Congestion-Aware Suspicious Object Detection System Using Information-Centric Networking

Xin QI*, Toshio SATO†, Keping YU‡, Zheng WEN§,
San Hlaing Myint¶, Yutaka KATSUYAMA‖, Kiyohito TOKUDA**, Takuro SATO††
* † ‡ ¶ ‖ ** Global Information and Telecommunication Institute, Waseda University, Tokyo, Japan
§ School of Fundamental Science and Engineering, Waseda University, Tokyo, Japan
†† Research Institute for Science and Engineering, Waseda University, Tokyo, Japan
Email: *samqixin@aoni.waseda.jp

2021 IEEE 18th Annual Consumer Communications & Networking Conference (CCNC) | 978-1-7281-9794-4/21/$31.00 ©2021 IEEE | DOI: 10.1109/CCNC49032.2021.9369510

***Abstract*—Deadly diseases and terrorist attacks are greatly threatening human safety, which challenges global security. To address this issue, urban surveillance systems are being applied at a rapid pace with mature but inefficient solutions in large scale networks. When a surveillance network is managing the data generated from multiple edge nodes, it is easy to create congestions due to concentrated data traffic and inefficient data delivery mechanism. In parallel, 5G technology, cope with explosive mobile data traffic growth and massive device connections, can realize a true "Internet of Everything" and build the social and economical digital transformation. In this paper, in the context of 5G technology, we propose an Information-Centric Networking (ICN) surveillance system based on our designed Suspicious Object Network System (SONS) over the concept of next-generation networking. In this solution, the edge nodes in the network distribute the computing and data storage requirements. We first describe the current surveillance issues and our proposed system architecture. Then we use simulation to verify and evaluate the system performance between legacy all-to-one centralized surveillance system and ICN based decentralized surveillance system.**

***Index Terms*—ICN, surveillance system, decentralization, 5G, next-gen network**

I. Introduction

Public security, as affect the daily life of each person, is vital in modern society. In recent years, public security in the world has become more and more complex. Traffic accidents, terrorist attacks, and frequent crimes have severely affected social stability. In order to curb crime and reduce public security risks, surveillance systems have been performed in important locations. It can realize the recording of monitoring audio and video. In short, the surveillance system is playing an essential role in early warning of dangerous behaviours and crime detection afterwards.

The current state of surveillance system uses a mature solution containing point-to-point connections between the central server and each monitoring node, on which live video data is transmitted. The center server is receiving multiple live feed during this process. The simple and effective way to release the transmitting capacity pressure is to increase the capacity by upgrading to higher level transmit standards. However, higher quality equipments comes with higher investments and energy consumption.

5G is a next-generation mobile standard that provides new applications and services with gigabit rates to ensure a better experience for end-users and significantly improve performance and reliability. The high speed and low latency promised by 5G will push society into a new era of smart cities and the Internet of Things (IoT). Industry stakeholders have identified several possible use cases for 5G networks, and ITU-R has identified these use cases as three critical categories 1) enhanced mobile broadband; 2) massive machine type communication; and 3) ultra-reliable and low-latency communication. Meanwhile, Information-Centric Networking (ICN) [1] is a next-generation network architecture designed for the future. It has the benefits of utilizing content store in the nodes to cache data and serve the consumers effectively. By naming the content, the content consumers can find the target data efficiently. In the context of 5G technology, applying ICN to the suspicious object network system can solve problems such as network congestion, high investment, and high energy consumption in TCP/IP.

Although just watching selected video data by using indices of time and location is already realized in regular IP-based video surveillance systems [2], this function is insufficient for large-scale video surveillance. Extraction of various indices by analysing video data is fist studied for video-on-demand systems [3] [4] to handle and retrieve huge amounts of videos. This approach is applied to the control surveillance video system. In the reference [5], video data from multiple surveillance cameras are processed by using indices about kinds of vehicles. However, they use a centralized storage server that limits scalability for multiple areas.

Reduction of data volume in videos is also researched including data compression and data extraction. For instance, the extraction of a key frame from the video is proposed in [3]. As the distributed process and storage platforms, the Hadoop [6] and Hbase [7] were proposed, however, these database-like approaches are suited for the bigdata analysis not appropriate for frequently updated video surveillance data. Another de-centralization approach is using the Information-Centric Network (ICN) platform [8] [9]. The ICN has the ability to use a namespace to process indices to realize efficient data transmitting. Though it also has a subject to handle frequently updated data [10], the ICN also has the possibility

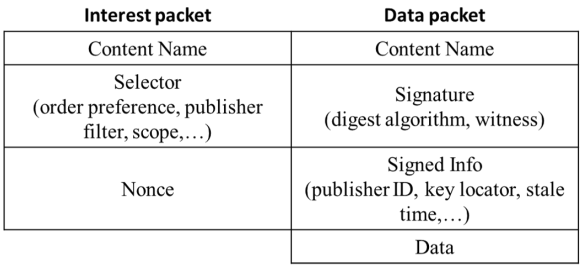

Fig. 1. Interest and Content packets.

to handle another issue for scalability against the number of surveillance areas.

The contribution of this paper are listed as follows.

1)We design a content-oriented surveillance system base on named data network architecture.

2)We introduce content names to the surveillance system data delivery mechanism.

3)A simulated evaluation is realized to compare transmitting capacity requirements between a legacy centralized network and a decentralized network.

The remaining of this paper is organized as follows. Section II presents preliminary studies, including researches on named data network, surveillance network systems and unidentified object detecting system. In section III, we present our design of a content-oriented surveillance system base on the named data network. Then in section IV, an evaluation with simulated scenarios. Finally, we conclude our work.

II. Preliminaries

A. ICN, CCN, NDN and 3N

There are many ICN projects, such like Content-Centric Networking (CCN) [1], Named Data Networking (NDN) [11] [12] and Named-Node Networking (3N) [13] [14]. Taking an example, NDN shares the idea of transferring the nowadays transformational architecture focusing on where, addresses and hosts, to what, content itself that consumers and applications care about.

In a named data network, the concept of addresses and hosts is weakened, replacing by data names. A consumer sends an interest packet when requiring a named data. An interest packet contains only the content name be routed to the content producer. Once the interest arrives at the producer, the target data is returned from the producer back to the consumer via the same route. By default, all the router nodes in the data delivery path cache the data in their content store. When other consumers are requiring the same data that matches it in the content store, the router node can return the data directly. The named data are addressable, route-able, authenticated and irrespective.

Introducing host-centric architecture back in ICN, 3N [13] [14] is designed to improve the packet loss and delay performance of ICN. 3N architecture applies to all its nodes and supports mobility [14] better than other ICN structure network.

TABLE I
Edge node input and output data size

	Input data	Output data
4K camera	3840x2160 pixel RGB images (24bit) @5 fps	80x286 compressed & cropped RGB image (24bit) 32KB
W-band imager	256x256 pixel images (58bit) @10 fps	256x256 compressed imager image (16bit) 640KB

B. Surveillance Network Data Delivery

In the current stage of surveillance system development, an IP-based video surveillance network [2] has realized reviewing interested video data by using indices like timestamps and location information. However, this function is insufficient for the large-scale video surveillance network. Various indices extraction by analysing video data was first studied in video-on-demand (VOD) systems [3] [4] to handle and retrieve huge volumes of video data. This approach is applied for controlling surveillance video system. In the reference [5], video data from multiple surveillance cameras are processed by using indices about various kinds of vehicles. However, they use a centralized storage principle that limits the scalability for multiple areas.

C. Suspicious Object Network System

With the principle of "ensuring a sufficient security level without stopping the flow of people", we have designed a suspicious object detection system (SONS) to recognize suspicious objects concealed by humans [15]. Figure 2 shows the overall structure of SONS. Before generating and transmitting data in the network, there are two screening procedures, primary screening, and secondary screening, take place. In the primary screening, we use a combination of sensors, multiple visible-light cameras and W-band active radar imagers, to recognize and identify whether a person is having suspicious objects (metal, etc.) from up to 15 meters away. In an area, there is a set of visible-light camera and an imager. The camera can detect a person within 6m 15m and the imager can detect a person within 1m 5m respectively. When the camera detects a person and confirms the person is moving into the imager detection zone, the imager starts to scan this person. If the imager detects suspicious objects that roughly categorized as metal, this person is identified as a suspicious person in the first screening process. If a suspicious person is recognized, our system will provide the person's information to the security personnel. The information includes sensor images and other data. Then the security can find and guide the suspicious person to a secondary screening place to perform more detailed inspections. In the secondary screening, hybrid imagers are used to further identify objects shape and materiel for what kinds of suspicious objects the suspicious person is carrying. During the two-stage screening process, each suspicious person's face and other information is recorded and tracked by our system and cached.

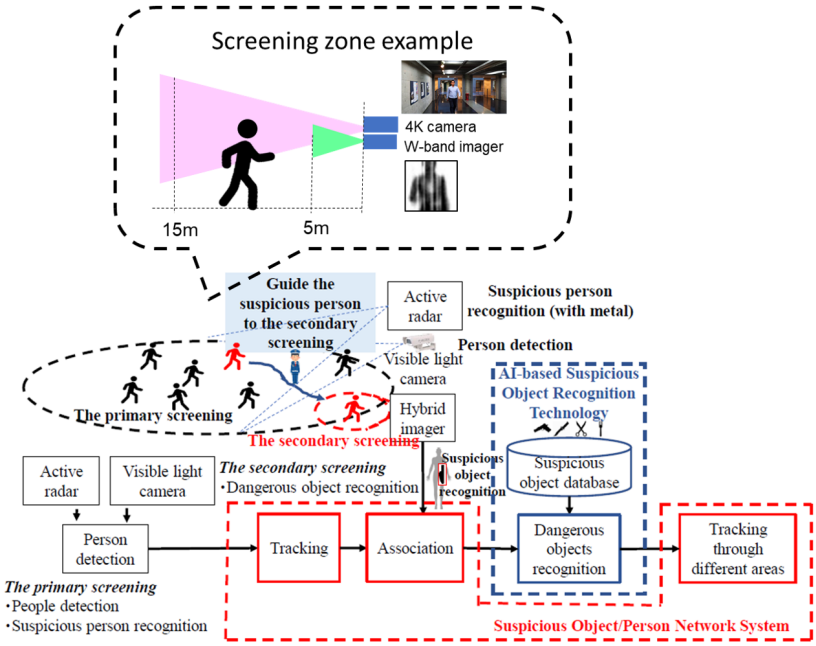

Fig. 2. SONS overall structure.

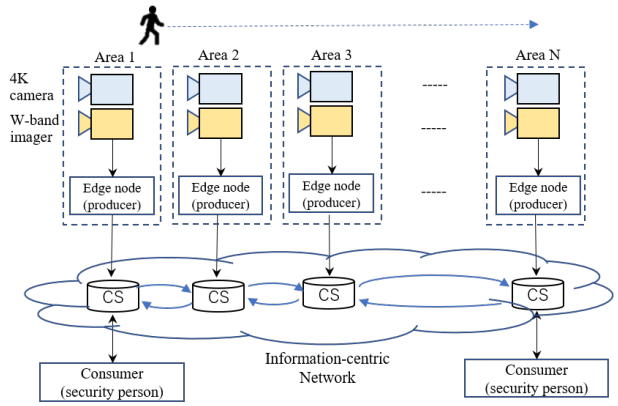

Fig. 3. System architecture.

III. System Design

A. Objective

The objective of this paper is to describe the design and simulation of the content-oriented surveillance network. The common TCP/IP based surveillance network usually uses video data delivery method to enable the center node processing the data and generate needed information. Our designed surveillance network focuses on decreasing the required processing power and network bandwidth consumption to provide a flexible surveillance topology.

In this section, our designed content-oriented surveillance network is described in three parts, system architecture, contents' naming structure and data delivery mechanism.

B. System Architecture

The system architecture is combined with several parts. From figure 3 we could see it is divided into multiple areas containing sensors like cameras and imagers that provide the raw data to their edge node content producer, which acts as a producer in the ICN network. The edge nodes are connected in the ICN network and a consumer can access and retrieve data from it.

When a person is moving among the areas covered by sensors, the sensors pick up the person's information and the edge node generates the data. In a legacy video surveillance system, the video data is transmitted directly to a central node or a user. In our system, the edge node can process and generate key information from the raw data and serve the information as contents. When one or multiple consumers are requesting the content, they can send an interest packet and retrieve the content from the original producer or nearest node's content store.

C. Content Naming

In the ICN based surveillance networks, a content name uses character "/" to separate different components and is designed as a hierarchical format. A sample name is shown in figure 4. The content name prefix structure can easily identify different layers, the place, person ID, suspicious object ID and timestamp. The first part of the name prefix identifies the kind of suspicious information and suspicious person. The second part identifies the location where this information is from, which can be area number indexed in the system. The third part identifies the person ID which generated in the surveillance system to record and track the person. The

fourth part identifies the suspicious object kind that the person is carrying, this ID can simply refer to a object category includes danger objects like knifes and guns. The last part is the timestamp, which indicates at what time this data is generated.

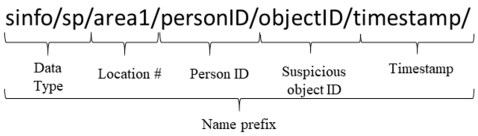

Fig. 4. Content name prefix.

D. Edge Nodes & Data Delivery

There are two parts involved in the edge node process, data input and output in the edge node. Table I shows the information of edge node input data, output data.

In the surveillance system, the edge node in each area responses to the data process and communication to the network. On the input side of the edge node, we have a 4K surveillance camera and a W-band imager. The 4K camera inputs a 3840x2160 resolution 24bit motion image flow (figure 6) at 5 frames per second. The W-band imager inputs a 256x256 resolution 58bit radio-based imager motion image flow at 10 frames per second. The RGB image is processed by person detection and tracking to select a keyframe and a region-of-interest. The W-band imager is utilized to detect and recognize concealed dangerous objects such as knives and guns. The recognition results are used for indices to make content data name. After standard processing, the edge node outputs compressed data to the consumer. The output data contains a 80x286 resolution RGB image (left of figure 7) that cropped and compressed from the input, a 256x256 resolution imager image (right of figure 7). The edge node transmits the full-size data to the consumer. The data contains a 32KB and 640KB images from the output data. The total content data volume is about 700KB.

In the legacy surveillance system designs, example shown in figure 8, the surveillance network usually works as a centralized topology. Every edge node delivers data to a central node or data center. It helps the data center to fully collect all the available data and process them together. This minimizes all data processing latency but it may suffer from data delivery latency in the first place.

We design a named content-based data delivery network for the surveillance system. The consumer only requests the necessary data from the edge nodes, which are suspicious person information in our scenario. This method should give us advantages of reducing data transmission congestion, using the content stores to serve multiple consumers requesting the same contents, and decentralizing data storing in the network. The required transmitting capacity *RTC* can be calculated:

$$RTC = N \times TUF \tag{1}$$

TABLE II
SIMULATION PARAMETERS

	Centralized System	Decentralized System	
Edge node numbers	100	100	100*10
Center node & consumerlink speed	100Mbps	100Mbps	1Gbps
Data size	700KB	700KB	700KB
Suspicious person appear rate	1/10	1/10	1/10
Simulation period	60s	60s	60s

where N is the number of areas, *TUF* is the utilization factor. After that, according to equation 1, *RTC* is calculated and shown in figure 9. It shows the rough estimation of required data transmission capacity for the total area number *N*, calculated by utilization factor of 50% and 80%. In a 100 Mbps link capacity network, it is obvious being not capable to transmit a $N>15$ areas data traffic.

IV. Simulation evaluation

A. Configuration & Parameter

The simulation uses ndnSIM [12] [16] simulator to achieve named data network structure and its data flow. Table II illustrates the parameters of the two configurations. To emulate the real application scenario of a high speed but also stable 5G environment. We didn't choose to have simulated 5G connections, but static wired connections. This let the surveillance system being able to adapt both legacy wired network and future next-gen network.

The edge node for each area has the same content producer function described in section II. In this simulation, there are 100 edge nodes response for 100 areas. Each edge node has a 100Mbps link to the center node. Every time a person is captured by the sensors in an area, the edge node generates a data packet with 700KB payload. As the topology in figure 8, all 100 edge nodes transmit data to the center node before identifying a suspicious person.

In one area, the decentralized system shares the same node number as the centralized system. The edge node number is also 100 and nodes are connected like in figure 3. The data links between each node is 100Mbps and the data payload size is 700KB. A data consumer in this decentralized network only requires suspicious person data from the edge nodes. Because the suspicious object appear rate is 1/10, the valid data is hitting about 1/10 of the total data in the surveillance network.

We have added a large-scale experiment with a total number of 10 surveillance facilities (each facility has 100 areas). In the large-scale experiment, the overall node number reaches 1000 and generate approximately 10 times traffic data of one surveillance facility (100-area). We will show the overall traffic figure which also indicates the overall link capacity usage and we look forward to tell potential differences between the ICN-based decentralized networks with different link load.

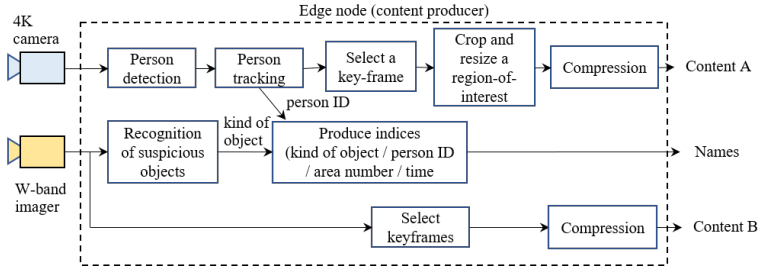

Fig. 5. Edge node (content producer) process flow.

Fig. 6. 4K camera input image.

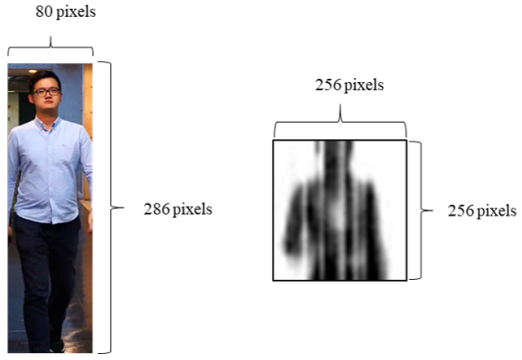

Fig. 7. RGB and imager output data.

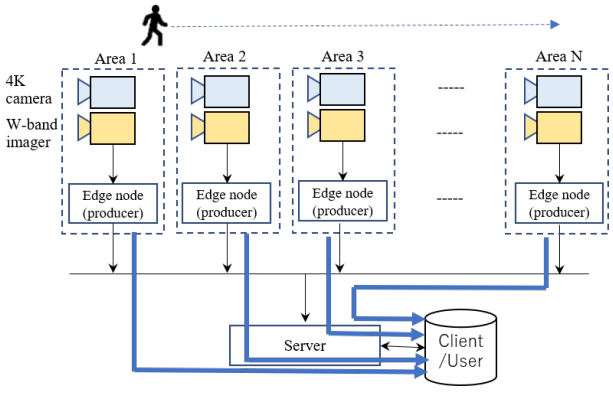

Fig. 8. Centralized data delivery.

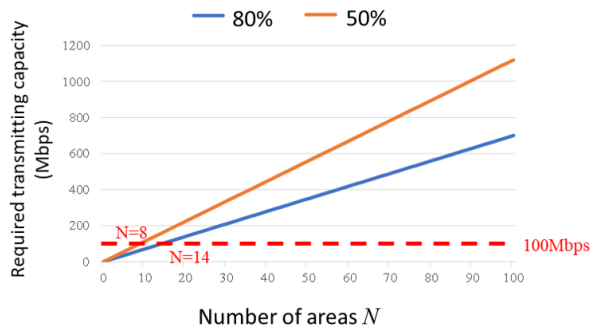

Fig. 9. Required transmitting capacity with different utilization factors.

B. Simulation Result

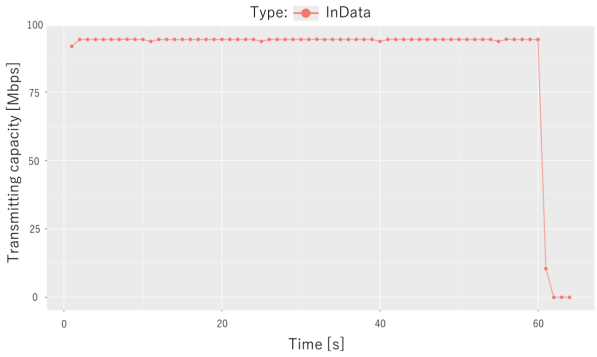

Fig. 10. Centralized system simulation result.

There are three results in this subsection. In each figures, InData represents the summary of data delivered to the consumer, OutInterest represents the summary of interests sent out from the consumer. In figure 10 shows the simulation result of a centralized network. We can see, during 60s of the simulation time, the link capacity is fully used under such high load. The edge nodes and their packets begin to timeout and being dropped. This results in losing data in the surveillance network. Figure 11 shows the result of a 100-area decentralized network simulation. The max transmitting capacity axis value is 60Mbps. The link capacity is utilized at about 60% to a 100Mbps link standard, which is at a reasonable rate for data transmission. The network has a certain amount of

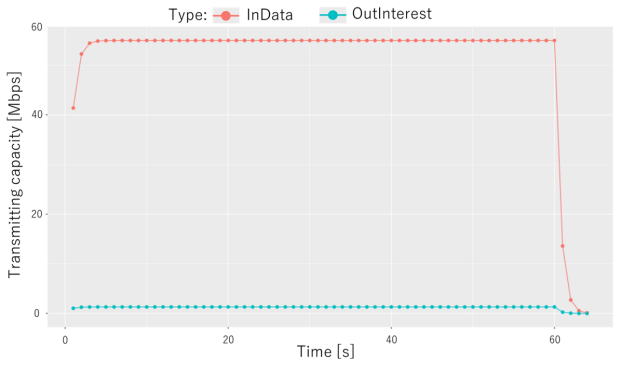

Fig. 11. Decentralized system simulation result.(100-area)

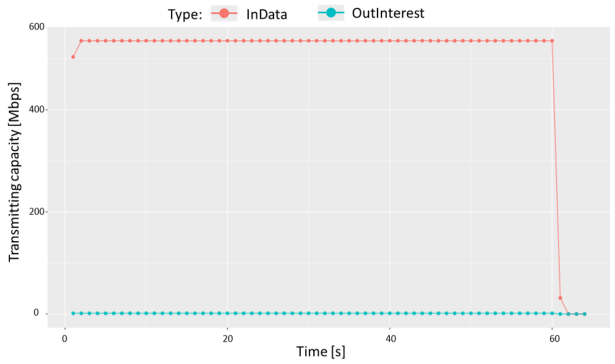

Fig. 12. Decentralized system simulation result.(1000-area)

transmitting redundancy. Figure 12 shows the overall result of a 1000-area decentralized network simulation. The max transmitting capacity axis value is 600Mbps, 10 times of that in the 100-area result. The overall link capacity utilization is about 60% to a 1Gbps link standard, similar to the previous result. This result further confirms that by increasing the total node number, the ICN network remains in high performance.

V. Conclusion

In this paper, we designed a content-oriented surveillance system with efficient data delivery that utilizes named contents. We proved ICN architecture can optimize system performance under multiple consumer scenarios. The useful content-only delivery mechanism optimizes network bandwidth usage and distributes part of the data processing to the edge nodes. The multiple simulations show great potential in large scale surveillance networks in large areas and larger groups of areas like subway stations, train stations and airports.

Acknowledgment

This research has been supported by research grant for expanding radio wave resources (JPJ000254) of Ministry of Internal Affairs and Communications under contract for “Research and development of radar fundamental technology for advanced recognition of moving objects for security enhancement”.

References

[1] V. Jacobson, D. K. Smetters, J. D. Thornton, M. F. Plass, N. H. Briggs, and R. L. Braynard, “Networking named content,” in *Proceedings of the 5th International Conference on Emerging Networking Experiments and Technologies*, ser. CoNEXT ’09. New York, NY, USA: ACM, 2009, pp. 1–12. [Online]. Available: http://doi.acm.org/10.1145/1658939.1658941

[2] M. Systems, “Xprotect vms 2020 r2, getting started guide - single computer installation,” available: https://doc.milestonesys.com/.

[3] H. D. Wactlar, T. Kanade, M. A. Smith, and S. M. Stevens, “Intelligent access to digital video: Informedia project,” *Computer*, vol. 29, no. 5, pp. 46–52, 1996.

[4] T. Sato, T. Kanade, E. K. Hughes, and M. A. Smith, “Video ocr for digital news archive,” in *Proceedings 1998 IEEE International Workshop on Content-Based Access of Image and Video Database*. IEEE, 1998, pp. 52–60.

[5] R. T. Collins, A. J. Lipton, H. Fujiyoshi, and T. Kanade, “Algorithms for cooperative multisensor surveillance,” *Proceedings of the IEEE*, vol. 89, no. 10, pp. 1456–1477, 2001.

[6] K. Shvachko, H. Kuang, S. Radia, and R. Chansler, “The hadoop distributed file system,” in *2010 IEEE 26th symposium on mass storage systems and technologies (MSST)*. Ieee, 2010, pp. 1–10.

[7] W. Zhang, Y. Zhang, L. Xu, and F. Gong, “Hbase based surveillance video processing, storage and retrieval,” in *2016 International Conference on Identification, Information and Knowledge in the Internet of Things (IIKI)*. IEEE, 2016, pp. 64–67.

[8] X. Qi, Z. Wen, T. Tsuda, W. Kameyama, K. Shibata, J. Katto, and T. Sato, “Content oriented surveillance system based on information-centric network,” in *2016 IEEE Globecom Workshops (GC Wkshps)*. IEEE, 2016, pp. 1–6.

[9] K. Okamoto, T. Mochida, D. Nozaki, Z. Wen, X. Qi, and T. Sato, “Content-oriented surveillance system based on icn in disaster scenarios,” in *2018 21st International Symposium on Wireless Personal Multimedia Communications (WPMC)*. IEEE, 2018, pp. 484–489.

[10] T. Ito, H. Noguchi, Y. Yamato, and T. Murase, “Transaction offloading for access management to live data of iot in information-centric network,” in *2018 IEEE 7th Global Conference on Consumer Electronics (GCCE)*. IEEE, 2018, pp. 287–288.

[11] L. Zhang, A. Afanasyev, J. Burke, V. Jacobson, P. Crowley, C. Papadopoulos, L. Wang, B. Zhang *et al.*, “Named data networking,” *ACM SIGCOMM Computer Communication Review*, vol. 44, no. 3, pp. 66–73, 2014.

[12] S. Mastorakis, A. Afanasyev, and L. Zhang, “On the evolution of ndnSIM: an open-source simulator for NDN experimentation,” *ACM Computer Communication Review*, Jul. 2017.

[13] J. LÓPEZ and T. SATO, “Seamless mobility in icn for mobile consumers with mobile producers,” *IEICE Transactions on Communications*, vol. advpub, 2017.

[14] X. Qi, Y. Su, K. Yu, J. Li, Q. Hua, Z. Wen, J. López, and T. Sato, “Design and performance evaluation of content-oriented communication system for iot network: A case study of named node networking for real-time video streaming system,” *IEEE Access*, vol. 7, pp. 88 138–88 149, 2019.

[15] K. Yu, X. Qi, T. Sato, Z. Wen, Y. Katsuyama, K. Tokuda, W. Kameyama, T. Sato *et al.*, “Design and performance evaluation of an ai-based w-band suspicious object detection system for moving persons in the iot paradigm,” *IEEE Access*, vol. 8, pp. 81 378–81 393, 2020.

[16] S. Mastorakis, A. Afanasyev, I. Moiseenko, and L. Zhang, “ndnsim 2: An updated ndn simulator for ns-3 revision 2,” *Univ. California, Los Angeles, Los Angeles, CA, USA, Tech. Rep. NDN-0028*, 2016.