A framework for discovering relevant patterns using aggregation and intelligent data mining agents in telematics systems

Bobby D. Gerardo\textsuperscript{a,}\textsuperscript{*}, Jaewan Lee\textsuperscript{b}

\textsuperscript{a}Institute of Information and Communications Technology, West Visayas State University, La Paz, Iloilo City, Philippines
\textsuperscript{b}School of Electronic and Information Engineering, Kunsan National University, 68 Miryong-dong, Kunsan, Chonbuk 573-701, South Korea

\textbf{A R T I C L E  I N F O}

\textbf{Article history:}
Received 29 April 2007
Received in revised form 21 April 2008
Accepted 19 May 2008

\textbf{Keywords:}
Telematics systems
Aggregation
Data mining agent
Flexible insurance

\textbf{A B S T R A C T}

The emerging technology in vehicle telematics drives several stakeholders in this field to consider services that could be beneficial for both clients and the telematics service providers. In particular this paper proposes a novel framework for insurance telematics in Korea using a mobile aggregation agent (AA) and intelligent data mining agent (IDMA). To our knowledge, this model is recent of its kind in this country and the base-line information from driver's characteristics serves as reference for the flexible insurance policies. We are able to present a use-case scenario and illustrative examples to demonstrate our model. With this flexible insurance framework, customers can manage their own insurance premiums and lower the cost of motoring. Promising applications of this system to business and industries have been recognized and discussed.

\textcopyright{} 2008 Elsevier Ltd. All rights reserved.

1. Introduction

Emerging technologies in telematics provide opportunities for data mining, cross-selling and even new products and services based on better perception of customer behavior. Several companies with promising visions and innovative ideas are beginning to see this wirelessly enabled and computer-enhanced set of solutions as an essential tool for doing business.

The current trends in vehicle telematics drives several stakeholders in this field to consider services that could be beneficial for both clients and the telematics service providers. Vehicle data recorders and telematics are not a future possibility, but currently available technologies (Chen et al., 2003; Munson et al., 2005; Nolte et al., 2005; Bisdikian et al., 2002). As costs of telematics devices decrease, many industries, including insurance, are taking a second look at utilizing these devices to improve services, and to develop innovative solutions.

On the other hand, recording devices such as black boxes (Chen et al., 2003; Bisdikian et al., 2002; ABI Research) can range from passive to real-time types where it can record vehicle conditions like speed, acceleration, position, and as well as crash data that includes information collected by the device prior to the accident.

Distributed environment implementation may utilize multi-agents to handle communication and messaging as part of the telematics systems. Multi-agents are intended for communication and cooperation in which they have the ability to behave socially, to interact and communicate with other agents like exchange information, receive instructions and give responses and cooperate when it helps them fulfill their own goals (Helm et al., 2006). The role of agents for distributed information management, which include resource discovery, information integrity and navigation assistance is perceived to be important. Multi-agents can then be employed in ubiquitous environment such as those intended for telematics application.
This paper proposes a novel framework for data aggregation and insurance service in vehicle telematics for Korean motorists using multi-agents. To our knowledge, this framework is new of its kind in this country and the baseline information from driver’s characteristics and patterns serves as reference for the development of flexible insurance policies which in return could benefit vehicle owners for a reduced premium.

The proposed mechanism of aggregation agent resides at the vehicle side. We assume that a vehicle is mobile and could have ability to transmit data in both real-time and non-real-time conditions. We will use transformation and visualization of data to discover driving patterns. Also, we propose an algorithm for flexible insurance policies based on expert opinions. Reduced premiums can then be achieved by correlating driving patterns to flexible insurance policies.

The main motivational factor for this study is to develop a functional telematics system that collects data from sensors, sends these to a data center, extracts information from the database using our proposed iDAMA, and transforms the data into knowledge that can be utilized for insurance telematics.

2. Related studies

2.1. Telematics services

Vehicle telematics allows the deployment of several new services and applications integrating wireless communication technology into an automobile. As a result, the vehicle acquires new capabilities and offers more services to its users.

Finnegan and Sirota (2004) and Moskowitz et al. (2002) presented some examples of telematics services and applications such as: (1) Navigation and traffic information system as a service where vehicle equipped with a telematics unit can direct a driver to a desired location, while providing real-time traffic information for a given route. In (2) voice recognition and wireless Internet connection service, drivers and their passengers can receive and send voice activated e-mails while on the road. Services of (3) which are on safety systems include collision avoidance systems, unsafe driving profiling and intelligent airbag deployment systems. (4) Security systems refer to service for vehicle antitheft and stolen vehicle tracking services. While (5) diagnostics and maintenance services provide remote diagnostics and/or maintenance systems, vehicle and driver monitoring.

One drawback of implementing insurance telematics is the cost of recording device. Earlier implementations of the systems did not progress well due to high cost of devices which eventually forced some companies like Progressive to stop the project (Finnegan and Sirota, 2004). Insurance companies must be aware of what is possible these days and in the future. Black boxes used for vehicle recording come into varieties such as in the form of trip logging, passive GPS tracking, and crash data recording devices.

2.2. Motivations

A classic method for data retrieval in vehicles is done through plugging-in of telematics device to a computer to obtain the data stored in it. Modern approach can be used using emerging technology for telecommunications. Insurance telematics is not yet popular in Korea, so there is a high probability that flexible insurance or pay as you drive scheme will be widely acceptable. We can develop proprietary scheme for Insurance telematics based on drivers’ behavior. This framework can encourage TSP and telecommunications collaboration which will yield rapid deployment of the system. Aggregation of information can yield comprehensive data warehouse that can be utilized for driving pattern discovery, traffic prediction and forecasting. This information can help road safety and reduce accidents.

The contributions of this paper include the design of three-layered architecture for the distributed system handling data aggregation from mobile vehicles and mining of data from distributed databases. Formulated discretization (delta function) mechanism to process continuous data and provide supervised mechanism for expert opinion on discrete classes is presented. We use Java 2 Micro Edition (J2ME) Wireless Toolkit to generate raw sensor data, which is based on connected limited device configuration (CLDC) and mobile information device profile (MIDP). In this paper, the wireless toolkit was used to develop a simulated telematics system generating driving attributes using random values for sensor data in a fixed time interval. Definitions, theorems and corollary are formulated as requisite and consequence of the development of our rule-based functions. These are later presented in pseudo-code form and converted to actual program codes in the implementation. We provided integration mechanism for combining data in different data centers. Also, we proposed and designed the distributed SQL queries for data extraction and its interoperability through CORBA architecture. This paper, proposes to develop intelligent data mining agent (iDAMA) integrating Chi-square function using Java for the implementation and testing of driving schemes. The iDAMA can generate models in vehicle telematics, which can be implemented for flexible insurance service.

2.3. Insurance telematics

Auto insurer initiatives can be traced back as early as 1998 where Progressive (Finnegan and Sirota, 2004) launched the pay as you drive service using GPS tracking device. In the same year Norwich Union (Finnegan and Sirota, 2004; United Kingdom Limited, 2004; Norwich Union, 2003) in UK piloted the same pay as you drive based on Progressive methodology.
Norwich initiative is using IBM Black Boxes for flexible insurance. This method of paying motor insurance will not only appeal to drivers who recorded low mileage but it could also attract high mileage drivers if they avoid busy city centers or drive at off-peak hours.

GMAC in USA also piloted an insurance discount based on mileage. Other initiatives were piloted by AXA of Ireland which gives discounts to young drivers if they can maintain a designated speed limit. Japan’s AIOI insurance had also initiated the pay as you drive service.

2.4. Multi-agents in vehicle telematics

As mentioned in Duri et al. (2004) mobile agents are used to tamper resistant hardware to protect privacy. Parties requesting access to private information receive mobile agents that encapsulate private data and access control policies. The agents are executed in a protected environment to access data. A paper on control system tasks using multi-agents (Helm et al., 2006) demonstrates that cooperative behavior can be achieved without extensive inter-agent communications. This means that multi-agent design for mobile object where resource is scarce is promising because overhead cost for communication can be reduced. This implies that multi-agent implementation in telematics system is potentially good.

3. System architecture

In this section, we present the proposed global architecture of the telematics system as shown in Fig. 1. There are ranges of wireless technologies for vehicle telematics that are described by Finnegan and Sirota (2004) where its viability will depend on the different target applications they were optimized for. Some examples are Bluetooth, ZigBee, UWB, and Wi-fi. On the other hand, InternetCar as cited by Chen et al. Chen et al. (2003) proposed architecture to connect car to the Internet.

Some other workable communication medium to transmit data from a mobile vehicle is described in the primer on real-time traffic system (Helm et al., 2006), which includes cellular technologies such as GSM, CDMA, GPRS, EDGE, WCDMA, TDMA, iDEN, and WiDEN. These cellular technologies are more suitable for real-time data delivery due to its availability and the types of network it can support.

Hence, in this framework we chose CDMA as a medium for data delivery of real-time in vehicle information because of its suitableness to the geographical implementation of the system. Data collection methods can be done through on board

Fig. 1. Global architecture for data aggregation in a telematics system.

Fig. 2. Architecture of the aggregation agent (AA).
sensors, acoustic, embedded, radar and video sensors. Moreover, floating car data (FDC) cited in Chen et al. (2003) and ABI Research can be used for more accurate recording of vehicle data. FDC refers to a set of protocols, services and data formats by which cars transmit information to a server (Chen et al., 2003). The use of FDC can be traced back to the mid-80’s using programs like Ali and Euroscout from Siemens and Socrates from Philips. In this program, vehicles are equipped with GPS and cellular modems to transmit speed and position data to remote data centers.

The aggregation function (AA) shown in Fig. 2 will act as coordinator to organize, extract and filter the sensor data. The design for AA is influenced and anchored on the FDC framework. The major task of AA is to prepare the data prior to data delivery. Three specific tasks undertaken by AA are extraction, filtering and communications. The extraction function of AA uses feature extraction and data fusion algorithm to collect data from sensor. While the filtering function removes noise, handle outlier values, and perform data transformation. The communication task of AA is to prepare the data in appropriate communication format so that in can be transmitted to the data center through CDMA channel.

Fig. 3 shows the process diagram of AA and DA. It shows the AA components, storing data in the management table, initiating the DA and generating the driving patterns. Finally, the generated patterns will be correlated to its corresponding insurance policies.

The components of data mining agent (DA) shown in Fig. 4 are summarization and rules generation functions. Summarization is extracting relevant data of selected or chosen cases from management table and tabulating all the events associated to it. Once the tabulation is done, the rule function will be initiated to calculate for the driving pattern based on the pre-defined patterns presented in Table 4. The parameters used for the rule functions are values from driving attributes.

As soon as the DA generated the driving patterns of a particular driver it will be correlated to pre-defined insurance schemes. In this scenario, the concerned driver can benefit lower premiums based on his driving characteristics. Our framework also considers important aspects of the system like data type and its availability whether it is public or private. Other aspects are method of storage, access controls, data security, transmission between users or system, and regulatory requirements that affect the data.

4. Illustration of data collected using vehicle telematics devices

Data A shows the sample data collected using a trip logging (Bisdikian et al., 2002) device, which shows the date and time that the car is started, stopping date, and total time of driving. In the parameter, distance traveled, throttle and speed are indicated. On the other hand, Data B shows the car mileage and as well as the average monthly mileage.
Data A:

<table>
<thead>
<tr>
<th>No.</th>
<th>Start Date</th>
<th>Start Time</th>
<th>Stop Date</th>
<th>Stop Time</th>
<th>Total Time</th>
<th>Parameter</th>
<th>Set Limit</th>
<th>Max Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2/9/04</td>
<td>12:22-40</td>
<td>2/9/04</td>
<td>2:16:16</td>
<td>0:56:30</td>
<td>Distance Traveled</td>
<td>N/A</td>
<td>55.6</td>
</tr>
<tr>
<td>2</td>
<td>2/9/04</td>
<td>1:23:40</td>
<td>2/9/04</td>
<td>1:23:41</td>
<td>0:00:01</td>
<td>Throttle</td>
<td>47</td>
<td>47.5</td>
</tr>
<tr>
<td>3</td>
<td>2/9/04</td>
<td>1:27:19</td>
<td>2/9/04</td>
<td>1:27:21</td>
<td>0:00:02</td>
<td>Throttle</td>
<td>47</td>
<td>61.6</td>
</tr>
<tr>
<td>4</td>
<td>2/9/04</td>
<td>1:27:48</td>
<td>2/9/04</td>
<td>1:29:02</td>
<td>0:01:16</td>
<td>Speed</td>
<td>44</td>
<td>72.7</td>
</tr>
</tbody>
</table>

Data B:

<table>
<thead>
<tr>
<th>Driver</th>
<th>Mileage</th>
<th>Avg. Monthly Mileage</th>
<th>Weekend mileage</th>
<th>%</th>
<th>Night Mileage</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bob</td>
<td>1805</td>
<td>233</td>
<td>607</td>
<td>.33</td>
<td>751</td>
<td>41</td>
</tr>
<tr>
<td>Betty</td>
<td>4826</td>
<td>21</td>
<td>2475</td>
<td>51</td>
<td>1604</td>
<td>30</td>
</tr>
</tbody>
</table>

Data C shows the passive GPS tracking (Bisdikian et al., 2002) data which includes driving date and time, distance traveled and location arrived. The table also shows the total driving time and the departure and arrival time. There are devices that can also monitor the real-time condition of the car. The information from this device includes dates, latitude and longitude, car location, and status of the car which is either running or stationary.

Data C:

Some papers (Chen et al., 2003; Duri et al., 2004; Finnegan and Sirota, 2004; Moskowitz et al., 2002) proposed other information to be gathered by the vehicle recording devices which can be used for analysis of vehicle crashes. This includes acceleration, speed, engine RPM and throttle, ignition cycle, air bag and safety belts.

Fig. 5 shows the programing flow chart for rendering the simulation. The terminator of the flow chart indicates the discovered patterns. In this simulation it refers to the driving scheme generated based on the characteristics of the driver. The patterns are computed using the intelligent data mining agent (IDMA) where Chi-square test for correlation is embedded into it.

The Netbeans 5.5 with JDK 6 (Netbeans) of Sun Microsystems is a Java based interface development environment used for coding and implementing the portion of the proposed system. Random generation of our test data and implementation of the Chi-square test were done using this IDE.

5. Use-case-scenario and evaluations

To illustrate the work around of our framework, we considered the data presented in Tables 1, 2 and 5 for the data processing and analysis. This is limited to and only includes information about position, braking pressure, speed, distance and total time of travel. Due to limitations of our study, only the mentioned attributes will be considered and tested. Other domains not covered here will remain the future task of this paper.

5.1. Use case

A use case diagram is shown in Fig. 6, where the actors are driver of the vehicle and its driving interaction and behavior are captured by the sensors actors such as position, braking pressure, speed, distance and time duration.

Once the data are aggregated at the vehicle side, the mobile phone actor will transmit the information to the data center and the receiver actor will receive it and then store in the database. The data accumulation is done by the aggregation agent. It has the function to extract data from different sensors and prepare this for data delivery to a data center.
The **driver** actor performs driving processes like acceleration, braking, stopping and road navigations. The distance is calculated based on the car mileage while the vehicle speed can be traced based on the speedometer. The pressure sensor will capture the amount of force when pedaling the brake. On the other hand, the position expressed in latitude and longitude is traced using a GPS device.

5.2. Obtaining base-line data prior to actual testing

The most appropriate method to get base-line data prior to actual implementations of the telematics system is through personal interview or survey questionnaires. Some questions to answers through this initial stage are: driving characteristics, willingness of the vehicle owner to use vehicle telematics, acceptability of the participants to share the information for research and commercial purposes. The data delivery which will comprise the data communication component of the

---

**Fig. 5.** Flow chart for data generation and IDMA.

**Fig. 6.** Use case diagram for data aggregation agent.
infrastructure can be coordinated with the telecommunications company that offers cellular service in the region. This can be included in the actual implementations of the system.

A randomized survey is recommended of at least 500 samples (0.001% of the population and the specific information to be learned are: (1) driving safety such as accidents, breakdown, and emergency services, (2) driving responsiveness, (3) positions such as where they usually go, (4) acceptability of the flexible insurance premium, and (5) determine how much is paid for the insurance and a possible discount that can be availed for variable scheme.

This research can be coordinated to the Telematics Service Providers and third party institutions that maybe interested for the telematics data like the insurance companies. A reward system can be given to the driver participants that are joining the study. For example, full disclosure of driving data for use by the third parties will give the concerned 20% discounts in one year insurance premium. Selected disclosure of data will correspond to lesser premium discounts.

5.3. Evaluations and testing

For illustration purposes, we show only one case-one event example to simplify the presentation of the concept. (Fig. 7) shows the driving characteristics of Driver 1 for a given event and the driving path as illustrated on the right side of the figure. Here, the data on the start and stop points can be determined. Other information in the figures refers to date and time, distance traveled and vehicle position.

The management table can be populated with several records based on all events accumulated on a time domain. Such table will also include the records of other drivers. The computed average for sensor data of 5 drivers is shown in Table 1.

Latitude and longitude values can be translated to actual location showing addresses relative to previous positions or starting point. The final position refers to the last position that the vehicle rested after the navigations. The values in the tables are average values for all the events in a certain time range. Heterogeneous devices would provide raw/unformatted data prior to uploading to a data center, hence, AA will be utilized to pre-process it.

Table 3 shows the transformed equivalent of the values in Table 1. The attribute classes are provided by the domain experts and the corresponding ranges are presented in Table 2. We can compute for the ranges of the given attribute using our proposed function in

![Fig. 7. Illustration of the destination points and total distance traveled. The final position is indicated by the latitude/longitude coordinates.](image)

**Table 1**
Illustration of cumulative data* derived from car sensors

<table>
<thead>
<tr>
<th>Driver</th>
<th>Position (final position)</th>
<th>Braking pressure (lbs)</th>
<th>Speed (km/h)</th>
<th>Distance (km)</th>
<th>Total time (h)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>37.57/126.97</td>
<td>Hard</td>
<td>87</td>
<td>451.11</td>
<td>91.21</td>
</tr>
<tr>
<td>2</td>
<td>37.48/126.89</td>
<td>Harder</td>
<td>85</td>
<td>358.23</td>
<td>85.70</td>
</tr>
<tr>
<td>3</td>
<td>37.56/126.99</td>
<td>Regular</td>
<td>58</td>
<td>147.68</td>
<td>70.32</td>
</tr>
<tr>
<td>4</td>
<td>37.54/126.95</td>
<td>Regular</td>
<td>44</td>
<td>512.12</td>
<td>95.65</td>
</tr>
<tr>
<td>5</td>
<td>37.53/126.82</td>
<td>Regular</td>
<td>76</td>
<td>342.50</td>
<td>72.33</td>
</tr>
</tbody>
</table>

* Cumulative data for 30 days.
\[
\delta = \frac{\max(I) - \min(I)}{\epsilon},
\]

where \( \delta \) is the discrete cut points, \( \max(I) \) is the highest score in the given attribute, \( \min(I) \) is the lowest score in the given attribute, \( \epsilon \) is the expert opinion on particular attribute. This will follow that the classes of the given attribute can be determined using the function in Eq. (2):

\[
f(a) = \begin{cases} 
1, & \text{if } a \leq C_1 \\
2, & \text{if } C_1 < a \leq C_2 \\
3, & \text{if } C_2 < a \leq C_3 \\
\vdots, & \\
r - 1, & \text{if } C_{r-2} < a \leq C_{r-1} \\
r, & \text{if } a > C_{r-1}
\end{cases}
\]

where \( a \) refers to the attribute that belongs to the range defined by the cut point, \( C_1 \) is calculated as \( \min(I) + \epsilon \); \( C_2 \) is calculated as \( C_1 + \epsilon \); \( C_3 \) is calculated as \( C_2 + \epsilon \); \( C_r \) is calculated as \( C_{r-1} + \epsilon \) and until \( \max(I) + \epsilon \). For example, if the \( \epsilon = 3 \), then the \( \delta \) for attribute speed is 14. In this case, the three classes formed are: regular (44–58), far (59–73) and farther (74–87).

Several factors that influence the choice of cut-points for the classes given in Table 2 are trends in driving, expert opinion, averaging, or using a function to compute for the classes. In this illustration we considered the domain expert opinion for the ranges.

Flexible and discounted insurance premiums can range from Policy 1 which is associated to regular premium to Policy 3 which is highly discounted rate. A “very safe driving” pattern can receive full benefits by getting highest premium discounts. On the other hand, risky driving pattern have no discount at all.

### 5.3.1. Insurance policies

- Policy 1: Regular premium for risky driving.
- Policy 2: Insurance premium discounted at 10% normal.
- Policy 3: Insurance premium discounted at 20% safe driving.

Number one concern of insurance companies is safety. This implies that the safer the driver is the lesser is the liability of the company for insurance claims. In addition, this will insure road safety and reduce road accidents. According to Bauer

### Table 2

<table>
<thead>
<tr>
<th>Attributes</th>
<th>Ranges</th>
<th>Descriptions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Position (From last starting point)</td>
<td>1</td>
<td>Near</td>
</tr>
<tr>
<td>2</td>
<td>Far</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Farther</td>
<td></td>
</tr>
<tr>
<td>Braking pressure</td>
<td>1</td>
<td>Moderate</td>
</tr>
<tr>
<td>2</td>
<td>Hard</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Harder</td>
<td></td>
</tr>
<tr>
<td>Speed</td>
<td>1</td>
<td>Regular</td>
</tr>
<tr>
<td>2</td>
<td>Fast</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Very Fast</td>
<td></td>
</tr>
<tr>
<td>Distance traveled</td>
<td>1</td>
<td>Short</td>
</tr>
<tr>
<td>2</td>
<td>Long</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Longer</td>
<td></td>
</tr>
<tr>
<td>Total time</td>
<td>1</td>
<td>Short</td>
</tr>
<tr>
<td>2</td>
<td>Regular</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Long</td>
<td></td>
</tr>
</tbody>
</table>

### Table 3

<table>
<thead>
<tr>
<th>Driver</th>
<th>Position</th>
<th>Braking pressure</th>
<th>Speed</th>
<th>Distance</th>
<th>Total time</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Near</td>
<td>Hard</td>
<td>Very fast</td>
<td>Longer</td>
<td>Long</td>
</tr>
<tr>
<td>2</td>
<td>Far</td>
<td>Harder</td>
<td>Very fast</td>
<td>Longer</td>
<td>Long</td>
</tr>
<tr>
<td>3</td>
<td>Far</td>
<td>Regular</td>
<td>Regular</td>
<td>Short</td>
<td>Regular</td>
</tr>
<tr>
<td>4</td>
<td>Near</td>
<td>Regular</td>
<td>Regular</td>
<td>Longer</td>
<td>Long</td>
</tr>
<tr>
<td>5</td>
<td>Near</td>
<td>Regular</td>
<td>Fast</td>
<td>Longer</td>
<td>Regular</td>
</tr>
</tbody>
</table>

\[
f(a) = \begin{cases} 
1, & \text{if } a \leq C_1 \\
2, & \text{if } C_1 < a \leq C_2 \\
3, & \text{if } C_2 < a \leq C_3 \\
\vdots, & \\
r - 1, & \text{if } C_{r-2} < a \leq C_{r-1} \\
r, & \text{if } a > C_{r-1}
\end{cases}
\]
et al. (2005), Berlin taxi accident fell by 66% after installing telematics tracking while European Union telematics found 28% reduction in accident and 40% in cost. Table 4 shows the driving patterns provided by the domain expert per driving category. Domain expert can base the patterns on the norms of data aggregated from the motorists.

As soon as the driving pattern is determined, a correlation value will be computed. The computation will be based from the records in Tables 4 and 5.

Table 5 shows the cumulative data for Driver 1. The events refer to driving events which constitute the start and stop mode of driving. The table entries correspond to the descriptive equivalent of the continuous data generated per event. The final position is computed relative to starting position.

The Chi-square test is used to compute the correlation and determine if the generated pattern do not differ significantly to the expected driving patterns. The equation is presented in Eq. (3), where \( O \) is the observed frequency, \( E \) is the expected frequency, and \( k \) is the number of cases

\[
\chi^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i}.
\]  

In our example, we are able to determine that the driving pattern of Driver 1 is “risky driving”. Hence, the insurance policy assigned to this pattern is “Policy 1”. This will imply that the driving pattern of driver 1 will be likely prone to danger or accident. In this scenario, the insurance company charges regular premium to Driver 1. Granting that the driving pattern is “normal”, and then insurance company can give a discount of 10% on insurance premium. On the other hand, 20% discount is given to driver with “safe driving” pattern.

Relevant patterns refers to the driving pattern discovered based on the driving events. It is highly relevant if the expected value as compared to the computed value in Chi-square test indicates a high significance value.

5.5. Applications

This framework has a promising applications ranging from academic, business to industrial uses. Researchers at the academy can formulate mechanism to improve the driving pattern discovery which can closely reveal driving characteristics. It is probable to collaborate with Insurance companies in Korea for the use, disposition of the information obtained from the study.

There are greater possibilities of commencing an initiative to encourage flexible insurance premium based on how safe the driver is. Also, insurance companies can support the initiative of developing a comprehensive algorithm for flexible policy scheme in Korea based on the data to be collected from actual deployment of the system. Also, we improve this scheme by continual extraction of data from samples for a certain period until the norm will be achieved. And in exchange, the study participants can be granted benefits like knowing their driving patterns or providing them reduced insurance premiums.

It is estimated that parties involve can benefit the scheme at several aspects. The clients can have choice of flexible insurance plan or pay as you drive options based on the driver’s characteristics. Companies can win the trust of the clients by providing them up-to-date data, diagnostics and road-side assistance and reduced insurance rates. In addition, they can accumulate substantial data for future traffic prediction, road safety, customer trends and for decision support.

6. Conclusion and recommendations

In this paper, we are able to formulate framework for data aggregation and Insurance service in vehicle telematics using multi-agents and perceived it to have promising applications ranging from academic, business to industrial uses. We are able
to present a use-case scenario and illustrative examples to demonstrate our model. With flexible insurance framework, customers can manage their own insurance premiums and as well as the cost of motoring. Safer driving will mean discounted insurance and as well as road safety.

However, there are some constraints that have to be addressed by this study in the future like the cost of the study including methodology, sampling and analysis. Another one is the willingness of the participant to cooperate in the study, which eventually affects their privacy and security. Others are investment requirements such as cost of telematics devices to be installed in a vehicle and availability of GPS.

Acknowledgement

This research was financially supported by the Ministry of Education, Science Technology (MEST) and Kroea Industrial Technology Foundation (KOTEF) through the Human Resource Training Project for Regional Innovation.

References