

Análise da Mobilidade em Redes Veiculares Usando Diferentes Tipos de Grafos

Gabriel R. Diniz¹, Felipe D. Cunha², Antonio A. F. Loureiro¹

¹Departamento de Ciência da Computação
Universidade Federal de Minas Gerais (UFMG)
Caixa Postal 702 – 30123-970 – Belo Horizonte – MG – Brasil

²Departamento de Ciência da Computação
Pontifícia Universidade Católica de Minas Gerais
R. Dom José Gáspar, 500 – 30535-901 – Belo Horizonte – MG – Brazil.

{dinizgabriel,loureiro}@dcc.ufmg.br, felipe@pucminas.br

Abstract. *Vehicle networks are a communication system between vehicles that allows the interaction and information exchange. One of the difficulties faced by these applications is the communication due to the frequent change in the network topology. For this reason, several studies have been done in order to understand the traffic and, consequently, the applications that aim to improve the experience of drivers and passengers in a traffic scenario. In the literature, the modeling of the vehicle mobility is often done using an instantaneous, or aggregate or temporal graph. This work compares the impact of these three approaches on the vehicle network analysis and discusses the advantages and disadvantages of each one. The goal is to understand how results may vary depending on the model used.*

Resumo. *As redes veiculares são um sistema de comunicação entre veículos que permite a interação e troca de informações destes automóveis. Uma das dificuldades enfrentadas por essas aplicações é a comunicação, pela mudança frequente na topologia da rede. Por este motivo, diversos estudos têm sido feitos com o intuito de entender o trânsito e, conseqüentemente, as aplicações que visam aprimorar a experiência dos motoristas e passageiros no tráfego. Na literatura, a modelagem da mobilidade veicular é frequentemente feita usando um grafo instantâneo, agregado ou temporal. Este trabalho compara o impacto dessas três abordagens nas análises das redes veiculares, além de discutir as vantagens e desvantagens de cada uma. O objetivo é entender como os resultados podem variar dependendo do modelo utilizado.*

1. Introdução

As redes veiculares permitem a interação e troca de informações entre veículos com o objetivo de melhorar as condições do trânsito e dos passageiros Alves et al. [2009]. Esses objetivos aparecem na forma de aplicações que possuem diversas finalidades para tornar a experiência no tráfego melhor como evitar colisões, diminuir engarrafamentos e promover entretenimento durante a viagem. Nessas aplicações, os veículos agem como sensores, compartilhando informações que podem ser usadas a favor dos motoristas e que permitem a criação de um Sistema de Transporte Inteligente (ITS – *Intelligent Transportation System*),

i.e., um sistema composto por ferramentas que visam prover serviços inovadores na área de transporte.

Uma das dificuldades enfrentadas por essas aplicações é a comunicação. Isso se dá pela frequente mudança na topologia da rede devido a diversos fatores como, por exemplo, a velocidade dos veículos, regras de trânsito a serem seguidas e eventos na pista devido a acidentes e engarrafamentos. Devido a esses e outros aspectos, os contatos que ocorrem entre os veículos costumam ter um tempo curto, dificultando a troca de informações. Para resolver esse problema, existem estudos na literatura que analisam a mobilidade veicular para identificar padrões na movimentação dos veículos. Esses estudos, em sua maioria, utilizam grafos para modelar a mobilidade e analisar a rede induzida pela movimentação dos veículos. Assim, é possível desenvolver soluções apropriadas para troca de informações em cada contato.

Ao observarmos os trabalhos já realizados na área de análise da mobilidade veicular, notam-se três abordagens diferentes de utilização de grafos nessa modelagem, i.e., grafos que modelam a mobilidade de forma instantânea, agregada e temporal. O objetivo deste trabalho é investigar o impacto destas diferentes abordagens encontradas na literatura ao se analisar a mobilidade dos veículos com esses tipos de grafos, em particular na utilização de grafos temporais. Esse é um aspecto fundamental, pois, dependendo do tipo de modelo usado, os resultados podem ser bastante diferentes e fazerem pouco sentido.

Este trabalho analisa esses três modelos presentes na literatura, investigando os resultados gerados por cada um deles. Além disso, o comportamento das métricas muda de acordo com a abordagem utilizada, o que mostra a necessidade de ter esse tipo de análise. Para isso, foram escolhidas duas bases de dados reais de mobilidade veicular onde se realizou a comparação desses modelos. Utilizando métricas de redes complexas e técnicas de estatística, mostramos que as diferentes formas de modelar a mobilidade dos veículos levam a diferentes resultados.

O restante deste trabalho está organizado da seguinte forma: na Seção 2 apresentaremos um resumo dos trabalhos relacionados encontrados na literatura. Na Seção 3, falaremos sobre a metodologia que utilizamos nesse trabalho, quais bases de dados foram trabalhadas e as técnicas utilizadas. Na Seção 4, apresentaremos os resultados de cada métrica avaliada. Por fim, a Seção 5 apresenta a conclusão do trabalho e os próximos passos.

2. Trabalhos Relacionados

Nos últimos anos, o volume de veículos tem se tornado cada vez maior. Isto adicionado à falta de planejamento urbano resultaram em diversos problemas no tráfego: crescimento do número de congestionamentos e intensificação do número de colisões. Com o objetivo de resolver estes problemas, diversos trabalhos têm estudado a mobilidade veicular através de *traces*¹ gerados por veículos. Basicamente, esses trabalhos utilizam três abordagens diferentes baseadas em grafos para processar os *traces*: grafo instantâneo (modela um instante específico no tempo), grafo agregado (modela um intervalo específico no tempo) ou grafo temporal (modela um cenário que evolui no tempo), explicados na Seção 3.

¹Arquivos gerados por veículos que normalmente contêm o horário e a localização GPS do veículo ao longo do dia

Na literatura, existem vários trabalhos que fazem a caracterização de *traces* com o objetivo de entender diferentes propriedades da mobilidade. No entanto, esses estudos utilizam um modelo específico de grafo.

Grafo instantâneo. Naboulsi and Fiore [2013] investigaram o *trace* da cidade de Colônia. Utilizando técnicas de redes complexas, eles verificaram a mobilidade veicular, componentes conexos e outras propriedades. Já Naboulsi and Fiore [2017] realizaram um trabalho semelhante analisando as mesmas métricas de mobilidade, mas também verificaram o *trace* da cidade de Zurique e compararam os resultados entre as cidades. Os autores identificaram uma limitação na topologia em termos de conectividade, disponibilidade, confiabilidade e navegabilidade. Hou et al. [2016] avaliaram o impacto da mobilidade em redes veiculares. Para isso analisaram momentos específicos do *trace* da cidade de Xangai e descobriram que a conectividade é afetada pela mobilidade somente quando a velocidade média dos veículos ultrapassa um limite específico. Pallis et al. [2009] investigaram a conectividade no *trace* da cidade de Zurique com o objetivo de analisar o desempenho dos protocolos de roteamento.

Grafo agregado. Guerber et al. [2018] analisaram os contatos no *trace* de Roma considerando medidas de grau, intermediação, proximidade e coesão. Para isso, eles realizaram uma agregação de todos os contatos existentes em um período de uma hora em um único grafo. Também agregando os contatos no período de uma hora, Cunha et al. [2014] geraram um grafo agregado, inserindo todos os contatos desse período no grafo. Cunha et al. [2016] estudaram como as lacunas encontradas na coleta de dados de redes veiculares podem afetar a análise de duração de contato, capacidade da rede e o número de contatos entre os veículos. O estudo foi realizado agregando o *trace* do período de 24 horas em um grafo. Diniz et al. [2017] juntaram os dados a cada uma hora para caracterizar a mobilidade em Roma. Na mesma direção, Qiu et al. [2018] demonstraram a natureza das comunicações veiculares, incluindo a duração do contato, velocidade do veículo e velocidade relativa entre os veículos na área urbana. Essa análise foi feita agregando os contatos dos veículos a cada quatro horas durante um período de 21 dias.

Grafo temporal. Qiao et al. [2017] avaliaram a evolução das características estruturais da mobilidade veicular na cidade de Pequim considerando a acessibilidade e conectividade na rede. Glacet et al. [2015] classificaram a conectividade da rede ao longo do tempo nas cidades de Bolonha e Colônia. Fiore et al. [2018] analisaram a mobilidade de Milão e identificaram uma relação que destaca a adequação da comunicação temporal para ambientes veiculares, o que pode fornecer diretrizes úteis para a avaliação e previsão do desempenho das redes veiculares.

Análise. No melhor do nosso conhecimento, encontramos um único trabalho na literatura que compara os diferentes tipos de grafos. Celes et al. [2018] realizaram uma comparação dos três tipos de grafos usados na modelagem da mobilidade. Entretanto, foram avaliadas poucas métricas e não há um estudo de como essas abordagens impactam os resultados dessas análises. Assim, os trabalhos existentes na literatura abordam diferentes tipos

de modelagens da mobilidade veicular e é necessário investigar como esses diferentes modelos influenciam as análises realizadas.

3. Metodologia

Esta seção discute a metodologia utilizada neste trabalho, apresenta as bases de dados analisadas juntamente com as métricas executadas e discute as três abordagens de modelagens investigadas e detalha como elas funcionam.

3.1. *Traces* de Mobilidade Veicular

Neste estudo, foram utilizadas duas bases de dados de mobilidade veicular de duas cidades diferentes: a cidade de São Francisco ² nos Estados Unidos e a cidade de Roma ³ na Itália. As duas bases de dados são de *traces* reais obtidos de táxis ao longo do dia. Esses *traces* são calibrados, i.e., há um registro dos veículos a cada segundo. Vale ressaltar que táxis são veículos que tem um comportamento diferente de veículos particulares, não tendo rotinas pré-estabelecidas. Os *traces* representam a mobilidade desses veículos em um dia (terça-feira) e foram escolhidos para representar o movimento durante a semana. A razão para avaliar apenas um dia é a quantidade de grafos gerados para os três modelos. Apesar de avaliarmos apenas para um dia, a metodologia não muda para realizar avaliações mais extensas. Além disso, segundo o estudo de rotinas de usuários no trabalho proposto em [Diniz et al., 2017], o movimento dos veículos é semelhante durante os dias da semana (segunda à sexta-feira).

3.2. Tipos de Modelagem

Na literatura existem três diferentes maneiras comumente encontradas para modelar a mobilidade veicular: grafos instantâneos, grafos agregados e grafos temporais. Antes de prosseguir, serão definidos os grafos que representam a mobilidade veicular. Nesses tipos de grafos, os vértices representam os veículos e as arestas representam a existência de contato entre eles (possibilidade de troca de informações). Para que um contato exista é necessário que dois veículos estejam dentro de um raio de comunicação. Neste trabalho, consideramos o raio de comunicação de cada veículo como 200 metros [Hui et al., 2010]. Novamente, outros raios podem ser considerados, mas a metodologia utilizada aqui não será alterada. Na Seção 5, ressaltamos como trabalhos futuros a aplicação da mesma metodologia com diferentes raios, para que possamos avaliar o impacto desse aspecto nos resultados.

a) *Grafo Instantâneo*: um grafo instantâneo é gerado a partir de um instante t do *trace* e a análise é feita sobre este instante. Neste tipo de abordagem, o fator temporal não é levado em consideração e a caracterização é baseada na análise deste único instante. Formalizando, $G_I = (V, E)$, tal que $V(t) = \{v_i(t)\}$ é o conjunto de vértices e $E(t) = \{e_{ij}(t) | v_i(t), v_j(t) \in V, i \neq j\}$ é o conjunto de arestas pertencentes ao grafo instantâneo no instante t . Neste trabalho, realizou-se a análise instantânea a cada 15 minutos pois, esse foi o mesmo tempo usado para gerar os grafos agregados. Assim será possível comparar de forma mais justa o resultado dessas modelagens.

²Esse *trace* contém dados de 500 táxis ao longo do mês de maio, ano 2009

³Esse *trace* contém dados de 320 táxis ao longo do mês de fevereiro, ano 2014

b) *Grafo Agregado*: o grafo agregado é a representação da mobilidade veicular durante um intervalo de tempo pré-estabelecido. Nesse tipo de grafo, todas as vértices e arestas existentes durante o intervalo escolhido são agrupados em um único grafo. Formalizando, $G_A = (V, E)$ tal que $V_A = V_t \cup V_{(t+1)} \cup V_{(t+2)} \dots \cup V_{(t+n)}$ e $E_A = E_t \cup E_{(t+1)} \cup E_{(t+2)} \dots \cup E_{(t+n)}$. Neste trabalho, utilizou-se um período de 15 minutos para gerar o grafo agregado. Este tempo foi escolhido por ser semelhante ao tempo usado no trabalho de [Celes et al., 2018] que compara essas abordagens.

c) *Grafo Temporal*: por fim, o grafo temporal é um grafo que analisa todo o período do *trace*. Para isso, em cada segundo um grafo é gerado e analisado, isso faz com que seja possível verificar o impacto ao longo do tempo no *trace*. Formalizando $G_{TVG} = (G_1, G_2, \dots, G_n)$ onde G_n representa o grafo do *trace* no instante n . Ressalta-se que o grafo instantâneo avalia um único instante do *trace* para representar todo o intervalo de tempo. Neste trabalho geramos um grafo instantâneo a cada 15 minutos para que seja possível comparar melhor a diferença entre as três métricas já que o grafo agregado foi gerado utilizando um intervalo de tempo também de 15 minutos. Já o grafo temporal leva em consideração todo o intervalo de tempo e mostra de forma mais minuciosa a variação das métricas durante o dia.

Faz-se necessário explicar como foi feita a modelagem dos pontos GPS de um instante t para um grafo G_t . A transformação foi feita da seguinte maneira: para cada instante de tempo t , um grafo G_t , vazio foi gerado. Em seguida, para cada veículo v com uma leitura GPS no instante t , um vértice foi adicionado no grafo G_t contendo o *id* e as coordenadas geográficas do veículo. Concluído esse processo, em cada grafo, realiza-se uma comparação entre cada par de vértices, somente se a distância geográfica entre eles têm menos de 200 metros, uma aresta entre estes dois vértices é adicionada ao grafo.

3.3. Métricas Analisadas

As métricas descritas a seguir são as frequentemente encontradas na caracterização de *traces*, como os trabalhos descritos na seção anterior. Dependendo da aplicação, é possível que uma métrica seja mais relevante. Este trabalho tem como objetivo avaliar essas métricas a partir de modelos diferentes, para podermos conhecer os resultados em cada um desses modelos. Optou-se por investigar essas métricas por serem métricas frequentemente usadas na literatura para caracterização de redes veiculares, inclusive nos trabalhos relacionados.

Diâmetro: após calcular o menor caminho entre todos os nós, o diâmetro é representado pelo maior dos menores caminhos. Encontrar o diâmetro tem grande importância como, por exemplo, nos algoritmos de roteamento *Open shortest path first* que se baseia no algoritmo de Dijkstra de menor caminho entre os nós.

Densidade: a densidade de uma rede é a relação entre o número de arestas existentes no grafo e o número de todas as arestas possíveis. Essa métrica consegue avaliar o relacionamento geral dos integrantes da rede, pois leva em consideração todos os contatos entre os veículos.

Número de arestas: representa a quantidade de contatos existentes entre os vértices do grafo. No caso de redes veiculares, significa que dois veículos estão próximos o suficiente para trocar informações.

Número de Componentes: um componente conexo de um grafo é composto por nós do grafo de forma que existe um caminho entre todos os nós pertencentes ao componente. Identificar os componentes, ajuda a identificar como as comunidades se movimentam ao longo do tempo e quais os locais de maior movimentação nas regiões analisadas, bem como a alcançabilidade da comunicação em cada grupo.

Número de Nós no Maior Componente: esta métrica identifica o número de nós no componente que possui mais nós. Isso nos permite avaliar a maior quantidade de nós conectados num dado instante.

Grau do Nó: o grau de um nó é dado pelo número de arestas incidentes neste nó, ou seja, o número de veículos com que ele consegue se comunicar diretamente. Estudar essa métrica ajuda a definir a distribuição do grau dos nós para geração de modelos que simulam a movimentação nas redes veiculares.

Centralidade de Intermediação: esta métrica quantifica a proporção de vezes que um nó faz parte do menor caminho entre outros dois nós. As medidas de centralidade ajudam a identificar os nós que são considerados influenciadores no ambiente em que estão. Nós com alto grau de centralidade de intermediação, caso removidos da rede podem gerar uma desconexão que impacta negativamente a distribuição dos dados.

Centralidade de Proximidade: em grafos conexos, existe uma medida de distância entre todos os nós, definida pelo comprimento de seus menores caminhos. O afastamento pode ser indicado como a soma das distâncias de seus menores caminhos entre todos os nós. A centralidade de proximidade é calculada pelo inverso do afastamento.

Coefficiente de Agrupamento: esta medida informa o grau com que os veículos tendem a formar grupos. Estes grupos normalmente possuem laços fortes e têm objetivos em comum e, geralmente, são chamados de comunidades. Isso pode facilitar o direcionamento de informações para determinado grupo de nós.

Correlação Temporal: é a probabilidade de uma aresta persistir em dois intervalos de tempo consecutivos em um grafo [Büttner et al., 2016]. Uma alta correlação indica que os contatos entre os veículos tendem a acontecer com frequência. Isso pode favorecer a criação de soluções personalizadas que permitem a resolução de problemas de tráfego.

4. Resultados

Esta seção apresenta os resultados obtidos neste estudo. O objetivo é comparar o impacto que os diferentes grafos usados nas modelagens têm sobre os resultados das análises das métricas. Vale ressaltar que as três abordagens são encontradas na literatura em diversos trabalhos, como mostrado na Seção 2.

4.1. Horário de Pico

Inicialmente, realizou-se uma análise do fluxo do trânsito ao longo do dia nas duas bases de dados selecionadas. Como isso, é possível entender melhor o comportamento do tráfego dessas cidades e compreender melhor os resultados obtidos nas próximas seções.

A Figura 1(a) mostra a movimentação dos veículos em Roma durante o dia. Pode-se notar que entre meia-noite e 4 horas da manhã, poucos veículos estão trafegando na cidade. Depois desse horário, o fluxo de automóveis começa a aumentar e por volta de

13 horas a rede atinge seu fluxo máximo. Após esse horário de pico, o fluxo diminui e se estabiliza entre os horários de 16 e 20 horas. Posteriormente, o movimento na rede volta a diminuir até o final da noite.

Já na cidade de São Francisco, percebe-se na Figura 1(b), que a movimentação dos automóveis vai se tornando mais intensa no período da madrugada e em parte da manhã. Por volta de 8 da manhã, o tráfego atinge o fluxo máximo e se mantém intenso até por volta de 17 horas. Após isso, a movimentação dos automóveis começa a diminuir e continua durante todo o período da noite.

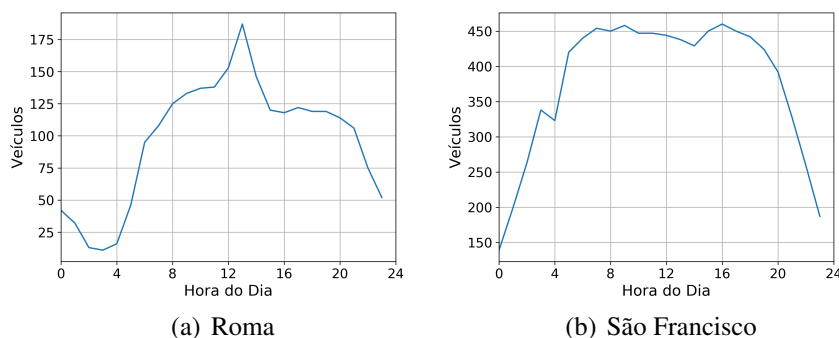


Figura 1. Fluxo de veículos durante o dia (em unidade).

4.2. Análise Comparativa

Esta seção apresenta os resultados para os três tipos de abordagens: Grafos Instantâneos, Grafos Agregados e Grafos Temporais, para os *traces* de Roma e São Francisco. Avaliamos os resultados para os modelos em nove métricas de redes complexas. Ressaltamos que a cada 15 minutos, geramos um grafo instantâneo. Apesar de o intuito deste tipo de grafo ser analisar um único instante do intervalo total de tempo para tomar como verdade, decidimos que seria mais justo que fosse gerado a cada 15 minutos para que a comparação com os grafos agregado e temporal fosse mais justa.

Primeiramente, será discutido o número de arestas nas três abordagens analisadas. Ao observar as Figuras 2(a) e 3(a) é possível ver a diferença dos resultados nas três métricas nas duas cidades. Olhando os resultados ao longo do dia, percebe-se como o grafo agregado apresenta valores de número de aresta superiores às outras duas abordagens. Isso também acontece quando se investiga a densidade da rede, que depende diretamente do número de arestas no grafo. Pode-se confirmar isso analisando as Figuras 2(b) e 3(b). No grafo instantâneo, os valores se aproximam mais da realidade, porém, ainda existe uma perda de informação a cada 15 minutos, já que um único segundo é analisado neste período. Somente no grafo temporal é possível analisar de forma clara e contínua a evolução dessas métricas no decorrer do dia.

Com um maior número de arestas, o grafo torna-se mais conexo e, consequentemente, o número de componentes é menor pois os nós se agrupam em componentes maiores. Isso acontece quando se analisa os componentes conexos no grafo agregado nas Figuras 2(e) e 3(e). Já no grafo instantâneo, é possível verificar de forma mais realista a evolução dos componentes conexos. Ainda assim, existe perda de informação significativa. Somente no grafo temporal é possível saber o que realmente acontece a cada instante

durante o dia. Como na métrica agregada existem menos componentes, cada componente possui mais nós e, conseqüentemente, o número de nós no maior componente aumenta. Isso pode ser observado nas Figuras 2(f) e 3(f). Outra métrica diretamente afetada pelo número de arestas é o grau médio dos nós. Observando as Figuras 2(d) e 3(d), podemos ver que no grafo agregado o grau dos nós é superior quando comparado aos outros dois grafos. Isso se dá pelo fato dos contatos existentes a cada 15 minutos serem agregados em um único grafo.

É importante ressaltar que a análise do número de componentes e a quantidade de nós em cada componente retrata o quão conectada a rede se encontra. Neste caso, é possível analisar como se comportam as soluções de disseminação e entrega de conteúdo na rede. Desta forma, a análise temporal consegue retratar de maneira mais realista esses padrões.

Outras três métricas que mostram o impacto dos diferentes tipos de modelagens são o coeficiente de agrupamento, a centralidade de intermediação e a centralidade de proximidade. Ao observarmos as Figuras 2(g), 3(g), 2(h), 3(h), 2(i) e 3(i), essas três métricas apresentam valores maiores no grafo agregado quando comparadas aos grafos instantâneo e temporal. Isso também está ligado ao número maior de arestas no grafo agregado. Em um grafo com mais arestas e menos componentes os nós possuem uma tendência maior de formar grupos pois, a chance de compartilharem o mesmo objetivo aumenta. Conseqüentemente, a possibilidade de participarem do menor caminho entre dois outros nós também aumenta, visto que uma quantidade maior de menores caminhos existirá em cada componente. Por fim, a distância de um nó a todos os outros nós (em número de saltos) também tende a ser maior. O nó de um componente tem mais nós para calcular a distância entre eles.

Por último, será discutido o diâmetro da rede. Ao analisarmos o resultado do diâmetro nas duas bases de dados nas Figuras 2(c) e 3(c) percebe-se que o impacto na cidade de Roma foi maior que na cidade de São Francisco. Na cidade de Roma, a abordagem agregada tem resultados de diâmetros maiores que na abordagem temporal. Já na cidade de São Francisco essa diferença não chega a ser tão expressiva. Na abordagem instantânea, é possível ver a curva de evolução da métrica, porém, no grafo temporal, vê-se de maneira mais precisa o desenvolvimento do diâmetro ao longo do dia.

4.3. Frequência vs. Duração

Dessa seção em diante, todas os resultados foram gerados usando o grafo temporal como abordagem escolhida. Essa decisão foi tomada, pois, na análise anterior notou-se que o grafo temporal tende a se aproximar mais da realidade que as duas outras abordagens. Foi investigada a relação entre frequência e duração dos contatos entre os automóveis. Com isso pode-se ter uma perspectiva melhor de como funciona a troca de informações entre os veículos ao longo do dia.

Inicialmente analisamos o comportamento dos contatos dos veículos na cidade de Roma. Como é possível observar na Figura 4(a), no período entre meia-noite e 6 : 00 da manhã, os contatos são mais duradouros e menos frequentes. Isso significa que os veículos tendem a manter o contato por mais tempo, mas raramente se reencontram neste período. A partir do horário de 6 : 00, a situação muda e os contatos passam a ter uma duração menor, porém, os reencontros são mais frequentes. Sendo assim, os veículos tendem a se

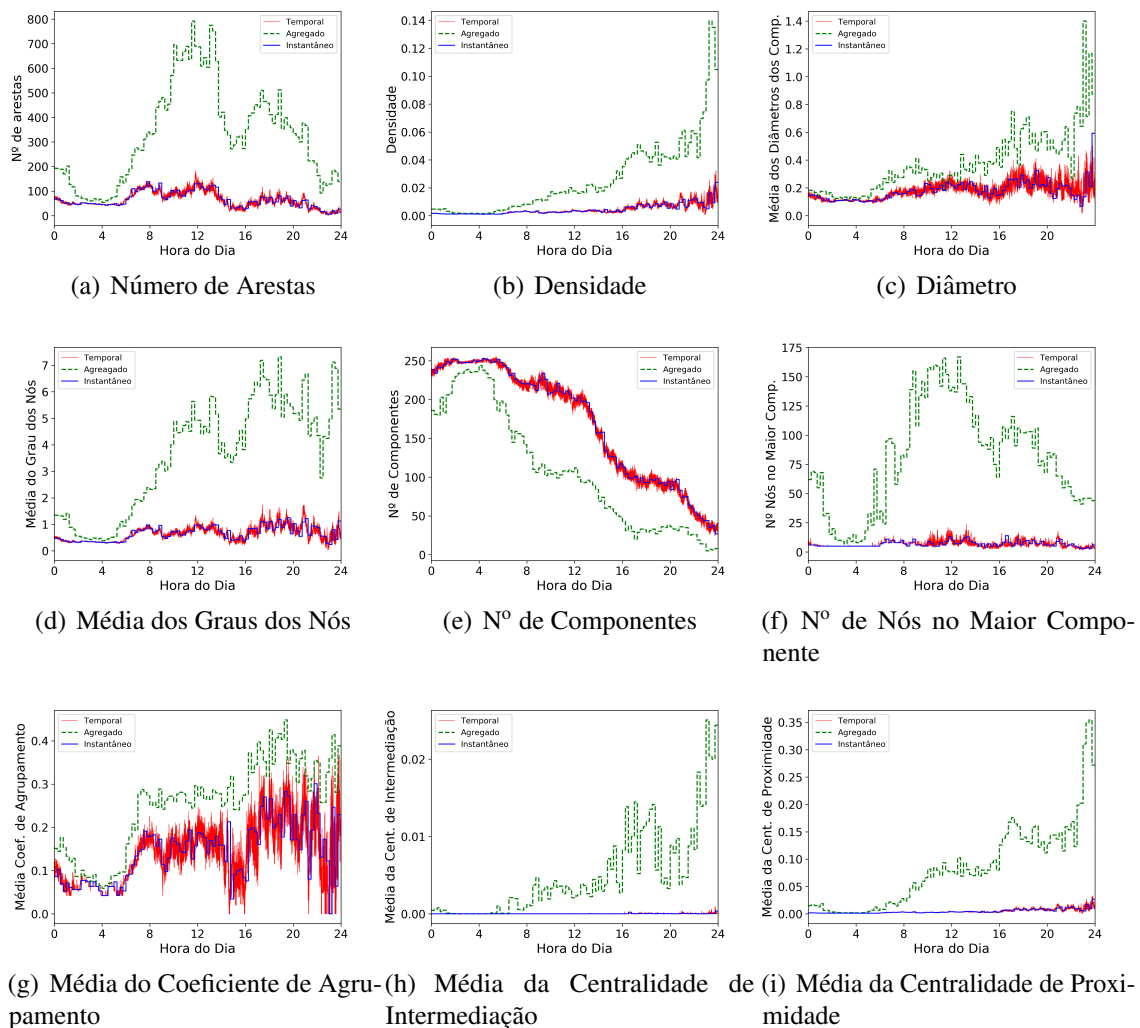


Figura 2. Comparando o impacto das diferentes abordagens em Roma.

manter menos tempo próximos, mas tendem a ter encontros recorrentes.

Na cidade de São Francisco, Figura 4(b), observamos que os contatos ocorrem em média uma única vez por hora e seu período conectado é próximo a 20 segundos na maioria do dia. Isso nos mostra que os veículos não se encontram frequentemente e não mantêm contatos muito duradouros ao longo do dia. Além disso, é possível observar que no período da noite os contatos têm uma duração maior do que ao longo do dia. Este tipo de informação sugere que diferentes soluções/estratégias devem ser desenvolvidas/consideradas para diferentes momentos.

Na cidade de Roma, os contatos se mantêm ativos durante mais tempo, fazendo com que seja possível trocar mais informações entre os pares de nós. Isso não acontece na cidade de São Francisco, onde os contatos devem ser melhor aproveitados já que não acontecem com frequência e têm um curto período de duração. Soluções personalizadas possivelmente serão necessárias visto que em cada ambiente, existe um comportamento diferente dos veículos.

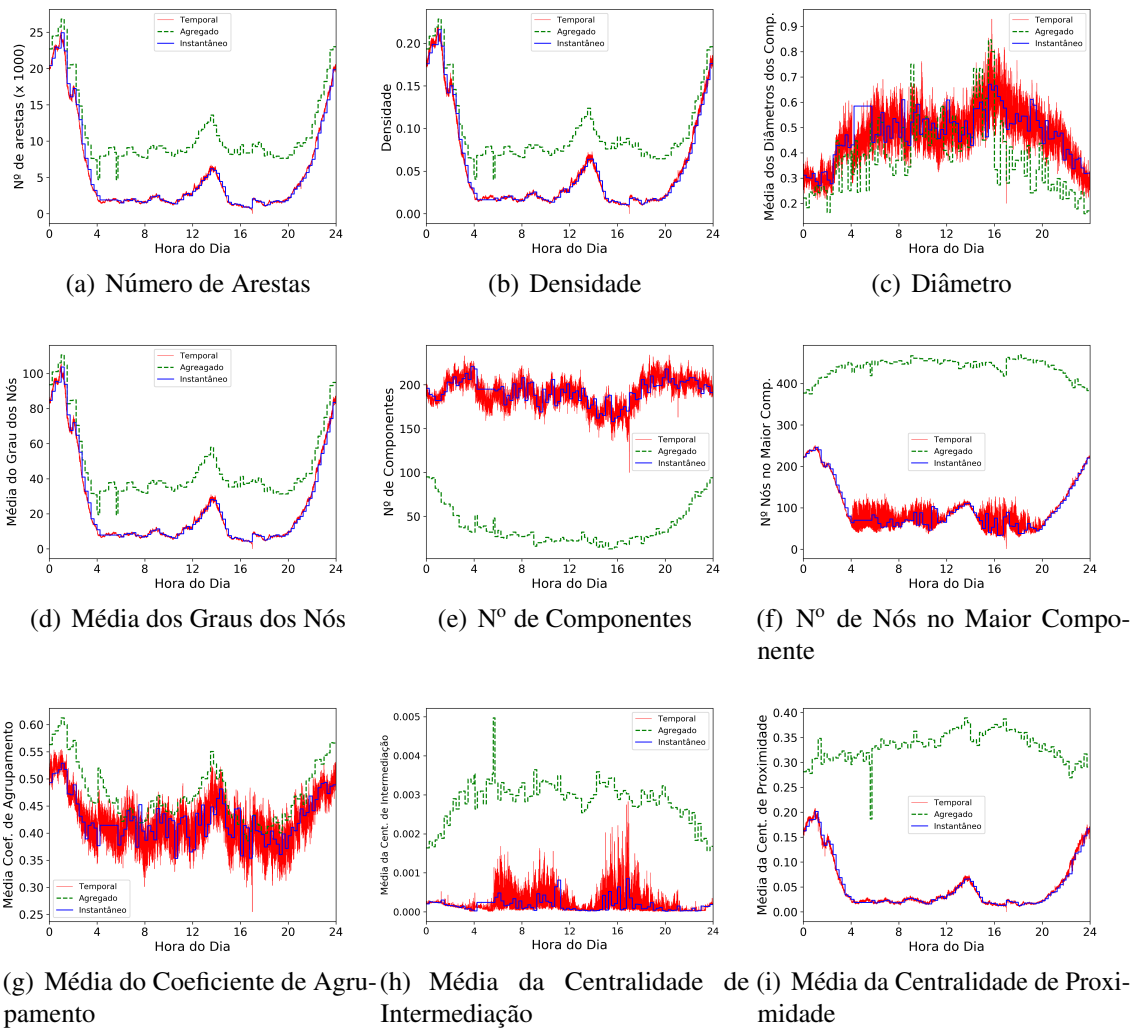


Figura 3. Comparando o impacto das diferentes abordagens em São Francisco.

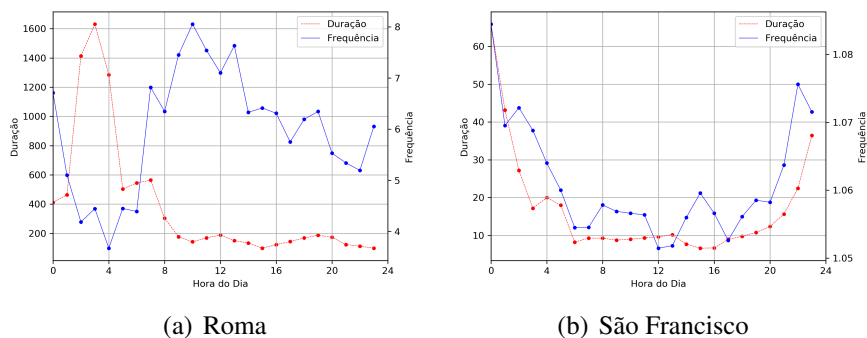


Figura 4. Duração vs. Frequência dos contatos ocorridos durante o dia.

4.4. Correlação Temporal

A correlação temporal é uma métrica que mede a probabilidade de uma aresta persistir por dois instantes consecutivos em um grafo temporal. Ao analisar esta métrica,

pode-se entender melhor qual a probabilidade de um contato entre os veículos se manter ao longo do dia. Se um contato tende a existir, pode significar que estes veículos possuem um relacionamento, como em uma rede social. Será discutida primeiro a análise realizada na cidade de Roma. Observando a Figura 5(a), é possível notar que os encontros ocorridos durante o dia tendem a ser recorrentes. Isso mostra que os relacionamentos criados pelos veículos são duradouros, fazendo que um automóvel esteja sempre em contato com seu círculo social. Observa-se que no período da noite, esta probabilidade diminui um pouco, mas ainda mantendo-se alta, com valores próximos à 0.8. Pressupõe-se que isso ocorre por ser o fim do expediente e os táxis começam a se deslocar para diferentes bairros, fazendo com que algumas arestas não existam mais.

Já na Figura 5(b) observa-se que a correlação temporal na cidade de São Francisco é alta e tende a manter valores mais constantes durante todo o período do dia. Isso significa que, também neste *trace*, os veículos criam relacionamentos que duram ao longo do dia, fazendo assim com que as arestas persistam. A correlação temporal pode ajudar no aprimoramento de aplicações para ITS (Inteligente Transportation System) pois permite entender o comportamento das arestas (conectividade) de um grafo. No caso de redes veiculares, como é necessário aproveitar da melhor forma possível o tempo de contato entre os veículos, podemos projetar soluções personalizadas para cada tipo de situação e local.

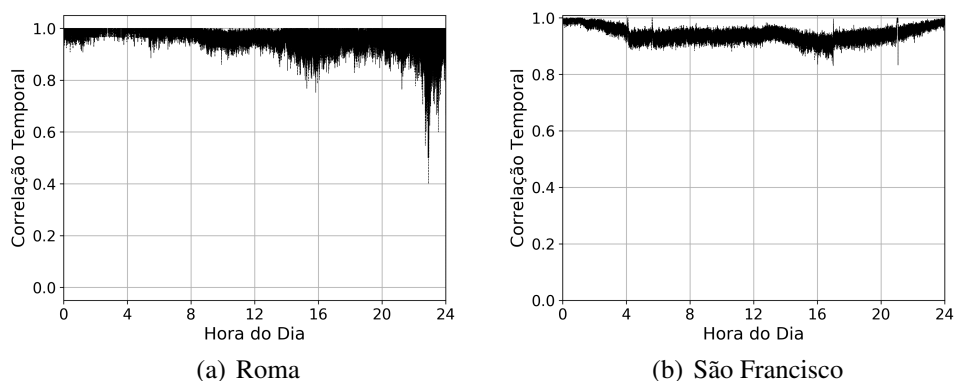


Figura 5. Correlação Temporal das arestas ao longo do dia.

4.5. Correlação de Pearson

Com o intuito de verificar como ocorre a correlação entre as métricas analisadas, foi calculada a Correlação de Pearson e gerado um mapa de calor. A Figura 6 apresenta os resultados obtidos. A seguir, a análise dos resultados será apresentada para cada cidade avaliada.

Roma: Analisando as correlações para o *trace* de Roma, pode-se verificar correlações muito fortes direta e inversa (valores entre 0.9 e 1 e entre -1 e -0.9) e as correlações fortes diretas e inversas (valores entre 0.8 e 0.9 e entre -0.9 e -0.8). Uma correlação direta significa que à medida que o valor de uma determinada variável avaliada aumenta, as variáveis correlacionadas também aumentam seu valor. Observando a Figura 6(a) podemos identificar duas correlações diretas fortes. A primeira é entre o coeficiente de agrupamento e o grau médio dos nós. Isso mostra que quando o grau dos

nós aumenta, também aumenta a probabilidade de este nó estar envolvido em um grupo com objetivos semelhantes. Isso acontece pois o nó aumenta seu círculo de contatos, aumentando a chance de algum vizinho ter os mesmos objetivos. Outra correlação direta forte analisada é entre a densidade da rede e a centralidade de proximidade. Este resultado era esperado visto que à medida que a densidade da rede aumenta, também aumentam as arestas presentes no grafo, fazendo assim com que os nós se tornem mais próximos uns dos outros.

Outra correlação observada é entre a densidade e o número de componentes. Ela é considerada inversa e forte, pois, à medida que a densidade da rede aumenta, mais arestas fazem parte do grafo, diminuindo assim o número de componentes.

São Francisco: Ao observarmos a Figura 6(b) podemos ver que existem várias correlações entre as métricas. Para melhor organizar a análise, primeiro serão discutidas as correlações diretas muito fortes. Existem cinco métricas que se correlacionam de forma muito forte e direta. Estas métricas são o Número de Arestas, o Grau Médio dos nós, o número de nós no Maior Componente, a Centralidade de Proximidade e a Densidade da rede. Isso é compreensível, visto que essas métricas estão diretamente ligadas. Por exemplo, se o número de arestas de um grafo aumenta, conseqüentemente o grau médio dos nós e a densidade da rede também aumentam. Mais arestas no grafo, significa também que o grafo tende a estar mais conectado. Com isso, o número de nós nos componentes também aumenta. Na mesma direção, mais nós em cada componente resultam também na aproximação dos mesmos, o que por sua vez ocasiona o aumento da centralidade de proximidade.

Outra métrica que também está relacionada diretamente com as cinco métricas citadas anteriormente é o Centralidade de Proximidade. Esta métrica também apresenta uma correlação forte com todas as cinco métricas citadas anteriormente. Isto já era esperado, pois em um grafo mais conectado, a probabilidade que os veículos que fazem parte do mesmo componente também compartilhem o mesmo objetivo é mais alta.

Nas correlações inversas, à medida que uma variável aumenta seu valor em relação a outra diminui. Observa-se que uma única métrica possui relacionamento inverso forte com outras métricas: o Diâmetro da rede. Ao observarmos as métricas citadas acima, vemos que elas se correlacionam de forma direta, pois à medida que um grafo fica mais conectado, seus vértices tendem a ser mais próximos uns dos outros. Isso explica a correlação inversa com o diâmetro da rede. Quanto mais conexo um grafo, menor vai ser o diâmetro, pois os nós terão que dar menos saltos para chegar aos outros vértices da rede.

5. Conclusão e Trabalhos Futuros

Este trabalho investigou o impacto das diferentes abordagens usadas na modelagem da mobilidade veicular na análise da rede. Para isso foram analisadas duas bases de dados, uma da cidade de Roma e uma da cidade de São Francisco, utilizando três tipos de grafos para modelagem encontradas na literatura: grafo instantâneo, grafo agregado e grafo temporal. Percebeu-se que os resultados gerados em cada grafo são diferentes. Isso mostra que a abordagem utilizada na modelagem da mobilidade pode afetar diretamente os resultados das análises de métricas de redes complexas e de outras métricas. Isso mostra claramente a importância de uma escolha rigorosa do modelo a se adotar no momento de se representar uma rede. No grafo agregado, os resultados acabam sendo “exagerados”

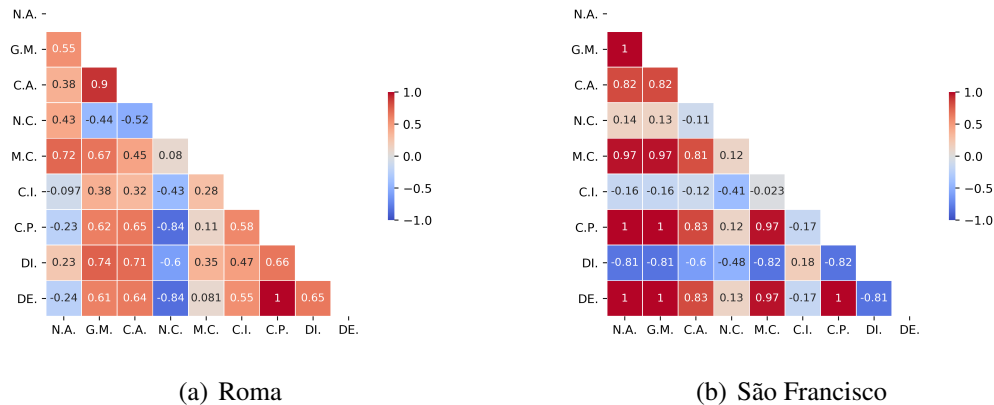


Figura 6. Correlação de Pearson entre as métricas analisadas. Número de Arestas (N.A.), Grau Médio (G.M.), Coeficiente de Agrupamento (C.A.), Número de componentes (N.C.), Número de Nós no Maior Componente (M.C.), Centralidade de Intermediação (C.I.), Centralidade de Proximidade (C.P.), Diâmetro (DI.) e Densidade (DE.)

pois todos os contatos existentes nos 15 minutos são agregados em um único grafo. Já no grafo instantâneo, é possível analisar a curva do comportamento de cada métrica ao longo do dia. Ainda assim, existe uma grande perda de informação nessa abordagem. Por fim, no grafo temporal, é possível ver claramente a evolução de cada métrica ao longo do tempo.

Como esperado, a análise do grafo temporal apresenta de forma clara e legítima como cada métrica se comporta ao longo do dia. Quando usado, esse tipo de grafo mostra, sem perda de informação, as características da rede. No entanto, o custo para computar essas métricas é proporcional à quantidade de grafos gerados, que no caso do modelo temporal usado neste trabalho foi de 1 segundo. Assim, deve-se avaliar com cuidado o modelo a ser usado em função da métrica de interesse. Este trabalho discute isso e serve para mostrar que podem existir tendências/resultados não necessariamente reais ou próximos da realidade dependendo do modelo utilizado.

Este trabalho mostrou também que é possível descobrir algumas propriedades que podem ajudar a desenvolver novas soluções personalizadas para melhorar o trânsito nas cidades. Como trabalhos futuros pretendemos analisar outros *traces* e outras métricas por mais dias, com o intuito de descobrir características que ajudem a aprimorar o tráfego. Além disso, pretende-se realizar os mesmos estudos variando o raio de comunicação escolhido para verificar o impacto nas modelagens e o tempo usado no grafo agregado. A partir da análise realizada, pretendemos propor novos serviços baseados nas propriedades e padrões encontrados.

Referências

- R. d. S. Alves, I. d. V. Campbell, R. d. S. Couto, M. E. M. Campista, I. M. Moraes, M. G. Rubinstein, L. H. M. Costa, O. C. M. Duarte, and M. Abdalla. Redes veiculares: Principios, aplicações e desafios. *Minicursos do Simpósio Brasileiro de Redes de Computadores, SBRC*, pages 17–24, 2009.

- K. Büttner, J. Salau, and J. Krieter. Temporal correlation coefficient for directed networks. *SpringerPlus*, 5, 12 2016. doi: 10.1186/s40064-016-2875-0.
- C. Celes, A. Boukerche, and A. A. Loureiro. On the temporal analysis of vehicular networks. In *2018 IEEE Symposium on Computers and Communications (ISCC)*, pages 01074–01079. IEEE, 2018.
- F. D. Cunha, A. C. Vianna, R. A. Mini, and A. A. Loureiro. Is it possible to find social properties in vehicular networks? In *2014 IEEE Symposium on Computers and Communications (ISCC)*, pages 1–6. IEEE, 2014.
- F. D. Cunha, F. A. Silva, C. Celes, G. Maia, L. B. Ruiz, R. M. Andrade, R. A. Mini, A. Boukerche, and A. A. Loureiro. Communication analysis of real vehicular calibrated traces. In *2016 IEEE International Conference on Communications (ICC)*, pages 1–6. IEEE, 2016.
- G. R. Diniz, A. A. Loureiro, and F. D. Cunha. On the characterization of vehicular mobility. *DIVANet '17 Proceedings of the 6th ACM Symposium on Development and Analysis of Intelligent Vehicular Networks and Applications*, pages 23–29, nov 2017.
- M. F. Fiore, C. Glacet, and L. Bedogni. Temporal reachability in vehicular networks. *INFOCOM 2018-IEEE Conference on Computer Communications*, pages 1–9, 2018.
- C. Glacet, M. Fiore, and M. Gramaglia. Temporal connectivity of vehicular networks: The power of store-carry-and-forward. In *2015 IEEE Vehicular Networking Conference (VNC)*, pages 52–59. IEEE, 2015.
- C. R. Guerber, E. L. Gomes, M. Fonseca, and A. Munaretto. Classificação geosocial de contatos para disseminação de dados em redes veiculares oportunistas. In *Anais do XXXVI Simpósio Brasileiro de Redes de Computadores e Sistemas Distribuídos*. SBC, 2018.
- X. Hou, Y. Li, D. Jin, D. O. Wu, and S. Chen. Modeling the impact of mobility on the connectivity of vehicular networks in large-scale urban environments. *IEEE Transactions on Vehicular Technology*, 65(4):2753–2758, 2016.
- P. Hui, J. Crowcroft, and E. Yoneki. Bubble rap: Social-based forwarding in delay-tolerant networks. *IEEE Transactions on Mobile Computing*, 10(11):1576–1589, 2010.
- D. Naboulsi and M. Fiore. On the instantaneous topology of a large-scale urban vehicular network: the cologne case. *Proceedings of the International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc)*, 07 2013. doi: 10.1145/2491288.2491312.
- D. Naboulsi and M. Fiore. Characterizing the instantaneous connectivity of large-scale urban vehicular networks. *IEEE Transactions on Mobile Computing*, 16(5):1272–1286, 2017.
- G. Pallis, D. Katsaros, M. D. Dikaiakos, N. Loulloudes, and L. Tassioulas. On the structure and evolution of vehicular networks. In *2009 IEEE International Symposium on Modeling, Analysis & Simulation of Computer and Telecommunication Systems*, pages 1–10. IEEE, 2009.
- L. Qiao, Y. Shi, and S. Chen. An empirical study on the temporal structural characteristics of vanets on a taxi gps dataset. *IEEE Access*, 5:722–731, 2017.
- Z. Qiu, S. Zhang, W. Zhou, and S. Yu. Empirical study on taxi’s mobility nature in dense urban area. In *IEEE INFOCOM 2018-IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, pages 232–237. IEEE, 2018.