Development of a Road Big-Data Storage Platform for Predicting the Driving Environment

1 Jung-Ah Ha, 2 Kyusoo Chong (Corresponding author)
1,2 Korea Institute of Civil Engineering and Building Technology
ksc@kict.re.kr

ABSTRACT

In this study, a platform that collects and stores various data on the roads to predict the driving environment was developed. Various open platforms were compared and analyzed for the development of a platform, and a road big-data storage platform was developed using Hortonworks, an open source distributed storage and processing platform. The collected data were classified into real-time data and non-real-time data according to their characteristics, and were stored in different ways. Non-real-time data were collected using Sqoop, which can take a large amount of data at a time, and the data that can be collected in real time through Open API were collected using Flume. The collected data were stored in the Hadoop-based HDFS from various processes, and were expressed on the web. Based on some collected data, it was found that the driving hours differed according to the precipitation, but it was believed that making an accurate prediction in a period which predict difficult requires data convergence. The platform that was developed in this study was a collection and storage model thus issued such as how to analyze the data collected through this system, how to predict the driving environment and what data is needed by users will addressed in the future.

Keywords: Hadoop, open API, big data, driving environment.

1. INTRODUCTION

Due to the recent surge in the use of the web, all kinds of sensor networks, various smart equipment, and SNS, the enormous amounts of data generated by these sources are continually exploding. Such data is normally stored in DBMS and is presented in the structured typical.

The current data, however, such as various sensor stream data and SNS, are semi-structured data.

Thus, the big-data processing technology was developed for the semi-structured data. The big data have huge volumes and wide variety, and they require processing and analysis at a high velocity. It is also important to create new values from the collected big data.

As the interest in big data is increasing, the need to utilize the public data provided by the government and public organizations as well as the private sector, which are regarded as part of big data, to improve the quality of people’s lives and to realize effective management is rising. Studies to maximize the public benefits through big-data analysis are actively being conducted. Also, many countries are introducing web-based data one-stop services for sharing and utilizing public data through government sites.

The big data addressed in this study are road traffic data for predicting the driving environment. These data can be utilized for the effective control of the city traffic circumstances by the existing systems. The fastest route search service of T-map, a popular mobile navigation platform in South Korea, is an additional service to the existing route search function. The service finds an optimal route leading to the user’s destination by using all the traffic data coming in real time. It also swiftly analyzes the traffic history and accident data stored by Korea Expressway Corporation (EX) for holidays and specific dates, and provides the users with the required data [1]. Also, the real-time optimal traffic route guidance system researched on by J. T. Lim et al. (2014) recommends routes by predicting the traffic situation in each part of the route through the analysis of the past traffic data thereof. It is also a system that was developed through big-data analysis [2].

Currently, the IT-based businesses are exerting efforts to find big-data businesses, and are strengthening their strategies to take the initiative in the global big-data market. While the traffic database in South Korea has grown significantly in quantity through the continuous extension of the traffic data system, its values are low owing to the absence of a system that reproduces meaningful data from the huge amount of collected data.

The construction of a big-data integrated analysis system using the traffic data of TMS (Traffic Monitoring System) and ITS (Intelligent Transportation System) run by the government will provide highly reliable integrated traffic data, contributing to the efficient and economical planning for the establishment of national policies [3]. A platform that collects and stores the data required for reproducing meaningful data and predicting the driving environment was developed in this study. Also, a platform for storing and managing the various road data generated from weather forecasts and open-type traffic data as open-type big data was developed.
2. STUDY REVIEW

2.1 Big-Data Studies

The cases that analyze and use big data are gradually expanding in all areas of the society. The traffic area is not an exception, and various analyses have been performed using various traffic data. One of the best cases of the application of big-data analysis results to everyday life in South Korea is that concerning the Seoul night bus routes.

For the selection of the Seoul night bus routes, Seoul City and KT co-worked to estimate the demand for such buses and the selected routes, and adjusted the bus intervals based on the locations and registered addresses of the people making a phone call with their cell phones at night. Based on the assumption that these night callers were going home at the corresponding times, the locations of their phone calls were set as the starting positions, and their actual addresses as the arrival positions. These separate data were linked in space, and additional six routes were selected based on the demand [4].

Daegu City constructed a traffic safety platform that can analyze and predict the current traffic situations and the future traffic accident risks using the traffic accident information for the past six years, the reports collected in real time, and the data provided from SNS and the web. Through the platform, the citizens can voluntarily provide information on the risk factors, including the traffic and pedestrian accidents on the roads, using a mobile application. The platform also provides data about traffic situations [5].

2.2 Big Data Open Platform

2.2.1 Horton Works

Horton works Data Platform (HDP) is an open-source distribution version run by Apache Hadoop. HDP provides the elements of the latest version of Apache. All the elements are official versions of the recently released Apache. As it is the policy of Hortonworks to provide patches only when necessary, HDP has almost perfect patches. Unlike the platforms constructed by other companies using Hadoop, Hortonworks is providing Apache Software Foundation with 100% codes. Hortonworks is selling professional technical assistance and educational services. It is providing all technologies as open sources at no cost [6].

The key element of HDP is Hadoop Distributed File System (HDFS). The previous YARN, a part of Hadoop 2.0, performs a process resource management function by MapReduce. It also provides arrays for as-extensive-as-possible interactions between the stored data in HDFS. Therefore, various applied programs can share common resource management in Hadoop.

2.2.2 Lumify

Lumify is an open-source project with which one can converge and analyze big data and make a visualization platform. The intuitive web-based interface is useful in investigating the correlations between data through the real-time visualization of sharing 2D and 3D graphs, the entire text search, dynamic histograms, conversational geographic maps, common workspaces, and the lines of analysis options. Lumify is useful in searches, graph visualization, link analysis, geospatial analysis, and multimedia analysis [7].

2.2.3 InfiniDB

InfiniDB is a high-performance analysis platform constructed for data warehousing, business intelligence and big-data analysis applications. InfiniDB offers considerable benefits by providing the cloumar-storage-type paradigm. It can search fast, analyze and combine all rows, and delete rows without requiring table reconstruction. InfiniDB is a GPLv2 open-source platform driven by the MySQL engine. It can perform not only ad-hoc BI but also dimensions and predictions analysis through the MySQL interface. InfiniDB can be installed in AWS Cloud and Apache Hadoop.

Horton works, a 100% open source, was used in this study for cost saving and for developing a storage platform for predicting the customized driving environment. Hortonworks was utilized because it has good Hadoop expandability and was believed to have more expertise as a third of the Hadoop developers are working in Hortonworks.

3. COLLECTED DATA

A big-data collection method for observing the road driving environment and for providing prediction data was studied to offer user-oriented road services. The prediction of the road driving environment requires weather, road environment, and traffic data. In this study, these data were classified into traffic data, road environment data, exterior environmental data, and road data, as shown in [Table 1].

The collected data were classified according to their characteristics before storage. Real-time collection was possible for the sites that provided data in Open API, but real-time collection was not performed for the other sites that did not provide data in Open API or that provided only statistical data. The data for real-time collection were stored using Flume, and the other data were stored using Sqoop, for the development of a storage platform.
Table 1: The itemized information needed for predicting the driving environment

<table>
<thead>
<tr>
<th>1st Category</th>
<th>2nd Category</th>
<th>3rd Category</th>
<th>Sites</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traffic Data</td>
<td>Traffic Flow Data</td>
<td>Velocity, traffic amount, density, traffic time, standby period, occupancy rate</td>
<td><a href="http://www.its.go.kr">http://www.its.go.kr</a></td>
</tr>
<tr>
<td>Traffic Data</td>
<td>Traffic Accident Data</td>
<td>Location, blind spot, casualty, damage data, recovery, dangerous area, frequent-accident area</td>
<td><a href="http://www.its.go.kr">http://www.its.go.kr</a></td>
</tr>
<tr>
<td>Traffic Data</td>
<td>Traffic Contingency Data</td>
<td>Location, situation type, object type, action status, renewal status</td>
<td><a href="http://www.its.go.kr">http://www.its.go.kr</a></td>
</tr>
<tr>
<td>Traffic Data</td>
<td>Traffic Control Data</td>
<td>Location, control type, object, time</td>
<td><a href="http://www.its.go.kr">http://www.its.go.kr</a></td>
</tr>
<tr>
<td>Road Environment Data</td>
<td>Road Contingency Data</td>
<td>Sharp curve, steep slope, narrowing area, dangerous area</td>
<td>-</td>
</tr>
<tr>
<td>Road Environment Data</td>
<td>Road Construction Data</td>
<td>Location, period, type, traffic availability</td>
<td><a href="http://www.its.go.kr">http://www.its.go.kr</a></td>
</tr>
<tr>
<td>Road Environment Data</td>
<td>Road Management Data</td>
<td>Location, authorization area, road type, road name, length, pavement type, number of lanes, land width</td>
<td>-</td>
</tr>
<tr>
<td>Exterior Environment Data</td>
<td>Weather Data</td>
<td>Temperature, weather, visibility, fog, wind speed, wind direction, humidity, precipitation</td>
<td><a href="http://www.kma.go.kr">http://www.kma.go.kr</a></td>
</tr>
<tr>
<td>Exterior Environment Data</td>
<td>Holiday Data</td>
<td>Chuseok, New Year’s Day, golden holidays</td>
<td>-</td>
</tr>
<tr>
<td>Road Data</td>
<td>Road Pavement Data</td>
<td>Crack rate, surface smoothness, pavement condition grade, repair history</td>
<td>-</td>
</tr>
<tr>
<td>Road Data</td>
<td>Slope Data</td>
<td>Slope precision investigation data, regular measurement data</td>
<td>-</td>
</tr>
<tr>
<td>Road Data</td>
<td>Bridge Tunnel Data</td>
<td>Bridge type, bridge load, design load, structure data</td>
<td>-</td>
</tr>
<tr>
<td>Road Data</td>
<td>Road Occupancy Data</td>
<td>Occupancy location, occupancy area, occupancy purpose, occupancy period</td>
<td><a href="http://road.cpermit.go.kr">http://road.cpermit.go.kr</a></td>
</tr>
<tr>
<td>Road Data</td>
<td>Road Signal Data</td>
<td>Road signal sheet, guidance site name, road name, signal status data</td>
<td>-</td>
</tr>
</tbody>
</table>

3.1 Flume

Kafka or Flume can be used to store the data collected in real time in the platform under development. As Flume provides a large amount of API to source, it is easily provided with Open API, which various sites offer.

Therefore, Flume was used to collect real-time data such as the data from the National Transport Information Center (NTIC) or Korea Meteorological Administration (KMA). Flume has an advantage of being able to store data in various data storages, such as RDBMS, HDFS, and NoSQL, among others.

3.2 Sqoop

Sqoop plays the role of designing the import algorithm for collecting large-capacity data and the expert algorithm for storing analysis results data. Sqoop is a specialized tool for data collection through SQL in the relation database, and has advantages in the parallel collection of as much data as the different distribution servers contain, or more data according to the Hadoop distribution environment. Each site stored data such as txt, xls, and csv, among others, in Hadoop through Sqoop.
Figure 1: Conceptual diagram of flume

Figure 2: Conceptual diagram of sqoop

Table 2: Data collected by apache flume

<table>
<thead>
<tr>
<th>No.</th>
<th>Site Address</th>
<th>Data</th>
<th>Storage Cycle</th>
<th>HDFS File Creation Period</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Twitter agent</td>
<td>Collect poltra keyword twit registered at Twitter</td>
<td>Upon event occurrence</td>
<td>10 minutes</td>
</tr>
<tr>
<td>2</td>
<td>ITS agent</td>
<td>Ilsan Seo-Gu and Bundang-Gu area road flow data</td>
<td>5 minutes</td>
<td>1 hour</td>
</tr>
<tr>
<td>3</td>
<td>KMA agent</td>
<td>KMA local forecast data</td>
<td>3 hours</td>
<td>3 hours</td>
</tr>
<tr>
<td>4</td>
<td>SK Planet agent</td>
<td>Bundang-Gu area weather data</td>
<td>1 hour</td>
<td>24 hours</td>
</tr>
<tr>
<td>5</td>
<td>TE agent</td>
<td>Traffic Event data</td>
<td>3 hours</td>
<td>6 hours</td>
</tr>
<tr>
<td>6</td>
<td>TA agent</td>
<td>Traffic Accident data</td>
<td>3 hours</td>
<td>6 hours</td>
</tr>
<tr>
<td>7</td>
<td>EX agent</td>
<td>Seoul-Busan Expressway safe-driving data Seoul-Busan</td>
<td>1 hour</td>
<td>6 hours</td>
</tr>
</tbody>
</table>

4. STORAGE PLATFORM DEVELOPMENT

A method of collecting and storing data in connection with road traffic and weather data collected outside was investigated to store road big data. Open API and RSS based on RESTful were used for the connection with outside data. A large number of data, such as traffic accident cause analysis and traffic flow data, were stored through the open-type big-data platforms that were developed in this study (e.g., HDFS, Mysql, and MongoDB).

4.1 Apache Flume Framework Agent

4.1.1 Twitter Agent

This is an agent that collects contingency/accident data occurring in Twitter. The contingency/accident data are the data continuously uploaded at the Twitter account of the NTIC. Each region has its own traffic information center account, and every traffic data tweet from all regions includes the keyword “#poltra.” Therefore, the Twitter agent filters and acquires only the tweets that include the keyword “#poltra.” The contents of the tweets collected by the Twitter agent have the following basic format:

[Traffic Events] Contingency contents (location, situation explanation, etc.) #poltra NTIC is the fastest source of road contingency data as it uploads construction, accident, contingency, and exception data on roads in real time. However, it is not operational for 24 hours; therefore, the data on the traffic events occurring at dawn cannot be collected.

4.1.2 ITS Agent

This is the agent for collecting the traffic data provided by ITS (Intelligent Transport System), a system being managed by NTIC. ITS is running an Open API server and is opening the public data. Each road is identified by a node ID and a link ID, which are composed based on the standard node link of ITS. Some roads may be limited in collection according to their circumstances. The roads that cannot respond to a collection request need to be classified as exceptions.

4.1.3 KMA Agent
The KMA agent is an agent that collects local weather forecast data from KMA. It requests for and receives weather forecast data from the KMA Open API server every 3 hours. The analysis of the local weather forecast data around Bundang-Gu collected every 3 hours between 17:00 July 27, 2015 and 23:00 July 28, 2015 revealed that the weather forecast had about 85.7% accuracy. Although KMA predicts the weather in two days, it is estimated that continuous collection every 3 hours will increase its forecast accuracy.

4.1.4 SK Planet Agent
SK Planet provides various Open API services, including maps, weather data, and payments, by running a developers’ center. The SK Planet agent provides various weather data in minutes or hours. Its storage platform reads the current weather data by the hour. As the SK Planet test grade (free of charge) has a restriction of 200 queries/day, weather data were collected every hour. As only the weather data of one area per query could be provided, the collection was restricted, and the weather around Bundang-Gu was collected in this study.

4.1.5 TE (Traffic Event) Agent
The TE agent collects the traffic contingencies provided by the ITS Open API servers and brings real-time construction data occurring nationwide. The TE agent has so small an amount of returned data without the occurrence of an overhead when the nation is assigned as a boundary that the collection of national contingency data is possible. However, as the construction data are event data with relatively long durations, data collection in short cycles may generate a large amount of duplicated data.

4.1.6 TA (Traffic Accident) Agent
The TA agent is an agent that retrieves the accident data provided by the Open API server of ITS. The accident data may be retrieved through Twitter, but the Twitter data and the accident data of ITS sometimes collide, which requires mutual supplementation.

4.1.7 EX Agent
The EX agent retrieves the safe-driving data of Seoul-Busan Expressway and the traffic amount of each entrance and exit of the operation center from EX [11].

The safe-driving data include message information for preventing collision, dazzling, and drowsy driving. The code for each expressway can be inputted as a request variable when making a query. When no code is inputted, the agent may retrieve the safe-driving data of all the expressways nationwide, but may experience overhead. Currently, the EX agent retrieves the safe-driving data only of Seoul-Busan Expressway.

The traffic amount of each entrance and exit of the operation center was provided by counting all the vehicles passing the operation center. Overhead may occur in the traffic amounts of the entrances and exits; therefore, only the amounts of Seoul-Busan Expressway were collected.

4.1.8 SD Agent
The SD agent retrieves the safe-driving data provided by ITS. These data are similar in type to those provided by EX. The safe-driving data of ITS include the standstill, rear-end collision, shoulder driving, and sleeping shelters. The SD agent requests for the nationwide expressway safe-driving data. As it receives more returned data compared to other agents and as the collection cycle is short (5 minutes), the memory channel capacity was expanded.

4.1.9 VMS Agent
The VMS agent retrieves the variable message sign data provided by ITS in real time. The data are varied and include the standstill and the living information required while driving on a highway.

<table>
<thead>
<tr>
<th>No.</th>
<th>Site Address</th>
<th>Data</th>
<th>Format</th>
<th>Period</th>
<th>Term</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>KMA National Climate Center</td>
<td>Weather observation data</td>
<td>CSV</td>
<td>2010-2014</td>
<td>1 hour</td>
</tr>
<tr>
<td>2</td>
<td>KoRoad</td>
<td>Traffic accident data</td>
<td>CSV</td>
<td>2011-2014</td>
<td>Each road</td>
</tr>
<tr>
<td>3</td>
<td>NTIC</td>
<td>Traffic flow data</td>
<td>CSV</td>
<td>2015 (2 days before the present)</td>
<td>5 minutes</td>
</tr>
<tr>
<td>4</td>
<td>NTIC</td>
<td>Construction/accident data</td>
<td>CSV</td>
<td>2015 (a month before the present)</td>
<td>Upon occurrence</td>
</tr>
<tr>
<td>5</td>
<td>NTIC</td>
<td>VMS data of each road</td>
<td>CSV</td>
<td>2015 (2 days before the present)</td>
<td>Upon occurrence</td>
</tr>
</tbody>
</table>
4.2 Apache Sqoop Framework Agent

4.2.1 KMA National Climate Center
KMA provides the past weather data through the National Climate Center. The meteorological observation data obtained since 1904 can be retrieved. The observation data were measured using the Automated Synoptic Observing System (ASOS) located outside each region, and include the pressure, temperature, humidity, wind, precipitation, and sunshine amount. The meteorological observation data measured in Seoul from 2010 to 2014 were collected and loaded in this study.

4.2.2 Road Traffic Authority (KoRoad)
KoRoad is running the Traffic Accident Analysis System (TAAS), which provides the past traffic accident history based on GIS [12]. EPDO data expressing the risk of a road in four levels (safe, caution, dangerous, and serious) by converging and analyzing the traffic accident data, real-time weather data, and contingency data for the past 3 years by road can be retrieved. The higher the EPDO grade is, the higher the road risk. The traffic accident data include the data on the accident-prone roads and the number of people injured in the accidents. The data between 2010 and 2014 were collected and loaded in this study.

4.2.3 NTIC Traffic Flow Data
NTIC is also providing the traffic flow data collected through the Flume agent in real time in the CSV file type. It provides the traffic flow data of all the nationwide links in the past in daily files, including the flow speed, flow time, and link ID. The traffic flow data of all the nationwide links in January 2015 were collected and loaded in this study.

4.2.4 NTIC Construction/Accident Data
The construction/accident data being provided by NTIC contain explanations of the contingencies and their locations (latitude and longitude). Different data regarding each contingency exist in column form; therefore, two spare columns were added in this study to organize a table.

4.2.5 NTIC VMS Data
NTIC collects the VMS data of all the roads nationwide. The VMS history data can be retrieved and downloaded at the ITS Open API site. The past VMS data were collected and loaded in this study.

4.2.6 EX
EX provides expressway public data information (traffic, construction, maintenance, general administration, toll, resting area, etc.) in Open API. The in- and out-traffic amount data of each operation center include the traffic amount data by car model and the traffic directions in the operation centers of each expressway. The number of columns is higher than in any other table due to the collection by car model. The in- and out-traffic amount data of every operation center on all the expressways nationwide between January and August 2015 were collected and loaded in this study.

4.3 Development Platform and Web UI
The collection and storage platform that was developed in this study used Flume and Sqoop for collection and HDFS for storage, as shown in the following figure. If a user requests for flow or contingency data through the web screen, the HODP (the name of the platform developed in this study) server requests for the data from the servers of NTIC or from others in real time, and receives the data. The web UI expresses the data requested for by the user through visualization.

Figure 3: Traffic data UI
The review of the weather data characteristics of Bundang-Gu in September 2015 revealed that the weather had the characteristics of late summer and early autumn in that it was sunny, there was a wide daily temperature range, and it was not raining often. The traffic characteristics of the days with rain and the days without rain were compared to analyze the data for predicting the Bundang-Gu area driving environment. September 5 was the only day with rain within the collected period, and the traffic data of that day were compared with those of September 12 and 19, which were expected to have similar characteristics as September 5.

5. ANALYSIS CONTENTS AND RESULTS

In this section, the data collected in real time were analyzed. To determine the analysis method that was available for predicting the driving environment, the traffic flow data of the Sunae subway station area in Bundang-Gu, among all the data collected in real time, were used before developing an analysis method. The real-time data that were available with regard to the Sunae subway station area were the traffic flow and weather data. These data for September 2015 were used for analysis.
The traffic time comparisons of September 5 (with rain) and September 12 (without rain) showed that there was a difference in traffic time between September 5 and 12. There was also a difference in traffic time between September 12 and 19. It was estimated that the traffic time on September 19 increased as many people were visiting their hometowns to celebrate Chuseok, a big South Korean holiday, on that day. The convergence and analysis of the event data, including the accident data, by collecting more data through the big-data collection and storage system is required in the future.

6. CONCLUSION

A road big-data collection and storage platform for predicting the driving environment was developed in this study. Although many studies predicted traffic data through big-data collection, few predicted and provided the traffic environment. Providing the traffic environment in addition to traffic data, and predicting traffic data, will allow the users use these data more conveniently.

Horton works, a 100% open source among the open-type big-data open platforms and considered a more specialized platform, was used in this study for the construction of the platform. The data to be collected for predicting the driving environment were largely divided into traffic data, road environment data, exterior environment data, and road data. Each of these data were divided into the smaller data required for predicting the driving environment. Some of the data were provided as statistical data and could be collected every month or year, while the other data were provided in real time. Therefore, different collection methods were applied according to the characteristics of the data. The data from the sites that provided statistical data were collected through Sqoop, which can retrieve a large amount of data at a time, as there was no need for real-time collection.

The data that can be collected in real time through Open API were collected through Flume. The collected data were stored in the Hadoop-based HDFS and were expressed on the web. The analysis of the collected traffic data and real-time weather data revealed that the traffic time had somewhat different results depending on the precipitation. However, it was estimated that various analyses, including traffic time analysis through the convergence of the event data, including accident data, are required for predicting data during a somewhat unpredictable period. This study had limitations in the analysis as the collected real-time data were not enough. More data need to be collected for the construction of the road big-data collection and storage system. Also, researches are required on how the collected data will be analyzed, how the driving environment will be predicted, and what data the users will need.

Table 4: Result of paired t-test

<table>
<thead>
<tr>
<th>Category</th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>Standard Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>9/5 vs. 9/12 travel time t-test</td>
<td>-19.26</td>
<td>30.87</td>
<td>6.44</td>
</tr>
<tr>
<td>Category</td>
<td>t</td>
<td>Degree of freedom</td>
<td>P-value</td>
</tr>
<tr>
<td>9/5 vs. 9/12 travel time t-test</td>
<td>-2.99</td>
<td>22</td>
<td>0.007</td>
</tr>
<tr>
<td>Category</td>
<td>Mean</td>
<td>Standard Deviation</td>
<td>Standard Error</td>
</tr>
<tr>
<td>9/12 vs. 9/19 travel time t-test</td>
<td>17.7</td>
<td>31.4</td>
<td>6.54</td>
</tr>
<tr>
<td>Category</td>
<td>t</td>
<td>Degree of freedom</td>
<td>P-value</td>
</tr>
<tr>
<td>9/12 vs. 9/19 travel time t-test</td>
<td>2.7</td>
<td>22</td>
<td>0.013</td>
</tr>
</tbody>
</table>

The traffic time comparisons of September 5 (with rain) and September 12 (without rain) and of September 12 and September 19 (both without rain) showed that there was a difference in traffic time between September 5 and 12.
ACKNOWLEDGEMENT
This research was supported by a grant from a Strategic Research Project (Development of driving environment prediction platform based on big data) funded by the Korea Institute of Civil Engineering and Building Technology.

REFERENCES

AUTHOR PROFILES
Jung-Ah Ha received the doctor’s degree in transport engineering at the Ajou University in Korea. Currently, she is a researcher specialist at Korea Institute of Civil Engineering and Building Technology. Her research interest covers intelligent transport systems, traffic data analysis and big data.

Kyusoo Chong finished the doctor’s coursework in transport engineering from Seoul National University in Korea. Currently, he is a research fellow at Korea Institute of Civil Engineering and Building Technology. His research interest covers road/traffic data analysis and big data.