International Journal of Engineering Science Invention (IJESI)
ISSN (Online): 2319 – 6734, ISSN (Print): 2319 – 6726
www.ijesi.org ||Volume 7 Issue 8 Ver I || Aug 2018 || PP 74-80

Development of integrated system for providing real-time driving environment information based on individual vehicle sensors using experimental vehicles

Intaek Jung[1]
[1](Department of Future Technology and Convergence Research, Korea Institute of Civil Engineering and Building Technology, Republic of Korea)

Abstract: *Currently, the data-collection system installed on Korean roads only produces information about specific road points or sections in which fixed data-collection devices are installed, meaning that data about many areas remains unrecorded and thus inaccessible. The cost of installation of sufficient fixed collection devices at regular intervals to collect a full set of data is prohibitory. To address this problem, this study developed an integrated system to provide real-time driving environment information by using vehicle sensing data in conjunction with public data collected from individual vehicle sensors in order to compensate for the limitations of the existing fixed collection point system.*

Keywords- *Vehicle sensors, Vehicle sensing data, Driving environment information, Experimental vehicles, Public data*

Date of Submission: 21-07-2018 Date of acceptance: 6-08-2018

I. INTRODUCTION

The driving environment on any road can change rapidly due to various event factors including bad weather, traffic accidents, congestion, and other incidents. It is very important to provide real-time driving environment information to drivers as major traffic accidents can result if they fail to anticipate dangerous conditions in advance. In order to provide real-time driving environment information, data collection devices are required. Currently, a variety of fixed data collection systems exist to collect driving-environment data at specific points or intervals along roads, including loop detectors, CCTV, RSE, video detectors, and road weather measurement equipment. This fixed collection system is limited in time and space because it only provides information on the point or road section where the detector is installed [1]. With respect to road weather information in particular, the information provided by a specific weather station applies to all the road sections in the corresponding administrative district rather than local and heterogeneous weather information being provided for each road section. Vehicle sensors include GPS, temperature sensors, radar, and cameras, and data produced include coordinates, external temperatures, radar frequency data, and images. A mobile data collection system using vehicle sensing data of this nature collected from individual vehicle sensors can overcome the limitations of the fixed data collection system. It is necessary, therefore, to develop an integrated system to provide real-time driving environment information using the aforementioned individual vehicle sensors as a distributed data source. This study utilized experimental vehicles equipped with the same sensors as regular vehicles in order to generate typical vehicle sensing data collected from the sensors of general vehicles on the road. In addition, the government has encouraged the public sector to use public data since opening it to the private sector in 2013 [2]. Therefore, this study aims to develop an analysis system for generating and displaying real-time driving environment information using a combination of vehicle sensing data collected from individual vehicle sensors, and public data collected through Open API. More specifically, an integrated system comprised of two parts, hardware (HW) and software (SW), are developed to collect, store, analyze, and visualize vehicle sensing data and public data. The system is designed to provide road event information for conditions including road icing, road rainfall/snowfall, road incident, and traffic congestion using driving environment information generated by the system such as road surface temperature, precipitation, and traffic density.

II. LITERATURE REVIEW

The Ministry of Land, Infrastructure, and Transport, the Road Weather Information System (RWIS) includes providing changing road weather information to drivers in real-time using various fixed sensors installed at the roadsides and on the road surfaces. However, RWIS's weather information is limited to certain road sections where these sensors are installed. With respect to the general weather information provided by the Korean Meteorological Administration, data sourced from stationary weather stations installed at specific points provides weather information for the corresponding administrative district. The current traffic information system based on the Intelligent Transportation System provides a variety of traffic information to drivers in real-time. Most data generation sources use fixed sensors with space constraints such as loop detectors and image sensors, and mobile sensors such as vehicle sensors are not utilized at all. In some systems, On Board Equipment (OBU) and Road Side Equipment (RSE) are employed to utilize a mobile sensor based on a probe vehicle. However, in order to overcome space constraints, RSE can only be installed at certain intervals between roads [3], forming a sound justification for the development of mobile sensing technologies.

System development in this study can be classified into two parts, HW development and SW development. The contemporary focus on big data has spurred remarkable technical development of HW and SW. Major system HW development trends include high-performance low-power processor cores, processes for large data processing and memory integrated computing technology, high-performance and low-power next generation memory devices, and the commercialization of artificial intelligence computers. The major technological development trend of system SW has been the massive outpouring of open source programs offering functionality from big data collection to visualization, and there is an increasing emphasis on the contemporary importance of open-source SW system development [1, 4-6]. The open source approaches to big data collection include Crawling, File Transfer Protocol (FTP), Really Simple Syndication (RSS), Streaming, and Log collectors (such as Flume, Scribe, and Chukwa). The open source approaches to big data distribution processing and storage include Hadoop Distributed File System (HDFS), Spark, and Storm. Open source for database building is divided into relational database and non-relational database solutions. Relational database solutions include MySQL and PostgreSQL, whereas functionality for building a non-relational database includes Hbase, MongoDB, and CassandraDB as NoSQL. Open source options for big data analysis include Mahout, Zeppelin, and R. Finally, Prefuse, D3.js, Node.js, and Matplotlib offer open source solutions for big data visualization.

Table 1 lists the data collection status of major domestic and foreign road traffic systems. Most of them have their own internal collection system, utilizing real-time data collection and data from external systems. Weather data other than road traffic information is collected by some systems, but to an insignificant extent. In addition, data collection using fixed sensors is the norm, and there is no recorded use of vehicle sensing data collected from individual mobile vehicle sensors [1].

Table 1. Data collection status of road traffic systems

	Domestic system				Foreign system			
	NTIC	ROADPLUS	UTIC	TOPIS	RITIS	NPMRDS	VICS	TCC
Real-time data collection	○	○	○	○	○	○	○	○
Internal collection system	○	○	○	○	×	○	○	○
Use of external data	○	○	○	○	○	○	×	○
Weather data collection	×	×	○	○	○	×	×	○
Vehicle sensing data	×	×	×	×	×	×	×	×

※ NTIC (National Transport Information Center), ROADPLUS (Expressway Traffic Information Center), UTIC (Urban Traffic Information Center), TOPIS (Transport Operation and Information Service), RITIS (Regional Integrated Transportation System), NPMRDS (National Performance Management Research Data Set), VICS (Vehicle Information and Communication System), TCC (Traffic Control Center)

III. METHODOLOGY

1. Conceptual diagram for system development

In this study, experimental vehicles were fitted with the same vehicle sensors as those installed in a regular vehicle, in order to enable the experimental vehicles to collect representative data. Vehicle sensors installed on these vehicles include a GPS, temperature sensors, radar, and cameras. The integrated system collects vehicle sensing data generated from all these sensors. Whereas the temperature sensor data is simply collected as raw data, the vehicle sensing data generated from the radar is processed and transmitted through the primary data processing module. This design allows for future expansion of the range of collection targets, as it is difficult for the system to store the raw data for large-capacity radar. The wireless communication protocol used for real-time data transmission is Message Queuing Telemetry Transport (MQTT). Public data includes data collected in real-time (synchronously) through Open API, and offline data collected asynchronously. The public data collected in real-time are weather data, traffic data, accident data, and construction data. The asynchronously collected public data is digital tachograph (DTG) data of for freight cars, included to facilitate big data system performance evaluation. These collected public data are used as input parameters for the road surface temperature estimation module in the system. The integrated system to be developed in this study consists of a number of tasks and tools including data collection, data storage, data processing, a driving environment analysis tool, and an information visualization tool. Here, the driving environment analysis tool includes an information generation module for generating travel environment information such as road surface temperature, precipitation, and traffic density. Figure 1 shows the overall concept diagram for the development of an integrated system for providing real-time driving environment information using a combination of vehicle sensing data and public data.

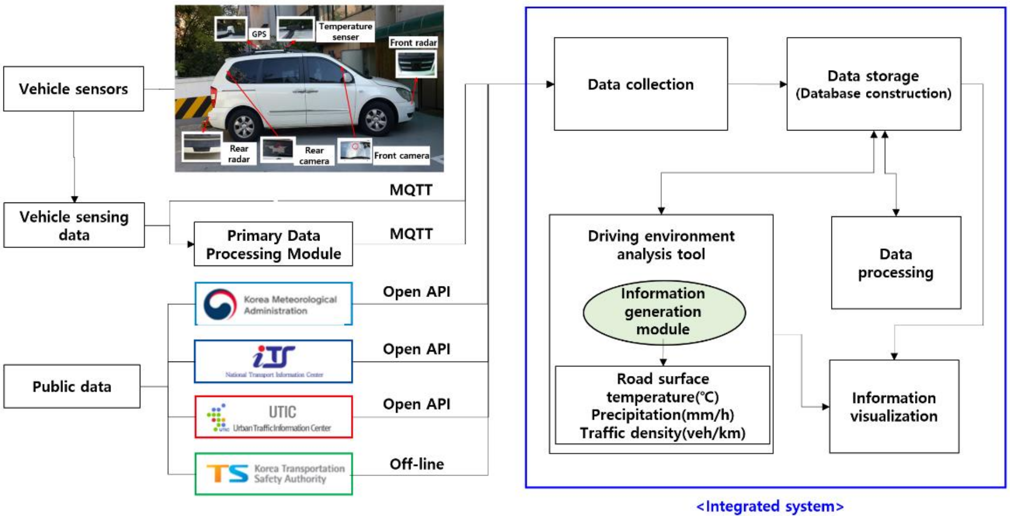


Figure 1. Structural conceptual diagram for integrated system development

2. Collection format of vehicle sensing data

Twenty-three types of vehicle sensing data, including GPS data, temperatures, and radar data, are generated from individual vehicle sensors (Table 2). Among these, the vehicle sensing data collected from the temperature sensor is the raw data, and that from the radar is the primary processed data. The collected vehicle sensing data is used as input data for generating driving environment information using the information generation module in the system.

Table 2. Collection format of vehicle sensing data

		Data name		Data type	Unit	Value
GPS		GPSTime		Time	(HMS)	8:24:29.000000
		Date		Data	(MDY)	9/17/2015
		Latitude		Float8	(+/- D M S)	37 39 38.41317
		Longitude		Float8	(+/- D M S)	126 43 00.58528
		H-Ell		Float4	(m)	38.983
		SDHoriz		Float4	(m)	0.929
		SDHeight		Float4	(m)	1.135
		VEast		Float4	(m/s)	0.000
		VNorth		Float4	(m/s)	0.000
		VUp Sol		Float4	(m/s)	0.00052
		Roll		Float4	(deg)	0.00000
		Pitch		Float4	(deg)	0.00000
		Heading		Float4	(deg)	0.00000
		SectionNo		Int	-	135 (0~65535)
		LaneNo		Int	-	2 (0~255)
		TotalLaneNo		Int	-	4 (0~255)
Temperature sensor		Outside temperature		Float4	(℃)	25.9
Radar	QC data	Point Reflectivity		Float4	(dBZ)	0.0000
		Point Radial Velocity		Float4	(mm/s)	0.0000
	Distribution of nearby vehicles	SVCount		Int	vehicles	4 (0~255)
		List of SVD(Surrounding Vehicle Data	Angle	Int	0.1deg	134
			Distance	Int	0.1m	200
			Speed	Int	0.1km/h	60

3. Communication protocol and information generation module

In this study, MQTT is applied as a wireless communication protocol to transmit vehicle sensing data collected from individual vehicles to the integrated system in real-time. MQTT is to a lightweight message transmission protocol based on push technology developed for optimal transmission in bandwidth-restricted communication environments such as Machine to Machine (M2M) and the Internet of Things (IOT). As shown in Figure 2, the MQTT protocol uses a message broker to send a specific message to the sender and subscribe the recipient to the message; messages are sent and received via the broker. The MQTT broker installed for this study is Mosquitto-1.4.15. The topic for message transmission is set as the travel environment data element name for each data type, such as temperature, rain, and density. MQTT is installed on the Linux Centos 6 operating system, and the MQTT QoS (Quality of Service) applied is simply level 0 because the message is transmitted and received.

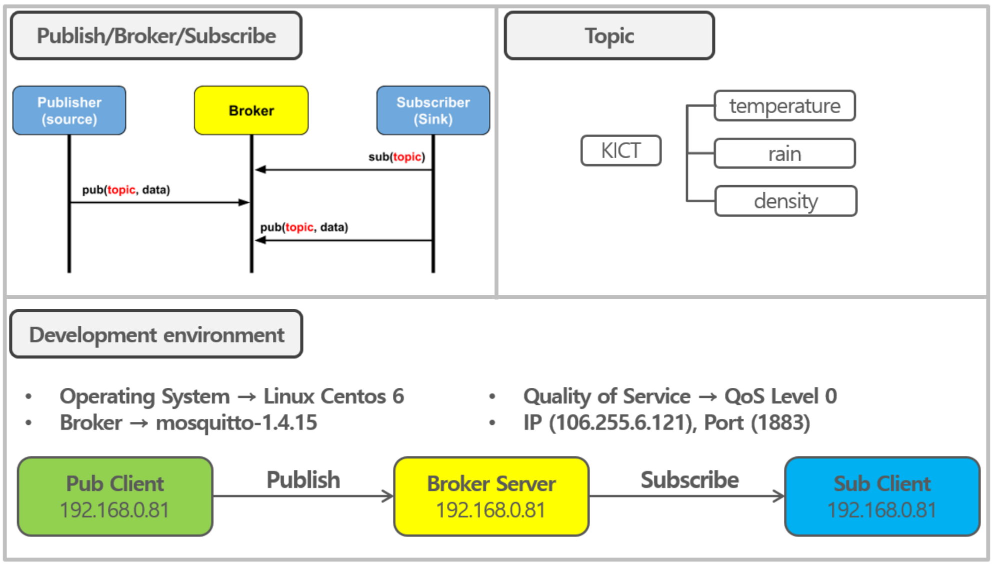


Figure 2. MQTT development environment

Using vehicle sensing data and public data, the information generation module of this study estimates three kinds of driving environment, specifically road surface temperature (℃), precipitation (mm/h), and traffic density (veh/km). First, the module for estimating the road surface temperature uses external temperature data collected from the temperature sensor data in the vehicle sensors dataset, and from the weather data in the public dataset. The weather data in use includes atmospheric temperature, humidity, and current weather conditions. The road surface temperature estimation model is a weighted k-nearest neighbors (KNN) model, as this model has the lowest estimation error of those available. The module for estimating precipitation uses point reflectivity and point radial velocity data, which are primarily processed from radar sensors. For the precipitation estimation model, a K-Band FMCW analysis model compatible with the installed radar is applied. Finally, the module for estimating the traffic density of a road point or section uses the distribution of nearby vehicles processed from the radar sensor dataset. The distribution data includes an SVCount and List of SVD (Angle, Distance, Speed). The traffic density estimation model applied is the Simulation-based Traffic Density Estimation Algorithm - Long and Short Term Memory (STREAM-LSTM) algorithm as this has the lowest error rate compared to the other models.

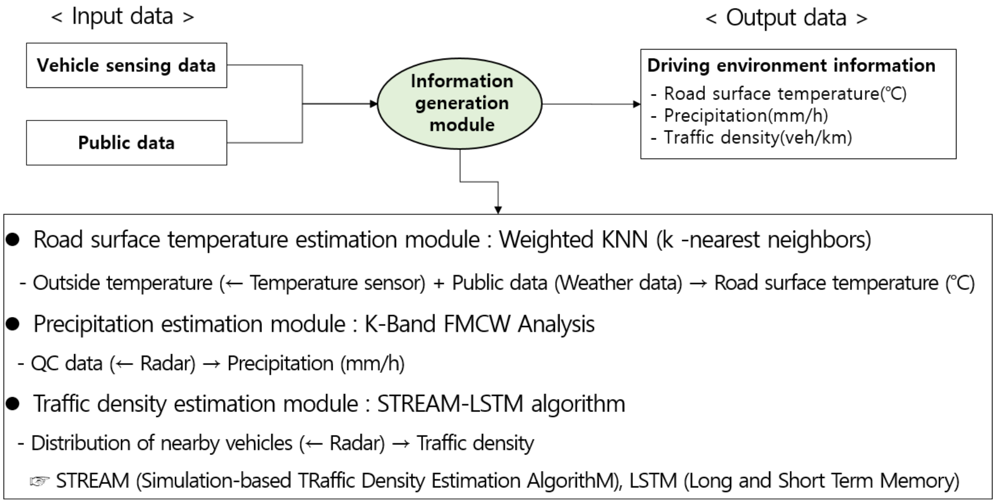


Figure 3. Structure diagram for generating driving environment information

4. Development of integrated system

The development of the integrated system in this study is divided into HW and SW development, as previously explained. First, the system HW is assembled, composed of a web server, network switch, system server, and server storage, as shown in Figure 4. The web server functions as a separate server for visualizing the generated travel environment information and a variety of service information. The network switch is a communication device that connects each server and storage device to a network; it only transmits data to the devices that need it, so bottlenecks do not occur as easily as they would if a simple hub was used. Server storage is responsible for backing up the database that is being built. The system server is developed to be small in size and have a parallel processing structure capable of driving 10 physical server nodes. The physical server nodes consist of one master node and nine slave nodes, and the physical server can expand the nodes according to the specifications desired by the user. If an HW fault occurs in one or more of these nodes, loss of node data occurs. In order to prevent such data loss, the HW the designed configuration is cluster-based.

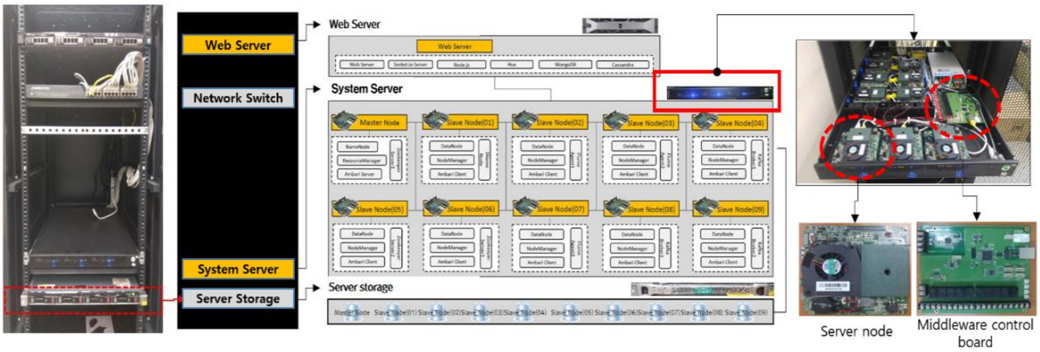


Figure 4. Development of System HW

Secondly, as shown in Figure 5, the system SW development involves the creation of individual programs for collection, storage, processing, analysis, and visualization of vehicle sensing data and public data using a variety of open source software. A collection interface for the collection software is developed using Kafka and Flume for real-time data and Sqoop for asynchronous data. The storage SW interface for storing collection data classifies data into the categories HDFS, HBASE DB, and Cassandra DB, according to its usage. Two data processing tools are developed, one based on in-memory processing using Spark and HDFS, and one based on disk processing using Yarn and HDFS. The smaller the data processing capacity is, the more advantageous the in-memory processing method is, and conversely, the larger the data processing capacity, the more advantageous the disk processing method is. For the processing SW a program is developed for matching the time and space units of various collected data by applying the grid index technique. For the analysis SW, a driving environment analysis tool is developed, based on the Zeppelin notebook for development and evaluation of development algorithms and the creation of a prediction model for the future driving environment. Finally, the information visualization SW developed for providing a variety of driving environment information and visualization is Geo Server, based on the web GIS engine program.

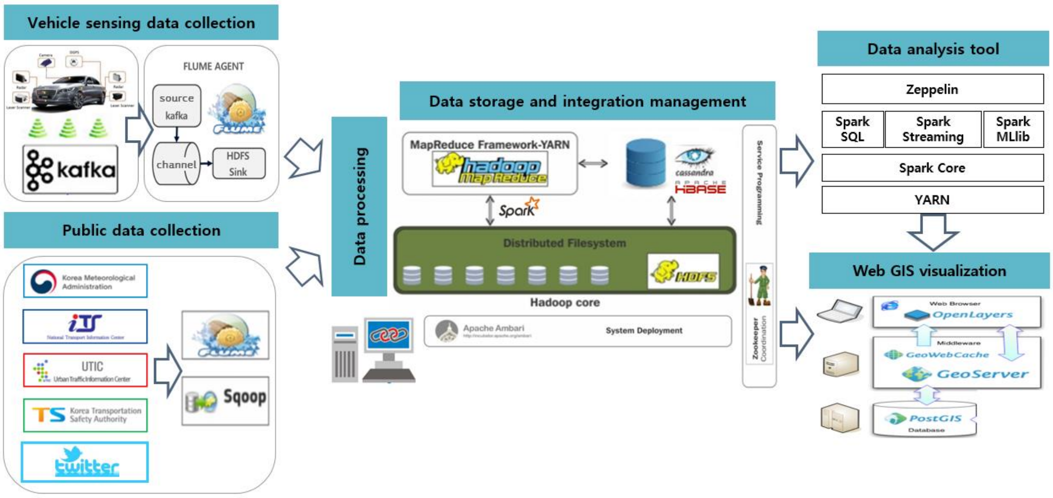
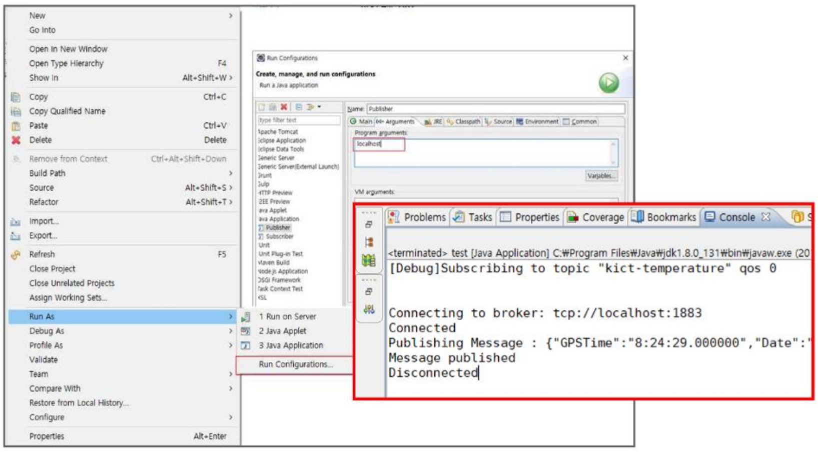


Figure 4. Development of system SW

IV. PERFORMANCE EVALUATION

The performance evaluation of the developed system is divided into HW and SW evaluations. The HW evaluation compares the performance of the system server with the server of the other organization (Korwa Institute of Science and Technology Information). The execution algorithm used for evaluating the performance of the two servers was the KNN algorithm [7]. The evaluation is run on 5 GB of data, and the system's performance is compared according to the number of the executor and the capacity of the executor in the Spark analysis environment. Results indicate that the performance of the development server is superior to that of the comparison server in all analysis environments.

The SW evaluation involves installing and testing MQTT for real-time data transmission and comparing the performance with two processing tools for the distributed storage of collected data. Paho was installed to implement the publisher and subscriber functionality of MQTT, and used to test real-time vehicle sensing data transmission (Figure 3). Results indicate that transmitting the vehicle sensing data of the experimental vehicle in real-time worked appropriately.

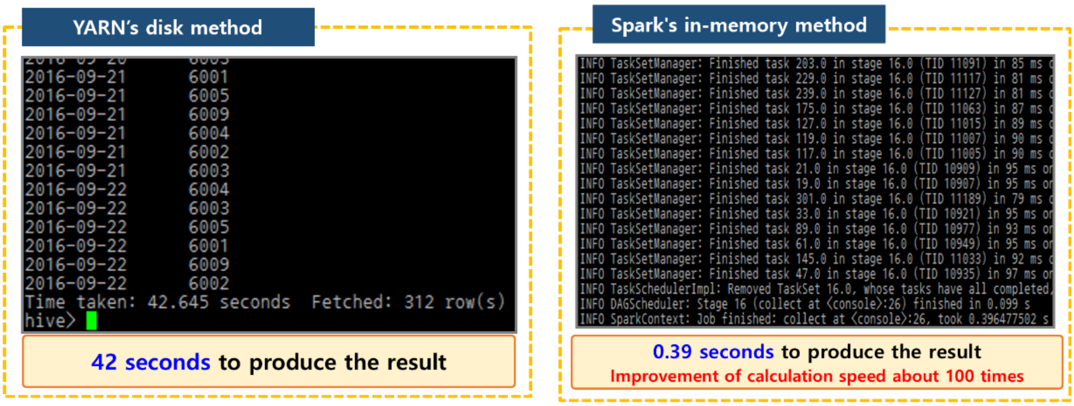


```
<terminated> test [Java Application] C:₩Program Files₩Java₩jdk1.8.0_131₩bin₩javaw.exe (20
[Debug]Subscribing to topic "kict-temperature" qos 0


Connecting to broker: tcp://localhost:1883
Connected
Publishing Message : {"GPSTime":"8:24:29.000000","Date":"
Message published
Disconnected
```

Figure 5. Results of MQTT test using Paho

Two processing tools are developed for the distributed storage of collected data: the first is an in-memory data processing tool using Spark and HDFS, and the second is a disk-based data processing tool using Yarn and HDFS. The performance comparison method implements the same aggregate query [avg (speed)] with Group By and Sort. Thirteen million data elements are used as the input data for the performance evaluation, to shorten test time among freight DTG data. The comparison (shown in Figure 6) indicates that the Spark in-memory method improves performance 100 times more than the YARN disk method. Future work includes a comparison of the two processing methods with varying input data sizes, and the establishment of quantitative application criteria for each method.



```
2016-09-21        6001
2016-09-21        6005
2016-09-21        6009
2016-09-21        6004
2016-09-21        6002
2016-09-21        6003
2016-09-22        6004
2016-09-22        6003
2016-09-22        6005
2016-09-22        6001
2016-09-22        6009
2016-09-22        6002
Time taken: 42.645 seconds  Fetched: 312 row(s)
hive>
```

42 seconds to produce the result

```
INFO TaskSetManager: Finished task 203.0 in stage 16.0 (TID 11091) in 85 ms
INFO TaskSetManager: Finished task 229.0 in stage 16.0 (TID 11117) in 81 ms
INFO TaskSetManager: Finished task 239.0 in stage 16.0 (TID 11127) in 81 ms
INFO TaskSetManager: Finished task 175.0 in stage 16.0 (TID 11063) in 87 ms
INFO TaskSetManager: Finished task 127.0 in stage 16.0 (TID 11015) in 89 ms
INFO TaskSetManager: Finished task 119.0 in stage 16.0 (TID 11007) in 90 ms
INFO TaskSetManager: Finished task 117.0 in stage 16.0 (TID 11005) in 90 ms
INFO TaskSetManager: Finished task 21.0 in stage 16.0 (TID 10909) in 95 ms
INFO TaskSetManager: Finished task 19.0 in stage 16.0 (TID 10907) in 95 ms
INFO TaskSetManager: Finished task 301.0 in stage 16.0 (TID 11189) in 79 ms
INFO TaskSetManager: Finished task 33.0 in stage 16.0 (TID 10921) in 95 ms
INFO TaskSetManager: Finished task 89.0 in stage 16.0 (TID 10977) in 93 ms
INFO TaskSetManager: Finished task 61.0 in stage 16.0 (TID 10949) in 95 ms
INFO TaskSetManager: Finished task 145.0 in stage 16.0 (TID 11033) in 92 ms
INFO TaskSetManager: Finished task 47.0 in stage 16.0 (TID 10935) in 97 ms
INFO TaskSchedulerImpl: Removed TaskSet 16.0, whose tasks have all completed
INFO DAGScheduler: Stage 16 (collect at <console>:26) finished in 0.099 s
INFO SparkContext: Job finished: collect at <console>:26, took 0.396477502 s
```

0.39 seconds to produce the result

Improvement of calculation speed about 100 times

Figure 6. Comparison result of two data processing methods

V. CONCLUSION AND FUTURE WORK

In this study, I developed an integrated system for providing real-time driving environment information based on individual vehicle sensors using experimental vehicles. Additionally, system HW and SW were developed that can utilize the vehicle sensing data collected from future instances of probe vehicles. The performance of the HW and SW components were evaluated separately. Analysis of the performance of the system HW showed it to be superior to that of other HW organizations given the same server specification. In the evaluation of the system SW, MQTT, a wireless communication protocol, was installed and performance tests conducted for the transmission of real-time vehicle sensing data. Results indicated that real-time data transmission was reasonable. The performance of different data processing tools for distributed storage of collected data differed depending on the usage variables: the disk method using Yarn is more advantageous when the data volume is larger, and the in-memory method using Spark is more advantageous for processing smaller data volumes. Future work will address upgrading the system HW and SW according to future data expansion, and further diversification of the performance evaluation method to assess these upgrades. It will also be necessary to expand the system in order to utilize the vehicle sensing data collected from a multitude of probe vehicles rather than just an experimental vehicle. Finally, it will be necessary to develop various service information contents using the generated driving environment information in combination with collected public data.

Acknowledgements

This research was supported by a grant from a fusion research project (Development of Driving Environment Observation, Prediction, and Safety Technology based on Automotive Sensors) funded by the Korean Institute of Civil Engineering and Building Technology.

REFERENCES

[1]. I. Jung, K. Chong, Development of Information Technology Infrastructures throughConstruction of Big Data Platform for Road Driving EnvironmentAnalysis, Journal of the Korea Academia-Industrialcooperation Society, 19(3), 2018, 669-678.

[2]. National Information Society Agency, Open Data Portal.

[3]. I. Jung, H. Sung, K. Chong, Potential Utilization of Observed Sensor Data From Individual Vehicles For Improving Road Traffic Safety, International Journal of Advances in Mechanical and Civil Engineering, 3(6), 2016, 27-31.

[4]. K. Shvachko, H. Kuang, S. Radia, R. Chansler, The Hadoop Distributed File System,Proceedings of the 2010 IEEE 26th Symposium on Mass Storage Systems and Technologies(MSST), 2010, 1-10.

[5]. A. Lakshman, P. Malik, Cassandra: a Decentralized Structured Storage System, ACM SIGOPS Operating Systems Review, 44(2), 2010, 35-40.

[6]. R. Padhy, M. Patra, S. Satapathy, RDBMS to NoSQL: Reviewing Some Next-Generation Non-Relational Database's, International Journal of Advanced Engineering Science and Technologies, 11(1), 2011, 15-30.

[7]. I. Jung, AADT Estimation of Unobserved Road Segments Using GPS Vehicle Trip Data, Seoul National University Ph.D thesis, 2016.

Intaek Jung "Development of integrated system for providing real-time driving environment information based on individual vehicle sensors using experimental vehicles "International Journal of Engineering Science Invention (IJESI), vol. 07, no. 8, 2018, pp 74-80