Journal of Harbin Engineering University
ISSN: 1006-7043

Vol 45 No. 3
March 2024

Cyber Strategic Technology Key Topics and Trend Analysis

Seung-Yeon Hwang [1],Jeong-Joon Kim [2*]

[1]Dept. of Computer Engineering, Anyang University, Anyang-si, Gyeonggi-do, Republic of Korea
[2]Corresponding Author, Dept. of Software, Anyang University, Anyang-si, Gyeonggi-do, Republic of Korea

Abstract

In response to evolving societal needs and advancements in science and technology, there is active engagement in research dedicated to forecasting and analyzing future technologies anticipated to arise across various scientific and technological domains. This paper utilizes the titles and summary information of about 4,200 papers published in journals in the field of information security to analyze key topics and trends in strategic technology in the field of cybersecurity among various fields of science and technology. In addition, Latent Dirichlet Allocation (LDA) is used to analyze key topics and keywords of cyber strategic technology, and Dynamic Topic Modeling (DTM) is used to analyze trends. By analyzing it on a yearly basis from 2010 to 2020, promising and unpromising fields could be identified. The results derived from the subject extraction engine developed in this study and the opinions of experts in each field are expected to achieve more reliable results, and the strategic technology of the future cyber strategy field is analyzed through cyber strategy technology prediction model research. Based on this accumulated strategic analysis information, it is expected that it will be able to predict new convergence strategies according to environmental changes and strengthen all of the world's cyber response capabilities and competitiveness.

Keywords:Latent Dirichlet Allocation, Dynamic Topic Modeling, Cyber Strategic Technology, Key Topics and Trend Analysis

1. Introduction

The continuous emphasis on the importance of establishing a basis for national science and technology strategy, through the anticipation of future societal shifts and the identification of emergent technologies, has been noted. In line with the Framework Act on Science and Technology, surveys forecasting science and technology trends have been systematically conducted and released every five years since 1994. The Korea Institute of Science and Technology Planning and Evaluation (KISTEP) issued the [5th Science and Technology Prediction Survey] in 2017[1]. This report anticipates the future societal landscape by considering both internal and external environmental changes. It forecasts and examines technologies expected to surface across all science and technology sectors by 2040, driven by the evolving demands of society and advancements in science and technology. Various methods have been used for technology prediction evaluation and future prospects conducted by research institutes such as KISTEP, such as expert brainstorming [2], Delphi [3], scenario techniques [4]), news or Google trend search, and thesis and patent keyword analysis. Based on this method, future trends are derived, issues expected to arise in the future due to the trend, and social needs necessary to solve major issues or problems discovered and to enable policy and science and technology responses are identified. In particular, the Delphi survey based on the knowledge and experience of experts has established itself as the most important methodology in the current science and technology prediction survey [5]. Research on cybersecurity detection, prevention, and response based on vulnerabilities at home and abroad is being conducted a lot by related agencies and companies, but studies on cybersecurity strategy (attack) types and characteristics analysis are insufficient. Therefore, the purpose of this study is to conduct in-depth research that can analyze the characteristics of each stage of cyber attack based on quantitative analysis according to classification and type of various attack types in the field of cyber strategy.

Most of the cyber strategy analysis and prediction-related technologies are being studied abroad, and there are several cases in Korea, but the level of technology development is low. Therefore, since the cyber strategy field is a very important issue directly related to national security, localization is needed to reduce overseas dependence and secure independent

technology. The results of this study can provide feedback on early detection of vulnerabilities to cyber strategy targets, and furthermore, it can be importantly used as a base technology for predicting new cyber strategy technologies to target those vulnerabilities.

This paper explains the research related to cyber strategy in Chapter 2, and introduces specific research contents in Chapter 3. Chapter 4 explains the research results and concludes in Chapter 5.

2. Related Works and Dataset

In this chapter, core topics and keywords in the field of cyber strategic technology are extracted, and technologies for trend analysis are described.

2.1 LDA (Latent Dirichlet Allocation) [6, 7]

LDA employs topic modeling to uncover hidden topics within a vast corpus of documents, by recognizing the underlying themes present. It operates under the premise that each document encompasses one or more thematic elements. The process involves selecting a specific word within each document and allocating it to the topic that shows the highest probability, based on the combined likelihood of the topic's presence in the document and the specific word's relevance. LDA plays a crucial role in various natural language processing (NLP) tasks, including document classification, summarization, assessing similarities, and evaluating relevance.

The conceptual framework of LDA, depicted as a probabilistic model, illustrates how it determines the presence of topics across documents. As presented in a hypothetical Figure 1, the probabilities for the terms 'gene,' 'dna' and 'genetic' to appear under a specific (yellow) topic are 0.04, 0.02, and 0.01, respectively. This suggests the topic's association with genetics. If the frequency of words related to this (yellow) topic surpasses those linked to other (blue and red) topics within a document, it implies that the document predominantly discusses genetics-related themes.

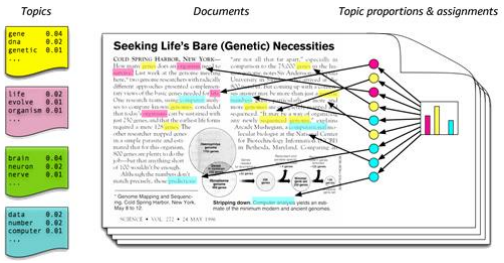

Fig. 1. LDA schematic diagram

2.2 DTM (Dynamic Topic Modeling) [8, 9]

Following the advent of LDA for topic analysis, its application broadened to encompass the examination of academic texts. This expansion paved the way for its utilization in detecting trends across various fields, highlighted by methodologies such as Topic Over Time (TOT) and Dynamic Topic Model (DTM). DTM, in particular, stands out as a dynamic approach for tracing the evolution of topics within a document corpus over time. Taking LDA as an example, imagine analyzing a decade's worth of documents from 2010 to 2019. Initially, the documents are segmented into annual batches for a detailed year-by-year topic evolution study. By designating 10 topics per year, LDA could potentially unveil 100 distinct topics across the decade. However, this approach encounters a limitation in trend analysis due to the uniqueness of each topic. For effective trend analysis, it's crucial to demonstrate how topics are sequentially interconnected through the years, ensuring that a topic in one year bears resemblance to its counterpart in the subsequent year. DTM addresses this challenge by enabling an analysis of how topics and keywords evolve over time, factoring in the continuity of topics year over year, thus facilitating a more nuanced understanding of thematic progression and keyword dynamics.

3. Research Contents

In this paper, research was conducted to analyze key technology elements for analyzing core topics and trends of cyber strategy technology and to identify and classify the characteristics of cyber strategy technology. Figure x shows the architecture for analyzing the core topics and trends of cyber strategic technology.

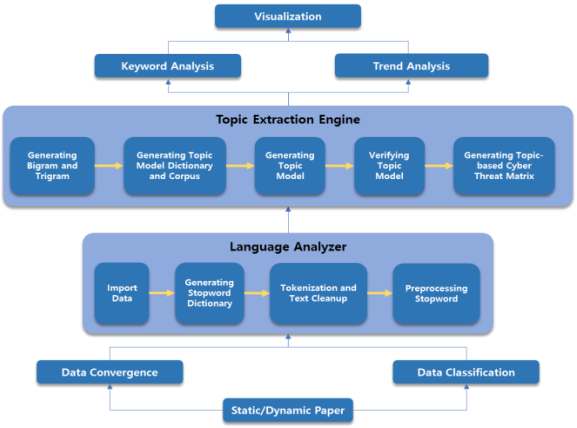

Fig. 2. Cyber Strategic Technology Key Topics and Trend Analysis Architecture

Table 1. Key topics of cyber strategy technology and trend analysis technology elements

Number	**Skill**	Content
1	**Data Convergence**	It refers to a module that fuses data collected for extracting key keywords and topics into a single file unit.
2	**Data Classification**	It refers to a module that classifies data by period and topic for core keyword and topic extraction.
3	**Import Data**	It refers to a module for loading preprocessing data fused and classified in a language analyzer.
4	**Generating Stopword Dictionary**	It refers to a module for constructing an idiom dictionary by identifying words that frequently appear in the collected data but are not necessary for analysis.
5	**Tokenization and Text Cleanup**	It refers to a module for decomposing each sentence into a word list from the loaded data and removing blank characters and punctuation marks.
6	**PreprocessingStopword**	It refers to a module for removing words unnecessary for extracting key keywords and topics by using the constructed terminology dictionary.
7	**Generating Bigram and Trigram**	It refers to a module for identifying and converting Bigram (two-year word) and Trigram (three-year word) for various types of words existing in the collected data sentence.
8	**Generating Topic Model Dictionary and Corpus**	It refers to a module that assigns a unique ID to each word present in the collected data and generates a mapping corpus in the form of (word ID, frequency).
9	**Generating Topic Model**	It refers to a module that generates a topic model by setting hyperparameters necessary for topic extraction based on the topic model dictionary and corpus.
10	**Verifying Topic Model**	It refers to a module that measures the accuracy and consistency of a model in order to confirm the appropriateness of the generated topic model.
11	**Generating Topic-based Cyber Threat Matrix**	It refers to a module for generating a cyber threat matrix based on the generated topic.
12	**Keyword Analysis**	It refers to a module for extracting key keywords in the topic

		through statistical analysis.
13	**Trend Analysis**	It refers to a module for analyzing and predicting key keywords and topic trends through time series analysis.
14	**Visualization**	It refers to a module for intuitively understanding and efficiently delivering extracted core keywords and topic trends.

3.1 Dataset and Experimental Scenario

This research utilized the titles and abstracts of approximately 4,200 articles released between 2010 and 2020. These publications are from the four leading organizations in the cybersecurity arena, namely ACMCCS, USENIX Security, IEEE Security & Privacy, and NDSS, renowned for their significant influence and authority. Figure x illustrates the organizational scheme of the dataset curated for the analysis of principal topics, keywords, and their evolution over time.

Various types of experiments can be conducted through the system developed in this study, and in this study, the total number of topics was designated as 10 to analyze the keywords and trends of each topic, and analysis was performed on a yearly basis from 2010 to 2020.

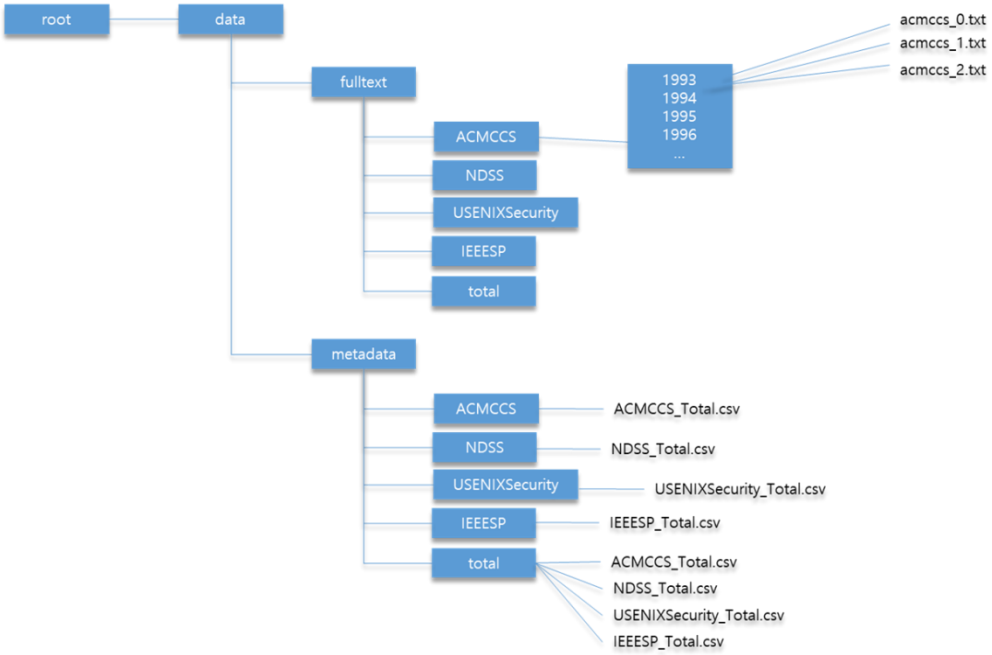

Fig. 3. Key topics of cyber strategy technology and data directory structure required for trend analysis

4. Research Results

4.1 Cyber Strategic Technology Key Topic and Kewords Analysis Result

Visualization data is generated to confirm key keyword information for each topic through keywords assigned to topics and topics, topics and keyword information assigned to papers, topic matrix, and lambda value changes. The format of the result data is shown in Table 2.

Table 2.The result data of the topic extraction engine.

Result data	Format
Keywords assigned to topic and topic	CSV
Topic and keyword information assigned to the paper	CSV
Topic Matrix	CSV
Visualization data	HTML

Since the method of interpreting the topic modeling results is similar by year, this paper explains the

analysis results of the last year in detail. Table 3 shows some of the keyword data allocated to the topic and topic of 2020.

Table 3.Shows some of the keyword data allocated to the topic and topic of 2020.

index_pos	topic_id	Freq	Total	Category	logprob	loglift	Relevance
1019	strategy	16.3474	32.69962	Topic1	-5.2529	1.3137	-1.9696
140	vulnerability	34.65463	148.1766	Topic1	-4.5015	0.5541	-1.9737
653	payment	10.35611	14.34599	Topic1	-5.7094	1.6811	-2.01415
3182	crash_report	7.788405	8.515328	Topic1	-5.9943	1.9178	-2.03825
4447	merchant	7.788368	9.322019	Topic1	-5.9943	1.8273	-2.0835
4117	circumvention	8.644312	11.617	Topic1	-5.89	1.7115	-2.08925
867	algorithm	12.92366	26.03654	Topic1	-5.4879	1.3066	-2.09065
3202	flash	6.932489	7.659412	Topic1	-6.1107	1.9073	-2.1017

Table 4 shows some of the topic and keyword data allocated to the paper of 2020.

Table 4.Some of the topic and keyword data assigned to the paper (2020)

index	topicId	topicWeight	topicWords	docId	year
0	Topic8	0.967616022	allegation, trusted, poisoning, twitter, label, path, sealer, tkperm, package, escrow	IEEEsp2020_7.pdf	2020
1	Topic3	0.964607	access, speech, attack, service, webassembly, device, macao, keyvalue, oram, delegation	IEEEsp2020_8.pdf	2020
2	Topic9	0.964965999	advice, droplet, filtering, netwarden, database, packet, besfs, seal, covert, intelligence	IEEEsp2020_9.pdf	2020
3	Topic2	0.983038008	cache, extension, server, device, packed, defense, packing, chaperone, blacklist, compliance	IEEEsp2020_10.pdf	2020
4	Topic1	0.979005992	strategy, vulnerability, payment, crash_report, merchant, circumvention, algorithm, flash, card, attack	IEEEsp2020_11.pdf	2020
5	Topic5	0.963451982	coverage, collision, stealthy, impact, investigation, document, protection, accident, vulnerability, malware	IEEEsp2020_12.pdf	2020
6	Topic4	0.969924986	unpublished_manuscript_inclusion_upcoming, prospective_author_requested_submit, event_described_paper, manager,	IEEEsp2020_15.pdf	2020

			keyless_entry, project, genetic, custos, phmon, finauth		
7	Topic1	0.977199018	strategy, vulnerability, payment, crash_report, merchant, circumvention, algorithm, flash, card, attack	IEEEsp2020_14.pdf	2020

Table 5 shows the topic matrix for 2020.

Table 5.Topic Matrix (2020)

topic_id	topic_words
Topic1	strategy, vulnerability, payment, crash_report, merchant, circumvention, algorithm, flash, card, attack
Topic2	cache, extension, server, device, packed, defense, packing, chaperone, blacklist, compliance
Topic3	access, speech, attack, service, webassembly, device, macao, keyvalue, oram, delegation
Topic4	unpublished_manuscript_inclusion_upcoming, prospective_author_requested_submit, event_described_paper, manager, keyless_entry, project, genetic, custos, phmon, finauth
Topic5	coverage, collision, stealthy, impact, investigation, document, protection, accident, vulnerability, malware
Topic6	firmware, isolation, grid, revocation, inverter, system, preview, attack, sender, email
Topic7	skill, functional, dongle, uiscope, block, mining, ecommerce, conversation, unsolicited, privacysensitive
Topic8	allegation, trusted, poisoning, twitter, label, path, sealer, tkperm, package, escrow
Topic9	advice, droplet, filtering, netwarden, database, packet, besfs, seal, covert, intelligence
Topic10	plane, censor, upload, proxy, nonce_leakage, vulnerability, fuzzing, ecdsa, enclave, trip

Figure 4 shows the topic and keyword visualization data of the 2020 paper.

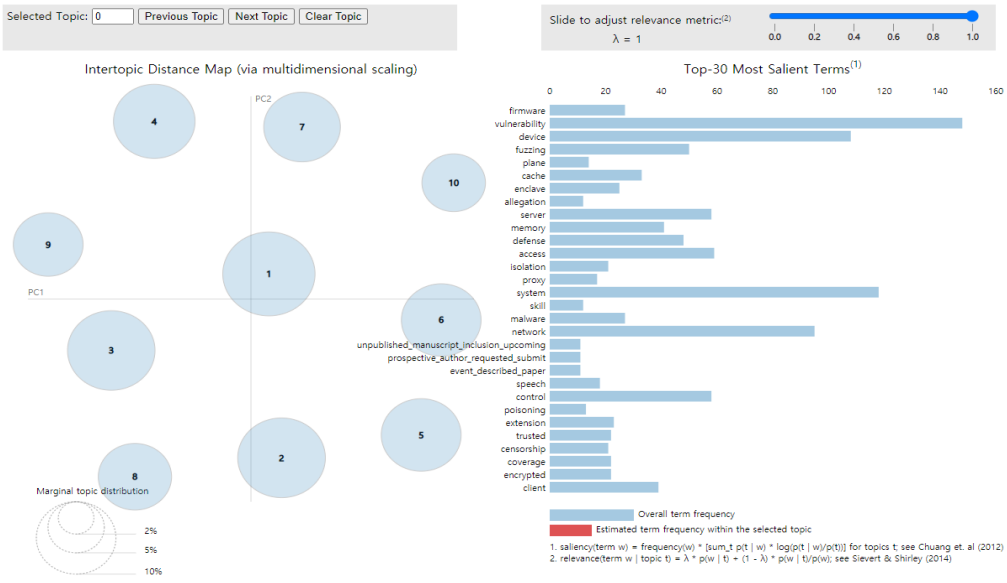

Fig. 4.Visualization results of topics and keywords

The circle output to the left of Figure 4 means each topic, and when selected, the keyword assigned to the topic is output to the right. In addition, the value can be adjusted using the lambda value adjustment bar in the upper right corner, and the order of the keywords output varies depending on the lambda value. Figure 5 shows keyword information when the lambda value of 2020 topic 1 is 1.

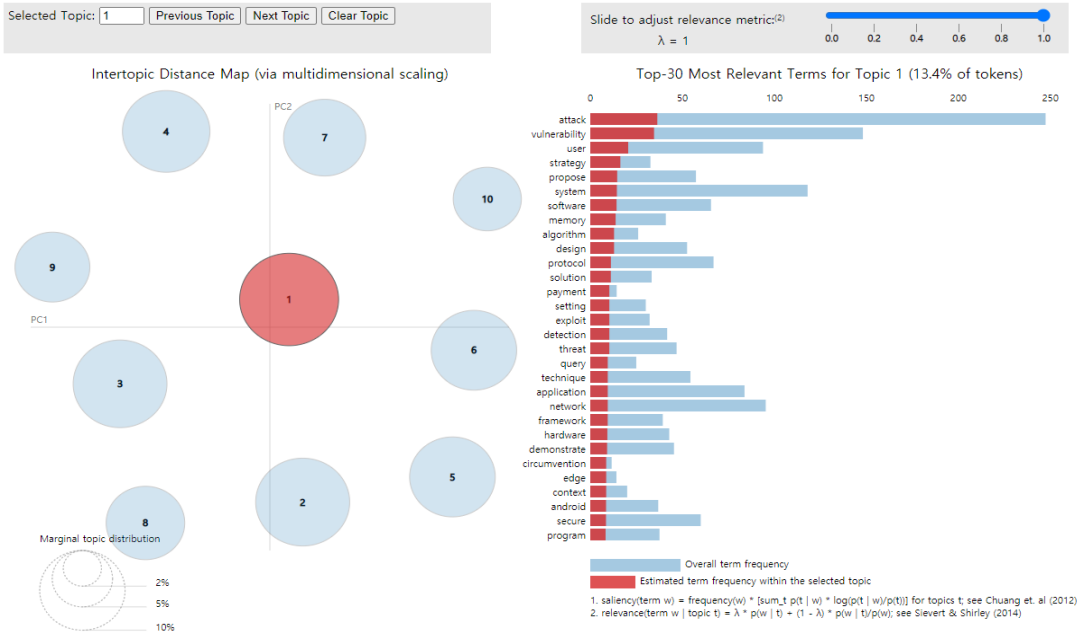

Fig. 5.When lambda is 1, keyword information in topic 1 (2020)

Figure 6 shows keyword information when the lambda value of 2020 topic 1, is 0.5.

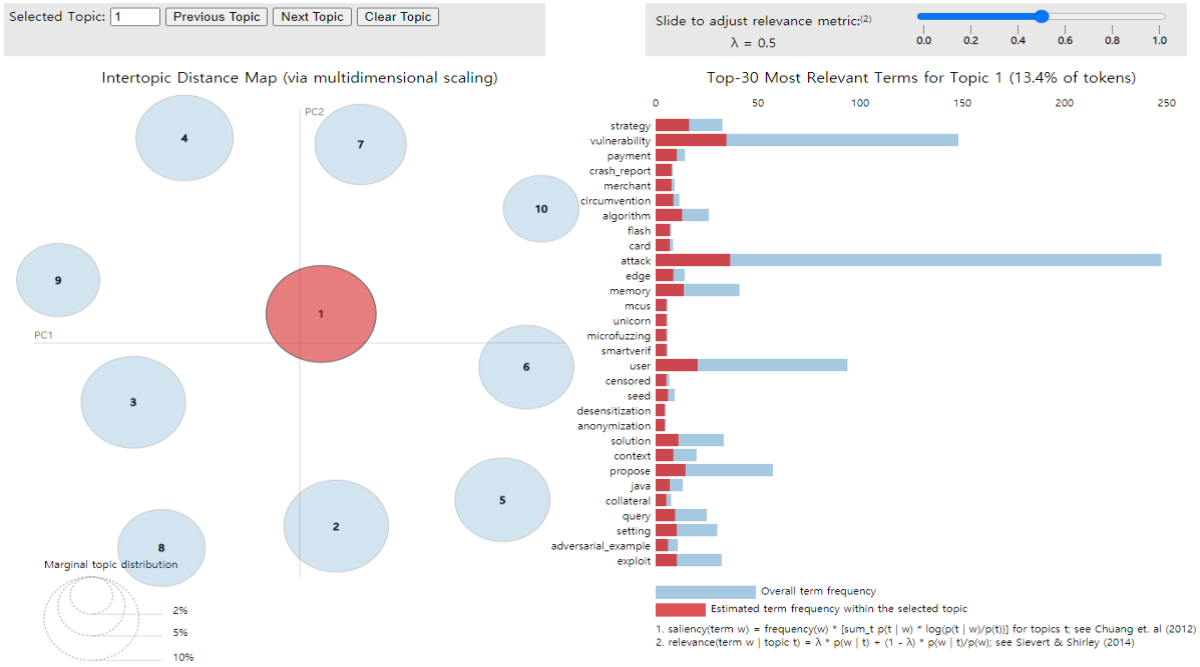

Fig. 6.When lambda is 0.5, keyword information in topic 1 (2020).

By adjusting the lambda value as shown in Figures x and x, keywords assigned to each topic may be identified for each year. As the lambda value approaches 0, a keyword specialized for the topic is derived, and as it approaches 1, a general keyword is derived. That is, a keyword having a large influence on the corresponding topic may be extracted by adjusting the lambda value.

Table x shows the results of 2020 core keyword analysis, and each attribute consists of topics, keywords, and estimation fields. The topic attribute refers to a total of 10 topics from Topic 1 to 10, which appeared in 2020. The keyword attribute represents a keyword appearing in the topic matrix among the result data of the topic extraction engine. The estimation field refers to a field that is estimated to belong to the topic by synthesizing keyword information allocated to the topic and topic, topic and keyword information allocated to the paper, and visualization data centering on the topic matrix results. The estimation field was determined based on the author's subjective opinion, and after that, more accurate and specific fields can be determined by synthesizing the opinions of experts by information security field.

Table 6. Results of analysis of key topics and keywords (2020)

Topic	Keyword	Presumption field
T1	strategy, vulnerability, payment, crash_report, merchant, circumvention, algorithm, flash, card, attack	SW/HW security
T2	cache, extension, server, device, packed, defense, packing, chaperone, blacklist, compliance	Malicious code, home city security, fintech security
T3	access, speech, attack, service, webassembly, device, macao, keyvalue, oram, delegation	Bio-recognition, cloud security
T4	unpublished_manuscript_inclusion_upcoming, prospective_author_requested_submit, event_described_paper, manager, keyless_entry, project, genetic, custos, phmon, finauth	Data security, authentication/authorization technology
T5	coverage, collision, stealthy, impact, investigation, document, protection, accident, vulnerability, malware	Malicious code, AI security
T6	firmware, isolation, grid, revocation, inverter, system, preview, attack, sender, email	Applied security, SW/HW security
T7	skill, functional, dongle, uiscope, block, mining, ecommerce, conversation, unsolicited, privacysensitive	Network security, fintech security
T8	allegation, trusted, poisoning, twitter, label, path, sealer, tkperm, package, escrow	Cryptographic technology
T9	advice, droplet, filtering, netwarden, database, packet, besfs, seal, covert, intelligence	Network and database security
T10	plane, censor, upload, proxy, nonce_leakage, vulnerability, fuzzing, ecdsa, enclave, trip	Threat analysis and control, ship, ocean, aviation security

4.2 Cyber Strategic Technology Keyword Trend Analysis Results

Like the topic extraction engine, the result data of the trend analysis system generates visualization data that allows users to check key keyword information for each topic through topic and topic allocated to topics, topic and keyword information assigned to papers, topic metrics, and lambda value change. On the other hand, since the trend analysis system uses DTM, topics 1 to 10 for each year are similar. Table x shows the results of core keyword analysis for each topic from 2010 to 2020, derived using DTM, and each attribute consists of topics, keywords, and estimation fields. The topic attribute refers to a total of 10 topics from Topic 1 to Topic 10, which appeared from 2010 to 2020. The keyword attribute represents a keyword appearing in the topic matrix among the result data of the trend analysis system. The estimation field refers to a field that is estimated to belong to the topic by synthesizing keyword information allocated to the topic and topic, topic and keyword information allocated to the paper, and visualization data centering on the topic matrix results. The estimation field was determined based on

the researcher's subjective opinion, and after that, more accurate and specific fields can be determined by synthesizing the opinions of experts in each field of information security.

Table 7.As a result of analyzing key keywords by topic from 2010 to 2020.

Topic	Keyword	Presumption field
T1	computation, efficient, protocol, secure, proof, key, cryptographic, encryption, implementation, scheme, server, encrypted, query	암호기술
T2	adversary, channel, cache, attacker, network, defense, attack, timing, packet, countermeasure, authentication, vulnerable, victim	유선네트워크보안, 무선네트워크보안, 악성코드
T3	software, policy, coverage, tool, flow, memory, fuzzing, bug, vulnerability, library, application, program, kernel	보안취약성
T4	binary, detection, sample, detecting, model, classifier, training, accuracy, spam, malware, classification, detect, technique, domain, malicious	악성코드
T5	smart_contract, password, website, user, payment, internet, account, bitcoin, site, content, online, service, transaction	핀테크보안
T6	author, organization, policy, challenge, technology, risk, issue, trust, science, cybersecurity, development, article, community, system, cyber	위협분석및관제
T7	android, device, image, permission, smartphone, firmware, mobile, location, access, smartphones, embedded	홈시티보안
T8	enclave, intel, phone, distance, dnns, voice, manager, processor, adversarial_example, hardware, trusted, audio	인증/인가기술
T9	router, anonymity, advertisement, workshop, society, prospective_author_requested_submit, relay, network, networking, routing, unpublished_manuscript_inclusion_upcoming, paper, node, event_described_paper	응용보안
T10	storage, auditing, forensic, provider, forensics, cloud, graph, monitoring, cloud_computing, audit, provenance, event	클라우드보안

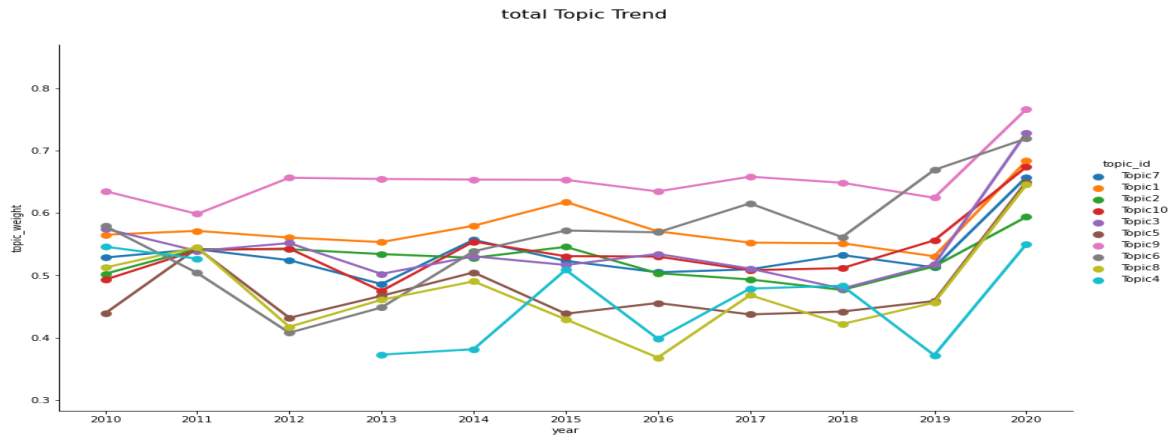

Fig. 7.Topic trends (2010~2020)

Most topics are on the rise over time. It can be seen that topic 1 belonging to the field of cryptographic technology is popular on average. It can be seen that Topic 2, which belongs to the fields of wired network security, wireless network security, and malicious code, decreased slightly until 2018 and then rose in 2019 and 2020. It can be seen that Topic 3, which belongs to the field of software and hardware security vulnerability, steadily appeared every year and then rose in 2019. Topic 4, which belongs to the field of malicious code, did not appear after 2011, but reappeared in 2013 and repeatedly rose and fall. It can be seen that Topic 5, which belongs to the fintech security field, rose in 2019 after repeating the rise and fall. Topic 6, which belongs to the field of threat analysis and control, decreased until 2012 and then steadily increased until 2020. It can be seen that Topic 7, which belongs to the home city security field, steadily appeared every year and then rose in 2019. It can be seen that Topic 8, which belongs to the field of authentication/authorization technology, rose in 2019 after repeating the rise and fall. It can be seen that Topic 9, which belongs to the field of applied security, appeared a lot every year and then rose further in 2019. Topic 10, which belongs to the cloud security field, decreased slightly in 2013, then rose again, and increased sharply from 2018. Figure 8 shows the trend of keyword assigned to topic 1 from 2010 to 2020.

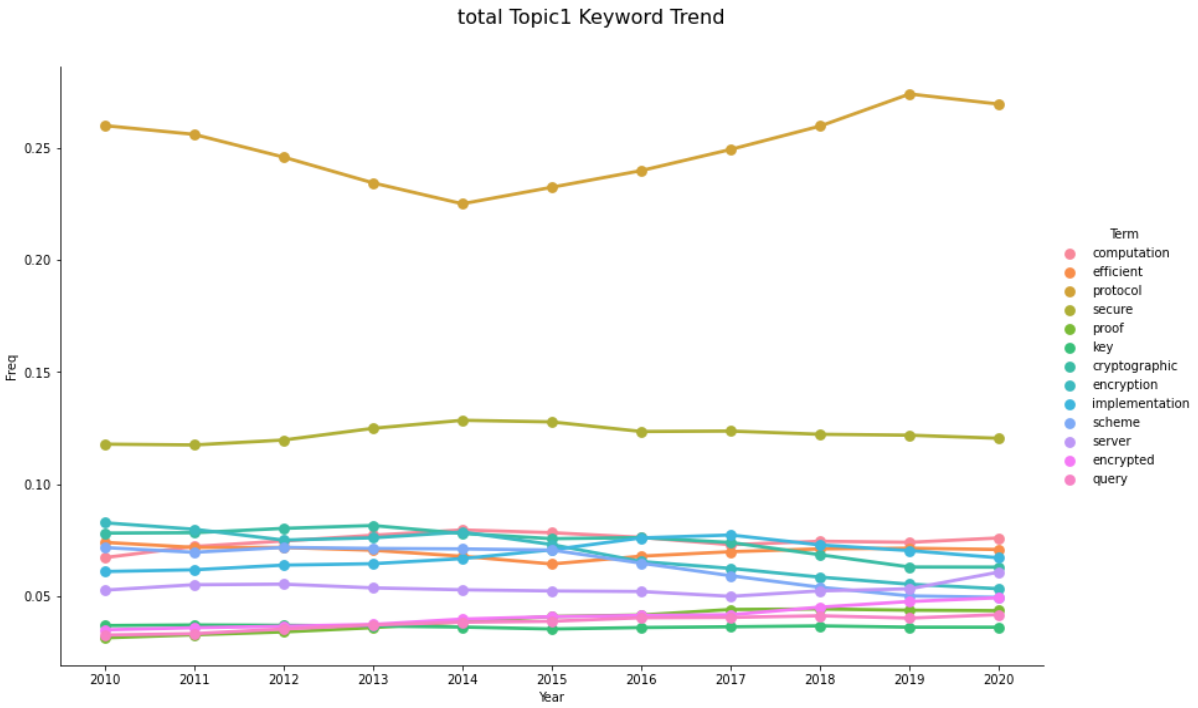

Fig. 8.Topic 1's keyword trend (2010-2020)

The protocol keyword decreased until 2014 and then increased again, and the secure keyword appeared at a constant frequency every year. Other keywords except the protocol keyword and the secure keyword appeared at a constant low frequency every year. Figure 9 shows the trend of keywords assigned to topic 2 from 2010 to 2020.

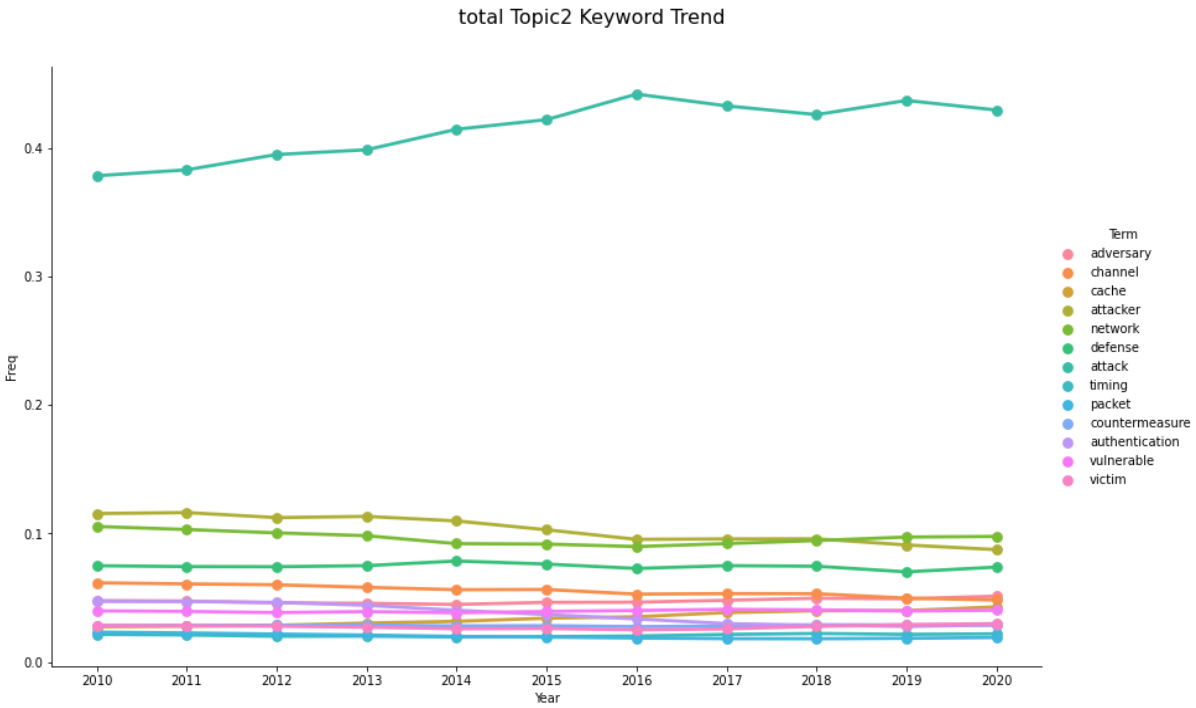

Fig. 9.Topic2's keyword trend (2010-2020)

Defense keywords are on a slight increase, and other keywords have appeared at a constant low frequency every year. Figure 10 shows the trend of keywords assigned to topic 3 from 2010 to 2020.

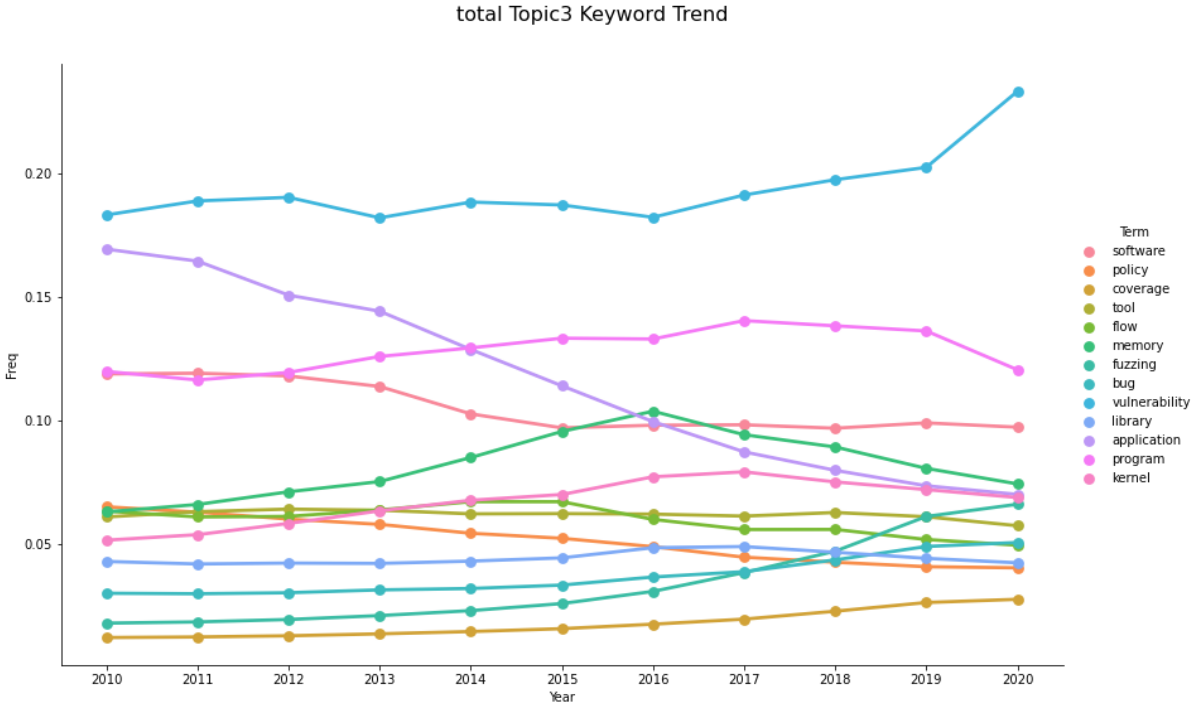

Fig. 10.Table 3 keyword trend (2010-2020)

The vulnerability keyword has appeared at a high rate every year and is steadily increasing. Application keywords appeared at a high rate, but are steadily decreasing. Software keywords appeared at an intermediate rate and maintain a constant rate every year. Memory keywords gradually increased until 2016 and then decreased. Other keywords have a low frequency of appearance, but they tend to increase slightly. Figure 11 shows the trend of keywords assigned to topic 4 from 2010 to 2020.

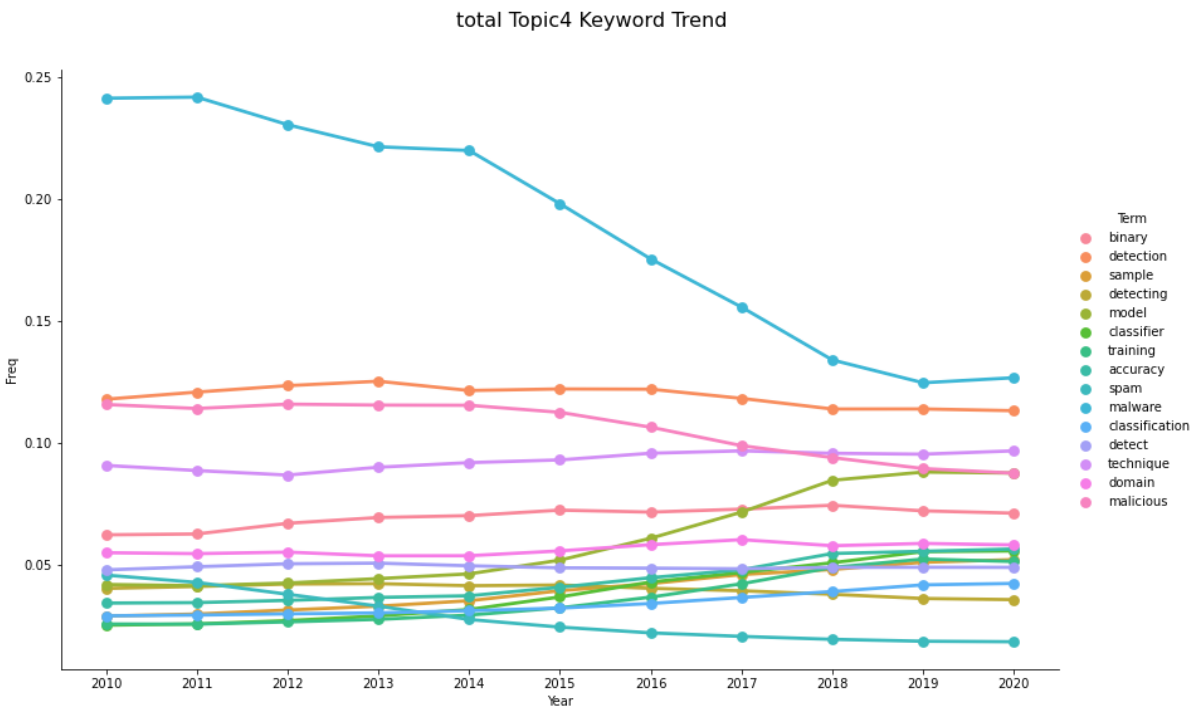

Fig. 11.Table 4 keyword trend (2010-2020)

Malware keywords appeared at high frequencies in 2010 but have steadily decreased since then. The detection keyword appeared at a constant rate every year. It can be seen that the frequency of appearance of model keywords was low until 2015 and then increased. In addition, keywords appeared at a constant low frequency every year. Figure 12 shows the trend of keywords assigned to topic 5 from 2010 to 2020.

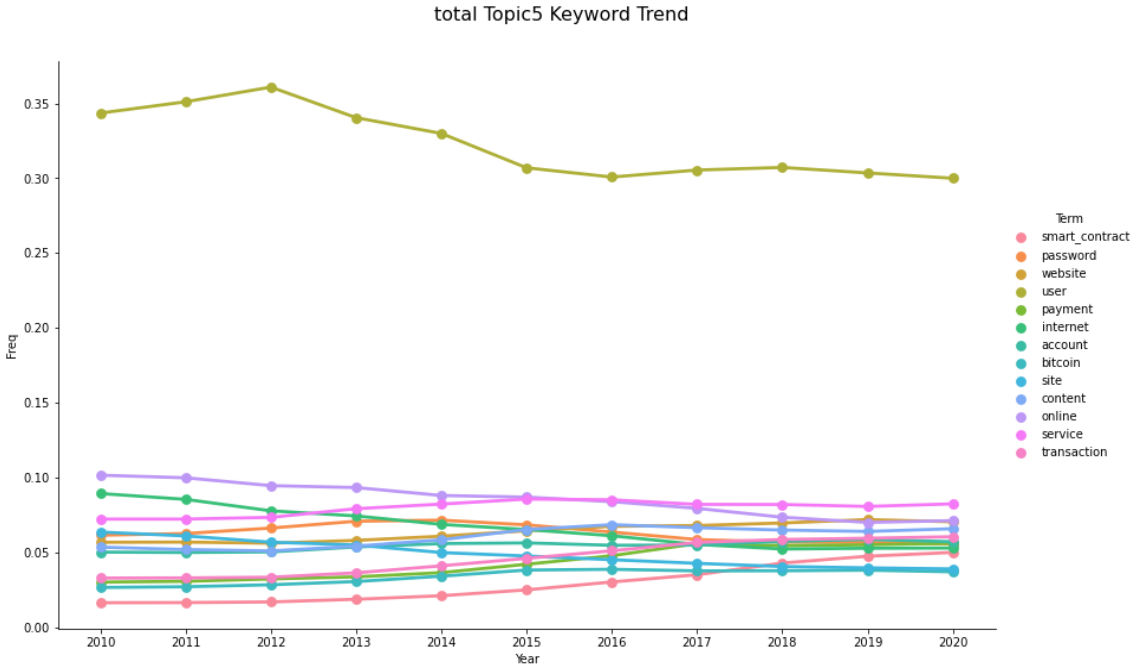

Fig. 12.Table 5 keyword trend (2010-2020)

The user keyword appears at a high frequency every year,but shows a trend of decreasing little by little. Other keywords appeared at a constant low frequency every year. Figure 13 shows the trend of keywords assigned to topic 6 from 2010 to 2020.

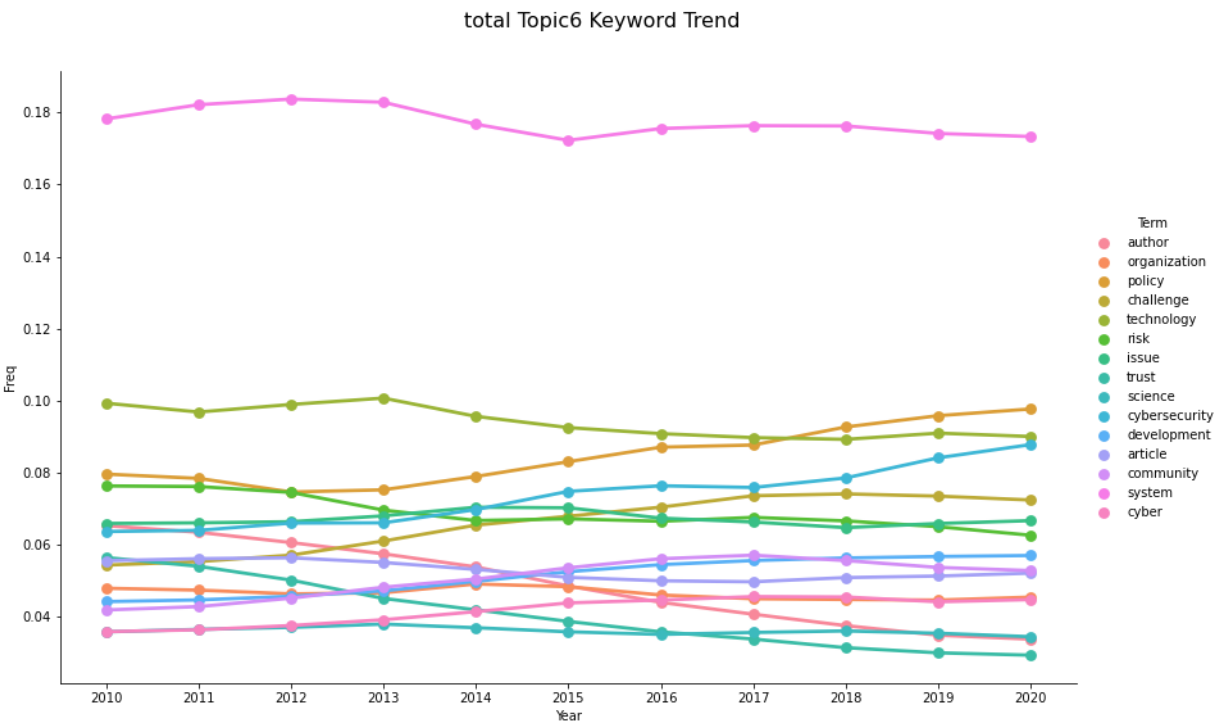

Fig. 13.Topic 6 keyword trend (2010-2020)

It can be seen that the system keyword appears constantly at a high frequency every year. Policy keywords and cybersecurity keywords appear at low frequencies every year, but they increase little by little. Science keywords appear at low frequencies every year and are steadily decreasing. Other keywords appear constantly at a low frequency every year. Figure 14 shows the trend of keywords assigned to topic 7 from 2010 to 2020.

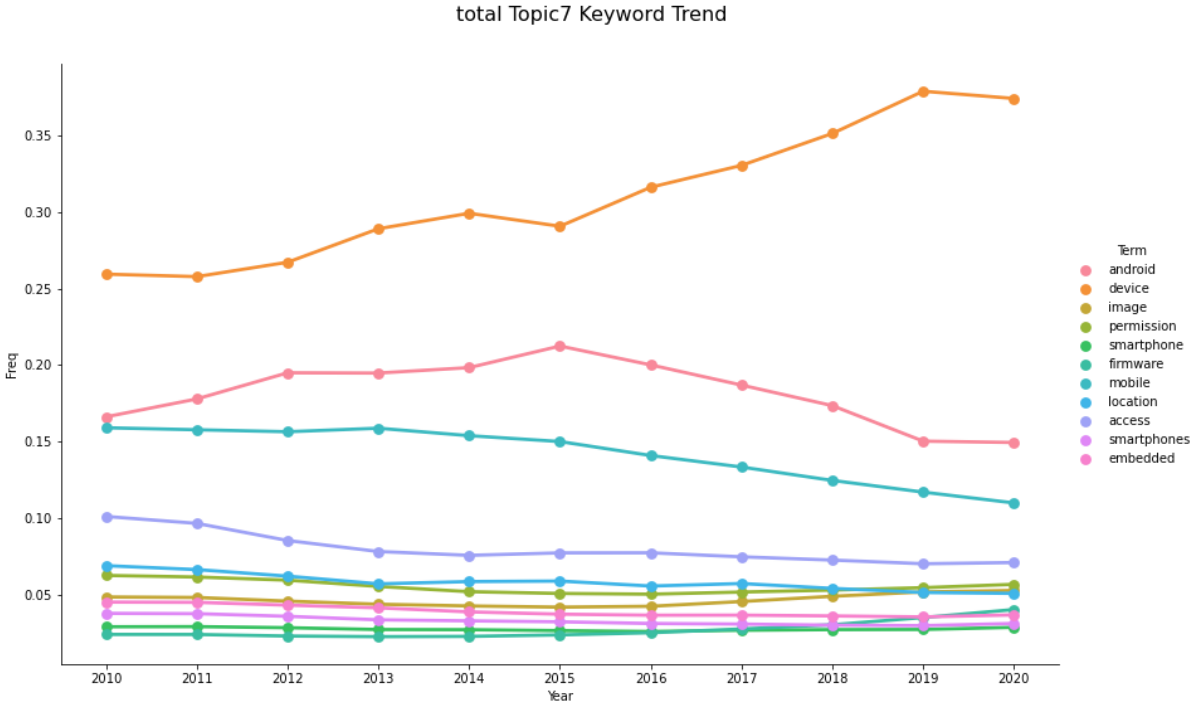

Fig. 14.Topic 6 keyword trend (2010-2020)

Device keywords have been on the rise every year since 2010. The android keyword increased slightly until 2015 and has since declined. Firmware keywords are on the decline every year. Other keywords appear constantly at a low frequency every year. Figure 15

shows the trend of keywords assigned to topic 8 from 2010 to 2020.

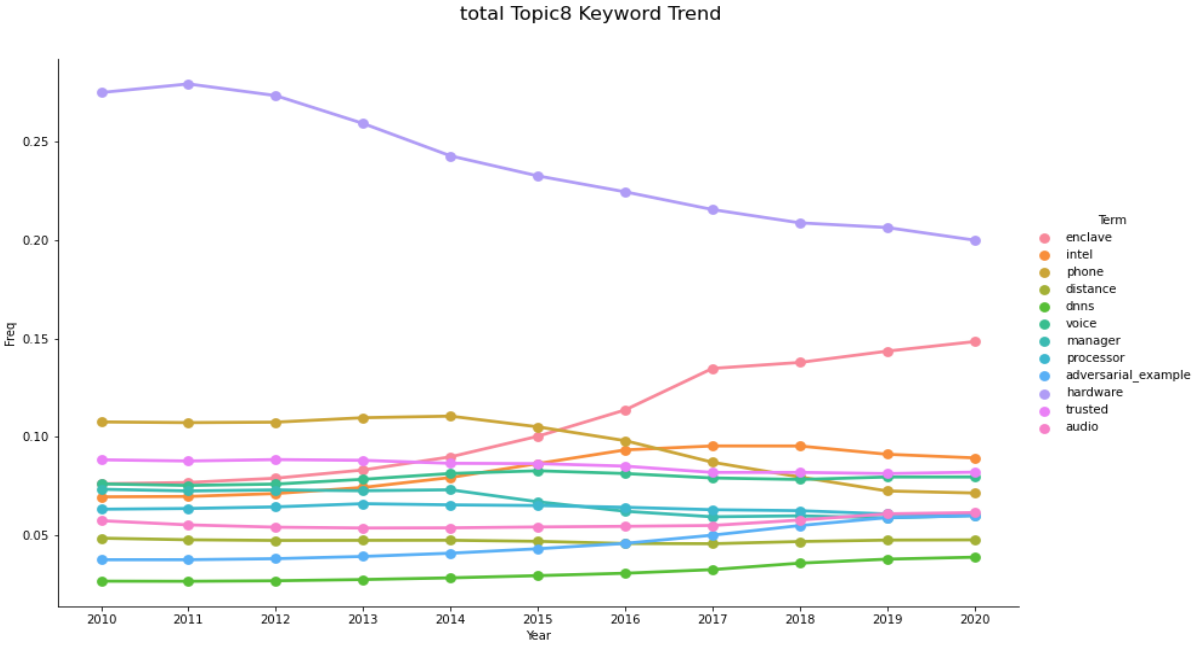

Fig. 15.Topic 8 keyword trend (2010-2020)

Hardware keywords appeared at high frequencies in 2010 and have since shown a steadily decreasing trend. The enclave keyword appeared at a low frequency in 2010 and then showed a steadily increasing trend. Other keywords appear constantly at a low frequency every year. Figure 16 shows the trend of keywords assigned to topic 9 from 2010 to 2020.

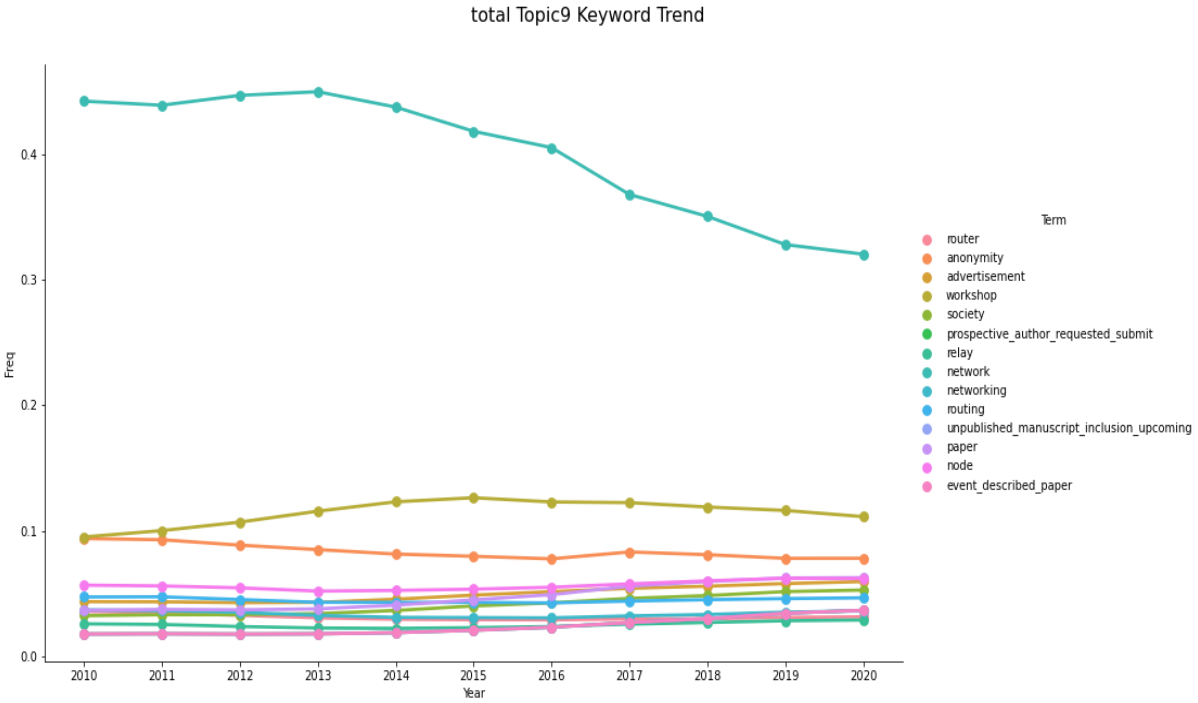

Fig. 16.Topic 9 keyword trend (2010-2020)

The network keyword appeared at a high frequency in 2010 and has since shown a steadily decreasing trend. Other keywords appear constantly at a low frequency every year. Figure 17 shows the trend of keywords assigned to topic 10 from 2010 to 2020.

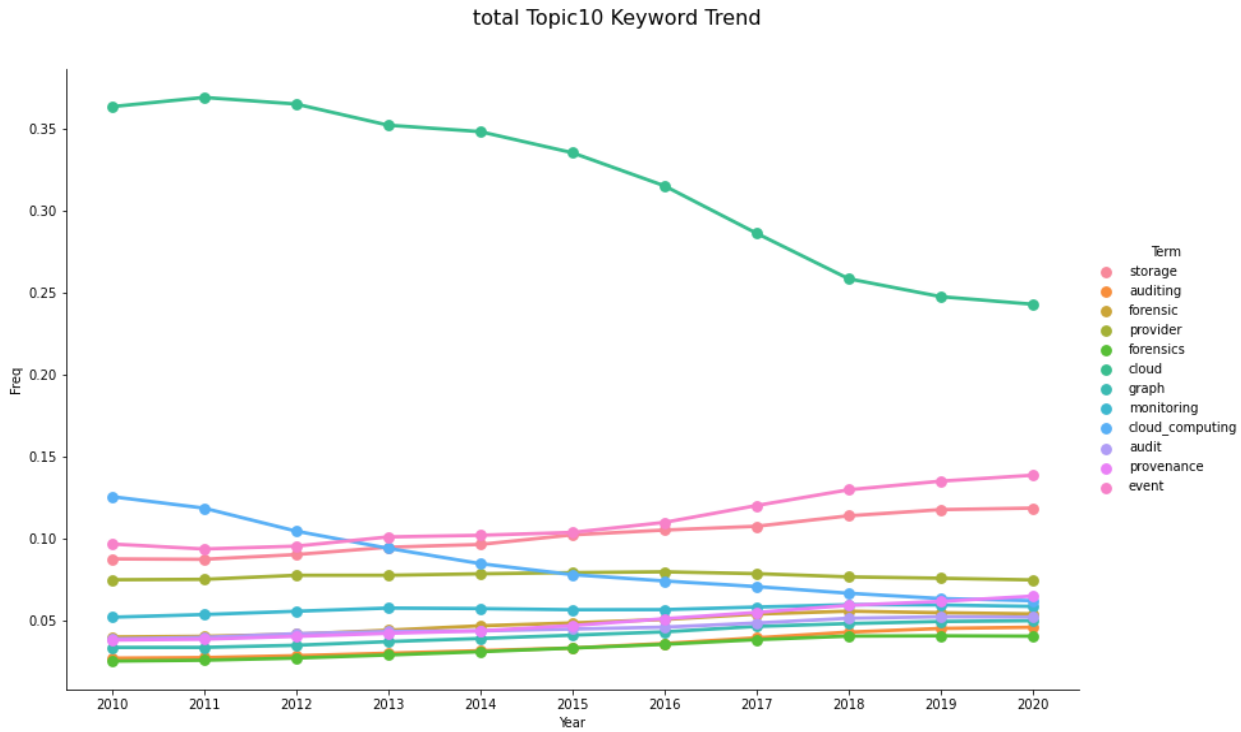

Fig. 17.Topic 10 keyword trend (2010-2020)

Cloud keywords appeared at a high frequency in 2010 and have shown a steadily decreasing trend since then. The probability keyword and storage keyword appear at a low frequency and are slightly increasing.

5. Conclusion

This paper presents an annual analysis of essential topics, keywords, and trends by examining the titles and abstracts of roughly 4,200 articles from the key information security societies (ACMCCS, USENIX Security, IEEE Security, NDSS) over an eleven-year period (2010-2020). Through this study, key topics and keywords of cyber strategic technology were identified, and promising fields were identified by analyzing the trends of key topics and keywords of cyber strategic technology.

The integration of data outcomes from the topic extraction engine developed in this study with insights from domain experts is anticipated to yield more dependable findings. Through research on cyber strategy technology prediction models, it is expected to strengthen global cyber warfare response capabilities and competitiveness by analyzing strategic technologies in the future cyber strategy field and predicting new convergence strategies according to environmental changes based on such accumulated strategic analysis information. By deriving basic analysis data on core cyber strategy technologies, it will be possible to increase the future prediction capability of cyber strategy technology by establishing a leading cyber strategy field information collection policy. The findings of this research are anticipated to serve as a foundational technological framework for creating AI-driven cybersecurity strategy systems.

Acknowledgement

This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIT) (No. 2022R1F1A1062953).

References

[1] Korea Institute of Science and Technology Evaluation and Planning (2017), The 5th Science and Technology Forecasting Report. DOI: https://doi.org/10.23000/TRKO201800037414

[2] Putman, V. L., & Paulus, P. B. (2009). Brainstorming, brainstorming rules and decision making. The Journal of creative behavior, 43(1), 29-40.

[3] Goodman, C. M. (1987). The Delphi technique: a critique. Journal of advanced nursing, 12(6), 729-734.

[4] Koutsoyiannis, D., Efstratiadis, A., &Georgakakos, K. P. (2007). Uncertainty assessment of future hydroclimatic predictions: A comparison of probabilistic and scenario-

based approaches. Journal of Hydrometeorology, 8(3), 261-281.

[5] Yang, J. S., & Kim, I. H. (2013). Development of drought vulnerability index using delphi method considering climate change and trend analysis in Nakdong river basin. KSCE Journal of Civil and Environmental Engineering Research, 33(6), 2245-2254.

[6] Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. Journal of machine Learning research, 3(Jan), 993-1022.

[7] Wei, X., & Croft, W. B. (2006, August). LDA-based document models for ad-hoc retrieval. In Proceedings of the 29th annual international ACM SIGIR conference on Research and development in information retrieval (pp. 178-185).

[8] David M. Blei and John D. Lafferty. 2006. Dynamic topic models. In Proceedings of the 23rd international conference on Machine learning (ICML '06). Association for Computing Machinery, New York, NY, USA, 113–120. DOI:https://doi.org/10.1145/1143844.1143859

[9] Wang, C., Blei, D., & Heckerman, D. (2012). Continuous time dynamic topic models. arXiv preprint arXiv:1206.3298.

[10] Xuerui Wang and Andrew McCallum. 2006. Topics over time: a non-Markov continuous-time model of topical trends. In Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining (KDD '06). Association for Computing Machinery, New York, NY, USA, 424–433. DOI:https://doi.org/10.1145/1150402.1150450