



PADRÃO DE MOBILIDADE DOS USUÁRIOS DO MOZILLA LOCATION SERVICE PARA A CIDADE DO RIO DE JANEIRO

Frank L. M. J. de Oliveira Queiroz

Paulo Cezar M. Ribeiro

Programa de Engenharia de Transportes
COPPE - Universidade Federal do Rio de Janeiro

RESUMO

Métodos tradicionais de coleta de dados em transportes são caros, demorados e são atualizados numa frequência baixa. Devido a popularização dos telefones celulares, um novo campo de pesquisa foi criado, o qual busca analisar dados gerados por estes dispositivos a fim de obter resultados semelhantes às tradicionais pesquisas domiciliares. Este trabalho buscou estudar o padrão de mobilidade de uma amostra de colaboradores do *Mozilla Location Service* para a cidade do Rio de Janeiro, apesar do grande viés, os resultados obtidos são relativamente condizentes com a realidade da cidade e mostra que os dados são promissores ao estudo em transportes.

ABSTRACT

Traditional methods of data collection in transport are expensive, time-consuming and are updated at a low frequency. Due to the popularization of mobile phones, a new field of research was created, which seeks to analyze data generated by these devices in order to obtain results similar to the traditional domiciliary surveys. This work sought to study the mobility pattern of a sample of users of the *Mozilla Location Service* for the city of Rio de Janeiro, despite the great bias, the results obtained are relatively consistent with reality and showing that they are promising to study in transport.

1. INTRODUÇÃO

Os métodos tradicionais de pesquisa em transporte acarretam em alto custo aos administradores públicos. A estrutura necessária, bem como a quantidade de profissionais para a realização desta pesquisa contribui para seu custo elevado. Neste cenário, diversas limitações são intrínsecas à estes métodos como, por exemplo, o uso de uma pequena amostra e limitação espacial e temporal dos dados coletados, isto é, não é possível obter um cronograma completo das viagens, tampouco variações sazonais (Calabrese *et al.*, 2013).

O fator custo é determinante na frequência de atualização de uma pesquisa de transporte. As pesquisas de origem-destino com base domiciliar para a cidade de São Paulo, por exemplo, é atualizado a cada 10 anos. Esta frequência também é adotada pela *Enquete Globale Transport* da cidade de Paris (Bonnell *et al.*, 2015). A baixa frequência de atualização limita a capacidade de resposta das políticas urbanas no que se relaciona ao rápido crescimento metropolitano e das mudanças socioeconômicas, de infraestrutura e de comportamento das viagens (Calabrese *et al.*, 2013).

Outra limitação reside na participação ativa daqueles que realizam deslocamentos e respondem a tais pesquisas, visto que as pessoas podem não lembrar corretamente as informações solicitadas e também podem ter razões para não reportar honestamente as informações (Mellegard, 2011). Estas razões se resumem, sobretudo, em fatores relacionados à segurança e privacidade.

O estabelecimento da estrutura, somado ao tempo de coleta e processamento dos dados, torna o método tradicional bastante demorado. Conforme Feriancic (2015), o tempo de realização fica no intervalo de 6 a 12 meses, levando em conta a concepção, testes, aplicação, tabulação e verificação de consistências.

A área de mobilidade vem recebendo crescente atenção de administradores públicos brasileiros. No ano de 2012, foi aprovada a Lei 12.587/12, que instituiu a Política Nacional de Mobilidade



Urbana. Esta lei previa um prazo de 3 anos para a elaboração do Plano de Mobilidade Urbana (PMU), e o não cumprimento acarretaria no impedimento do envio de recursos orçamentários federais destinados à mobilidade urbana. Entretanto, o prazo se exauriu e um número ínfimo de cidades conseguiu atender às exigências. Segundo estudo da NTU (2015), mais de 70% das capitais e cidades brasileiras acima de 500 mil habitantes, e 95% do total de municípios acima de 50 mil habitantes, não conseguiram finalizá-lo. De acordo com o mesmo estudo, os principais problemas enfrentados pelos municípios na elaboração dos planos, são referentes à falta de recursos financeiros e a precária estrutura de pessoal.

Em contrapartida, vivemos em uma era de alta tecnologia, na qual os *smartphones* possuem dispositivos e aplicativos que geram uma quantidade massiva de dados locais, surgindo como uma grande ferramenta aos planejadores urbanos a fim de modificar o cenário supracitado. Dentre as vantagens do uso desta tecnologia pode-se citar o grande tamanho da amostra, assim como dispensa a instalação de equipamentos no veículo ou a abordagem de motoristas (Qiu e Cheng, 2008).

Existem, no entanto, limitações tais como a falta de atributos demográficos e socioeconômicos, a não aleatoriedade da amostra da população e o fato de que o conjunto de dados não foi projetado para propósitos referentes a transportes. Consequentemente, o formato dos dados não vem de maneira fácil a ser utilizada, o que restringe sua utilidade (Calabrese *et al.*, 2013). Apesar de tais desvantagens, estudos relevantes obtiveram resultados positivos por meio de uso de tais dados.

Visando contribuir aos métodos alternativos para obtenção de dados voltados ao planejamento de transporte, este estudo busca analisar a adequabilidade dos dados do *Mozilla Location Service* (MLS) na verificação do padrão de mobilidade de seus colaboradores. Foi realizado o estudo de caso para a cidade do Rio de Janeiro.

Após esta introdução o artigo segue com o referencial teórico no capítulo 2, métodos de distribuição de viagens no capítulo 3, uma explanação acerca dos dados fornecidos pelo MLS no capítulo 4, no capítulo 5 é explicado método de processamento dos dados, posteriormente, os resultados obtidos são apresentados no capítulo 6, o capítulo 7 conclui o estudo e propõe recomendações de pesquisas futuras.

2. USO DA TELEFONIA PARA ESTUDOS EM TRANSPORTE

Com o surgimento da era digital, foram criados métodos alternativos de pesquisas em transportes baseados na tecnologia da informação. Podem-se destacar trabalhos que utilizam dados locais do Twitter (McNeill *et al.*, 2016), ou até mesmo fazem mineração de texto da mesma rede social a fim de inferir localizações (Sparks *et al.*, 2016). Há também trabalhos que buscam estudar padrões de mobilidade por meio de dados de bilhetagem eletrônica (Guerra *et al.*, 2015). Além disso, há trabalhos que estudam o fluxo de veículos de carga por meio do RFID (Rojas e Ribeiro, 2012). O presente estudo se atem aos dados locais de telefonia móvel.

As tecnologias para localizar um telefone são baseadas no dispositivo (GPS, entre outros) ou soluções baseadas na rede (antenas) (Qiu e Cheng, 2008). Os trabalhos que utilizaram a rede são os mais antigos e de suma importância a área de estudo. Pode-se dizer que o trabalho que



representa a gênese destes estudos foi elaborado por White e Wells (2001). Estes autores utilizaram registros de chamada para gerar a matriz OD para o condado de Kent, Inglaterra.

Os registros de chamada são mais conhecidos pela sigla CDR (*Call Detail Record*). Tais registros contém a posição da antena que cobre a área na qual está sendo realizada a chamada, bem como o horário, duração e o identificador do dispositivo. Estas informações são necessárias para a operadora de telefonia fazer a cobrança mensal do proprietário da linha, por este motivo estes dados também são chamados de *billing data*. Os resultados obtidos por White e Wells (2001) não foram satisfatórios e os autores perceberam a necessidade de uma amostra de CDR muito maior para construir uma matriz OD semelhante àquelas obtidas por métodos tradicionais.

Asakura *et al.* (2002) utilizaram o sistema PHS (*Personal Handy Phone System*) para coletar dados de espectadores de um evento em Osaka, Japão. A vantagem da adoção do sistema PHS é que não há a necessidade de realização de uma chamada para coleta de dados, isto é, bastava o dispositivo estar ligado para acontecer o intercâmbio de dados, tal como o funcionamento de um GPS. Os autores fizeram importantes descobertas acerca da mobilidade urbana sob grandes eventos, dentre elas descobriu-se que atrações no entorno do estádio após um evento são efetivos na redução do congestionamento em estações ferroviárias ou metroviárias.

Os trabalhos acima representam a gênese dos estudos de telefonia móvel voltados à mobilidade urbana. Do começo do século até os dias atuais, os estudos se intensificaram, sobretudo pela popularização do telefone celular. Surgiram estudos que se voltaram a elaboração de algoritmos de *data mining* a fim de estimar fluxos origem-destino, bem como também prever a alocação de tráfego (Wang *et al.*, 2013; Toole *et al.*, 2014). Outros trabalhos se ativeram não apenas aos algoritmos de mineração de dados, mas também aos meios de validação de informações obtidas por meio dos dados de telefonia móvel (Bar-Gera, 2007; Calabrese *et al.*, 2013).

Bekhor *et al.* (2013) utilizaram dados de sinal de telefonia (*signalling data*) para estudar padrões de viagens de longa distância em Israel. Estes dados são praticamente da mesma família do *billing data*, entretanto, tem a diferença de que não há necessidade de realização de uma chamada para a geração do registro, a localização é armazenada continuamente. Embora possa parecer que o *signalling data* é melhor que o CDR, deve-se levar em conta que há um *tradeoff* entre estes dados, enquanto o CDR tem a precisão de antenas, o dado permanente tem precisão de um conjunto de dezenas de antenas. Por estudar o território nacional a precisão fornecida pelos dados contínuos foi satisfatória, levando a bons resultados e a percepção de padrões tanto em dias úteis, quanto em fins de semana.

Bonnel *et al.* (2015) utilizaram dados de chamadas telefônicas para a estimativa da matriz OD para a cidade de Paris. Os resultados foram comparados com dados obtidos por fontes tradicionais de pesquisa em transporte na França. Houve boa correlação com os dados da *Enquête Globale Transport*, documento que pode ser comparado aos planos diretores de transporte brasileiros.

Este trabalho, ao contrário dos supracitados, não utilizou CDRs, tampouco os dados passivos, haja vista que, por questões relacionadas à privacidade, a complexidade de obtenção desses dados é alta. Com isso, neste estudo utilizou-se uma fonte aberta de dados de telefonia, o



Mozilla Location Service. O objetivo desta pesquisa é analisar a adequabilidade de tais dados aos estudos de transporte.

3. MODELOS DE DISTRIBUIÇÃO DE VIAGENS

O conhecimento da geração de viagens nas zonas de tráfego não é suficiente para a modelagem e tomada de decisão em planejamento de transportes, é necessário conhecer movimentos interzonais. As viagens geradas e atraídas por uma zona de tráfego, são distribuídas às demais zonas da área de estudo por meio de métodos de distribuição.

Duas categorias básicas de métodos de distribuição são predominantes no planejamento de transporte urbano. As categorias são: métodos de fator de crescimento e método gravitacional.

3.1. Fator de crescimento

Estes métodos envolvem o dimensionamento de uma matriz existente pela aplicação de fatores multiplicativos às células da matriz, estes fatores são, geralmente, derivados de uma produção ou atração prevista.

3.2. Método gravitacional

Os modelos gravitacionais aplicados em transportes surgiram de suposições de que os padrões de interação espacial são regidos por leis análogas à Lei de Gravitação Universal de Newton, onde a força de gravidade que age entre dois corpos é diretamente proporcional à massa dos dois corpos e inversamente proporcional ao quadrado da distância entre eles.

Em um espaço geográfico associado a um espaço econômico, as distâncias devem ser vencidas com um mínimo custo de transporte. Neste cenário, surgem fatores que tendem tanto a facilitar a interação entre os elementos que compõem o espaço, quanto dificultar tal interação. Esses fatores podem ser entendidos como forças de atração à determinada área, tal força terá grande intensidade quando haver baixos custos de transporte, energia elétrica abundante, facilidades de comunicação, etc. Tal força terá menor intensidade quando os custos de transporte forem elevados ou devido à falta de mão de obra qualificada.

Na formulação inicial, o modelo considerou apenas a população das cidades estudadas, entretanto esta análise foi considerada muito simplista sendo incluído o uso de total de viagens de um O_i para um D_j ao invés de população total. O modelo foi ainda mais generalizado assumindo que o efeito ou “separação” poderia ser modelado melhor por uma função decrescente, a ser especificada, da distância ou custo de viagem entre as zonas (Ortuzar e Willumsen, 2011). Isto pode ser escrito como:

$$T_{ij} = \alpha O_i D_j f(c_{ij})$$

Em que T_{ij} : Fluxo entre i e j

α : Fator de proporcionalidade

O_i : Viagens com origem i

D_j : Viagens com destino j

$f(c_{ij})$: função de fricção, retenção ou impedância

Neste trabalho o método gravitacional foi utilizado para o cálculo do fluxo entre zonas, os dados de origem e destino foram obtidos por meio dos dados do MLS, já a impedância será obtida pela distância entre as zonas estudadas.



4. MOZILLA LOCATION SERVICE

Este capítulo trata de uma explanação acerca dos dados utilizados por este estudo.

4.1. Serviços de localização

Os serviços de localização são de grande importância nos dias atuais. Por meio deles há a possibilidade de estimar a localização de um indivíduo e a partir disso desenvolver aplicativos que se baseiam em localização, como a capacidade de visualizar o tráfego, encontrar restaurantes próximos, postos de gasolina ou marcar fotos com local de onde foi tirada.

Como são baseados em GPS, em determinadas situações, os aplicativos que fazem uso destes serviços podem sofrer uma interrupção do serviço ou apresentar algum defeito momentâneo, por exemplo, no interior de prédios ou em dias nublados. Visando contornar isso, foram criados novos meios de se obter dados locais, como Wi-Fi, antenas telefônicas, bluetooth, entre outros. O MLS utiliza sinais emitidos por Wi-Fi públicos e antenas telefônicas para estimar a posição de um indivíduo ao redor do mundo, é um serviço gratuito e pode ser utilizado no desenvolvimento de aplicativos.

4.2. Método de coleta de dados utilizado pelo MLS

Para estimar a localização de um indivíduo, é necessário saber com exatidão a posição das antenas telefônicas e pontos de Wi-Fi, de posse dessas posições a localização de um aparelho é calculada com base na intensidade do sinal que chega ao dispositivo. Entretanto, obter a localização de todos os pontos requeridos seria um trabalho oneroso e demorado. Então o Mozilla criou uma plataforma na qual os indivíduos podem colaborar com a base de dados de localização de pontos Wi-Fi e antenas. A cada dia a localização destes pontos é atualizada, tornando-se cada vez mais exata a todo instante em que um aparelho, com aplicativos Mozilla instalados, passa pela área de cobertura

O método de coleta de dados é baseado em *Crowdsourcing*, isto é, são fomentados por um grande número de usuários. Tais dados são gerados pelos dispositivos que tem aplicativos Mozilla instalados, os aplicativos são Mozilla Stumbler ou o Mozilla Firefox. Estes apps são capazes de estimar a área de cobertura de uma antena de telefonia celular, bem como a posição do dispositivo por meio do GPS, com estas duas informações pode-se estimar a posição da antena. Todos estes dados são compilados para criar uma espécie de mapa global de antenas e wi-fi públicos. Desta forma, qualquer dispositivo que acessar uma das redes identificadas poderá obter a sua geolocalização aproximada.

4.3. Dados fornecidos pelo MLS

A MLS disponibiliza os dados de localização atualizada de antenas telefônicas diariamente. Os dados são do tipo .csv (*comma separated values*) e tem informações relativas às coordenadas geográficas do dispositivo, cobertura da antena, horário de carregamento dos dados para a base MLS, entre outros.

Tabela 1 – Dados disponibilizados pela MLS

Tipo da rede de telefonia
ID da antena
Coordenadas geográficas
Área de cobertura da antena
Hora de carregamento



Devido ao fato de os dados apresentarem o horário de carregamento, não se pode afirmar que o horário disponibilizado é o horário no qual o deslocamento foi realizado. Entretanto, após a análise preliminar dos dados, percebeu-se que para haver coleta e armazenamento deve haver deslocamento, pois o aplicativo desliga automaticamente quando o aparelho não está em movimento. Portanto, entende-se que todas as detecções dos dados do MLS são referentes a viagens executadas, mas não exatamente no horário registrado.

Durante uma viagens, o aplicativo permanece ativo, coletando estimativas de localização de pontos Wi-Fi e antenas. Após o fim da viagem, o aplicativo carrega automaticamente as localizações do dispositivo, bem como as estimativas de área de cobertura. Como as informações são carregadas todas de uma vez, o horário de carregamento é o mesmo, o que possibilita coletar a viagem de um dispositivo. Esta característica de carregamento dos dados é utilizada para superar a limitação de não haver ID do dispositivo nos dados disponibilizados.

A Figura 1 é referente a uma viagem coletada por meio do MLS.

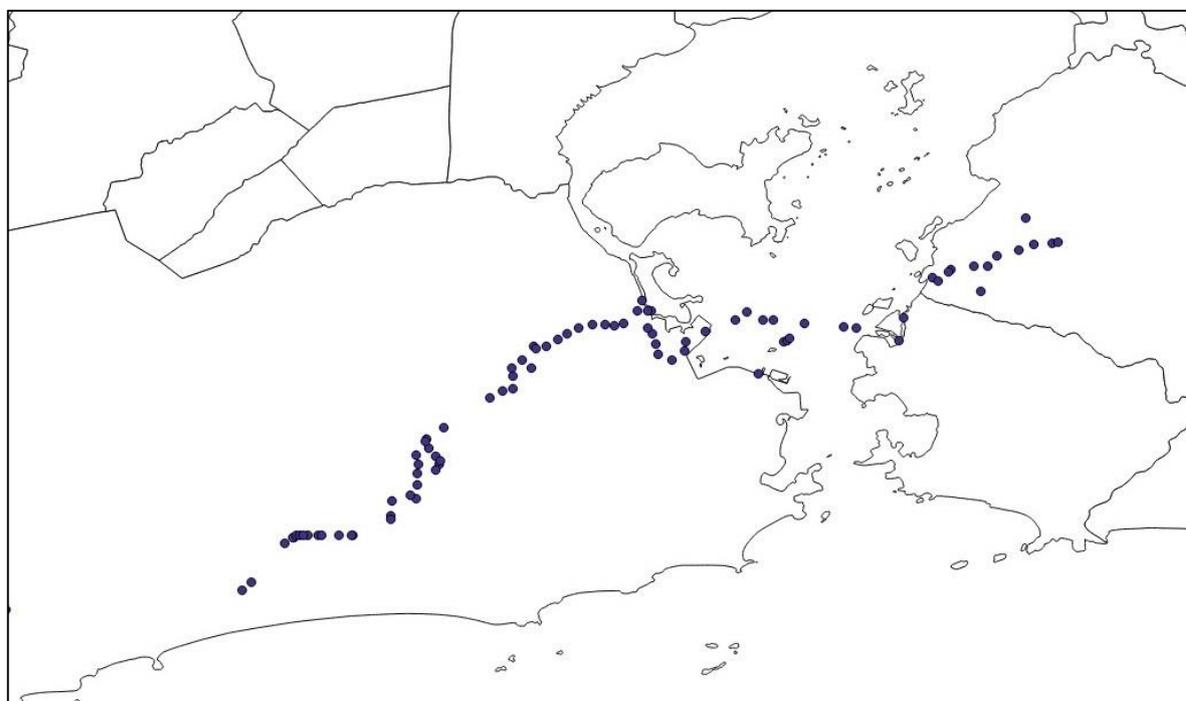


Figura 1 - Caminho feito por um dispositivo

As limitações são explicadas devido ao fato de que tais dados não tem como objetivo o estudo em transportes, no entanto, isto não esgota a possibilidade de estudo dos mesmos, tendo em vista que os dados provenientes de CDR também não são ideais ao propósito, mas tiveram bons resultados nos trabalhos citados anteriormente.

4.4. Estudo preliminar dos dados fornecidos pelo MLS

As Figuras 2 e 3 representam um estudo preliminar feito com tais dados, a fim de verificar sua validade ao propósito deste estudo. Tais imagens representam mapas de calor das detecções de dispositivos que colaboram com a MLS no dia 15 de maio de 2017. Os mapas retratam a quantidade de detecções em determinada área, ou seja, quanto maior a intensidade da cor escura, maior o número de detecções da área em questão.



Na base de dados existem um grande número de horário de carregamento correspondentes a um único registro. Considerou-se que tais registros são referentes a viagens curtas e que foram carregadas em um curto espaço de tempo após o deslocamento, ou seja, o horário de carregamento se assemelha ao horário de movimento.

Adotando a mesma consideração feita por Bonnel *et al.* (2015), o horário de carregamento foi dividido em duas partes, hora em que teoricamente o indivíduo está em casa (ou proximidades) e, em um segundo grupo, hora que os indivíduos teoricamente estão no trabalho ou estudo (ou proximidades). Os horários adotados foram:

- Domicílio: dezenove às sete (19h – 7h)
- Trabalho ou estudo: sete às dezenove (7h – 19h)

A Figura 2 retrata a faixa horária na qual teoricamente o indivíduo proprietário do telefone celular está em casa, já a Figura 3 considera o horário no qual o indivíduo está teoricamente no trabalho, ou estudo.

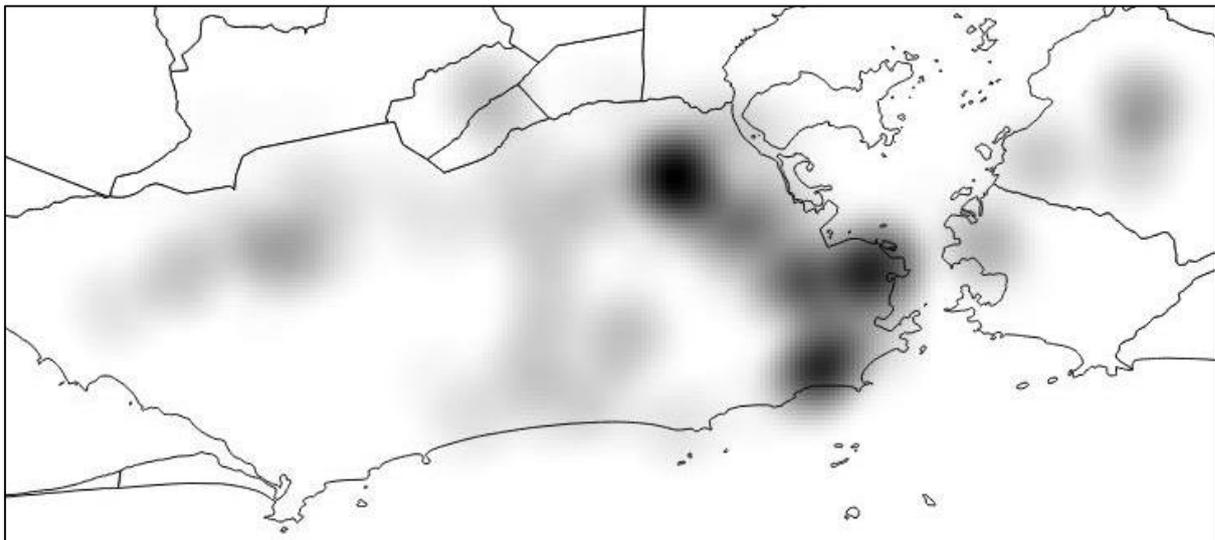


Figura 2 - Detecções no horário de casa

Na Figura 2 percebe-se uma grande concentração de detecções nos bairros da zona Sul do Rio de Janeiro, esta concentração é ainda maior em alguns bairros da zona norte. Foram detectados dispositivos na zona oeste, entretanto a concentração é bastante reduzida se comparada às áreas citadas anteriormente. Na Figura 3 percebe-se uma grande concentração de detecções na região da Grande Tijuca e em bairros da região central.

Tais informações possibilitam inferir que os indivíduos detectados no horário de casa, provavelmente trabalham ou estudam na região central do Rio de Janeiro. Este resultado, proveniente de um estudo preliminar, levou a acreditar que tais dados pudessem ter utilidade à análise da mobilidade dentro de uma cidade. A partir daí foi elaborado o método para filtragem e estudo destes dados explanado no item 5.

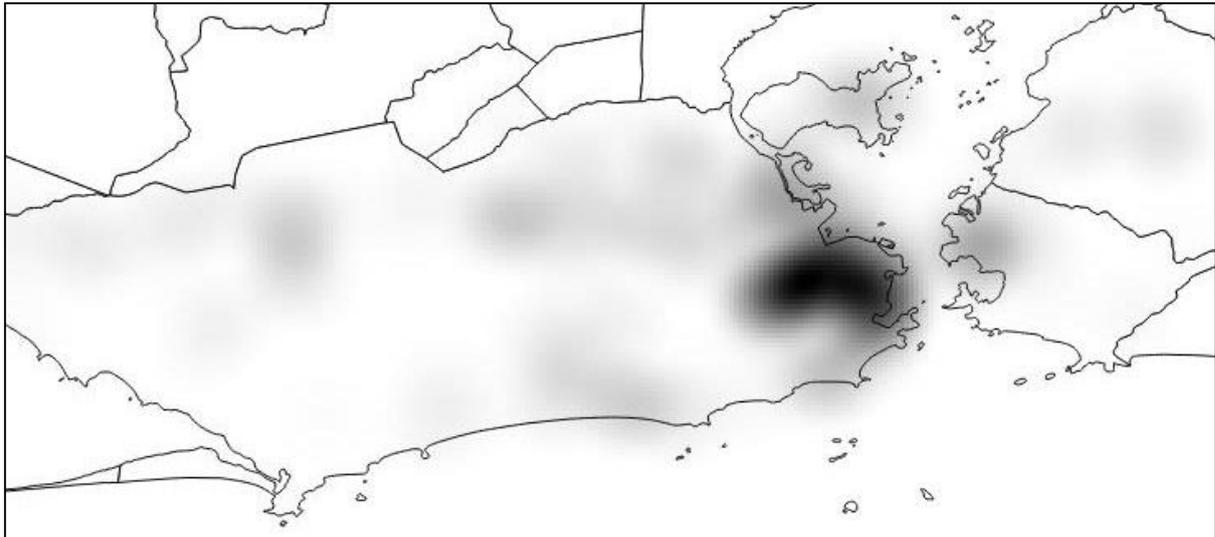


Figura 3 - Detecções no horário de trabalho ou estudo

5. MÉTODO E MATERIAIS

O processamento de dados neste trabalho adotou o seguinte procedimento:

1. Extrair da base de dados MLS os dados de localização da cidade por meio de coordenadas de latitude e longitude

A base de dados do MLS tem 20 milhões de registros de localização de antenas que são atualizados diariamente, os dados disponíveis nestes registros estão contidas na Tabela 1. Este número é espalhado pelo mundo, portanto, este item inicial é necessário para a filtragem dos dados para a cidade do Rio de Janeiro.

2. Converter hora do registro de UTC Unix para dia, mês e ano para a cidade do Rio de Janeiro. Os dados são disponibilizados no formato Unix que é um formato de tempo em segundos cujo início é 1/1/1970. A conversão é necessária para organizar os dados de maneira eficiente, além de facilitar as consultas no banco de dados.

3. Estabelecer a Área de Planejamento (AP) de cada ponto extraído da base de dados. O estabelecimento da AP, conforme Figura 4, é necessário para a obtenção das produções e atrações de viagens no nível de precisão mostrado.

4. Consulta ao banco de dados

A consulta filtrará os dados de localização do banco de dados Mozilla, obedecendo as seguintes considerações:

- Caso o horário de carregamento tenha um número maior do que quatro repetições, considera-se como uma viagem e extrai-se o primeiro (origem) e o último (destino) registro;
- Caso o horário de carregamento tenha quatro ou menos repetições, considera-se o dispositivo como estacionário (viagem muito curta) e extrai-se apenas uma localização;
- Para as detecções com quatro ou menos repetições, será adotado os horários prováveis de casa, trabalho ou estudo, estes horários são: Casa: dezenove às sete (19h – 7h) / Trabalho ou estudo: sete às dezenove (7h – 19h)



Tais considerações foram implementadas em um algoritmo SQL o qual possibilitou a extração das produções e atrações de cada Área de Planejamento do Rio de Janeiro.

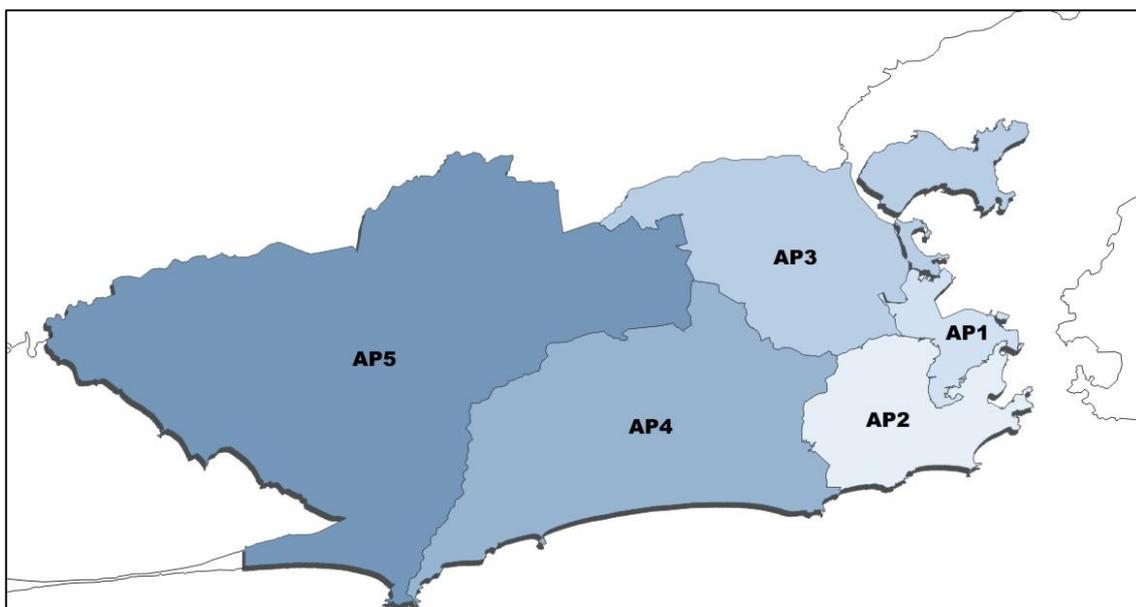


Figura 4 - Áreas de Planejamento do Rio de Janeiro

5. Obtenção dos fluxos entre zonas

Tendo em vista que se possui a atração, produção e impedância (distância) entre zonas para todas as AP do Rio de Janeiro, o método gravitacional de distribuição de viagens torna-se o mais indicado. Utilizou-se o software TransCAD para o cálculo de tais fluxos, os resultados estão dispostos no capítulo abaixo.

6. RESULTADOS OBTIDOS

Para este estudo optou-se por extrair a matriz OD de um dia. O dia analisado é 24 de maio de 2017, quarta-feira de uma semana comum, ou seja, sem feriados prolongados ou eventos que possam afetar a mobilidade na cidade. Dos dispositivos que possuem aplicativos Mozilla instalados, as detecções se dividiram por macrozona conforme o gráfico abaixo.



Gráfico 1 - Percentual de detecções por macrozona

Percebe-se que a maior concentração está na região Central (AP1) do Rio de Janeiro, isto se deve ao fato de que comumente moradores de outras Áreas de Planejamento exercem suas atividades na macrozona correspondente aos bairros centrais. A matriz que corresponde aos



fluxos calculados é mostrada na Tabela 2. Optou-se por dar destaque aos movimentos entre macrozonas. A tabela abaixo representa a percentagem de fluxo entre as área de planejamento.

Tabela 2 – Fluxo percentual entre Áreas de Planejamento

	AP1	AP2	AP3	AP4	AP5
AP1	-	-	-	-	-
AP2	17%	-	-	-	-
AP3	12%	13%	-	-	-
AP4	3%	7%	9%	-	-
AP5	10%	7%	13%	9%	-

Percebe-se que o maior fluxo de viagens entre a zona Sul (AP2) e Centro (AP1), apresentando uma proporção de 17% dos deslocamentos totais entre macrozonas. Há percentagem significativa de viagens entre o Centro e AP3, bem como Centro e AP5. A região correspondente à AP4 apresentaram baixa taxa de viagens em relação à região central.

Os fluxos entre AP3 e AP2 também se destacaram, sendo praticamente o somatório entre viagens das Áreas de Planejamento 4 e 5. A AP4 foi a área que apresentou os menores fluxos origem-destino, merece destaque o número relativamente alto de viagens para a AP3. No que toca à AP5 percebe-se que os fluxos relacionados à AP3 e Centro se destacam. De fato o acesso entre estas áreas é relativamente bom, visto que além da Avenida Brasil que faz a ligação rodoviária, existe transporte público por trilhos.

7. CONCLUSÃO E RECOMENDAÇÕES

Os métodos tradicionais de pesquisa em transporte tem execução cara e demorada, por isso são de difícil atualização. Devido ao crescente uso dos aparelhos celulares, vários estudos propuseram procedimentos para estimativa de matrizes OD, alocação de tráfego, dentre outras informações com base em dados produzidos por estes aparelhos. Tais estudos utilizaram precipuamente o CDR das operadoras de telefonia móvel, dados que, devido à condições de privacidade, são difíceis de serem obtidos.

Este estudo usou dados do MLS a fim de extrair os fluxos de viagens de seus colaboradores entre macrozonas do Rio de Janeiro. Usando um dia de detecção foi observado que os resultados são promissores no que diz respeito ao padrão de mobilidade da cidade, visto que foi observado um alto fluxo entre AP1 e AP2, bem como para AP3 e AP1, em consonância com a realidade do município. Dentre as recomendações de pesquisa futura pode-se citar a implementação de um algoritmo de filtragem mais criterioso do que o que foi adotado neste trabalho. Além disso, podem ser realizados trabalhos com o objetivo de estudar os padrões de mobilidade nos finais de semana, haja vista que o MLS também disponibiliza tais dados. Também é importante a análise de outra cidade a fim de verificar qual é a influência da quantidade de dados na qualidade final da informação obtida. O estudo da cidade de São Paulo é uma sugestão válida. Por fim, recomenda-se a montagem de matrizes para um número maior de dias, com o intuito de aprimorar os resultados obtidos nesta pesquisa. Tais pesquisas podem fornecer um entendimento mais profundo do potencial dos dados do *Mozilla Location Service* para estudos em transporte.



REFERÊNCIAS BIBLIOGRÁFICAS

- Asakura, Y., Hato, E., & Sugino, K. (2005) Simulating Travel Behaviour Using Location Positioning Data Collected with a Mobile Phone System. In *Simulation Approaches in Transportation Analysis* (pp. 183-204). Springer US.
- Associação Nacional das Empresas de Transportes Urbanos - NTU (2015) Revista NTU Urbano Edição Nº13.
- Bar-Gera, H. (2007) Evaluation of a Cellular Phone-Based System for Measurements of Traffic Speeds and Travel Times: A Case Study from Israel. *Transportation Research Part C: Emerging Technologies*, 15(6), 380-391.
- Bekhor, S., Cohen, Y., & Solomon, C. (2013) Evaluating Long-Distance Travel Patterns in Israel By Tracking Cellular Phone Positions. *Journal of Advanced Transportation*, 47(4), 435-446.
- Bonnel, P., Hombourger, E., Olteanu-Raimond, A. M., & Smoreda, Z. (2015) Passive Mobile Phone Dataset to Construct Origin-Destination Matrix: Potentials and Limitations. *Transportation Research Procedia*, 11, 381-398.
- Caceres, N., J. P. Wideberg e F. G. Benitez (2007) Deriving Origin-Destination Data From a Mobile Phone Network. *IET Intelligent Transportation System*. Pag. 15-26.
- Calabrese, F., Diao, M., Di Lorenzo, G., Ferreira, J., & Ratti, C. (2013) Understanding Individual Mobility Patterns From Urban Sensing Data: A Mobile Phone Trace Example. *Transportation research part C: emerging technologies*, 26, 301-313.
- Ferriancic, G., Celeiro, F. R., & Silva, L. N. B. (2015) Planejamento da Mobilidade com Big Data de Telefonia Móvel. In *20º Congresso Brasileiro de Transporte e Trânsito*, Santos.
- Guerra, A. L., Barbosa, H. M., & de Oliveira, L. K. (2014) Estimativa de Matriz Origem/Destino Utilizando Dados do Sistema de Bilhetagem Eletrônica: Proposta Metodológica. *TRANSPORTES*, 22(3), 26-38.
- Mellegård, E. (2011). Obtaining Origin/Destination-Matrices from Cellular Network Data.
- McNeill, G., Bright, J., & Hale, S. A. (2016) Estimating Local Commuting Patterns From Geolocated Twitter Data.
- Qiu, Z., & Cheng, P. (2008) State of the Art and Practice: Cellular Probe Technology Applied in Advanced Traveler Information System. In *86th Annual Meeting of the Transportation Research Board*, Washington, DC (No. 0223)
- Ortuzar, J. D. D.; Willumsen, L. G. (2011) Modelling Transport - Fourth edition. New York: Wiley&Sons.
- Rojas, A., & Ribeiro, P. C. M. (2012) Freight Truck Assignment from RFID Flow Analysis. In *19th ITS World Congress*.
- Sparks, K. A., Li, R. G., Thakur, G. S., Stewart, R. N., & Urban, M. L. (2016) Facility Detection and Popularity Assessment from Text Classification of Social Media and Crowdsourced Data. In *Proceedings of the 10th Workshop on Geographic Information Retrieval* (p. 2). ACM.
- Toole, J. L., Colak, S., Sturt, B., Alexander, L. P., Evsukoff, A., & González, M. C. (2015) The Path Most Traveled: Travel Demand Estimation Using Big Data Resources. *Transportation Research Part C: Emerging Technologies*, 58, 162-177.
- Wang, M. H., Schrock, S. D., Vander Broek, N., & Mulinazzi, T. (2013) Estimating Dynamic Origin-Destination Data and Travel Demand Using Cell Phone Network Data. *International Journal of Intelligent Transportation Systems Research*, 11(2), 76-86.
- White, J., & Wells, I. (2002) Extracting Origin Destination Information from Mobile Phone Data.