimension of a feature space can be significantly reduced depending on semantic features. The platform using ELM can classify semantic attribute problems to train SVM models.

2.3 Interplanetary file system (IPFS)

IPFS is a peer-to-peer file-sharing technology characterized by its distributed dash tables. IPFS applies the same file system for all computing devices. It is similar to the global information network as it uses a BitTorrent cluster of nodes for exchanging items. IPFS allows for a high-throughput, content-addressed block storage model and content-related hyperlinks. IPFS combines a distributed hash table and provides block-switching incentives and a self-certified namespace. In IPFS, nodes exchange trustworthy information. By using distributed content delivery, potential Hypertext Transfer Protocol scenarios of a distributed denial of service attack can be prevented. Images or videos published by IPFS are shared in the blockchain network, enabling distributed computing (Fig. 3).[5]

2.4 Disinformation screening

In this study, we used policies and platforms for screening disinformation as shown in Fig. 4. We assumed that government decision-making and platforms based on AI and blockchain technology identify information owners, filter out disinformation, and alert people to disinformation. We constructed a system architecture (Fig. 5) and collected data to test the constructed system for this assumption.

3. Principles for Screening Disinformation

3.1 Thailand Verification Center

According to Thailand's National Development Strategy for "stability, prosperity, and sustainability", people are using social media more than before to obtain and use information online. Smart cities have been constructed under the Four Color Chrysanthemum Smart City

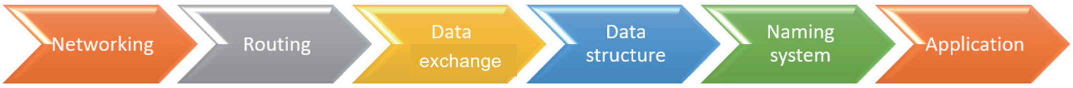

Fig. 3. (Color online) Process of IPFS.

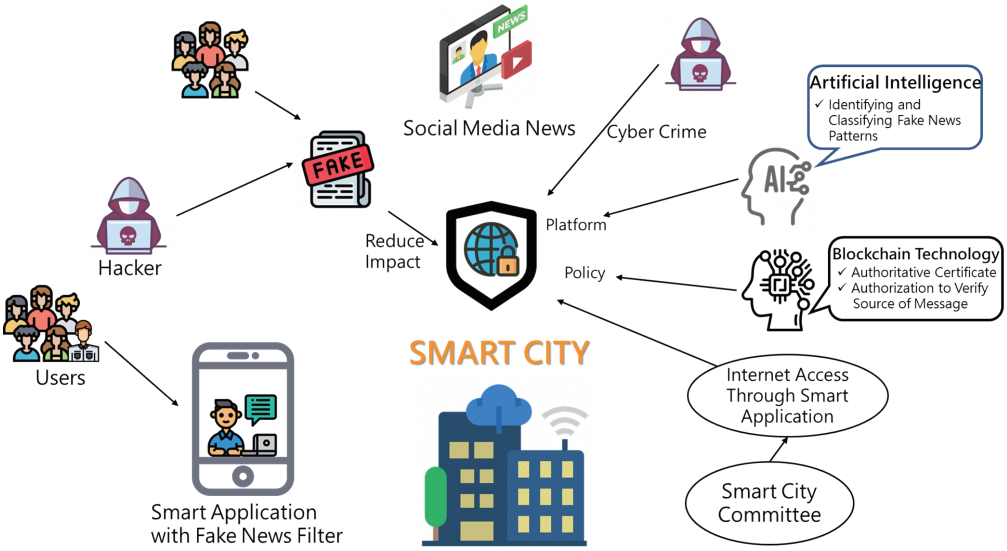
Social Media News
Cyber Crime
Artificial Intelligence
✓ Identifying and Classifying Fake News Patterns
FAKE
Hacker
Reduce Impact
Platform
Policy
Blockchain Technology
✓ Authoritative Certificate
✓ Authorization to Verify Source of Message
Users
SMART CITY
Internet Access Through Smart Application
Smart City Committee
Smart Application with Fake News Filter

Fig. 4. (Color online) Design policy and platform to screen disinformation.

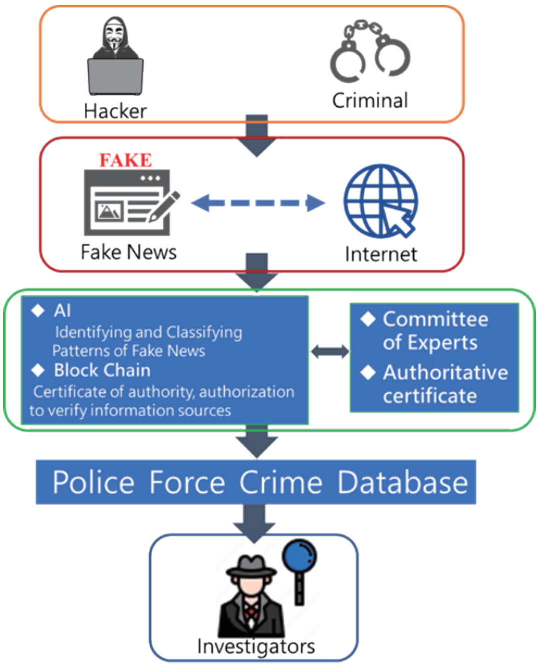
Hacker
Criminal
FAKE
Fake News
Internet
◆ AI
Identifying and Classifying Patterns of Fake News
◆ Block Chain
Certificate of authority, authorization to verify information sources
◆ Committee of Experts
◆ Authoritative certificate
Police Force Crime Database
Investigators

Fig. 5. (Color online) System architecture constructed in this study.

Project in Thailand, where information technology is critical to managing and developing appropriate plans, policies, and practices. Therefore, the use of recent digital technologies is essential, which may also increase the possibility of fraud with disinformation. The Thailand Verification Center provides ‘a fact check’ service based on evidence. The center uses a strict verification methodology with the principles of openness, transparency, prudence, and responsibility. All verifications are made through discussions and meetings, and verification reports are released after being reviewed by three auditors. All fact checks are made public, and if any error is found, the report is corrected immediately. The Thailand Verification Center prevents disinformation from entering people's computers and smartphones using ELM. Blockchain technology is used to store critical information, and part of the information is transferred into a non-fungible token to identify the owner and source of the information. As a result, useful and valid information is acquired. The Thailand Verification Center will adopt quantum computing and new blockchain technologies such as an electronic wallet system.(6)

3.2 Taiwan FactCheck Center

By defining abbreviations and acronyms in headings or titles, unless they are unavoidable, the Taiwan FactCheck Center verifies facts and information under the principles of professionalism, transparency, and impartiality. The negative impact of disinformation is seriously considered to improve information literacy and benefit social development. The Taiwan FactCheck Center was established by the Media Education Watch Foundation and the Quality News Development Association of Taiwan in 2018 and began to issue verification reports.(7) To achieve a long and stable development, the center was transformed into the New Taipei City Taiwan Fact Checker Education Foundation in 2020. The board of directors reviews and discusses issues.(2) The Taiwan FactCheck Center mainly relies on traditional verification methods. It takes a few minutes or longer to verify disinformation. On the basis of evidence, parties, experts, and authorities who disseminate information are verified. The 2018 Kansai airport incident was reviewed by the Taiwan FactCheck Center.(7) There is the issue of public trust, so it is necessary to gain credibility from the public.

3.3 MyGoPen

MyGoPen was founded in 2015 for clean networks without disinformation. It conducts fact-checking in a fair, transparent, and nonpartisan manner and promotes accurate information delivery. It provides fact-checking, rumor-busting, data analysis, and fast-checking robot services.(3) MyGoPen checks documents, arguments, and reviews and provides the results using the LINE Quick Reference Service. Incorrect or suspicious messages in social and mainstream media and misguided messages on policies, articles with inappropriate titles, scams, fake accounts, phishing links, incorrect healthcare knowledge, and rumors are monitored. MyGoPen validates information by using a pictorial search tool such as Google or Yandex to screen the illegal use of images in different contexts.(8) It searches the websites of communities and the sources of text messages on the basis of time-based retrieval filtering. Authorities or local police cooperate to verify the information. For videos, MyGoPen cuts them into multiple frames using a network authentication tool such as InVID retrieve. The origin of reference of the information is investigated for discrepancies. MyGoPen was developed to mitigate the information divide and realize a clean Internet through information sharing. MyGoPen was accredited by the

International Fact-Checking Network, which is committed to building a cleaner Internet environment in Taiwan. In 2020, MyGoPen became Facebook's third-party finding program to prevent the proliferation of disinformation more effectively and develop robust digital literacy. To verify disinformation and malicious rumors, official documents, content, and web pages are monitored.

3.4 Cofacts

Cofacts is a platform that promotes open collaboration with the public. Suspicious information is inspected and discussed between users. Cofacts uses a chatbot for each participant to contribute to screening and monitoring disinformation. Results are forwarded by the LINE chatbot. In its collaborative system, a participant can be an editor and/or publisher of the result.[4] Cofacts provides fact-checking data and technical reports. It is open-sourced and shares data with the public.[9]

4. Construction of System

To construct a new system for screening disinformation, we compared the configurations of the previously mentioned systems. Table 1 presents the comparison results of the Thailand Verification Center, Taiwan FactCheck Center, MyGoPen, and Cofacts. Different from the others, the Thailand Verification Center verifies disinformation using algorithms. The other three systems screen disinformation manually. While the Thailand Verification Center verifies disinformation in online materials such as social media texts and website content, the other systems screen it on the basis of user-generated content. The Thailand Verification Center processes the information fastest.[10]

On the basis of the configurations of the four systems, we developed a new verification system. We applied the cybercrime theory of everyday activity proposed by Cohen and Felson in 1979.[11] The theory is based on studies on motivational and competent offenders , appropriate objects, and the suppression of absenteeism of offenders. We designed the system by considering people, events, time, place, and facts and Locard's exchange principle. To enhance the reliability

Table 1
Comparison of Thailand Verification Center, Taiwan FactCheck Center, MyGoPen, and Cofacts.

Configuration	Thailand Verification Center	Taiwan FactCheck Center	MyGoPen	Cofacts
Type of verification	People/algorithm	People	People	People
Verification method	Advisory committee, AI, blockchain technology	Advisory committee, experts	Editors, experts, readers	Media, users
Message source	Social media, Websites	Whistleblowers, Concern/controversy	Reports, replies	Whistleblowers, Concern/controversy
Verification speed	Fast	Slow	Slow	Intermediate
Platform usage	High	Low	Low	Intermediate
Speed	High	Low	Low	Intermediate

of the system, we analyzed false information on the basis of confirmed cybercrimes and distinguished the sources of disinformation from the “dark figure of crime” such as public reports, police discovery, and victim reports. In screening disinformation, time, place, fact, and possible damage must be considered. At the same time, personal information including the sate of mind, age, family, and different professions must be analyzed to simulate the dissemination of disinformation. We used the deductive method to prove and identify the investigation process (Figs. 6 and 7).[12.13]

To enhance the reliability of the constructed system, the following processes were adopted: crime process analysis, forensic process analysis, and case reports and archiving in the input, process, output process (Fig. 8).[5] In the crime process analysis, after obtaining the relevant information, the 5W1H principle is applied to confirm the elements of crimes and define the investigation process considering people, events, time, places, and objects. The motivation of the crime is also analyzed to confirm the verification process. Forensic process analysis is performed to determine crime elements, collect evidence, and confirm the authenticity of information. Case reports and archiving are conducted after the completion of the verification to review feedback and relevant case reports. A criminal record database is used to refer to similar cases and increase the efficiency of the verification.

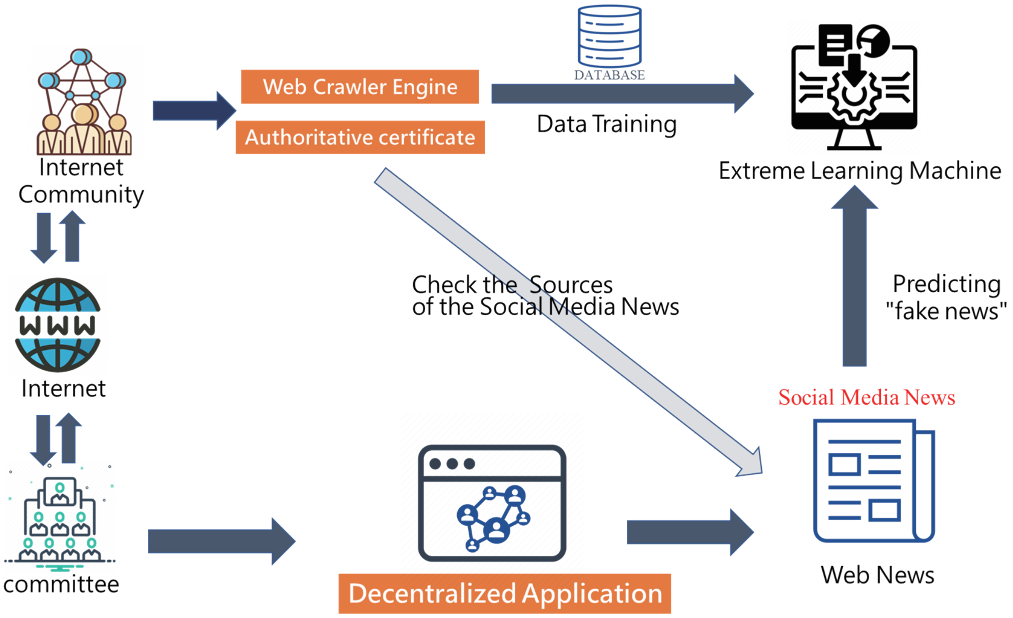

Fig. 6. (Color online) Flowchart of disinformation verification.

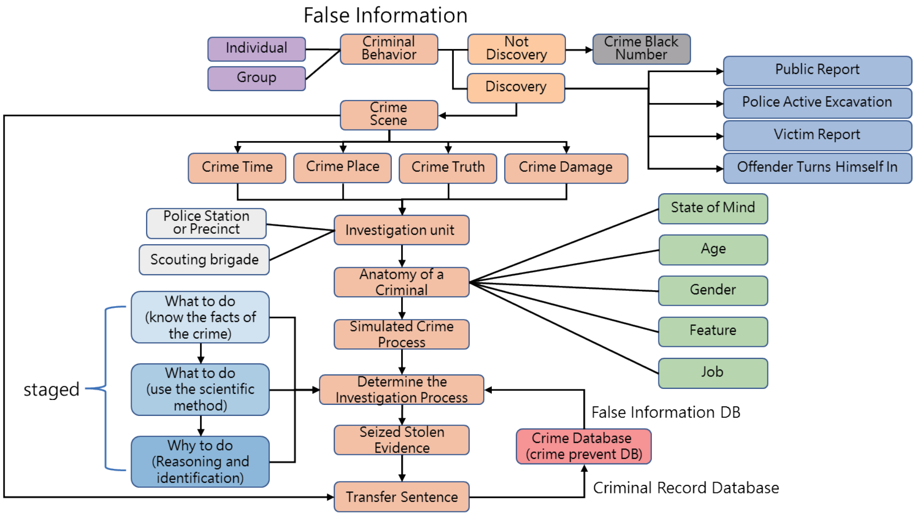

Fig. 7. (Color online) Framework for verification of disinformation of constructed system in this study.

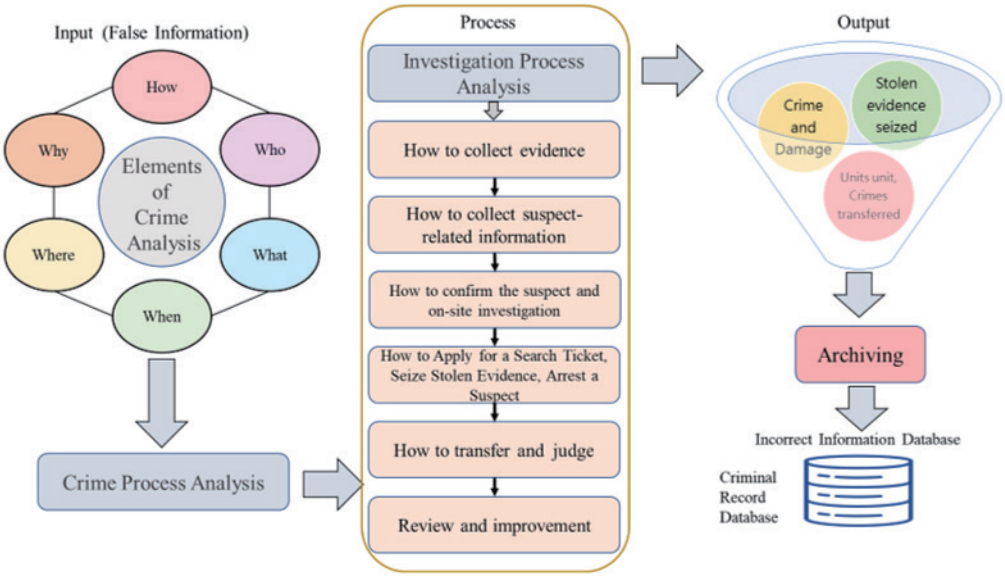
Input (False Information)
How
Why
Who
Elements of Crime Analysis
Where
What
When
Crime Process Analysis
Process
Investigation Process Analysis
How to collect evidence
How to collect suspect-related information
How to confirm the suspect and on-site investigation
How to Apply for a Search Ticket, Seize Stolen Evidence, Arrest a Suspect
How to transfer and judge
Review and improvement
Output
Crime and Damage
Stolen evidence seized
Units unit, Crimes transferred
Archiving
Incorrect Information Database
Criminal Record Database

Fig. 8. (Color online) Process of verifying disinformation of constructed system.

5. Conclusion

To screen disinformation effectively, we constructed a system by reviewing previous systems, namely, the Thailand Verification Center, Taiwan FactCheck Center, MyGoPen, and Cofacts. The constructed system integrated the criminal record database to verify disinformation.

Disinformation misleads the public and harms entities or individuals for economic benefits or political influence on them. The constructed system can be used as a reference to develop similar systems to prevent the dissemination of disinformation.

References

1 Executive Yuan of Taiwan: https://english.ey.gov.tw/News3/9E5540D592A5FECD/df8e5cf1-d2ca-4e5f-83b3-501aeaf1dcfb (accessed April 2024).

2 F. C. C. Santos: J. Media **4** (2023) 679. https://doi.org/10.3390/journalmedia4020043

3 P. Akhtar, A. M. Ghouri, H. U. R. Khan, M. A. ul Haq, U. Awan, N. Zahoor, Z. Khan, and A. Ashraf: Ann. Oper. Res. **327** (2023) 633. https://doi.org/10.1007/s10479-022-05015-5

4 S. K. Rath, M. Sahu, S. P. Das, S. K. Bisoy, and M. Sain: Electronics **11** (2022) 2707. https://doi.org/10.3390/electronics11172707

5 Netskope: https://www.netskope.com/blog/interplanetary-file-system-a-decentralized-place-to-host-phishing-content (accessed April 2024).

6 Special Branch Bureau: https://pcscenter.sbpolice.go.th/en (accessed April 2024).

7 Taiwan FactCheck Center: https://tfc-taiwan.org.tw/ (accessed April 2024).

8 MyGoPen: https://www.mygopen.com/ (accessed April 2024).

9 Cofacts: https://cofacts.tw/ (accessed April 2024).

10 Thailand Forensic Center https://www.book721.com/index2.php (accessed May 2024).

11 ACADEMIA: https://www.academia.edu/8897451/Theorizing_Cybercrime_Applying_Routine_Activities_Theory#:~:text=In%20their%20seminal%20work%2C%20Cohen%20and%20Felson%20%281979%29,of%20targets%20had%20driven%20increases%20in%20property%20crime (accessed May 2024).

12 X. Cao: LNEP **16** (2023) 47. https://doi.org/10.54254/2753-7048/16/20231102

13 K. F. Hyde: Qual. Mark. Res. **3** (2000) 82. https://doi.org/10.1108/13522750010322089

About the Authors

Mu-Chuan Chen received his B.S. degree in innovation management from the National Open University in 2012 and his M.S. degree in information management from Yuanpei University of Medical Technology in 2016. He is a Ph.D. candidate in computer science and engineering at Tatung University. He has been a police officer in Taiwan since 1989. His interests include the prevention and identification of online fraud and disinformation. (muchuan1968@gmail.com)

I-Long Lin received his B.S. degree from Central Police University, Taiwan, in 1983, and his M.S. and Ph.D. degrees from Tamkang University and National Taiwan University of Science and Technology, Taiwan, in 1989 and 1998, respectively. From 1983 to 2011, he was a professor at Central Police University, Taiwan. From 2012 to 2021, he was a professor at Yuanpei University of Medical Technology, Taiwan. Since 2021, he has been a professor at Tatung University. His research interests include digital evidence, forensics, and cybersecurity. (cyberpaul@gm.ttu.edu.tw)

Hung-Cheng Yang received his bachelor's degree in traffic management in 1984 and his master's degree in information management in 2003 from the Central Police University in Taiwan. He has been studying for his Ph.D. degree in computer science and engineering at Tatung University since 2021. Since 1984, he has worked for the police. His research interests include digital evidence, forensics, and traffic accident prevention using big data. (yangyeh5046@gmail.com).