2015 15th International Conference on Control, Automation and Systems (ICCAS 2015)
Oct. 13-16, 2015 in BEXCO, Busan, Korea

A Big Data System Design to Predict the Vehicle Slip

Joohyoung Jeon[1*], Woosik Lee[2], Hyo Joo Cho[3], Hongchul Lee[4]

[1]School of Industrial and Management Engineering, Korea University,
Seoul, 136-713, Korea (joohyoung@korea.ac.kr) * Corresponding author
[2] Department of Mechanical Engineering, Korea University,
Seoul, 136-713, Korea (WeeingChiking@gmail.com)
[3] Department of Mechanical Engineering, Korea University,
Seoul, 136-713, Korea (djowlrid@korea.ac.kr)
[4]School of Industrial and Management Engineering, Korea University,
Seoul, 136-713, Korea (hclee@korea.ac.kr)

Abstract: In this paper, we propose a system storing sensor data generated from the vehicle using Big data system and analyzing real-time sensor output to predict the car slip. A large amount of data will be distributed and stored by MongoDB which is one of NoSQL Database, and analyzed to control the vehicle by Hadoop Map/Reduce. The proposed system can be applied to general vehicle and also to the field of driverless vehicle or driverless vehicle platooning.

Keywords: Car-Slip, Big data, NoSQL, Extended Kalman Filter.

1. INTRODUCTION

Car Registrations in South Korea increased to about 20million in 2014 and 220 thousand of road traffic accidents were reported to occur each year [1]. Among the road traffic accidents, the car slip account for a slippery road can cause critical damage to both driver and pedestrian. Car slip can be caused by many factors like each wheel's linear velocity, cornering force and rigidity modulus [2]. By the influence ratio of these factors, various types of the car slip occur [3]. On account of various influence factors and diverse resultant slip, it is hard to interpret cause and effect relation between them. Also the large size of data to interpret makes the analysis harder.

To solve these issues, big data processing applications are have been studied and NoSQL (Not only SQL) database makes it possible to provide a mechanism for storage and retrieval of data that is modeled in means other than the tabular relations used in relational databases.

In this paper, we propose a car slip prediction system analyzing real time output data by using Map/Reduce model and at the same time, storing massive amount of data extracted from the sensors measuring the car slip cause. There are many systems for recovery from a wheel slip condition but few systems for prediction before the slip occurs and the proposed system highly expected to enhance the safety of driver and pedestrian.

The remainder of this paper is organized as follows. Technologies related to car slip and component of big data processing system are presented in Section 2. And section 3 shows a research model. In section 4, the proposed system is presented. Finally, the conclusion of this research is presented in Section 5.

2. BACKGROUNDS

2.1 Mechanical Engineering Field

Many researches have conducted about the safety of vehicle. The Anti-lock Brake system (ABS) is an automobile safety system which prevents the wheels from locking a hard braking [4]. The Traction Control System (TCS) enables the driver to trace corners on a paved road smoothly and safely by preventing excessive lateral acceleration [5]. VSC is a vehicle stability control system that improves active safety performance when the vehicle skids [6], and *Bosch., Ltd.* suggested VDC [7]. Also, *Continental TEVES., Ltd.* developed ESP. Steadily, there are many researches about system development and performance evaluation of ESP for enhancement of driving stability while cornering.

Techniques for the fast recovery position by observing vehicle's slippage also have been studied in various fields. There is interesting research which estimates longitudinal wheel slip and detecting immobilized conditions by using Encoders, IMU, and GPS [8]. Another research predicts vehicle's sideslip angle 0.5 second in advance applying neural network [9]. However, these researches have a limitation which only focused on recovery of vehicle's stability after slip and it is insufficient that study about prediction before the vehicle's slip occur.

2.2 NoSQL Database

The explosive increase of the data has led to the limitations of the physical storage device and it made a new computing environment systems so called "Cloud System" and "Distributed System."

In these Big Data Environment, NoSQL(Not-Only Structure Query Language) database has emerged. NoSQL is no need to pre-defined of the Schema unlike traditional relational database and it is easy to support horizontal-scale up and manage unstructured data.

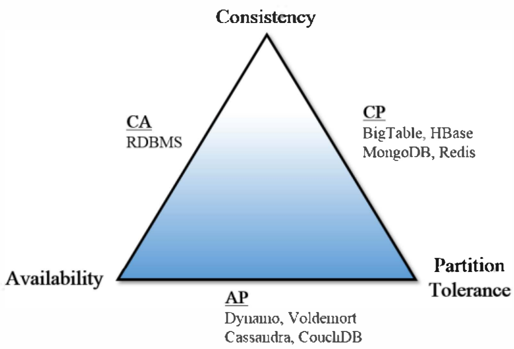

Fig. 1 The CAP theory of NoSQL

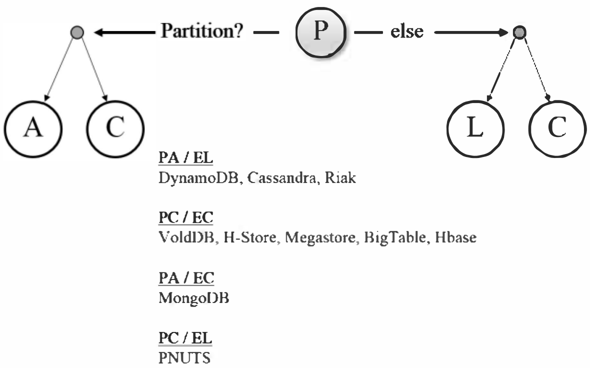

Fig. 2 The PACELC theory of NoSQL

Fig. 1~2 show two theories to explain the NoSQL. In Fig. 1 CAP theory is classified according to the characteristics of the selection of three network properties of the Consistency, Availability, and Partition-Tolerance. Whereas in Fig. 2 PACELC theory points out that CAP theory is placed for Partition-Tolerance features as optional.

According to the PACELC theory, Network Partition will not be selected because the problems that inevitably occur in a distributed network environment [10].

PACELC theory was selected as a priority classification of Network Partition situation of NoSQL.

When Network Partition situation occurs (in Partition State), NoSQL select one of the Availability or Consistency. In normal network state, it also select one of Latency or Consistency. Each of the sub-functions is trade-off.

Table 1. The functional classification of NoSQL.

Function	Name
Column/ Column Families	Hadoop/HBase, Cassandra, Hypertable, Accumulo, Informix *et al.*
Document-oriented	Elasticsearch, MongoDB, Couchbase, RethinkDB *et al.*
Key Value/Tuple	DynamoDB, Riak, Redis, Aerospike *et al.*
Graphic Database	Neo4J, Infinite Graph, *et al.*

Table. 1 shows a functional classification of NoSQL. Column/Column Families will query the data in column basis. Document-oriented database system is useful for storing data that is insert as a time period. Key Value/Tuple NoSQL Database manages the date using a pair of key-value. Graphic Database is useful for storing graphic information and it supports an ontology-based queries [11].

2.3 Briefly Introduction to MongoDB

The MongoDB is developed based on the C++ programming language in *10Gen* and it is open source database system. According to PACELC theory, MongoDB will select Availability in Network Partition situation. If due to network failures or expansion fail-over occurs, MongoDB is replicated asynchronously to the slave node from the master node. Because of this asynchronous replication, a degradation in the MongoDB service does not occur even when the network error or network expansion occurs. If degradation of database occurs because of the network failover, it can lead to delayed response of the entire system.

The MongoDB supports a distributed system environment, such as *HDFS*. It is possible to storing a large amount of sensor data by using the *Auto-Sharding* functions. *Auto-Sharding* is storing the data to multiple lower PC. Therefore, it is much cheaper than a relational database [12]. It also can store a large amount of data at very high speed since the In-Memory based database [13]. In consideration of the availability of the database, performance, and cost, MongoDB is selected to be the most suitable database system to store sensor data.

2.4 Map/Reduce

Map/Reduce is an open-source framework, which emerged from a *Google's* paper published in 2004 [14]. It is useful framework to split a large amount of data (Map) and combine the data (Reduce). Hadoop Map/Reduce job is divided into a set of Map task and a set of Reduce task.

Map task job extracts the required data from the data by loading specific dataset, analysis, conversion, and filtering. Reduce task job is processing a portion of the Map task outputs. It copies the intermediate data processed by the Mapper task to classify and aggregate data. Due to this method Map/Reduce is useful for the analysis of large amounts of data in real time [15].

3. RESEARCH MODEL

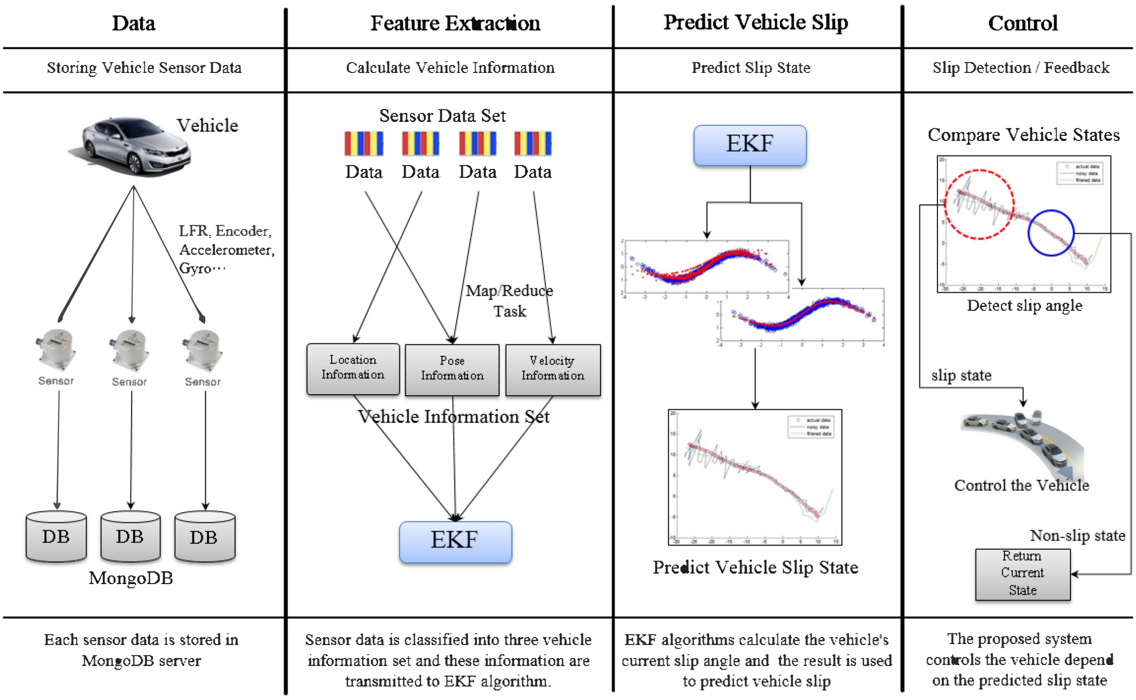

Fig. 3 The Research Model

Fig. 3 is a research model proposed in this paper. The research model is a kind of equation. When the independent variables are inserted in the research model, the dependent variables are calculated. In this paper, the proposed research model is to predict the slip of the vehicle.

Data Region of the left in Fig. 3 refers to the independent variables which are input to the system. The independent variables are consists of the data set generated from the vehicle sensors. Extraction and calculation of the data is necessary to determine position, pose, and velocity information of the vehicle from the sensor data.

For the extraction and calculation of data, we use the Map/Reduce system. Since the system must derive the vehicle information from the sensor data, Mapper Job divides the sensor data and stores it to temporary space.

After Map Job is processed, Reducer Job collects Maps to generate the three vehicle information. When the vehicle information is generated by the Reduce Job, the resulting vehicle information is input to the *EKF* (*Extended Kalman Filter*) algorithm. It calculates the slip angle of the vehicle through the received position, pose, and velocity information of vehicle.

The sideslip angle calculated by *EKF* is used as a variable to predict vehicle's slip. When predicted sideslip angle is over the threshold value, *"5mm/Meter"* that determined by system, the system automatically controls the vehicle.

However, when if sideslip angle is not over the threshold value, system returns calculated vehicle state and it is used to machine learning. Learned data develops more accurate data by comparing with calculated at later variable.

In summary, this proposed research model, the independent variables are sensor data and the dependent variables are vehicle's control signal or state.

4. DESIGN OF SYSTEM

Fig. 4 shows the data flow diagram of proposed system. The dashed line represents the data packet flow and the solid line indicates the vehicle control signal. The proposed big data system is consist of three systems. One is data storage area, another is data analysis area, and the other is vehicle control center.

4.1 MongoDB Database system

MongoDB database system store the data generated by the car sensors (LFR, Encoder, Accelerometer, Gyro, Handle, GPS) in the collection by chronological order. The database system administrator can control the *Config server*, *Shard server*, and *ChunkSize* using the built-in *MongoS* applications. If there is limitation to capacity in a single-node, the administrator or system divides the server by using the *Auto-Sharding* function of *MongoS*. By using *ChunkSize* command, it can be calculated the default size of the data to be stored in *Auto-Sharding* environment. In addition, MongoDB will automatically create a *Replica-set* to back up the stored data.

The data can be stored in MongoDB is normally raw-data sets which is generated by the car sensors. Therefore, for a precise prediction, raw sensor data should be refined. For data refinement, the following system of Map/Reduce processes the data.

4.2 Hadoop Map/Reduce System

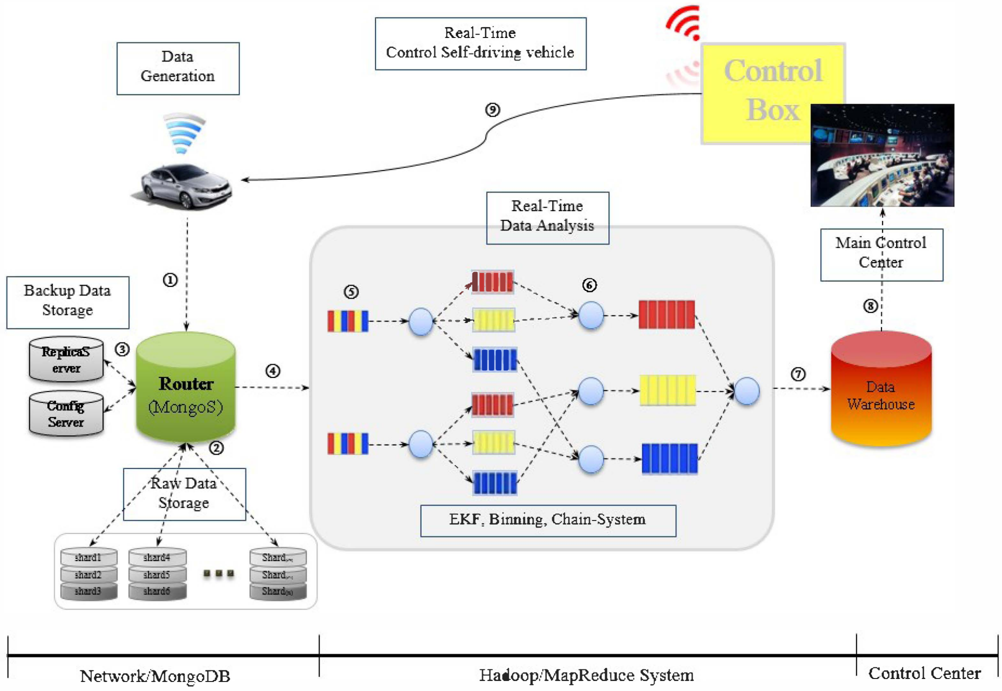

Fig. 4 The Data Flow Diagram

Hadoop Map/Reduce system is used to analyze the data stored in MongoDB. It extracts and calculates the necessary data from the sensor dataset by using Map Task (Mapper Job). In addition to transmit the vehicle information to EKF algorithm on the Map/Reduce System, Map Task corrects the position, pose, and velocity information of the vehicle in real-time basis.

Current vehicle state value that is returned to the time t through *EKF* is used to the basis of data to be analyzed in the time $t+1$. In this case, by comparing the state value of t and the state value of $t+1$, Reduce task transmits a warning signal to the control center when the compared result is out of *5mm/Meter*.

4.3 Control Center

Control center controls the vehicle. If the warning signal is generated in the Hadoop Map/Reduce system and sent to the control center. When the warning signal is transmitted to the control center, it immediately controls the vehicle to prevent the car slip. The Control Box embedded in the vehicle upon receiving the control signal from the Control center controls the vehicle. Hence, the control signal is to be secured thoroughly in a unique frequency band. In addition, it stores the vehicle sensor values and car status values in the Data Warehouse. And the control signal must be encrypted utilizing the *WPA2-Enterprise* technology.

5. CONCLUSION

This paper represents the car slip prediction system based on big data system and sensor data set from the driving vehicle. The proposed system aims to predict sideslip angle and to prevent vehicle slippage by applying *Extended Kalman Filter* based on the sensor data set from the driving car. Due to reduction of car slip accident, it is expected to decrease the damage on human life and property. Furthermore, it is able to analyze the data in real-time using Map/Reduce function of MongoDB and when the data is stored, the storage is expanded horizontally. Thus it will be the required system for vehicle platoon system.

In future works, we will research on prediction system that gives driver more accurate prediction slip information based on actual vehicle's data analysis and we will conduct a study to replace MongoDB with *IBM's Cloudant*.

REFERENCES

[1] National Police Statistical information of Traffic. *Websites*: http://www.police.go.kr/portal/main/contents.do?menuNo=200193#aco02

[2] W. Leroquais and B. d' A. Novel, "Modeling and control of wheeled mobile robots not satisfying ideal velocity constraints: The unicycle case," in *Proc. 35th IEEE Conf.Decis. Control,* Dec. 1996, vol. 2, pp. 1437–1442.

[3] Trachtler and Ansgar, "Integrated vehicle dynamics control using active brake, steering and suspension systems." *International Journal of Vehicle Design* 36(1) pp. 1-12, 2004.

[4] Suarez, H, J. W. Barlow, and D. R. Paul, "Mechanical properties of ABS/polycarbonate

blends." *Journal of applied polymer science* 29.11 pp. 3253-3259, 1984.
[5] Tanaka, Tadao, and Keiji ISODA, "Traction Control System." *JSME international journal.* Ser. 3, Vibration, control engineering, engineering for industry 35.1, pp. 116-120, 1992.
[6] Toyota, co. "Vehicle stability control system". *Jsae.or.jp.* Retrieved, 2011.
[7] Van Zanten, Anton T., Rainer Erhardt, and Georg Pfaff. *VDC,* "the vehicle dynamics control system of Bosch." No. 950759. *SAE Technical Paper*, 1995.
[8] C.Ward and K. Iagnemma, "Model-based wheel slip detection for outdoor mobile robots," in *Proc. IEEE ICRA,* Apr. 2007, pp. 2724–2729.
[9] Du, Xiaoping, et al. "A prediction model for vehicle sideslip angle based on neural network." *Information and Financial Engineering (ICIFE)*, 2010 2nd *IEEE International Conference on. IEEE*, 2010.
[10] D. J. Abadi, "Consistency tradeoffs in modern distributed database system design: CAP is only part of the story." *IEEE Computer*, 45(2):37-42, 2012.
[11] Stefan Edlich, "List of NoSQL Databases." *Websites*: http://nosql-database.org. 2009.
[12] Y. Liu, Y. Wang, and Y. Jin, "Research on the Improvement of MongoDB Auto-Sharding in Cloud Environment." *7th International Conference on Computer Science and Education,* 2012.
[13] Banker, K, "MongoDB in Action." *Manning Publications* 2012.
[14] Dean, J. and Ghemawat, S, "MapReduce: Simplified data processing on large clusters." *In Proceedings of the 6th USENIX Symposium on Operating Systems Design and Implementation,* pp. 137–150, 2004.
[15] D. Miner and A. Shook, "MapReduce Design Patterns: Building Effective Algorithms and Analytics for Hadoop and Other Systems." *O'Reilly Media*, 2012.