

Localização de Usuários Móveis Baseada em *Fingerprint* de Rádio Frequência: Redução de Espaço de Busca usando Parâmetros de Atraso de Onda das Redes Celulares

Guilherme H. S. Silva¹, Thiago A. L. Domingues¹
Gabriel W. A. Silva¹, Daniel C. Cunha¹

¹Centro de Informática – Universidade Federal de Pernambuco (UFPE)
50740-560 – Recife – PE – Brazil

{ghss, tald, gwas, dcunha}@cin.ufpe.br

Abstract. *Radiolocation techniques based on radio frequency fingerprinting have proven to be an interesting alternative to global positioning systems. One of the attractions of the fingerprint technique is the reduction of power consumption through simplified implementations of its steps, such as the mobile user position prediction step. This work presents a modification in the search space reduction technique using the timing advance (TA), a wave delay parameter of cellular networks. Results showed that the prediction time was reduced by an average of 72.63% compared to the original technique, while the training time increased by an average of 60.18%, without compromising the accuracy of the location.*

Resumo. *As técnicas de radiolocalização baseadas em fingerprint de rádio frequência têm se mostrado uma alternativa interessante frente aos sistemas de posicionamento global. Um dos atrativos da técnica de fingerprint é a redução do consumo de energia por meio de implementações simplificadas de suas etapas, como, por exemplo, na etapa de predição da posição do usuário móvel. Este trabalho apresenta uma modificação na técnica de redução de espaço de busca que utiliza o timing advance (TA), parâmetro de atraso de onda das redes celulares. Os resultados mostraram que o tempo de predição foi reduzido, em média, 72,63% em relação à técnica original, enquanto o tempo de treinamento sofreu um acréscimo médio de 60,18%, sem comprometer a acurácia da localização.*

1. Introdução

Nos últimos anos, os avanços dos sistemas de comunicações tem contribuído significativamente para o crescimento da utilização de smartphones para acessar uma variedade de serviços de Internet [Deville et al. 2014]. Tais serviços são representados por uma diversidade de aplicações, quais sejam, segurança em transações de cartões de crédito [Hammad and Faith 2017], métodos de pagamento baseados em localização [Salomon and Mahaffey 2019], assistência ao motorista na estrada [Trogh et al. 2019] e rastreamento de doenças [Nyhan et al. 2019]. Nesse contexto, a informação sobre a localização dos usuários se torna crucial para a qualidade e a confiabilidade das aplicações mencionadas.

O processo de geolocalização mais utilizado em sistemas de localização ao redor do mundo é o sistema de posicionamento global (GPS, *global positioning system*) [Peral-Rosado et al. 2018]. Apesar de ter excelentes resultados em termos de acurácia, o GPS possui severas limitações quando utilizados em ambientes densamente urbanos e sob condições climáticas adversas [Beekhuizen et al. 2013]. Além disso, o hardware do GPS é um dos módulos que mais consome energia em dispositivos móveis, comprometendo a autonomia de suas baterias [Chen et al. 2019].

Devido às limitações do GPS, outras alternativas como, por exemplo, as técnicas de radiolocalização tem recebido bastante atenção nas últimas décadas [Peral-Rosado et al. 2018]. Incluídas nesta categoria, as técnicas de *fingerprinting* de rádio frequência (RF) apresentam algumas vantagens em relação ao GPS. Primeiramente, as técnicas baseadas em *fingerprint* de RF utilizam, em geral, a intensidade do nível de sinal recebido (RSSI, *received signal strength indicator*) para oferecer boas estimativas de posição, mesmo em ambientes com obstrução de visada direta entre a estação móvel (EM) e a estação rádio base (ERB) [Zhang et al. 2019]. Em outras palavras, diferente do GPS, a localização por *fingerprint* de RF não depende de linha de visada entre transmissor e receptor, o que não impede, por exemplo, a sua utilização em condições climáticas desfavoráveis. Uma segunda vantagem da radiolocalização por *fingerprint* de RF está no seu baixo consumo de energia, o que a torna atrativa para aplicações de Internet das Coisas [Nouichi et al. 2019, Kose et al. 2019].

De maneira geral, as técnicas de *fingerprinting* de RF apresentam estimativas de localização com boa acurácia de maneira estável [Vo and De 2016]. No entanto, além de um bom desempenho em termos de acurácia, é imprescindível que esses sistemas de localização consumam pouca memória e, adicionalmente, encontrem o usuário de forma rápida, ou seja, com um tempo de predição pequeno. Com a adoção do *fingerprinting*, características como o tamanho da área geográfica e a quantidade de ERBs envolvidas na predição podem impactar nessas métricas. Considerando que um dos métodos utilizados para reduzir o tempo de predição aplica a redução no espaço de busca, o presente trabalho apresenta uma proposta de técnica de redução de espaço de busca que utiliza o *timing advance* (TA), parâmetro de atraso de onda das redes celulares. Além disso, a técnica proposta é comparada a outro método de redução de espaço de busca denominado filtragem por TA. A comparação é realizada em termos de erro de estimativa de posição, tempo de treinamento da técnica e tempo de predição da localização do usuário.

O artigo está organizado como se segue. A Seção 2 apresenta os conceitos básicos referentes à localização baseada em *fingerprint* de RF, em especial a construção do mapa de rádio e a redução do espaço de busca associado ao mapa de rádio, e como funciona a filtragem por TA. A proposta de redução de espaço de busca é descrita na Seção 3. Na Seção 4, os resultados são apresentados e, por fim, comentários finais são realizados na Seção 5.

2. Fundamentação Teórica

A predição da posição de uma EM por *fingerprint* utiliza um conjunto de características de um sinal de RF presentes em uma rede de comunicações sem fio [Vo and De 2016, Zekavat and Buehrer 2019]. Um *fingerprint* pode ser classificado como *fingerprint* de referência, caso faça parte de um conjunto de informações coletadas (ou geradas) na área

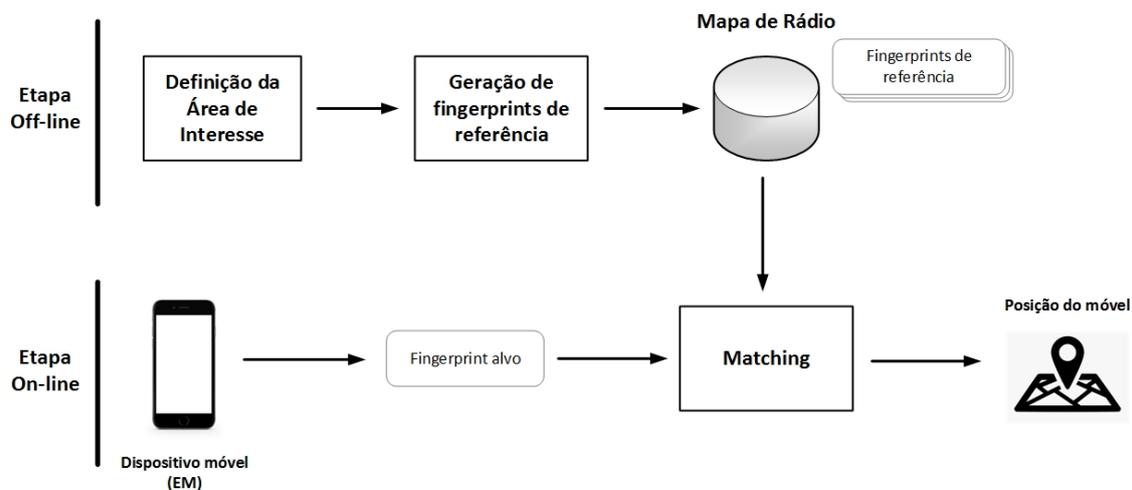


Figura 1. Fluxograma de um sistema de localização baseado em *fingerprint* de RF.

de interesse, ou como um *fingerprint* alvo, caso seja o conjunto de informações aferidas de uma EM, cuja localização se deseja estimar. A estimativa desta posição da EM é feita através da comparação de seu *fingerprint* alvo com um conjunto conhecido de *fingerprints* de referência, analisando qual destes mais se assemelha ao *fingerprint* da EM.

Uma arquitetura geral e simplificada de um sistema de localização baseado em *fingerprint* de RF pode ser vista na Figura 1. De modo geral, esse tipo de sistema de localização possui duas etapas principais, chamadas de *off-line* e *on-line*. Na etapa *off-line*, também conhecida como fase de treinamento, ocorre a escolha da área de interesse onde a EM pode estar localizada. Definida a área de interesse, ocorre a aquisição e/ou a geração de *fingerprints* de referência para a construção de uma base de dados, conhecida por mapa de rádio. Já na etapa *on-line*, também conhecida como fase de teste ou predição, ocorre a estimativa de posição da EM por meio da comparação de *fingerprints*.

A técnica de redução de espaço de busca consiste basicamente na diminuição do conjunto de *fingerprints* de referência para reduzir o tempo de execução da etapa *on-line*. Esta etapa relacionada à restrição do conjunto será detalhada na Subseção 2.2. Após a redução do conjunto de busca, a predição da posição da EM é realizada no passo de *matching*, associando o *fingerprint* de referência que mais se assemelha ao *fingerprint* alvo. Maiores detalhes serão descritos na Seção 3.

2.1. Construção do Mapa de Rádio

O conceito principal da técnica *fingerprinting* de RF consiste na suposição de que cada pequena divisão da área de interesse selecionada possua características únicas (“impressões digitais”) em relação aos níveis de sinais das ERBs. Dessa forma, ao se comparar informações de sinais da EM que se deseja descobrir a localização com as informações disponíveis no mapa de rádio, será possível inferir quais são as coordenadas geográficas (latitude e longitude) da EM.

Para a construção do mapa de rádio, primeiramente é preciso mapear toda a área de interesse de tal forma que possamos definir pontos de possíveis posições da EM. No entanto, registrar todos os pontos da região demandaria uma grande quantidade de

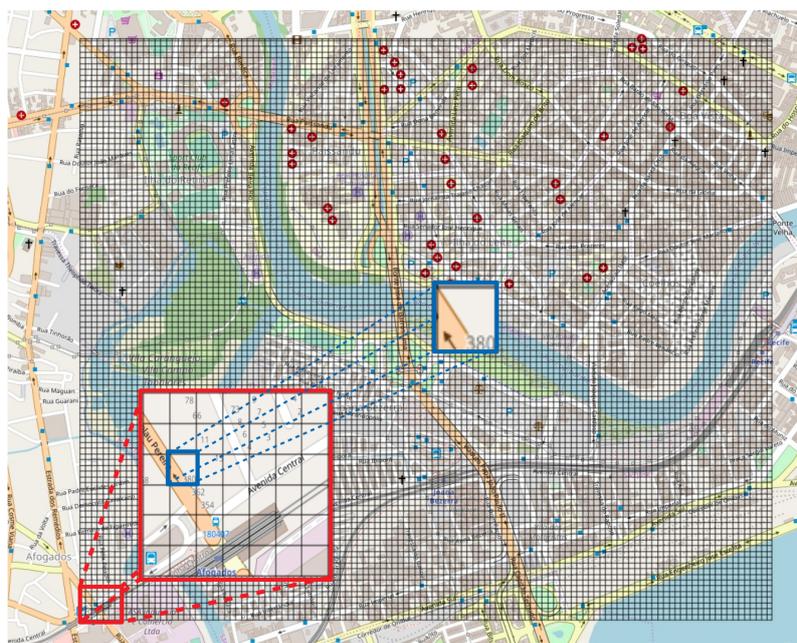


Figura 2. Exemplo de *grid* regular cuja resolução é definida como o tamanho do lado do quadrado azul indicado. Neste exemplo, a resolução vale 20 metros.

memória para armazenar os dados. Isto posto, vamos assumir que a região de interesse é um espaço bidimensional coberto por um *grid* de pequenas regiões denominadas células¹. Tais células são representadas pelas coordenadas geográficas de seu centro. A organização desse *grid* pode ocorrer de maneira regular ou irregular. Nos *grids* regulares, a área de interesse é dividida de tal forma que todas células sejam espaçadas uniformemente. A Figura 2 ilustra a cobertura da região de interesse com um *grid* regular de resolução de 20 m, em que a resolução é definida como o tamanho do lado do quadrado que representa a célula do *grid*. Já nos *grids* irregulares, são utilizadas outras formas de organização, como, por exemplo, assumir que as células são formadas por pontos que cobrem as malhas viárias da região, como mostrado em [Bittencourt et al. 2018]. Independentemente da forma de organização do *grid*, as células que o compõem podem ser vistas como uma lista de possíveis localizações para a EM. Porém, é necessário que estabeleçamos um *fingerprint* para cada uma das células do *grid*, ou seja, cada célula terá um conjunto de informações de RSSI e um par de coordenadas geográficas.

A geração dos *fingerprints* de referência para a construção do mapa de rádio pode ser feita de várias maneiras. Uma delas é por meio de modelos de propagação, sem a necessidade de medições de campo, porém os valores de RSSI são gerados com menos acurácia [Saunders and Aragón-Zavala 2007]. Outra possibilidade é por meio de medições realizadas em campo e registro de coordenadas geográficas. Por meio de uma técnica conhecida como *wardriving*, tais medições e registros de pontos da região são coletados em relação a cada uma das ERBs (geralmente, as ruas são percorridas com um veículo que possui um *scanner* de RF). Entretanto, o trabalho de aferir medições de RSSI para toda a área de interesse é um tanto dispendioso [Mondal et al. 2015]. Além disso, não há garantia de que as medições serão realizadas em todas as células do *grid*, pois

¹Não confundir células do *grid* de localização com as células da rede de telefonia celular.

podem estar situadas em propriedades privadas ou em locais de difícil acesso. Logo, a base de dados não é, em geral, suficiente para construção de todo o mapa de rádio.

Para suprir essa deficiência, modelos de propagação podem ser utilizados para complementar os pontos não obtidos por meio da medição de campo. Outra alternativa é utilizar os dados coletados pela técnica de *wardriving* para treinar algoritmos de aprendizado de máquina com a finalidade de prever os valores de RSSI que a medição em campo não foi capaz de fornecer [Bittencourt et al. 2018, Timoteo et al. 2016]. Dentre as técnicas de regressão que podem ser aplicadas, uma das mais simples é a técnica de regressor k -NN [Noi and Kappas 2018], onde cada regressor utiliza as coordenadas geográficas (latitude e longitude) como atributos e estima o valor de RSSI para cada ERB. Dessa forma, com a combinação das saídas de cada regressor, é possível gerar os *fingerprints* de referência que completarão o mapa de rádio.

2.2. Redução do Espaço de Busca

Assumindo a adoção de um *grid* regular e dependendo de sua resolução, o mapa de rádio construído na etapa *off-line* pode se tornar muito extenso. Por exemplo, considere uma área de interesse representada por um quadrado com 10 *km* de lado e que o *grid* regular adotado para esta região tenha uma resolução de 5 *m*. Dito isso, o *grid* será então formado por 4 milhões de células, o que pode representar um espaço de busca custoso, dependendo da plataforma computacional utilizada na etapa *on-line*.

Definido o problema, algumas técnicas de localização baseadas em *fingerprint* de RF, após a construção do mapa de rádio, se utilizam de alguma abordagem para reduzir o espaço de busca. Conforme já mencionado, a redução de espaço de busca consiste em limitar o conjunto de possíveis células candidatas do mapa de rádio que pode conter a EM procurada. A intenção da redução do espaço de busca é evitar que sejam utilizadas todas as células do mapa de rádio na etapa de *matching*, o que aumentaria consideravelmente o custo computacional da etapa de predição. As técnicas de redução de espaço de busca geralmente empregam não apenas os valores de RSSI medidos em relação às ERBs, mas também alguns parâmetros adicionais. Nesse contexto, é possível utilizar parâmetros da rede celular que podem servir como alguma métrica de similaridade entre o *fingerprint* alvo e os *fingerprints* de referência. Aplicando uma técnica de redução de espaço de busca e tomando o exemplo citado acima, podemos reduzir este espaço de 4 milhões de células para 80.000 células. Isso significa uma redução de 98% dos *fingerprints* de referência usados para predição da localização.

Na literatura, podemos encontrar algumas estratégias para a redução do espaço de busca, como, por exemplo, a técnica de filtragem de CDB [Campos and Lovisolo 2009] e uma baseada em algoritmos genéticos [Campos and Lovisolo 2019]. Uma terceira técnica, denominada filtro por *timing advance* (TA) foi proposta em [Timoteo et al. 2016] e aprimorada em [Bittencourt et al. 2018]. TA é um parâmetro de atraso de onda disponível nas redes celulares GSM (2G, segunda geração) que tem a finalidade de permitir que a ERB controle os atrasos de sinal em sua comunicação com a EM. O sinal do TA é transmitido por meio de um dos canais de controle do sistema como um número inteiro entre 0 e 63. Cada valor do TA corresponde a uma distância aproximada de 550 *m* entre a ERB e a EM. Por exemplo, TA= 0 significa que a EM encontra-se até 550 *m* da ERB; TA= 1, que a EM está entre 550 *m* e 1.100 *m* e assim por diante.

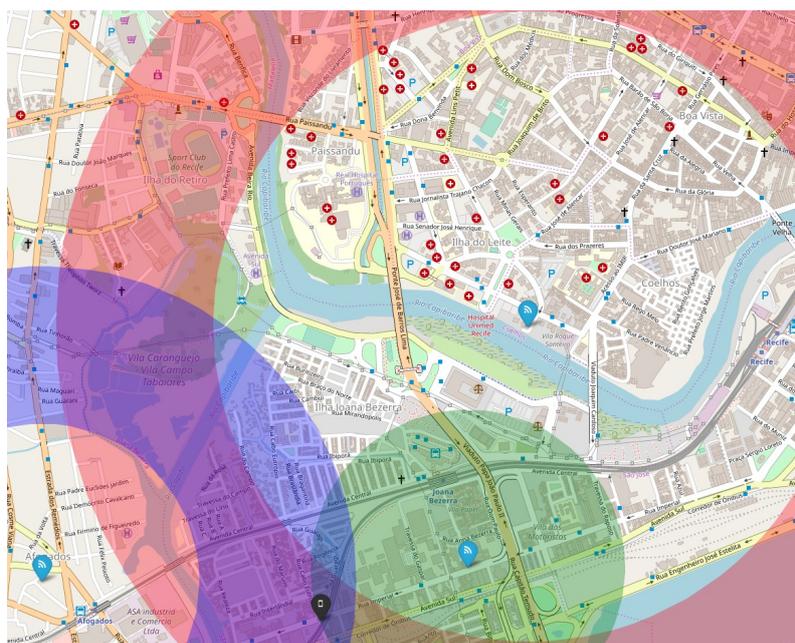


Figura 3. Indicação da sobreposição de três regiões (duas coroas circulares e um círculo) referentes a três ERBs, cujas localizações (centros das regiões) estão indicadas em cor azul. O ponto em cor preta indica a EM cuja posição se deseja encontrar.

Parâmetros de atraso de onda também são empregados nas redes celulares de terceira (3G) e quarta gerações (4G). Nas redes W-CDMA (3G), o parâmetro de atraso de onda é denominado *propagation delay* (PD) e cada valor inteiro do PD representa uma distância de 234 m. Já nas redes LTE (4G), ambos os parâmetros (TA e PD) são empregados. A utilização de parâmetros de atraso de onda nas redes 3G e 4G nos permite afirmar que técnicas de redução de espaço de busca na localização baseada em *fingerprint* de RF podem ser aplicadas não apenas em redes GSM, mas também nas redes celulares mais evoluídas.

A técnica de filtragem por TA se baseia na ideia de que, a partir da combinação dos valores de TA, é possível estabelecer uma área de interseção das coroas circulares (ou eventualmente, um círculo) de cada ERB, que delimita uma região na qual a EM possui alta probabilidade de estar presente. A Figura 3 mostra um exemplo de como seria essa região, caracterizada pela interseção de duas coroas circulares e um círculo. Como temos um conjunto de células do *grid* que mapeia pontos da região de interesse, o problema de redução de espaço de busca se reduz a encontrar as células cujas coordenadas estejam compreendidas na região de interseção na qual consta o ponto procurado (cor preta). Contudo, realizar toda essa operação na etapa de predição implica em um custo computacional elevado, uma vez que será necessário consultar todas as células do mapa de rádio.

3. Técnica Proposta

Para evitar a complexidade associada à localização de cada nova EM (redução do espaço de busca na etapa de predição), propõe-se neste trabalho uma técnica de localização baseada em *fingerprint* de RF que considera uma nova abordagem para a redução do espaço

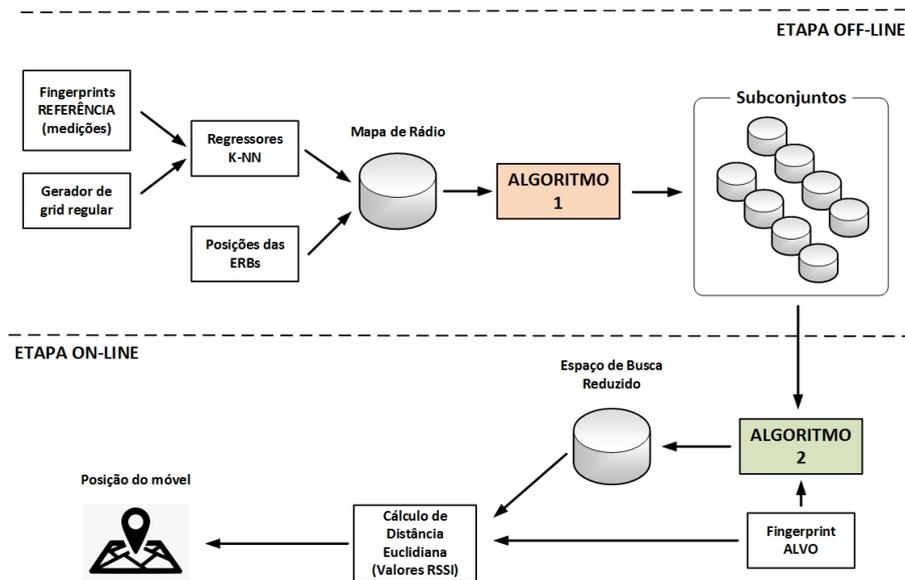


Figura 4. Fluxograma da técnica de localização proposta, com divisão da redução do espaço de busca nas etapas *off-line* (Algoritmo 1) e *on-line* (Algoritmo 2).

de busca. A Figura 4 ilustra o fluxograma da proposta, que consiste em dividir o esforço da redução do espaço de busca em dois algoritmos, sendo o Algoritmo 1 na fase de treinamento e o Algoritmo 2 na fase de predição.

Inicialmente, as etapas de construção do mapa de rádio são semelhantes às aquelas descritas na Subseção 2.1. Durante a fase de treinamento, o Algoritmo 1 é o responsável por gerar vários subconjuntos com a finalidade de armazenar as células do *grid* referentes a cada ERB e cada valor possível de TA. Isto significa que, considerando que existem N ERBs no sistema e M possíveis valores de TA, serão criados NM subconjuntos de células, sendo cada subconjunto associado a uma ERB específica e um valor de TA possível. Por exemplo, considere a Figura 5, na qual temos uma suposta rede celular com $N = 3$ ERBs. Para cada ERB, temos a geração de $M = 63$ subconjuntos (um para cada valor de TA). A geração dos 189 subconjuntos (63 valores de TA por cada ERB) representa a saída do Algoritmo 1 na Figura 4.

Na etapa de predição, dados os valores de TA extraídos da EM a ser localizada (*fingerprint* alvo), inicialmente é necessário identificar os subconjuntos gerados pelo Algoritmo 1 que possuem os mesmos valores de TA para cada ERB. De posse desses subconjuntos, o Algoritmo 2 (indicado na Figura 4) é responsável por gerar o espaço de busca reduzido. Por exemplo, considere novamente a Figura 5, na qual a EM procurada possui os valores de TA 0, 1 e 2 para as ERBs 1, 2 e 3, respectivamente. A partir dos subconjuntos TA=0 (ERB 1), TA=1 (ERB 2) e TA=2 (ERB 3), o espaço de busca reduzido é gerado. Finalmente, o *fingerprint* alvo pode ser buscado no espaço reduzido para obtenção de suas coordenadas geográficas.

O último passo é fazer o *matching* entre o fingerprint alvo e os fingerprints de referência usando a distância Euclidiana como função de similaridade. Sendo f_i o i -ésimo fingerprint do espaço de busca reduzido, m o fingerprint alvo e d_i como a distância

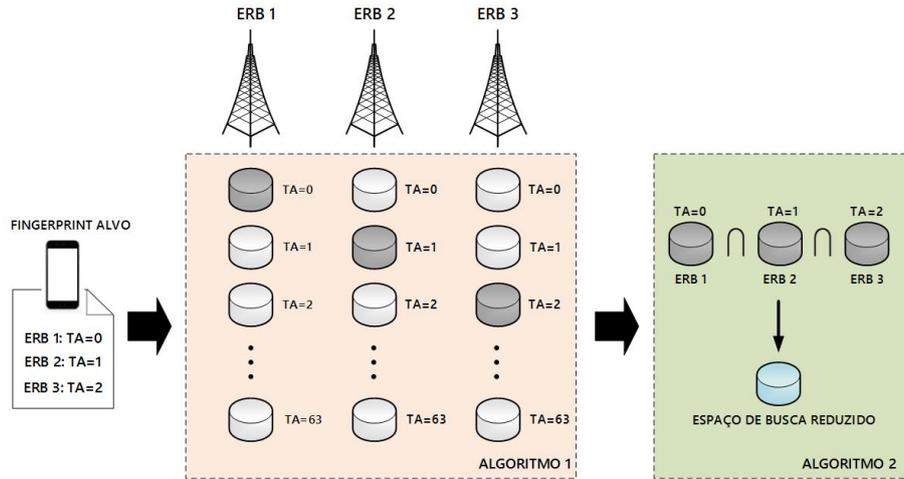


Figura 5. Detalhamento da técnica de localização proposta para um cenário hipotético com 3 ERBs: Algoritmo 1 (etapa *off-line*) e Algoritmo 2 (etapa *on-line*).

Euclidiana entre f_i e m (relativa às N ERBs), que pode ser expressada por

$$d_i = \sqrt{(f_i^{(1)} - m^{(1)})^2 + (f_i^{(2)} - m^{(2)})^2 + \dots + (f_i^{(N)} - m^{(N)})^2}, \quad (1)$$

, a posição estimada será a que obtiver o menor valor da distância Euclidiana dentre todos os *fingerprints* de referência contidos no espaço de busca reduzido.

4. Resultados

Para avaliar a complexidade da técnica proposta neste trabalho, foi realizado um experimento com dados coletados em campo com a ajuda de uma operadora de telecomunicações da região. Os dados foram aferidos de uma rede GSM com portadora na frequência de 1,8 GHz. A base de amostras utilizada possui 2.956 pontos para um trecho urbano de Recife-PE com aproximadamente 5 km² de área [Timoteo et al. 2016]. A área é delimitada pelo quadrilátero definido pelas seguintes coordenadas geográficas (latitude e longitude): [-8.059338, -34.885067], [-8.059338, -34.90683], [-8.07757599, -34.885067] e [-8.07757599, -34.90683]. A base de dados foi construída por meio da técnica de *wardriving*, na qual os valores de RSSI e TA são aferidos para cada uma das ERBs consideradas na área de interesse, por meio de um *scanner* de RF. Adicionalmente, são obtidas, por meio de GPS, as coordenadas geográficas dos pontos de medição, assim como das ERBs. A Figura 6 mostra o mapa do trecho urbano da cidade de Recife-PE escolhido para a coleta dos dados, cujos pontos estão indicados pela trajetória na cor azul. Adicionalmente, a localização das seis ERBs presentes na área de interesse estão identificadas na Figura 6.

A Tabela 1 ilustra um subconjunto de amostras extraídas da base de dados gerada na técnica de *wardriving*. As amostras coletadas são formadas por 14 dimensões (colunas). As duas primeiras, denotadas por **lat** e **lon**, são as longitudes e latitudes dos pontos aferidos. Logo após, são dispostas seis medições de RSSI (**RSSI1** a **RSSI6**), expressas em *dBm*, referentes a cada ERB da região. Por fim, as seis últimas colunas (**TA1** a **TA6**) representam os valores de TA, um para cada ERB.

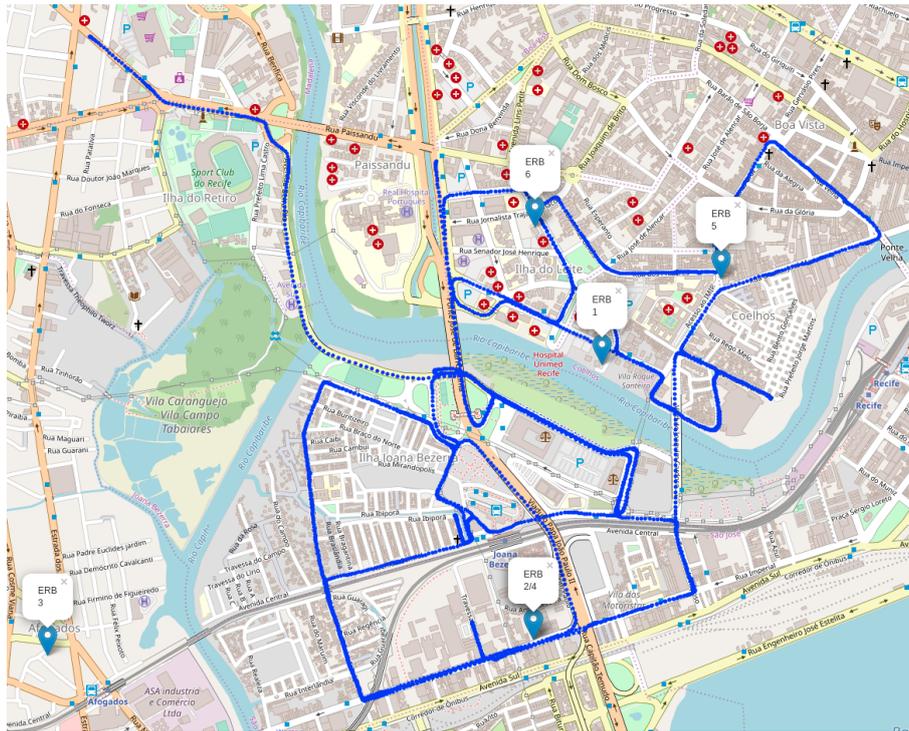


Figura 6. Localização das ERBs e trajetória dos dados coletados (em cor azul) no mapa da área de interesse na região da Ilha do Leite, Recife-PE.

Na fase de treinamento, as coordenadas geográficas coletadas foram normalizadas e utilizadas para treinar um conjunto de seis regressores k -NN [Noi and Kappas 2018]. Para cada um dos seis regressores, as coordenadas geográficas foram utilizadas como *features* e os valores de RSSI em relação a cada uma das seis ERBs foram os valores alvo. A escolha do regressor k -NN se deve ao fato de ser uma técnica de baixa complexidade [Bittencourt et al. 2018].

A Figura 7 ilustra o mapa de calor referente aos níveis de potência do sinal (valores de RSSI) coletados em relação à ERB 6 (ponto em cor azul). A escala situada à direita da figura indica a variação do nível do sinal, na qual a cor amarela representa o nível mais fraco e a cor vermelha, o nível mais forte. Podemos observar que pontos situados mais próximos à ERB possuem níveis de sinal mais intensos, enquanto pontos mais distantes apresentam níveis mais fracos. Também é possível notar o direcionamento da antena da ERB 6, uma vez que os pontos vermelhos estão localizados em direção à parte inferior do

Tabela 1. Exemplos de amostras da base de dados construída por meio da técnica de *wardriving*, contendo as coordenadas geográficas, assim como os valores de RSSI e TA para cada ERB.

lat	lon	RSSI1	RSSI2	RSSI3	RSSI4	RSSI5	RSSI6	TA1	TA2	TA3	TA4	TA5	TA6
-8.077576	-34.899261	123.790	88.9566	122.256	126.89	132.790	138.056	2	0	1	0	3	2
-8.077570	-34.899296	123.065	104.540	120.965	137.140	135.140	139.015	2	0	1	0	3	2
-8.077567	-34.899223	122.756	88.723	119.223	124.256	135.290	132.923	2	0	1	0	3	2
-8.077550	-34.899174	119.123	82.057	121.656	126.856	133.090	133.923	2	0	1	0	3	2
-8.077547	-34.899319	127.590	110.890	130.923	138.290	135.823	135.123	2	0	1	0	3	2

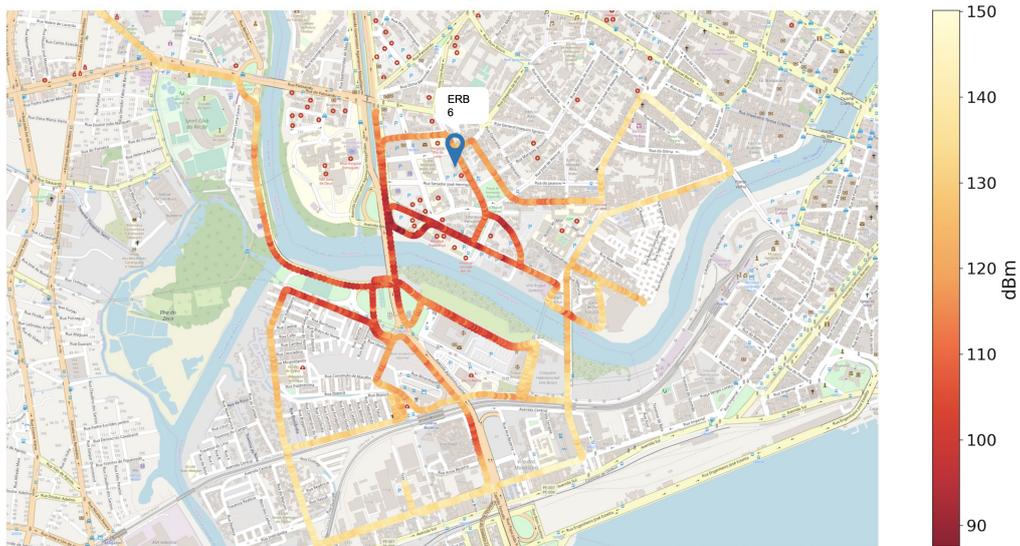


Figura 7. Potência do sinal recebido na EM em relação a ERB 6 no mapa da área de interesse na região da Ilha do Leite, Recife-PE.

mapa, indicando que os valores de RSSI são mais elevados neste trecho.

Face ao exposto, foram implementadas duas técnicas de localização baseadas em *fingerprind* de RF, com modificação apenas da técnica de redução de espaço de busca. No primeiro caso, foi realizada a filtragem por TA, enquanto no segundo, foi implementada a proposta de redução apresentada neste trabalho. Para ambas as técnicas, calculamos o erro médio da estimativa de posição da EM, obtido com base no método de validação cruzada *k-fold* [Kuhn and Johnson 2013]. De maneira geral, este método consiste na divisão do conjunto total de dados em k porções (*folds*), das quais $(k - 1)$ delas compõem a base de treinamento e uma forma a base de teste. Feita essa alocação de *folds*, o regressor é treinado e aplicado à base de teste para o cálculo do erro. Ao final deste processo, uma iteração é finalizada. Em seguida, uma nova iteração é realizada e um novo erro é calculado. Este procedimento é repetido até que cada um dos k *folds* sejam utilizados como base de teste. Ao final, o erro médio é obtido a partir dos erros obtidos em cada uma das k iterações. Neste trabalho, o método de validação cruzada foi empregado para $K = 10$ iterações.

O erro médio da estimativa de posição da EM é definido por

$$\bar{e} = \frac{1}{k} \sum_{j=1}^k e_j \quad (2)$$

em que e_j é o erro médio calculado para o j -ésimo *fold*, que é dado por

$$e_j = \frac{1}{T} \sum_{t=1}^T |P^{(t)} - P_s^{(t)}| \quad (3)$$

sendo $P^{(t)}$ e $P_s^{(t)}$, respectivamente, as posições real e predita da t -ésima instância do conjunto de teste, enquanto T representa o número total de instâncias do *fold*. A diferença

Tabela 2. Valores de erro médio de estimativa de posição da EM e de desvio padrão (ambos em metros) para as duas técnicas de localização implementadas, considerando diferentes resoluções do *grid* de localização.

Res. grid (m)	Filtragem por TA		Técnica proposta	
	Erro médio \bar{e} (m)	Desv. pad. σ_e (m)	Erro médio \bar{e} (m)	Desv. pad. σ_e (m)
5	60,92	60,37	60,92	60,37
10	63,71	61,03	63,71	61,03
15	65,95	62,30	65,95	62,30
20	67,54	61,58	67,54	61,58

entre $P^{(t)}$ e $P_s^{(t)}$ em metros é obtida por meio da distância geodésica entre as posições predita e real da EM. É importante ressaltar que $P^{(t)}$ é previamente conhecida por se tratar de uma instância coletada no *wardriving*. A Tabela 2 apresenta o erro médio \bar{e} e seu desvio padrão σ_e , ambos medidos em metros, para as duas técnicas estudadas e diferentes resoluções de *grid* de localização (5, 10, 15 e 20 m). Conforme é possível observar na Tabela 2, os resultados obtidos em cada técnica são similares, mostrando que a mudança na técnica de redução de espaço de busca não interferiu diretamente na acurácia da localização baseada em *fingerprint* de RF.

Considerando que a proposta apresentada tem como foco a diminuição da complexidade de implementação, serão consideradas duas métricas de avaliação para esta finalidade. A primeira delas é o tempo de treinamento, que consiste no tempo que o algoritmo necessita para construir o mapa de rádio e será denotado por t_T . Foi assumido que o tempo de treinamento foi aferido sem considerar os tempos de treinamento dos regressores k -NN e de construção do *grid*. Tal decisão se justifica pelo fato de ambas as técnicas de localização aqui abordadas utilizarem o mesmo *grid* de localização e o mesmo número de regressores (parâmetro obtido pelo método *grid search* [Huazhou et al. 2018], onde o número de vizinhos é escolhido dinamicamente de acordo com a base de dados). Sendo assim, assumiremos que o tempo de treinamento é composto pelo tempo de associação dos *fingerprints* de referência com o *grid* regular e pelo tempo de construção dos subconjuntos baseados na quantidade de ERBs e nos valores de TA (*vide* Algoritmo 1). Por fim, a segunda métrica é o tempo de predição, definido como o tempo necessário para o algoritmo prever a posição de todos os pontos do conjunto de teste e será denotado por t_p .

A Tabela 3 indica o valor médio \bar{t}_T e o desvio padrão σ_T dos tempos de treinamento das técnicas de *fingerprint*, considerando filtragem por TA e a técnica de redução de espaço de busca proposta neste trabalho, para diferentes resoluções de *grid* de localização. É possível observar que a nossa proposta demanda um tempo maior de treinamento, o que se justifica pela adição da etapa de construção dos subconjuntos referentes às ERBs e aos valores de TA. Para cada resolução considerada, temos a variação percentual $\Delta(\%)$ do tempo de treinamento. Valores positivos de Δ indicam aumento percentual, enquanto valores negativos representam redução. Em média, é observado um acréscimo de 60,18% no tempo de treinamento para as resoluções de *grid* indicadas. De fato, quando a quantidade de valores de TA aumenta (maior extensão da área de interesse), a utilização da técnica proposta tende a elevar o tempo de treinamento da localização. No experimento realizado

Tabela 3. Tempos de treinamento (valor médio \bar{t}_T e desvio-padrão σ_T , medidos em segundos) da localização baseada em *fingerprint* de RF, considerando filtragem por TA e a técnica proposta, para diferentes resoluções de *grid* (Δ é o aumento percentual em \bar{t}_T da técnica proposta em relação ao filtro por TA).

Res.grid (<i>m</i>)	Filtragem por TA		Técnica proposta		$\Delta(\%)$
	Tempo tr. \bar{t}_T (s)	Desv. pad. σ_T (s)	Tempo tr. \bar{t}_T (s)	Desv. pad. σ_T (s)	
5	29,26	0,42	46,24	0,30	58,03
10	7,71	0,10	12,31	0,11	59,66
15	3,84	0,04	5,89	0,01	53,38
20	1,78	0,02	3,02	0,004	69,66

neste trabalho, o valor máximo de TA foi fixado em 10, em virtude do tamanho da área de interesse.

A segunda métrica avaliada foi o tempo de predição da etapa *on-line* da localização. A Tabela 4 mostra o valor médio \bar{t}_p e o desvio padrão σ_p dos tempos de predição das técnicas de *fingerprint*, considerando filtragem por TA e a técnica de redução de espaço de busca proposta neste trabalho, para diferentes resoluções de *grid* de localização. Podemos notar que, independente da resolução considerada, o tempo médio de predição da proposta apresentada é menor do que o tempo médio de predição da filtragem por TA. Para as resoluções consideradas, a técnica proposta promoveu uma redução média de 72,63% em \bar{t}_p quando comparado à filtragem por TA. Uma vez que o esforço de criação de subconjuntos foi retirado da etapa de predição (*on-line*) e transferido para a etapa de treinamento (*off-line*), a diminuição apresentada no tempo de predição é justificada. Esses resultados são expressivos, uma vez que, com a técnica de redução proposta, é possível utilizar resoluções de *grid* menores (o que implica em maior acurácia da localização) sem onerar o tempo de predição. Por exemplo, considerando a filtragem por TA, o tempo de predição para obter as posições do conjunto de teste em um *grid* regular com resolução de 20 *m* é equivalente ao tempo de predição em um *grid* de 10 *m* com a técnica proposta.

Por fim, apesar do tempo de treinamento ser diretamente proporcional ao tamanho da área de interesse, o aumento do tempo desta etapa pode se tornar irrelevante, dado que o sistema apenas necessita ser retreinado quando houver uma degradação no modelo de predição. No caso de sistemas de localização baseados em *fingerprint* de RF, este processo se faz necessário quando há mudanças de ERBs ou adição de características ao sistema, como mudança da topografia da região, mudanças climáticas e modificações no gerenciamento da rede [Viel et al. 2017].

5. Conclusões e Trabalhos Futuros

Neste trabalho, foi proposta uma técnica de redução de espaço de busca que se destina a reduzir o tempo de predição da posição da EM em um processo de localização baseado em *fingerprint* de RF. Foi realizada uma comparação da técnica de redução de espaço de busca que utiliza filtragem por TA com a nossa proposta. A filtragem por TA se baseia em gerar, na etapa *on-line*, os subconjuntos de pontos com base nos valores de TA do usuário móvel. De forma diferente, a nossa proposta divide o esforço de geração desses subconjuntos nas etapas *off-line* e *on-line*. Os resultados obtidos comprovaram a hipótese de que a

Tabela 4. Tempos de predição (valor médio \bar{t}_p e desvio-padrão σ_p , medidos em s) da localização baseada em *fingerprint* de RF considerando a filtragem por TA e a técnica proposta para diferentes resoluções de *grid* (Δ é a redução percentual em \bar{t}_p da técnica proposta em relação ao filtro por TA).

Res. grid (m)	Filtragem por TA		Técnica proposta		$\Delta(\%)$
	Tempo pr. \bar{t}_p (s)	Desv. pad. σ_p (s)	Tempo pr. \bar{t}_p (s)	Desv. pad. σ_p (s)	
5	244,62	0,53	68,96	0,20	- 71,80
10	62,67	0,15	16,77	0,07	- 73,24
15	28,82	0,01	7,94	0,01	- 72,44
20	16,74	0,01	4,51	0,07	- 73,05

modificação proposta neste trabalho promove a diminuição do tempo de predição na etapa *on-line* ao preço de um aumento no tempo de treinamento. Para as resoluções de *grid* consideradas, o tempo médio de predição sofreu uma diminuição de aproximadamente 72,63 % quando comparadas as duas técnicas de redução de espaço de busca. Como perspectivas de trabalhos futuros, é possível explorar a técnica proposta em redes celulares mais recentes, como, por exemplo, as redes LTE e 5G, que possuem parâmetros de atraso de onda menores, podendo assim permitir a obtenção de resultados melhores, tanto em acurácia como em custo computacional.

Referências

- Beekhuizen, J., Kromhout, H., Huss, A., and Vermeulen, R. (2013). Performance of GPS-devices for environmental exposure assessment. *Journal of Exposure Science and Environmental Epidemiology*, 23(5):498.
- Bittencourt, G. P., Urbano, A. A., and Cunha, D. C. (2018). A proposal of an RF fingerprint-based outdoor localization technique using irregular grid maps. In *Proc. of the IEEE Wireless Communications and Networking Conference (WCNC 2018)*, pages 1–6.
- Campos, R. and Lovisolo, L. (2019). Genetic algorithm-based cellular network optimisation considering positioning applications. *IET Communications*, 13(7):879–891.
- Campos, R. S. and Lovisolo, L. (2009). A fast database correlation algorithm for localization of wireless network mobile nodes using coverage prediction and round trip delay. In *Proc. of the IEEE Vehicular Technology Conference (VTC 2009)*, pages 1–5.
- Chen, K., Tan, G., Cao, J., Lu, M., and Fan, X. (2019). Modeling and improving the energy performance of gps receivers for location services. *IEEE Sensors Journal*, 20(8):4512–4523.
- Deville, P., Lenard, C., Martin, S., and Gilbert, M. (2014). Dynamic population mapping using mobile phone data. *Proc. of the National Academy of Sciences*, 111(45):15888–15893.
- Hammad, A. and Faith, P. (2017). Location based authentication. US Patent 9,721,250.
- Huazhou, Chen, Z., Liu, K., Cai, L., Xu, A., and Chen (2018). Grid search parametric optimization for ft-nir quantitative analysis of solid soluble content in strawberry samples. *Vibrational Spectroscopy*, 94(1):7–15.

- Kose, M., Tascioglu, S., and Telatar, Z. (2019). RF fingerprinting of IoT devices based on transient energy spectrum. *IEEE Access*, 7:18715–18726.
- Kuhn, M. and Johnson, K. (2013). *Applied Predictive Modeling*. Springer, 1nd edition.
- Mondal, R. U., Ristaniemi, T., and Turkka, J. (2015). Genetic algorithm optimized grid-based RF fingerprint positioning in heterogeneous small cell networks. In *Proc. of the IEEE Int. Conf. on Location and GNSS (ICL-GNSS 2015)*, pages 1–7.
- Noi, P. T. and Kappas, M. (2018). Comparison of random forest, k-nearest neighbor, and support vector machine classifiers for land cover classification using sentinel-2 imagery. *Sensors*, 18(1):1–20.
- Nouichi, D., Abdelsalam, M., Nasir, Q., and Abbas, S. (2019). IoT devices security using RF fingerprinting. In *Proc. of the Int. Conf. on Advances in Science and Engineering Technology (ASET 2019)*, pages 1–7.
- Nyhan, M. M., Kloog, I., Britter, R., Ratti, C., and Koutrakis, P. (2019). Quantifying population exposure to air pollution using individual mobility patterns inferred from mobile phone data. *Journal of Exposure Science & Environmental Epidemiology*, 29(2):238.
- Peral-Rosado, J. A., Raulefs, R., López-Salcedo, J. A., and Seco-Granados, G. (2018). Survey of cellular mobile radio localization methods: From 1G to 5G. *IEEE Commun. Surveys & Tutorials*, 20(2, Second Quarter 2018):1124–1148.
- Salomon, A. and Mahaffey, K. P. (2019). Mobile communications device payment method utilizing location information. US Patent App. 10/181,118.
- Saunders, S. and Aragón-Zavala, A. (2007). *Antennas and Propagation for Wireless Communication Systems*. John Wiley & Sons, 2nd edition.
- Timoteo, R. D., Silva, L. N., Cunha, D. C., and Cavalcanti, G. D. (2016). An approach using support vector regression for mobile location in cellular networks. *Elsevier Computer Networks*, 95:51–61.
- Trogh, J., Plets, D., Surewaard, E., Spiessens, M., Versishele, M., Martens, L., and Joseph, W. (2019). Outdoor location tracking of mobile devices in cellular networks. *EURASIP Journal on Wireless Communications and Networking 2019*, (115):1–18.
- Viel, A., Gallo, P., Montanari, A., Gubiani, D., Dalla Torre, A., Pittino, F., and Marshall, C. (2017). Dealing with network changes in cellular fingerprint positioning systems. In *Proc. of the IEEE Int. Conf. on Localization and GNSS (ICL-GNSS 2017)*, pages 1–6.
- Vo, Q. D. and De, P. (2016). A survey of fingerprint-based outdoor localization. *IEEE Communications Surveys & Tutorials*, 18(1):491–506.
- Zekavat, R. and Buehrer, R. M. (2019). *Handbook of Position Location: Theory, Practice and Advances*. John Wiley & Sons, 2nd edition.
- Zhang, H., Zhang, Z., Zhang, S., Xu, S., and Cao, S. (2019). Fingerprint-based localization using commercial lte signals: a field-trial study. In *Proc. of the IEEE 90th Vehicular Technology Conf. (VTC-2019 Fall)*, pages 1–5.