

Classificando Comportamentos Sociais em Redes Veiculares

Davidysson A. Alvarenga¹, Felipe D. Cunha², Aline C. Viana³
Raquel A. F. Mini¹, Antônio A. F. Loureiro²

¹ DCC – Pontifícia Universidade Católica de Minas Gerais (PUC-Minas)

²DCC – Universidade Federal de Minas Gerais (UFMG)

³INRIA – France

davidysson@ufmg.br, fdcunha@dcc.ufmg.br, aline.viana@inria.fr
raquelmini@pucminas.br, loureiro@dcc.ufmg.br

Abstract. *Vehicular networks are networks comprised by vehicles traveling cities and highways. During their trajectories, these vehicles interact with other vehicles in order to make safer and enjoyable traffic. These interactions may be influenced by several factors. To mention a few: vehicle speed, roads condition and time of day. Moreover, driver behavior and its interests can influence in many vehicle features. In this context, the Vehicular Social Networks arise as a new perspective to vehicular networks. These are networks in which vehicles “socialize” and share common interests. In this work, we evaluate the behavior of vehicles in two mobility scenarios, in order to classify them according to the interactions performed, identifying common interests and similar routines. Thus, we use metrics of complex networks and statistical techniques. Results prove the existence of routines and human features in Vehicular Networks.*

Resumo. *Redes veiculares são redes formadas por veículos que transitam nas cidades e rodovias. Durante suas trajetórias estes veículos interagem uns com outros veículos com o objetivo de tornar o tráfego mais seguro e a viagem agradável. Essas interações podem ser influenciadas por diversos fatores. Dentre eles pode-se citar: a velocidade do veículo, as condições das vias e o período do dia. Além disso, o comportamento do condutor e seus interesses podem influenciar em muitas características do veículo. Nesse contexto, surgem as Redes Sociais Veiculares como uma nova perspectiva para as redes veiculares. Essas são redes nas quais os veículos “socializam” e compartilham interesses comuns. Neste trabalho serão avaliados o comportamento dos veículos em dois cenários de mobilidade, com o objetivo de classificá-los de acordo com as interações realizadas, identificando interesses comuns e rotinas semelhantes. Para isso, foram utilizadas métricas de redes complexas e técnicas estatísticas. Resultados comprovam a existência de rotinas e propriedades sociais nas Redes Veiculares.*

1. Introdução

Redes Veiculares (*Vehicular Networks* - VANETs) são um tipo de rede formadas por veículos, que interagem entre si diretamente ou por meio de infraestrutura instaladas nas vias públicas. Esses veículos possuem capacidade de processamento, comunicação e sensoriamento. No entanto, devido à grande mobilidade dos veículos, estas redes possuem uma alta variação na sua topologia. Outros fatores que cooperam para esta variação são os limites de velocidade e direção impostos aos veículos pelas vias públicas.

Durante suas trajetórias os veículos interagem entre si, e diversos fatores podem influenciar nestas interações. Dentre eles podemos citar: o horário do dia, a velocidade dos veículos, o local, engarrafamentos, a existência de eventos, etc. Por exemplo, em horários de pico (tráfego denso), as vias públicas podem ser vistas como um lugar popular, onde é mais suscetível ocorrer interações. Outro fator que pode influenciar nas interações é o comportamento do condutor e seus interesses comuns com outros condutores. Este traz consigo uma nova perspectiva para as redes veiculares: as Redes Sociais Veiculares (*Vehicular Social Networks* - VSNs) [25].

Neste contexto, este trabalho se propõe a investigar as VSNs, estudando a influência do comportamento humano na mobilidade veicular, identificando padrões, rotinas semelhantes e interesses comuns. Além disso, pretende-se discutir as melhorias que esse conhecimento pode oferecer no estabelecimento das redes veiculares. Para isso, foram analisadas duas bases de dados de veículos usando métricas de redes complexas e técnicas estatísticas. Os resultados apontam que é possível identificar classes de comportamentos nas redes sociais veiculares.

O restante do trabalho está organizado da seguinte forma: a Seção 2 apresenta um breve resumo dos trabalhos relacionados encontrados na literatura. A Seção 3 descreve a metodologia utilizada neste trabalho, as bases de dados selecionadas e as métricas e técnicas utilizadas. A Seção 4 apresenta os resultados quantitativos, de acordo com cada métrica avaliada. A Seção 5 discute sobre como aplicar estas métricas na elaboração de novos serviços para VANETs. Finalmente, a Seção 6 apresenta as conclusões e as perspectivas futuras.

2. Trabalhos Relacionados

Estudos recentes tem comprovado a influência do comportamento humano na mobilidade dos veículos [9, 12, 5]. Nesses estudos pode-se averiguar como as interações ocorrem entre os veículos, a frequência dessas interações, os horários de maior probabilidade de ocorrência e também como utilizar melhor essas interações em projetos de novos serviços e aplicações para redes veiculares.

Normalmente esses estudos se apoiam na teoria de redes complexas, na análise de traces de mobilidade de veículos e em técnicas estatísticas [21]. Assim, os autores exploram a avaliação das métricas em grafos que representam contatos entre os veículos numa determinada janela de tempo de avaliação. A Tabela 1 apresenta os trabalhos que caracterizaram cenários de mobilidade veicular e que buscam entender melhor a evolução topológica da rede.

Pode-se observar que em [9, 20, 15], os autores avaliaram apenas pequenos cenários analíticos, nos quais foram utilizados geradores sintéticos de mobilidade. De forma mais ampla, o comportamento urbano das grandes cidades foi estudado em [18, 23, 16], nos quais os autores avaliaram as características da mobilidade realística dos veículos que circulam dentro de alguns municípios.

Por outro lado, em [13, 12] os autores estudaram cenários com a mobilidade de táxis reais em algumas cidades, extraindo propriedades e comportamentos da rede. Entretanto, por questões de privacidade, na literatura não se encontra dados reais de veículos particulares, o que pode impossibilita uma melhor extração de comportamentos humanos. Desta forma, em [22], os autores constroem e analisam a estrutura da rede utilizando fotografias de algumas regiões das cidades.

Autores	Traces Avaliados	Métricas Avaliadas	Observações
Fiore e Härrri (2008) [9]	Cenários analíticos	Duração de encontros, densidade, grau dos vértices, <i>clusters</i> e coeficiente de agrupamento	modelos distintos; raio fixo na comunicação; não utiliza grafos temporais e aleatórios; área geográfica delimitada; duração limitada.
Pallis et (2009) [18]	Zurique	Grau dos vértices, duração de encontros, diâmetro, centralidade de vértices, <i>clusters</i> , coeficiente de agrupamento, comunidades e correlação entre as métricas	Varia o raio da comunicação, não utiliza grafos temporais e aleatórios; Avalia apenas algumas horas do dia; área geográfica delimitada.
Loulloudes, Pallis e Dikaiakos (2010) [13]	Xangai e Los Angeles	grau dos vértices, diâmetro, densidade, centralidade de vértices, duração e frequência de encontros, <i>clusters</i> , coeficiente de agrupamento e comunidades	Mobilidade realística em Los Angeles; não varia o raio da comunicação; não utiliza grafos temporais e aleatórios; área geográfica delimitada; duração limitada.
Rezende et al. (2011) [20]	Cenários analíticos	Diâmetro, coeficiente de agrupamento e grau dos vértices	Região urbana/rodovia; densidade fixa; não varia o raio de comunicação; não compara com grafos aleatórios; varia o tamanho da janela temporal; área geográfica delimitada; duração limitada.
Liu et al. (2012) [12]	São Francisco e Xangai	Grau dos vértices, distância média e coeficiente de agrupamento	Seleciona apenas o maior componente do grafo; não varia o raio da comunicação; compara com grafos aleatórios; não teve avaliação de janelas temporais.
Uppoor, Naboulsi e Fiore (2012-2013) [23, 16]	Colônia	Densidade e distribuição de veículos, tempo de viagem, distância percorrida, duração de encontros componentes, grau dos vértices, centralidade de intermediação e assortatividade	Analiza apenas sobre o tráfego de veículos; varia o raio de comunicação; compara com grafos aleatórios; não varia o tamanho da janela temporal, duração limitada.
Thakur, Hui e Helmy (2013) [22]	Connecticut, Londres, Seattle, Sydney, Toronto e Washington	Coeficiente de agrupamento, densidade, grau dos vértices e distância média	Modelagem dos grafos utilizando <i>webcams</i> como vértices e as vias como arestas; não avalia variação no raio de comunicação; compara com grafos aleatórios; não varia o tamanho da janela temporal; avalia apenas quatro horários do dia.
Monteiro et al. (2013) [15]	Cenários analíticos	Grau dos vértices, distância média, coeficiente de agrupamento e conectividade	Região urbana/rodovias; Variação de densidade; Raio de comunicação fixo; Não utiliza grafos temporais e aleatórios; área geográfica delimitada; duração limitada.
Cunha et al. (2013-2014) [5, 4]	São Francisco e Zurique	Grau dos vértices, distância média, coeficiente de agrupamento, densidade e persistência de arestas	Raio de comunicação fixo; compara com grafos aleatórios; não varia tamanho da janela temporal.

Tabela 1. Trabalhos recentes encontrados na literatura que apresentam caracterização da mobilidade veicular.

Contudo, parte destes trabalhos apresentam algumas limitações. Alguns não analisaram as bases em sua plenitude, uma vez que isolaram a área central das cidades, ignorando as regiões periféricas. Outros computaram poucas métricas, não relacionando os resultados com outros modelos. Estudos recentes mostraram que, com a análise de todas essas bases de dados, é possível encontrar características humanas no comportamento dos veículos [5, 4]. Apesar destas análises estarem apoiadas em bases de dados de táxis ou mesmo em bases sintéticas, foi possível verificar a existência de interesses comuns e comportamentos similares em uma rede veicular. Esses comportamentos reforçam a ideia das Redes Sociais Veiculares [25], redes nas quais os veículos interagem, compartilham interesses comuns e comportamentos semelhantes.

3. Metodologia

Nesta seção serão apresentados os detalhes de toda a metodologia utilizada na avaliação das bases de dados. O foco dessa avaliação é entender melhor como acontecem as interações entre os veículos, destacando os comportamentos e as características sociais dos mesmos. Em primeiro momento, será descrito os detalhes das bases de dados avaliadas. Em seguida, será descrito o processo de geração dos grafos e as métricas utilizadas.

3.1. Bases de Dados Avaliadas

A seleção das bases de dados utilizadas neste trabalho foi definida por critérios que envolveram: mobilidade diversificada ao longo das 24 horas do dia, registros em mais de um dia da semana, granularidade semelhante, quantidade semelhante de veículos em circulação por faixas horárias e disponibilidade pública. A primeira base de dados foi obtida a partir da extração de coordenadas de localização geográfica durante a simulação de um modelo de mobilidade realística na cidade de Helsinque. A segunda base de dados contém dados reais da mobilidade de táxis dentro da cidade de Roma.

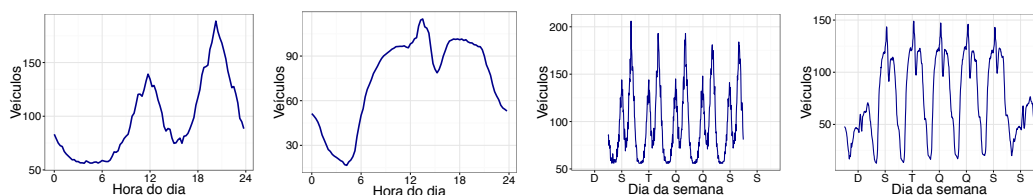
Parâmetro	Valores
Pessoas	2 mil
Grupos	8 grupos em 4 distritos
Carro próprio	50%
Locais	Casa, trabalho e áreas de lazer
Meios de transporte	Ônibus, carro ou táxis
Simulador	The ONE
Filtro utilizado	Apenas veículos

Tabela 2. Parâmetros do modelo *Working Day Movement Model*.

- **Helsinki:** representa a mobilidade realística gerada pelo modelo *Working Day Movement Model* (WDM) [7]. Assim, a mobilidade de 2.000 pessoas foi gerada utilizando veículos diversos (carros, ônibus ou táxis), se locomovendo entre 3 locais distintos (residência, trabalho e pontos de lazer), ao longo das 24 horas do dia, dentro de um intervalo de 1 semana útil. A simulação do tráfego foi realizada no *software* The ONE, utilizando os parâmetros descritos na Tabela 2.
- **Roma:** esta base de dados contém a mobilidade real de 320 táxis que trabalham na região central da cidade. Os registros foram coletados durante 4 semanas em 2014 [2].

A Figura 1 apresenta a evolução do tráfego diário e semanal para as duas bases. É possível observar um trânsito maior durante os horários de pico, com uma redução no horário de almoço e nas madrugadas. Além disso, para a base de Helsinki é possível verificar que não existe tráfego durante os finais de semana. Este é um modelo que considera apenas dias úteis, desprezando finais de semana e feriados. Ao olhar para o tráfego semanal em Roma, nota-se uma redução aos finais de semana, o que pode ser explicado pelo contexto dos dias.

Analisando a evolução do tráfego sob as cidades, a Figura 2 apresenta o *heatmap* do tráfego. Pode-se destacar Roma, onde às 6 am. o tráfego é bem esparso, e o mesmo apresenta-se muito mais denso às 6 pm, onde os veículos passam a utilizar vias paralelas aos congestionamentos. Além disso, nota-se que as duas cidades possuem características de tráfego bem distintas para os mesmos horários do dia. Esse é um comportamento que indica características peculiares em cada cidade, e que as mesmas devem ser consideradas no projeto de soluções para essas redes.



(a) Helsinki diário. (b) Rome diário. (c) Helsinki semanal. (d) Roma semanal.

Figura 1. Evolução do tráfego ao longo do dia e da semana.

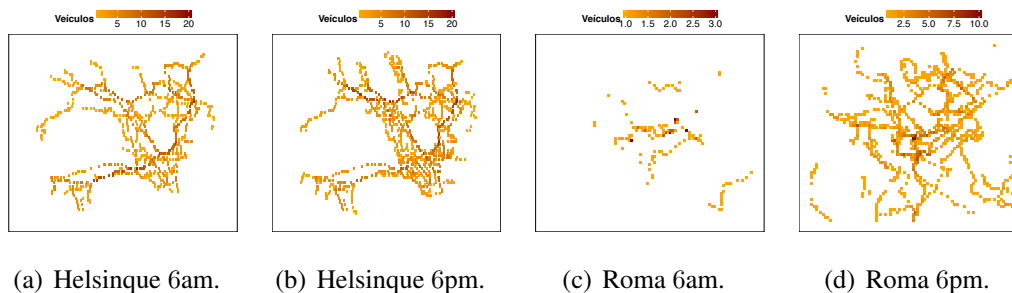


Figura 2. Mapa de calor representando a mobilidade de veículos em horários de pico.

3.2. Grafos Temporais

Um grafo temporal no tempo t é um grafo não dirigido, definido como um grafo $G(t) = (V, E)$, onde V representa o conjunto com todos os veículos V_i e E representa o conjunto de arestas E_{ij} . Em $G(t)$, existe uma aresta $E_{ij}(t)$ entre os veículos V_i e V_j durante o tempo t , se $i \neq j$. A avaliação de cada métrica deve levar em conta cada um dos grafos temporais $G(t)$, com as interações que acontecem em cada período t . Por exemplo, se t estiver definido para 15 minutos, os registros ocorridos entre 08:00 – 08:14 serão agrupados em t_0 e os registros ocorridos entre 08:15 – 08:29 horas serão agrupados em t_1 .

Durante a criação dos grafos temporais, o critério para inserção de arestas entre dois vértices é definido pela distância entre as coordenadas dos veículos. Se essa distância for menor ou igual ao valor do raio da comunicação, os vértices serão conectados. Todas as arestas que pertencerem ao mesmo intervalo de avaliação serão agrupadas em um grafo $G(t)$. Na direção de entender melhor as características destas arestas e verificar as propriedades existentes, cada grafo $G(t)$ será comparado a outro grafo aleatório G^R gerado pelo algoritmo de urna [11], com mesma quantidade de arestas e distribuição de graus dos vértices. Por meio deste processo, foi possível verificar a existência de estruturas sociais no grafo temporal, que é inexistente no grafo aleatório.

3.3. Métricas Avaliadas

Nesta seção serão descritas as métricas escolhidas para avaliação das bases de dados. Dentre as métricas conhecidas de redes complexas [3], neste trabalho optou-se por agrupá-las de forma que as métricas que representam o mesmo comportamento pertençam a mesma categoria. Assim, quatro grupos foram definidos: métricas que mensuram a influência dos vértices sob seus vizinhos, métricas relacionadas ao fenômeno “Mundo Pequeno”, métricas de centralidade e métricas de semelhança de contatos. Os valores finais de cada métrica foram calculados utilizando o valor médio observado em cada intervalo de tempo analisado.

Vértices vs. Vizinhos: O grau do vértice representa a quantidade de interações que o mesmo faz com sua vizinhança. Além disso, vértices que apresentam grau elevado são os mais ativos na rede, e podem atuar como um canal para a troca de informações [6]. Tal métrica pode ser vista como uma medida de centralidade, indicando a popularidade do vértice no grafo. Em um nível global, um alto grau pode indicar uma alta concentração de veículos próximos uns aos outros, comum em horários de pico ou engarrafamentos. Por outro lado, o grau médio dos vizinhos (*average neighbor degree*) representa a média dos graus de todos vizinhos que cada vértice possui. A avaliação da sua correlação com os

graus dos vértices mostra a tendência de um vértice se ligar a outros com graus similares. Nesse sentido, a rede é considerada como assortativa quando a correlação for crescente. Nessas redes, veículos que compartilham rotinas estarão ligados uns aos outros e com graus semelhantes [17].

Outra importante métrica que se pode avaliar no contexto de comportamentos semelhantes é a presença de comunidades no grafo. No contexto de redes veiculares, uma comunidade representa indivíduos que possuem interesses comuns. No grafo pode-se destacar um subconjunto de vértices fortemente conectados entre si. Neste trabalho, para avaliação dessa métrica optou-se em executar o algoritmo *Walktrap* [19], aplicado em cada componente existente no grafo temporal $G(t)$ a fim de detectar conjunto de veículos que formam as comunidades. Ao quantificar as conexões entre os veículos, pode-se avaliar também a densidade de arestas no grafo. Essa pode apresentar valores elevados em regiões urbanas em relação às áreas periféricas ou rurais, devido ao tráfego lento ou à existência de engarrafamentos. Valores baixos de densidade ajudam a identificar regiões cuja conectividade pode ser interrompida mais facilmente, exigindo estratégias que busquem minimizar tal problema.

Centralidade: A centralidade de um vértice no grafo indica o quão importante aquele vértice é em relação aos demais. Do ponto de vista de distância e comunicação, o quão mais central um vértice for, mais caminhos mínimos passarão por ele e mais comunicações o mesmo irá intermediar. Uma forma de definir essa centralidade é a métrica centralidade de intermediação (*betweenness centrality*), que representa a frequência com que um veículo faz parte de um caminho mais curto entre quaisquer dois outros veículos. Essa métrica é importante em situações nas quais dois vértices não são vizinhos e dependem dos outros vértices para realizar sua troca de dados. A posição da intermediação indicará a influência do vértice ao levar em consideração o controle exercido por ele sobre as interações entre os dois outros.

Mundo Pequeno: O fenômeno “mundo pequeno” (*small-world*) [14] define a premissa que qualquer pessoa pode se conectar à outra por meio de uma sequência de relacionamentos em um grau médio de 5 a 6 pessoas [14, 26], o que numa rede representaria o número de conexões que separam duas pessoas. Além disso, redes que apresentam esse fenômeno possuem uma grande tendência de formar comunidades, ou seja, formar agrupamentos de nós que compartilham interesses comuns. Para avaliar e classificar uma rede como “mundo pequeno” utiliza-se as métricas coeficiente de agrupamento e a distância média. Assim, essas redes tendem a apresentar uma distância média baixa e um elevado coeficiente de agrupamento. A distância média representa o número de saltos necessários para a conexão entre um par quaisquer de veículos. Valores altos podem indicar atrasos em transmissões de dados, pois os veículos estão fisicamente distantes, fora do raio da comunicação, e provavelmente não visitam locais próximos. Por outro lado, valores baixos indicam lugares próximos ou pontos em comum em suas rotas, podendo representar interesses comuns. Considerando o coeficiente de agrupamento, nas redes veiculares, os veículos que possuem interesses comuns tendem a ficar mais agrupados, o que contribui para um aumento no valor do coeficiente de agrupamento.

Semelhança de Contatos: No contexto das redes sociais veiculares, a semelhança de contatos pode ser avaliada por duas métricas: a sobreposição topológica e a persistência de arestas. A sobreposição topológica (*topological overlap*) retrata a tendência de existir vértices com vizinhos partilhados, enquanto que a persistência de arestas (*edge persistence*) define o número de vezes que dois vértices se encontram na mesma janela temporal,

ao longo do período. Esta métrica permite identificar as relações regulares entre entidades. De acordo com [24], a avaliação destas métricas permite a classificação das interações em quatro tipos de relacionamentos. Essa classificação baseia-se na comparação com grafos aleatórios, de forma que é possível quantificar se os relacionamentos em uma rede veicular são aleatórios ou obedecem a um padrão. Como observado na Tabela 3, na classe de Amigos (*friends*), são encontradas características sociais sobre rotinas semelhantes entre os veículos tanto na persistência de arestas quanto na sobreposição topológica. Na classe de Conhecidos (*acquaintance*), essa característica é encontrada somente na sobreposição topológica. Na classe de Pontes (*bridges*), a característica social é encontrada somente na persistência de arestas. E finalmente, na classe Aleatória (*random*), não há presença de tais características nas duas métricas.

Classes	Sobreposição Topológica	Persistência de Arestas
Amigos	Social	Social
Conhecidos	Social	Aleatório
Pontes	Aleatório	Social
Aleatória	Aleatório	Aleatório

Tabela 3. Classificação da semelhança entre os contatos [24].

4. Resultados

Nesta seção serão apresentados os resultados da análise efetuada sobre as bases de dados. Para essa análise foi considerada a metodologia e as métricas definidas na Seção 3. Primeiramente, a Seção 4.1 define os parâmetros e detalhes da análise. Em seguida, os resultados serão apresentados de acordo com as propriedades definidas na Seção 3.3.

4.1. Parâmetros

A análise dos dados foi efetuada considerando os valores ao longo do dia e da semana. O objetivo principal desta análise é identificar comportamentos e propriedades que descrevem rotinas e interesses dos condutores numa rede veicular. A partir desse entendimento será possível propor serviços e protocolos que atendam melhor a demanda dos seus usuários.

Com o intuito de verificar o impacto da avaliação nas interações entre os veículos, neste trabalho o raio de comunicação foi configurado de acordo com definições do protocolo 802.11p. No entanto, foi avaliada uma variação no raio da comunicação, com valores de 100 e 150 metros, a fim de verificar o impacto nos resultados, uma vez que um incremento em seu valor pode agregar mais contatos antes inalcançáveis.

Considerando a janela de tempo para agregação de contatos, foram considerados os seguintes valores: 5, 15, 30 e 60 minutos, conforme discussões realizadas em trabalhos relacionados [8, 10]. Ao considerar valores pequenos de janela, foi possível encontrar um grafo mais esparsos com muitos componentes. Em contrapartida, janelas de tempo com grande intervalo agrupam uma grande quantidade de contatos, o que pode gerar grafos mais densos, retratando um comportamento homogêneo nas métricas. Com isso, a variação é feita para mostrar como os resultados das métricas podem ser influenciados por tal parâmetro.

A Figura 3 apresenta o gráfico analisando o tamanho da janela versus número de componentes e densidade de arestas do grafo. É possível observar que o valor de 15

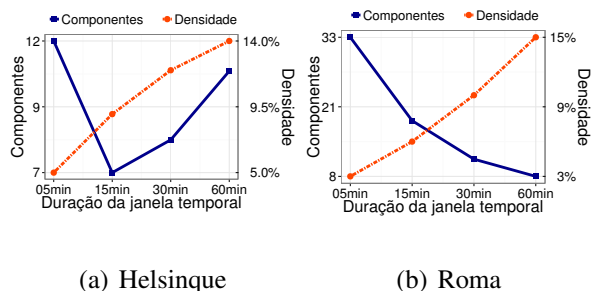


Figura 3. Relação entre densidade de arestas e componentes no grafo.

minutos apresenta uma boa proporção de arestas e componentes no grafo. No intuito de melhor caracterizar as variações de comportamentos, neste trabalho optou-se por definir o tamanho de 15 minutos como padrão para o tamanho da janela de tempo.

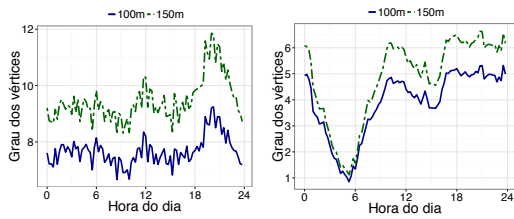
Além disso, a correlação entre algumas métricas foi avaliada. O objetivo desta análise é identificar o quão dependente uma métrica é de outra, e se as métricas apresentam o mesmo comportamento, quando o tráfego varia. Para isso, foi utilizada a Correlação de *Pearson* que mede o grau da relação linear entre duas variáveis, indicando a direção da correlação dentro do intervalo de -1 a 1 . Onde o valor é positivo quando uma variável aumenta junto com a outra, ou negativo, quando uma diminui à medida que a outra aumenta. Com isso, o nível da correlação pode ser classificado por três intervalos: fraco (de $0,0$ a $\pm 0,3$), moderado (de $\pm 0,3$ a $\pm 0,7$) e forte (de $\pm 0,7$ a $\pm 1,0$) [1].

4.2. Vértices vs. Vizinhos

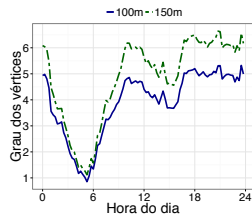
Ao avaliar o grau dos vértices, é importante verificar que essa métrica sofre variações ao longo do dia, diante das mudanças na quantidade de veículos em circulação. A Figura 4 apresenta o seu valor médio. Observa-se em Helsinque, Fig. 4-(a), o valor do grau médio permanece constante durante o dia e aumenta ao final, momento de maior densidade de veículos em circulação. Em contrapartida, em Roma (Fig. 4-(b)), observa-se que o valor do grau acompanha a evolução do tráfego. Por serem táxis, veículos que circulam muito, os valores tendem a ser mais semelhantes. Nas duas bases, o aumento do raio da comunicação agregou mais contatos e, conseqüentemente, o número médio dos graus.

Quando se analisa a correlação entre os graus dos vértices e seus vizinhos, a Figura 5 mostra que os vizinhos dos vértices possuem graus parecidos com os graus dos vértices, uma vez que o Coeficiente de *Pearson* apresentou índices de forte correlação positiva. Isso indica a presença de comportamentos semelhantes dentro das redes, podendo classificá-las como redes assortativas. Redes nas quais os vértices possuem uma forte tendência de encontrar vértices altamente com graus semelhantes conectados entre si.

Quando se verifica a presença de interesses comuns na rede através da busca por comunidades no grafo, a Figura 6, produzida utilizando o algoritmo *Walktrap* [19], mostra que os grafos temporais apresentam comunidades bem definidas. Quando se compara com os grafos aleatórios é possível verificar que a aleatoriedade na geração das arestas quebra as comunidades existentes, deixando comunidades mal definidas e vértices isolados.

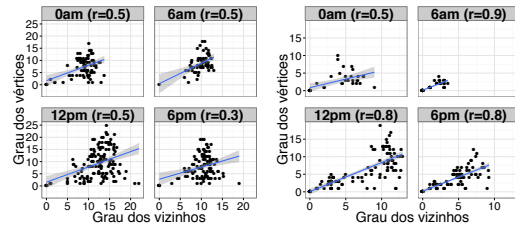


(a) Helsinque



(b) Roma

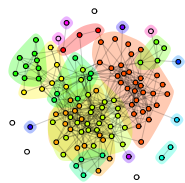
Figura 4. Grau médio dos vértices variado o raio de comunicação.



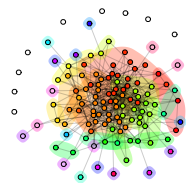
(a) Helsinque

(b) Roma

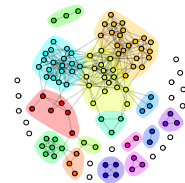
Figura 5. Correlação entre o grau dos vértices e o grau dos vizinhos.



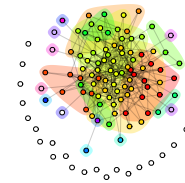
(a) Helsinque real.



(b) Helsinque aleatório.



(c) Roma real.



(d) Roma aleatório.

Figura 6. Comunidades nos grafos temporais e aleatórios.

4.3. Centralidade dos vértices

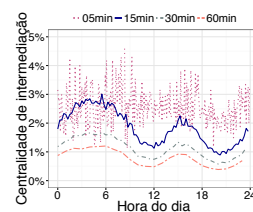
A centralidade de um veículo define a distância dele em relação a sua distância para os demais veículos do grafo. O quão mais central for o veículo, menor será a sua distância para os demais. A Figura 7-(a) e (b) apresenta os valores para centralidade de intermediação (*betweenness*). Pode-se observar que os valores dos grafos reais são maiores que dos grafos aleatórios, o que comprova que os grafos aleatórios descaracterizam as conexões dos grafos originais. Quando se observa o comportamento destes valores ao variar o tamanho da janela na Figura 7-(c) e (d), nota-se que valores menores para a janela de avaliação apresentam maiores valores de centralidade. Isso ocorre, pois, com janelas menores, a rede apresenta mais desconexões e uma quantidade maior de pequenos componentes.



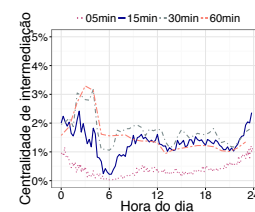
(a) Helsinque



(b) Roma



(c) Helsinque



(d) Roma

Figura 7. Centralidade de intermediação variando o tamanho da janela temporal.

4.4. Mundo pequeno

As Figuras 8 e 9 apresentam respectivamente os valores para distância média e coeficiente de agrupamento. Essas métricas retratam a presença de características sociais com rotinas semelhantes e interesses em comum. Baixos valores pra distância e altos valores pra o coeficiente de agrupamento indicam que a rede apresenta o fenômeno *small-world*, uma vez que existem vértices que formam atalhos na rede reduzindo a média da distância global e existe presença de comunidades formando grupos com interesses em comum.

Analisando a distância média, Figuras 8-(a) a (d), observa-se que os grafos aleatórios apresentam valores um pouco menor que as bases reais, devido à ruptura das conexões que ocorre no grafo aleatório. Ao variar o raio de comunicação, observa-se que a existência de mais contatos não impacta no valor da distância média em nenhuma das bases. Nas Figuras 9-(a) e (b) são apresentados os valores do coeficiente de aglomeração para Helsinque e Roma. Pode-se notar que as duas bases apresentam valores maiores quando comparados a grafos aleatórios. Nota-se também que durante as madrugadas os valores são mais altos indicando que os veículos circulam formando grupos provavelmente utilizando as mesmas vias. Pode-se observar também, nas Figuras 9-(c) e (d) a variação dos valores de coeficiente em cada rede. Nota-se que a rede Helsinque, por representar mobilidade de um dia de trabalho apresentam valores maiores do coeficiente quando comparados com a base de Roma. Esta apresenta uma grande variabilidade de valores, alternando entre 20% a 40%.

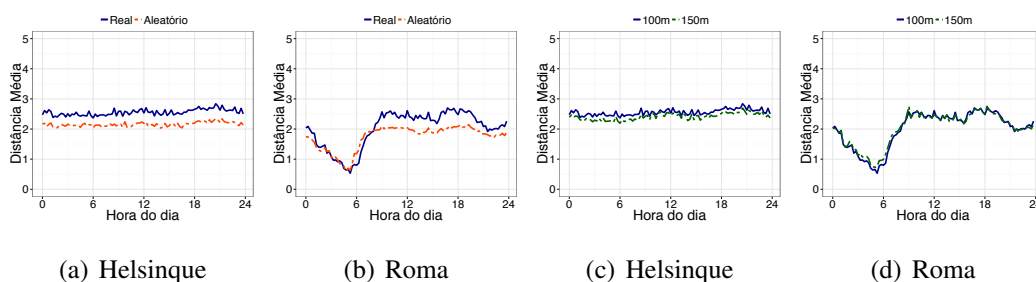


Figura 8. Distância média entre os vértices ao longo do dia.

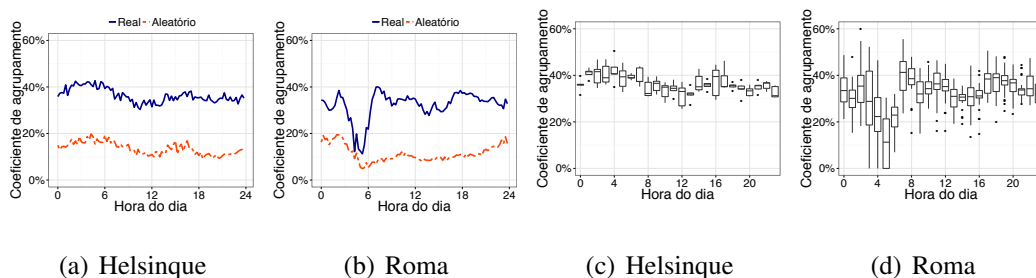


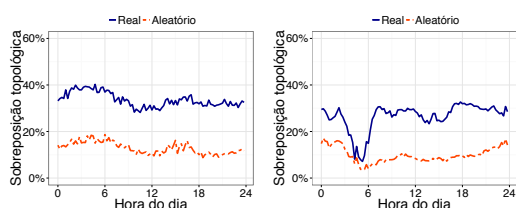
Figura 9. Coeficiente de agrupamento ao longo do dia.

4.5. Semelhança de contatos

A análise da semelhança de contatos pode constatar que ambos os cenários apresentaram rotinas semelhantes e interesses em comum durante o processo de avaliação das métricas

de sobreposição topológica e persistência de arestas, apresentadas nas Figuras 10 e 11. Os resultados obtidos na avaliação da sobreposição topológica nos grafos temporais foram maiores que os grafos aleatórios, da mesma forma que o coeficiente de agrupamento apresentado na Figura 9. Isso indica a existência de grupos de vértices que compartilham interesses em comum. Entretanto ao analisar a repetição dos encontros, Fig. 11-(a) e (b), nota-se que os grafos temporais em Roma apresentaram valores menores que os grafos aleatórios, enquanto que em Helsinque os resultados maiores, indicando que os veículos pessoais repetem mais encontros nos mesmos horários ao longo dos dias, o que não ocorre com táxis. Esses são veículos que retratam o interesse de várias pessoas no mesmo veículo.

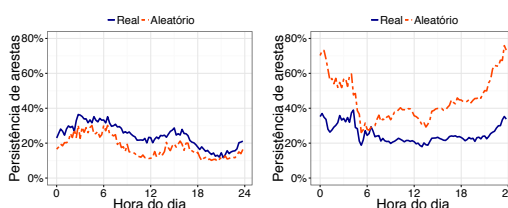
Assim, a Tabela 4 indica que a base de Helsinque pode ser classificada como uma rede de amigos, pois ela apresenta características sociais com rotinas semelhantes tanto na sobreposição topológica e quanto na persistência de arestas. Por outro lado, o cenário de Roma pode ser classificado como uma rede de conhecidos, uma vez que os resultados dos grafos temporais não resultaram valores melhores que os grafos aleatórios na avaliação da persistência de arestas.



(a) Helsinque

(b) Roma

Figura 10. Sobreposição topológica ao longo do dia.



(a) Helsinque

(b) Roma

Figura 11. Repetição de encontros ao longo do dia.

Base de dados	Sobreposição Topológica	Persistência de Arestas	Classe
Helsinque	Social	Social	Amigos
Roma	Social	Aleatório	Conhecidos

Tabela 4. Classificação das interações social/aleatório baseado na classificação proposta em [24].

4.6. Correlações entre métricas

Com o objetivo de verificar a correlação entre as métricas e assim identificar métricas que possuem o mesmo comportamento, a Figura 12-(a) e (b) apresenta os resultados da correlação de *Pearson* para as métricas sobreposição topológica, coeficiente de aglomeração, grau dos vértices e centralidade de intermediação. Interessante observar que as métricas sobreposição topológica e coeficiente de aglomeração apresentam uma correlação forte, em torno de 0.9 e 1. Ambas são métricas que retratam a presença de interesses comuns entre os vértices, o que explica esse resultado. Entretanto, com as demais métricas a correlação do coeficiente de clusterização e sobreposição topológica é baixa.

Outra correlação pode ser observada entre a centralidade de intermediação e o grau do nó. De fato, quanto maior for o grau de um nó, maior também será a chance de

se ter mais caminhos passando por ele. Comportamento esse explicado pelos valores de 0.7 a 0.9 de correlação entre eles.

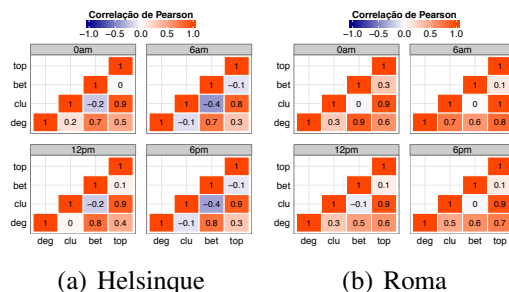


Figura 12. Grade de correlações entre as métricas.

5. Aplicações em VSNs

Este trabalho apresenta uma análise de duas bases sociais que descrevem a mobilidade de veículos em um cenário, caracterizando-as como uma rede social veicular com o objetivo de compreender melhor as interações entre os veículos, seus comportamentos e as rotinas dos condutores. Entretanto, esta análise foi executada considerando parâmetros e suposições que caracterizam as interações entre os veículos. Através destas interações foram gerados grafos e computadas métricas de redes complexas. Assim, pelos resultados foi possível identificar propriedades e características que indicam fortemente a presença de comportamentos sociais nas redes veiculares.

Muitas são as possibilidades de aplicar essas propriedades e comportamentos no projeto de soluções para rede veicular. No projeto de infraestrutura para a rede, pode-se fazer o uso de métricas como densidade de arestas e distância que ajudam na deposição de ponto de acesso na direção de melhorar a cobertura da rede. Através da densidade de arestas por região é possível identificar locais mais visitados, e pela distância média é possível mensurar a latência na transmissão. Em outra perspectiva, é possível projetar protocolos de comunicação cientes da variação do tráfego e da probabilidade de encontros entre os veículos fazendo uso da métrica grau dos nós [15]. Esta métrica pode auxiliar na definição da probabilidade de retransmitir o dado, o que pode garantir uma melhor cobertura.

Em outra direção, é possível aplicar as métricas para o projeto de aplicações que venham auxiliar os condutores durante o trajeto, tornando a viagem mais segura e agradável. Métricas como coeficiente de agrupamento e sobreposição topológica podem ser usadas para identificar condutores com interesses semelhantes e que visitam os mesmos lugares para o projeto de aplicativos que incentivam caronas e compartilhamento de veículos. Além disso, baseado-se no histórico de visitação será possível criar aplicativos para sugestão de destinos.

6. Conclusões e Trabalhos Futuros

Este trabalho apresentou uma caracterização de duas bases de dados que representam a mobilidade de veículos, buscando encontrar propriedades e características que retratam a interferência do comportamento humano na mobilidade veicular. A base realística representa a mobilidade de veículos pessoais sob o modelo *Working Day Movement Model*,

e a segunda base representa a mobilidade de táxis na cidade de Roma. Esta análise foi realizada considerando os conceitos e métricas de redes complexas e técnicas estatísticas.

Durante a análise das bases, foi possível identificar características semelhantes às redes sociais. Apesar das bases serem de naturezas distintas, onde não há como fazer uma comparação direta entre os resultados, diversas métricas analisadas mostraram a existência de grupos de veículos formando comunidades de pessoas com indícios de rotinas semelhantes e interesses comuns. Estes comportamentos podem ser explicados pela existência de veículos que frequentam os mesmos lugares nos mesmos períodos do dia. Além disso, os resultados mostraram que o cenário de mobilidade realística apresenta características de uma rede de amigos: onde há grupos de pessoas que se encontram com uma frequência. Em contrapartida, o cenário com mobilidade de táxis em Roma apresentou características de uma rede de pessoas conhecidas: onde há grupos de pessoas que compartilham interesses, mas seus encontros não são frequentes.

Em trabalhos futuros, pretende-se avaliar os resultados das interações entre os veículos do ponto de vista espacial, explorando mais as interações entre eles e os locais visitados. Também pretende-se avaliar outras bases com um número maior de veículos em circulação, efetuando filtros e ajustes espaciais e temporais entre os registros, com o objetivo de possibilitar uma comparação de bases de mesma natureza. Além disso, pretende-se incluir análises de estabilidade entre as comunidades para verificar como é o comportamento delas ao longo do tempo, gerando janelas deslizantes na agregação de contatos temporais. Por fim, pode-se estender esse trabalho buscando aplicar o uso das métricas para o projeto de novas soluções para redes veiculares.

Agradecimentos

Os autores agradecem CAPES, CNPq e FAPEMIG pelo apoio financeiro.

Referências

- [1] Richard A Becker, John M Chambers, and Allan R Wilks. The new s language. *Pacific Grove, Ca.: Wadsworth & Brooks, 1988*, 1, 1988.
- [2] Lorenzo Bracciale, Marco Bonola, Pierpaolo Loreti, Giuseppe Bianchi, Raul Amici, and Antonello Rabuffi. CRAWDAD data set roma/taxi (v. 2014-07-17). Downloaded from <http://crawdad.org/roma/taxi/>, July 2014.
- [3] L da F Costa, Francisco A Rodrigues, Gonzalo Travieso, and Paulino Ribeiro Villas Boas. Characterization of complex networks: A survey of measurements. *Advances in Physics*, 56(1):167–242, 2007.
- [4] Felipe D Cunha, Aline Carneiro Vianna, Raquel A F Mini, and Antonio A F Loureiro. Is it possible to find social properties in vehicular networks? In *Computers and Communication (ISCC), 2014 IEEE Symposium on*, pages 1–6, June 2014.
- [5] Felipe D Cunha, Aline Carneiro Vianna, Raquel A F Mini, and Antonio A F Loureiro. How effective is to look at a vehicular network under a social perception? In *Wireless and Mobile Computing, Networking and Communications (WiMob), 2013 IEEE 9th International Conference on*, pages 154–159. IEEE, 2013.
- [6] Elizabeth M Daly and Mads Haahr. Social network analysis for information flow in disconnected delay-tolerant manets. *Mobile Computing, IEEE Transactions on*, 8(5):606–621, 2009.
- [7] Frans Ekman, Ari Keränen, Jouni Karvo, and Jörg Ott. Working day movement model. In *Proceedings of the 1st ACM SIGMOBILE workshop on Mobility models*, pages 33–40. ACM, 2008.

- [8] Tanja Falkowski and Myra Spiliopoulou. Observing dynamics in community structures. In *Symposium on Network Analysis in Natural Sciences and Engineering*, page 15, 2006.
- [9] Marco Fiore and Jérôme Härrri. The networking shape of vehicular mobility. In *Proceedings of the 9th ACM international symposium on Mobile ad hoc networking and computing*, pages 261–272. ACM, 2008.
- [10] Theus Hossmann, Thrasyvoulos Spyropoulos, and Franck Legendre. Know thy neighbor: Towards optimal mapping of contacts to social graphs for dtn routing. In *INFOCOM, 2010 Proceedings IEEE*, pages 1–9. IEEE, 2010.
- [11] Norman Lloyd Johnson and Samuel Kotz. Urn models and their application; an approach to modern discrete probability theory, 1977.
- [12] Xin Liu, Zhuo Li, Wenzhong Li, Sanglu Lu, Xiaoliang Wang, and Daoxu Chen. Exploring social properties in vehicular ad hoc networks. In *Proceedings of the Fourth Asia-Pacific Symposium on Internetware*, page 24. ACM, 2012.
- [13] Nicholas Loulloudes, George Pallis, and Marios D Dikaiakos. The dynamics of vehicular networks in urban environments. *arXiv preprint arXiv:1007.4106*, 2010.
- [14] Stanley Milgram. The small world problem. *Psychology today*, 2(1):60–67, 1967.
- [15] Romeu Monteiro, Susana Sargento, Wantanee Viriyasitavat, and Ozan K Tonguz. Improving vanet protocols via network science. *arXiv preprint arXiv:1301.0965*, 2013.
- [16] Diala Naboulsi and Marco Fiore. On the instantaneous topology of a large-scale urban vehicular network: the cologne case. In *Proceedings of the fourteenth ACM international symposium on Mobile ad hoc networking and computing*, pages 167–176. ACM, 2013.
- [17] Mark E J Newman. Assortative mixing in networks. *Physical review letters*, 89(20):208701, 2002.
- [18] George Pallis, Dimitrios Katsaros, Marios D Dikaiakos, Nicholas Loulloudes, and Leandros Tassioulas. On the structure and evolution of vehicular networks. In *Modeling, Analysis & Simulation of Computer and Telecommunication Systems, 2009. MASCOTS'09. IEEE International Symposium on*, pages 1–10. IEEE, 2009.
- [19] Pascal Pons and Matthieu Latapy. Computing communities in large networks using random walks. In *Computer and Information Sciences-ISCIS 2005*, pages 284–293. Springer, 2005.
- [20] Cristiano Rezende, Azzedine Boukerche, Richard W Pazzi, Bruno P S Rocha, and Antonio A F Loureiro. The impact of mobility on mobile ad hoc networks through the perspective of complex networks. *Journal of Parallel and Distributed Computing*, 71(9):1189–1200, 2011.
- [21] Francisco A Rodrigues. Caracterização, classificação e análise de redes complexas. *Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos*, 2007.
- [22] Gautam S Thakur, Pan Hui, and Ahmed Helmy. The structure and traffic flow anatomy of the planet-scale urban vehicular mobility. *Networking Science*, 3(1-4):13–23, 2013.
- [23] Sandesh Uppoor and Marco Fiore. Insights on metropolitan-scale vehicular mobility from a networking perspective. In *Proceedings of the 4th ACM international workshop on Hot topics in planet-scale measurement*, pages 39–44. ACM, 2012.
- [24] Pedro O S Vaz de Melo, Aline C Viana, Marco Fiore, Katia Jaffrès-Runser, Frédéric Le Mouël, and Antonio A F Loureiro. Recast: Telling apart social and random relationships in dynamic networks. In *Proceedings of the 16th ACM international conference on Modeling, analysis & simulation of wireless and mobile systems*, pages 327–334. ACM, 2013.
- [25] Anna Maria Vegni and Valeria Loscri. A survey on vehicular social networks. *Communications Surveys & Tutorials, IEEE*, 17(4):2397–2419, 2015.
- [26] Duncan J Watts and Steven H Strogatz. Collective dynamics of small-world networks. *nature*, 393(6684):440–442, 1998.