Proceedings of APCC2015 copyright © 2015 IEICE 14 SB 0087

Serviec Optimization in Vehicular Content-Centric Networking

Liang Shan
CE Dept. Sungkyunkwan Univ.
Suwon, South Korea
ethan080916@gmail.com

Do-hyung Kim
CS Dept. KAIST Univ.
Daejeon, South Korea
Mr.dohkim@gmail.com

***Abstract*—The Content Centric Networking is scalable and efficient in the Internet, it raises concerns when deployed in a mobile large scale network like the Vehicular Ad-hoc Network (VANET). However, many applications in multi-hop ad-hoc vehicular networks are push-based and require fast communications. The CCN's pull-based transport in such cases may backfire, underperforming and costly without careful design. According to method for reducing the network cost caused by overheads and congestion. In this paper we firstly propose a service discovery scheme then optimize the number of vehicles which cache the matching data for a certain interest in the Content Store (CS), while satisfying the requirement of delay. We model this problem as optimization problem and we prove this problem is NP-Complete. Then we develop a heuristic algorithm to solve this problem according to the normal distribution property.**

Keywords—Content Centric Networking, Vehicle-to-Vehicle Communication, Optimization.

I. Introduction

Content Centric Networking (CCN) [1] is a pioneering work for the future Internet, which points out a new direction to design vehicular ad hoc networks. CCN focuses on retrieving desired data rather than establishing a session between two end nodes. It also eliminates the use of node addresses and retrieves data by using application data names directly[3-6]. CCN is scalable and efficient, it raises concerns when deployed in a mobile large scale network like the Vehicular Ad-hoc Network (VANET). However, many applications in multi-hop ad-hoc networks are push-based and require fast communications. The CCN's pull-based transport in such cases may backfire, underperforming and costly without careful design.

Face to these challenges and realize this application, such local resources should be discovered first before provided. We refer to delay bounded service discovery as locating distributed resources and services in the network within a certain delay bound. By using this way, we can alleviate the congestion, lighten overhead and avoid wasting resources, which caused by broadcast storm problem and continually retransmission.

We propose this scheme can guarantees that Interests issued from vehicles can be answered within a given query delay bound while minimizing the number of vehicles which has the matching data, so as to minimize the bandwidth and deployment cost. Specifically, we define the data encounter ratio and data encounter guarantee ratio, which reflects the fraction of vehicles that can be served by these vehicles which have the corresponding data to the Interest in the required delay and the confident interval to guarantee this encounter ratio, respectively. Then we formulate the problem as an optimization problem that minimizes the number of vehicles while ensure data encounter ratio and the data encounter guarantee ratio requirements mentioned above. We prove this problem is NP-Complete. Then we design a heuristic algorithm to iteratively add the vehicle that can maximize the expected increment of vehicles which can be encounter with the vehicles as a data mule within the required delay.

The remainder of this paper is organized as follows. In Section 2 after we introducing some basic concepts we give our naming scheme, then we formally define the problem under consideration and present the algorithms for optimizing the number of vehicles in Data Encounter Set. The simulation results are presented in Section 3. Finally, we conclude the paper in Section 4.

II. Design and Optimizing

A. Basic Concepts

There are three kinds of communication modes in vehicular networks, infrastructure, Ad-hoc and hybrid. We can also classify Ad-hoc mode into two kinds of communication modes: one is that some vehicles have better processing capacity and more energy than the others, which is similar with the sensor and actor networks. Such communication mode may lead to the redundant deployment cost. The other mode is all the vehicles have the same level of processing capacity and energy. We use the latter communication mode in this paper.

Above all, we use the Ad-hoc communication mode without any infrastructure and all vehicles have the same level processing capacity and energy in the network, which illustrated in figure 1.

Vehicles perform three roles in the network, which are data publishers, mules and data consumers. The publishers produce and store data about specific events. The mules cache all the data they can collect for a later redistribution. Consumers request specific data using Interest packets targeting the publishers or the mules. Multiple data can be processed into a new one which carries different information. That processed data can be fetched by consumers. Each vehicle can act at -

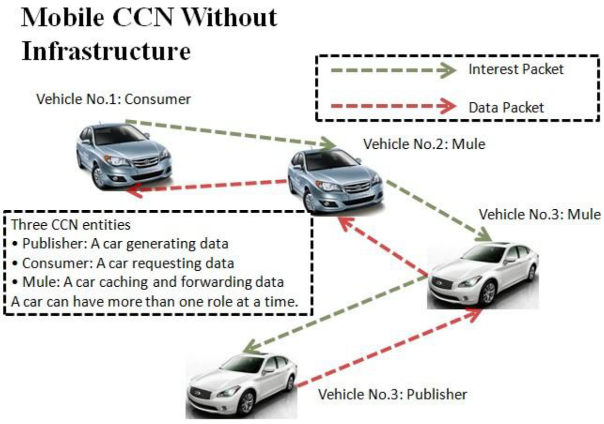

Fig. 1. V2V Communication Architecture in CCN.

least one of all roles in the network.

B. *Naming Scheme*

The We propose the following structure of data names for V2V traffic information exchanges, which base on the naming design in [2]:

/IDv/Locationc/locationd/Direction/IDt/Datatype/Timestamp/Nonce.

Here the IDv is the unique ID for each vehicle, which can be based on the license of each vehicle. Locationc is the geographical location of the consumer represented by longitude and latitude coordinates (x, y), which can be fetched by some applications such as iTouchMap equipped in smart device. Locationd is the geographical location of destination for Interest. Direction is the direction of headway represented by a geographical angle clockwise from the due north. IDt is the traffic ID for data. Data type component indicates the meaning of the data itself, e.g. closed lane, vehicle speed, etc. Timestamp component uses the same UNIX timestamp format that is used for the data.

C. *Definition and Preliminaries*

For a specified Interest packet, there are some vehicles can provide the matching Data correspond to the Name. These vehicles form a set, in which each vehicle caches the matching Data in the CS and can provide service to the Interest. During the operation time for retrieving matching data, some vehicles will add into the set. The purpose of our design is to optimize the number of the vehicles in the Data Encounter Set.

In the networks, vehicles can communicate with each other when two vehicles come within the communication range R. And we refer to such communication opportunity as an encounter event. The N vehicles in the networks and the set of all the location sequences of the Interest packets in the vehicles in N between time ta and tb is denoted by P(ta, tb). The vehicles cache the matching data in the CS compose a set denoted by E. Finally, D is a delay bound for retrieving data and the interest will fail if not been answered within this D time bound.

D. *Formulation*

In this part, I will prove the Data Encounter Set problem is a NP-complete problem. The interest packet v is said to be in set E with a delay D if there is any u ∈ E that can be found by v within delay D. It is denoted as v ∈ M (E, D), where M (E. D) is called data encounter set.

Definition 1 (Data Encounter Ratio)

Data encounter ratio is the proportion of the interest v ∈ N that satisfies v ∈ M (E, D). It is denoted as:

$$\delta(E, D) \triangleq \frac{\|M\,(E, D)\|}{\|N\|} \tag{1}$$

,where the ‖ * ‖ represents the size of a set.

Definition 2 (Optimal Data Encounter Problem)

$$\min_E \ \|E\| \quad s.t.\ \ \delta(E, D) > \gamma \tag{2}$$

, where γ is a required data encounter ratio.

Based on our routing design, for an interest packet, we can discuss whether its trace is known or not. According to the both situations, we have the Conditional Optimal Data Encounter problem and the Unconditional Data Encounter problem.

Definition 3 (Conditional Optimal Data Encounter Problem)

The conditional optimal data encounter problem is an optimal data encounter problem which the P(t, t+D) is given as a condition. The conditional optimal data encounter problem is NP-Complete.

To proof a problem is NP-Complete we have to proof two things: firstly, the problem is in the NP class. Secondly, it is NP-Hard problem, since the problem is NP-Complete. Here we firstly show the conditional optimal data encounter problem is in the NP problem set. We can choose one subset E from N that has k elements. Then we check how many interests will be covered by E scanning P(t, t + D) and determine whether the constraint in eq.4 is satisfied. This testing phase can be executed in polynomial time. Therefore, conditional optimal data encounter problem is in the NP class. We prove the conditional optimal data encounter problem is NP-hard by reducing the Dominating Set Problem to it, where the dominating set problem is a classical NP-Complete decision problem mentioned in [11]. In graph theory, the dominating set for a graph G = (V, E) is a subset T of V such that every vertex in V that is not in T will share some edges with at least one vertex in T. And the dominating set problem is defined as whether ‖ T ‖ ⩽ K for a given Graph and input integer K exists. Here we reduce dominating set problem to conditional optimal data encounter problem. Given the input of the Dominating Set Problem G and K, we can construct the input of conditional optimal data encounter problem as follows:

Each interest packet in a vehicle v_i corresponds to one vertex v_i of G. For each edge $e_{ij} \in$ G, we create P(t, t + D) so that there is a encounter event between v_i and v_j in the future at a random time $t_{ij} < D$. We can see the reduction can be performed in Polynomial time, and the answer to the new optimal data encounter problem is also the answer to the dominating set problem. Thus the conditional optimal data encounter problem is NP-Hard. Since conditional optimal data encounter problem is NP-Complete.

Definition 4 (Data Encounter Guarantee Ratio)

The data encounter guarantee ratio is a probability with the condition for data encounter ratio δ(E, D) > γ:

$$\zeta\ (E,\ D,\ \gamma) = \Pr\{\delta(E, D) > \gamma\} \tag{3}$$

Definition 5 (Unconditional Optimal Data Encounter Problem)

$$\min_E \ \ \|E\| \qquad s.t.\ \ \zeta\ (E, D, \gamma) > \varpi \tag{4}$$

, where □ is a required data encounter guarantee ratio, given a historical trace P (t-T, t)

Definition 6 (Encounter Indicator)

The Encounter Indicator Ii(E, D) is set to 1 when vi is covered by E in next time period of D and 0 otherwise.

$$I_i(E, D) \triangleq \begin{cases} 1 & \text{if } v \in M\ (E, D) \\ 0 & \text{otherwise} \end{cases} \tag{5}$$

Definition 7 (Pre-encounter Probability)

The Pre-encounter Probability of a set of interests in vehicles A to include a specific interest in some vehicle vi in the next time period of D is the probability that for all vj∈A, vi∈M(vj ,D). It is denoted:

$$g_i(\ A,\ D) \triangleq \Pr\{\forall v_j \in A, v_i \in M(\ v_j,\ D)\ \} \tag{6}$$

We can prove the random variable ε(E, D) follows a normal distribution when the number of interest packets to infinity.

$$\varepsilon(E, D)\ \sim N\left(\sum_{i=1}^{n} g_i\ (E, D),\ \sum_{i=1}^{n} g_i\ (E, D)\ \left(1 - g_i(E, D)\right)\right) \tag{7}$$

, where $g_i(\ E,\ D) \triangleq \Pr\{v_i \in M(\ E,\ D)\ \}$

According to the Venn Diagram, For any interest in a vehicle vi, its probability gi (E, D) of being found by any vehicle which has the matching data in E can be computed as:

$$h_i(E, D) = \sum_{i=1}^{\infty} (-1)^{(i+1)} \sum_{A \in E, \|A\| = i} g_i(A, D) \tag{8}$$

We assume these random variables Ii are independent from each other. The ε(E, D) can be viewed as the sum of all the indicators Ii:

$$\varepsilon(E, D) = \sum I_i \tag{9}$$

Here let $Y_i = I_i - h_i$. For any δ > 0.

$$1 > h_i\,(1 - h_i) = \ E\ Y_i^2 > E|Y_i|^{2+\delta} \tag{10}$$

Let $I_n = \sum_{i=1}^{n} Y_i$ and $E_n^2 = \mathrm{Var}\ I_n = \sum_{i=1}^{n} \delta_i^2$ then we have:

$$\frac{1}{E_n^{2+\delta}} \sum_{i=1}^{n} E|Y_i|^{2+\delta} \leq \frac{1}{E_n^{2+\delta}} \sum_{i=1}^{n} \mathrm{Var}\, Y_i = \frac{1}{E_n^{\delta}} \tag{11}$$

, if n → ∞, En → ∞ (which is true given h_i is not all 0 or 1), the Lyapunov condition is satisfied. According to the Lyapunov central limit theory and we have:

$$\sum_{i=1}^{n} \frac{Y_i}{E_n} \sim N\,(0, 1),\ \text{i.e.}$$

$$\varepsilon\ (E, D)\ \sim N\left(\sum_{i=1}^{n} h_i\,,\ \sum_{i=1}^{n} h_i\ (1 - h_i)\right) \tag{12}$$

, therefore the Random variable ε (E, D) follows a normal distribution when the number of interest packets to infinity.

III. Algorithms

In this section, we first use a greedy algorithm for solving the conditional optimal data encounter problem. Then according to the NP-complete feature, we give a heuristic algorithm to solve the unconditional optimal data encounter problem.

Algorithm. For unconditional optimal data encounter Problem

```
Input: p(t-T, t), D, γ, □
Output: E
E ← ∅
calculate all gi (A, D) from p(t − T, t).
While ture do
for all vj ∈ E do
w (vj) ←
end for
vm ← vi that maximize w (vj)
E ← E ∪ vm
if historical data satisfied
if the cooperative checking satisfied then
return E
end if
end if
end while
```

We use the greedy iterative approach for solving the conditional optimal data encounter problem. Then the same strategy can be used for dealing with the unconditional optimal data encounter problem. Firstly, we should know how to select the best vehicle in data encounter set in this situation, then how to stop the process.

For the first problem we introduce the Data Encounter Ratio Factor which can be denoted as:

$$w_i(E, D, v_j) = h_i(E \cup v_j, D) - h_i(E, D). \quad (13)$$

, where w_i can be calculated based on historical data to choose the best vehicle in data encounter set. Then we can stop the process according to the feature of normal distribution and the historical data.

IV. Simulation Results

In this section, we present our simulation results. We use an NS-3 based NDN simulator in version 3.20 to evaluate the performance of our design. The network size is 20km × 20km and the transmitting range for each vehicle is 150m. The other metric for the performance evaluation is the size of E. We take the variables of γ, □ and the delay D as the factors. The default settings are: D = 10min, γ = 0.95, □ = 0.9.

We summarize the data encounter set problem to an optimization problem for minimize the data encounter set size. We observe the data encounter set size is related to the required encounter ratio, guarantee ratio and the delay.

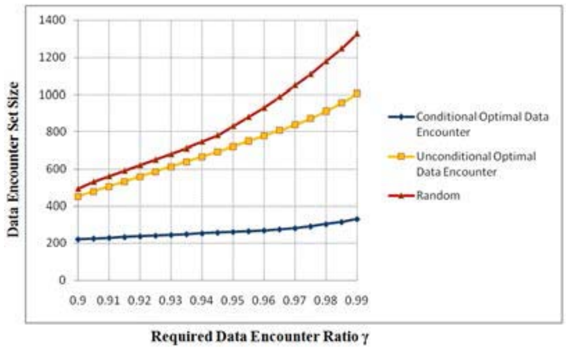

Fig. 2. Data Encounter Set Size, D = 10min, □ = 0.9.

We present the data encounter set size changes along the data encounter ratio and data encounter guarantee ratio for required delay in figure 2.

The unconditional optimal data encounter algorithm average performs over 25% better than the method for adding a vehicle into the encounter set randomly, for using the historical data and the feature of the normal distribution. The data encounter set size changes along the delay with required data encounter ratio and required data encounter guarantee ratio is shown in figure 3.We can clearly observe that the encounter set size decreases very fast until the delay increasing to 10 minutes.

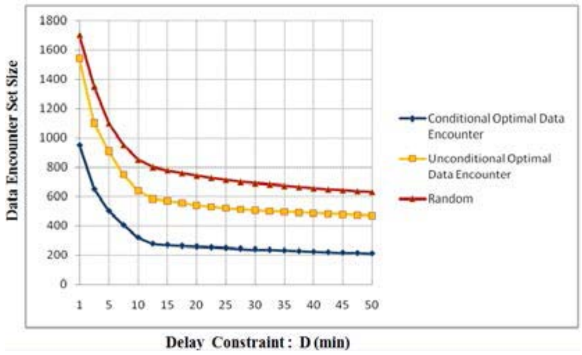

Fig.3. Data Encounter Set Size, γ = 0.95, □ = 0.9

We can clearly observe that the encounter set size decreases very fast until the delay increasing to 10 minutes. Then the set size goes to be stable, which means the data encounter set size is very close to the optimal solution when the delay constraint value is 10min with the known conditions γ = 0.95, □ = 0.9 in such networks.

V. Conclusion

We provide two contributions in this paper. Firstly, we present a position and direction hybrid data forwarding scheme for vehicle-to-vehicle communication in ad hoc content centric network, specifically face to the intersection problem. Secondly, we increase the network performance by optimizing the number data encounter set size. We prove this problem is NP complete problem and we use a heuristic algorithm to solve this problem. The simulation result shows our design over 25% better than the random method.

VI. References

[1] V. Jacobson, D. K. Smetters, J. D. Thornton, M. F. Plass, N. H. Briggs, and R. L. Braynard, "Networking. Named Content," in ACM CoNEXT, December 2009.

[2] Lucas Wang, Lixia Zhang et al "Data Naming in Vehicle-to-Vehicle Communications," in INFOCOM 2012 Workshop on Emerging Design Choices in Name-Oriented Networking, March 2012

[3] G. Arnould et al., A self-organizing content centric network model for hybrid vehicular ad-hoc networks, in: ACM DIVANet, 2011.

[4] G. Grassi et al., Vehicular inter-networking via named data, in: ACM HotMobile 2013.

[5] M. Amadeo et al., CRoWN: content-centric networking in vehicular ad hoc networks, IEEE Communications 16 (9) (2012) 1380–1383.

[6] M. Amadeo et al., Content-centric networking: is that a solution for upcoming vehicular networks? in: ACM VANET, 2012.