Development of the Absent Detection System for Online Integrity Evaluation

Il-Kwon Lim, Hyun-Jun Park, Han-Jin Cho

With the development of information and communication technology, modern society has developed into a super connective society. As a result, online evaluation is widely used in modern society where it is common to see work remotely. In Korea, Public Procurement Service and the Ministry of SMEs and Startups are used as representative. If the online evaluation is performed by the Public Procurement Service, video evaluation is performed in the same way as offline face-to-face evaluation, but the sincerity of the evaluation committee can be a problem on-line. Thus, Public Procurement Service administrator individually measured the sincerity according to the assessor's The Absent Detection Rate. However, in this paper, we implemented the Absent Detection System for Online Integrity Evaluation using Spark and OpenCV technology to automatically measure the Absent Detection Rate. This allows us to measure the integrity of our members more accurately and efficiently.

Key words: Online Evaluation, Big Data, OpenCV, Spark, The Absent Detection

I. INTRODUCTION

The development of information and communication technology has evolved into a super-connected society that connects all human beings, objects, and environments. Technological advancements such as artificial intelligence and block chains are underway based on connecting everything. The foundation of this development is the Internet usage rate. In 2016, the number of Internet users in Korea is 43 million, and the number of users worldwide is 340 million [1].

Therefore, information and communication technologies such as e-government system, electronic payment, internet banking, smart city, smart home using mobile Internet, and mobile web service are widely used and used in civil society in the modern society. In the modern society where it is common everyday to see work remotely, such as video conferencing, smart work, and BYOD (Bring Your Own Device), utilizing the information and communication technology, the evaluation of the projects and tasks using online is widely used.

Recently, the most widely used institutions for online evaluation in Korea are *Ministry of SMEs and Startups* and *Public Procurement Service* ("PPS" hereinafter*)*. In the case of *Ministry of SMEs and Startups*, on-line evaluation of SMEs' R&D projects is performed, while on-line evaluation of online evaluation is conducted through PPS's The E-Ordering System. In the case of the online image evaluation service, basically, evaluation is performed using a personal PC at home using a webcam for PC and a video conferencing application, and the process is the same as the offline face evaluation [2-5].

In this paper, *The Absent Detection Rate* is the ratio of the actors who vacate the evaluation committee or other evaluation participant while confirming each other's appearance on online screen only. In the case of the online evaluation, the PPS captures the assessment of the evaluation committee by unit time and confirms *the Absent Detection Rate*. In this paper, we present the Apache Spark framework, which is widely used for real-time data collection, and the OpenCV technology, which is widely used in computer vision technology, and the Hadoop cluster, which is a standard technology for big data processing. We developed the absent detection system for online integrity for evaluating the loyalty of the participant in online evaluation.

The composition of this paper is as follows. Chapter 1 introduces the background and necessity to write a thesis, and Chapter 2 describes research trends of big data and computer vision technology to be implemented in thesis. Section 3 describes the application program for system implementation and system design based on it. In Chapter 4, we showed the key sources, implementation methods, and implementation screens for infrastructure configuration and development to implement the system. Section 5 describes the significance of the system implemented in this paper and its future work.

II. RELATED RESEARCH

2.1 Apache Spark for real-time data processing

Hadoop and NoSQL for handling big data are non-relational DBMSs that are designed differently from traditional relational database management systems ("RDB" hereinafter) and have the advantage of being able to handle large amounts of data flexibly. Technology is utilized as a standard file system and a distributed processing platform as well as being utilized as an industry standard technology as various ecosystems are researched and developed. Hadoop refers to the implementation of Hadoop Distributed File System ("HDFS" hereinafter) and MapReduce. However, Hadoop has some problems in handling real-time data. First, MapReduce generates large amounts of disk I/O and network traffic in intermediate data transfers.

Revised Manuscript Received on May 22, 2019.

Secondly, it is not suitable for performing repeated tasks according to the processing of the two-stage structure. Third, it is not asynchronous with sporadic data operations. Therefore, spark must be used for fast repetitive operations and real-time batch processing [6, 7].

Apache Spark (“Spark” hereinafter) provides a structure and interactive data mining tool that supports fast processing data for repetitive data operations. Accordingly, various operations are supported through the RDD (Resilient Distribute DataSet), which is an abstraction object concept of the data set. When the RDD cannot be used, it is processed in cooperation with the Hadoop file system. RDD represents users' work with DAG (Data Acyclic Graph), and job scheduling is managed through Mesos or YARN [6]

2.2 Computer Vision Technology and OpenCV

Computer vision is a technology that extracts features from images or image data to analyze objects and motion. Computer vision can be expressed as computer vision. In recent years, advanced technology has been studied every year by introducing deep learning technology to machine learning technology and image recognition [8].

The OpenCV library (Open Source Vision Library, “OpenCV” hereinafter) is an open source computer vision and machine learning software library. It is the de facto standard because it is the most used and most used program worldwide [9, 10]. In this paper, since it is aimed to quickly recognize whether a person is a simple person or not, rather than using accurate user recognition and object judgment, it is aimed to provide an administrator with the ability to use the OpenCV technology, which is a virtually standardized technology of computer vision, so that it can be quickly judged and processed.

III. SYSTEM DESIGN

3.1 Evaluation committee's real-time data sitting image collection system design

In this paper, when the manager starts the evaluation, the evaluation committee accesses the evaluation page of the system. Accordingly, the webcam is activated and the sitting image is transmitted to the evaluation committee every 5 seconds, and the image collected from the system is received to determine whether or not the user is seated through the image analysis. Images are classified as unstructured data because they are not composed of RDB or a set of rules.

In this paper, we use Apache Kafka (“Kafka” hereinafter) technology to collect images in real time. Kafka is a distributed messaging system that guarantees high performance and stability. It can process the system down or error-free even if many messages are generated at a certain point in time due to the system characteristics. In order to collect images in real time, we use Kafka to store and process the collected image data in HDFS. This process is shown in figure 1.

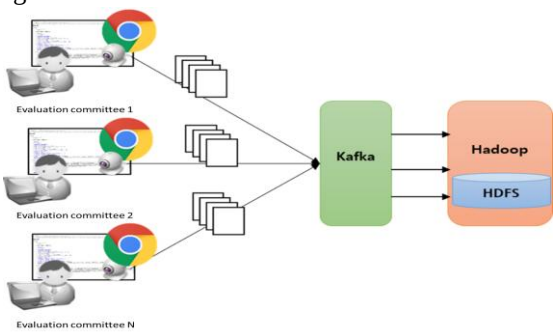

Figure 1 Design of a collection system for the Absent Detection Rate detection system

3.2 System design for analysis of collected sitting images

Kafka and Spark Streaming techniques are used to process the sitting image analysis in real time, which is generated by the evaluation committee in 5 seconds when performing online evaluation. The system is largely divided into Producer and Consumer areas. Producer sends the evaluation information and sitting image generated every 5 seconds to Kafka in the Producer area and imports messages (messages sent to the Producer) loaded in Kafka in the Consumer field and stores the analysis results in HDFS. This is implemented in the Spark Streaming application. The image processing and analysis design of the system is as shown in figure 2 and the execution process is as shown in the following table 1.

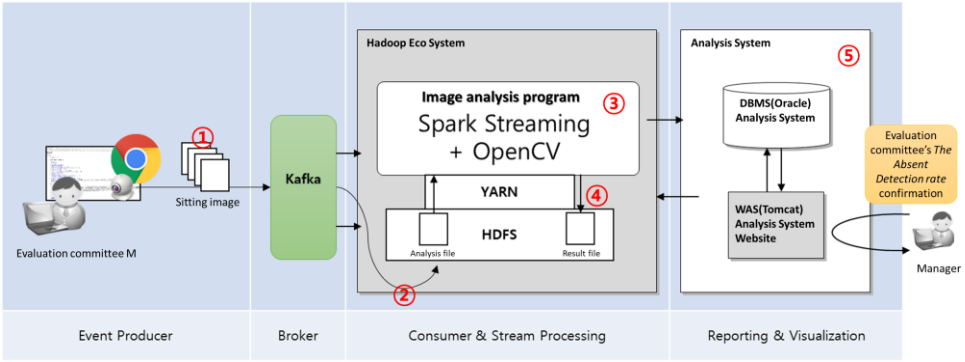

Figure 2 Overall System Design

Table 1 Processes running on the system

Num.	*Execution Process*
1	*When the procurement evaluation task is performed on the analysis system, the seat image of the evaluator is transmitted to the server once every 5 seconds through the webcam*
2	*Upload captured image to Hadoop cluster HDFS*
3	*Determine whether the uploaded images are seated by face analysis through OpenCV*
4	*Analysis results Upload images to Hadoop cluster HDFS*
5	*Save analysis results to RDB and provide them to users through web application*

IV. SYSTEM IMPLEMENTATION

4.1. System configuration

Hadoop which consists of MapReduce and HDFS for distributed processing is installed and Zookeeper is installed HA configuration to increase stability of distributed processing. Hadoop cluster consisting of three data nodes and two name nodes are configured. The following figure 3 shows the system configuration screen.

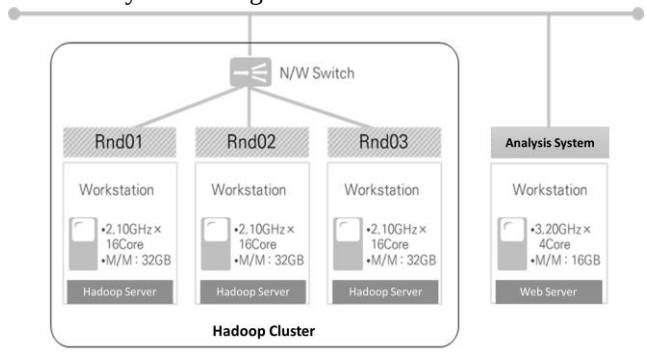

Figure 1 System Configuration

4.2. System implementation

Since the virtual evaluation system is developed for evaluating the seats at the online evaluation, the evaluation committee implemented the web-based evaluation system so that the evaluation committee implemented the function to send the sitting image to the generated image analysis program every 5 seconds through the webcam. The Web-based evaluation system performs the functions of the Producer, a key component of the Kafka architecture. Create a stream message and send the message to the topic. Due to the nature of the system, we need to use the webcam function in the web browser, so we configured the web server as HTTPS. It is a Java-based web program that consists of an application divided into a server and a screen. The screen was implemented using the AngularJS (ver 1.6.1) framework based on Single Page Application (SPA), and the server implemented the Korea e-government standard framework (ver 3.5.1) We implemented a Java web application in which the screen and the server communicate with JSON (JavaScript Object Notation).

4.3. System implementation screen

Figure 4 is a web screen for transmitting a sitting image when the implemented virtual online evaluation is performed. The transmitted image is stored in HDFS as shown in figure 5.

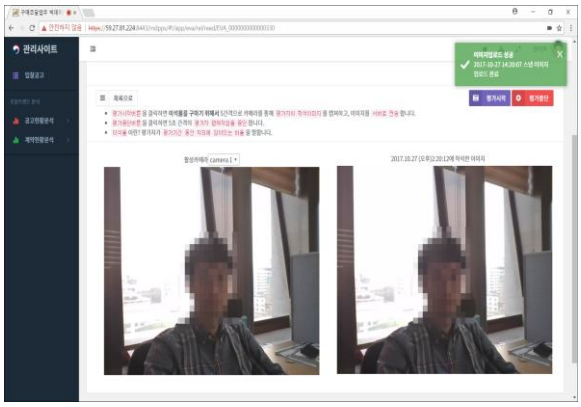

Figure 2 Sitting image transfer screen

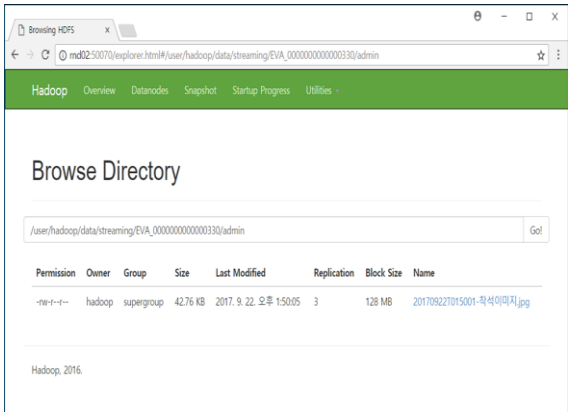

Figure 3 Save the evaluation result image HDFS

Next, figure 6, 7 shows the result screen after the evaluation of the virtual web-based online evaluation system.

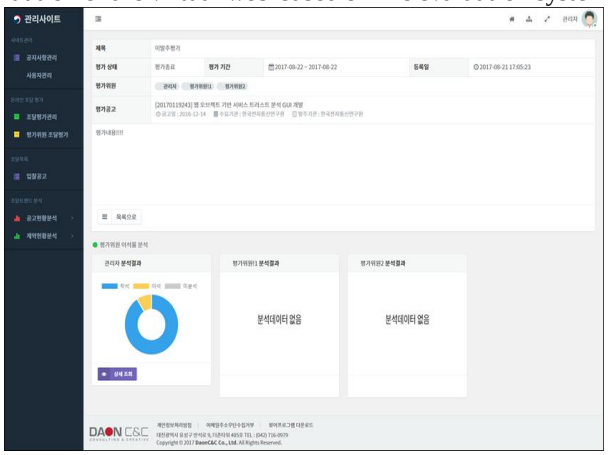

Figure 4 The Absent Detection Rate Result Confirmation Screen After Evaluation (1/2)

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

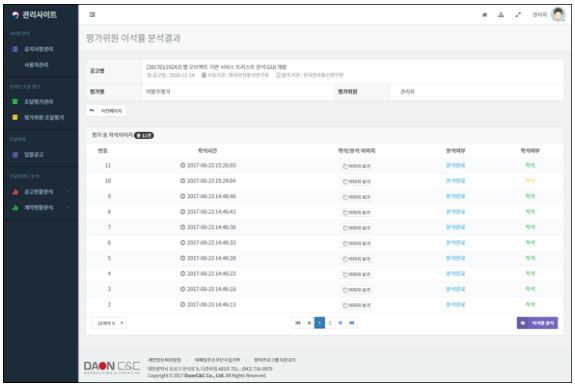

Figure 5 The Absent Detection Rate Result Confirmation Screen After Evaluation (2/2)

V. CONCLUSIONS AND FUTURE WORK

In this paper, we have developed a system to measure *the Absent Detection Rate* Result and to measure *the Absent Detection Rate* Result in order to evaluate the trustworthiness of the evaluation committee in online evaluation. In the real-time evaluation based on the online evaluation, which is the same as the face-to-face evaluation, *the Absent Detection Rate* Result of the evaluation committee is evaluated in real time using the loyalty evaluation system developed in this paper. We have developed a system that automatically measures and notifies the manager. In this paper, Kafka and Spark streaming technology was used to collect and store the evaluations of the evaluation committee in real time by using webcam, and HDFS and Map Reduce, spark technology.

In the case of the online evaluation system of the e-ordering system of the PPS, since the administrator checks the collected webcam image to measure *the Absent Detection Rate* Result to measure *the Absent Detection Rate* Result of the evaluation committee, The ABSENT DETECTION SYSTEM FOR ONLINE INTEGRITY EVALUATION is expected to increase work time and efficiency. However, this paper uses Big Data technology only for real - time image processing and analysis, and if deep-learning and machine learning that are being actively studied are utilized recently, it will be used for user classification and further user authentication Can be. In addition, this paper will analyze *the Absent Detection Rate* Result by analyzing only the seating rate of the simple evaluation committee member, but it will be able to distinguish the state of sleeping or doing something if more detailed image analysis is done. In addition, we will be able to measure the learner sincerity of the online education system through extension and application of this technology.

Therefore, it is expected that future research will be able to utilize the study of classification and user recognition of the evaluation committee by collecting more sample image data and mechanically learning the collected images. In addition, it is possible to measure whether or not the evaluation committee evaluates faithfulness in a true sense, rather than simply measuring *the Absent Detection Rate* implemented in this paper.

ACKNOWLEDGMENT

This work was supported by the 2018 Far East University Research Grant(FEU2018S04).

REFERENCES

1. 2017 Korea Internet White Paper, *Korea Internet & Security Agency*, 2017
2. Online written evaluation committee manual of *Ministry of SMEs and Startups, Korea Technology and Information Promotion Agency for* SMEs, 2018
3. E-ordering system web page, Introduction of proposal and proposal evaluation, web: rfp.g2b.go.kr, 2015
4. E-ordering system's video evaluation manual, *Public Procurement Service*, 2017
5. E-ordering system's manual, Public Procurement Service, 2017
6. J. H., Eun, T. H., Kim, S. W., Lee, C. H., Jung, H. M., Jung, Next Generation Real Time Big Data Distribution System Trends - Focused on Sparks and Storms -, Information and Communication Technology Promotion Center, Weekly Technology Trend (2014)
7. Apache Spark web page, web:spark.apache.org, 2018
8. Y. G., Jung, S. Y., Ji, Journal of Communications of the Korean Institute of Information Scientists and Engineers, Special Issue (2014), 15-21
9. H., Jang, Computer Vision and Artificial Intelligence, Information and Communication Technology Promotion Center, Weekly Technology Trends (2016), 14-23
10. OpenCV web page, web:opencv.org, 2018