Abstract

Vehicle density is an important metric in monitoring traffic conditions to improve the safety and efficiency of roads. However, obtaining reliable traffic density results that faithfully reflect actual road conditions can be challenging. This paper presents a fully-distributed approach to obtaining estimations of traffic density in real-time with a high degree of accuracy and minimal errors. This is achieved by using the coverage area provided by each vehicle to determine which vehicle should be selected for the collection of traffic information capable of providing accurate estimations for the traffic density along the road segment between two junctions. The proposed scheme does not require any network infrastructure, and it is well-suited to urban environments. Simulation results demonstrate that the proposed density estimation strategy is more accurate and requires less transmission overhead than existing methods.

Introduction

Vehicular communication is an important area of research in the field of intelligent transportation systems. Vehicles with wireless communication capability are able to exchange traffic-related information and establish mutual awareness among one another to improve road safety and efficiency. The cooperative exchange of information, such as traffic density, can increase the situational awareness of drivers and the ability to detect problematic road traffic conditions. It also improves the efficiency of data delivery by avoiding the congested networks. However, the dynamic nature of vehicles in this environment makes the underlying network topology changes frequently, which can cause rapid changes in the density of traffic. Thus, the efficient collection and provision of appropriate road traffic data is vital to the estimation of road conditions with the aim of enhancing road safety in the vehicular environment.

Traffic density is among the most common metrics used to monitor road traffic conditions. A number of traffic information systems [1]-[3] have been designed for the collection and distribution of road traffic information in a centralized manner. However, most of these systems tend to leave out a major fraction of the road by covering only the specific locations in which sensors are deployed. This is due to the high cost of deploying and maintaining sensors across an entire road network. A number of studies have suggested one alternative decentralized approach [4]-[7] in which road traffic information is obtained without the need for telecommunication infrastructures. Vehicles in this scheme act as mobile sensors cooperatively sharing observations concerning road traffic conditions in its vicinity, which are then made available to all participating vehicles in the environment. This provides a more detailed view of road conditions, such as traffic density, which is qualitatively superior to systems using only static traffic sensors.

This paper proposes a distributed density estimation scheme referred to as Coverage Based Density Estimation strategy for VANET. The main objective is to provide accurate estimations related to traffic density within an urban road environment. This is achieved by employing the area covered by each vehicular node in the estimation of traffic density for a particular segment of the road between two junctions. Based on the transmission range of the node, it is possible to determine the area covered by each node for use in the selection of the node that is the farthest from the source and provides the most comprehensive coverage area in that neighborhood to act as the point from which traffic information should be collected. By applying this technique throughout the road segments that are being tracked, independent estimations of the traffic density in that segment of the road can be obtained independently by each selected node, based on the mutual awareness established among one another. Simulation results demonstrate the effectiveness of the proposed strategy as well as high accuracy in the estimation of traffic density with far less transmission overhead than that required by existing methods.

The remainder of this paper is organized as follows: Section II outlines previous related work. The problem statement is formulated in Section III. A detailed description of the proposed density estimation strategy is presented in Section IV. Section V presents simulation results and an evaluation of performance. Finally, conclusions are drawn in Section VI.

Related Work

The various traffic density estimation schemes that have been proposed can be classified as centralized and decentralized.
The centralized approach relies on fixed infrastructure, such as inductive loop detectors, roadside RADAR, infrared sensors, and surveillance cameras to be installed for the collection of traffic information [8]-[11]. This requires the deployment of a large number of sensors for monitoring traffic conditions. Traffic density can be estimated by comparing camera images or collecting information such as the number of vehicles passing a sensor. However, these approaches tend not to be very reliable and provide limited coverage despite high deployment and maintenance costs [15]. Thus, traffic information can only be obtained from streets on which sensors are implemented. A number of decentralized approaches [2], [16], [20] have also been proposed, wherein the vehicles utilizing the network cooperate with each other in gathering traffic information and estimating traffic density on the urban roads. These approaches have been widely used with routing protocols that require information pertaining to road traffic density [12]-[15].

In [2], a decentralized scheme was proposed wherein a table is compiled using data related to the current position and average velocity of vehicles. The table is periodically updated and exchanged with neighbor vehicles. Consequently, every vehicle becomes familiar with the position and average velocity of neighboring vehicles. However, this method is intended only for highway scenarios and depends on the direction of traffic for the exchange of tables among vehicles. In [16], data aggregation was used to gather and disseminate traffic information in real-time; however, it periodically disseminates the position and velocity of individual vehicles instead of average velocities in a given segment. This helps drivers to find suitable routes and important information concerning road conditions; however, it is suitable only for highway scenarios and does not consider the direction of vehicles in the exchange of traffic information among vehicles.

Other methods [19], [20] are based on counting the number of vehicles in a particular geographical location via clustering or grouping mechanisms. The decentralized traffic density estimation scheme in [19] divides each road into a number of fixed size cells based on the transmission range of vehicles. The cell density is calculated by having the number of vehicles in the cell counted by the vehicle that is located closest to the cell center. This vehicle is responsible for the estimation of cell density and the exchange of this information among other cells along the road until it reaches the designated junction. In this manner, the traffic density of each cell is combined with that of other cells to obtain the average density on the road. In [20], a neighbor-based scheme was proposed for the estimation of vehicle density based on the number of vehicles in the vicinity of the probe vehicle without the requirement of fixed size cells. However, dividing the road into cell poses various problems related to reliability. The following section discusses these problems in depth.

**Problem Statement and Formulation**

Numerous researchers [12]-[15] have used the idea of dividing the road segment into multiple fixed size cells and then sending probe packets from a source junction to its neighboring junctions for the collection of information. However, the fact that the location of nodes on the road is arbitrary and node density varies means that a node located close to the central point of each cell used to obtain network information does not necessarily exist. In this situation, the node selected in each cell may deviate considerably from the central point of that particular cell, such that it does not cover all the area within that cell. Thus, nodes located in the uncovered area are not included in the information collection process. More specifically, a portion of the road segment is not covered by the selected node, such that only a portion of the traffic observations are used in the estimation of traffic density. Thus, information collected along that road segment would not reflect the actual situation along that particular road, thereby compromising the accuracy of density estimation. This phenomenon can be formulated as follows:

If \( P(i) \) represents the center point of the cell \( i \), then we assume the point in \( P(i) \) is the origin point \( O_i = (0,0) \) of the Cartesian coordinate system as illustrated in Figure 1. The width and the height of the cell is \( W \) and \( H \), the corner points of cell \( i \) form a set \( S \).

\[
S = \{ \left( \frac{W}{2}, \frac{H}{2} \right), \left( -\frac{W}{2}, \frac{H}{2} \right), \left( \frac{W}{2}, -\frac{H}{2} \right), \left( -\frac{W}{2}, -\frac{H}{2} \right) \}.
\]

![Figure 1. Example of a cell used for traffic density estimation](image)

When the distance of any point \((x, y)\) within the cell to any point in set \( S \) is longer than the radius of the transmission range \( r \), this means that some of the area...
within this cell will not be able to be covered from the transmission range of that particular point. For example, the transmission range of node \( k \) located close to the top-right corner of the cell in Figure 2 does not cover all of the area in this cell.

Based on the result of above discussion, when the condition in (1) cannot be satisfied by a given point within a cell, then the area within the cell will not be fully covered by the transmission range of that particular point and the accuracy of density estimation is compromised.

\[
\forall (a, b) \in S \quad s.t. \quad \sqrt{(a-x)^2 + (b-y)^2} \leq r \quad (1)
\]

Generally, the width of a road is considerably less than the radio transmission range of a node. One simple way to improve the accuracy of density estimation is to divide the cell into even smaller areas. That is, instead of selecting the node that is located closest to the central point of a cell, the interquartile range \((Q_3 - Q_1)\) illustrated in Figure 3 can be used to identify the point from which information should be collected when a node close to the central point of a cell cannot be located.

\[
Q_1 = \left( \frac{-W}{4}, 0 \right), \quad Q_3 = \left( \frac{W}{4}, 0 \right) \quad (2)
\]

Thus, calculation of the central region of a cell as the average value of the interquartile range proceeds as follows:

\[
O = \frac{(Q_3 + Q_1)}{2} \quad (3)
\]

The central point of a cell \( O \) is replaced with interquartile range \( Q_1 \) and \( Q_3 \) as the point from which information should be collected when a node close to \( O \) cannot be found. This increases the probability that the entire area within a cell will be covered by the node that is selected. This approach reduces the influence of inaccuracy induced by the uncovered area; however, the degree to which the area is divided is proportional to the amount of data that must be sent to determine the traffic conditions. This is due to the fact that the use of interquartile range decreases the progress toward the destination. Thus, there is a tradeoff between costs in estimation error and costs in data transmission. The accuracy of this method relies heavily on the location and number of nodes located within a cell. Therefore, this paper proposes a solution in which coverage area is used to reduce error in estimation.

The Proposed Solution

### A. Assumptions

In the following, we consider an urban environment comprising roads and junctions in which a set of nodes is moving in the environment. We assume that each node in the network has a unique ID and is aware of its own physical location based on positioning devices, such as a GPS. Nodes are equipped with digital road maps, which include details related to the location of roads and junctions, to determine where nodes are located. We also assume that each node is equipped with a short range wireless transceiver with an identical transmission radius, allowing it to communicate with nearby neighbors. The coverage area of each node is determined by its transmission radius and the width of the road is far smaller than the transmission range of a node. It is also assumed that two nodes can communicate with each other directly only as long as they are within each other’s transmission range. Each node is required to maintain a neighbor table and to inform nodes in its vicinity of its current position, direction, and velocity through the transmission of periodic beacon messages. The node that relays probe packets is referred to as the relay node.

### B. Coverage Based Density Estimation

In this work, we propose a density estimation strategy based on coverage area to address the problems described in the previous section. This strategy can adapt to changes in network topology caused by the mobility of nodes, which
makes it suitable for any mobile environment. Instead of dividing road segment between two junctions into multiple fixed size cells, it uses the coverage area of each node to identify the point that should be selected as the point from which traffic information is collected. The maximum area covered by a relay node depends on the distance between the source and the relay node. Calculating the difference in coverage area makes it possible to decide which node is the better relay node candidate. Thus, estimating the traffic density of a road segment involves sending a periodical probe packet, referred to as Information Collection (IC), from the node at the source junction to the node at the destination junction to obtain traffic information to estimate traffic density for that road segment as illustrated in Figure 4.

Let us consider a scenario in which an application seeks information related to traffic density along the segment of a road between the source and destination junction. The node that is located at the source junction initiates a search for all nodes within its transmission coverage area. The node farthest from the source that provides the most coverage among its neighbors toward destination junction is selected as the relay node for the forwarding of IC packet. The periodic beacon signals sent by neighboring nodes help to establish mutual awareness among the nodes and maintains up-to-date information about neighboring nodes. Thus, the node selected as relay node is able to observe the number of nodes in its vicinity for use in the independent estimation of density for a given road region. The IC packet is delivered to the node with maximum coverage toward the destination junction as the point where the information should be collected to acquire traffic density along the given road segment until it reaches its designated junction. The IC packet is updated by each relay node and a node from the designated junction duplicates the IC packet and sending it back toward the source junction in a greedy manner. In so doing, each relay node should cover most, if not all, of the area along the given road segment between the source and destination junction.

Let us consider two nodes, \( i \) and \( j \), located in positions \((x_i, y_i)\) and \((x_j, y_j)\) in which the maximum distance between \( i \) and \( j \) is denoted as \( d(i, j) \) as shown in Figure 5, where \( d(i, j) \) is defined as follows:

\[
d(i, j) = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}
\]

(4)

The radio communication range of each node is \( r \); if \( d(i, j) > 2r \), then \( j \) is not considered in the relay node selection process. The width of the road is \( R_w \) and the length from \( j \) to the top boundary of the given road segment is \( B_{top(j)} \). To calculate the coverage area of \( j \), we need to remove the intersection area of \( i \) and \( j \) and the area beyond the road boundary.

Knowledge of the circumference of the transmission coverage of node \( i \) makes it possible to derive the spatial intersection of the covered areas denoted as \( i_{int} \), between two nodes located a distance of \( d(i, j) \) apart, can be derived as follows:

\[
i_{int} = 2\int_{d(i,j)/2}^{r} \sqrt{r^2 - x^2} \, dx + 2\int_{d(i,j)/2}^{r} \sqrt{r^2 - x^2} \, dx
\]

(5)

To calculate the area of coverage that is intersected with node \( j \) beyond the top road boundary denoted as \( J_{top} \), we assume \( B_{top} \) is the \( y \)-axis, the center of \( i \) is \((x_{i, top}, y_{i, top})\) and the center of \( j \) is \((x_{j, top}, y_{j, top})\). By using their point of intersection, \( J_{top1}, J_{top2} \) and \( J_{top3} \) at coordinates \((x_{j, top1}, y_{j, top1}),(x_{j, top2}, y_{j, top2})\) and \((x_{j, top3}, y_{j, top3})\), this intersection area can be derived as follows:
\[ \hat{j}_{\text{top, int}} = \begin{cases} \int_{x_{j,\text{top}}}^{\infty} \left( \sqrt{r^2 - (x-x_{j,\text{top}})^2} + y_{j,\text{top}} \right) dx, & y_{j,\text{top}} > 0 \\ 0, & y_{j,\text{top}} \leq 0 \end{cases} \]

To calculate the area of coverage that is intersected with node \( j \) beyond the bottom road boundary denoted as \( \hat{j}_{\text{bot, int}} \), we assume \( B_{\text{bot}} \) is the \( x \)-axis, the center of \( i \) is \((x_{i,\text{bot}}, y_{i,\text{bot}})\) and the center of \( j \) is \((x_{j,\text{bot}}, y_{j,\text{bot}})\). By using their point of intersection, \( J_{\text{bot1}}, J_{\text{bot2}} \) and \( J_{\text{bot3}} \) at coordinate \((x_{j,\text{bot1}}, y_{j,\text{bot1}})\), \((x_{j,\text{bot2}}, y_{j,\text{bot2}})\) and \((x_{j,\text{bot3}}, y_{j,\text{bot3}})\), this intersection area can be derived as follows:

\[ \hat{j}_{\text{bot, int}} = \begin{cases} \int_{x_{j,\text{bot1}}}^{\infty} \left( \sqrt{r^2 - (x-x_{j,\text{bot}})^2} + y_{j,\text{bot}} \right) dx, & y_{j,\text{bot}} > 0 \\ 0, & y_{j,\text{bot}} \leq 0 \end{cases} \]

The area of coverage beyond the top road boundary is denoted as \( \hat{j}_{\text{top}} \), and given by:

\[ \hat{j}_{\text{top}} = \begin{cases} 2rB_{\text{top}}(j), & B_{\text{top}}(j) < r \\ 0, & B_{\text{top}}(j) \geq r \end{cases} \]

The area of coverage beyond the bottom road boundary is denoted as \( \hat{j}_{\text{bot}} \), and given by:

\[ \hat{j}_{\text{bot}} = \begin{cases} 2rR_{\text{bot}}(j), & R_{\text{bot}}(j) > 0 \\ 0, & R_{\text{bot}}(j) \leq 0 \end{cases} \]

The coverage of node \( j \) is denoted as \( \hat{j}_{\text{cov}} \) and can be derived as follows:

\[ \hat{j}_{\text{cov}} = \begin{cases} \pi r^2 - \hat{j}_{\text{int, bot}} - \hat{j}_{\text{bot, int}} - \hat{j}_{\text{top, int}} - \hat{j}_{\text{bot}}, & 0 < d(i,j) < 2r \\ 0, & d(i,j) \geq 2r \end{cases} \]

The node that intersects \( i \) forms set \( l \). From Equation (10), we can calculate the area of coverage formed by all nodes in set \( l \). We then select the node from set \( l \) that is farthest from \( i \) with the largest \( \hat{j}_{\text{cov}} \) to act as the relay node for the collection of traffic information. The node that is selected is responsible for estimating traffic density within its coverage area by consulting its neighbor table and calculating the number of nodes situated within its area of coverage. Thus, the node that is selected inserts traffic density data into the IC packet for propagation along the road segment to the next relay node ahead of the selected node. By doing so, the selected node provides additional coverage over what has already been covered by node \( i \). By repeating this process until the designated junction is reached, the proposed density estimation strategy ensures that most of the area is monitored and maximizes the likelihood that all nodes located along the road segment between the source and the destination junction are included in the estimation of traffic density. Once a node is selected as a node for relaying IC packets from set \( l \), the selected node in set \( l \) is omitted from the selection process.

**Simulation**

**A. Simulation Setting**

Simulations were conducted in an NS2 simulator and the mobility model of vehicles are created by using TraNS [21]. In the simulation, a topology containing 16 junctions and 24 bi-directional roads was spread over a simulation area of 1500m x 2000m obtained from the TIGER database [22]. The speed limit of each road is set at 80km/h. Each node was equipped with an 802.11 interface with rate of 0.25s. The total simulation time was 600s with a node transmission range of 250m. The source road segments used to estimate traffic density were randomly selected from across the road map. The number of nodes varied between 75 and 300 and the interval of Hello messages was 1s, thereby allowing the updating of position information every second. Various numbers of vehicles were used to obtain various traffic densities. The performance of the proposed strategy was compared with that of Interquartile Range and IFTIS [19] using various metrics, including estimation accuracy and overhead.

**B. Simulation Results**

A road segment between two road junctions was selected to demonstrate the proposed density estimation strategy in an urban environment. We began by comparing the density estimated over time with various node densities in an urban environment...
environment. Figure 6 presents the estimations obtained in a low-density environment. Density estimation using the proposed Coverage Based strategy was very close to the actual density value, thereby demonstrating the accuracy of the estimates in which the focus was on finding the farthest node with the best coverage in its neighborhood as the point from which traffic data should be collected. The Interquartile Range is very close to the actual traffic values; however, both Interquartile Range and IFTIS uses cell formation techniques for the estimation of traffic density, and is therefore prone to a drop in accuracy in cases where the cell density is low or when a node located closest to the cell center cannot be found. Consequently, some of the nodes are not included in density estimation. Furthermore, when traffic density is low, connections between nodes are not always available, which decreases the likelihood of finding a node located close to the center of the cell.

![Figure 6. Density Estimation - Low Density](image)

As shown in Figure 7, density estimation using the proposed Coverage Based strategy is very close to the actual density values. Using the Interquartile Range provides higher accuracy in this situation than in a low-density environment because an increase in traffic density provides more nodes to select from, which increases the probability that a node close to the cell center can be obtained. It also increases the chances of finding a reference point when the nodes in a cell deviate far from the center. The traffic density estimations of IFTIS vary considerably from the actual traffic values due to the fact that a portion of the cells is left uncovered when a node cannot be located near the center of the cell. Consequently, the node located at the uncovered area are left out in the density estimation process, which ultimately decreases the accuracy of the estimation results.

![Figure 7. Density Estimation - High Density](image)

Table 1 shows the overhead of different methods under different traffic scenarios. Overhead here is defined as the number of beacons sent from the source to the destination junction. The proposed Coverage Based strategy has the least overhead on the network followed by Interquartile Range and IFTIS.

<table>
<thead>
<tr>
<th>Table 1. Overhead</th>
</tr>
</thead>
<tbody>
<tr>
<td>Method</td>
</tr>
<tr>
<td>IFTIS</td>
</tr>
<tr>
<td>Interquartile Range</td>
</tr>
<tr>
<td>Coverage Based</td>
</tr>
</tbody>
</table>

Conclusions

In this paper, we propose a fully distributed density estimation strategy capable of providing traffic density estimations in real-time using the coverage area provided by each node to estimate traffic density for a road segment between two junctions. This approach makes it possible to select the appreciate relay node for the collection of traffic information and decreases the likelihood of nodes being excluded from the process of traffic density estimation. Simulation results demonstrate that the proposed density estimation strategy is more accurate than existing schemes under a variety of scenarios. These findings demonstrate the effectiveness of using coverage area to determine the point from which traffic information should be collected in reducing estimation error.

Acknowledgments

The authors would like to thank the Ministry of Science and Technology for supporting this research, which is part of the project numbered 101-2221-E-006-259-MY3 and 104-2221-E-006-049.
References


Biographies

CHUN-CHIH LO received his M.S. degree in Computer Science from Tunghai University, Taichung, Taiwan in 2007. He is currently working toward the Ph.D. degree in Computer Science and Information Engineering at National Cheng Kung University. His research interests revolve around wireless communication, with particular attention to wireless networking problems presented by vehicular networking.

YAU-HWANG KUO received the Ph.D. degree in computer engineering from National Cheng Kung University (NCKU), Tainan, Taiwan in 1988. He is a Distinguished Professor with the Department of Computer Science and Information Engineering, NCKU. In his career, he has been consistently active in the fields of academia, education accreditation, and government policy planning. He was a Deputy Executive Secretary of the Science and Technology Advisory Group in the Executive Yuan (Cabinet), the
Director of the Computer Center in the Ministry of Education. His research interests include wireless broadband communication, cloud computing, computational intelligence, intelligent information analytics, and context-aware computing.