

Análise da Cobertura Espacial de uma Rede de Sensores Baseada em Ônibus Urbanos

Pedro Cruz¹, Rodrigo S. Couto² e Luís Henrique M. K. Costa¹ *

¹Universidade Federal do Rio de Janeiro - PEE/COPPE/GTA

²Universidade do Estado do Rio de Janeiro - PEL/DETEL/FEN

{cruz, luish}@gta.ufrj.br, rodrigo.couto@uerj.br

Abstract. *Smart Cities can employ urban buses with sensors to enlarge their coverage, with low infrastructure needs. In this context, this paper analyses the coverage of a mobile wireless sensor network where mobility is provided by urban buses. We propose a model and an optimization problem to maximize the coverage for a given number of sensing buses. The problem is then applied to a real data set of buses from Rio de Janeiro, to evaluate the maximum coverage for a limited number of sensors. Results show that if 18% of the fleet are equipped with sensing nodes, it can cover at least 94% of the streets of Rio de Janeiro that have some sort of bus circulation, the equivalent to 5,606 km of streets.*

Resumo. *Cidades inteligentes podem utilizar ônibus urbanos equipados com sensores para aumentar a cobertura espacial, com pouca necessidade de infraestrutura. Nesse contexto, este trabalho estuda a cobertura de uma rede de sensores móveis com mobilidade provida por ônibus urbanos. Para tal, propõe-se um modelo de cobertura e formula-se um problema de otimização que maximiza a região coberta para um dado número de ônibus que realizam sensoriamento. O problema é aplicado a dados reais dos ônibus da cidade do Rio de Janeiro, determinando a maior área sensoriável com um número limitado de sensores. Os resultados mostram que 18% da frota cobre pelo menos 94% das ruas do Rio de Janeiro com circulação de ônibus, ou o equivalente a 5.606 km de ruas.*

1. Introdução

A coleta de dados é um importante aspecto da computação urbana e um fator determinante para se construir cidades inteligentes [Kamiński et al., 2016]. O paradigma de Internet das Coisas (*Internet of Things* - IoT) incrementa objetos do dia-a-dia com capacidades de comunicação e processamento [Gubbi et al., 2013], [Santos et al., 2016], criando uma importante ferramenta para o sensoriamento de cidades inteligentes [Zanella et al., 2014]. Esses objetos inteligentes podem ser instalados no ambiente urbano para coletar dados. Por meio de pontos de coleta espalhados pela cidade, os dados podem ser enviados para uma nuvem ou disponibilizados diretamente para os usuários finais, criando uma rede de sensores sem-fio (*Wireless Sensor Network* - WSN).

Por outro lado, são necessários muitos sensores para cobrir toda a área de uma cidade [Du et al., 2015]. A implementação de uma WSN com sensores estáticos por toda

*Este trabalho foi parcialmente financiado pela CAPES, CNPq, FAPERJ e pelos processos nº 15/24494-8 e nº 15/24490-2, da Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP).

uma cidade pode ser inviável financeiramente, pois os sensores são capazes de coletar dados ininterruptamente das áreas nas quais estão instalados, mas não coletam dados de áreas mais distantes. Uma forma de cobrir espacialmente toda a cidade com uma quantidade reduzida de sensores é ter sensores que se movem pela cidade. Desta forma cada sensor pode cobrir uma área maior, porém ao custo de uma menor frequência de sensoriamento de cada região [Liu et al., 2005]. Além disso, os sensores são capazes de realizar comunicação oportunística com os pontos de coleta, reduzindo a quantidade de pontos de coleta necessários para a recepção dos dados [Ekici et al., 2006]. Uma solução de baixo custo para dotar os sensores de mobilidade é acoplar os nós de sensoriamento em ônibus urbanos, dentro do modelo da IoT. Dessa forma, a mobilidade é adquirida com custos desprezíveis e os pontos de ônibus podem servir como pontos de coleta para a rede, como mostrado em [Cruz et al., 2017c].

Existe uma grande gama de aplicações que podem se alimentar de dados coletados por uma rede urbana de sensores móveis. Cada aplicação pode ter requisitos específicos em termos de completude de dados, temporalidade de dados [Pipino et al., 2002] e de granularidade de informação [Zanella et al., 2014]. No caso da cobertura espacial, a completude e a granularidade da informação estão relacionadas ao conhecimento da área completa da cidade e do espalhamento das medidas que são obtidas. Adicionalmente, a cobertura de novas regiões da cidade pode ser capaz de eliminar ruídos e incertezas sobre dados de suas regiões vizinhas [Sta, 2016].

Este trabalho analisa a cobertura espacial de uma rede de sensores com mobilidade baseada em ônibus urbanos. Um modelo de cobertura é proposto e um Problema de Programação Inteira-Mista (*Mixed-Integer Linear Programming* - MILP) é formulado, com o intuito de escolher os ônibus capazes de maximizar a área coberta pelos ônibus. Com essa formulação, é possível encontrar um resultado ótimo ou, a um menor custo computacional, encontrar um resultado subótimo e estimar o erro máximo para o resultado ótimo. Dados reais de posição dos ônibus da cidade do Rio de Janeiro são coletados, pré-processados e aplicados ao MILP. Com os resultados da solução do problema, estima-se que com cerca de 18% dos ônibus do Rio de Janeiro seja possível cobrir pelo menos 5.606 km das vias da cidade. Essa área corresponde a 94% da região coberta caso todos os ônibus estejam equipados com nós de sensoriamento. Adicionalmente, estima-se que uma rede experimental que cubra 40% da região originalmente coberta custa menos de R\$ 13.000,00 em equipamentos e R\$ 120,00/mês em tráfego. Ao final, é realizada uma análise do efeito da escolha dos ônibus na quantidade de vezes que cada ponto de interesse é visitado ao longo de um dia.

Este trabalho está organizado da seguinte forma. A Seção 2 analisa os trabalhos relacionados. A Seção 3 descreve a rede considerada neste trabalho, examina o modelo de cobertura adotado para esta rede e apresenta um modelo em MILP para a maximização da cobertura da rede quando há um número limitado de ônibus com nós de sensoriamento. A Seção 4 descreve os experimentos realizados, que mostram a análise de cobertura do Rio de Janeiro em função do número de nós de sensoriamento. A Seção 5 conclui o trabalho e apresenta os trabalhos futuros.

2. Trabalhos Relacionados

A utilização de veículos no sensoriamento urbano já foi considerada em outros trabalhos. O projeto BusNet [Zoyza et al., 2007] utiliza ônibus urbanos para detectar buracos nas estradas por onde os veículos passam. O projeto, porém, não realiza um estudo da cobertura do sensoriamento e é limitado à aplicação de manutenção de estradas.

Existem alguns projetos relacionados à utilização de veículos urbanos para coleta de dados em um sentido mais amplo. Alguns exemplos são os projetos Mosaic [Dong et al., 2015] e Opensense [Marjovi et al., 2015], que utilizam veículos para monitorar a qualidade do ar em ambiente urbano. O Projeto Sensing-Bus [Cruz et al., 2017c] utiliza ônibus urbanos para coletar dados gerais sobre a cidade.

Fiore *et al.* [Fiore et al., 2016] exploram modelos de processamento de sinais para estimar a acurácia de medidas obtidas por veículos e pedestres. O trabalho considera que todos os veículos são equipados com nós de sensoriamento e utiliza a densidade de dispositivos distribuídos em regiões de uma cidade para definir a qualidade do sensoriamento em cada região.

Zhu *et al.* [Zhu et al., 2013] e Du *et al.* [Du et al., 2015] analisam a eficiência da aplicação de uma rede veicular no monitoramento de tráfego urbano. Du *et al.* propõem modelos para utilizar sensores presentes em automóveis para estimar o volume e a velocidade do tráfego. Zhu *et al.* estudam a cobertura e a qualidade dos dados gerados por uma frota de aproximadamente 4.000 táxis na cidade de Xangai. O presente trabalho analisa uma rede similar, porém não limitada ao monitoramento do tráfego urbano, utilizando apenas os ônibus urbanos e estudando como diferentes números de nós de sensoriamento afetam a cobertura da rede.

O trabalho de Zhao *et al.* [Zhao et al., 2015] estuda a cobertura oportunista da cidade utilizando diferentes modelos de mobilidade. Zhao *et al.* caracterizam a cobertura de sub-regiões a partir da probabilidade de um veículo equipado com um nó de sensoriamento visitar a sub-região. O presente trabalho tem como principal diferença a modelagem de uma rede com rotas conhecidas e determinísticas, e não um modelo estocástico.

Nos trabalhos anteriores [Cruz et al., 2017b] e [Cruz et al., 2017a], o atraso na entrega de dados coletados por uma rede de sensores baseadas em ônibus urbanos é discutido. Esses trabalhos se baseiam na arquitetura do SensingBus, que possui pontos de coleta espalhados pela cidade, implementando uma entrega oportunística. Assim, o atraso na entrega de dados coletados é relacionado com o número e o posicionamento dos pontos de coleta da rede. O presente trabalho analisa a cobertura espacial da rede estudada no trabalho anterior, sem considerar o atraso na entrega dos dados. Ambos os trabalhos podem ser aplicados a uma mesma rede, com o intuito de reduzir custos ou aumentar a qualidade do serviço oferecido por ela.

3. Modelo de Cobertura da Região Sensoriada

Este trabalho está no contexto do projeto SensingBus [Cruz et al., 2017c]. O projeto utiliza uma rede de três camadas para realizar a coleta, pré-processamento, processamento e entrega dos dados aos usuários finais. Nessa rede, ônibus funcionam como nós de sensoriamento, circulando pela cidade e coletando dados. Os dados são entregues a

Tabela 1. Notações utilizadas no trabalho.

Notação	Descrição	Tipo
\mathcal{I}	Trechos de ruas a serem cobertos	Conjunto
\mathcal{I}_s	Trechos de ruas que são cobertos	Conjunto
\mathcal{J}	Ônibus que circulam pela cidade	Conjunto
\mathcal{J}_s	Ônibus equipados com nós de sensoriamento	Conjunto
\mathcal{N}_i	Ônibus que passam pelo trecho i	Conjunto
\mathcal{T}_j	Trechos de ruas pelo qual o ônibus j passa	Conjunto
a_i	Valor binário indicando o comprimento do trecho i	Parâmetro
p	Número total de ônibus a serem escolhidos como nós de sensoriamento	Parâmetro
x_j	Valor binário que indica se o ônibus j é escolhido como nó de sensoriamento	Variável
y_i	Valor binário indicando se o trecho i é coberto	Variável
C	A cobertura total possível da cidade pela rede estudada	Variável

pontos de coleta espalhados pelos pontos de ônibus da cidade, são pré-processados e re-passados a um serviço de nuvem que processa e provê a usuários finais acesso aos dados.

No SensingBus, os contatos entre os nós de sensoriamento e os pontos de coleta são oportunistas, já que assume-se que não existem pontos de coleta cobrindo toda a cidade. Ainda assim, o modelo de cobertura adotado neste trabalho não é limitado a redes oportunistas, podendo ser aplicado a uma rede na qual os nós de sensoriamento nos veículos estejam sempre conectados.

No modelo considerado, os ônibus seguem trajetos que são compostos de trechos de ruas. Um trecho de rua é definido como um fragmento atômico de uma rua. Um trecho de rua pode ser o intervalo compreendido entre duas esquinas consecutivas, por exemplo. Assim sendo, um trajeto de um ônibus é definido por uma sequência de trechos de rua pelos quais esse mesmo ônibus passa. Como os ônibus possuem trajetos diferentes pela cidade, cada ônibus é capaz de cobrir um determinado conjunto de trechos de ruas. Alguns trechos fazem parte do trajeto de mais de um ônibus, ao mesmo tempo em que alguns outros trechos não fazem parte do trajeto de nenhum ônibus. Esses últimos não podem ser sensoriados pela rede estudada e estão fora do escopo deste trabalho, devendo receber alguma forma complementar de sensoriamento.

3.1. Cobertura em Função de Trechos de Ruas

Neste trabalho, considera-se que um trecho rua é coberto quando pelo menos um ônibus equipado com um nó de sensoriamento contenha esse trecho de rua no seu trajeto. Dessa forma, a cobertura da cidade se dá em termos da cobertura de cada trecho de rua da cidade, como definido na Equação 1, na qual \mathcal{I}_s é o conjunto de trechos de rua cobertos, \mathcal{J} é o conjunto dos ônibus que circulam pela cidade e \mathcal{T}_j é o conjunto de trechos de rua pelo qual o ônibus j passa ao longo de seu trajeto. A Equação 1 define o conjunto de trechos cobertos pela rede estudada, denominado \mathcal{I}_s , como a união de todos os trechos de rua cobertos por cada ônibus j . A Tabela 1 contém a notação utilizada na Equação 1 e no restante deste trabalho.

$$\mathcal{I}_s = \bigcup_{j \in \mathcal{J}} \mathcal{T}_j. \quad (1)$$

A Figura 1 exemplifica essa situação. Nessa figura, o trajeto do ônibus X passa pelo conjunto de trechos $T_X = \{D, H, L, P\}$; o trajeto do ônibus Y passa pelo conjunto de trechos $T_Y = \{A, H, O\}$ e o trajeto do ônibus Z passa pelo conjunto $T_Z = \{C, F, I, P\}$. Assim, o conjunto \mathcal{I}_s de trechos cobertos é a união $T_X \cup T_Y \cup T_Z = \{A, C, D, F, H, I, L, O, P\}$.

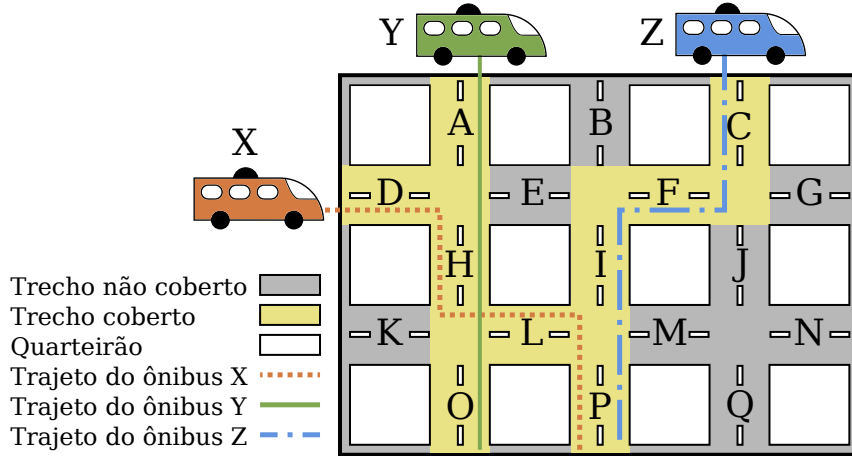


Figura 1. Cobertura de trechos de ruas por nós de sensoriamento embarcados em ônibus urbanos.

No caso geral, os trechos podem possuir comprimentos diferentes. Por isso, a cobertura deve levar em conta também o comprimento de cada trecho que é coberto. A Equação 2 define a cobertura da rede quando cada trecho i possui um comprimento a_i . Nessa equação, a cobertura é definida como o somatório dos comprimentos de todos os trechos de rua cobertos pela rede:

$$C = \sum_{i \in \mathcal{I}_s} a_i. \quad (2)$$

É possível associar uma variável y_i a cada trecho de rua $i \in \mathcal{I}$, de forma que:

$$y_i = \begin{cases} 1, & \text{se } i \in \mathcal{I}_s \\ 0, & \text{se } i \notin \mathcal{I}_s. \end{cases}$$

Assim, y_i indica se o trecho i está coberto ou não, e a Equação 2 é reescrita como:

$$C = \sum_{i \in \mathcal{I}} a_i y_i. \quad (3)$$

Se todos os ônibus da cidade fizerem parte da WSN, obtém-se a máxima cobertura possível. Entretanto, essa situação pode gerar de um lado custos proibitivos de instalação e manutenção dos sensores. Por outro lado, pode-se ter mais nós sensoriando uma área que necessário, desperdiçando recursos. Assim, pode ser necessário estabelecer um número de ônibus para receber nós de sensoriamento e, então, o problema é de escolher quais ônibus devem ser responsáveis pelo sensoriamento. Seja \mathcal{J} o conjunto dos ônibus que circulam por uma cidade, p o limite do número de ônibus que possam ser equipados com nós de sensoriamento, função de um orçamento máximo desejado. Deve ser escolhido um conjunto $\mathcal{J}_s \subset \mathcal{J}$ para receberem nós de sensoriamento de forma que $|\mathcal{J}_s| = p$. Como os ônibus podem cobrir diferentes conjuntos de trechos de ruas, com diferentes comprimentos, a escolha de quais ônibus recebem nós de sensoriamento pode afetar a cobertura da rede, posto que o conjunto \mathcal{I}_s é justamente o conjunto de trechos

de rua pelos quais os ônibus em \mathcal{J}_s passam. Para cada trecho de rua $i \in \mathcal{I}$, pode-se definir um conjunto \mathcal{N}_i dos ônibus que passam por i . Se para algum ônibus $j \in \mathcal{N}_i$, $j \in \mathcal{J}_s$ é considerado coberto. Assim sendo, existe um problema de escolher $\mathcal{J}_s \subset \mathcal{J}$ com cardinalidade p de forma a maximizar a cobertura dada por 2.

3.2. Formulação em Programação Mista

O problema é modelado como um problema de programação linear mista. As convenções de notação utilizadas podem ser encontradas na Tabela 1.

$$\text{maximizar: } \sum_{i \in \mathcal{I}} a_i y_i \quad (4)$$

$$\text{sujeito a: } \sum_{j \in \mathcal{N}_i} x_j \geq y_i, \quad \forall i \in \mathcal{I}; \quad (5)$$

$$\sum_{j \in \mathcal{J}} x_j = p; \quad (6)$$

$$x_j = (0, 1), \quad \forall j \in \mathcal{J}; \quad (7)$$

$$y_i = (0, 1), \quad \forall i \in \mathcal{I}. \quad (8)$$

O objetivo do problema é maximizar a cobertura espacial dos nós de sensoriamento. A equação 4 define a cobertura como a soma dos comprimentos dos trechos de ruas que são cobertos pelos ônibus escolhidos. A Equação 5 define que um trecho j só pode ser considerado coberto se pelo menos um dos ônibus i que passam por j for escolhido. É importante notar que, por ser um problema de maximização, não são necessárias restrições para limitar inferiormente o valor das variáveis y_i , pois essas variáveis contribuem positivamente para a função objetivo, assumindo sempre seu maior valor possível. A Equação 6 garante que o número de ônibus escolhidos seja p . As Equações 7 e 8 caracterizam as variáveis x_j e y_i como binárias. Com a resolução do problema, cada variável x_j igual a 1 indica a escolha do ônibus j para receber um nó de sensoriamento e cada variável x_j igual a 0 indica a rejeição do ônibus j para receber um nó de sensoriamento.

A próxima seção demonstra como o problema de escolher o conjunto de ônibus a serem equipados com nós de sensoriamento, de forma a maximizar a cobertura da rede, é equivalente ao Problema de Localização de Cobertura Máxima [Church e Velle, 1974].

3.3. Problema de Localização de Cobertura Máxima

O problema estudado neste trabalho pode ser modelado como um Problema de Localização de Cobertura Máxima [Church e Velle, 1974] (*Maximal Covering Location Problem* - MCLP). Nesse problema, há um conjunto \mathcal{I} de demandas distribuídas no espaço. Para cada demanda $i \in \mathcal{I}$, há um valor a_i associado, relativo ao benefício que se tem ao atender essa demanda. Essas demandas devem ser atendidas por um conjunto \mathcal{J} de candidatos a instalações, também distribuídos no espaço. Para cada demanda $i \in \mathcal{I}$ existe um conjunto $\mathcal{N}_i \subset \mathcal{J}$ de todos os candidatos a instalações que podem suprir a demanda i , caso sejam instaladas.

A transformação do problema estudado em um MCLP consiste em transformar os trechos de ruas em demandas e os ônibus em candidatos a instalações. O benefício de atender a uma demanda é equivalente ao comprimento de um trecho de rua, enquanto o conjunto de instalações que podem atender uma determinada demanda é o conjunto de ônibus cujo trajeto passa por um determinado trecho de rua.

Na definição do MCLP, uma instalação pode suprir uma demanda se a menor distância d_{ij} entre a demanda i e a instalação j for menor do que um valor S estabelecido. No problema estudado por este trabalho, não existe uma distância entre um ônibus e um trecho de rua. Entretanto, a transformação é possível, pois a distância entre uma demanda e um candidato a instalação serve apenas para a construção dos conjuntos \mathcal{N}_i . No MCLP, se $d_{ij} < S$, então $j \in \mathcal{N}_i$; no problema estudado, se i faz parte do trajeto de j , então $j \in \mathcal{N}_i$. Como é possível converter os problemas de forma simétrica, eles são equivalentes. Portanto, o problema estudado é NP-Difícil, assim como o MCLP [Current e Schilling, 1990].

A seguir, dados reais dos ônibus da cidade do Rio de Janeiro são coletados, pré-processados e utilizados como entrada para o MILP formulado.

4. Análise Experimental

O modelo proposto neste trabalho é utilizado para analisar o sensoriamento das ruas da cidade do Rio de Janeiro. São coletados dados reais da posição dos ônibus pela cidade, que são pré-processados para que representem trechos de ruas. Em seguida, os dados servem de entrada para o MILP formulado na Seção 3.2. A partir da solução desse problema, são obtidas soluções sobre os conjuntos de ônibus capazes de maximizar a cobertura da rede, para diferentes números de ônibus equipados com nós de sensoriamento. Esses passos são detalhados a seguir.

4.1. Coleta de dados

A Prefeitura da Cidade do Rio de Janeiro fornece uma API que entrega dados de sensores de GPS dos ônibus em circulação pela cidade, atualizados uma vez por minuto [IPLANRIO, 2016]. Cada dado consiste de uma tupla contendo a identificação do ônibus, as coordenadas de posição e o instante em que as coordenadas foram adquiridas pelo GPS, para cada ônibus da cidade. A API fornece também outras informações, como a velocidade dos ônibus e as linhas de cada ônibus. Porém, esses dados apresentam incompletudes ou inconsistências. A correção dessas imperfeições está fora do escopo deste trabalho, portanto, tais dados são desconsiderados. Cada uma dessas tuplas é denominada posição de um ônibus.

Neste trabalho foram coletados dados da API por um período de 24 horas, entre 0:00h do dia 10 de maio de 2017 e 0:00h do dia 11 de maio de 2017. Ao realizar, para cada ônibus, a ordenação temporal de suas posições, obtém-se uma aproximação do seu trajeto, minuto a minuto. Esse conjunto de posições é chamado conjunto coletado.

4.2. Pré-processamento dos dados

Os dados coletados precisam ser pré-processados para estimar, com o modelo proposto, a cobertura da cidade pelos ônibus. Inicialmente, é necessário dividir em trechos as ruas que fazem parte dos trajetos dos ônibus. Adicionalmente, é necessário mapear os

Tabela 2. Atributos dos conjuntos de dados gerados.

Atributo	Valor	Conjunto
Total de posições coletadas (#)	5.496.878	Coletado
Número de ônibus no conjunto original (#)	6.075	Coletado
Número de posições removidas durante a filtragem (#)	1.384.925	Coletado
Número de posições após a filtragem (#)	4.111.953	Coletado
Número de ônibus removidos durante a filtragem (#)	328	Coletado
Número de ônibus após a filtragem (#)	5.747	Coletado
Número de posições após a estimação (#)	52.250.671	Estimado
Número de trechos de rua (#)	95.992	Estimado
Soma dos comprimentos dos trechos de rua (km)	5.655,002	Estimado

ônibus nos trechos de rua que fazem parte de seus trajetos. Por último, deve-se calcular o comprimento de cada trecho de rua. Finalmente, é possível construir os conjuntos de dados necessários para a resolução do problema formulado na Seção 3.2. Assim, em linhas gerais, o pré-processamento visa transformar as posições coletadas nos conjuntos e parâmetros indicados na Tabela 1.

Durante o pré-processamento, são realizados três procedimentos: filtram-se as posições que se assemelham a ruídos; a seguir, utiliza-se a API Google *Snap To Roads* [Google, 2017] para obter uma estimativa corrigida dos trajetos coletados e uma divisão desses trajetos em trechos de rua; por último, estimam-se os comprimentos dos trechos de rua obtidos no segundo procedimento. O primeiro procedimento transforma o conjunto de dados coletado. O conjunto de dados criado durante o segundo procedimento e alterado durante o terceiro procedimento é denominado conjunto estimado de dados.

O primeiro procedimento de pré-processamento, a filtragem das posições, é necessário pois dados de GPS possuem erros [Kaplan e Hegarty, 2006] e porque alguns ônibus, apesar de estarem parados, mantêm seus aparelhos de GPS ligados. Assim, os erros criam nos dados de posição a ilusão de que há uma trajetória em torno do local onde tais ônibus estão estacionados (p.ex., garagens e pontos finais). Assim, é realizada uma filtragem nos trajetos coletados: se a distância entre duas posições consecutivas for menor do que um limite estabelecido, a segunda é descartada, até que não haja mais duas posições consecutivas com distância menor do que o limite. A primeira posição é sempre mantida e o cálculo de distância é sempre feito a partir de uma posição que não pode mais ser removida, para evitar que um trajeto de um ônibus que de fato se deslocou pela cidade não seja descartado. Escolhe-se o limite de 10 m, pois esse é um erro típico para aparelhos de GPS de uso comercial [Kaplan e Hegarty, 2006]. Após essa filtragem, alguns ônibus têm todas as suas posições filtradas e, portanto, não fazem mais parte do conjunto de dados estudado. A Tabela 2 contém alguns dos atributos dos conjuntos de dados obtidos.

Ao longo de um minuto, um ônibus pode passar por diversos trechos de rua. Dessa forma, o conjunto de dados coletados não possui o trajeto completo dos ônibus em termos de trechos de rua. O segundo procedimento do pré-processamento tem por objetivo estimar os trajetos completos dos ônibus, estimando todos os trechos de rua pelos quais um ônibus passou. Esse procedimento é realizado com a API Google *Snap To Roads*. Essa API, a partir de uma sequência de coordenadas geográficas que fazem parte da trajetória de um veículo, estima a trajetória do veículo ajustada à topologia das vias mais próximas. A resposta é entregue na forma de uma sequência de coordenadas geográficas associadas, cada uma, a um identificador de lugares, denominado *Place ID*. Os *Place IDs*

retornados pela API *Snap To Roads* identificam os trechos de rua que compõem o trajeto estimado. Assim, uma sequência de posições do conjunto de dados coletado pode ser transformada em uma sequência de posições estimadas, ajustadas para a topologia das ruas. Uma posição estimada é uma tupla com o identificador do ônibus, as coordenadas geográficas e o *Place ID* que identifica o trecho de rua ao qual aquela coordenada geográfica pertence.

Um ônibus pode estar em diferentes posições geográficas no mesmo trecho de rua. Dessa forma, diversos pontos consecutivos de um mesmo trajeto podem possuir o mesmo *Place ID*, representando a passagem do veículo pelo trecho de rua. Uma consideração importante é que os trajetos estimados possuem, em geral, mais posições do que os coletados. Isso ocorre pois duas posições coletadas para um ônibus são separadas no tempo por um minuto e, no caso geral, um minuto de trajeto de um ônibus gera um trajeto estimado com mais do que duas posições.

O terceiro e último procedimento do pré-processamento é o cálculo do comprimento de cada trecho de rua que foi capturado. Tal cálculo é realizado utilizando, para um mesmo trajeto, a distância entre a primeira e o última coordenadas de uma sequência posições com o mesmo *Place ID*. Caso o mesmo *Place ID* apareça em mais de um trajeto, o maior valor calculado é utilizado.

A Figura 2 mostra a CCDF (função complementar de distribuição cumulativa) dos comprimentos obtidos para cada trecho de rua. O maior comprimento tem cerca de 2,5 km e menos de 1% dos trechos de rua possuem comprimentos maiores do que 400 m. O conjunto contendo os trajetos estimados e os comprimentos dos trechos de rua é chamado de conjunto estimado. A Tabela 2 contém alguns dos atributos dos conjuntos de dados estimados.

4.3. Execução dos experimentos

Os dados obtidos nas fases anteriores são utilizados como entrada para o problema formulado na Seção 3.2. O problema é executado utilizando a ferramenta IBM ILOG CPLEX 12.5.1. A ferramenta trabalha definindo limites superiores para a função objetivo (ou seja, para a cobertura) que não necessariamente são viáveis (ou seja, que não necessariamente são mapeáveis em uma configuração de sensores nos ônibus). Devido à complexidade do problema, que impossibilita sua solução para o caso ótimo, a ferramenta é configurada para retornar resultados assim que a diferença entre a melhor solução encontrada e o limite superior encontrado seja de, no máximo, 2.0% do limite superior, encontrando um resultado pessimista. A diferença de 2.0% foi escolhida pois é a menor diferença obtida para todos os valores testados, considerando a melhor máquina da infraestrutura utilizada, um Intel Xeon E5-2650 com 264 GB de memória. Como número de ônibus equipados com nós de sensoriamento, são utilizados os valores de 2, 4, 8, 16, 32, 64, 128, 256, 512, 1024, 2048 e 4096, de um total de 5747 ônibus. Os resultados obtidos estão detalhados a seguir na Seção 4.4.

4.4. Resultados

Os resultados dos experimentos (Figura 3) mostram a proporção de cobertura em percentual, para diferentes números de ônibus equipados com nós de sensoriamento. A proporção de cobertura é calculada com relação à cobertura realizada caso todos os ônibus

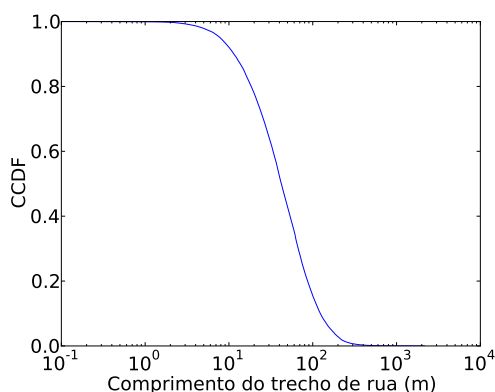


Figura 2. Distribuição dos comprimentos dos trechos de rua capturados com a ferramenta Snap to Roads.

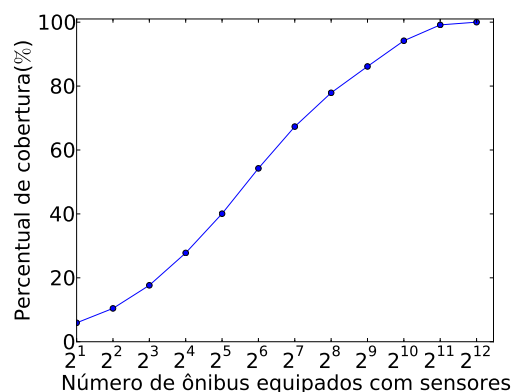


Figura 3. Porcentagem da cobertura da rede com relação à cobertura de todos os ônibus comparada com a quantidade de ônibus equipados com sensores.

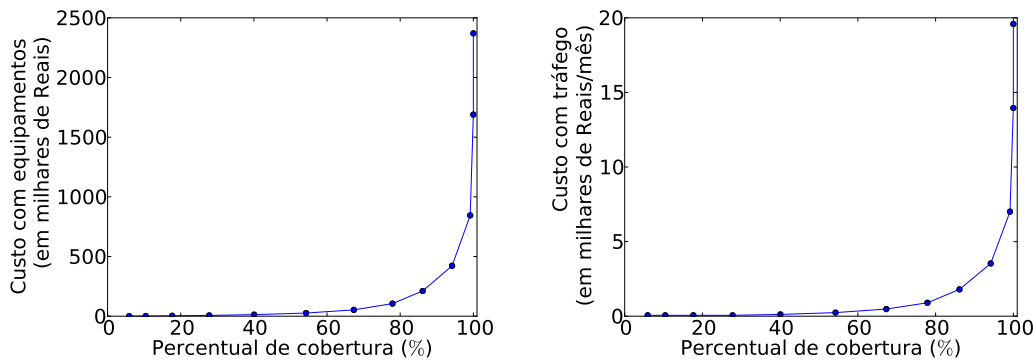
sejam equipados com nós de sensoriamento. O eixo horizontal da figura está em escala logarítmica, para facilitar a visualização.

Pelos resultados expostos na Figura 3, é possível inferir que 1024 ônibus equipados com nós de sensoriamento, ou seja, 18% da frota, podem cobrir pelo menos 94% das ruas do Rio de Janeiro onde há circulação de ônibus, ou o equivalente a cerca de 5606 km de ruas. Para fins de comparação, a Prefeitura da Cidade do Rio de Janeiro informa que o total de logradouros da cidade é 10.577 km [Passos, 2017]. Outra observação é que com 32 ônibus é possível cobrir pelo menos 40% da região coberta por todos os ônibus. Isso significa que é possível iniciar, a um custo menor do que o custo da rede completa, uma rede que ofereça alguns serviços de sensoriamento. A partir dessa rede inicial, pode-se realizar um desenvolvimento incremental para suportar mais serviços, já havendo uma rede com cobertura significativa da cidade.

Utilizando como referência o protótipo elaborado em [Cruz et al., 2017c], é possível calcular um orçamento aproximado para os equipamentos da rede para diferentes percentuais de cobertura, já que esta é função do número de ônibus com equipamentos instalados. Para evitar a complexidade da comunicação oportunística, que não é o foco deste trabalho, assume-se que o nó de sensoriamento realiza a entrega de dados através de uma rede M2M (*Machine-to-Machine*). Portanto, recebe uma interface GPRS/M2M com um cartão SIM no lugar de uma interface 802.11b/g/n. A Tabela 3 contém estimativas de preços para os equipamentos baseadas em uma pesquisa em lojas virtuais na Internet, considerando preço unitário. Pode-se esperar que, ao se comprar em quantidades maiores, os custos sejam mais baixos. Esse custo é denominado custo inicial de equipamentos da rede e acontece na ocasião da instalação da rede. Utilizando a quantidade de dados gerada pelo protótipo a cada segundo, é possível concluir que cada nó de sensoriamento gera cerca de 142 MB/mês. É necessário, então, um plano de dados de uma operadora para a entrega dos dados. Uma nova estimativa é feita a partir de lojas virtuais e é encontrado um plano M2M com franquia de 2,5GB ao custo de R\$ 59,90/mês, podendo ser compartilhada por mais de um equipamento. A Figura 4(b) mostra os custos mensais com o tráfego da rede, para diferentes valores de cobertura.

Tabela 3. Preços unitários dos equipamentos do protótipo da rede.

Módulo	Equipamento	Fabricante	Preço (R\$)
Controlador	Arduino UNO R3	Arduino	44,90
Receptor GNSS	GS-96U7	Guangzhou Xintu	249,90
Memória Persistente	GS-96U7	Guangzhou Xintu	
Banco de Sensores	Sensor de Umidade DHT11	DFRobot	12,90
	Sensor de Temperatura DHT11	DFRobot	
	Sensor de Intensidade Luminosa GL5528	GBK Robotics	9,90
	Sensor de Intensidade de Chuva YI-83	GBK Robotics	12,90
Interface GPRS	GSM GPRS SIM800L	Coreboard	74,90
	Cartão SIM	TIM	7,00
Preço Total			412,40

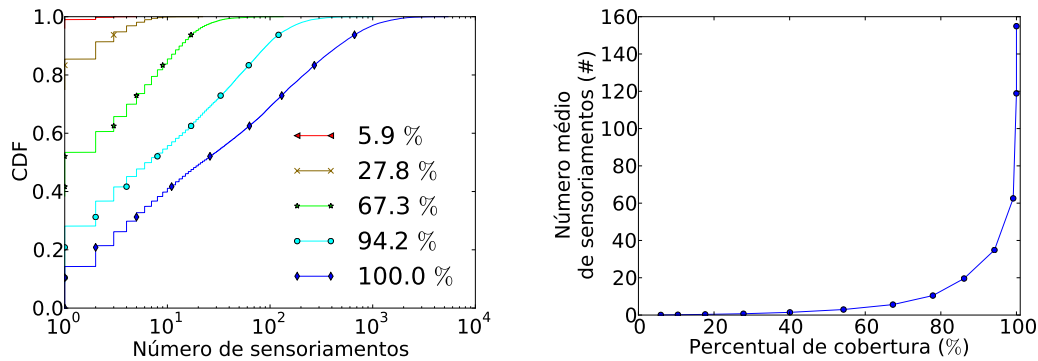


(a) Custo inicial dos equipamentos da rede para diferentes percentuais de cobertura. (b) Custo mensal do tráfego da rede para diferentes percentuais de cobertura.

Figura 4. Custos de instalação e de tráfego da rede estudada.

A Figura 4(a) apresenta o custo estimado dos equipamentos necessários para obter diferentes percentuais de cobertura da cidade, com relação à cobertura máxima possível. Pode-se observar que, com menos de R\$14.000,00, é possível adquirir equipamentos suficientes para cobrir pelo menos 40% da cobertura da rede completa. A Figura 4(b) ilustra o custo mensal do tráfego da rede, para diferentes percentuais de cobertura. É possível observar que, para cobrir cerca de 40% da cobertura completa, há um custo de aproximadamente R\$ 120,00 mensais. Estas estimativas não contam com todos os custos operacionais da rede, como pessoal e reposição de equipamentos defeituosos.

Os experimentos executados possuem foco na cobertura espacial da rede. No sensoriamento móvel, há um compromisso entre cobertura temporal e cobertura espacial [Liu et al., 2005]. Assim, é importante saber quantas vezes um mesmo trecho é sensoriado ao longo de um dia. Quanto mais vezes um trecho é sensoriado, maior a probabilidade de se detectar um evento de interesse e também maior a eficácia de algoritmos de correção de erros e de predição de valores [Sta, 2016], [Zhao et al., 2015]. Para estimar o efeito do número de ônibus na frequência de sensoriamento de cada trecho de rua quando os ônibus são escolhidos para a maximização da cobertura, calcula-se o número de vezes que cada trecho de rua é visitado por algum ônibus contido no conjunto solução do problema executado. A Figura 5(a) mostra a CDF do número de vezes que um mesmo



(a) CDF dos sensoriamentos de trechos por um dia, para diferentes percentuais de cobertura. (b) Média de sensoriamentos de cada trecho por um dia em função do percentual de cobertura.

Figura 5. Frequência de sensoriamento dos trechos de rua ao longo de um dia.

trecho é sensoriado ao longo de um dia para os percentuais de cobertura de 5,9%, 27,8%, 67,3%, 94,2% e 100%. Esses percentuais de cobertura são obtidos utilizando 2, 16, 128, 1024 e 5747 ônibus, respectivamente. O eixo horizontal está em escala logarítmica.

Outra forma de visualizar a relação do número de ônibus com a frequência com a qual cada trecho de rua é sensoriado está ilustrado na Figura 5(b), que exibe no eixo horizontal o número de ônibus equipados com sensores e, no eixo vertical, a média de sensoriamentos sofrida por cada trecho de rua, no intervalo de um dia. Observa-se que o crescimento se acentua quando a cobertura é cerca de 50% da cobertura total da rede. Adicionalmente, a cobertura da rede cresce mais rapidamente com o aumento do número de ônibus com sensores quando comparada com o crescimento no aumento do número de sensoriamentos diários dos trechos em função do número de ônibus. Esse resultado reforça a noção de que é possível desenvolver a rede de sensoriamento de forma incremental, atendendo inicialmente os requisitos de alguns serviços para uma determinada região da cidade e expandindo a área e a gama de serviços que podem ser oferecidos. O trabalho [Cruz et al., 2017a] visa reduzir os custos de tráfego da rede, que podem aumentar bastante, de acordo com a Figura 4(b).

Utilizando requisitos de intervalo de sensoriamento de [Zanella et al., 2014] e os intervalos de sensoriamento esperados para a rede, conclui-se que uma rede com 1024 ônibus é capaz de suportar aplicações como monitoramento da qualidade do ar ou controle de tráfego para cerca de 40% dos trechos de rua cobertos.

5. Conclusões e Trabalhos Futuros

A cobertura de toda uma área urbana com uma rede de sensores pode ser proibitivamente custosa. Uma opção para reduzir os custos e ainda assim obter uma significativa cobertura espacial é utilizar nós móveis de sensoriamento. Cada nó é responsável por sensoriar uma região maior, ao mesmo tempo em que a rede de coleta não precisa cobrir a cidade inteira, podendo ser acessada pelos nós de sensoriamento de forma oportunística. Uma forma barata de dotar de mobilidade os nós de sensoriamento é embarcá-los em ônibus urbanos. Além de, em geral, circularem por grande parte das cidades, o custo energético de carregar um nó de sensoriamento é desprezível para um ônibus.

Neste trabalho foi analisada a cobertura espacial de uma rede de sensores baseada em ônibus urbanos. Foi proposto um modelo de cobertura de acordo com os trechos de rua a sensoriar, os ônibus equipados com nós de sensoriamento e seus trajetos. Um problema de programação inteira foi formulado para maximizar a cobertura da rede assumindo uma limitação no número de ônibus que são equipados com sensores, visando a limitação do custo. Dados de trajetos de ônibus da cidade do Rio de Janeiro foram coletados, pré-processados e utilizados como entrada para o problema formulado. Os resultados mostraram que 18% da frota consegue cobrir pelo menos 94% das ruas do Rio de Janeiro onde circulam ônibus. Adicionalmente, os resultados mostraram que com 32 nós de sensoriamento cobre-se pelo menos 40% da cidade. Assim, é possível desenvolver uma rede inicial com custo de R\$ 14.000,00 em equipamentos e R\$120,00/mês em tráfego, e desenvolver outros serviços de forma progressiva. Observou-se também que o número de vezes que um trecho qualquer é sensoriado ao longo do dia depende do número de ônibus. Há uma relação entre o número de ônibus, a cobertura da rede e o número de vezes que um trecho é visitado ao longo do tempo.

Pretende-se explorar no futuro os trajetos das linhas de ônibus com o intuito de maximizar a cobertura da rede, assim como investigar heurísticas para uma resolução mais rápida do problema e analisar com mais detalhes a frequência com a qual um mesmo trecho é sensoriado.

Referências

- Church, R. e Velle, C. R. (1974). The maximal covering location problem. *Papers in regional science*, 32(1):101–118.
- Cruz, P., Couto, R. S. e Costa, L. H. M. K. (2017a). An algorithm for sink positioning in bus-assisted smart city sensing. *Future Generation Computer Systems*.
- Cruz, P., Couto, R. S. e Costa, L. H. M. K. (2017b). Um algoritmo de posicionamento de pontos de coleta para uma rede de sensores baseada em ônibus urbanos. *XXXV Simpósio Brasileiro de Redes de Computadores e Sistemas Distribuídos*.
- Cruz, P., da Silva, F. F., Pacheco, R. G., Couto, R. S., Velloso, P. B., Campista, M. E. M. e Costa, L. H. M. K. (2017c). Sensingbus: um sistema de sensoriamento baseado em ônibus urbanos. *XXXV Simpósio Brasileiro de Redes de Computadores e Sistemas Distribuídos*.
- Current, J. R. e Schilling, D. A. (1990). Analysis of errors due to demand data aggregation in the set covering and maximal covering location problems. *Geographical Analysis*, 22(2):116–126.
- Dong, W., Guan, G., Chen, Y., Guo, K. e Gao, Y. (2015). Mosaic: Towards city scale sensing with mobile sensor networks. Em *IEEE International Conference on Parallel and Distributed Systems (ICPADS)*, p. 29–36. IEEE.
- Du, R., Chen, C., Yang, B., Lu, N., Guan, X. e Shen, X. (2015). Effective urban traffic monitoring by vehicular sensor networks. *IEEE Transactions on Vehicular Technology*, 64(1):273–286.
- Ekici, E., Gu, Y. e Bozdogan, D. (2006). Mobility-based communication in wireless sensor networks. *IEEE Communications Magazine*, 44(7):56 – 62.

- Fiore, M., Nordio, A. e Chiasserini, C.-F. (2016). Driving factors toward accurate mobile opportunistic sensing in urban environments. *IEEE Transactions on Mobile Computing*, 15(10):2480–2493.
- Google (2017). Google maps API snap to roads. Disponível em <https://developers.google.com/maps/documentation/roads/snap>.
- Gubbi, J., Buyya, R., Marusic, S. e Palaniswami, M. (2013). Internet of things (IoT): A vision, architectural elements, and future directions. *Future Generation Computer Systems*, 29(7):1645–1660.
- IPLANRIO (2016). Descrição do dataset conjunto GPS ônibus. Disponível em <http://data.rio/dataset/gps-de-onibus>.
- Kamienski, C., Biondi, G. O., Borelli, F. F., Heideker, A., Ratusznei, J. e Kleinschmidt, J. H. (2016). Computação urbana: Tecnologias e aplicações para cidades inteligentes. *Minicursos SBRC*.
- Kaplan, E. D. e Hegarty, C. J. (2006). Understanding gps-principles and applications second edition. artech house. *Inc., MA*, p. 153–173.
- Liu, B., Brass, P., Dousse, O., Nain, P. e Towsley, D. (2005). Mobility improves coverage of sensor networks. Em *Proceedings of the 6th ACM international symposium on Mobile ad hoc networking and computing*. MobiHoc.
- Marjovi, A., Arfire, A. e Martinoli, A. (2015). High resolution air pollution maps in urban environments using mobile sensor networks. Em *International Conference on Distributed Computing in Sensor Systems (DCOSS)*, p. 11 – 20. IEEE.
- Passos, I. P. (2017). Rio em síntese. Disponível em <http://www.data.rio/pages/rio-em-sntese-2>.
- Pipino, L. L., Lee, Y. W. e Wang, R. Y. (2002). Data quality assessment. *Communications of the ACM*, 45(4):211–218.
- Santos, B. P., Silva, L. A., Celes, C. S., Borges, J. B., Neto, B. S. P., Vieira, M. A. M., Vieira, L. F. M., Goussevskaia, O. N. e Loureiro, A. A. (2016). Internet das coisas: da teoria à prática. p. 1–50.
- Sta, H. B. (2016). Quality and the efficiency of data in “smart-cities”. *Future Generation Computer Systems*.
- Zanella, A., Bui, N., Castellani, A., Vangelista, L. e Zorzi, M. (2014). Internet of things for smart cities. *IEEE Internet of Things Journal*, 1(1):22–32.
- Zhao, D., Ma, H., Liu, L. e Li, X.-Y. (2015). Opportunistic coverage for urban vehicular sensing. *Computer Communications*, 60:71–85.
- Zhu, Y., Li, Z., Zhu, H., Li, M. e Zhang, Q. (2013). A compressive sensing approach to urban traffic estimation with probe vehicles. *IEEE Transactions on Mobile Computing*, 12(11):2289–2302.
- Zoysa, K. D., Keppitiyagama, C., Seneviratne, G. P. e Shihan, W. W. A. T. (2007). A public transport system based sensor network for road surface condition monitoring. Em *ACM SIGCOMM (NSDR)*. ACM.