A Novel Traffic Information Estimation Method Based on Mobile Network Signaling

Ling-Chih Kao
Telecommunication Laboratory of Chunghwa Telecom
Taoyuan, Taiwan, ROC
Email: lckao@cht.com.tw

Zsehong Tsai
Graduate Institute of Communication Engineering
National Taiwan University
Taipei, Taiwan, ROC
Email: ztsai@ntu.edu.tw

Abstract*—Two commonly used methods for traffic information rely on Vehicle Detector (VD) and Global Positioning System-Based Vehicle Probe (GVP); however, they have some confinements, such as high cost for construction and maintenance, and limited coverage. For the sake of overcoming dilemmas happened on VD and GVP, Cellular-Based Vehicle Probe (CVP) comes into being. However, the current applications for CVP mainly focus on arteries or freeways, where traffic information for longer distance is derived from two Inter-Visitor Location Register Location Area Update (Inter-VLR LAU) events with various Location Area Code (LAC) borders, and the one for shorter distance is from two consecutive handover events. The perplexity of available CVP techniques comes about is that there are no two Inter-VLR LAU events with various LAC borders and few handover events in scenic spots. In order to expand the applications for CVP to scenic spots, a cost-effective and flexible method utilizing mobile network signalling called Enhanced CVP (ECVP) is proposed. The key concept for ECVP is that we adopt Inter-VLR LAU events at the origin and all kinds of communication events at the destination to retrieve traffic information. The inaccuracy of ECVP consists in the uncertainty of event occurred time at the destination. Therefore, with a view to acquiring more accurate traffic information, three novel Reinforced CVP (RCVP) algorithms, inclusive of Fixed r percent samples CVP (F-RCVP), Dynamic r percent samples (D-RCVP), and Dynamic r percent samples with Discarding former samples (DD-RCVP), are presented. Numerical results show that F-RCVP is suitable for scenic spots that the LAC border only contain samples resulting from cars. By contrast, if the samples consist of both cars and motorcycles, it is recommended that D-RCVP and DD-RCVP are introduced.

***Keywords*-CVP, GVP, handover, LAU, traffic information, VD**

I. Introduction

Nowadays, there are two methods in common use acquiring traffic information. The first one is by Vehicle Detector (VD), in which the VD traffic information center gathers traffic information from all VDs installed on roads. Because the life cycle of VD is limited, about five years, the maintenance cost is very high, especially in suburbs. Also, the coverage offered by VD mainly focuses on arteries or freeways. The second one is from Global Positioning System-Based Vehicle Probe (GVP), where the GPS traffic information center collects traffic information from fleets or taxis with GPS devices on roads. The amount of data is unstable due to the limitation of districts and predefined routes for the fleets or taxis. With the popularization of mobile network and for the sake of conquering the difficulties occurred in the previous two methods, Cellular-Based Vehicle Probe (CVP) emerges [1], [2], [3], [4], [5]. The comparisons for various techniques are summarized as follows [6]. Although the accuracy for CVP is not as good as that for VD or GVP, it has the advantage that no specific client or software is required in the mobile equipment. Furthermore, CVP adopts mobile network signaling among subscribers and mobile networks to estimate traffic information, and its coverage is widespread. Additionally, the cost is low and maintenance is easy compared with VD or GVP. There are benefits and defects for the three techniques, and here are the comparisons for them. For the data source, CVP is the most widespread among the three techniques. The data source for CVP is from mobile network signaling in mobile networks, the one for VD is by devices installed on roads, such as infrared rays, induction coil, image, and radar, and the one for GVP is via cars equipped with GPS devices. For CVP, because mobile network signaling is captured in the equipments' room of mobile operators, the maintenance is easier than that for VD or GVP. Because the cost of VD is proportional to the number of VDs, the one for GVP is proportional to the number of cars, and the one for CVP is independent of the number of terminals, the cost of VD or GVP is higher than that for CVP. Unlike CVP, the data source for VD or GVP is from outdoors; therefore, the life cycle for CVP is longer than that for VD or GVP. There are fewer people in scenic spots than in urban areas; as a result, the cost is high if VDs or GVPs are deployed in scenic spots. In spite of superiority about data source, maintenance, cost, life cycle and applicability in scenic spots for CVP, its accuracy is not so good as that for VD or GVP, and the algorithm is more complex compared with VD or CVP. Every technique has its applicability; hence, it is vital to make the most of possible techniques to economically provide comprehensive traffic information. The suitable technology for arteries or freeways is from VD, and the one for scenic spots or alternative roads toward the airport is by CVP. For the rest of roads that can't be fulfilled by VD or CVP, the best way is via GVP.

The current available solutions for CVP are applied in freeways. As shown in Fig. 1, when a mobile equipment moves from a Location Area Code (LAC) to the other one, it will automatically execute an Inter-Visitor Location Register Location Area Update (Inter-VLR LAU) procedure, in which

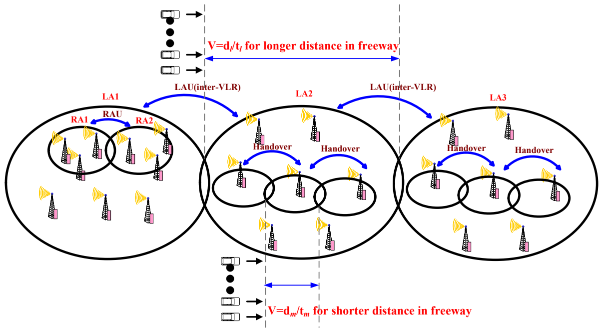

Fig. 1. The current available solutions for CVP.

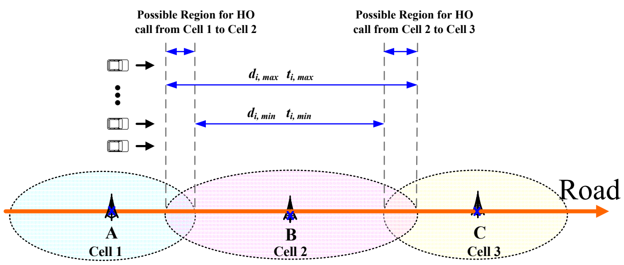

Fig. 2. Difficulty for the CVP Algorithm of Freeway.

Fig. 3. The map from Lugu to Xitou.

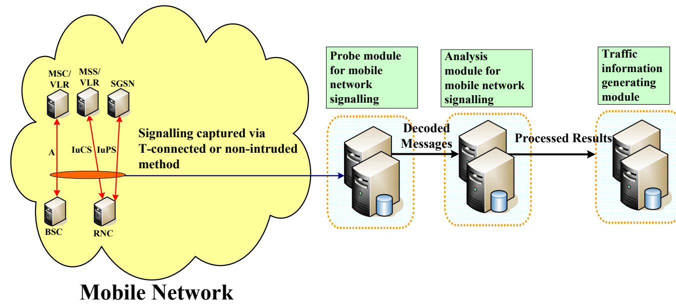

Fig. 4. The architecture of proposed ECVP system.

it will camp on a particular cell at certain time. That cell has its corresponding longitude and latitude. When a mobile equipment has a ongoing voice call and moves from a cell to the other one, the base station subsystem (BSS)/Universal Terrestrial Radio Access Network (UTRAN) will automatically report a new serving cell to the core network for it, which is known as a handover procedure. Using the two Inter-VLR LAU events with various LAC borders constitutes the algorithm of freeways for a longer distance, and the average speed is defined as the distance of two Inter-VLR LAU events with various LAC borders divided by the time difference of them [1], [2], [3], [4]. For a shorter distance, two handover events shall be adopted, and the average speed is defined as the distance of two handover events divided by the time difference of them [4], [5]. The difficulty for this algorithm lies in that the actual location of Inter-VLR LAU or handover event is a variable depending on radio environments, speed, etc. As shown in Fig. 2, the distance of two consecutive handover events is range from $d_{l,min}$ to $d_{l,max}$. For the sake of obtaining the most reasonable location, field testings are required. Is the algorithm of freeway described above still applied to scenic spots? The answer is no because two dilemmas occurred on arterial roads in scenic spots. One of the dilemmas is only a single Inter-VLR LAU event; therefore, we can't use two Inter-VLR LAU events with various LAC borders to get traffic information. The other is that the number of two consecutive handover events in scenic spots is less than that in cities; hence, we can't use two different handover events to obtain traffic information. As shown in Fig. 3, the map of Xitou in Taiwan, if one wants to acquire travel time from Lugu to Xitou, a novel algorithm shall be investigated.

The remainder of the paper is organized in the following. In Section II, the architecture and functionality of proposed CVP System is described. In Section III, the reinforced CVP algorithms are derived. Numerical results are discussed in Section IV. Conclusions are drawn in Section V.

II. Architecture and Functionality of Proposed CVP System

In order to efficiently and cost-effectively obtain traffic information from the mobile network, one needs a total solution. This section develops a traffic information estimation system called Enhanced CVP (ECVP) utilizing mobile network signaling.

The ECVP system portrayed in Fig. 4 is composed of three major functional blocks: probe module, analysis module, and traffic information generating module. All of them are developed by ourselves and hence cost-effective. First, the probe module [7], [8] takes charge of capturing and decoding signaling between mobile network equipments of circuit-switched and packet-switched mobile networks [9], [10]. The capture method we adopt is T-connected or non-intruded, and therefore it has no impact on the performance of mobile network. The involved interfaces include A [11], [12] between Base Station Controller (BSC) and MSS, IuCS [13], [14] between Radio Network Controller (RNC) and MSS, and IuPS [13], [14] between RNC and Serving GPRS Support Node (SGSN). Because mobile network signaling can't be directly applied to generate traffic information, one needs a analysis module to convert mobile network signaling to useful events in near real-time. The analyzed events cover LAU, voice call, video call, Short Message Service (SMS), Supplementary Service (SS), service request, attach, and activate PDP context, where the content of every event comprises Cell ID, Occurred Time, International Mobile Subscriber Identification Number (IMSI), International Mobile Equipment Identification Number (IMEI), Location update type, and mobility behavior. At the final stage, a traffic information generating module based on those events obtained from analysis module is required.

How do we derive traffic Information in scenic spots by mobile network signaling? I illustrate the proposed algorithm using the map from Lugu to Xitou as shown in Fig. 3, whose total distance from the origin A to the destination B is about sixteen point two kilometers. First, each selected event at the origin A is Inter-VLR LAU, the occurred time of which is just when the subscriber enters this Inter-VLR LAC border. Second, selected events at destination B are any kind of events, including voice call, video call, SMS, SS, and LAU, whose occurred time is indefinite. Finally, the average travel time from A to B is defined as

$$\sum_{i=1}^{N} \frac{t_{d,i} - t_{o,i}}{N} \tag{1}$$

where $t_{d,i}$ is the time of all events for subscriber i at the destination B in the recent ten minutes, $t_{o,i}$ is the time of all Inter-VLR LAU events for subscriber i at the origin A corresponding to those events for $t_{d,i}$, and N is the number of all possible samples. Because the event occurred time at the destination for each mobile equipment is indefinite, not all samples are useful. With a view to getting the optimal result for equation 1, the three reinforced algorithms will be discussed in the next section.

III. The Reinforced CVP Algorithms

The perplexity of traditional CVP techniques encounters while applied on arterial roads in scenic spots lies in that the number of two consecutive Inter-VLR LAU events with various LAC borders or handover events is not sufficient. Therefore, the principal concept for the ECVP algorithm we proposed in the previous section is that we take advantage of Inter-VLR LAU events at the origin to retrieve all the possible communication events and make the the occurred time definite. That is, when subscribers enter the Inter-VLR LAC border, they will automatically perform Inter-VLR LAUs, so the occurred time is definite for every subscriber, instead of waiting for a period of time. Moreover, for the sake of solving the problem that there are no possible Inter-VLR LAU events at the destination, we adopt all kinds of communication events. However, the derived problem is that the event occurred time at the destination is uncertain. As a result, one needs a comprehensive method to obtain more accurate traffic information. Depending on the ways to deal with the possible samples, three Reinforced CVP (RCVP) algorithms will be illustrated in the followings.

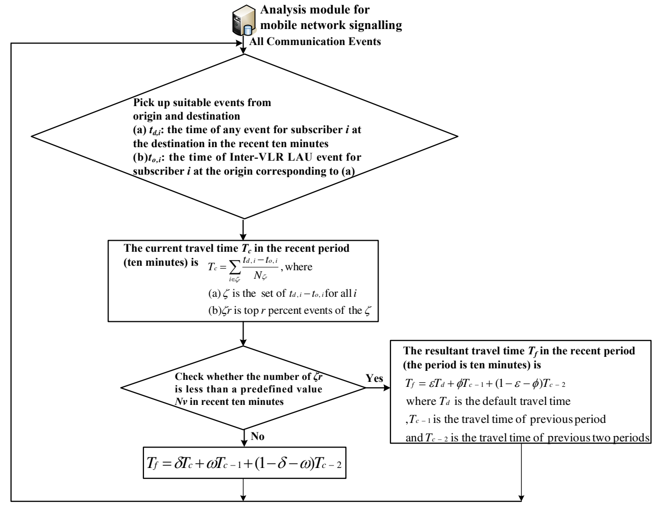

Fig. 5. The F-RCVP algorithm under fixed r.

A. Algorithm 1-Fixed r

Because communication events may be occurred when subscribers arrive at the destination and wait for a period of time, the intuition method is to select the former samples with faster arriving time. Based on this concept, the first proposed RCVP algorithm called Fixed r percent samples CVP (F-RCVP) [15] is shown in Fig. 5 and explained below.

- Step 1: Calculate all possible $t_{d,i}$ and $t_{o,i}$ based on the data obtained from the analysis module as described in the previous section.
- Step 2: The current travel time T_c in the current time period from the origin to destination is defined by

$$\sum_{i\in\zeta_r} \frac{t_{d,i} - t_{o,i}}{N_{\zeta_r}} \tag{2}$$

  where ζ is the set of $t_{d,i} - t_{o,i}$ for all i, ζ_r is the top r percent of ζ, and N_{ζ_r} is the number of ζ_r.
- Step 3: The travel time based on T_c may be fluctuated a lot; therefore, we use the weighted average of travel time across three time periods. First, we need to check whether the number of ζ_r is less than a predefined value N_v in recent ten minutes. If the answer is yes, it means that the number of samples in the current time period is not enough. Hence, we need to abate the influence of them to T_c. The resultant travel time T_f is defined as

$$T_f = \delta T_d + \omega T_{c-1} + (1 - \delta - \omega) T_{c-2} \tag{3}$$

  where T_d is the default travel time which is usually the travel time under smooth traffic condition, T_{c-1} is the travel time of the previous time period, and T_{c-2} is the travel time of the previous two time periods. Otherwise, T_f is defined as

$$T_f = \varepsilon T_c + \theta T_{c-1} + (1 - \varepsilon - \theta) T_{c-2} \tag{4}$$

  Note that the weighting factor in T_f mentioned above is based on the historical data, and $0 < \delta, \omega, \varepsilon, \theta < 1$.

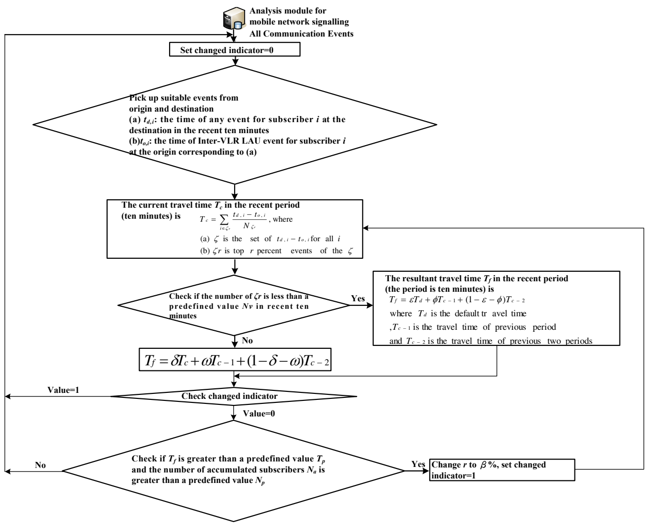

Fig. 6. The D-RCVP algorithm under dynamic r.

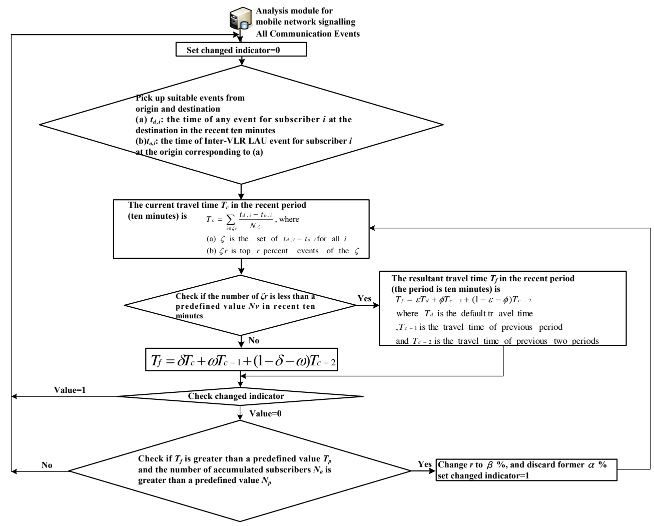

Fig. 7. The DD-RCVP algorithm under dynamic r with discarding former samples.

- Step 4: After obtaining T_f, then go to Step 1 to calculate the travel time for the next time period.

B. Algorithm 2-Dynamic r

The algorithm for F-RCVP only pertains to arterial roads in scenic spots where samples only result from cars. If samples originated from both cars and motorcycles, the algorithm requires modifying because the subscribers riding motorcycles may not be congested on the road while in heavy traffic conditions. For this purpose, more samples shall be allowed to calculate the travel time, and the refined algorithm called called Dynamic r percent samples CVP (D-RCVP) [15] shown in Fig. 6 is as follows.

- Step 1: Set changed indicator to zero, it means that T_f is calculated one time in the current time period. If T_f is calculated two times in the current time period, changed indicator is changed to one.
- Step 2 to Step 4: The procedures for Step 2 to Step 4 are the same with those in Step 1 to Step 3 of the algorithm 1.
- Step 5: Check the value of changed indicator. If the value is 1, go to the next time period; otherwise, check if T_f is greater than a predefined value T_p and the number of accumulated subscribers N_a between the origin and destination is greater than a predefined value N_p. If the answer is no, go to the next time period; otherwise, change r to $\beta\%$ and set changed indicator to one. At the same time, re-calculate T_f based on Step 2 to Step 4 with new r. Note that

$$N_a \doteq N_b - N_d - N_t \tag{5}$$

where N_b is the number of accumulated subscribers traverses the origin, N_d is the number of subscribers traverses the destination, and N_t is the number of subscribers where the sojourn time exceeds T_f.

C. Algorithm 3-Dynamic r, Discard Former Samples

Although the algorithm for D-RCVP permits more samples, the samples arising from motorcycles still exist. In order to conquer this problem, the Step 5 in the D-RCVP algorithm is enhanced by

- Step 5: Check the value of changed indicator. If the value is 1, go to the next time period; otherwise, check if T_f is greater than a predefined value T_p and the number of accumulated subscribers between the origin and destination N_a is greater than a predefined value N_p. If the answer is no, go to the next time period; otherwise, change r to $\beta\%$, discard former $\alpha\%$ samples and set changed indicator=1. At the same time, re-calculate T_f based on Step 2 to Step 4 with new r.

The ultimate algorithm called Dynamic r percent samples with Discarding former samples CVP (DD-RCVP) is shown in Fig. 7.

IV. Numerical Results

Based on the various RCVP algorithms proposed in the previous section, comprising F-RCVP, D-RCVP, and DD-RCVP, we can obtain the travel time from Lugu to Xitou.

A. Algorithm F-RCVP

Take the days on Feb. 11th 2013 and Feb. 14th 2013, under the conditions that $r = 10$, $N_v = 10$, $t_d = 25$ minutes, $\delta = 0.2$, $\omega = 0.5$, $\varepsilon = 0.8$, and $\theta = 0.1$, for examples, the horizontal axis is time and vertical axis is travel time T_f. Note that r, N_v, t_d δ, ω, ε, and θ mentioned above depend on historical data. From T_f on Feb. 11th 2013 shown in Fig. 8, we can know that traffic condition on that day is smooth and the travel time is about 25 minutes. By contrast, from T_f on Feb. 13th 2013 shown in Fig. 9, we can know that the traffic is congested from 10:40 to 15:50, and the travel time is more than thirty minutes. However, the congested travel time

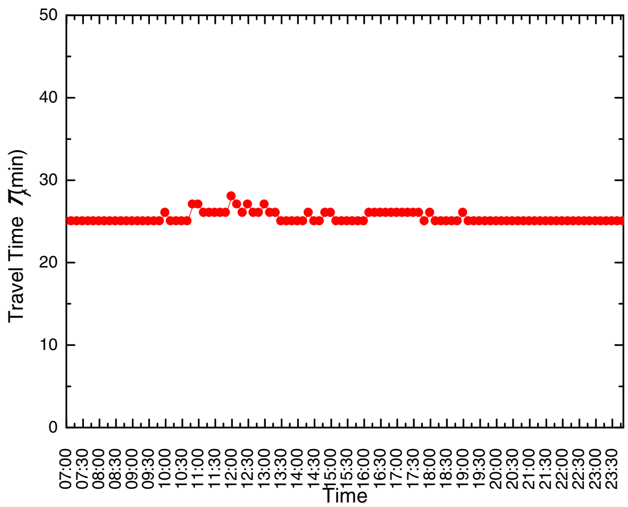

Fig. 8. The travel time from Lugu to Xitou on Feb. 11th 2013 under F-RCVP algorithm.

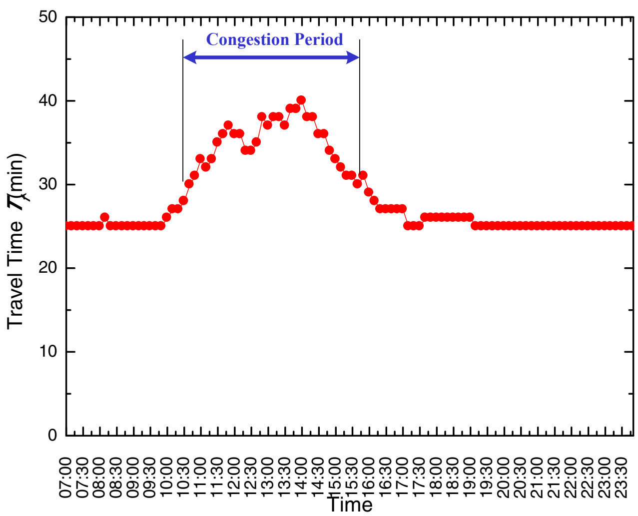

Fig. 9. The travel time from Lugu to Xitou on Feb. 13th 2013 under F-RCVP algorithm.

need modifying as the LAC border contains the samples of motorcycles.

B. Algorithm D-RCVP

In order to understand how N_a affects T_f, we add N_a in Fig. 8 and Fig. 9, and re-shown in Fig. 10 and Fig. 11 respectively. From these two figures and historical data, we can come to conclusion that the traffic from Lugu to Xitou is congested under the condition that $N_a \geq 800$ and $T_f \geq 30$.

After understanding the value of N_a and T_f under traffic congestion condition, we can then utilize it for D-RCVP to determine r dynamically. According to historical data, $r = \beta\%$ is set to 50%, and refined T_f is shown in Fig. 12. From this figure, we can know that the congestion time is from fifty

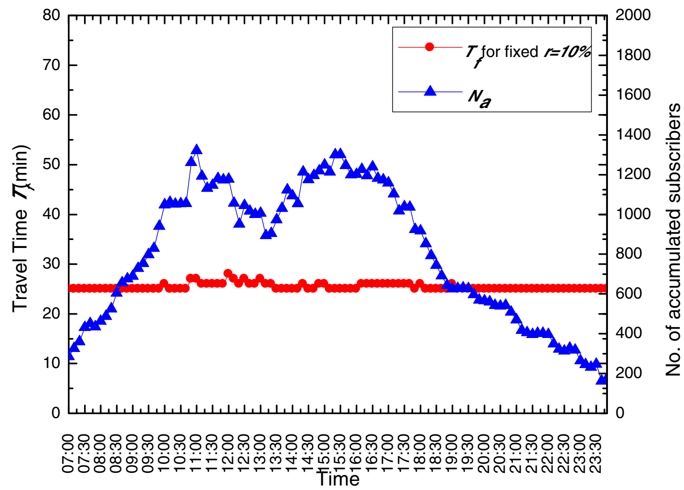

Fig. 10. The travel time and the number of accumulated subscribers N_a from Lugu to Xitou on Feb. 11th 2013 under F-RCVP algorithm.

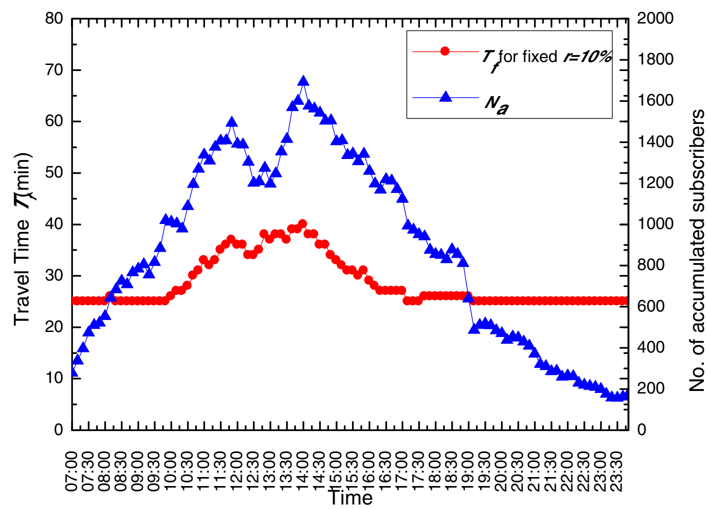

Fig. 11. The travel time and the number of accumulated subscribers N_a from Lugu to Xitou on Feb. 13th 2013 under F-RCVP algorithm.

minutes to seventy minutes. By contrast, the congestion time for F-RCVP is from thirty minutes to forty minutes.

C. Algorithm DD-RCVP

For the sake of suppressing the influence of motorcycles on the travel time while in traffic congestion, we need to discard the the former $\alpha\%$ samples obtained from D-RCVP. Based on historical data, the reasonable value for α is 20, and the resultant T_f is shown in Fig. 13. From this figure, we can know that the congestion time is from sixty-five minutes to ninety-five minutes. According to the data from radio broadcast system on that day, the result retrieved from DD-RCVP is close to the real situation.

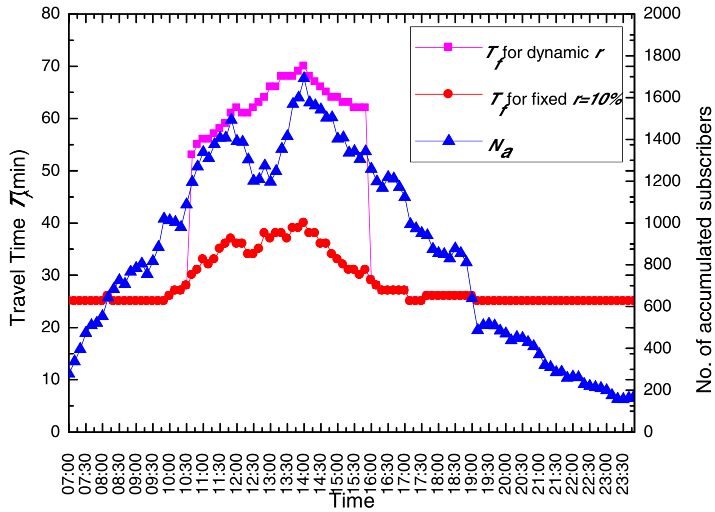

Fig. 12. The travel time and the number of accumulated subscribers N_a from Lugu to Xitou on Feb. 13th 2013 under D-RCVP algorithm.

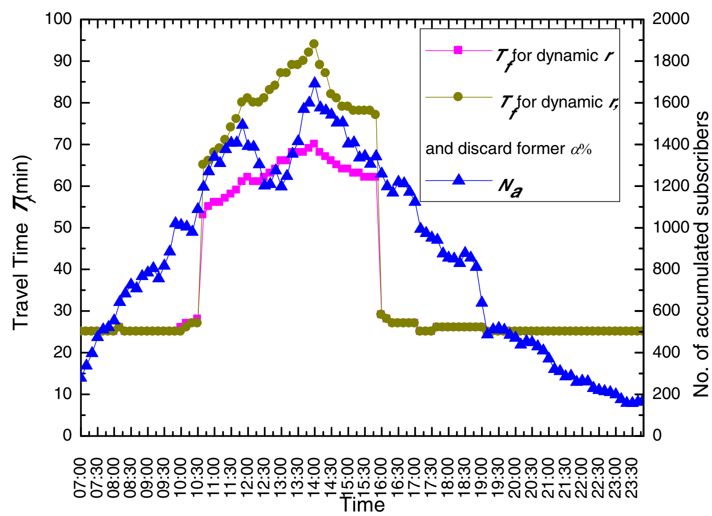

Fig. 13. The travel time and the number of accumulated subscribers N_a from Lugu to Xitou on Feb. 13th 2013 under DD-RCVP algorithm.

V. Conclusion

In this paper, we developed a CVP system, ECVP, which directly utilizes mobile network signaling to estimate traffic information characterizing with low cost, easy maintenance, and widespread coverage. The system is composed of three major functional blocks: probe module, analysis module, and traffic information generating module. All of them are developed by ourselves and hence cost-effective.

Compared with available CVP solutions using LAU or handover events to get traffic information, we adopt Inter-VLR LAU events at the origin and any kind of events at the destination to retrieve traffic information, which fits to scenic spots. For the purpose of getting more accurate and reliable traffic information, three reinforced algorithms including F-RCVP, D-RCVP, and DD-RCVP are proposed. Owing to the uncertainty about the occurred time for the events at the destination, F-RCVP truncates all samples to former r percent, which is suitable for the situation where the LAC border only contains samples of cars. Contrarily, if the samples cover both cars and motorcycles, it is advised that D-RCVP and DD-RCVP are used. The dissimilarity between D-RCVP and F-RCVP is that r is changed from the fixed to the dynamic. Moreover, r depends on the accumulated subscribers between the origin and destination, N_a, which is set to the former β percent samples. The reason for migrating D-RCVP to DD-RCVP is to eliminate the influence of motorcycles on traffic information while in traffic congestion. Therefore, the former α percent ones among the β percent samples are discarded. The algorithm we proposed is the world-first solution applied on arterial roads in scenic spots adopting mobile network signaling, and has been launched in Chunghwa Telecom HiNet's service.

References

[1] Puntumapon, K. and Pattara-atikom, W., "Classification of Cellular Phone Mobility using Naive Bayes Model,"*IEEE VTC*, 2008.

[2] Bolla, R. and Davoli, F., "Road Traffic Estimation from Location Tracking Data in the Mobile Cellular Network ,"*IEEE WCNC*, 2000.

[3] Knuuttila, "Method and System for Collecting Traffic Data," US patent, 10/838,915, May 2004.

[4] David Gundlegard and Johan M Karlsson, "Generating Road Traffic Information from Cellular Networks - New Possibilities in UMTS,"*ITS Telecommunications Proceedings*, 2006.

[5] Gundlegard, D. and Karlsson, J.M., "Handover Location Accuracy for Travel Time Estimation in GSM and UMTS,"*IET Intelligent Transport Systems*, 2009.

[6] H. S. ZHANG, et al., "An Architecture of Traffic State Analysis Based on Multi-Sensor Fusion,"*Proceedings of 2005 IEEE International Conference on Intelligent Vehicles Symposium*, 2005.

[7] L.-C. Kao, et al., "Mobile Network Diagnosis and Location Services via Stream-based Signalling,"*IEEE APNOMS*, pp. 1–4, 2011.

[8] Ling-Chih Kao, et al., "The System and Method of Multi-interface Signaling Trace for the UMTS Mobile Network," China patent, ZL200810090775.7, March 2011.

[9] "Digital cellular telecommunications system; Network architecture," Technical Specification 3GPP TS 03.02 version 7.1.0, (2000-01), 2000.

[10] 3GPP, "3rd Generation Partnership Project; Technical Specification Group Services and Systems Aspects; Network architecture," Technical Specification 3GPP TS 23.002 version 7.6.0, (2008-12), 2008.

[11] 3GPP, "3rd Generation Partnership Project; Technical Specification Group GSM EDGE Radio Access Network; Mobile-services Switching Centre - Base Station System; (MSC - BSS) interface; Layer 3 specification," Technical Specification 3GPP TS 08.08 version 7.7.0, (2000-10), 2000.

[12] 3GPP, "3rd Generation Partnership Project; Technical Specification Group Core Network; Mobile radio interface Layer 3 specification," Technical Specification 3GPP TS 04.08 version 7.21.0, (2003-12), 2003.

[13] 3GPP, "3rd Generation Partnership Project; Technical Specification Group Radio Access Network; UTRAN Iu interface; Radio Access Network Application Part (RANAP) signalling," Technical Specification 3GPP TS 25.413 version 7.10.0, (2009-03), 2009.

[14] 3GPP, "3rd Generation Partnership Project; Technical Specification Group Core Network and Terminals; Mobile radio interface Layer 3 specification; Core network protocols; Stage 3," Technical Specification 3GPP TS 24.008 version 7.15.0, (2010-03), 2010.

[15] L.-C. Kao, "Traffic Information Estimation of Arterial Roads in Scenic Spots Adopting Mobile Network Signalling,"*AP-RASC'13*, Sept. 2013.