

CONFERENCE PROCEEDINGS MITA 2019

The **15th** International Conference
on Multimedia Information Technology
and Applications (**MITA2019**)

June **27** - July **1**, 2019

VNUHCM-UEL, Ho Chi Minh City, Vietnam

Organized by

Korea Multimedia Society (KMMS), Korea

Vietnam National University-Ho Chi Minh City

University of Economics and Law (UEL), Vietnam

www.icmita.org



Poster Session B

When & Where: June 28, 15:30 ~ 16:30, A Hall Lobby, A1 Building

Chair: Prof. Mikio Hasegawa (Tokyo University of Science, Japan)

- PB-1. Directionality Change Study according to Array Change of Ultrasonic Transducer for Completely in the Canal Hearing Aids** 246
Jin Wan Seok*, Sungdae Na**, Qun Wei*** and Myoung Nam Kim*
*Kyungpook National University, Korea
**Kyungpook National Chilgok University Hospital, Korea
***Keimyung University, Korea
- PB-2. Deep Learning Based Pedestrian Detection Embedded System For Traffic Light Control** 248
Pu-Lum Son* and Joon-Jae Lee*
*Keimyung University, Korea
- PB-3. A Decentralized Copyright Management System Centered on Copyright Holders** 252
Hye-Bin Kim*, ByeongGyu Jeong* and Sang Uk Shin*
*Pukyong National University, Korea
- PB-4. Analysis of Fish Behavior Responses in Chemical Stress Using Permutation Entropy and Fractal Dimension Algorithms** 256
Quach Kha Quang*
*Dong Thap University, Vietnam
- PB-5. Analysis of the Core Technique for Copyright Infringement Sites Tracking and Automatic Evidence Collecting** 259
Seul-Ki Choi* and Jin Kwak*
*Ajou University, Korea
- PB-6. Home IoT Devices and Network Domination with AI Speaker Vulnerabilities** 263
Ju-Seong Ko* and Jin Kwak*
*Ajou University, Korea
- PB-7. PCB Conformal Coating Inspection System** 267
Hyeong-Jo Yoon*, Ndayishimiye Fabrice* and Joon-Jae Lee*
*Keimyung University, Korea
- PB-8. Design of a parking management system with user convenience** 271
Jeong-Min Oh*, So-Jeong Park*, Tae-Seong Yun*, Jeong-Gook Koh* and Tae-Wook Kim**
*Tongmyong University, Korea
**Autonics, Korea

Analysis of the Core Technique for Copyright Infringement Sites Tracking and Automatic Evidence Collecting

Seul-Ki Choi

ISAA Lab., Dept. of Computer Engineering
Ajou University
Suwon, Republic of Korea
skchoi.isaa@gmail.com

Jin Kwak

Dept. of Cyber Security
Ajou University
Suwon, Republic of Korea
security@ajou.ac.kr

Abstract— With the development of IT technology, information sharing technology such as the Internet has made it possible to obtain various information easily. However, there are adverse effects such as illegal circulation of copyrighted content. Therefore, research to prevent such illegal distribution is necessary. In this paper, we analyze the core techniques required for tracking copyright infringement sites and collecting evidence of illegal distribution. It also proposes the entire process of the system using core techniques.

Keywords—copyright infringement; copyright protection; core technique;

I. INTRODUCTION

Recently, it is actively utilizing IT technologies such as the Internet, selling and distributing contents online. With this, we have the benefit of being able to easily access content such as music, movies, and broadcasts anytime, anywhere. However, using this information sharing technology maliciously, copyright infringement cases are increasing rapidly. According to an annual report released by Korea Copyright Protection, the number of content infringed by copyright infringement in 2017 reached 997.1 million, which amounted to 2.564 trillion won[1]. The table below shows the first issue related to copyright protection by five types of copyright.

TABLE I. COPYRIGHT PROTECTION ISSUE BY TYPE (ONLY 1ST RANK)

Type	Description of the 1 st issue
Music	Illegal Distribution Music using SNS
Movie	The outflow of simultaneous movies in IPTV
Broadcasts	Illegal Distribution using mobile apps
Webtoon	Illegal copy and distribution webtoon site
Publishing	Distribution of illegal copies for online publications
Game	Game distribution and infringement through overseas

Various efforts and studies have been conducted to prevent illegal distribution of contents. Despite these efforts, however, there are increasing cases of unauthorized distribution of content by opening a site with a server outside the country. Especially in recent years, not only movies and broadcasting, but also illegal Webtoon sites are at serious level.

Therefore, in this paper, we analyze the core techniques required for a system that can track and collect evidence against copyright infringement sites for copyright protection.

II. REALTED WORKS

A. ICOP in Korea copyright protection agency

In 2004, ETRI developed ‘digital fingerprinting’ technology to track the illegal distribution of digital content[3], thereby enabling technology to monitor and track the initial spread of illegal content. The Korea copyright protection agency operates the ICOP based on this result[2].

The ICOP (Illegal Content Obstruction Program) is a system that automatically monitors illegally distributed copyrighted works in web hard based on web crawling technology. In 2017, the existing system was reorganized and the development of a lightweight web hard monitoring search program was promoted and is currently in operation. The Korea Copyright Supervisor has supplemented the illegal duplicate search function and constructed a linkage system to transmit the monitoring result of the emergency response work to the system that manages the overall situation of coping with piracy. The illegal copy search function is to automatically perform a series of processes such as accessing a web hard site, collecting post information (crawling), and capturing a post screen. This is similar to the way people monitor their illegal posts directly on the web hard site.

B. Toon Radar in Naver webtoon

Naver webtoon platform developed a technology to track user illegally copied and distributed by reading user

identification information embedded in webtoon through 'Toon Rader AI' applied artificial intelligence technology[4]. Machine learning for image tracking, artificial neural network with convolution structure is used to detect illegal distributors. At this time, because human beings are verifying the information that artificial intelligence finds, automated data detection technology is needed.

C. Watch Tower in Daum webtoon

Webtoon platform Daum webtoon analyzed patterns such as user's content subscription time to construct a technical system to prevent copyright. Based on the results of the analysis, Daum webtoon developed a 'Watch Tower' technology that can identify users who are suspected of illegal copying and who upload them to illegal sites[5].

III. ANALYSIS OF CORE TECHNIQUE

In this chapter, we propose core technique for tracking copyright infringement sites and collecting evidence of illegal distribution.

We also analyze the necessities and characteristics of each component technique and propose an automated entire process using each core technique.

A. List of Core Technique

The table below shows the technique needed to track copyright infringement and collect evidence.

TABLE II. CORE TECHNIQUE BY CATEGORY

No.	Category	Core Technique
1	Search	Automatic search technique for content distribution sites
2	Anonymization	Anonymization-based access technique
3	Analysis	Image-based multi-character combination language recognition technique
4		Illegal distribution content detection technique
5		High-speed analysis technique
6	Profiling	Profiling technique for copyright infringement
7		Lifecycle profiling technique for copyright infringement sites
8	Evidence generation	Evidence generation technique for illegal distribution of contents
9	Storage and management	Evidence storage and management technique for illegal distribution contents
10		Data storage and management technique regarding the characteristics of copyright infringement
11	Automation	Automatic technique for entire process

B. Necessities and feature of Core Technique

(No.1) Automatic search technique for content distribution sites

- There is a need for a technique for automatically searching for sites that distribute content illegally.
- This technique is used to automatically search and collect addresses of sites that distribute content based on SNS and real-time search terms.

(No.2) Anonymization-based access technique

- Some sites use technique to block access and search of bots that collect information about illegal content.
- An anonymization technique is needed to bypass the above-mentioned access blocking technique.
- An anonymity can be strengthened by using a relay node with features that are difficult to trace such as Tor.

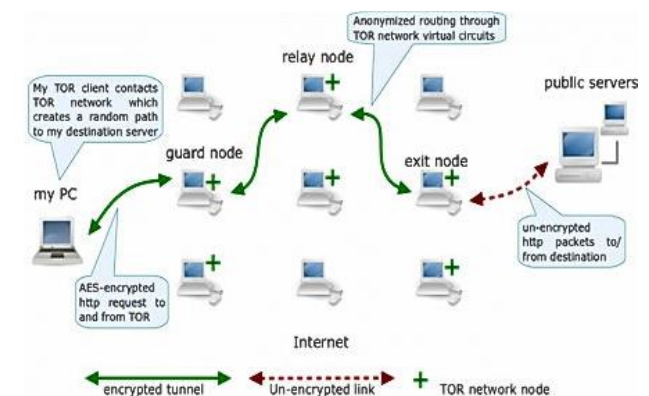


Fig. 1. Anonymization technique implemented as Tor

(No.3) Image-based multi-character combination language recognition technique

- Some illegal content also destroys the form of the language to prevent the collection system from searching the title or contents.
- By using language or special symbols in other countries to destroy the form of language, people can understand, but the collection system is not easily understood.

e.g.) It can transform the word 'collect' to 'co1lect'.

(No.4) Illegal distribution content detection technique

- There is a need for technique to detect copyright infringement on the contents posted on the site.
- Content descriptions can be written and published so that only people can understand it in context.
- It is difficult to detect illegal distribution contents by a simple method of comparing only the copyright protection list.

e.g.) A description of the movie "John Wick" can be written as follows. 'A movie where the hero of the Matrix, who became homeless, continues to shoot.'

(No.5) High-speed analysis technique

- Speed similarity comparison technique and data clustering technique are needed.
- Since content search is basically a problem that requires a number of comparisons, it is necessary to have a technique capable of high-speed comparative analysis through work such as clustering in advance.

(No.6) Profiling technique for copyright infringement

- Copyright infringement sites periodically change the URL for site access to avoid crackdowns.
- However, since the URL is changed, but the site maintains its basic form, it needs profiling technique to keep track of these sites quickly.

(No.7) Lifecycle profiling technique for copyright infringement sites

- Copyright infringement sites are operated through various processes to avoid crackdown on illegal distribution.

- There is a lifecycle for the operation of infringing sites, such as a countermeasure to avoid crackdown on illegal distribution while maintaining the operating phase after the site is created.

- By analyzing and profiling these lifecycles, it is possible to understand the infringement site (such as response to crackdown on illegal distribution) and to track in detail.

(No.8) Evidence generation technique for illegal distribution of contents

- There is a need for a technique for automatically collecting evidence of illegal distribution of content.
- Because the methods of collection of evidence may vary according to the types of illegal distribution contents, research for classifying the characteristics of contents by type is also necessary.

e.g.) Content such as movies can be illegally distributed in the form of downloading or streaming. In the case of books, you can also use the method of posting scanned files. In this way, illegal distribution methods can be different for each type of content, and the method of collecting evidence according to illegal distribution methods can also be changed.

(No.9) Evidence storage and management technique for illegal distribution contents

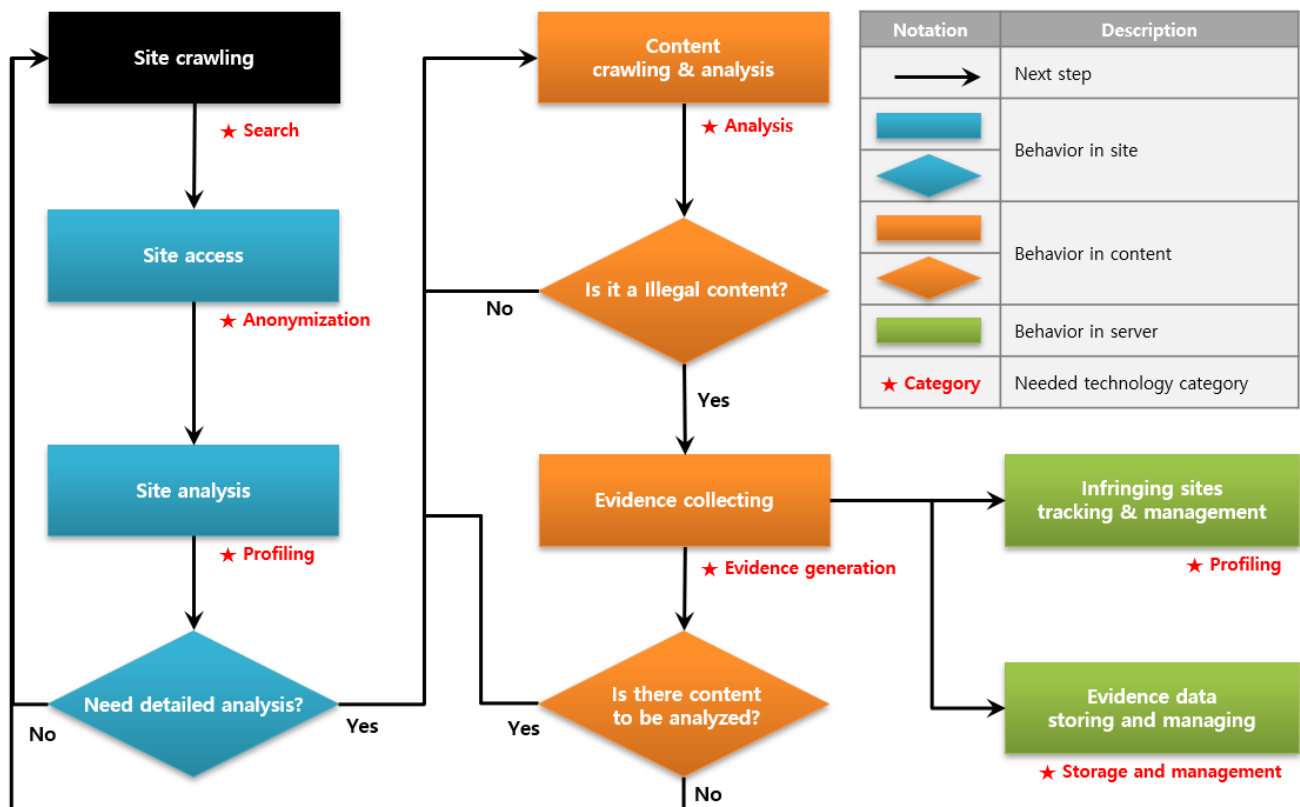


Fig. 2. Automate the entire process of tracking copyright infringement sites and collecting evidence

- There is a need for technique that can quickly and accurately collect and manage evidence of illegal distribution.
- Because the number and size of evidence data collected at the infringing site can be enormous, techniques are needed to effectively manage it.

(No.10) Data storage and management technique regarding the characteristics of copyright infringement

- Techniques are needed to store and manage data collected from infringing sites in order to perform profiling on the characteristics of copyright infringement.
- By performing profiling through feature analysis of copyright infringement, it is possible to improve the efficiency and accuracy of tracking and evidence collection techniques for copyright infringement sites.
- In addition, it can be used in research to predict the new types of copyright infringement by analyzing trends in copyright infringement.

(No.11) Automatic technique for entire process

- It is not efficient for people to directly analyze and track illegal distribution contents and copyright infringement sites that are generated and operated hourly.
- There is a need for a technique that can automatically run the entire process using core technologies from No.1 to No.10.
- By automating the entire process, it can be used to track and manage illegal distribution of content 24 hours a day.

Fig. 2. shows the whole process of tracking and collecting evidence against copyright infringement sites using core techniques analyzed in this paper.

IV. CONCLUSION

Internet technologies that provide a variety of information provide us with various benefits. However, due to the indiscreet use of information sharing technology, contents that should be copyright protected are illegally distributed.

In this paper, therefore, we analyze the core techniques required to track copyright infringement sites and collect evidence of illegal distribution. We also proposed the overall operation of the system using core techniques. Through the follow-up study on the technologies analyzed in this paper, we will try to protect the rights of the creators of contents that should be protected by copyright by developing an automated system for copyright infringement site tracking and evidence collection.

Acknowledgment

This research project was supported by Ministry of Culture, Sports and Tourism(MCST) and from Korea Copyright Commission in 2019.

References

- [1] Korea Copyright Protection Agency, "Annual Report on Copyright Protection," KCOPA, 2018
- [2] Korea Copyright Protection Agency, "Center of Copyright Protection Technology Korea Copyright Protection Agency," KCOPA STORY – Official blog of Korea copyright protection agency, July 2018
- [3] W.Y. Yoo, Y.H. Seo, J.G. Choi, and J.H. Park, "Technical Trends of Digital Fingerprinting & Buyer-Seller Watermarking," Electronics and Telecommunications Trends, Vol. 19 No.3, pp. 96-106, June 2004.
- [4] National IT Industry Promotion Agency, "Digital Contents Weekly Report," Vol 94. May 2018.
- [5] ZDNet Korea, "Daum webtoon 'Watch Tower' copyright protection role," <http://www.zdnet.co.kr/view/?no=20180615151526>, June 2018.