Information Sciences 180 (2010) 62–70

Contents lists available at ScienceDirect

Information Sciences

journal homepage: www.elsevier.com/locate/ins

Collaborative real-time traffic information generation and sharing framework for the intelligent transportation system

Wei-Hsun Lee [a,b], Shian-Shyong Tseng [a,c,*], Wern-Yarng Shieh [d]

[a] Department of Computer Science, National Chiao Tung University, 1001 Ta Hsueh Road, Hsinchu 300, Taiwan, ROC
[b] Telecommunication Laboratories of Chung-hwa Telecom, 12, Lane 551, Min-Tsu Road, Sec. 5 Yang-Mei, Taoyuan 326, Taiwan, ROC
[c] Department of Information Science and Applications, Asia University, 500 Liufeng Road, Wufeng, Taichung 413, Taiwan, ROC
[d] Department of Computer and Communication Engineering, St. John's University, 499, Sec. 4, Tam King Road, Tamsui, Taipei 251, Taiwan, ROC

ARTICLE INFO

Article history:
Received 3 October 2008
Received in revised form 2 September 2009
Accepted 4 September 2009

Keywords:
Collective intelligence
Traffic status prediction
Smart traffic agent
Intelligent transportation system (ITS)
Knowledge-based system

ABSTRACT

Real-time traffic information collection and data fusion is one of the most important tasks in the advanced traffic management system (ATMS), and sharing traffic information to users is an essential part of the advance traveler information system (ATIS) among the intelligent transportation systems (ITS). Traditionally, sensor-based schemes or probing-vehicle based schemes have been used for collecting traffic information, but the coverage, cost, and real-time issues have remained unsolved. In this paper, a wiki-like collaborative real-time traffic information collection, fusion and sharing framework is proposed, which includes user-centric traffic event reacting mechanism, and automatic agent-centric traffic information aggregating scheme. Smart traffic agents (STA) developed for various front-end devices have the location-aware two-way real-time traffic exchange capability, and built-in event-reporting mechanism to allow users to report the real-time traffic events around their locations. In addition to collecting traffic information, the framework also integrates heterogeneous external real-time traffic information data sources and internal historical traffic information database to predict real-time traffic status by knowledge base system technique.

1. Introduction

Real-time traffic information plays an important role in several applications of intelligent transportation system (ITS), such as advanced traffic management systems (ATMS), advanced traveler information system (ATIS), commercial vehicle operation (CVO) and emergency management system (EMS), etc. It is one of the most useful features especially for users on the road, because it can assist and support travelers and drivers to plan the trip before travelling and decide on the route to take in order to reduce travel time and improve travel safety. Moreover, the traffic information collected can also support traffic management administrators in making decisions, taking appropriate actions to alleviate congestions, and improving the global performance of the traffic network. However, with one-way broadcasting mechanism, most traffic information services lack collective intelligence of gathering real-time traffic information from the service terminal devices and user-reporting mechanism, and continuous real-time traffic information for the whole traffic network is also hardly available due to cost, sensor coverage, and several other issues such as data collection and transmission techniques.

* Corresponding author. Address: Department of Computer Science, National Chiao Tung University, 1001 Ta Hsueh Road, Hsinchu 300, Taiwan, ROC. Tel.: +886 3 5731966.
E-mail addresses: leews@cht.com.tw (W.-H. Lee), sstseng@asia.edu.tw (S.-S. Tseng), shiehwy@mail.sju.edu.tw (W.-Y. Shieh).

0020-0255/$ - see front matter © 2009 Published by Elsevier Inc.
doi:10.1016/j.ins.2009.09.004

On the other hand, mobile personal devices such as mobile phones, PDAs are equipped with location capability by connecting to GPS module by wired or wireless scheme. More and more personal navigation device (PNDs) and ultra-mobile PC (UMPC), which are equipped with GPS are now capable of communication in order to retrieve real-time traffic information. All these devices have the features of mobility, communication, location-aware and user-interface capabilities. With the front-end devices having these features, many modern mobile applications such as location-based service (LBS), telematics, and ITS-related applications can be provided. For example, LBS, a kind of service providing appropriate location-aware information for users, accomplish the business processes by exchanging information between the mobile front-end devices and the back-end systems. On the basis of such high penetration of mobile communication devices with location capability, the goal of this paper is to propose a collaborative traffic information generation and sharing framework for real-time traffic information collection, fusion and distribution.

The applications of traffic information are generally divided into three stages, which are data collection and cleansing, data fusion and integration, and data distribution. In the data collection and cleansing stage, raw data are collected from different data sources, filtered and analyzed in the back-end system. The collected data are then integrated with the geographical information system (GIS) and external traffic information data sources in the data fusion and integration phase. In the data distribution phase, several interfaces including web, radio, and mobile subscription are provided for users, where the mobile subscription scheme suitable for users with mobile communication devices can retrieve real-time traffic information provided by the traffic information center (TIC) according to the device location.

Traditionally, there are three schemes for collection of real-time traffic information [5]: site-based, sensor-based and probing vehicle-based data collection schemes. Site-based measurement collects vehicle license plate information and arrival times at various checkpoints through automatic vehicle identification (AVI) technologies, matches the license plates between consecutive checkpoints, and computes travel times from the difference between arrival times. Vehicle-based methods analyze the raw data collected from fleet of probe vehicles by matching the vehicle tracks with geographical information system (GIS). Sensor-based scheme collects raw data from the stationary sensors like loop detectors, transponders or radio beacons installed at arterial roads. However, these traffic information collection methods have drawbacks and limitations. For example, site-based and sensor-based methods have the spatial coverage problem due to the fixed and limited number of sensors or AVI devices. Vehicle-based scheme [1,8,12] has cost, spatial and temporal coverage problems due to the very high cost for maintaining a dedicated fleet of urban network traffic probing vehicles. Moreover, the cost of real-time transmission for the whole traffic network in each data collection scheme is also very high [4].

As information exchanging and sharing through a public platform, e.g., the Wikipedia website, is the major trend of Web 2.0 technology, several wiki-like collective intelligence models or applications have been proposed [3,7,9,10] recently. Our idea is to create a wiki-like environment for exchanging and sharing real-time traffic information. Accordingly, a collaborative real-time traffic information collection, data fusion and distribution framework including a front-end smart traffic agent (STA), real-time traffic information exchange protocol (TIEP) and traffic data fusion and distribution back-end center (TIC) is proposed in this paper. In addition to two-way real-time location-aware traffic information exchange with TIC through TIEP, STA is also developed as a smart interface between the terminal and the driver for displaying real-time and predicted traffic information, and providing real-time traffic event-reporting interface to share local traffic information with other users. Several versions of STAs are developed for various location-aware front-end devices such as GPS smart phone, PDA, PND, or UMPC, which have two-way real-time traffic information exchange capability, and built-in event-reporting mechanism to allow users to report the real-time traffic events around their locations. TIEP is an XML-based traffic information exchange interface between STA and the back-end TIC. The back-end TIC includes data collection and cleansing, data fusion and integration, and a knowledge-based system for traffic status prediction. Several external traffic information data sources including national traffic information center [17], national freeway traffic information [19], and police station real-time traffic information [18] are integrated into the TIC by adopting the optimal weight traffic information fusion scheme [13,14] to enhance the traffic information prediction capability. A knowledge-based system with expert heuristic rules is also implemented in the TIC to predict short-term traffic status by combining real-time collected traffic information, historical traffic database, and multiple external traffic information data sources.

The rest of this paper is organized as follows. Section 2 introduces the design details of STA and TIEP according to the request-response model. In Section 3, the proposed collaborative traffic information generation and distribution framework is discussed, which includes traffic information collection and cleansing, traffic information generation for urban network and freeway, and external traffic information data sources connection. The integration and information fusion of heterogeneous multi-sources external traffic information data sources is discussed in Section 4, where optimal weighting scheme developed from the Shannon entropy theory is applied. In Section 5, a knowledge-based system for predicting the traffic status is discussed. Finally, concluding remarks are presented in Section 6.

2. Smart traffic agent (STA) and traffic information exchange protocol (TIEP)

STA, a software agent for the front-end devices, provides an interface between the traveler and the TIC. It is installed on the front-end device with a two-way real-time data exchange channel involving TIEP for (1) automatically echoing the local traffic information detected by STA to the TIC, (2) automatically downloading the local real-time and predicted traffic information, and (3) manually reporting the traffic events or traffic status to the TIC by the traveler. The TIEP is defined in XML format based upon the request-response model, where the request message is integrated with the echoing traffic information to be sent to the TIC, which includes GPS coordinates, moving direction, speed, traffic status information and scope of request. The frequency and scope of exchange concerning the local traffic information is determined by rules embedded in the STA, which is calculated using the traveling speed and the vehicle location. For example, the higher the moving speed of the STA, the larger the scope of request and the shorter the frequency in order to retrieve more up-to-date and more suitable traffic information.

The TIC response message consists of a list of traffic messages which are divided into two parts: text message information and average speed information. The response messages received by the STA can be employed to update the map-based user interface, which helps the travelers easily catch the local real-time and predicted traffic information. The text messages inform the travelers of the local traffic messages consisting of ⟨*no., message type, data source, message, severity, coordinate*⟩ such as traffic events, CMS. For example, text message ⟨*3, 'events', 'TIC', 'car accident', 5, x1, y1*⟩ indicates that a car accident reported by STA users happened at the location $(x1, y1)$ which has the severity level 5, and text message ⟨*5, 'CMS', 'EIOT', 'Link A1(North) strongly congested, please detour!', 4, x2, y2*⟩ informs that a changeable message sign (CMS) system shows "*Link A1(north) is strongly congested, please detour!*" message at the location $(x2, y2)$ where the message comes from the data source "EIOT". On the other hand, the average speed information contains the current average speed and predicted traffic condition consisting of ⟨*no., type, link/intersection id, avg. speed, current status, predicted status, coordinates*⟩, which informs the current average traveling speed and predicted traffic status of the links around the STA. Real-time traffic status and predicted traffic status are combined in the TIC response message by considering the predicted STA location. As illustrated in Fig. 1, after an STA request is sent from location A, TIC responds the real-time (dotted small rectangle) and predicted (dotted large rectangle) traffic messages centralized by predicted location B by considering the moving direction, moving speed, and response packet size (response time). In the dotted small rectangle centralized by predicted STA location at point B, the TIC responding message contains only real-time traffic information (average speed, traffic events) since the predicted information makes less sense to the current status. In the large dotted rectangle, both real-time and predictive traffic information is contained in the response message to inform drivers of the current status and predicted traffic status around the STA.

There is a wiki-like user-centric traffic information sharing mechanism for users to manually report local traffic events or traffic status to the TIC by the user interface provided by the STA, as shown in Fig. 2. The reported traffic event or status will not be distributed instantly until it is confirmed by another traveler (through the same mechanism) or confirmed by the external data sources. The intention of this double-check mechanism is to prevent the system from receiving fake messages.

3. Collaborative traffic information generation and distribution framework

As shown in Fig. 3, the architecture of the collaborative traffic information generation and distribution framework includes three parts: location-aware mobile device with STA, TIC for data collection and fusion, and external traffic information data sources, where various versions of STAs are developed for different front-end mobile devices depending on different hardware and operating system. The generated real-time traffic information can be distributed by STA (request and response model), radio broadcast network (RDS), and web site. The TIC, is connected to external real-time traffic information data sources, including E-IOT national traffic events database [17], TANFB freeway traffic information [19], PRS public traffic service [18], and parking slot databases.

3.1. Traffic information collection and cleansing

Based on the TIEP, the traffic information is exchanged between local STAs and the TIC. Each record of data exchange collected by the request-response model between STA and TIC is treated as a traffic information sample, which will be cleansed and analyzed to get the local traffic status at the STA location. The local traffic information uploaded by the STA includes moving speed, moving direction, GPS coordinates, STA state, and scope of request. The TIC then answers the request by

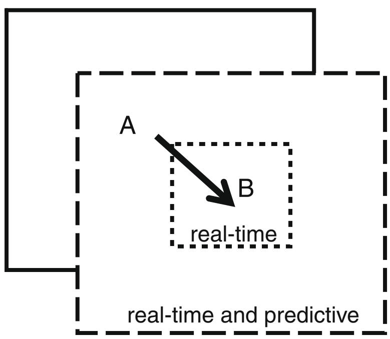

Fig. 1. TIC real-time and predicted traffic information response for STA request.

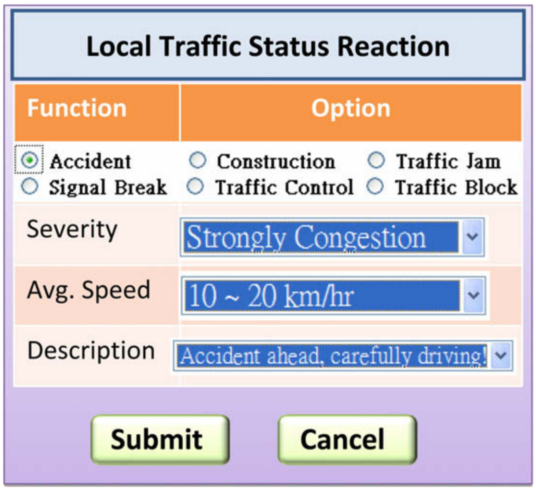

Fig. 2. User-centric traffic status reporting user interface in STA.

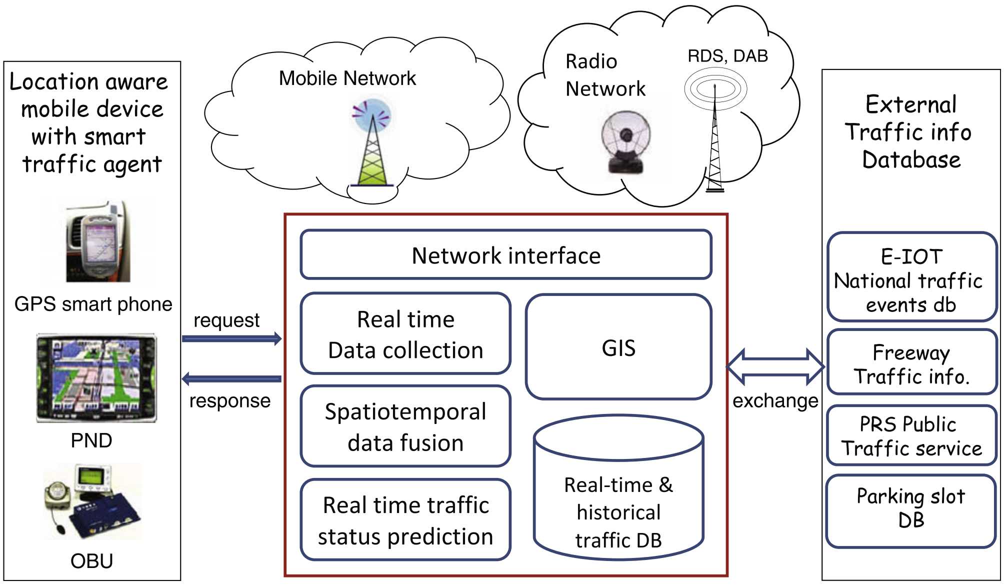

Fig. 3. Collaborative traffic information generation and distribution framework.

querying the traffic information database for the traffic status around the GPS coordinates of the STA. The cleansing process filters out the useless raw data, e.g., incomplete data, incorrect data, outlier, or noise data. The location information of the STA (link, address) is transformed from the GPS data in the collected sample with the assistance of traffic network database and GIS. A traffic information sample $S_k(0_i^j, V, D)$ transformed from the STA uplink request U_k of a vehicle, as illustrated in Eq. (1), consists of object id (0_i^j), speed (V), direction (D) of the vehicle when it communicates with the TIC at time t at GPS coordinates (x, y), where 0_i^j is a spatiotemporal network object spatially indexed by network object id i (transformed from the location to address interpolation) and temporally indexed by time zone j (transformed from timestamp t).

$$U_k(X, Y, t, V, D) \stackrel{Gis}{\rightarrow} S_k(O_i^j, V, D) \tag{1}$$

3.2. Freeway traffic information generation

The focus of traffic information for the freeway area is different from that for the urban area. Travelers are more concerned about the traffic event messages and ramp-to-ramp traveling time than the route path choice on the freeway. On the contrary, in the urban area, travelers are more concerned about the route path choices for avoiding the congestion. Since the topology in the freeway is different from that in the urban network, the traffic information generation schemes use different sampling strategies to generate the traffic information depending on the location of the traffic information in the freeway area or urban area. For example, for a vehicle equipped with STA traveling through a freeway, a series of samples uploaded from the same STA are collected by the TIC, and each sample can be transformed into the format of (M_x, T_x), where M_x indicates the mileage and T_x indicates the uploaded timestamp. By selecting the entry sample (M_s, T_s) and exit sample (M_e, T_e) of an STA, the average traveling speed $(\overline{V_{s,m}})$ between these two ramps can be calculated by $(M_e - M_s)/(T_e - T_s)$, which is regarded as a case. For any two ramps k and j, the overall average traveling speed can be calculated by the arithmetic mean value $(\overline{V_{k,j}})$ for all the cases of (k,j) as shown in Eq. (2).

$$\overline{V}_{k,j} = \frac{n \cdot (|M_k - M_j|)}{\sum_{i=1}^{n}(T_{k,i} - T_{j,i})} \tag{2}$$

3.3. Urban network traffic information generation

Urban network consists of a set of network objects, each of which is either a link or an intersection where traffic congestions occur on some network objects in which the traffic demand cannot be fully serviced. The traffic status of the network can be represented by a set of network objects which constitute the network. Average speed of a link in a temporal period can be calculated by the arithmetic mean value shown in Eq. (2) of all the samples spatially falling into the link and temporally falling into the temporal period. However, in urban area, it seems unreasonable to represent the link traffic status by the average speed for all the links because the service levels of road grades of different categories are different. We normalize the link traffic status by the categories of the links, as shown in Table 1, the service levels of the three road categories in urban area of Taiwan are defined from levels A–F [2]. The link traffic information can be transformed into the service level by mapping the average speed to the road category and service level defined in Table 1. For example, average speed of 30 km/h indicates that the traffic status is good (level B) in street (category III), but slight congestion (level D) in the expressway or arterial road (category I).

On the other hand, the average intersection delay can represent traffic information for the intersection, and the delay between two consecutive links, which is mostly caused by signal delay and queuing delay, can be classified by *TD*, *LTD* and *RTD* patterns according to the three possible directions from one link to another [4]. Eq. (3) shows the general format of intersection delays (*TD*/*LTD*/*RTD*), where P is the pattern type, SO_{id} and SI_{id} are the two consecutive links where vehicle leaves out the link SO_{id} and comes into the link SI_{id}, T_{id} is the temporal id, D_{avg} is the average delay time of this intersection, and *Sup*, *Con* are the support and confidence of the pattern, respectively.

$$[TD/LTD/RTD] : (P, SO_{id}, SI_{id}, T_{id}, D_{avg}, Sup, Con) \tag{3}$$

For example, (‘*RTD*’, ‘*L*1’, ‘*L*2’, ‘*W,P*’, 40, 0.2%, 75%) represents that in the peak hours of workday, it takes 40 seconds to make a right-turn from link ‘*L*1’ to link ‘*L*2’, with the support being 0.2% and confidence being 75%. Intersection delay patterns can be discovered by sequential pattern mining or spatial and temporal sequence mining on all samples of intersection delay in a journey containing two consecutive samples with different links in the historical traffic information database. Fig. 4 shows an example of *RTD* pattern: a probing vehicle driving north and then turning right to east, it reports TIS at location A of Link L_a and reports consecutively TIS at location B of Link L_b. The symbols of the TIS format (*T,L,X,Y,D,V*) in Fig. 4 stand for timestamp (T), link id (L), coordinates (X, Y), direction (D) and speed (V). The distance d_a, d_b in Fig. 2 stands for the distance from A or B to the intersection of links L_a or L_b, respectively. Assuming that in the short time interval between T_b and T_a, the vehicle is driving at the speed of V_a at link L_a and V_b at the link L_b. Then the right-turn delay (*RTD*) from L_a to L_b can be estimated by subtracting travel time of d_a and d_b from elapsed time between two TISs $(T_b - T_a)$.

The status of each intersection delay in an intersection is also classified into six service levels (A–F) by normalizing each intersection delay, i.e., by equally dividing the intersection delay samples into six levels according to the historical traffic information database. With the traffic status of all the links and intersections, the whole network status can be easily represented on the map by coloring the objects in the user interface.

Table 1
Service level classification for the three road categories in Taiwan [2].

Road grade	I	II	III
Free flow speed	55 (kph)	45 (kph)	40 (kph)
Service level	Avg. speed (kph)	Avg. speed (kph)	Avg. speed (kph)
A (90%)	>51	>43	>33
B (70%)	51–39	43–32	33–25
C (50%)	39–34	32–27	25–20
D (40%)	34–29	27–23	20–16
E (33%)	29–21	23–17	16–10
F (25%)	<21	<17	<10

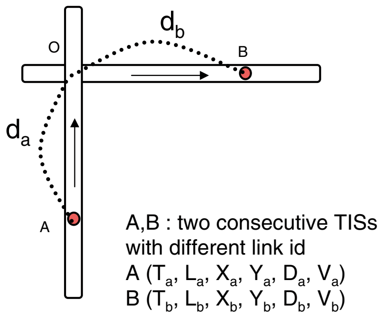

Fig. 4. Intersection delay example: RTD.

3.4. External traffic information data sources connection

In addition to the traffic information collected from LBS, external real-time traffic information data sources including national traffic information center (E-IOT) [17], Taiwan Area National Freeway Bureau (TANFB) freeway traffic information system [19], public traffic service of police radio station (PRS) [18], and several parking slot databases are integrated in the framework to retrieve more driving assistance information such as real-time CMS (changeable message sign) information, and parking slot information. Accordingly, the traffic information in the framework is sufficient for not only reporting real-time traffic information but also predicting the traffic status and can be further fed back to these external data sources for sharing the information with travelers not equipped with STA.

3.5. Robustness of the system

In the proposed framework, huge number of STAs may request to a centralized TIC and overload it, so that the robustness of the TIC must be considered. For the robustness issue, several data servers can be set up in the TIC in order to handle large amount of STAs, and the load distribution can be done by distribution rule embedded in the STA or the corresponding data server assigning rule defined in the TIEP. The principle of these distribution methods is to averagely share the load to the data servers in the TIC.

4. Multi-sources heterogeneous traffic information fusion

The traffic information collected in the framework is classified by traffic events (message), and traffic status (link average speed, intersection status). The collected STA-centric echoing samples can be transformed into freeway traffic status, urban network link and intersection traffic status as discussed in Section 3. The generated traffic status and user-reported traffic events are stored in the real-time traffic database, and the out-of-date data will be periodically moved into the historical traffic database.

In order to integrate the required external real-time traffic information data sources with the generated traffic information in the framework, an optimal weighting combination scheme developed from Shannon entropy theory [11] for traffic data fusion approach [13–15] is adopted. The internal traffic information in the framework and each of the external traffic data source is regarded as a data source (S_i). The weighted combination fusion equation for a network object is shown in Eq. (4), where O_j can represent any attribute of a network object such as link average speed, intersection delays (*TD*/*LTD*/*RTD*), and W_i is the weight for the data source S_i.

$$\begin{aligned} \overline{O_j} &= \sum_i W_i \cdot S_i(O_j), \\ \sum_i W_i &= 1 \end{aligned} \tag{4}$$

The Shannon entropy is employed to measure the uncertainty and randomness of the collected data and external data sources, and the optimal weighting approach [14] tries to minimize the overall uncertainty by assigning the lower weight to higher entropy data sources. The entropy with conditional probability of data source S_i, denoted as $h_i(\gamma)$ which is adopted from [13], is shown in Eq. (5), where $P(\theta_j|\gamma)$ is the probability of occurrence of the state θ_j given the observed state γ. The probability can be summarized from the historical traffic database; for example, the probability of service level '*C*' on road L_j can be calculated by the speed samples which are located in L_j and falls in the range of level '*C*' in the road category of L_j divided by the total samples located in L_j. The weight assignment equation in optimal weighting approach [14] is shown in Eq. (6), where S is the set of all data sources. It shows that the weight $W_i(\gamma)$ has a negative relationship with $h_i^2(\gamma)$, so that data source i will be assigned a lower weight W_i if the entropy h_i (uncertainty) is high.

$$h_i(\gamma) = -\sum P(\theta_j|\gamma)\log P(\theta_j|\gamma) \tag{5}$$

$$W_i(\gamma) = \frac{1}{h_i^2(\gamma)\sum_{i\in S}\frac{1}{h_i^2(\gamma)}} \tag{6}$$

However, there are several traffic characteristics in temporal dimension such as peak hour, weekday, or midnight. Without considering temporal traffic characteristics, the probability calculation in the entropy $h_i(\gamma)$ equation defined in Eq. (5) may lose the precision for the "optimal weight" calculation. For example, traffic status characteristics on workdays are different from those on holidays, and traffic characteristics during AM peak hours are not the same as those during PM peak hours. In order to take the traffic status characteristics in temporal dimension into consideration, we divide the collected samples in the historical traffic information database into five temporal periods: {AM peak hour, PM peak hour, workday, holiday, midnight}. The probability of occurrence of the state θ_j given the observed state $\gamma, P(\theta_j|\gamma)$ is modified into $P(\theta_j^k|\gamma)$ to extend the temporal dimension consideration, where k is the index for temporal periods. The entropy defined in Eq. (5) is then modified to consider temporal dimension, as shown in Eq. (7), where $h_i^k(\gamma)$ denotes the entropy of network object i at temporal index k.

$$h_i^k(\gamma) = -\sum P(\theta_j^k|\gamma)\log P(\theta_j^k|\gamma) \tag{7}$$

It is easy to map the events from the original entropy definition (Eq. (5)) to the new definition (Eq. (7)) by mapping the event timestamp to one of the five pre-defined temporal periods stated above. The advantage of adding this temporal dimension consideration is that the effects of the traffic events happened in different temporal periods can be considered differently. For example, a serious traffic accident event happened in the midnight may slightly affect the traffic flow, but the same event may seriously affect the traffic flow if it happened in the AM/PM peak hours.

5. Knowledge-based traffic status predicting system

The traffic event messages collected from the STA and external traffic data sources are used not only for informing current traffic status but also for predicting the traffic status. For each STA traffic information request, the prediction result is sent to the STA accompanied with real-time traffic information, as illustrated in Fig. 1. The traffic status prediction is implemented by the knowledge-based system, in which expert heuristic rules are elicited to enhance the precision of traffic status prediction. There are many rules for traffic events to be considered for traffic status prediction on an STA request, such as event type, severity, event elapsed time, and spatial consideration related to the STA. For example, a traffic accident may have severe impact on the surrounding STAs, but the severity will be reduced as the distance increases.

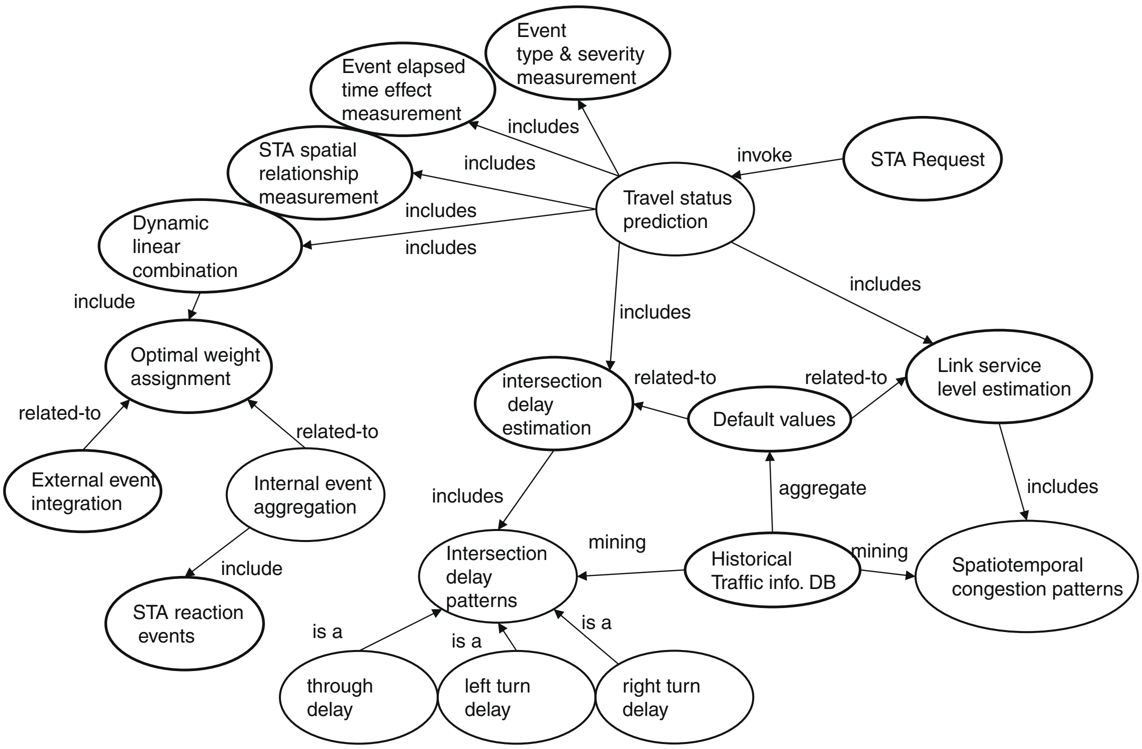

Fig. 5. Design ontology for traffic status prediction knowledge base system.

The design ontology of traffic status prediction knowledge base system is illustrated in Fig. 5, which includes several design concepts and relationships between these concepts, such as "*is-a*", "*includes*", "*modify*", and "*related-to*". These relationships connecting concepts in the ontology represent some interactions among them. For example, the "*traffic status prediction*" concept includes several concepts such as "*dynamic linear combination*", "*intersection delay estimation*", "*link service level estimation*", and three traffic event factor measurement concepts. When the system receives an STA traffic status request, the "*traffic status prediction*" concept is then "*invoked*" to process the whole traffic status prediction process, which again invokes those concepts included in the "*traffic status prediction*" concept. The optimal weight assignment module integrates the internal STA reported events and the external traffic event information sources by optimal weighting linear combination scheme. On the other hand, the real-time link service level and intersection delay estimation modules retrieve the current link service level and intersection delay information from the real-time traffic information database. "*Traffic status prediction*" returns the current traffic status around the STA (real-time area in Fig. 1) and predicts the traffic status (real-time and prediction area in Fig. 1) by fine-tuning the current traffic status and considering the localized STA and traffic event impacts consideration modules including "*event type and severity measurement*", "*event elapsed time effect measurement*", and "*event location to STA spatial relationship measurement*".

The "*event type and severity measurement*" module evaluates the effect of the event to the STA by its type and severity, and tunes the real-time evaluation result; for example, the service level prediction of the link and related intersections surrounding the traffic accident event may be downgraded one or two levels compared with the current link service level and intersection delay estimation. The "*event elapsed time effect measurement*" and "*event location to STA spatial relationship measurement*" modules tunes the link service level and intersection delay by considering the temporal distance (event elapsed time from the STA request time) and spatial relationship (distance from the event to the STA), respectively. The traffic status prediction knowledge base prototype system has been implemented by the expert system shell, DRAMA [16], a New Object-oriented Rule Model (NORM) [6] knowledge-modeled rule base system platform implemented using pure Java language, includes DRAMA Server, Console, Knowledge Extractor, and Rule Editor. Heuristic rules donated by the traffic domain experts are categorized into several knowledge classes (KC), and stored at knowledge base of DRAMA server.

6. Conclusion

Real-time traffic information plays an important role in ITS, especially for travelers, fleet managers, and ATMS administrators. However, the availability of real-time traffic information suffers from the cost, coverage, and real-time issues in the traditional data collection techniques, so compromise solution must be made among these considerations. A wiki-like collaborative real-time traffic information generation and sharing framework involving a real-time traffic status prediction knowledge base system is proposed in this paper. By this collective traffic information generation scheme, more real-time traffic data will be collected cost-effectively and accurately, and the spatiotemporal coverage is better than the traditional traffic information collection scheme. Front-end STA, TIEP protocol, and back-end TIC constitute the traffic information sharing framework, where STA and TIEP are designed for the location-aware real-time traffic information exchange between the STA and TIC, which enables STA to automatically echo the local traffic information detected by STA to the TIC, automatically download the local real-time and predicted traffic information, and report the traffic events or traffic status to the TIC by the traveler.

In additional to real-time traffic information collection and generation from STA, the TIC back-end system adopts the Shannon entropy-based optimal weighting approach [14] to integrate external real-time traffic information data sources to minimize data uncertainty. A rule-based knowledge system is implemented to predict the short-term traffic status by fine-tuning the real-time traffic status by considering the event impact factor. Heuristic rules donated by domain experts improve the prediction accuracy. The methodology proposed in this paper turns the traditional government centralized data collection and distribution scheme into wiki-like distributed collective intelligence scheme, which may become the main stream for real-time traffic information collection and sharing scheme in the near future.

Acknowledgements

This work was partially supported by National Science Council of the Republic of China under Grant No. NSC95-2520-S009-008-MY3. The authors wish to express their appreciation to the Telecommunication Labs of Chunghwa Telecom Co. Ltd. for co-working the proposed framework.

References

[1] E. Chung, M. Sarvi, Y. Murakami, R. Horiguchi, M. Kuwahara, Cleansing of probe car data to determine trip OD, in: Proceedings of 21st ARRB and 11th REAAA Conference, Cairns, Australia, 2003.

[2] Tzu-Jung Cheng, Tyh-ming Lin, Yu-Ching Ou-Yang, 2001 Highway Capacity Manual in Taiwan, Government Publication No.: 1009001056, 2001, Available: <http://www.iot.gov.tw/>.

[3] V. Furtado, L. Ayres, M. de Oliveira, E. Vasconcelos, C. Caminha, J. D'Orleans, M. Belchior, Collective intelligence in law enforcement – the wikicrimes system, Elsevier Information Sciences 180 (1) (2010) 4–17.

[4] Wei-Hsun Lee, Shian-Shyong Tseng, Sheng-Han Tsai, A knowledge base real-time travel time prediction system for urban network, Elsevier Expert Systems with Applications 36 (3, Part 1) (2009) 4239–4247.

[5] H.E. Lin, R. Zito, A review of travel-time prediction in transport and logistics, in: Proceedings of the Eastern Asia Society for Transportation Studies, vol. 5, 2005, pp. 1433–1448.
[6] Y.T. Lin, S.S. Tseng, C.F. Tsai, Design and implement of new object-oriented rule base management system, Elsevier Expert Systems with Applications 25 (3) (2003) 369–385.
[7] I. Lykourentzou, K. Papadaki, D.J. Vergados, D. Polemi, V. Loumos, CorpWiki: a self-regulating wiki to promote corporate collective intelligence through expert peer matching, Elsevier Information Sciences 180 (1) (2010) 18–38.
[8] T. Nakata, J. Takeuchi, Mining traffic data form probe-car system for travel time prediction, in: Proceedings of ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Seattle, WA, USA, August 2004, pp. 817–822.
[9] Debbie Richards, A social software/Web 2.0 approach to collaborative knowledge engineering, Elsevier Information Sciences 179 (15) (2009) 2515–2523.
[10] Martijn C. Schut, On model design for simulation of collective intelligence, Elsevier Information Sciences 180 (1) (2010) 132–155.
[11] C. Shannon, W. Weaver, The Mathematical Theory of Communication Urbana, University of Illinois, 1964.
[12] J.S. Yang, Travel time prediction using the GPS test vehicle and Kalman filtering techniques, in: Proceedings of American Control Conference, Portland, OR, USA, 2005, pp. 2128–2133.
[13] Hsin-Chieh Wu, The Application of Entropy of Data Fusion of Traffic Information, Thesis of Master of NCTU, 2004, pp. 16–19.
[14] Ruei-Feng Wu, The Optimal Weighting Approach for Traffic Data Fusion, Thesis of Master of NCTU, 2006, pp. 11–27.
[15] Yifeng Zhou, Henry Leung, Minimum entropy approach for multisensor data fusion, in: Proceedings of IEEE Signal Processing Workshop on Higher-Order Statistics (SPW-HOS '97), 1997, p. 0336.
[16] DRAMA, Coretech Inc., 2003, Available: <http://www.ctknow.com.tw/rule.htm>.
[17] National traffic information center, Institute of Transportation MOTC, Available: <http://www.e-iot.iot.gov.tw/>.
[18] Police station public traffic service, Police radio station, Available: <http://www.prs.gov.tw/>.
[19] TANFB Traffic information system, Taiwan Area National Freeway Bureau, MOTC, Available: http://www.1968.nfreeway.gov.tw/.