

Sistema Distribuído para Gerenciamento de Informação e Distribuição de Conhecimento em Redes Veiculares

Ademar T. Akabane^{1,2}, Richard W. Pazzi²,
Edmundo R. M. Madeira¹, Leandro A. Villas¹

¹Instituto de Computação – Universidade de Campinas (UNICAMP), Brasil
takeo@lrc.ic.unicamp.br, {edmundo, leandro}@ic.unicamp.br

²Instituto de Tecnologia – Universidade de Ontário (UOIT), Canadá
richard.pazzi@uoit.ca

Abstract. *In Intelligent Transportation System (ITS), the periodic information exchange between vehicles is of vital importance to provide services in a vehicular environment. Typically, ITS applies this practice to extract knowledge about the vehicle traffic condition and to distribute knowledge. Although the solutions designed to use this approach reach the main goal, distribution of knowledge, they do not worry about selection of the most relevant vehicle in the network to perform the task of knowledge extraction. In these solutions the extraction can be carried out by any vehicle, consequently, it could easily cause a network overhead, thus, generating a highly redundant traffic of knowledge in the network. With this in mind, we propose the SDGI, a system for information management and knowledge distribution, which applies the egocentric betweenness measure, with the purpose of selecting the most relevant vehicle to carry out the knowledge generation. The results have shown that the SDGI outperforms solutions found in the literature at different requirements.*

Resumo. *Sabe-se que, em Sistemas de Transporte Inteligentes (STIs), a constante troca de informação entre os veículos é de vital importância para prover serviços em um ambiente veicular. Tipicamente, os STIs fazem uso dessa prática para extrair o conhecimento sobre a condição do tráfego veicular e distribuir o conhecimento. Apesar das soluções que aplicam essa abordagem atingirem o objetivo principal, distribuição do conhecimento, elas não se preocupam em selecionar o veículo mais relevante na rede para realizar a tarefa de extração do conhecimento. Nessas soluções a extração pode ser realizada por qualquer veículo, conseqüentemente acarretará na sobrecarga da rede, devido as elevadas transmissões de mensagens redundantes. Com isso em mente, foi proposto SDGI, um sistema para o gerenciamento de informação e distribuição de conhecimento que aplica a métrica de centralidade de intermediação egocêntrica, com o propósito de selecionar o veículo mais relevante para realizar a extração do conhecimento. Os resultados mostraram que o SDGI tem o melhor desempenho em diferentes requisitos em comparação com as soluções da literatura.*

Introdução

Uma das tecnologias promissoras para aplicações de Sistemas de Transporte Inteligentes (STIs) é a redes veiculares (VANETs), pois as VANETs possibilitam a comunicação veículo-com-veículo (V2V) e veículo-com-infraestrutura (V2I), como as *Road Side Units* (RSUs) [Yousefi et al. 2006, Akabane et al. 2017a]. Além disso, devido à topologia altamente dinâmica das VANETs, uma pilha de protocolos para a comunicação inter-veicular, conhecida como IEEE 1609 *Wireless Access in Vehicular Environments* - WAVE, tem sido desenvolvida para atender um conjunto de requisitos, tais como tempo curto de

comunicação e elevada partição da rede [Campolo et al. 2011]. Além do mais, o padrão IEEE 802.11p DSRC (*Dedicated Short-Range Communication*) tem atribuído múltiplos canais de comunicação inter-veicular tais como canal de controle (CCH) e canais de serviços (SCHs) [Campolo et al. 2011].

Para muitas aplicações de STIs, o constante compartilhamento de informações locais entre os veículos vizinhos de 1-salto de comunicação é essencial para criar o conhecimento sobre as condições de tráfego veicular [Lochert et al. 2010, Yu et al. 2012, Yuan et al. 2014]. Esse tipo de compartilhamento é bem conhecido como *beaconing* e na maioria das vezes ocorre por meio do canal de controle com uma frequência de transmissão geralmente entre 1Hz e 10Hz [Sommer et al. 2015]. As informações contidas por padrão no pacote *beacon* são, por exemplo: identificação do veículo, posição atual do veículo, velocidade média, direção de deslocamento, entre outras [Campolo et al. 2011]. Já o compartilhamento de informação de caráter global é dada por meio de um dos canais de serviços.

Diferentes STIs foram propostos para lidar com gerenciamento de informação local e distribuição do conhecimento gerado sobre as condições de tráfego veicular [Lochert et al. 2010, Yu et al. 2012, Yuan et al. 2014]. Esse tipo de sistema extrai o conhecimento por meio do processamento das informações locais agregadas recebidas das vizinhanças de 1-salto de comunicação. No entanto, nesses sistemas foram observados uma grande lacuna, a ausência de nomeação do veículo para realizar as tarefas de agregação e de extração do conhecimento. Pois se as tarefas forem executadas por todos os veículos que estão participando da rede, pode ocasionar em um alto tráfego de conhecimento redundante. Além do mais, as soluções propostas por [Lochert et al. 2010, Yu et al. 2012] não aplicam nenhum mecanismo de supressão de *broadcast* durante a distribuição do conhecimento, sobrecarregando ainda mais a rede.

Com intuito de superar as lacunas destacadas anteriormente, foi proposto SDGI um Sistema Distribuído para Gerenciamento de Informação e distribuição de conhecimento. Por meio da prática do *beaconing*, o SDGI coleta as informações locais necessárias para aplicar a métrica de centralidade de intermediação egocêntrica, com o propósito de selecionar o veículo mais relevante no momento, para agregar as informações locais e/ou extração do conhecimento. A relevância é definida como a importância de um veículo em função dos fluxos de informação que passam por ele, ou seja, reflete o quanto o veículo intermediário é importante para a sequência do fluxo de informação na rede. Uma das vantagens dessa métrica é o emprego das informações locais (egocêntrica) disponíveis para realização do cálculo. Além dessa vantagem, no trabalho de [Akabane et al. 2017b] foi confirmado que a métrica de centralidade de intermediação egocêntrica, em ambiente altamente dinâmico, possui alta correlação comparada com a métrica de centralidade de intermediação sociocêntrica. Adicionalmente foi aplicado um mecanismo de supressão de *broadcast* para evitar o tráfego de conhecimento redundante.

O principal objetivo desse trabalho é reduzir a sobrecarga da rede apresentada nas soluções da literatura. Os experimentos do sistema proposto foram realizados por meio de simulações comparando com os sistemas da literatura [Lochert et al. 2010, Yu et al. 2012, Yuan et al. 2014] em diferentes requisitos. As análises de desempenho foram feitas sobre duas perspectivas: (i) desempenho dos sistemas em relação à distribuição do conhecimento e (ii) taxa de utilização do canal de controle nas trocas de informações locais.

O restante deste artigo está organizado da seguinte maneira. Na Seção 2 é descrita uma visão geral dos trabalhos existentes na literatura para gerenciamento de informação e distribuição do conhecimento gerado. A solução proposta neste trabalho é apresentada na Seção 3. Enquanto que as avaliações detalhadas dos resultados dos experimentos são

descritas na Seção 4. Por fim, na Seção 5 é apresentada a conclusão do trabalho.

Trabalhos Relacionados

Os sistemas descritos a seguir empregam a troca periódica de informações locais entre as vizinhanças de 1-salto de comunicação, com intuito de criar a base de conhecimento local. Além do mais, todos eles são independentes de uma infraestrutura de comunicação, ou seja, empregam apenas a comunicação entre veículos.

Os autores [Lochert et al. 2010] propuseram um sistema de agregação probabilística, chamado de Probabilistic. O sistema aplica uma abordagem de agregação hierárquica chamada de *soft-state sketches* baseada em uma versão modificada de *Flajolet-Martin sketches* [Flajolet and Martin 1985]. A principal característica dessa abordagem está no fato que as informações agregadas não possuem valores específicos sobre o local sensoriado (por exemplo, a quantidade de veículos de uma determinada via), ao invés disso, esta possui uma aproximação probabilística. A principal vantagem dessa abordagem é a capacidade de combinar múltiplos valores agregados sobre o mesmo contexto, para a geração do conhecimento. Porém, uma das desvantagens desse sistema está no fato de não possuir nenhum mecanismo de seleção do veículo para agregar as informações locais e gerar o conhecimento. Pois na abordagem adotada todo veículo pode ser candidato a realizar ambas tarefas. Como consequência gerando alto tráfego de conhecimento redundante na rede. Além do mais, durante a distribuição do conhecimento não é aplicado nenhum mecanismo de supressão de *broadcast*.

No trabalho de [Yu et al. 2012] foi proposto o Catch-up, um sistema de encaminhamento de informações agregadas adaptativo com base em atraso. Este sistema tem como objetivo principal permitir o encontro das informações agregadas de diferentes fontes para geração do conhecimento. A fim de aumentar a probabilidade do encontro das informações agregadas é inserido um atraso adaptativo no encaminhamento da mensagem que pode ser um dos tipos, *RUN* (curto) ou *WALK* (longo). O atraso adaptativo é calculado com base nas informações locais recebidas. O sistema foi projetado com base no algoritmo de aprendizagem distribuído, isto é, cada veículo aprende por meio das observações locais e estipula um atraso baseado nos resultados dessa aprendizagem. A vantagem desse sistema é o emprego da agregação probabilística e um atraso adaptativo no encaminhamento das informações agregadas. No entanto, pode-se observar que não existe nenhum mecanismo de supressão de *broadcast* durante a distribuição do conhecimento e qualquer veículo pode ser agregador de informação e gerador do conhecimento.

Os autores [Yuan et al. 2014] propuseram o DA2RF (*Data Aggregation Algorithm by Restricting Forwarders*) um mecanismo de agregação de dados com foco principalmente na seleção dos veículos que continuarão o processo de encaminhamento do conhecimento gerado na agregação. Para isso, todo veículo recebe uma das duas etiquetas (*encaminhador* ou *não-encaminhador*) de acordo com as etiquetas das vizinhanças. Essa etiqueta, como o próprio nome diz, determina se o veículo é encaminhador do conhecimento ou não. Todo veículo será não-encaminhador se houver um veículo encaminhador imediatamente à sua frente e atrás. A identificação dos veículos na vizinhança é dada pela troca periódica de informações locais. Uma das vantagens dessa solução é o mecanismo aplicado para supressão de *broadcast* durante o processo de distribuição do conhecimento. No entanto, não há nenhum mecanismo de seleção de veículo para agregar informações recebidas das vizinhanças e gerar o conhecimento. Dessa forma, qualquer veículo pode contribuir na geração do conhecimento, assim produzindo conhecimento redundante na rede.

Observa-se que nos sistemas anteriormente apresentados apontam algumas

limitações como: (i) soluções que empregam apenas a técnica de agregação da informação ([Lochert et al. 2010, Yu et al. 2012]). Outra aplicando alguma técnica de agregação da informação e também um mecanismo de supressão de *broadcast* ([Yuan et al. 2014]). Contudo, todos os sistemas apresentados possuem a ausência de mecanismo de seleção do veículo para realizar as tarefas de agregação da informação e de geração do conhecimento. Com base na deficiência observada, foi proposto um sistema para gerenciamento de informação e distribuição de conhecimento, chamado SDGI. Este aplica a métrica de centralidade de intermediação egocêntrica para a seleção do veículo mais relevante na rede, a fim de desempenhar as tarefas anteriormente mencionadas, com intuito de reduzir a sobrecarga da rede. Além disso, aplica-se o mecanismo de supressão de *broadcast* para evitar o tráfego de conhecimento redundante.

SDGI

Neste trabalho foi proposto o SDGI, um sistema distribuído para gerenciar informações e distribuir conhecimento sobre as condições de tráfego veicular por meio das VANETs, sem a necessidade do apoio de uma infraestrutura externa, como as RSUs. A proposta emprega técnica de agregação de dados, métrica de centralidade aplicada em redes sociais e mecanismo de supressão de *broadcast* para alcançar o objetivo principal.

Um dos grandes desafios deste tipo de sistema é a seleção do veículo mais relevante na rede, devido à sua alta dinamicidade, para continuar o processo de agregação de dados e/ou geração de conhecimento sobre o local sensoriado. Pois, se todos os veículos que estiverem participando da rede realizarem essas tarefas, isto acarretará em inundações de conhecimento redundante na rede. Com isso em mente, a medida de centralidade de intermediação egocêntrica foi aplicada para a seleção do veículo mais relevante para realizar a agregação da informação local e/ou a geração do conhecimento. Essa métrica egocêntrica foi escolhida devido a sua simplicidade de utilizar apenas as informações locais disponíveis (vizinhos de 1-salto de comunicação). Além de atribuir a relevância do veículo em relação ao fluxo de informação que passa por ele.

O restante da seção está dividido da seguinte forma: a métrica aplicada para a seleção do veículo mais relevante para a agregação de informação e/ou geração do conhecimento é apresentada na Subseção 3.1. Enquanto que, na Subseção 3.2 estão descritos a técnica empregada na agregação da informação e o procedimento de geração do conhecimento. Além do mecanismo de supressão de *broadcast* aplicado neste trabalho.

Centralidade de Intermediação Egocêntrica

Antes de apresentar a métrica de Centralidade de Intermediação Egocêntrica (CIE), faz-se necessária uma compreensão da definição de rede egocêntrica, uma vez que a métrica de CIE é calculada sobre a topologia deste tipo de rede. Por definição, uma rede egocêntrica é uma subrede local constituída de um nó central (*ego*), juntamente com os nós vizinhos (*alters*) que estão diretamente conectados a ele, além das interconexões entre os *alters* [Marsden 2002, Everett and Borgatti 2005], veja a Figura 1. Na figura o *ego* está sendo representado pelo nó n e os *alters* estão sendo representados pelos nós enumerados de 1 a 5. A seguir, será dada uma definição formal de uma rede egocêntrica e o cálculo matemático da métrica de CIE.

Definição: Considere um grafo não direcionado $G = (V, E)$ no qual V corresponde um conjunto de vértices (v) ou nós, E corresponde um conjunto de arestas ou links (e , onde $e \in E \subseteq V \times V$ é identificado por um par de vértices). Já os vizinhos de um vértice v' é o conjunto de vértices $v \in V$ atingíveis em r salto(s). Considere que N_n^r seja um conjunto de vértices atingíveis em r -salto a partir de *ego* (n), i.e., $N_n^r = \{v' \in$

$V|v' \neq n \wedge 1 \leq d(n, v') \leq r\}$, onde $d(n, v')$ denota 1-salto entre n and v' . Assim, 1ª-ordem do vértice n consiste de um grafo não direcionado $G = (V_n^1, E_n^1)$, onde o conjunto de vértices corresponde $V_n^1 = \{N_n^1 \cup \{n\}\}$ e o conjunto de arestas corresponde $E_n^1 = \{(i, j) \in E_n^1 | i, j \in V_n^1\}$.

Matematicamente, uma matriz de adjacência simétrica ($B_{n \times n}$) pode representar o *link* de intercomunicação entre os nós, no qual n é o número de nós de 1-salto de comunicação. Dessa forma, cada elemento da matriz de adjacência, $b_{i,j}$, é representado da seguinte maneira:

$$b_{ij} = \begin{cases} 1 & \text{se existe um link de comunicação entre } i \text{ and } j \\ 0 & \text{caso contrário} \end{cases}$$

Para demonstrar o cálculo da métrica de CIE juntamente com a matriz de adjacência, foi utilizado como exemplo um grafo bem-conhecido da Figura 2 [Marsden 2002].

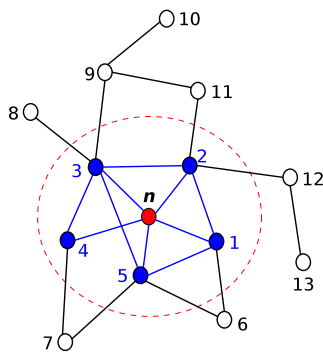


Figura 1. Exemplo de uma rede egocêntrica.

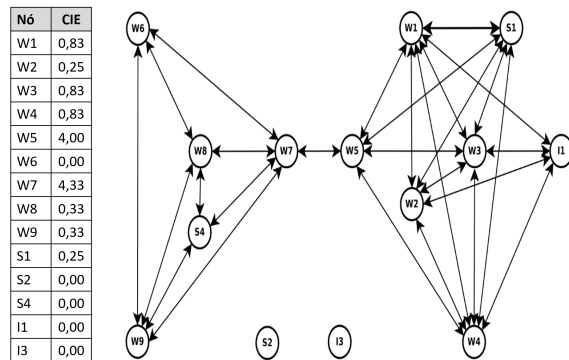


Figura 2. Exemplo ilustrativo com o valor de CIE para cada nó.

De acordo com os autores [Marsden 2002, Everett and Borgatti 2005], o valor de CIE de cada nó pode ser calculado por meio da contagem da quantidade de pares de *alters* não-adjacentes que estão diretamente ligados ao *ego*. Assim, por meio da expressão matemática, $B^2[1 - B]_{i,j}$, o valor de CIE do *ego* é inversamente proporcional à soma dos valores de entrada [Everett and Borgatti 2005], onde 1 na expressão corresponde uma matriz com todos os elementos iguais a 1. Para maior clareza, foi realizado o cálculo do valor de CIE sobre a perspectiva apenas do nó $W4$ da Figura 2 como exemplo. A matriz de adjacência do nó $W4$, abaixo à esquerda, representa uma visão completa de todos os *links* de conexão entre o $W4$ e os *alters*, além das interconexões entre os *alters*. Uma vez que a matriz $W4$ é simétrica, somente os valores inteiros diferentes de zero acima da diagonal principal necessitam ser analisados. Dessa forma, os valores de entrada resultante da $W4^2[1 - W4]$ são 4, 3 e 4, como demonstrado na matriz abaixo à direita. Assim, o valor de CIE para o nó $W4$ é 0,83 ($1/4 + 1/3 + 1/4$). Uma vez feito o cálculo de CIE, esta informação é adicionada no pacote *beacon*. O cálculo de CIE é feito toda vez que existe uma atualização na matriz de adjacência. A Figura 2 possui uma tabela que apresenta o valor de CIE para cada nó do grafo.

$$W4 = \begin{matrix} & W4 & I1 & S1 & W3 & W1 & W2 & W5 \\ \begin{matrix} W4 \\ I1 \\ S1 \\ W3 \\ W1 \\ W2 \\ W5 \end{matrix} & \begin{bmatrix} 0 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 1 & 0 & 0 \end{bmatrix} \end{matrix} \quad W4^2[1-W4] = \begin{matrix} & W4 & I1 & S1 & W3 & W1 & W2 & W5 \\ \begin{matrix} W4 \\ I1 \\ S1 \\ W3 \\ W1 \\ W2 \\ W5 \end{matrix} & \begin{bmatrix} * & * & * & * & * & * & * \\ * & * & \mathbf{4} & * & * & * & \mathbf{3} \\ * & * & * & * & * & * & * \\ * & * & * & * & * & * & * \\ * & * & * & * & * & * & * \\ * & * & * & * & * & * & \mathbf{4} \\ * & * & * & * & * & * & * \end{bmatrix} \end{matrix}$$

No trabalho de [Marsden 2002] foi demonstrado que as medidas de centralidade de intermediação egocêntrica e sociocêntrica possuem alta correlação nas redes estáticas. Porém, no trabalho de [Akabane et al. 2017b] foi demonstrado que essa alta correlação também ocorre em redes altamente dinâmicas, como as VANETs.

Como mencionado, a métrica de CIE foi aplicada para selecionar o veículo mais relevante para continuar o processo de agregação da informação local e/ou a geração de conhecimento. Pode-se observar na tabela do exemplo da Figura 2 que existem alguns nós com o mesmo valor de CIE, por exemplo: 0,83. Supondo que a Figura 2 seja um grafo de *links* de comunicação inter-veicular em um determinado instante, e o nó *I1* necessita transmitir o seu conhecimento local agregado. Entretanto, antes disso, o nó *I1* deve selecionar o *alter* que possui o maior valor de CIE para dar continuidade no processo de agregação e/ou na geração do conhecimento. Neste caso, observa-se que o *I1* possui três *alters* (*W1*, *W3* e *W4*) com valor de CIE igual a 0,83. Constatado isso, o indicador de potência do sinal foi aplicado como critério de desempate e o modelo utilizado foi *Two-Ray Interference* [Sommer et al. 2012], como demonstrado na Equação 1.

$$L_{TRI}[dB] = 20 \log(4\pi \frac{d}{\lambda} |1 + \Gamma \exp^{\varphi}|^{-1}) \quad (1)$$

onde λ é o comprimento de onda, d é a distância Euclidiana entre dois veículos, Γ é o coeficiente de reflexão e φ é a interferência dos raios. Para o cálculo da interferência dos raios é dada da seguinte forma:

$$\varphi = 2\pi \frac{d_{los} - d_{ref}}{\lambda}, \quad \begin{cases} d_{los} = \sqrt{d^2 + (h_t - h_r)^2} \\ d_{ref} = \sqrt{d^2 + (h_t + h_r)^2} \end{cases} \quad (2)$$

onde d_{los} e d_{ref} descrevem a distância de propagação e a distância de reflexão, respectivamente. As variáveis h_t e h_r representam a altura em que a antena do transmissor e do receptor, respectivamente, estão em relação ao solo. Neste caso foi utilizada a mesma altura do teste real feito no trabalho de [Sommer et al. 2012] para ambas antenas, de 1,495 m. Já o coeficiente de reflexão pode ser calculado da seguinte forma:

$$\Gamma = \frac{\sin \theta_i - \sqrt{\varepsilon - \cos \theta_i}}{\sin \theta_i + \sqrt{\varepsilon - \cos \theta_i}}, \quad \begin{cases} \sin \theta_i = \frac{h_t + h_r}{d_{ref}} \\ \cos \theta_i = \frac{d}{d_{ref}} \end{cases} \quad (3)$$

onde ε representa a reflexão em relação ao solo com valor de 1,02 [Sommer et al. 2012] e θ é o ângulo entre o solo e o raio refletido. Por fim, o valor do comprimento de onda foi fixado em 0,051 m de acordo com IEEE 802.11p [IEEE 2013].

Dando continuidade no exemplo, uma vez selecionado o *alter* após o critério de desempate, por exemplo o *W3*, este realiza a agregação das suas informações com as recebidas, enquanto que os *alters* restantes descartam as informações agregadas recebidas e assim por diante. Esse processo de agregação de informação local será realizado até

atingir o nó $W7$, pois, temporariamente, ele possui o maior valor de CIE na rede. Uma vez agregadas todas as informações locais, o nó $W7$, se necessário, gera o conhecimento e inicia-se a distribuição do conhecimento produzido. Na subseção seguinte será explicada a técnica de agregação de dados, além do procedimento para a geração do conhecimento e o mecanismo de supressão de *broadcast* aplicado.

Agregação de dados e Procedimento de Geração do Conhecimento

SDGI compartilha periodicamente as informações locais, com os vizinhos de 1-salto de comunicação, por meio de pacotes *beacons* através do canal de controle, a fim de criar a base de conhecimento local. Além das informações já contidas no pacote *beacon*, foram adicionadas mais duas informações necessárias: o valor atual do CIE e as informações agregadas.

A base de conhecimento local é criada por meio da agregação das informações locais recebidas da vizinhança, além da computação do peso das vias. Uma vez criada a base é necessário o compartilhamento desta com o vizinho mais relevante de acordo com o critério de seleção apresentado na Subseção 3.1.

A Fusão de dois valores agregados pode ser representada de forma geral da seguinte maneira: $A_r := \partial(A_1, A_2)$, onde ∂ é a função de agregação que tem como entrada dois valores agregados ou não (A_1 e A_2) e os combinam gerando um novo valor agregado (A_r). Como o trabalho está focado no gerenciamento de informação e distribuição de conhecimento sobre a condição de tráfego veicular, logo a função de agregação aplicada é dada da seguinte maneira:

$$v_{agr_i}^{med} = \frac{v_1 c_1 + v_2 c_2}{c_1 + c_2} \quad (4)$$

onde $v_{agr_i}^{med}$ representa a velocidade média agregada de uma determinada via i . O v_1 e o v_2 representam os dois valores de entrada, da via i , a serem agregados. O c_i é um contador que indica o número de informações que contribuiu para o gerar o valor agregado. Já o peso da via i (w_i) é calculado da seguinte forma:

$$w_i = \frac{v_{agr_i}^{med}}{v_{via_i}^{max}}, \begin{cases} w_i : \text{peso da via } i \\ v_{agr_i}^{med} : \text{velocidade média agregada da via } i \\ v_{via_i}^{max} : \text{velocidade limite da via } i \end{cases} \quad (5)$$

Após computar todas as informações locais agregadas, o veículo de maior valor de CIE classifica o peso da via de acordo com a Tabela 1. Os níveis de serviços e as classificações dos tráfegos foram baseados no *Highway Capacity Manual* (HCM) [Elefteriadou 2016].

Tabela 1. Classificação de cada nível de serviço [Elefteriadou 2016].

Nível de serviço	Classificação do tráfego	Peso da via
A	Tráfego livre	(1,0 ~ 0,9]
B	Tráfego razoavelmente livre	(0,9 ~ 0,7]
C	Limite da capacidade da via	(0,7 ~ 0,5]
D	Congestionamento leve	(0,5 ~ 0,4]
E	Congestionamento crítico	(0,4 ~ 0,33]
F	Tráfego totalmente colapsado	(0,33 ~ 0,0]

Terminada a etapa de classificação, caso seja identificado algum evento, nesse caso via(s) congestionada(s) (nível D, E ou F da Tabela 1), cria-se uma mensagem (ou conhecimento) contendo a(s) identificação(ões) dessa(s) via(s) e inicia-se o processo de distribuição desse conhecimento no canal de serviço. Todos os veículos da vizinhança que recebem a mensagem agendam a retransmissão para continuar o processo de distribuição desse conhecimento. Toda vez que um veículo recebe um conhecimento para ser distribuído, este verifica se está localizado em uma zona de preferência [Villas et al. 2014, Akabane et al. 2015], caso afirmativo, ele transmite primeiro. Devido ao mecanismo de supressão de *broadcast* implementado, assim que os veículos receberem a mesma mensagem agendada, eles cancelam a retransmissão, dessa forma evitando retransmissões de mensagens redundantes na rede.

Na Figura 3 é apresentado o fluxo de operação realizado pelo SDGI. Enquanto que o SDGI recebe as informações locais, este ou insere ou agrega as informações locais na base de conhecimento (Bloco 1 da Figura 3). Passado essa etapa, este computa o peso das vias (Legenda (a)), além disso, caso o veículo possuir o maior valor de CIE (Legenda (b)), este classifica o peso das vias de acordo com a Tabela 1 e se constatar via(s) congestionada(s) (Legenda (c)) uma mensagem é gerada e distribuída com o conhecimento produzido (Legenda (d)). Caso o veículo não possuir o maior valor de CIE este seleciona o próximo veículo, de acordo com a Subseção 3.1, para realizar a agregação da informação e/ou geração do conhecimento e envia a sua informação local agregada (Legenda (e)).

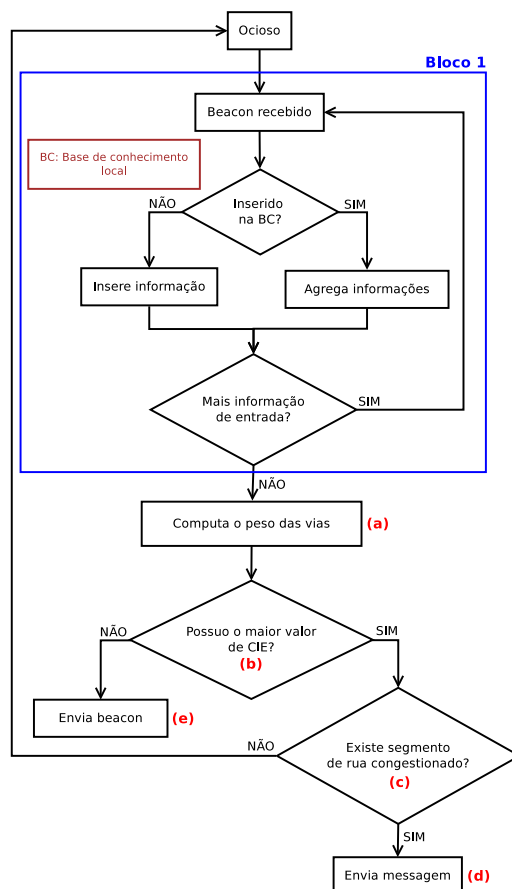


Figura 3. Fluxo de operação do SDGI para geração do conhecimento.

Avaliação de Desempenho

Nesta seção é apresentada a metodologia científica aplicada neste trabalho, os simuladores empregados e também os resultados comparativos da abordagem proposta com as soluções da literatura.

Metodologia de Simulação

A avaliação de desempenho do SDGI foi realizada por meio de simulações com apoio das seguintes ferramentas: Veins 4.5 - *Vehicular Network Simulations*¹, OMNet++ 5.0 - *Network Simulation Framework*² e SUMO 0.29.0 - *Simulation of Urban MObility*³. O modelo de mobilidade dos veículos empregado foi o *Krauss β car-following* [Brackstone and McDonald 1999]. Neste modelo, o comportamento de cada veículo é calculado em relação ao estado (velocidade, aceleração e posição) dos veículos da vizinhança. Cada veículo é capaz de se localizar com auxílio do GPS e a informação das vias pode ser extraída com auxílio do mapa digital.

Para o cenário, foi empregado um segmento de mapa da região urbana da cidade de Erlangen - Alemanha, (Figura 4), e as densidades utilizadas nos experimentos foram de 100, 150, 200, 250 e 300 veículos por km². Além disso, cada veículo foi configurado com uma potência de transmissão de 0.98 mW, *bitrate* de 6 Mbps e o raio de comunicação de 200 m com o padrão de comunicação 802.11p. Por fim, a Tabela 2 descreve os principais parâmetros de simulação aplicados nos experimentos.



Figura 4. Cenário Erlangen - Alemanha.

Tabela 2. Parâmetros.

Parâmetro	Valor
Densidade dos veículos	100 a 300 veículos/km ²
Camada MAC	802.11p
Potência de transmissão	0.98 mW
Bitrate	6 Mbps
Raio de transmissão	200 m
Frequência de transmissão (<i>Beacon</i>)	1 Hz
Tempo de simulação	350 s
Intervalo de confiança	95 %

Em seguida são apresentadas as métricas de avaliação empregadas em todas as soluções:

1. **Mensagens transmitidas:** total de mensagens transmitidas para realizar a entrega da conhecimento gerado;
2. **Cobertura:** porcentagem de mensagens entregues para os veículos que estão dentro do cenário;
3. **Atraso:** tempo gasto para realizar a entrega das mensagens para os veículos que estão dentro do cenário;
4. **Colisões de pacote:** número total de colisões de pacotes durante a transmissão das mensagens;
5. **Taxa de ocupação do canal:** porcentagem de utilização do canal de controle durante a troca de informações locais;
6. **Beacons transmitidos:** total de *beacons* transmitidos na rede;
7. **Beacons transmitidos por veículos:** total de *beacons* transmitidos por cada veículo.

¹<http://veins.car2x.org/>

²<https://omnetpp.org/>

³<http://sumo.sourceforge.net/>

Análise dos resultados das simulações

As análises dos resultados foram divididas em dois grupos. Primeiro foi analisado o desempenho das soluções na distribuição do conhecimento, Subseção 4.2.1. Enquanto que na Subseção 4.2.2 foi avaliada a taxa de ocupação do canal de controle durante a troca de informações locais.

Análise dos Resultados em Relação à Distribuição do Conhecimento

Na Figura 5 são apresentados os resultados dos experimentos com referência a quatro métricas empregadas em função das densidades de veículos. Mais especificamente, na Figura 5(a) é exposto o resultado da cobertura para todas as soluções analisadas. A solução Probabilistic apresenta a menor cobertura para todas as densidades analisadas, alcançando uma média de 80 %. Tal resultado é decorrente da sobrecarga da rede (Figura 5(b) e Figura 5(d)), pois nessa solução todos veículos realizam as tarefas de agregação de informação, geração e distribuição do conhecimento produzido. Além do mais, durante o processo de distribuição de conhecimento não é aplicado nenhum mecanismo de supressão de *broadcast*, resultando em uma alta taxa de transmissão de conhecimento redundante (veja a Figura 5(b)). As consequências dessa alta taxa resultaram em elevadas colisões de pacotes na rede (veja a Figuras 5(d)) e maior atraso na entrega do conhecimento (veja a 5(c)). Observa-se também uma ligeira queda na cobertura em altas densidades (250 e 300 veículos/km²). Isso se deve ao fato da alta sobrecarga na rede e alta taxa de colisão, assim danificando o conteúdo da informação trafegada.

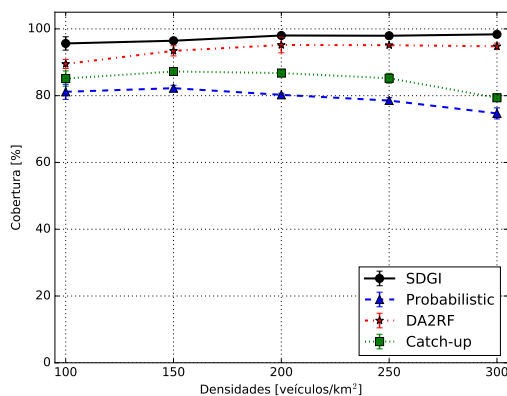
Outra solução analisada é o Catch-up, na qual a principal estratégia é a inserção de um atraso adaptativo no encaminhamento da mensagem para aumentar a probabilidade do encontro das informações agregadas. Observa-se que essa estratégia diminuiu a quantidade de transmissões e consequentemente as colisões, vejam as Figuras 5(b) e 5(d). Sendo assim, Catch-up alcança melhores resultados, comparado com Probabilistic, reduzindo em média 10 % tanto nas mensagens transmitidas, quanto nas colisões de pacotes, além de aumentar a taxa de cobertura em 5 %, veja a Figura 5(a). Da mesma forma que em Probabilistic, em Catch-up conforme eleva a densidade dos veículos pode-se observar uma ligeira queda na cobertura, pois ainda assim esta solução possui alta taxa de transmissão de conhecimentos e colisões de pacotes. Sabe-se que ambos, Probabilistic e Catch-up, não aplicam nenhum mecanismo de seleção de veículo mais relevante para realizar as tarefas de agregação das informações locais e geração do conhecimento, com isso elevando a carga da rede e mantendo o atraso próximo à solução Probabilistic, veja a Figura 5(c).

Já o DA2RF emprega um mecanismo de seleção dos veículos que continuarão o processo de encaminhamento do conhecimento gerado na agregação. Como pode-se observar na Figura 5(a), essa abordagem melhora a cobertura em média de 18 % e 15 % comparando com as soluções Probabilistic e Catch-up, respectivamente. Com esse mecanismo é possível observar uma diminuição na taxa de transmissão de conhecimentos, em média de 30 % e 20 % (Figura 5(b)), e colisões de pacotes, em média de 30 % e 25 % (Figura 5(d)), quando comparado com Probabilistic e Catch-up, respectivamente. Apesar de aplicar apenas o mecanismo de supressão de *broadcast* e desconsiderar a seleção dos veículos mais relevantes para realizar as tarefas de agregação e geração do conhecimento, ainda assim, DA2RF introduz um atraso muito próximo das outras soluções já analisadas, veja a Figura 5(c).

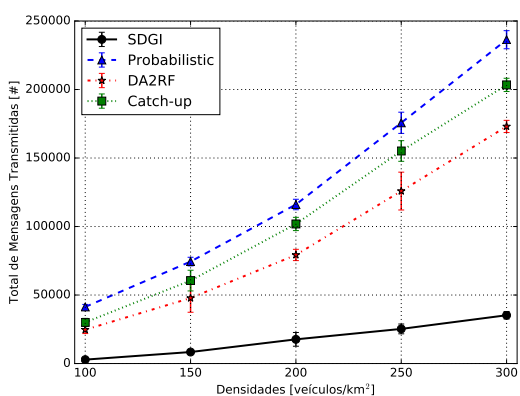
Por fim, como o SDGI aplica a métrica de CIE para a seleção dos veículos mais relevantes na rede, além do mecanismo de supressão de *broadcast* no processo de

distribuição do conhecimento, assim pode-se observar melhores resultados em todas as métricas avaliadas. SDGI reduz drasticamente a quantidade de conhecimentos transmitidos na rede (veja a Figura 5(b)), apresentando uma redução, em média, de mais de 85 %, 80 % e 70 % comparado com Probabilistic, Catch-up e DA2RF, respectivamente. Como consequência dessa redução, os conhecimentos trafegados na rede podem atingir maior quantidade de veículos em todas as densidades analisadas, resultando em uma alta taxa de cobertura, próximo à 98 %, em média (veja a Figura 5(a)). Além do mais, o mecanismo de supressão de *broadcast* implementado juntamente com a métrica de CIE reduziram as colisões de pacotes (veja a Figura 5(d)), apresentando uma redução em média acima de 75 %, 70 % e 50 % comparado com Probabilistic, Catch-up e DA2RF, respectivamente. Sendo assim, o SDGI apresenta o menor atraso médio, entre todas as soluções analisadas, na entrega dos conhecimentos, em torno de 0,15 segundos.

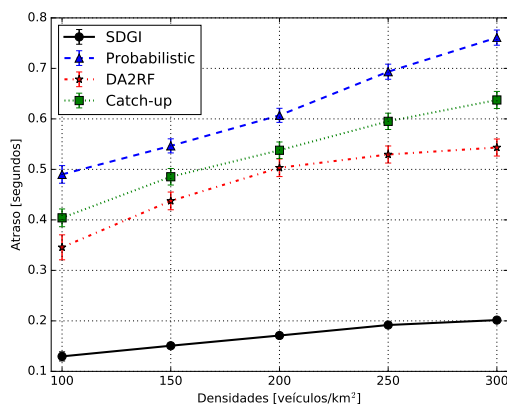
Como conclusão da análise dos resultados, pode-se destacar a importância, para o sistema de gerenciamento de informação e distribuição do conhecimento, de um mecanismo de seleção do veículo mais relevante para realizar a tarefa de agregação de informação e geração do conhecimento. Visto que, sem esse tipo de mecanismo acarretará em uma sobrecarga na rede e uma baixa cobertura. Além do mais, pode-se destacar a métrica de CIE como uma opção viável para o mecanismo de seleção dos veículos mais relevantes em VANETs.



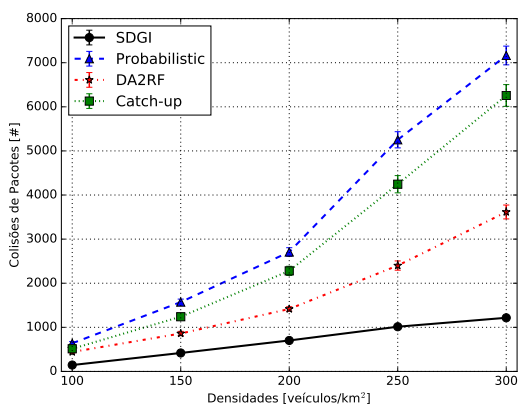
(a) Cobertura vs Densidade.



(b) Total de mensagens transmitidas vs Densidade.



(c) Atraso vs Densidade.



(d) Colisões de pacotes vs Densidade

Figura 5. Análise dos resultados em relação às trocas de mensagens.

Análise dos Resultados em Relação ao Canal de Controle

Nesta subseção descreve a análise dos resultados sobre a taxa de ocupação do canal de controle. Essa análise se faz necessária para avaliar a relação entre frequência de transmissão dos pacotes *beacons* adotada e a taxa de ocupação do canal. Em todas as soluções analisadas, os pacotes *beacons* foram empregados com intuito de trocar informações entre vizinhos de 1-salto de comunicação. Além do mais, em todas as soluções foram aplicadas a mesma frequência de transmissão (1Hz), conforme descrito nos trabalhos de [Lochert et al. 2010, Yu et al. 2012, Yuan et al. 2014], inclusive no SDGI. Dessa forma os resultados apresentados a seguir valem para todas as soluções analisadas neste trabalho.

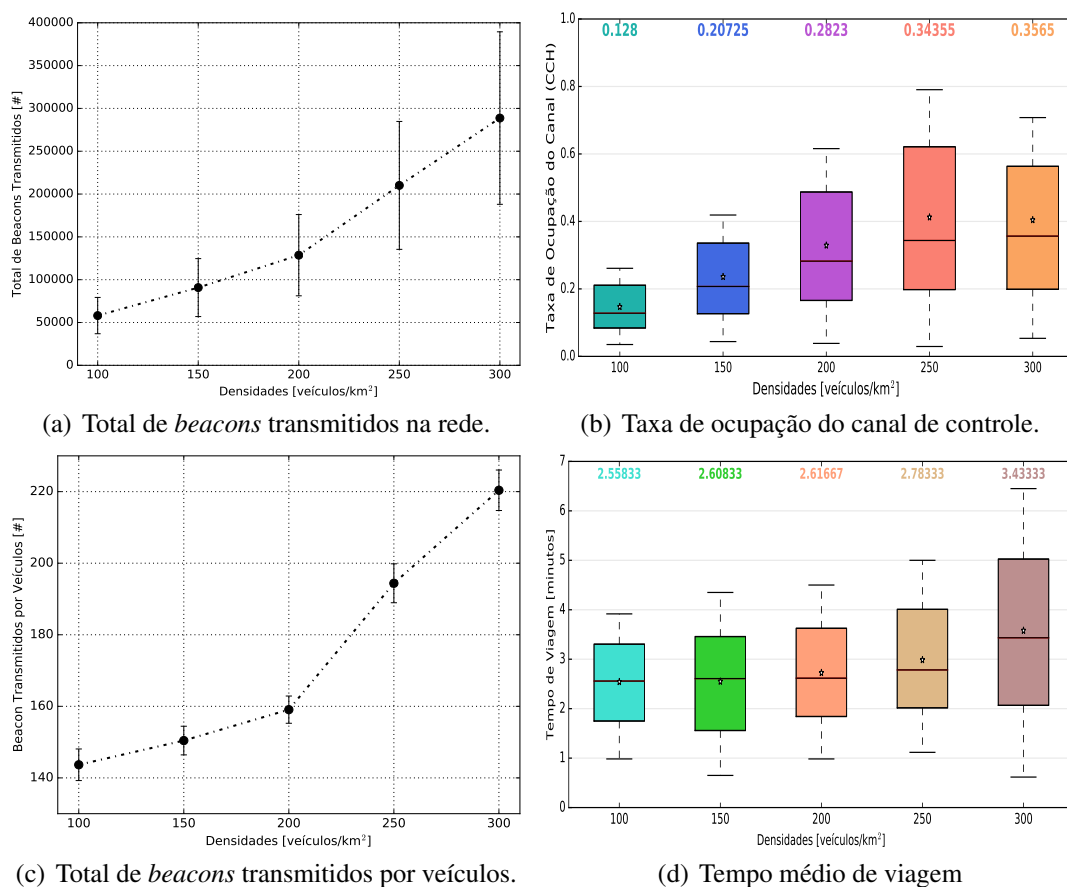


Figura 6. Análise dos resultados em relação ao canal de controle.

Na Figura 6 são apresentados os resultados das métricas empregadas na análise de desempenho do canal de controle em função das densidades de veículos. Em particular, na Figura 6(a) é apresentada uma visão macroscópica da quantidade total de pacotes *beacons* transmitidos na rede. Pode-se observar que a quantidade de pacotes transmitidos aumenta linearmente conforme aumentam as densidades de veículos, como esperado. Essa expectativa é válida, pois conforme elevam-se as densidades dos veículos durante as simulações, maior será o número de pacotes *beacons* transmitidos na rede. Observar-se que a taxa de ocupação do canal de controle, Figura 6(b), apresenta o mesmo comportamento, no valor médio, em relação à quantidade total de pacotes *beacons* transmitidos na rede, Figura 6(a). Esse comportamento observado também é esperado, pois a taxa de ocupação do canal aumenta conforme eleva a quantidade de pacotes que estão sendo

transmitidos na rede. Entre as maiores densidades, como 200, 250 e 300 veículos/km², observa-se uma taxa de ocupação do canal em média de 28 %, 34 % e 35 %, respectivamente, veja a Figura 6(b).

Uma visão microscópica da quantidade de pacotes *beacons* transmitidos na rede pode ser observada na Figura 6(c). Nela é apresentado o número médio de pacotes *beacons* transmitidos por cada veículo em função da densidade. Sabe-se que a quantidade de pacotes *beacons* transmitidos por cada veículo está diretamente relacionado com o tempo de viagem que ele permanece no cenário durante as simulações. Para que fosse possível a comparação, foi gerado o resultado do tempo médio de viagem para cada densidade, veja a Figura 6(d). Como a frequência de transmissão dos pacotes *beacons* aplicada foi de 1 Hz, pode-se observar que a quantidade de pacotes transmitidos (Figura 6(c)), em média, por cada veículo está de acordo com o tempo médio que ele permanece na simulação (Figura 6(d)).

Como conclusão da análise desses resultados, verifica-se que a taxa de transmissão de pacotes *beacons* na frequência de 1 Hz pode ser adequada para o cenário avaliado, juntamente com o modelo de mobilidade aplicado neste trabalho. Dado que a taxa de ocupação do canal de controle apresentou uma utilização de 35 %, em média, para a densidade máxima analisada (300 veículos/km²), veja a Figura 6(b).

Conclusão

Muitas aplicações de Sistemas de Transporte Inteligentes foram propostas para lidar com o gerenciamento de informação e distribuição do conhecimento sobre as condições de tráfego veicular. Neste tipo de sistema o conhecimento é gerado a partir do processamento das informações locais agregadas. Porém nas soluções encontradas na literatura, todos os veículos participam na geração do conhecimento, assim, acarretando em uma sobrecarga na rede. Observada essa lacuna, foi proposto SDGI, um sistema distribuído para gerenciamento de informação e distribuição de conhecimento. SDGI emprega a métrica de centralidade de intermediação egocêntrica com intuito de selecionar o veículo mais relevante para realizar a agregação da informação local e/ou a geração do conhecimento. Além de aplicar um mecanismo de supressão de *broadcast* durante o processo de distribuição do conhecimento.

Os resultados dos experimentos mostraram que SDGI tem melhor desempenho, em todas as métricas avaliadas, em comparação com as soluções Probabilistic, DA2RF e Catch-up. Como trabalho futuro, pretende-se consumir o conhecimento gerado, realizando o cálculo de rotas alternativas para os veículos que possivelmente entrarão em vias congestionadas. Dessa forma abordando as três fases do sistema: sensoriamento do ambiente, geração e distribuição do conhecimento e por fim o consumo desse conhecimento produzido.

Agradecimentos

Os autores agradecem a Fundação de Amparo à Pesquisa do Estado de São Paulo - FAPESP (2015/07538-1, 2015/24494-8, 2015/25588-6 e 2016/24454-9), Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - CAPES (401802/2016-7) e o Natural Sciences and Engineering Research Council of Canada (NSERC) pelo apoio financeiro.

Referências

Akabane, A. T., Gomes, R. L., Pazzi, R. W., Madeira, E. R., and Villas, L. A. (2017a). Apolo: A mobility pattern analysis approach to improve urban mobility. In *GLOBECOM 2017-2017 IEEE Global Communications Conference*, pages 1–6. IEEE.

- Akabane, A. T., Pazzi, W. R., Madeira, E. R. M., and Villas, L. A. (2017b). Applying egocentric betweenness measure in vehicular ad hoc networks. In *Network Computing and Applications (NCA), 2017 IEEE 16th International Symposium on*. IEEE.
- Akabane, A. T., Villas, L. A., and Madeira, E. R. M. (2015). An adaptive solution for data dissemination under diverse road traffic conditions in urban scenarios. In *2015 IEEE Wireless Communications and Networking Conference (WCNC)*, pages 1654–1659. IEEE.
- Brackstone, M. and McDonald, M. (1999). Car-following: a historical review. *Transportation Research Part F: Traffic Psychology and Behaviour*, 2(4):181–196.
- Campolo, C., Molinaro, A., and Vinel, A. (2011). Understanding the performance of short-lived control broadcast packets in 802.11 p/wave vehicular networks. In *Vehicular Networking Conference (VNC), 2011 IEEE*, pages 102–108. IEEE.
- Elefteriadou, L. A. (2016). The highway capacity manual: A guide for multimodal mobility analysis. *TR News*, (306).
- Everett, M. and Borgatti, S. P. (2005). Ego network betweenness. *Social networks*, 27(1):31–38.
- Flajolet, P. and Martin, G. N. (1985). Probabilistic counting algorithms for data base applications. *Journal of computer and system sciences*, 31(2):182–209.
- IEEE, S. A. (2013). Ieee guide for wireless access in vehicular environments (wave) architecture. *IEEE Std*, pages 1609–0.
- Lochert, C., Scheuermann, B., and Mauve, M. (2010). A probabilistic method for cooperative hierarchical aggregation of data in vanets. *Ad Hoc Networks*, 8(5):518–530.
- Marsden, P. V. (2002). Egocentric and sociocentric measures of network centrality. *Social networks*, 24(4):407–422.
- Sommer, C., Joerer, S., and Dressler, F. (2012). On the applicability of two-ray path loss models for vehicular network simulation. In *Vehicular Networking Conference (VNC), 2012 IEEE*, pages 64–69. IEEE.
- Sommer, C., Joerer, S., Segata, M., Tonguz, O. K., Cigno, R. L., and Dressler, F. (2015). How shadowing hurts vehicular communications and how dynamic beaconing can help. *IEEE Transactions on Mobile Computing*, 14(7):1411–1421.
- Villas, L. A., Boukerche, A., Maia, G., Pazzi, R. W., and Loureiro, A. A. (2014). Drive: An efficient and robust data dissemination protocol for highway and urban vehicular ad hoc networks. *Computer Networks*, 75:381–394.
- Yousefi, S., Mousavi, M. S., and Fathy, M. (2006). Vehicular ad hoc networks (VANETs): challenges and perspectives. In *ITS Telecommunications Proceedings, 2006 6th International Conference on*, pages 761–766. IEEE.
- Yu, B., Xu, C.-Z., and Guo, M. (2012). Adaptive forwarding delay control for vanet data aggregation. *IEEE Transactions on Parallel and Distributed systems*, 23(1):11–18.
- Yuan, Y., Luo, J., Yan, W., Zhao, T., and Lu, S. (2014). Da2rf: A data aggregation algorithm by restricting forwarders for vanets. In *Computing, Networking and Communications (ICNC), 2014 International Conference on*, pages 393–397. IEEE.