

GTE: Um Sistema para Gerenciamento de Trânsito Escalável baseado em Compartilhamento Oportunista

Allan M. de Souza¹, Leonardo C. Botega², Leandro A. Villas¹

¹Instituto de Computação – UNICAMP

allanms@lrc.ic.unicamp.br, leandro@ic.unicamp.br

²Centro Universitário Eurípides de Marília – UNIVEM

botega@univem.edu.br

Abstract. *Urban mobility has become one of the most challenging issue in large urban centers. As a consequence, traffic congestion has become a daily problem. Several Traffic Management Systems (TMS) have been proposed to improve overall traffic efficiency. However, the proposals exchange inefficiently traffic information, which can lead to network overload. In order to overcome the mobility problem and improve the efficiency in dealing with vehicle traffic, this work introduces a scalable traffic management system based on opportunistic content sharing, named GTE. Simulation results indicate that GTE outperforms the assessed solutions in different scenarios and in different key requirements of TMS.*

Resumo. *Mobilidade em grandes centros urbanos tornou-se um grande desafio, como consequência dos congestionamentos recorrentes. Dessa forma, vários Traffic Management Systems (TMS) têm sido propostos para melhorar a eficiência do tráfego. Entretanto, muitas das soluções propostas não são escaláveis, utilizando-se de métodos limitados para troca de informações, o que pode sobrecarregar a rede. Além disso, outras soluções não são capazes de fornecer um conhecimento preciso sobre as condições de tráfego. Dessa forma, para minimizar o problema de mobilidade e melhorar a eficiência do tráfego, introduzimos GTE - Gerenciamento de Tráfego Escalável, um TMS totalmente distribuído e escalável baseado em compartilhamento oportunista. Os resultados mostram que o GTE supera as soluções avaliadas em diferentes requisitos essenciais de um TMS.*

1. Introdução

Congestionamento tornou-se um problema diário em grandes centros urbanos, resultando-se do grande crescimento populacional juntamente com o aumento do número de veículos, excedendo a capacidade da infraestrutura de transporte [de Souza et al. 2016, Djahel et al. 2015]. Diante desse cenário, outros problemas podem emergir, incluindo degradação da qualidade de vida, impactos negativos na economia regional e danos ambientais [Karagiannis et al. 2011, de Souza et al. 2016].

A fim de maximizar a eficiência da infraestrutura de transporte disponível, minimizando o congestionamento e seus danos, Sistemas de Gerenciamento de Trânsito (*TMS – Traffic Management System*) foram propostos. TMSs concentram-se na integração de tecnologias de sensoriamento, comunicação e processamento de dados para coletar e explorar dados relacionados ao trânsito, fornecidos por veículos, sensores, subsistemas e até pessoas, com o objetivo de identificar eventos que podem degradar a eficiência do trânsito e prover serviços para minimizar os efeitos desses eventos [Djahel et al. 2015].

A base para a implantação de um TMS eficiente é a Rede Veicular, a qual é uma rede *ad hoc* móvel composta por veículos equipados com sensores, unidades de processamento e interfaces de comunicação sem fio [Karagiannis et al. 2011]. Devido a essas características,

os veículos podem comunicar entre si criando uma rede *ad hoc*, ou eles podem se comunicar com *Roadside Units* (RSUs), as quais são entidades com capacidade de processamento, sensoriamento e comunicação localizadas próximo à infraestrutura de transporte e são utilizadas para maximizar a capacidade da rede e seu gerenciamento, bem como prover acesso a Internet. Em suma, as Redes Veiculares permitem a troca de informações entre os veículos e infraestrutura, e também fornecem uma plataforma para sensoriamento e atuação para os TMSs [Djahel et al. 2015].

Diferentes TMSs foram propostos para lidar com as necessidades de mobilidade urbana e melhorar a eficiência do tráfego [Araujo et al. 2014, de Souza et al. 2015, Wang et al. 2015, Pan et al. 2016, Doolan and Muntean 2016]. Algumas soluções [de Souza et al. 2015, Wang et al. 2015] utilizam uma abordagem centralizada para detectar iminência de um congestionamento como também para sugerir rotas alternativas. Entretanto, essas soluções não são escaláveis e podem apresentar um *overhead* indesejado, especialmente em cenários de alta densidade de veículos, como por exemplo as altas densidades de veículos durante horários de pico [de Souza et al. 2015, Wang et al. 2015, Pan et al. 2016, Doolan and Muntean 2016]. Por outro lado, outras soluções como [Araujo et al. 2014] utilizam uma abordagem distribuída, porém, não possuem conhecimento da real condição de tráfego corrente, o que pode incorrer em orientações de rotas errôneas.

Dessa forma, devido às limitações apresentadas pelos TMSs atuais, tais como, falta de escalabilidade, *overhead* introduzido em altas densidades, e a falta de conhecimento sobre as reais condições de tráfego, torna-se desejável uma abordagem escalável, a qual lida com esses problemas. Entretanto, um desafio chave nesta abordagem é como fornecer um conhecimento preciso sobre a real condição de tráfego corrente para todos os veículos em circulação, um vez que soluções simples podem potencialmente sobrecarregar a rede e introduzirem um *overhead* indesejado.

Visando superar o desafio supracitado, propomos GTE (Gerenciamento de Tráfego Escalável), um TMS distribuído e escalável projetado para maximizar a eficiência do tráfego sem sobrecarregar a rede. GTE concentra-se em fornecer um conhecimento preciso sobre as condições de tráfego em áreas críticas (áreas com congestionamentos recorrentes) em vez de fornecer conhecimento geral sobre as condições de tráfego. Além disso, GTE utiliza uma abordagem de *Conteúdo Flutuante*, a qual atua como um serviço baseado em localização para disseminar informações geo-localizadas ao invés de disseminar a informação para toda a rede [Castro et al. 2008, Hyytiä et al. 2011]. De forma geral, um conteúdo flutuante é associado a uma área geográfica, onde o mesmo é responsável por se manter disponível para os veículos localizados nesta área, bem como para os veículos que irão adentrar na mesma, utilizando um compartilhamento de informação oportunista. Portanto, ao aproximar-se de uma área crítica ou adentrá-la, os veículos receberão informações precisas sobre a condição de tráfego corrente, permitindo-lhes detectar vias congestionadas e calcular rotas alternativas para evitá-las.

Neste trabalho, é apresentada uma série extensiva de experimentos que mostram a necessidade de um TMS escalável para maximizar a eficiência do trânsito. Adicionalmente, resultados dos experimentos com GTE indiciam um desempenho superior quando comparado a soluções existente.

O restante do trabalho encontra-se organizado da seguinte forma. A Seção 2 fornece uma visão geral dos TMS existentes e apresenta alguns trabalhos relacionados. A Seção 3 descreve a solução proposta. A análise de desempenho da solução proposta é apresentada na Seção 4. Por fim, a Seção 5 apresenta as conclusões e indica trabalhos futuros.

2. Trabalhos Relacionados

Os TMSs atuais podem ser amplamente classificados em (i) Centralizados; (ii) Distribuídos; e (iii) Híbridos. TMSs centralizados têm o propósito de fornecer um melhor gerenciamento do tráfego, utilizando uma abordagem centralizada para detectar congestionamentos e sugerir rotas alternativas [Wang et al. 2015, de Souza et al. 2015]. Para isso, estas soluções baseiam-se em informações reportadas por veículos periodicamente como posição, velocidade média, destino, etc. TMSs distribuídos [Araujo et al. 2014] utilizam uma abordagem cooperativa para detectar congestionamentos e evitá-los baseado em informações compartilhadas entre os veículos. TMSs híbridos [Doolan and Muntean 2016, Pan et al. 2016] utilizam uma abordagem centralizada para detectar congestionamentos e construir um conhecimento sobre a condição de tráfego, porém, em seguida utilizam uma abordagem distribuída para calcular rotas alternativas, onde o conhecimento é encaminhado para os veículos, para que os mesmos possam calcular novas rotas.

Contudo, TMSs centralizados não são escaláveis, pois podem sobrecarregar a rede em determinadas densidades de veículos e introduzir um *overhead* indesejado [Wang et al. 2015, de Souza et al. 2015, Doolan and Muntean 2016, Pan et al. 2016]. TMSs híbridos, por utilizarem uma entidade central para agregar o conhecimento do tráfego, também pode não ser escalável, e assim, introduzir uma latência indesejada para o sistema se troca de informações não for feita de maneira eficiente. Por fim, TMSs distribuídos [Araujo et al. 2014] apesar de apresentarem uma alta escalabilidade, podem incorrer no mesmo problema. Adicionalmente, eles não possuem um conhecimento preciso da condição de tráfego corrente [Araujo et al. 2014].

Araújo et al. [Araujo et al. 2014], propuseram CARTIM, um TMS distribuído para minimizar congestionamentos. Os veículos mensuram os níveis de congestionamento em suas vias baseado informações compartilhadas por mensagens *beacons* enviadas por todos os veículos. Dessa forma, baseado nas informações recebidas e utilizando um sistema baseado em lógica fuzzy, cada veículo é capaz de classificar o nível de congestionamento em sua via. Posteriormente, os veículos compartilham suas classificações com seus vizinhos, para cooperativamente detectar um congestionamento. Além disso, quando uma via congestionada é detectada, os veículos calculam uma rota alternativa utilizando uma heurística que baseia-se na distância até o destino. Entretanto, essa heurística pode incorrer em uma má orientação de rota, uma vez que os veículos não possuem conhecimento sobre a condição de tráfego nas vias.

Doolan and Muntean [Doolan and Muntean 2016] introduziram EcoTrec, um TMS híbrido que baseia-se no roteamento periódico dos veículos. Por ser uma solução híbrida, EcoTrec utiliza um servidor central para coletar informações dos veículos e construir um conhecimento sobre a condição de tráfego do cenário todo, em seguida, esse conhecimento é disseminado para os veículos para que eles possam calcular uma rota baseada no menor consumo de combustível. Para enviar informações para o servidor, EcoTrec utiliza um protocolo de disseminação, o qual não implementa nenhum mecanismo de supressão de *broadcast*, onde, por sua vez pode introduzir um *overhead* para o sistema em cenários densos.

Em nosso trabalho anterior [de Souza et al. 2015], propomos CO-OP, um TMS centralizado que classifica a condição de tráfego no cenário todo baseado no algoritmo *K-Nearest Neighbor* (KNN). Igualmente o EcoTrec, no CO-OP todos os veículos precisam enviar suas informações para um servidor central, para que o mesmo possa detectar congestionamentos e sugerir rotas alternativas. Quando um congestionamento é detectado, CO-OP emprega um roteamento cooperativo, o qual sugere rotas alternativas para os veículos que vão passar pelo

congestionamento detectado. Este roteamento cooperativo aplica um balanceamento de carga para distribuir melhor o fluxo de veículos em todo o cenário. Entretanto, como EcoTrec, CO-OP por introduzir um *overhead* indesejado em altas densidades de veículos.

Os trabalhos supradescritos apresentam algumas limitações como baixa escalabilidade, dependência de uma RSU ou servidor central, sobrecarga da rede [de Souza et al. 2015, Doolan and Muntean 2016, Wang et al. 2015, Pan et al. 2016], e a ausência de conhecimento sobre a real condição de tráfego para uma melhor orientação de rotas [Araujo et al. 2014]. Neste contexto, para superar tais limitações apresentadas, propomos um TMS distribuído baseado em compartilhamento oportunista, o qual não sobrecarrega a rede e também não introduz uma latência indesejada. Além disso, o TMS proposto fornece um conhecimento preciso sobre a condição de tráfego.

3. Gerenciamento de Trânsito Escalável

GTE é um TMS baseado em conteúdo flutuante, proposto para melhorar a eficiência do tráfego. Este emprega uma abordagem de compartilhamento oportunista para fornecer e armazenar informações sobre as reais condições de tráfego em regiões críticas. Para isso, GTE não depende de nenhuma RSU ou entidade central, conseqüentemente os veículos são aptos a detectar vias congestionadas e calcular uma rota mais eficiente.

Definição: *Considere um ambiente veicular representado por um grafo simples e ponderado $G = (V(G), E(G))$, onde o conjunto $V = \{v_1, v_2, \dots, v_i\}$ representa o conjunto de interseções do cenário (vértices), o conjunto $E = \{e_{01}, e_{12}, \dots, e_{ij}\}$ representa o conjunto de ruas do cenário (arestas) as quais conectam as interseções $E(G) \subseteq V(G) \times V(G)$, onde a aresta e_{ij} é definida por um par de vértices subsequentes $(v_i, v_j) \in V(G)$. Além disso, $W = \{w_{01}, w_{12}, \dots, w_{ij}\}$ é o conjunto de pesos que representa a condição de tráfego em cada rua do cenário, na qual $w : E \rightarrow \mathbb{R}_+^*$. Seja $N = \{n_1, n_2, \dots, n_i\}$ o conjunto de veículos em circulação e $R = \langle e_{ij}, \dots, e_{mn} \rangle \mid R \subseteq E(G)$ uma rota $\forall n \in N$.*

Um desafio chave em prover um TMS distribuído, é como construir e entregar para todos os veículos um conhecimento preciso sobre a condição de tráfego do cenário, sem sobrecarregar a rede e conseqüentemente produzindo um baixo *overhead*. Pois, as informações geradas por cada veículo precisam ser entregues para todos os outros veículos na rede, portanto, caso este processo não for feito de forma eficiente, muitas transmissões em um breve intervalo podem sobrecarregar a rede, produzindo um *overhead* indesejado e degradar a eficiência do sistema [de Souza et al. 2016]. Conseqüentemente, os veículos não terão um conhecimento preciso sobre a condição de tráfego, incapacitando-os de detectar vias congestionadas. Sendo assim, para lidar com esse problema, GTE constrói um conhecimento preciso sobre as reais condições de tráfego para em áreas críticas, ou seja, áreas conhecidas por possuir congestionamentos recorrentes ou alto número de incidentes de tráfego. Além disso, GTE usa uma abordagem de compartilhamento oportunista para prover esse conhecimento como um conteúdo flutuante por toda área crítica. Portanto, sempre que um veículo entrar em uma área crítica, o mesmo receberá o conteúdo flutuante, permitindo-o detectar vias congestionadas em sua rota corrente.

Áreas críticas possuem uma localização (x_c, y_c) e uma área $C = (h \times l)$, os quais definem seu centro (ponto crítico) e a área potencialmente afetada por incidentes dentro da mesma. Para cada conteúdo flutuante de cada área crítica, GTE atribui uma *Área de Ancoragem* (AA) definida por um raio $r_a \in \mathbb{R}_+^* \mid r_a \geq \frac{h}{2} \geq \frac{l}{2}$, o qual é baseado no ponto crítico (x_c, y_c) . Cada AA é responsável por manter o conteúdo flutuante dentro dela, tornando-o disponível para os veículos que estão localizados dentro da mesma e para os outros veículos que irão passar por ela. Essas AAs podem ter tamanhos diferentes e podem ser utilizadas para

diferentes propósitos [Hyytiä et al. 2011]. Além disso, elas podem ter locais conhecidos ou dinâmicos, ou seja, elas podem surgir assim que um incidente de trânsito ocorre, a fim de fornecer conhecimento sobre o incidente para os veículos que estão próximos do incidente, assim como para os veículos que irão passar por essa área.

A Figura 1 apresenta as AAs para dois conteúdos flutuantes. Particularmente, AAs (AA_1 and AA_2) possuem tamanhos diferentes e seus respectivos conteúdos armazenados. Os veículos n_1, n_2 e n_3 possuem o conteúdo flutuante fornecido pela AA_1 , enquanto os veículos n_4, n_5, n_6 e n_7 possuem o conteúdo flutuante fornecido pela AA_2 . É importante notar que, se um veículo estiver em local onde tem interseção de duas ou mais AAs, ele potencialmente pode possuir mais de um conteúdo flutuante, neste caso o veículo n_8 possui os conteúdos flutuante fornecidos pelas duas AAs. Apesar das áreas críticas conhecidas

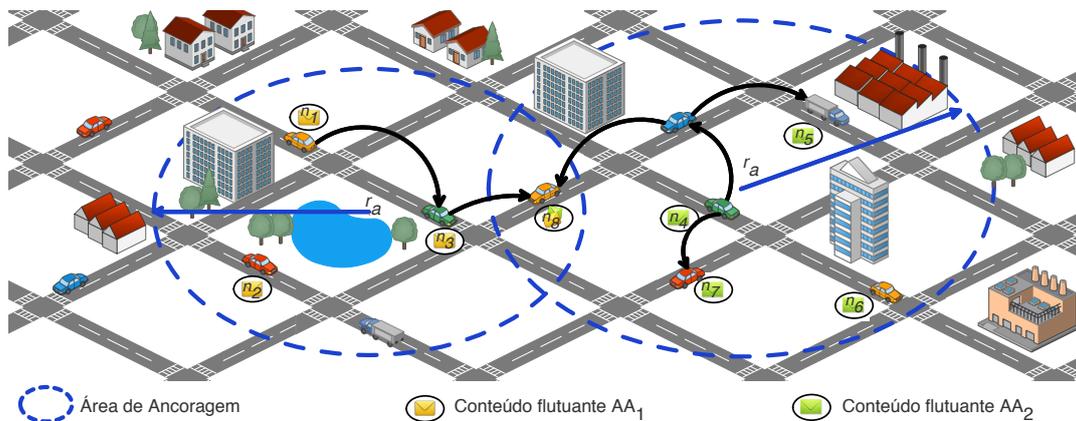


Figura 1. Conteúdos flutuantes e suas respectivas áreas de ancoragem.

pele GTE, o mesmo não possui nenhum conhecimento sobre as condições de tráfego dentro delas. Neste caso, GTE emprega um mecanismo eficiente para suprir essa necessidade, construindo um conhecimento preciso sobre as condições de tráfego dentro dessas áreas, os quais, em seguida são disponibilizados como conteúdo flutuante, garantindo um compartilhamento oportunista e a sobrevivência do mesmo nesta área. Por fim, ele também fornece um mecanismo para cada veículo detectar vias congestionadas, permitindo-os a ter uma melhor orientação de rota. Sendo assim, a Subseção 3.1 descreve como este conteúdo flutuante é criado, Subseção 3.2 descreve como GTE garante a disponibilidade e sobrevivência de cada conteúdo flutuante, e Subseção 3.3 descreve como os veículos detectam vias congestionadas e computam rotas alternativas para si mesmos para evitar vias congestionadas.

3.1. Criação do conteúdo flutuante

GTE baseia-se em informações periódicas compartilhadas por todos os veículos através de mensagens *beacons*, as quais possuem a posição atual do veículo $e_{ij} \in R$, velocidade média s_m , direção e um identificador para informar se o veículo armazena algum conteúdo flutuante e se o mesmo está atualizado.

Dado uma área crítica, GTE define um conjunto de *veículos produtores* $N' \subseteq N$, os quais estão localizados dentro da área crítica e são responsáveis por construir o conhecimento sobre a condição de tráfego naquela área. Entretanto, existem alguns desafios a serem considerados: (i) como garantir a cobertura total de uma área crítica, uma vez que o raio de comunicação pode ser muito menor do que a área crítica, e (ii) como construir o conhecimento sem sobrecarregar a rede.

Neste caso, GTE segmenta toda área crítica em k regiões menores chamadas de *subáreas*. para primeiro construir um conhecimento sobre o tráfego de cada subárea, e em seguida agregar o conhecimento de cada subárea em um único conhecimento, o qual representa a condição de tráfego de toda área crítica. A quantidade de subáreas é definida baseada no tamanho do raio de comunicação dos veículos [de Souza and Villas 2016].

Antes de construir o conhecimento de cada subárea, primeiramente cada veículo produtor $n_i \in N'$ cria um conhecimento local sobre o conhecimento do tráfego agregando as mensagens de *beacons* de seus vizinhos. Este conhecimento local é representado pelo conjunto $E^{n_i} = \{e_{ij}, \dots, e_{mn}\} \mid E^{n_i} \subset E(G)$, o qual é composto pelas vias dentro do raio de comunicação do veículo produtor e suas respectivas velocidades médias. Após criar seu conhecimento local E^{n_i} , cada veículo produtor compartilha-o com seus vizinhos localizados na mesma subárea, a fim de criar um conhecimento sobre toda a subárea, a função de agregação é dada por:

$$S_i = \frac{1}{n} \sum_{n_i \in N'} e_{ij}, \forall e_{ij} \in E^{n_i} \quad (1)$$

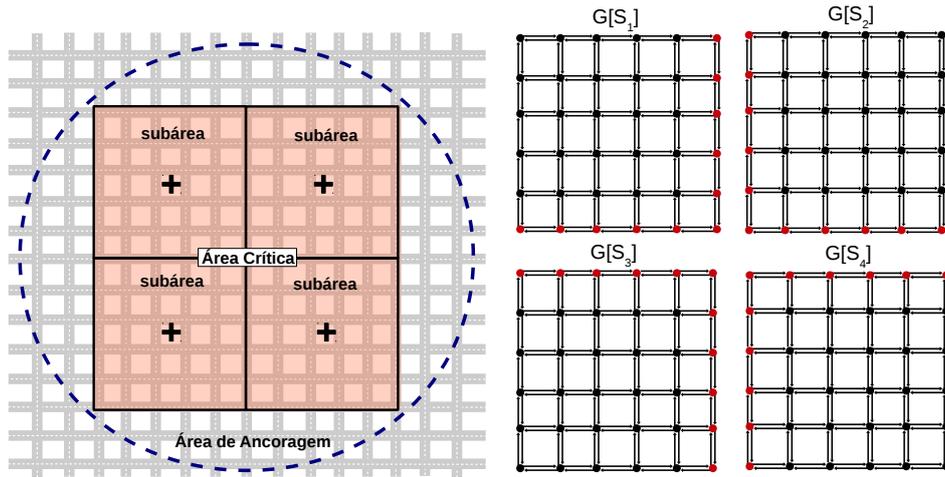
onde, S_i com $i \in \{1, 2, \dots, k\}$ representa o conhecimento sobre a condição de tráfego de cada subárea formado pela agregação do conhecimento E^{n_i} recebido de todos os veículos produtores dentro dela. Além disso, baseado na velocidade média de cada via, cada veículo atribui um peso para a mesma utilizando a seguinte fórmula:

$$w_{ij} = 1 - \max(e_{ij})^{-1} \times e_{ij} \quad \forall e_{ij} \in E^{n_i} \quad (2)$$

onde $\max(e_{ij})$ e e_{ij} representam a velocidade máxima e média de cada via.

Posteriormente, todo veículo localizado na mesma subárea cria um subgrafo $G[S_i] \subset G$ induzido pelas arestas com o conhecimento agregado S_i , onde o conjunto $E(G[S_i]) \subset E(G)$ representa o conjunto de vias pertencentes a cada subárea, enquanto o conjunto $V(G[S_i]) \subset V(G)$ representa o conjunto de interseções que conectam as vias da subárea $e_{ij} \in E(G[S_i])$. Portanto, para criar o conteúdo flutuante com o conhecimento sobre o tráfego de toda área crítica, é necessário integrar cada subgrafo $G[S_i]$ induzido pelas arestas de cada subárea em um único supergrafo G' . Sendo assim, cada subgrafo de cada subárea precisa ser encaminhado para todas as outras subáreas. Para isso, GTE emprega um protocolo de disseminação de dados entre as subáreas baseado em [de Souza et al. 2016], tal protocolo utiliza um mecanismo de supressão de *broadcast* e outro para minimizar a quantidade de colisões de pacotes evitando a resincronização introduzida pelo padrão IEEE 802.11p utilizado em Redes Veiculares.

Graças à segmentação da área crítica juntamente com o protocolo de disseminação de dados para agregar o conhecimento de todas as subáreas, GTE maximiza a cobertura reduzindo o número de transmissões, e ainda, garante pelo menos uma transmissão dentro de cada subárea. Para isso, GTE seleciona um veículo de cada subárea para iniciar o processo de disseminação, este veículo é selecionado baseado em sua posição corrente e também com base em seu último conhecimento gerado, para saber se o mesmo encontra-se atualizado. Portanto, todos os veículos na mesma subárea agendam uma transmissão de uma mensagem que contém o conhecimento agregado da subárea com um identificador da subárea em questão, cada transmissão é agendada baseada na distância euclidiana em que os veículos encontram-se do centro de sua subárea, assim, permitindo que veículos mais próximos do centro transmitam primeiro, e devido ao mecanismo de supressão de *broadcast* implementado, quando veículos que haviam agendado a transmissão da mesma mensagem recebem uma mensagem com o mesmo identificador, os mesmos cancelam a transmissão, evitando a transmissão de mensa-



(a) Áreas de ancoragem, área crítica, k -subáreas e seus respectivos centros (b) Subgrafo induzido pelas arestas de cada subárea.

Figura 2. Área crítica, subáreas, AA e subgrafo induzido pelas arestas da subárea.

gens redundantes. Note que, como todas as subáreas são distintas, as transmissões de uma subárea não cancelam transmissões de outra subárea.

Entretanto, se tal segmentação não fosse utilizada, essa mesma abordagem de disseminação geraria lacunas no conhecimento geral da área crítica, uma vez que a transmissão de um determinado veículo poderia potencialmente cancelar transmissões de outros veículos com informações diferentes, onde essas informações seriam importantes para fornecer um conhecimento preciso sobre a condição de tráfego em toda área crítica [de Souza et al. 2016]. Além disso, o processo de encaminhamento empregado pelo GTE não sobrecarrega a rede [de Souza et al. 2016].

Depois de receber todos os conhecimentos (subgrafos $G[S_i]$) de cada subárea, GTE cria o conteúdo flutuante o qual contém um subgrafo $G' \subseteq G$ dado pela Equação 3, a qual representa toda a área crítica, contendo o conhecimento preciso sobre a condição de tráfego, um raio de ancoragem r_a para determinar sua geolocalização, uma variável temporal para informar quando o conteúdo foi criado e um identificador para identificar a qual área crítica o mesmo pertence.

$$G' = \bigcup_{i=0}^k G[S_i] \quad (3)$$

Por questão de clareza, a Figura 2 descreve cada subárea de uma área crítica com o subgrafo G' criado pelo GTE. Em particular, a Figura 2(a) apresenta a segmentação implementada pelo GTE apresentando a área de ancoragem, área crítica, k -subáreas e seus respectivos centros. Adicionalmente, a Figura 2(b) apresenta cada subgrafo $G[S_i]$ induzido pelas arestas de cada subárea. Os pontos vermelhos indicam os vértices em comum em dois ou mais subgrafos, estes vértices são utilizados para fazer a união dos subgrafos $G[S_i]$ em um supergrafo G' . É importante salientar que este processo de união dos subgrafos resulta em um grafo simples, ou seja, as arestas múltiplas geradas no processo de união são removidas.

3.2. Compartilhamento oportunista

O compartilhamento oportunista é essencial para garantir a sobrevivência de um determinado conteúdo flutuante em sua AA. Ele mantém cada conteúdo flutuante em sua AA, tornando-o

disponível para os veículos dentro da área crítica, e também para os veículos que irão chegar nesta área. Além da AA, GTE implementa uma outra área, nomeada *Área de Encaminhamento* (AE), a qual é definida por $r_f \in R_+^* \mid r_f > r_a$ e é responsável por enviar de volta o conteúdo flutuante armazenado em veículos que estão deixando a AA. Em outras palavras, veículos que saíram da AA mas ainda possuem o conteúdo flutuante, precisam procurar ativamente por veículos que estão indo em direção da AA para encaminhar o conteúdo flutuante.

A fim de evitar transmissões desnecessárias, quando um veículo que armazena um conteúdo flutuante deixa sua AE, ele descarta o conteúdo para evitar o encaminhamento para veículos que estão muito distantes, uma vez que eles não possuem interesse naquele conteúdo neste momento, e a condição de tráfego de tal área crítica possivelmente pode mudar até o veículo chegar nessa área. Além disso, assim que os veículos se aproximaram de uma área crítica, os mesmos irão receber um conteúdo flutuante pertencente a esta área com um conhecimento atualizado sobre a condição de tráfego.

O processo de encaminhamento garante a sobrevivência do conteúdo em sua AA e o mesmo é baseado no interesse que o veículo possui em tal conteúdo. Resumidamente, o interesse depende se o veículo vai passar por uma área crítica, e se o mesmo possui um conhecimento atualizado sobre o local. Estas informações são obtidas através das mensagens de *beacons* periodicamente compartilhadas por todos os veículos. Portanto, quando um veículo que armazena um conteúdo flutuante atualizado recebe uma mensagem de *beacon* de um outro veículo, e eles encontram-se dentro de uma das áreas AA ou AE, o mesmo irá encaminhar o conteúdo baseado em uma probabilidade p

$$p = \begin{cases} 1 & \text{se } \exists e_{ij} \in R \mid e_{ij} \in E(G') \\ \theta(R) & \text{Caso contrário} \end{cases} \quad (4)$$

onde, $\theta(R) \in [0, 1]$ é uma função decrescente que fornece a probabilidade de encaminhamento quando veículos não irão passar pela AA.

3.3. Detecção de vias congestionadas e orientação de rota

O conteúdo flutuante disponível, permite que veículos que o possuem detectem vias congestionadas e melhorem sua rota, calculando rotas alternativas para evitar as vias congestionadas detectadas. Cada veículo pode detectar vias congestionadas e classificar o nível de congestionamento baseado no peso de cada via w_{ij} , $\forall e_{ij} \in E(G')$. A classificação é baseada no *Level-Of-Service* (LOS) do *Highway Capacity Manual* (HCM)¹. Tal manual apresenta medições de qualidade utilizadas para descrever a porcentagem de condições operacionais dentro de um fluxo de veículos, definindo seis níveis de serviços diferentes, onde LOS A representa a ausência de congestionamento e LOS F representa um alto índice de congestionamento. A Tabela 1 apresenta a relação entre o peso das vias w_{ij} e a classificação do tráfego baseado no LOS do HCM.

Como o conteúdo flutuante apenas fornece conhecimento sobre a condição de tráfego nas áreas críticas, os veículos conseguem apenas detectar vias congestionadas e calcular rotas alternativas para vias dentro da área crítica. Portanto, cada veículo verifica apenas um subconjunto de vias de sua rota completa R contido no conhecimento fornecido pelo conteúdo flutuante. Seja $R' = \{e_{ij}, \dots, e_{xy}\} \mid R' \subseteq R \cap E(G')$ este subconjunto. Primeiramente, um veículo precisa verificar se sua rota corrente vai passar por alguma via congestionada, verificando se a via $e_{ij} \in R'$ possui um peso maior que um *threshold* ε , o qual indica a existência

¹<http://hcm.trb.org>

Tabela 1. Relação entre o peso da via e sua classificação

Peso da via	LOS	Classificação
(0, 0.15]	A	Tráfego livre
(0.15, 0.33]	B	Tráfego livre
(0.33, 0.50]	C	Congestionamento Leve
(0.50, 0.60]	D	Congestionamento Leve
(0.60, 0.70]	E	Congestionada
(0.70, 1.00]	F	Congestionada

de algum congestionamento. Esta verificação é dada pela função $\sigma(n)$

$$\sigma(n) = \begin{cases} 1 & \text{se } \exists e_{ij} \in R' \mid w_{ij} \geq \varepsilon \\ 0 & \text{Caso contrário} \end{cases} \quad (5)$$

onde, o valor 1 significa que um veículo possui a via e_{ij} em sua rota R' que possui peso w_{ij} maior que o *threshold* ε definido, e 0 caso contrário.

Sempre que um veículo vai passar por uma via congestionada, uma rota alternativa deve ser calculada. Portanto, o veículo calcula esta rota alternativa baseado na sua posição corrente e a última via $e_{ij} \in R'$ utilizando um algoritmo de k -menores caminhos com escolha probabilística (PkSP – *Probabilistic k Shortest Paths*) baseado no peso das vias. PkSP utiliza a distribuição de Boltzmann [Kirkpatrick et al. 1983] para selecionar uma rota entre um conjunto de k -rotas calculadas, reduzindo assim a possibilidade de criar um congestionamento em outra área em um futuro próximo. Sendo assim, é importante notar que a rota alternativa completa é dada por

$$R_{alternativa} = pksp(e_{ij}, e_{xy}) \cup R' \setminus R \quad (6)$$

onde, $pksp(e_{ij}, e_{xy})$ é a rota alternativa calculada pelo veículo de sua posição atual até a última via $e_{ij} \in R'$ que o mesmo possui conhecimento sobre o tráfego e $R' \setminus R$ é o restante da rota original.

4. Análise dos resultados

Esta seção descreve a análise de performance do GTE. A Subseção 4.1 introduz as ferramentas utilizadas na simulação. As Subseções 4.2 e 4.3 apresentam a avaliação de escalabilidade do sistema e a eficiência do tráfego respectivamente. A análise de escalabilidade foi feita baseada no custo da rede avaliando o mecanismo segmentação e agregação implementado pelo GTE juntamente com a abordagem de compartilhamento oportunista. Além disso, tanto para a análise de escalabilidade, quanto para a análise de eficiência do tráfego, comparamos GTE com outros TMSs existentes na literatura.

4.1. Metodologia

Para as simulações, utilizamos o simulador de rede OMNeT++ 4.3² e o *framework* SUMO (*Simulator for UrbanMOBility*)³, version 0.25.0, para gerenciar o cenário e a mobilidade dos veículos. Para a rede veicular, utilizamos o *framework* Veins 4.3⁴ o qual implementa o padrão IEEE 802.11p e atenuação de sinal utilizando obstáculos.

Para o cenário de simulação, utilizamos uma região de 25 km² da cidade de Colônia, Alemanha, obtido do projeto TAPASCologne⁵ o qual tem o objetivo de reproduzir as reais

²<http://omnetpp.org/>

³<http://http://www.dlr.de/>

⁴<http://veins.car2x.org/>

⁵<http://kolntrace.project.citi-lab.fr/>

condições de tráfego da cidade em questão. Além disso, para garantir a ocorrência de um congestionamento foi utilizado apenas o intervalo das 11:00 – 13:00, onde sabemos que há vários congestionamentos na área selecionada. A partir dessas informações, definimos uma área crítica de 9 km² como pode ser observado na Figura 4.1, onde é apresentado o cenário utilizado juntamente com a condição de tráfego durante o horário selecionado e as regiões de 25 km² e 9 km². Por fim, a Tabela 2 descreve os principais parâmetros utilizados em nossa simulação.

Tabela 2. Parâmetros	
Parâmetros	Valores
Potencia de transmissão	2.2 mW
Raio de comunicação	300 m
Bit rate	18 Mbit/s
Cenário size	25 km ²
Cenário	Colônia, Alemanha
# veículos	≈ 20.000
# Áreas críticas	1
C	9 km ²
r_a	2 km
r_f	2.5 km
ε	0.5

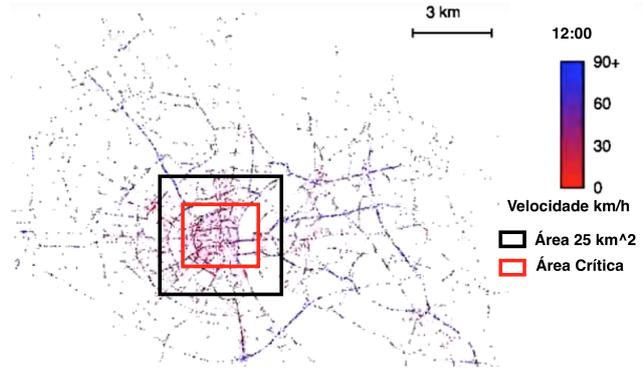


Figura 3. Colônia Alemanha.

4.2. Análise de escalabilidade

Com o foco de analisar a escalabilidade dos TMSs, realizamos simulações variando a densidade de veículos de 20% – 100% da densidade original da área de 25 km² utilizada do cenário. Sendo assim, as métricas de avaliação são: (i) Cobertura: porcentagem de mensagens entregue para os veículos dentro da área crítica; (ii) Mensagens transmitidas: total de mensagens transmitidas para realizar a entrega do conteúdo; (iii) Colisões de pacote: número total de colisões de pacotes durante a transmissão do conteúdo; (iv) Atraso: tempo médio gasto para realizar a entrega das mensagens.

A Figura 4 apresenta os resultados para as métricas avaliadas em função da porcentagem de veículos. Em particular, a Figura 4(a) apresenta os resultados de cobertura. O CO-OP, por utilizar uma abordagem centralizada não escalável, possui menor cobertura para todas as porcentagens de densidade, atingindo uma cobertura de 60% com 20% da densidade total e aproximadamente 40% com 100% da densidade total.

Tais resultados são consequência da sobrecarga da rede (veja Figuras 4(b) e 4(c)), pois além de cada veículo ter de enviar suas informações para um servidor central, para cada veículo que vai passar por um congestionamento detectado, o CO-OP deve sugerir uma rota alternativa, aumentando drasticamente o número de mensagens transmitidas (veja Figura 4(b)). Além disso, CO-OP não implementa nenhum mecanismo de agregação ou supressão de *broadcast* para reduzir o número de mensagens transmitidas, então essa grande quantidade de mensagens transmitidas pelos veículos e servidor geram muitas colisões (veja Figura 4(c)), o que consequentemente geram outras transmissões, sobrecarregando ainda mais a rede. Essa sobrecarga e não escalabilidade do CO-OP pode ser vista na Figura 4(d), pois é introduzido um alto *overhead*, onde o atraso médio de entrega de mensagens com 100% da densidade original é de aproximadamente 10 minutos.

De outra forma, EcoTrec apresenta uma abordagem híbrida, onde, apesar dos veículos terem que enviar suas informações para um servidor central como o CO-OP, depois do servidor ter criado o conhecimento sobre o tráfego, esse conhecimento é disseminado para os veículos, para que os mesmos possam calcular uma rota alternativa. Sendo assim, EcoTrec consegue reduzir o número de mensagens transmitidas e colisões de pacotes em relação ao

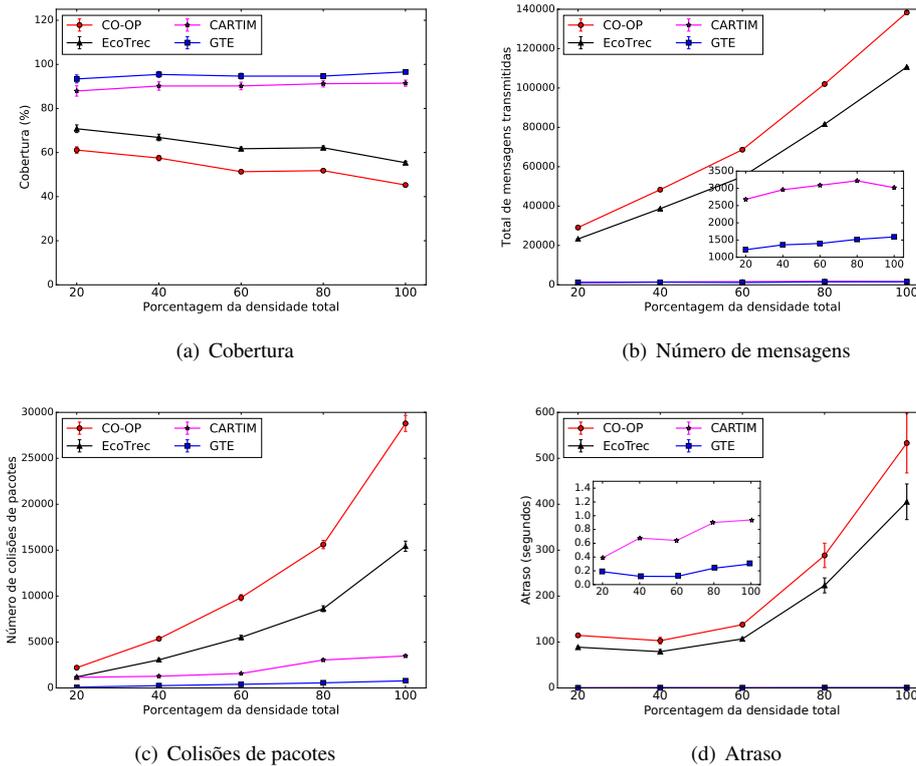


Figura 4. Resultados avaliação de escalabilidade.

CO-OP em aproximadamente 30% e 40% respectivamente (veja Figuras 4(b) e 4(c)), consequentemente, EcoTrec consegue um ligeiro aumento na cobertura, atingindo 75% com 20% da densidade total de veículos. Entretanto, assim como o CO-OP, o EcoTrec não implementa nenhum mecanismo para reduzir o número de mensagens transmitidas, sendo assim, o mesmo ainda introduz um alto *overhead* para o sistema, apresentando um atraso médio de aproximadamente 7 minutos com 100% da densidade original (veja Figura 4(d)).

Utilizando uma abordagem distribuída, CARTIM apresenta uma solução escalável, onde o mesmo atinge uma cobertura de aproximadamente 90% para todas as porcentagens de densidade. Além disso, como CARTIM não cria um conhecimento sobre a condição de tráfego, apenas detecta o nível de congestionamento da via em que o veículo está, o mesmo apresenta uma redução significativa na quantidade de mensagens transmitidas, reduzindo em até 95% e 92% quando comparado com CO-OP e EcoTrec, respectivamente. Entretanto, CARTIM não implementa nenhum mecanismo de agregação, sendo assim, muitos veículos na mesma via podem gerar mensagens redundantes sobre o congestionamento da mesma (veja Figuras 4(b) e 4(c)), adicionalmente, CARTIM não verifica o interesse dos veículos nas mensagens disseminadas (classificação do congestionamento), onde muitos veículos podem receber uma mensagem sem ter interesse na mesma. Contudo, como resultado de sua escalabilidade, CARTIM não introduz um *overhead* indesejado para o sistema, apresentando um atraso médio de menos de 1 segundo para todas as porcentagens de densidade (Figura 4(d)).

Por fim, os resultados do GTE mostram sua eficiência e escalabilidade. Gerando o conhecimento apenas das áreas críticas, as quais realmente degradam a eficiência do tráfego, juntamente com a abordagem de segmentação da mesma em áreas menores, assim como os mecanismos de agregação de conhecimento e de compartilhamento oportunista, o qual evita a transmissão de mensagens redundantes e desnecessárias, entregando o conhecimento apenas

para veículos que realmente possuem interesse no mesmo (veículos dentro da AA).

Adicionalmente, GTE reduz drasticamente o número de mensagens transmitidas, apresentando uma redução de mais de 50% quando comparado com CARTIM, 97% em relação ao EcoTrec e 98% em relação ao CO-OP. Consequentemente, GTE apresenta uma cobertura maior que 95% para todas as porcentagens da densidade da densidade original. Além disso, a eficiência do mecanismo de agregação implementado reduz a transmissão de mensagens redundantes, assim reduzindo também o número de colisões de pacotes (veja Figura 4(c)). Sendo assim, o GTE não produz um *overhead* indesejado para o sistema, apresentando um atraso médio de menos de 0,5 segundos para todas as porcentagens de densidade. Portanto, o GTE consegue construir um conhecimento preciso, transmitindo um menor número de mensagens, maximizando a cobertura e utilização da rede, sem introduzir um *overhead* indesejado.

É importante salientar a baixa escalabilidade apresentada pelas soluções CO-OP e EcoTrec, onde quanto maior a densidade de veículos, maior é o número de mensagens transmitidas e colisões de pacotes, os quais impactam diretamente no desempenho do sistema, assim minimizando a cobertura e introduzindo um *overhead* indesejado.

4.3. Análise de gerência do tráfego

Nesta seção, avaliamos a eficiência da gerência de tráfego do GTE comparando-o com três soluções CARTIM [Araujo et al. 2014], EcoTrec [Doolan and Muntean 2016], CO-OP [de Souza et al. 2015]. É importante salientar que TAPASCologne é o tráfego normal de veículos, sem aplicar nenhum mecanismo para gerenciar o tráfego. Sendo assim, as métricas avaliadas foram: (i) Tempo de viagem; (ii) Tempo de congestionamento; e (iii) Distância percorrida. Para obtermos o congestionamento real do cenário, todas as soluções foram simuladas com 100% da densidade total de veículos (para a região selecionada de 25 km²).

A Figura 5 apresenta os resultados para todas as métricas avaliadas. Em particular, a Figura 5(a) apresenta os resultados para o tempo de viagem, onde podemos ver que o tráfego original do cenário de TAPASCologne tem um tempo de médio de viagem aproximadamente 28 minutos, onde, em 68% desse tempo os veículos permanecem presos em congestionamentos (veja Figura 5(b)). Como na mobilidade original nenhuma alteração na rota é feita, TAPASCologne apresenta a menor distância percorrida, onde a mesma é de aproximadamente 5,5 km².

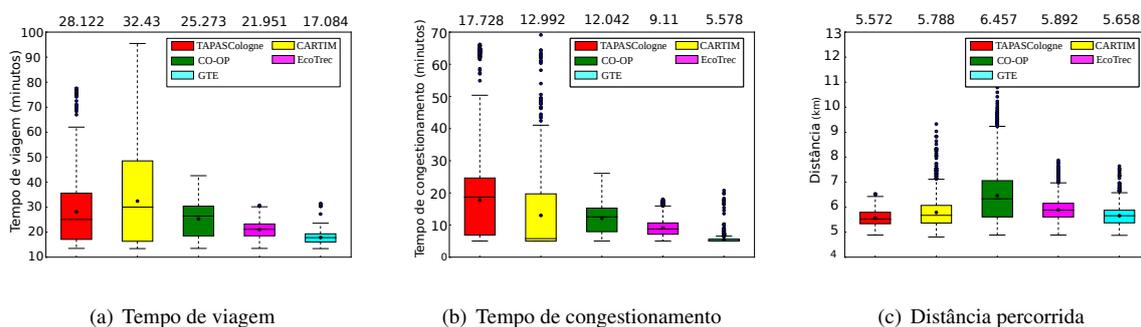


Figura 5. Resultados análise gerência de tráfego.

CARTIM, apesar de implementar uma solução distribuída, ele aumenta o tempo de viagem em aproximadamente 15% como consequência do seu mecanismo para evitar congestionamentos, o qual não possui conhecimento sobre as reais condições de tráfego. Consequentemente, CARTIM pode criar congestionamentos em outras áreas e também terá que evita-lás, assim, aumentando a distância percorrida (veja Figura 5(c)). Entretanto, CARTIM reduz o tempo de congestionamento em 25% quando comparado com o tráfego real (TAPASCologne).

Em contra partida, CO-OP emprega uma solução centralizada não escalável, portanto, devido ao *overhead* introduzido, o mesmo pode fazer detecções errôneas, como também muitos veículos podem não receber uma rota alternativa, como resultado da sobrecarga da rede. Dessa forma, CO-OP apresenta uma redução de apenas 15% em relação a mobilidade original (TAPASCologne). Entretanto, mesmo com o *overhead* introduzido, CO-OP atinge uma redução de 30% no tempo de congestionamento (veja Figura 5(b)). Utilizando uma abordagem diferente de CARTIM e CO-OP, EcoTrec implementa uma solução híbrida, a qual apresenta uma baixa escalabilidade devido a grande quantidade de mensagens transmitidas entre veículos e servidor central. Porém, diferente do CO-OP, os veículos são responsáveis por calcular suas rotas alternativas com base na condição de tráfego fornecida pelo servidor. Sendo assim, EcoTrec introduz um *overhead* menor que o CO-OP, portanto, o mesmo consegue uma melhor gerência do tráfego, reduzindo em aproximadamente 28% o tempo de viagem e em 45% o tempo de congestionamento, quando comparado com a mobilidade original (TAPASCologne). Para realizar essa melhor gerência do tráfego, tanto CO-OP quanto EcoTrec aumentam a distância percorrida pelos veículos, porém, devido ao sugestão de rota cooperativa implementada pelo CO-OP, o mesmo aumenta a distância percorrida em apenas 10%, enquanto EcoTrec apresenta um aumento de 20% (veja Figura 5(c)).

Por fim, GTE apresenta o menor tempo de viagem e tempo de congestionamento, apresentando uma redução de 36% e 68%, comparado com a mobilidade original, 44% e 57% quando comparado com CARTIM, 30% e 55% quando comparado com CO-OP e, 15% e 40% quando comparado com EcoTrec (veja Figura 5(b) e Figura 5(a)). Tais reduções são resultados da sua alta escalabilidade, a qual não introduz nenhum *overhead* para a solução, e também devido ao conhecimento preciso entregue para os veículos dentro de áreas críticas.

Graças à sua alta escalabilidade, e conseqüentemente seu baixo *overhead* GTE consegue prover o conhecimento da área crítica para os veículos potencialmente *on-the-fly*, assim, os veículos conseguem detectar congestionamentos e agir antes das outras soluções [de Souza et al. 2015, Doolan and Muntean 2016]. O CARTIM não introduz *overhead* para o sistema, porém o mesmo não tem conhecimento sobre a condição de tráfego e, portanto, os veículos acabam tendo uma má orientação de rota.

O baixo *overhead* é resultado de sua abordagem totalmente distribuída e dos mecanismos eficientes de agregação e encaminhamento de conhecimento, os quais conseguem prover um conhecimento preciso sobre a condição do tráfego para todos os veículos em uma área crítica, permitindo-os a calcular rotas alternativas para melhorar a eficiência do tráfego. O alto *overhead* introduzido pelas soluções CO-OP e EcoTrec é resultado de ambos utilizarem um servidor central para construir o conhecimento sobre condição de tráfego.

5. Conclusão

Congestionamento se tornou um problema recorrente afetando vários aspectos na sociedade. Porém, muitas soluções para lidar com esse problema apresentam limitações, como a falta de escalabilidade, ou até mesmo a falta de conhecimento sobre o tráfego. Sendo assim, foi proposto GTE, um TMS totalmente distribuído e escalável, baseado em compartilhamento oportunista para lidar com congestionamentos, permitindo que os veículos possuam um conhecimento sobre a condição de tráfego, para que os mesmos possam detectar vias congestionadas e calcularem uma rota alternativa para melhorar a eficiência do tráfego. GTE foi comparado com outras soluções existentes para avaliar sua escalabilidade e eficiência em gerenciar o tráfego.

Os resultados mostraram uma performance superior do GTE em todas as métricas avaliadas, em comparação com as soluções CARTIM, CO-OP e EcoTrec. Como trabalho futuro,

pretende-se analisar como o tamanho da área crítica, área de ancoragem e área de encaminhamento impacta no desempenho da solução. Além disso, pretendemos estender a solução para lidar com congestionamentos causados por eventos inesperados, e também sugerir rotas seguras.

6. Agradecimentos

Os autores gostariam de agradecer o apoio financeiro concedido da FAPESP por meio do processo nº 2015/07538-1, Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP).

Referências

- Araujo, G., Queiroz, M., Duarte-Figueiredo, F., Tostes, A., and Loureiro, A. (2014). Um protocolo de identificação e minimização de congestionamentos de tráfego para redes veiculares. In *Simpósio Brasileiro de Redes de Computadores e Sistemas Distribuídos - (SBRC)*.
- Castro, A. A. V., Serugendo, G. D. M., and Konstantas, D. (2008). Hovering information – self-organizing information that finds its own storage. In *Spring, 2008 Autonomic Communication*, pages 111 – 145.
- de Souza, A. M., Guidoni, D., Botega, L. C., and Villas, L. A. (2015). Co-op: Uma solução para a detecção, classificação e minimização de congestionamentos de veículos utilizando roteamento cooperativo. In *Simpósio Brasileiro de Redes de Computadores e Sistemas Distribuídos - (SBRC)*.
- de Souza, A. M. and Villas, L. A. (2016). A fully-distributed traffic management system to improve the overall traffic efficiency. In *Proceedings of the 19th ACM International Conference on Modeling, Analysis and Simulation of Wireless and Mobile Systems, MSWiM '16*, pages 19–26, New York, NY, USA. ACM.
- de Souza, A. M., Yokoyama, R., Boukerche, A., Maia, G., Cerqueira, E., Loureiro, A. A., and Villas, L. A. (2016). Icarus: Improvement of traffic condition through an alerting and re-routing system. *Computer Networks*, 110:118 – 132.
- Djahel, S., Doolan, R., Muntean, G.-M., and Murphy, J. (2015). A Communications-Oriented Perspective on Traffic Management Systems for Smart Cities: Challenges and Innovative Approaches. *IEEE Communications Surveys Tutorials*, 17(1):125–151.
- Doolan, R. and Muntean, G. M. (2016). Ecotrec: A novel vanet-based approach to reducing vehicle emissions. *IEEE Transactions on Intelligent Transportation Systems*, PP(99):1–13.
- Hyytiä, E., Virtamo, J., Lassila, P., Kangasharju, J., and Ott, J. (2011). When does content float? characterizing availability of anchored information in opportunistic content sharing. In *INFOCOM, 2011 Proceedings IEEE*, pages 3137–3145.
- Karagiannis, G., Altintas, O., Ekici, E., Heijenk, G., Jarupan, B., Lin, K., and Weil, T. (2011). Vehicular networking: A survey and tutorial on requirements, architectures, challenges, standards and solutions. *Communications Surveys Tutorials, IEEE*, 13(4):584–616.
- Kirkpatrick, S., Gelatt, C. D., Vecchi, M. P., et al. (1983). Optimization by simulated annealing. *science*, 220(4598):671–680.
- Pan, J., Popa, I. S., and Borcea, C. (2016). Divert: A distributed vehicular traffic re-routing system for congestion avoidance. *IEEE Transactions on Mobile Computing*, PP(99):1–1.
- Wang, M., Shan, H., Lu, R., Zhang, R., Shen, X., and Bai, F. (2015). Real-time path planning based on hybrid-vanet-enhanced transportation system. *Vehicular Technology, IEEE Transactions on*.