



Universidade Federal de Pernambuco
Centro de Informática

Bacharelado em Engenharia da Computação

**Avaliação de um modelo de *grid* irregular
para localização *outdoor* utilizando
simulação**

Felipe Sousa de Andrade

Trabalho de Graduação

Recife
4 de julho de 2018

Universidade Federal de Pernambuco
Centro de Informática

Felipe Sousa de Andrade

**Avaliação de um modelo de *grid* irregular para localização
outdoor utilizando simulação**

Trabalho apresentado ao Programa de Bacharelado em Engenharia da Computação do Centro de Informática da Universidade Federal de Pernambuco como requisito parcial para obtenção do grau de Bacharel em Engenharia da Computação.

Orientador: *Daniel Carvalho da Cunha*

Recife
4 de julho de 2018

*Dedico este trabalho aos meus pais, que sempre colocaram
a educação dos filhos em primeiro lugar.*

Agradecimentos

Agradeço de modo especial aos meus pais por sempre terem dado suporte aos meus estudos e me motivado ao longo de toda a minha formação. Agradeço à minha irmã, que sempre se disponibilizou a tirar minhas dúvidas e me encorajar nos estudos. Agradeço a todos os meus amigos que me apoiaram ao longo da minha graduação na realização de projetos em equipe ou nos estudos para as provas. Agradeço a todos os professores que fizeram parte da minha graduação, por todos os valiosos ensinamentos que me foram passados.

Agradeço ao meu orientador, professor Daniel Cunha, e a Gustavo Bittencourt pelas orientações, ideias e explicações importantes para a realização deste trabalho.

Resumo

Recentemente, foi proposta uma técnica de localização *outdoor* baseada em *fingerprint* de RF (Rádio Frequência), que utiliza um modelo de *grid* irregular capaz de se adequar à malha viária de ambientes urbanos. O objetivo desse novo modelo de *grid* é reduzir o número de células geradas em cada *grid*, em comparação com o modelo de *grid* regular, melhorando significativamente o desempenho computacional e a precisão do algoritmo de localização. Este trabalho consiste em uma validação dessa técnica com outras bases de dados de medições de sinais de RF, geradas utilizando o módulo LENA LTE no simulador ns-3, um simulador de redes de computadores de eventos discretos. Os cenários de interesse deste estudo foram: bases de dados nas quais uma fração das medições corresponde a pontos que estão fora das vias urbanas e bases de medições realizadas em regiões com diferentes distribuições espaciais e morfologias de edificações.

Os resultados encontrados indicam que os *grids* irregulares são mais precisos que os regulares em cenários nos quais a porcentagem de medições em vias urbanas corresponde a 80% do total ou mais. Essa melhoria na precisão já não é mais observada quando essa porcentagem corresponde a 70% ou menos. Nessa avaliação, foram comparados *grids* regulares e irregulares com aproximadamente o mesmo número de células. Além disso, com respeito à distribuição de edifícios na região de cobertura, os resultados indicam que a densidade de ocupação urbana não influencia na precisão dos *grids* irregulares em comparação com os regulares.

Palavras-chave: localização *outdoor*, *RF fingerprinting*, simulação de redes celulares, *grid* irregular, modelo de propagação

Abstract

Recently, an RF (Radio Frequency) fingerprint-based outdoor location technique using an irregular grid model which is capable of fitting to the urban road network has been proposed. This new grid model aims to reduce the number of cells generated in each grid in comparison to the regular grid model, expressively improving the computational cost and the precision of the location algorithm. This study consists of validating this technique by using other RF measurement databases, generated by simulation using LENA module in ns-3, a discrete event network simulator. This study focus on the following scenarios, the first in which a portion of the measurements corresponds to points off the streets, and the second with measurements from regions with different building topology.

The results indicated that the irregular grid model performs better than the regular one in scenarios where at least 80% of the measurements correspond to points on the streets. The irregular grid model does not provide an improvement on the precision of the regular model in scenarios in which that percentage is no more than 70%. This analysis compared regular and irregular grids with nearly the same number of cells. Additionally, with respect to the distribution of buildings in the coverage area, the results indicated that the degree of urbanization does not affect the precision of the irregular grid model in comparison to the regular one.

Keywords: outdoor location, RF fingerprinting, mobile network simulation, irregular grid, propagation loss

Sumário

1	Introdução	1
1.1	Motivação	1
1.2	Objetivos	2
1.3	Contribuições do Trabalho	2
1.4	Estrutura do Documento	3
2	Conceitos Básicos	4
2.1	Técnicas de Localização Baseadas em <i>Fingerprint</i> de RF	4
2.2	Modelos de <i>Grids</i>	6
2.3	Modelagem de Sinais de RF Através de Simulação	8
2.3.1	Simuladores	8
2.3.2	Trabalhos relacionados	10
3	Método de Avaliação por Simulação	11
3.1	Definição dos Cenários de Simulação	11
3.1.1	Visualização dos Cenários Simulados	11
3.1.2	Parâmetros de Simulação	11
3.1.3	Tipos de Ambientes Simulados	17
3.2	Geração das Bases de Dados	17
3.3	Avaliação do Algoritmo de Localização	20
4	Experimentos e Análise dos Resultados	22
4.1	Influência de Medições Fora das Vias Urbanas	24
4.2	Influência da Densidade de Ocupação Urbana	29
4.2.1	Ambiente Suburbano	29
4.2.2	Ambiente Rural	31
5	Conclusão e Trabalhos Futuros	34
5.1	Contribuições	34
5.2	Trabalhos Futuros	34

Lista de Figuras

2.1	Diagrama de blocos de técnicas de localização baseadas em <i>fingerprint</i> de RF	5
2.2	Formato de um <i>grid</i> regular com resolução r	7
2.3	Topologia de rede do módulo LENA	9
3.1	Representação gráfica de um cenário do <i>Dual Stripe</i> original	12
3.2	Posicionamento dos edifícios num cenário urbano	13
3.3	Posicionamento dos edifícios, usuários e estações base ao longo de um cenário urbano	15
3.4	Posicionamento dos edifícios, usuários e estações base ao longo dos cenários suburbano e rural	18
3.5	Células dos <i>grids</i> regular e irregular com resolução de 15 m	21
4.1	Erro médio em um cenário urbano com 100% de medições em vias urbanas	25
4.2	Erro médio em um cenário urbano com 90% de medições em vias urbanas	26
4.3	Erro médio em um cenário urbano com 80% de medições em vias urbanas	27
4.4	Erro médio em um cenário urbano com 70% de medições em vias urbanas	27
4.5	Erro médio em um cenário urbano com 60% de medições em vias urbanas	28
4.6	Erro médio em um cenário suburbano com 100% de medições em vias urbanas	30
4.7	Erro médio em um cenário suburbano com 80% de medições em vias urbanas	30
4.8	Erro médio em um cenário suburbano com 60% de medições em vias urbanas	31
4.9	Erro médio em um cenário rural com 100% de medições em vias urbanas	32
4.10	Erro médio em um cenário rural com 80% de medições em vias urbanas	32
4.11	Erro médio em um cenário rural com 60% de medições em vias urbanas	33

Lista de Tabelas

2.1	Formato da base de dados original	4
3.1	Parâmetros de configuração das estações base	14
3.2	Desvio padrão da distribuição log-normal que calcula o sombreamento do sinal	16
3.3	Parâmetros de simulação para cada tipo de ambiente	17
3.4	Amostras das bases de dados geradas através de simulação	19
4.1	Número de células e taxa de redução em cada um dos <i>grids</i> gerados	23

CAPÍTULO 1

Introdução

1.1 Motivação

Técnicas de localização são amplamente utilizadas por diversas aplicações desenvolvidas para dispositivos móveis, a fim de prover algum serviço para seus usuários. Como exemplos, podemos listar serviços de navegação, transporte de passageiros e entrega de conteúdo direcionado a determinadas regiões, como previsão do tempo, divulgação de eventos e propaganda. Embora o GPS (*Global Positioning System*) seja um dos sistemas mais utilizados por essas aplicações para obter localizações *outdoor* dos seus usuários, esse sistema apresenta alto consumo de bateria, além de perda de precisão em ambientes urbanos com alta densidade de edificações [1, 2, 3]. A perda de precisão nesses ambientes se dá, principalmente, pela inexistência da linha de visada, causada por obstáculos entre o emissor e o receptor do sinal [3]. Essas duas características do GPS representam desvantagens críticas desse sistema, uma vez que grande parte dos usuários dessas aplicações se encontra em ambientes urbanos e deseja que seus dispositivos móveis tenham a maior autonomia de bateria possível. O crescente número de sensores embutidos nos aparelhos celulares modernos tem favorecido cada vez mais a localização desses dispositivos móveis com alta precisão e com baixo consumo de energia, especialmente através de técnicas baseadas em *fingerprinting* [2].

Uma dessas técnicas foi proposta recentemente em [1] e seu objetivo é a localização *outdoor* baseada em *fingerprint* de RF (Rádio Frequência), utilizando um modelo de *grid* irregular capaz de se adequar à malha viária de ambientes urbanos. Nessa técnica, a área de cobertura é dividida em células que remontam as vias urbanas da região de interesse. Os valores dos sinais de RF são estimados para cada uma dessas células, gerando uma espécie de assinatura para cada célula. Posteriormente, as posições dos usuários são estimadas com base na similaridade dos seus sinais de RF com os das células do *grid*. Esse modelo de *grid* parte da hipótese de que, em ambientes urbanos, os usuários de dispositivos móveis tem grande probabilidade de se encontrar em ruas ou avenidas, ou em calçadas e passarelas próximas a essas vias. O uso desse novo modelo de *grid* resultou em uma redução significativa do número de células do *grid*, em comparação ao *grid* regular, o que melhorou o desempenho computacional e a precisão do algoritmo de localização. Esses resultados foram obtidos utilizando 2756 medições reais de sinais de telefonia móvel GSM (*Global System for Mobile communications*) na faixa de 1,8 GHz, obtidas através de um *drive-test* numa região urbana de Recife-PE.

Essa técnica é muito importante para a área de localização *outdoor*, porque proporciona uma redução significativa no número de células geradas em cada *grid*. Essa redução permite trabalhar com *grids* de alta resolução, que levam a uma melhora significativa na precisão da estimativa de localização. Por essa razão, este trabalho consiste em uma validação dessa técnica

de localização sobre diferentes cenários. Esta avaliação permitiu estimar em quais cenários o modelo de *grid* irregular proposto em [1] apresenta melhor precisão na localização, quando comparado com o modelo de *grid* regular. Esses testes foram realizados sobre bases de dados de medições de sinais de RF geradas através de simulação.

1.2 Objetivos

Embora a técnica apresentada em [1], mencionada anteriormente, tenha sido validada para diferentes resoluções do *grid* (tamanho do lado de cada célula quadrangular do *grid*), ela foi testada apenas para uma base de dados, na qual todas as medições foram realizadas em vias urbanas. Por essa razão, o objetivo deste trabalho é validar essa técnica com outras bases de dados de medições de sinais de RF, geradas utilizando o módulo LENA LTE no simulador ns-3, a exemplo do trabalho feito por Han, Ma e Zhang [4], que usaram a mesma ferramenta para validar o modelo de *grid* irregular definido pelos mesmos.

Com isso, foi possível avaliar os resultados do algoritmo quando uma fração dos dados de treino e dos dados de teste corresponde a pontos que estão fora das vias urbanas. Utilizando a biblioteca LENA no simulador ns-3, foram testadas diferentes distribuições de medições em vias urbanas e fora dessas vias, a fim de mensurar em quais cenários o uso do *grid* irregular proporciona melhor precisão na localização, quando comparado com *grids* regulares com aproximadamente o mesmo número de células.

Analogamente, utilizando simulação, foi estimado para quais tipos de regiões, em termos de densidade de ocupação urbana, o *grid* irregular apresenta melhor precisão na localização. Para tanto, foram simulados ambientes urbanos, suburbanos e rurais. Para cada um desses tipos de ambientes, foi definida uma densidade média de edificações e um número médio de andares nesses edifícios, como descrito no Capítulo 3.

1.3 Contribuições do Trabalho

Como resultado da avaliação do impacto das medições fora das vias urbanas sobre a técnica definida em [1], concluiu-se que o modelo de *grid* irregular é mais preciso que o de *grid* regular em cenários com porcentagem de medições em vias urbanas maior ou igual a 80% do total. Em cenários nos quais essa porcentagem corresponde a 70% ou menos, essa melhoria na precisão já não é mais observada. Nessa análise, foram comparados *grids* regulares e irregulares com aproximadamente o mesmo número de células.

Também foi avaliado o impacto do nível de urbanização da região de cobertura sobre a precisão da técnica de localização avaliada. Os resultados encontrados indicam que a densidade de ocupação urbana não influencia no desempenho dos *grids* irregulares em comparação com os *grids* regulares. Adicionalmente, os resultados indicam que a técnica de localização baseada em *fingerprint* de RF tem um desempenho pior em ambientes suburbanos do que nos ambientes urbanos, independente do modelo de *grid* utilizado.

1.4 Estrutura do Documento

O próximo capítulo aborda conceitos básicos importantes para o entendimento deste trabalho. Ele apresenta técnicas de localização baseadas em *fingerprint* de RF, bem como alguns dos modelos de *grid* utilizados por essas técnicas. Em seguida, é apresentada uma visão geral sobre simulação de sinais de RF e sobre trabalhos que, assim como esse, utilizaram simulação para validar técnicas de localização.

O Capítulo 3 descreve o método de avaliação por simulação utilizado neste trabalho. Essa descrição inclui os parâmetros usados na simulação e as características de cada tipo de ambiente simulado. Além disso, o processo de geração de bases de dados de medições a partir das saídas do simulador é apresentado, bem como o processo de avaliação da técnica de localização em questão para cada uma dessas bases geradas.

Em seguida, no Capítulo 4, são descritos os experimentos realizados neste estudo, bem como os resultados de cada um deles. Esses resultados são divididos da seguinte forma: número de células geradas em cada *grid*, influência de medições realizadas fora das vias urbanas sobre o desempenho do modelo de *grid* irregular e influência da densidade de ocupação urbana sobre o desempenho do mesmo modelo. Por fim, o Capítulo 5 conclui este trabalho sintetizando as contribuições trazidas pelo mesmo e descrevendo possíveis trabalhos futuros.

Conceitos Básicos

Este capítulo apresenta conceitos básicos que serão abordados ao longo do restante deste trabalho, bem como trabalhos relacionados a esses conceitos. A Seção 2.1 explica técnicas de localização baseadas em *fingerprint* de RF, incluindo a técnica definida em [1], que é avaliada neste estudo. A Seção 2.2 descreve modelos de *grids* utilizados por essas técnicas, como o modelo de *grid* irregular, definido em [1]. Por fim, a Seção 2.3 trata de simulação de sinais de RF em sistemas de comunicações móveis, com foco especial no módulo LENA LTE do simulador ns-3.

2.1 Técnicas de Localização Baseadas em *Fingerprint* de RF

Algoritmos de localização baseados em *fingerprint* de RF são utilizados quando se tem como entrada uma base de dados com medições de sinais de RF. Cada instância dessa base de dados corresponde a uma medição realizada por um dispositivo móvel, também chamado de estação móvel. Essas medições geralmente são obtidas através de experimentos realizados em campo, conhecidos como *drive-test* e que percorrem uma determinada rota dentro da região de interesse. No entanto, como mencionado em [1], o fato desses testes não serem automáticos inviabiliza a coleta de dados em toda a região de cobertura. Além disso, algumas áreas dentro dessa região podem não ser abertas ao público.

Cada medição contém, além da posição do usuário, níveis de sinais de RF de cada uma das estações bases, medidas naquele ponto, como RSRP (*Reference Signal Received Power*) e RSRQ (*Reference Signal Received Quality*). O primeiro valor, também conhecido como RSSI (*Received Signal Strength Indication*), corresponde à potência do sinal recebido pela estação móvel e geralmente é medida de *dBm*. O segundo valor é uma medida de qualidade do sinal que também leva em conta interferência do canal e ruído térmico e geralmente é medida em *dB* [5]. A técnica definida em [1], avaliada nesse trabalho, utiliza apenas os níveis de RSRP, assim como as técnicas apresentadas em [6] e [4]. A Tabela 2.1 ilustra o formato da base de dados utilizada em [1].

Tabela 2.1 Formato da base de dados original

lat	lon	RSSI 1	RSSI 2	RSSI 3	RSSI 4	RSSI 5	RSSI 6
-8.077546	-34.899288	-66.825	-35.6	-64.875	-72.575	-79.525	-78.45
-8.077545	-34.899265	-66.2	-33.65	-71.05	-71.75	-71.4	-84.55
-8.077541	-34.899246	-68.225	-32.3	-66.475	-71.15	-79.5	-83

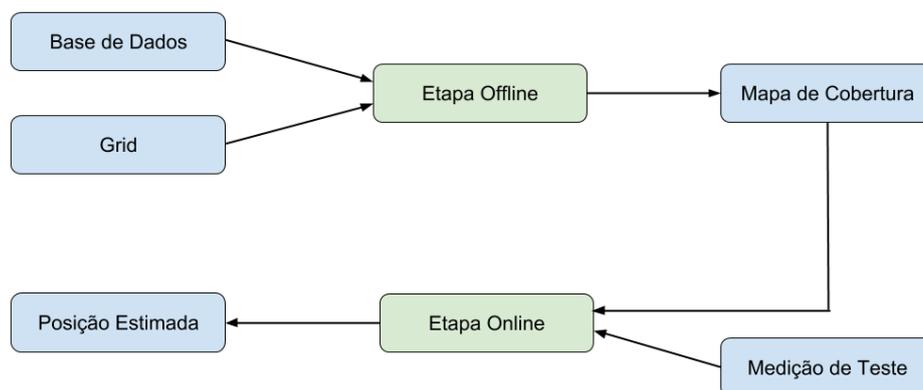


Figura 2.1 Diagrama de blocos de técnicas de localização baseadas em *fingerprint* de RF.

Em todas essas técnicas, a região de interesse é dividida em células que formam um *grid* e o algoritmo de localização consiste em duas etapas: etapa *offline* e etapa *online*. Na etapa *offline* é feita uma previsão dos níveis de sinais de RF em cada uma das células, que resulta num mapa de cobertura. Na etapa *online* é feita a estimativa da posição de uma medição de sinais de RF cuja posição é desconhecida. Nessa estimativa, é retornada a posição da célula do mapa de cobertura cujos níveis de sinais de RF é mais similar aos níveis da medição de teste. Nessas técnicas, é esperado que cada célula do *grid* tenha um vetor único de níveis de sinais, que funciona como uma assinatura ou uma impressão digital daquela célula. A Figura 2.1 apresenta um diagrama de blocos das etapas *offline* e *online* dessas técnicas de localização.

Conforme mencionado em [1], a maioria dos estudos realizados na área de localização baseada em *fingerprint* se concentra na melhoria das técnicas de aprendizagem de máquina utilizadas na etapa *offline* do mapa de cobertura. Há ainda estudos cujo foco são as métricas de similaridade aplicadas na etapa *online*, como é o caso do artigo [6], que avalia métodos que utilizam duas dessas métricas: distância de Mahalanobis e divergência de Kullback-Leibler. No entanto, o grande diferencial da técnica proposta em [1] é o modelo de *grid* usado para construção do mapa de cobertura. Detalhes sobre diferentes modelos de *grid* utilizados em técnicas de *fingerprint* de RF serão apresentados na Seção 2.2.

Na técnica de localização *outdoor* baseada em *fingerprint* de RF proposta em [1], os sinais de RF utilizados são os valores de RSRP de um determinado conjunto de estações base. Na etapa *offline*, é utilizado o regressor W-KNN (*Weighted K Nearest Neighbors*) para prever os valores de potência recebida em cada célula do *grid*. Em linhas gerais, esse regressor vai atribuir a cada célula do *grid* a média ponderada das potências recebidas nas K medições mais próximas a essa célula. Nessa média ponderada, o peso de cada medida corresponde ao inverso da distância euclidiana da mesma à célula em questão. Na etapa *online*, a distância euclidiana é empregada como métrica de similaridade para encontrar a célula cujos valores de potência recebida mais se aproximam dos da medição de teste. Esses algoritmos e métricas foram escolhidos em [1] por apresentarem uma implementação simples, mas que se mostraram eficazes em outros problemas de localização.

Outra característica relevante da técnica utilizada em [1] é o uso do TA (*Timing Advance*),

um parâmetro inteiro com valor positivo que pode ser obtido através do canal de controle do protocolo GSM. Esse parâmetro corresponde a uma estimativa da distância entre o usuário e a estação base e é dada em intervalos de 550 m. Se o usuário se encontra a uma distância de até 550 m da estação base, então $TA = 1$. Se o usuário se encontra a uma distância de 550 m a 1100 m, então $TA = 2$, e assim por diante. Esse parâmetro é utilizado na etapa *online* para filtrar células do mapa de cobertura, a fim de reduzir o espaço de busca. Essa redução tem o objetivo de diminuir o custo computacional da etapa *online*, reduzindo o número de células a serem consideradas no cálculo da métrica de similaridade, mas também ajuda a diminuir o erro médio da estimativa de localização, uma vez que evita que células com valores de TA muito diferentes dos valores de TA da medição em questão sejam escolhidas pela técnica.

2.2 Modelos de Grids

Os modelos de *grids* utilizados para a construção do mapa de cobertura nas técnicas de localização baseadas em *fingerprint* de RF são, genericamente, divididos em duas categorias: regulares e irregulares. Essas duas categorias são apresentadas a seguir, bem como alguns exemplos de *grids* irregulares.

Os *grids* regulares são os mais simples de serem construídos, pois consistem numa divisão da área de cobertura em uma malha quadriculada de células. O tamanho do lado dessas células quadradas é chamado de resolução do *grid* e é um parâmetro que influencia diretamente na complexidade de espaço e tempo do algoritmo, bem como na precisão da estimativa de localização do mesmo. A diminuição da resolução do *grid* tende a diminuir o erro médio da estimativa de localização, porém aumenta o número de células do *grid*, o que aumenta o custo computacional das etapas *offline* e *online* em termos de memória e processamento. A Figura 2.2 apresenta o formato de um *grid* regular com resolução r .

Grids irregulares tem sido propostos na literatura em busca de melhorar a precisão de técnicas de localização baseadas em *fringerprint* de RF sem aumentar significativamente o custo computacional das mesmas ou até mesmo reduzindo esse custo. Esses modelos de *grid* geralmente se baseiam em alguma característica do problema em questão para modificar a distribuição das células ao longo do *grid* e/ou o tamanho das mesmas.

Um desses modelos, proposto por Bittencourt, Urbano e Cunha [1], se ajusta à malha viária da região de interesse, ou área de cobertura. A hipótese desse modelo é que, em áreas urbanas densamente ocupadas, usuários localizados em ambientes *outdoor* estão localizados em ruas ou avenidas, ou em calçadas ou passarelas próximas a essas vias. Desse modo, o processo de construção desses *grids* é dividido em dois passos. O primeiro passo, executado apenas uma vez, consiste na obtenção dos dados da malha rodoviária da região de interesse utilizando a plataforma *OpenStreetMaps* (OSM), seguida de um processamento para extração dos segmentos que definem as vias dessa região. O segundo passo, executado uma vez para cada valor de resolução desejada, consiste em uma interpolação dos segmentos gerados no primeiro passo. Nessa interpolação, é gerada uma lista de células para cada segmento, de modo que a distância entre duas células adjacentes seja, no máximo, igual à resolução desejada. O *grid* irregular resultante consiste na união das células geradas pela interpolação de cada segmento.

A desvantagem do uso dos *grids* irregulares é que a construção dos mesmos é mais com-

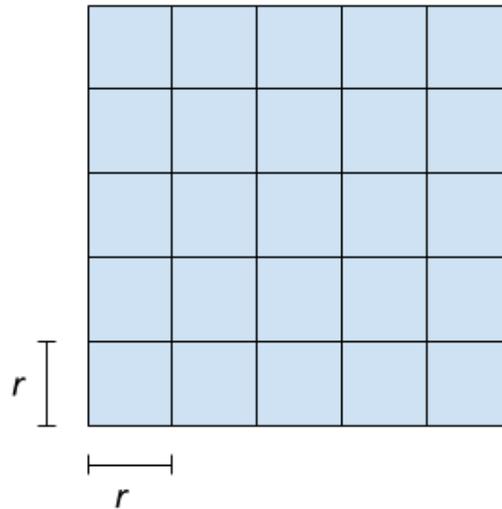


Figura 2.2 Formato de um *grid* regular com resolução r .

plexa e exige informações prévias sobre a malha viária da área de cobertura, como descrito no parágrafo anterior. No entanto, essa complexidade extra se justifica pelos resultados trazidos pelo uso desses *grids*, que foram apresentados em [1]. Esses resultados são: redução no número de células geradas em cada *grid* e melhoria na precisão do algoritmo de localização.

Outro modelo de *grid* irregular, proposto por Kim, Min e Yeo [7], é utilizado em técnicas de localização baseadas em *fingerprint* de valores de potência recebida de redes WiFi. Esse modelo de *grid* busca calibrar a base de dados, agrupando *fingerprints* medidos que são similares. Dessa forma, no *grid* resultante, as células vizinhas tem *fingerprints* estatisticamente distintos, o que torna a predição da localização da estação móvel na etapa *online* mais precisa. Cada célula do *grid* gerado representa um grupo de medições da base de dados e o seu tamanho é determinado pela distância média dessa célula para *fingerprints* de grupos vizinhos.

Há ainda outro modelo de *grid* irregular proposto por Han, Ma e Zhang [4] que divide a região em torno de cada estação base em duas ou três áreas com bordas hexagonais cujos centros correspondem à posição da estação base. Na área mais interna, foi observado que o nível da potência do sinal recebido é mais estável e, por isso, as células do *grid* resultante têm uma área maior. Já na área mais externa, percebe-se maior oscilação da potência recebida e, portanto, são definidas células com área menor, a fim de melhorar a precisão da estimativa de localização. Nos casos em que são definidas três áreas, são utilizadas células com tamanho médio na área intermediária, uma vez que a estabilidade do mesmo sinal de RF tende a ser intermediária.

Um diferencial do modelo de *grid* irregular proposto em [1] quando comparado com os demais modelos citados anteriormente é que o primeiro utiliza estritamente informações da malha viária da região de cobertura para a construção dos *grids* enquanto os outros se baseiam em características dos sinais de RF para a definição dos *grids*. Uma vez que o modelo proposto em [1] não depende de características dos sinais de RF da base de dados, é esperado que o

mesmo seja mais robusto com relação a sinais ruidosos e/ou com distribuições que não sigam um padrão convencional.

2.3 Modelagem de Sinais de RF Através de Simulação

Técnicas de localização baseadas em *fingerprint* de RF geralmente são avaliadas utilizando medições reais obtidas através de *drive-tests*. Um das desvantagens desse método é que a obtenção dessas bases ocorre de forma não automatizada e exige tempo e recursos financeiros. Além disso, é inviável obter medições de RF para todos os pontos de uma certa área de cobertura, uma vez que muitos terrenos são de propriedades privadas. A fim de viabilizar a avaliação dessas técnicas para diferentes cenários, muitos pesquisadores fazem uso de simulação para a obtenção de bases de dados para treinamento e teste.

2.3.1 Simuladores

Conforme descrito em sua documentação [8], o projeto ns-3 é um simulador de redes de computadores de eventos discretos que foi desenvolvido para uso educacional e de pesquisa. Esse projeto foi desenvolvido em C++ e é *open-source*, não sendo mantido oficialmente por nenhuma companhia, e está aberto para contribuições de pesquisadores. Um grande diferencial desse simulador, quando comparado a outros usados para o mesmo fim, é que o ns-3 consiste em um conjunto de bibliotecas que podem ser combinadas entre si, ou com outras bibliotecas externas, para definição do cenários de simulação. Dessa forma, para utilizar essa ferramenta, é necessário o desenvolvimento de código em C++ ou Python, uma vez que a mesma não provê uma interface gráfica para descrição dos cenários a serem simulados, como as interfaces providas por outros simuladores.

Uma das extensões do simulador ns-3 é o módulo LENA, que fornece bibliotecas para a simulação de redes de comunicações móveis. Como descrito em [5], esse módulo provê dois componentes principais: o modelo LTE e o modelo EPC. O modelo LTE é responsável pela simulação das camadas do Protocolo de Comunicação LTE (*Long Term Evolution*), protocolo adotado nas redes de telefonia celular de 4ª Geração (4G). Portanto, esse modelo se concentra na simulação da comunicação via rádio entre as estações base e as estações móveis (denominadas de eNBs e UEs dentro do módulo LENA, respectivamente). O modelo EPC, por sua vez, é responsável pela simulação da comunicação entre unidades centrais da rede de telefonia móvel. Logo, esse modelo se concentra na comunicação das estações base entre si e com as centrais de comutação e controle, e dessas centrais com a internet. A topologia desses dois modelos é apresentada na Figura 2.3. Por conta da natureza do problema tratado pela técnica de localização avaliada neste trabalho, apenas o modelo LTE do módulo LENA foi utilizado para descrição dos cenários de simulação.

O módulo LENA é apropriado para a geração das bases de dados de sinais de RF pois uma das saídas possíveis para a simulação são os dados da camada física, o que inclui o valor da potência recebida do sinal de RF em Watts. Arquivos que incluem dados de outras camadas de comunicação também podem ser gerados por esse módulo, o que inclui dados como o total de *bytes* transmitidos e o tempo médio de transmissão para cada canal lógico. No entanto, esses

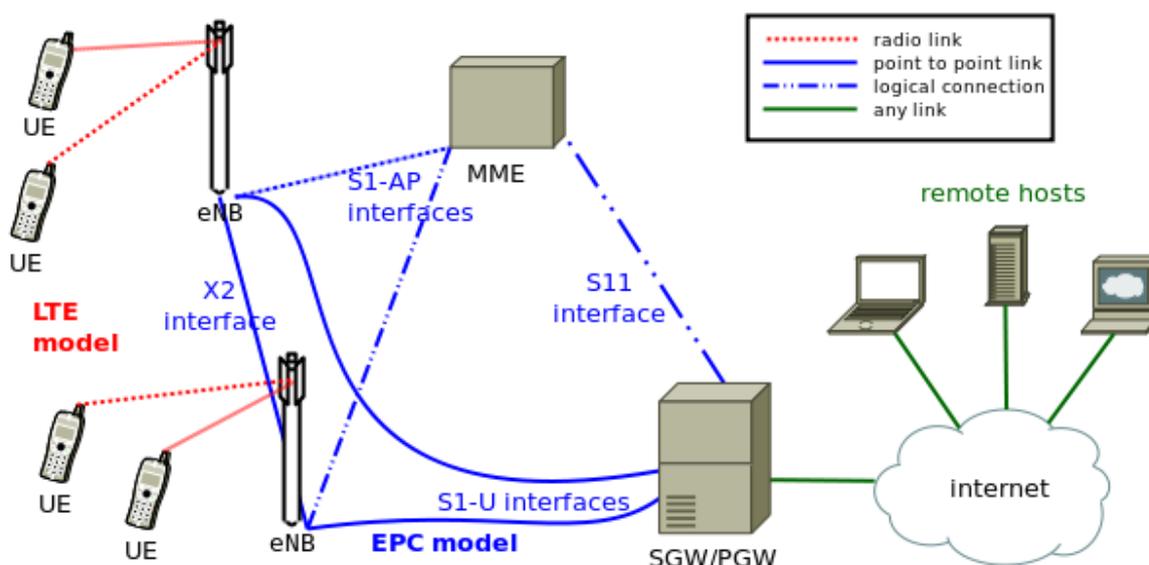


Figura 2.3 Topologia de rede do módulo LENA. Fonte: "<http://networks.cttc.es/mobile-networks/software-tools/lena/>", acessado em 22/06/2018.

outros dados estão fora do escopo desse trabalho e, por isso, não foram utilizados.

Esse módulo permite adicionar edifícios na simulação, através do *Buildings Model*, que deve ser usado juntamente com algum *Building Propagation Model*, para mensurar adequadamente a perda de caminho para nós *indoor*. Com essas ferramentas, é possível definir o tipo do edifício, que pode ser residencial, empresarial ou comercial; o tipo da parede, que pode ser de madeira, concreto com janelas ou concreto sem janelas; o número de quartos nos eixos x e y e o número de andares. Detalhes sobre as características dos edifícios inseridos na simulação e do modelo de propagação escolhido serão apresentados no próximo capítulo.

No módulo LENA, é possível definir a posição e a mobilidade de cada um dos nós da simulação, o que inclui estações base, estações móveis e edificações. A fim de simplificar esse processo, esse módulo também fornece classes que definem a posição e a mobilidade de um conjunto de nós de acordo com algumas topologias mais comuns em redes móveis. Esse é o caso da classe *LteHexGridEnbTopologyHelper*, que posiciona um conjunto de estações base tri-setoriais de acordo com um *grid* hexagonal. Outro exemplo é a classe *GridBuildingAllocator*, que permite o posicionamento de um conjunto de edifícios de acordo com um *grid* retangular bidimensional.

As duas classes mencionadas anteriormente foram utilizadas neste trabalho, juntamente com o modelo de mobilidade *ConstantPositionMobilityModel*, para posicionar as estações base e os edifícios apropriadamente e com posição constante ao longo da simulação. Em contrapartida, as estações móveis foram posicionadas utilizando o modelo de mobilidade *SteadyState-RandomWaypointMobilityModel*, que permite posicionar um conjunto de usuários em posições aleatórias dentro de uma determinada região retangular, com uma velocidade aleatória dentro de um determinado intervalo e com uma direção aleatória. Detalhes sobre os parâmetros de simulação utilizados são apresentados no próximo capítulo.

Outra importante característica que pode ser adicionada à simulação é o modelo de desva-

necimento do sinal. O módulo LENA fornece um *script* em MATLAB que gera como saída um arquivo com coeficientes que pode ser utilizado posteriormente pelo simulador para o cálculo do desvanecimento. Esse arquivo gerado contém um coeficiente para cada bloco de recurso, para cada instante de tempo da simulação. Os parâmetros para o gerador são: frequência usada, velocidade do usuário, duração da simulação e número de blocos de recursos. Internamente, esse *script* utiliza o modelo de desvanecimento de Rayleigh, que leva em conta o efeito Doppler causado pelo movimento do usuário em relação à estação base, entre outros fatores.

O módulo LENA fornece uma classe que permite criar um REM (*Radio Environment Map*), que consiste em um *grid* 2D de valores que representa a razão sinal ruído (SINR) do *downlink* em relação à estação base com maior potência recebida em cada ponto [5]. No entanto, a geração desse *grid* tem um alto custo de memória e processamento, o que limitou a resolução máxima usada para os *grids* gerados nesse trabalho. Esses mapas foram utilizados para visualização dos cenários simulados, conforme descrito no próximo capítulo.

2.3.2 Trabalhos relacionados

Han, Ma e Zhang [4] também fizeram uso do módulo LENA no simulador ns-3 para validar o algoritmo de localização baseado no modelo de *grid* irregular definido pelos mesmos, que foi descrito na seção anterior. Para definir os cenários simulados, ele se basearam em documentos que descrevem padrões de projeto e parâmetros de simulação definidos pela 3GPP (*3rd Generation Partnership Project*), projeto responsável por gerar especificações e relatórios de redes e telefonia móvel. Um desses documentos [9], também foi utilizado neste trabalho para definição dos cenários simulados, que serão descritos no próximo capítulo. Alguns desses parâmetros dizem respeito, por exemplo, à topologia de estações base, distância entre essas estações, número de estações, frequência do sinal de RF, modelo de propagação, modelos de desvanecimento, entre outros.

Mondal, Turkka, Ristaniemi e Henttonen [6] também fizeram uso de simulação para avaliar o desempenho de técnicas de localização baseadas em *fingerprint* de RF, sendo as métricas de similaridade da etapa *online* o foco dessa avaliação, como mencionado na seção anterior. Os cenários simulados por eles também se basearam em padrões definidos pela 3GPP em [9]. O estudo realizado em [6] simula ambientes rurais e urbanos, e considera apenas as medições *outdoor* no cálculo do erro médio da estimativa da posição. Um diferencial deste presente trabalho em relação ao estudo realizado em [6] é que, além dos ambientes rurais e urbanos, este trabalho também simula ambientes suburbanos. Os parâmetros utilizados em cada um desses ambientes são apresentados no próximo capítulo. Outro diferencial deste trabalho em relação ao estudo realizado em [6] é o uso das medições *indoor* tanto na etapa *offline* quanto na etapa *online*, uma vez que um dos objetivos deste trabalho é avaliar o impacto das medições fora das vias urbanas no modelo de *grid* proposto em [1] e esse conjunto de medições inclui medições *indoor*.

Método de Avaliação por Simulação

Este capítulo descreve o método utilizado neste trabalho para avaliação da técnica de localização em questão por meio de simulação. A Seção 3.1 explica como os cenários simulados são definidos. A Seção 3.2 descreve como as bases de dados são geradas através de múltiplas execuções do simulador. Por fim, a Seção 3.3 explica o processo de avaliação do desempenho do algoritmo de localização sobre os cenários simulados.

3.1 Definição dos Cenários de Simulação

3.1.1 Visualização dos Cenários Simulados

Conforme descrito no capítulo anterior, é possível gerar um REM para um determinado cenário de simulação. A região sobre a qual o REM é calculado neste trabalho coincide com a região para a qual é definido o mapa de cobertura do método de localização baseado em *fingerprint*. Além disso, o módulo LENA fornece métodos que geram arquivos plotáveis com informações sobre as estações base, estações móveis e edificações.

Esses arquivos gerados pelo simulador foram combinados em um *script*, utilizando a biblioteca gráfica *Gnuplot*, a fim de obter uma visualização dos cenários simulados. Tal visualização permite conferir se os diferentes componentes foram posicionados corretamente durante a simulação e se os modelos de propagação e desvanecimento foram configurados adequadamente. Essa última informação pode ser obtida a partir dos valores do REM, que correspondem à razão sinal ruído com respeito à estação base com melhor sinal em cada ponto, dando uma noção da qualidade do sinal em cada ponto. Esse método de visualização foi utilizado para gerar algumas das imagens presentes neste trabalho. Esse é o caso da Figura 3.1 e da Figura 3.2, por exemplo.

3.1.2 Parâmetros de Simulação

Além das bibliotecas fornecidas pelo módulo LENA, há ainda códigos de exemplo que demonstram como usar cada uma dessas bibliotecas. Um desses códigos de exemplo segue a topologia de rede e os parâmetros de simulação definidos em [10], que se chama *Dual Stripe Model*. Esse nome diz respeito ao formato dos quarteirões criados nesses cenários de simulação, que correspondem a dois prédios, sendo cada um deles uma linha reta de apartamentos, como ilustrado da Figura 3.1. Nesse modelo, além das estações base externas, chamadas de *MacroUEs*, são definidas estações base internas, chamadas de *Femtocells*. Essas *Femtocells* emitem uma potência mais baixa e são posicionadas aleatoriamente dentro de alguns apartamentos. Esse código de exemplo foi utilizado como ponto de partida para o simulador desenvolvido neste

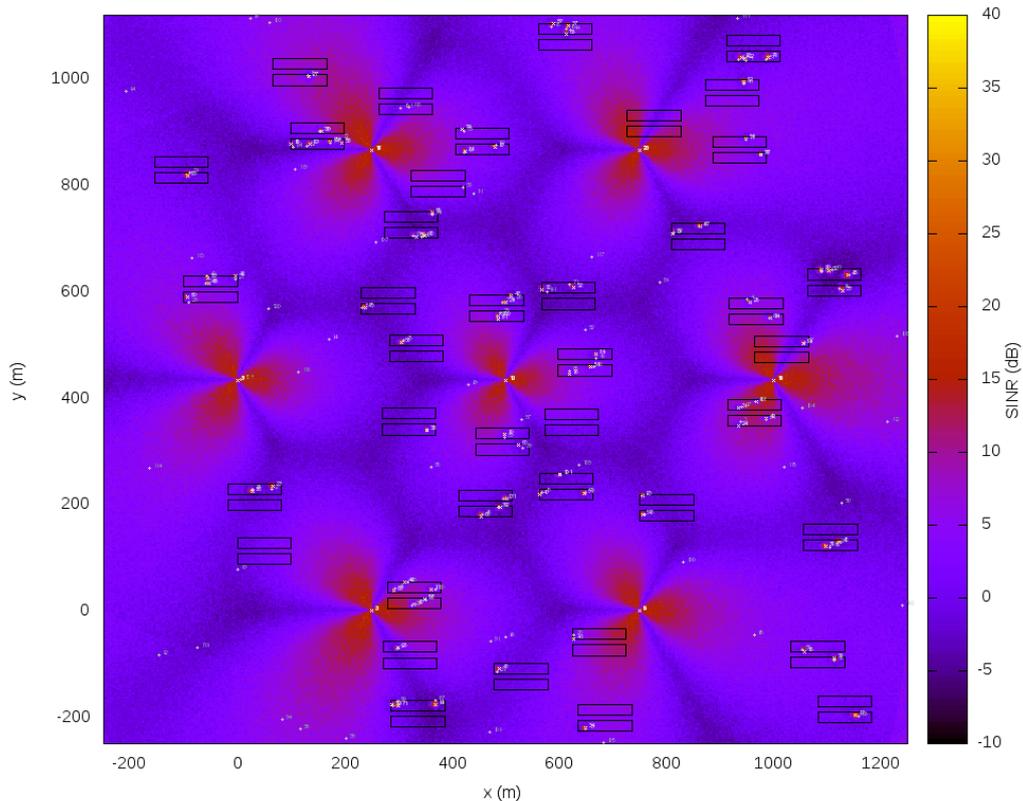


Figura 3.1 Representação gráfica de um cenário do *Dual Stripe* original, com 40 quarteirões e 21 estações base (7 torres tri-setoriais). Os retângulos correspondem aos edifícios e os pontos, aos usuários.

trabalho. A parte referente às *Femtocells* foi removida para simplificar as simulações e para que as bases de dados geradas tivessem o mesmo formato da base utilizada por em [1], na qual todas as medições contém o valor de RSRP de todas as estações base.

Utilizando a classe *GridBuildingAllocator*, foram criados quarteirões com 6 edifícios cada, seguindo uma topologia de *grid* 2×3 . Cada edifício contém 3 apartamentos na dimensão x e 5 apartamentos na dimensão y , todos com dimensões $10\text{ m} \times 10\text{ m}$. O espaçamento entre dois edifícios adjacentes num mesmo quarteirão é sempre 10 m . O número de andares em cada quarteirão é escolhido aleatoriamente dentro de uma faixa de valores que depende do tipo de ambiente simulado, conforme especificado na próxima seção. Todos os edifícios foram definidos como residenciais com paredes de concreto com janelas (configurações padrão do *Buildings Model*). A partir disso, foi implementada a classe *GridBlockAllocator*, que preenche o cenário simulado com quarteirões seguindo um modelo de *grid* 16×7 . As avenidas que separam esses quarteirões tem 25 m de largura. As dimensões utilizadas na criação e posicionamento desses edifícios foram definidas com base em parâmetros de simulação encontrados em [10] e [9]. Com essas dimensões, a região de cobertura consiste em um retângulo com dimensões $1495\text{ m} \times 1340\text{ m}$, com uma área de 2 Km^2 . A Figura 3.2 mostra o resultado do posicionamento desses quarteirões.

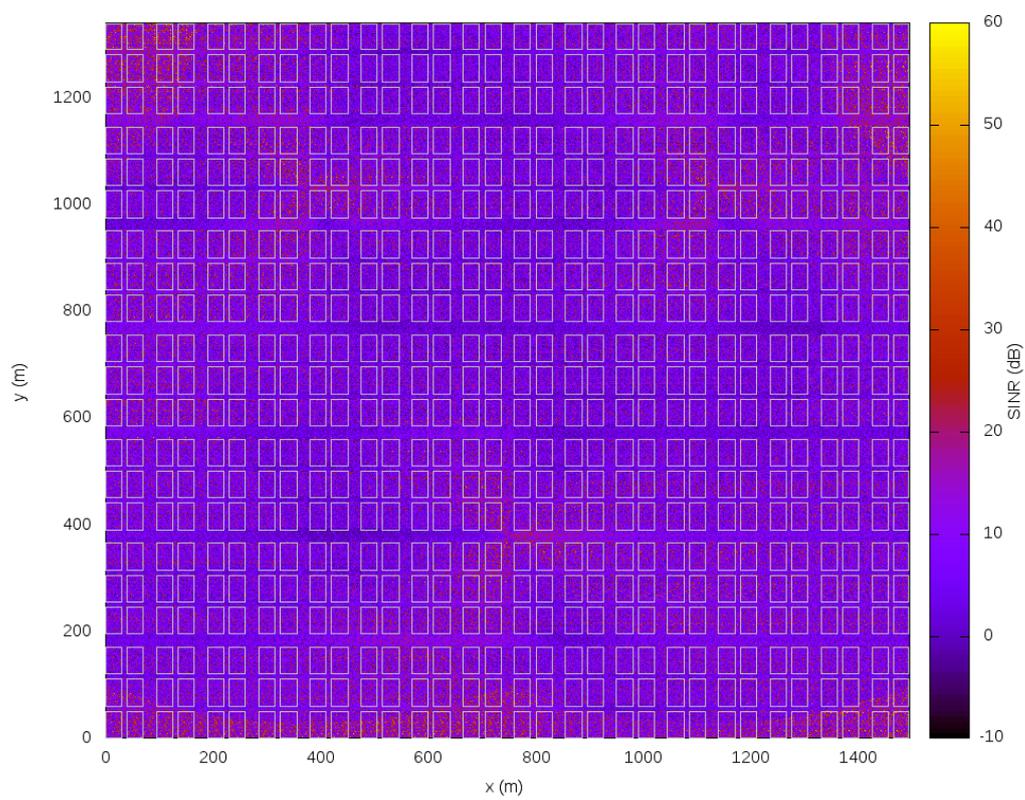


Figura 3.2 Posicionamento dos edifícios num cenário urbano, indicados por retângulos cinza.

Tabela 3.1 Parâmetros de configuração das estações base

Parâmetro	Valor
Modelo de Antena	Parabólica
Potência Emitida	46 dBm
Frequência da Portadora	2 GHz
Altura da Torre	30 m
Blocos de Recurso	25

Para posicionamento das estações base, foi utilizada a classe *LteHexGridEnbTopologyHelper*, que utiliza uma topologia hexagonal para posicionar estações base tri-setoriais. Foram posicionadas 3 estações tri-setoriais, que atuam como 9 estações base distintas, uma vez que, em cada ponto da região de cobertura, é possível obter 9 valores distintos de potência recebida. A distância entre duas estações adjacentes nos três tipos de ambientes simulados é de 750 m, um valor que está entre os valores sugeridos em [10] e [9] para simulações de ambientes urbanos e rurais. Dessa forma, essa distância aproxima bem os três tipos de cenários simulados. Os parâmetros usados na definição das estações bases são apresentados na Tabela 3.1, e foram definidos com base nas especificações presentes em [10].

Considerando que a técnica de localização avaliada nesse trabalho tem como foco os usuários *outdoor*, as estações móveis foram posicionadas ao longo de ruas do cenário simulado. Além disso, os trechos de ruas selecionados para esse posicionamento formam um trajeto que imita o percurso de um *drive-test*, com o intuito de gerar bases de dados semelhantes às que habitualmente são coletadas para esse tipo de estudo. Além disso, a ausência de amostras em determinados trechos da área de cobertura obriga a etapa de treino do *grid* a estimar os valores de potência recebida para alguns pontos, que é uma característica importante em técnicas de localização baseada em *fingerprint*. Dessa forma, a ausência de amostras em determinados trechos do cenário simulado é importante para avaliar apropriadamente a técnica em questão.

Como um dos objetivos desse estudo é avaliar o impacto das medições localizadas fora das ruas e avenidas no desempenho da técnica proposta em [1], uma fração dos usuários foi posicionada de modo a satisfazer essa condição, distribuídos aleatoriamente em cinco quarteirões adjacentes ao trajeto que simula o *drive-test*. Essas medições incluem estações móveis posicionadas dentro dos edifícios e nos espaços entre edifícios adjacentes (fora das avenidas). Bases de dados reais que incluem medições realizadas em vias urbanas e medições realizadas fora dessas vias podem ser obtidas para estudos como esse através de um único *drive-test*, ou de múltiplos *drive-tests* realizados na mesma região.

Utilizando o modelo de mobilidade *SteadyStateRandomWaypointMobilityModel*, fornecido pelo módulo LENA, todas as estações móveis foram posicionadas aleatoriamente dentro de regiões predefinidas com uma velocidade de 3 Km/h em uma direção escolhida aleatoriamente. Em todos os cenários simulados, foram criadas 3000 estações móveis, o que resulta no mesmo número de medições de potência recebida. Foram avaliados cenários com diferentes frações de medições localizadas fora das vias urbanas, como mostrado no próximo capítulo. A Figura 3.3 apresenta o posicionamento dos usuários e das estações base ao longo do cenário de simulação, para uma configuração com 20% dos usuários posicionados fora das vias urbanas.

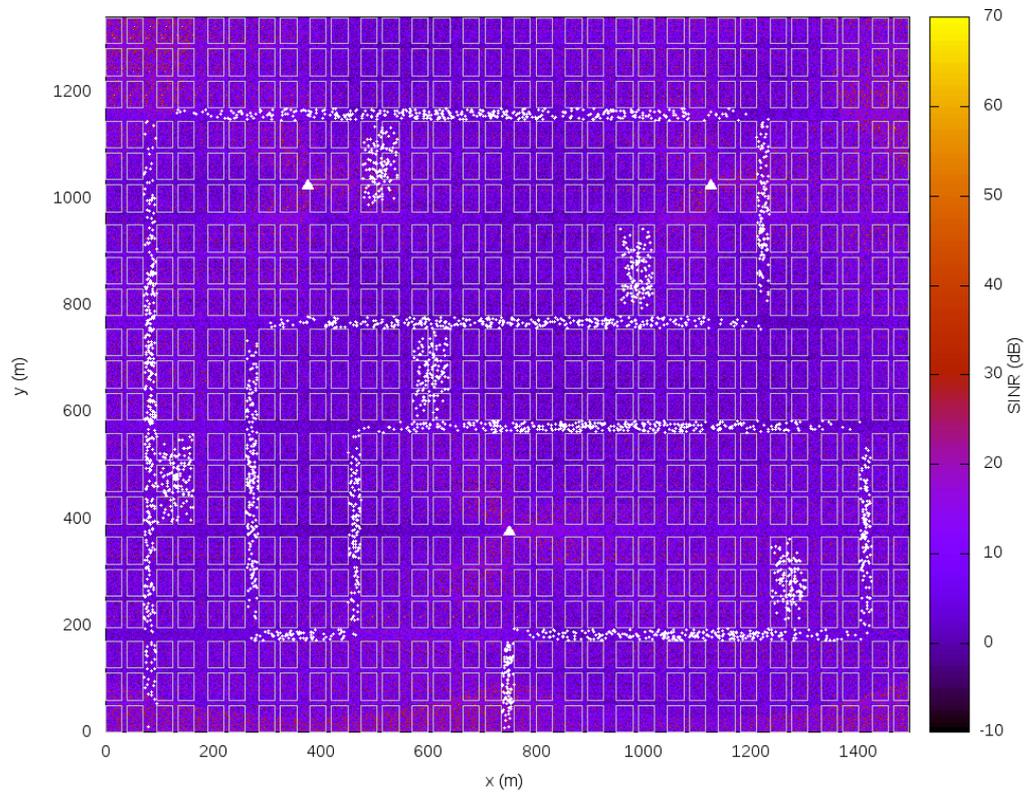


Figura 3.3 Posicionamento dos edifícios (retângulos), usuários (pontos) e estações base (triângulos) ao longo de um cenário urbano, com 20% dos usuários posicionados fora das vias urbanas.

Tabela 3.2 Desvio padrão da distribuição log-normal que calcula o sombreamento do sinal

Ambiente	σ (dB)
<i>Outdoor</i>	4
<i>Indoor</i>	3
Paredes Externas	10

O modelo de propagação utilizado foi o implementado pela classe *HybridBuildingsPropagationLossModel*, fornecida pelo módulo LENA, que consiste em uma combinação de modelos de propagação amplamente conhecidos, com o intuito de simular comunicações *outdoor*, *indoor*, de *outdoor* para *indoor* e de *indoor* para *outdoor*. Conforme descrito na documentação da própria classe, os modelos de propagação utilizados internamente pela mesma são: Okumura Hata, COST231, ITU-R P.1411 (para comunicações de curta distância) e ITU-R P.1238 (para comunicações *indoor*).

Esse modelo leva ainda em consideração propriedades do *Buildings Model* para estimar as perdas de caminho causadas pelas paredes internas e externas dos edifícios, além dos efeitos de sombreamento, que são estimados de acordo com uma distribuição log-normal com média nula e desvio padrão variável. Esse desvio padrão assume valores diferentes para ambientes *outdoor*, *indoor* e para as paredes externas dos edifícios (em caso de comunicação entre um nó *outdoor* e um nó *indoor*), como mostrado na Tabela 3.2. O valor estabelecido para a perda de potência para cada parede interna atravessada pelo sinal de RF foi de 10 dB para essa simulação. Os valores desses parâmetros foram escolhidos com base nos valores estabelecidos pela 3GPP em [10] e [9].

Os modelos de desvanecimento utilizados neste trabalho foram dois dos fornecidos pelo módulo LENA, cujos parâmetros foram definidos com base em especificações da 3GPP apresentadas em [11]. Um dos modelos é o *Extended Typical Urban model* (ETU), que é específico para ambientes urbanos. O segundo é o *Extended Pedestrian A model* (EPA), que é mais genérico. No entanto, ambos assumem que os usuários são pedestres a uma velocidade de 3 Km/h, a mesma velocidade definida para todos os usuários nas simulações deste trabalho. De acordo o tipo de ambiente escolhido para a simulação, um desses dois modelos de desvanecimento é utilizado, como mostrado na Tabela 3.3.

Após a adição do modelo de desvanecimento, foi possível observar oscilações nos níveis de RSRP recebidos pelo mesmo usuário ao longo da simulação. A explicação para esse fenômeno é que esse modelo considera a propagação multi-percurso, na qual as diferentes componentes do sinal de RF enviado chegam ao mesmo destino com fases que variam com o tempo, o que pode resultar na soma ou no cancelamento dessas componentes. Dessa forma, o valor da potência resultante varia consideravelmente ao longo do tempo. A forma como esses dados foram tratados, para obtenção de um valor significativo do nível de RSRP é apresentada na próxima seção.

Tabela 3.3 Parâmetros de simulação para cada tipo de ambiente

Parâmetro	Tipo de Ambiente		
	urbano	suburbano	rural
Altura Média dos Edifícios	20 m	10 m	5 m
Mínimo de Andares	1	1	1
Máximo de Andares	9	5	2
Densidade de Edifícios	100%	60%	10%
Modelo de Desvanecimento	ETU	EPA	EPA

3.1.3 Tipos de Ambientes Simulados

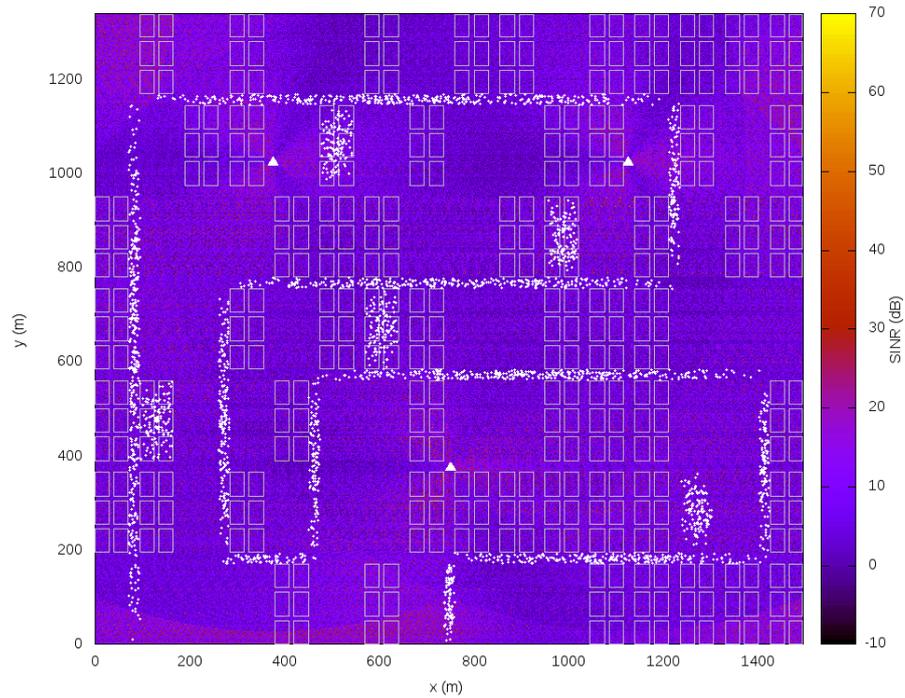
Com o objetivo de avaliar o desempenho do *grid* irregular proposto em [1] sobre ambientes com diferentes densidades de ocupação urbana, foram definidos três tipos de ambientes simulados: urbano, suburbano e rural. A classe utilizada para modelar a propagação do sinal de RF recebe parâmetros referentes ao tipo de ambiente e ao tamanho da cidade. Esses parâmetros foram definidos adequadamente para cada um dos três tipos de ambiente simulados. Um outro parâmetro recebido por essa classe é a altura dos edifícios, que foi definida de acordo com os valores sugeridos pela 3GPP em [9] para a altura média dos edifícios. Essas alturas médias e outros parâmetro que variam de acordo com o tipo de ambiente são apresentados na Tabela 3.3. A densidade de edifícios, mostrada nessa tabela, corresponde a uma probabilidade utilizada pela classe *GridBlockAllocator* para decidir se deve ser criado um quarteirão, para cada posição do *grid* de dimensões 16×7 .

As Figuras 3.3, 3.4a e 3.4b mostram exemplos de cenários de simulação correspondentes a cada um dos três tipos de ambientes. Conforme mostrado nessas figuras, as posições das vias urbanas são as mesmas nos três tipos de ambientes simulados, assim como a trajetória de medições é a mesma. Esses parâmetros foram fixados a fim de simplificar o processo de posicionamento dos edifícios e das estações móveis na definição dos cenários de simulação.

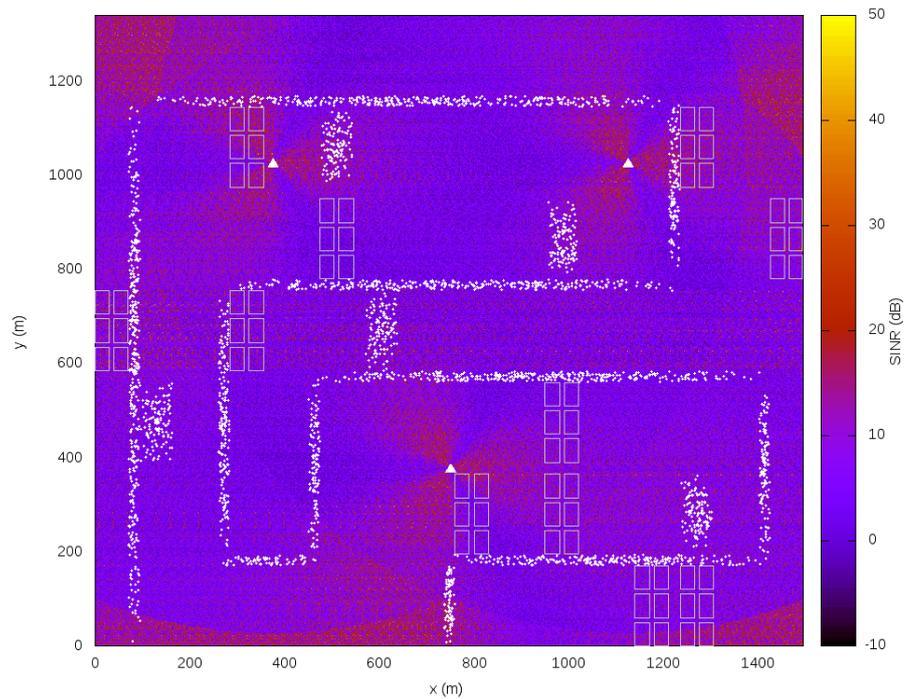
3.2 Geração das Bases de Dados

Uma vez que todos os usuários são posicionados de forma aleatória dentro de regiões predefinidas do cenário simulado, foi adicionada à simulação deste trabalho um método que imprime as posições de cada usuário, juntamente ao seu identificador dentro do simulador, o IMSI. Esse método foi implementado com base no método fornecido pelo módulo LENA que gera um arquivo plotável com informações sobre as estações móveis. Essas posições são informadas em coordenadas cartesianas (x, y) , onde $0 \text{ m} \leq x \leq 1495 \text{ m}$ e $0 \text{ m} \leq y \leq 1340 \text{ m}$. Esses limites correspondem aos limites da região de cobertura, descrita na seção anterior. O método que imprime as informações dos usuários é chamado apenas uma vez para cada cenário simulado.

Além disso, foi adicionada à simulação deste trabalho um método para imprimir informações sobre cada estação base, juntamente ao seu identificador dentro do simulador, o *CellId*. Esse método foi implementado com base no método fornecido pelo módulo LENA que gera um arquivo plotável com informações sobre as estações base. Dentre as informações impressas, a



(a) Ambiente suburbano



(b) Ambiente rural

Figura 3.4 Posicionamento dos edifícios (retângulos), usuários (pontos) e estações base (triângulos) ao longo dos cenários suburbano e rural, com 20% dos usuários posicionados fora das vias urbanas.

Tabela 3.4 Amostras das bases de dados geradas através de simulação

x	y	rsrp 1	rsrp 2	rsrp 3	rsrp 4	rsrp 5	rsrp 6	rsrp 7	rsrp 8	rsrp 9
78.6	509.1	-106.4	-85.3	-91.7	-99.4	-89.7	-78.6	-139.8	-142.7	-122.2
86.1	188.4	-101.9	-92.2	-88.5	-100.8	-97.4	-77.4	-133.9	-135.4	-124.9
82.9	508.3	-96.9	-87.9	-87.1	-100.6	-91.6	-75.7	-134.1	-141.9	-116.8

mais relevante é a posição da estação base, que é utilizada nas etapas *online* e *offline* para o cálculo do TA, que é importante para filtrar células do mapa de cobertura para reduzir o espaço de busca da etapa *online* [1]. Essas posições são informadas em coordenadas cartesianas, assim como as posições de cada usuário.

A saída gerada por um simulador que utiliza o módulo LENA no ns-3 inclui dados da camada física, como os níveis de RSRP. Por padrão, é impressa uma amostragem para cada milissegundo da simulação. Esses níveis devem ser reportados para camadas superiores da pilha de comunicação a cada 200 ms e o valor reportado deve ser a média de todas as amostras obtidas durante esse período [5]. Desse modo, para tornar as base de dados simuladas mais realistas, cada simulação executada nesse estudo teve um tempo de duração de 200 ms.

Uma restrição desse simulador é que, para cada usuário, são informados os níveis de potência recebida apenas da estação base à qual este usuário está conectado. Para contornar essa limitação, são executadas nove simulações para cada cenário de simulação, uma para cada estação base da região simulada. Na primeira simulação, todos os usuários se conectam na estação base 1, na segunda, todos se conectam na estação base 2, e assim sucessivamente. Dessa forma, é gerado um arquivo de saída para cada estação base, com os valores de RSRP correspondentes a essa estação base.

A Tabela 3.4 ilustra o formato desejado para as bases de dados geradas através de simulação, na qual os valores de exemplo foram arredondados para uma casa decimal, a fim de reduzir a largura da tabela. A fim de alcançar esse formato, todos os dados gerados nos passos anteriores são agregados em um único arquivo de medições. As duas primeiras colunas, posições x e y do usuário, são obtidas a partir do arquivo correspondente aos usuários, mencionado anteriormente. Em seguida, esse agregador é responsável por calcular a média dos 200 valores de RSRP impressos pelo simulador (uma amostra a cada 1 ms durante um período de 200 ms) para cada par ($IMSI, CellId$). Por fim, o agregador converte as medições de RSRP de Watts para dBm utilizando a Equação 3.1.

$$P_{dBm} = 10 \times \log_{10} \frac{P_W}{1mW} \quad (3.1)$$

O módulo LENA permite a configuração de uma semente aleatória que afeta todos os componentes aleatórios dos cenários de simulação. No caso dos cenários descritos neste trabalho, essa semente influencia nas alturas dos edifícios, nas posições e direções de movimento dos usuários e em componentes aleatórias dos modelos de propagação, por exemplo. Para cada um dos cenários avaliados neste trabalho, foi gerada apenas uma base de dados, utilizando sempre a mesma semente aleatória. A geração de múltiplas bases de dados para cada cenário avaliado permitiria analisar o desvio padrão do erro médio da técnica para cada *grid* avaliado. No

entanto, o custo computacional de realizar múltiplas gerações de bases de dados e múltiplas avaliações da técnica de localização inviabilizaram essa análise.

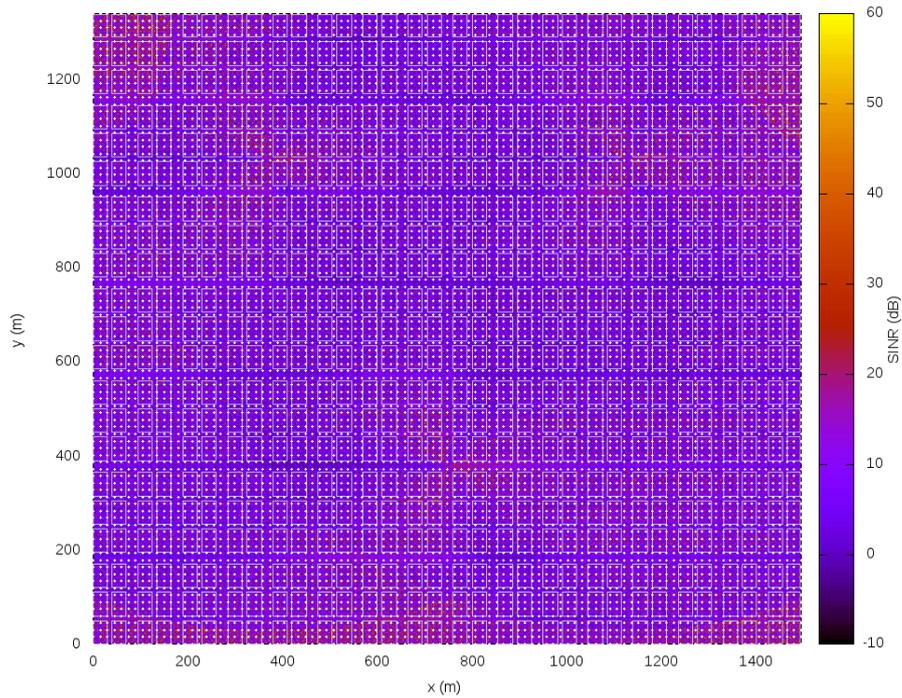
3.3 Avaliação do Algoritmo de Localização

Para a avaliação do algoritmo de localização proposto em [1], foi utilizada a mesma implementação usada pelos autores do artigo, com algumas adaptações necessárias para que o algoritmo lide com as posições dadas em coordenadas cartesianas (x, y) das bases de dados simuladas, em vez das coordenadas geográficas (lat, lon) da base de dados original. Essa implementação consiste em duas fases: geração e avaliação dos *grids*.

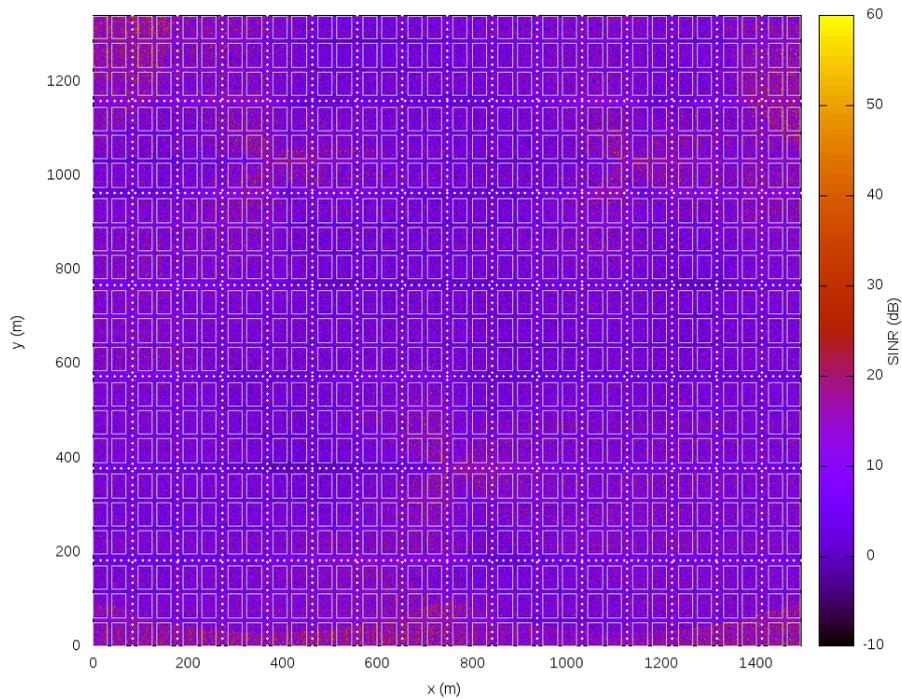
Na fase de geração dos *grids*, as classes que representam as coordenadas e fazem a interpolação de pontos adjacentes de um segmento foram adaptadas para trabalharem com coordenadas cartesianas. Na geração dos *grids* regulares, foram informados os limites do retângulo correspondente à região de cobertura dos cenários simulados. Na geração dos *grids* irregulares, foram informadas as coordenadas de início e fim de cada avenida dos cenários simulados. Foram gerados *grids* regulares e irregulares com todas as resoluções inteiras de 1 a 20 metros.

Ao final da geração dos *grids*, o algoritmo informa o número de células em cada um deles, informação importante para a estimativa da complexidade de processamento e memória da técnica de localização ao utilizar cada um dos *grids*. Ao final da geração dos *grids*, foi criado um arquivo plotável para cada um deles, que permite a visualização das posições dos centros das células do *grid* sobre o cenário simulado. As Figuras 3.5a e 3.5b mostram as células de um *grid* regular e irregular, respectivamente. Nessas representações gráficas, fica clara a redução do número de células do *grid* irregular em relação ao *grid* regular.

Na fase de avaliação dos *grids*, ajustes necessários para lidar com coordenadas cartesianas foram feitos na representações das estações base, das medições e no cálculo das distâncias. Foram avaliados *grids* regulares e irregulares gerados na fase anterior, com todas as resoluções inteiras de 5 a 20 metros. Os *grids* com resoluções inferiores a 5 metros não foram avaliados por conta do alto custo computacional da técnica de localização, quando aplicada sobre os *grids* regulares com essas resoluções. Os resultados dessas avaliações são apresentados no próximo capítulo.



(a) Grid regular



(b) Grid irregular

Figura 3.5 Células dos *grids* regular e irregular com resolução de 15 m. Os centros das células são representados por pontos brancos, enquanto os edifícios são representados por retângulos cinza.

Experimentos e Análise dos Resultados

Este capítulo descreve os experimentos realizados para avaliar a técnica de localização proposta em [1] e apresenta os resultados desses experimentos. Primeiramente, são apresentados os números de células geradas em cada um dos *grids*, informações que também serão utilizadas nas seções mais adiante neste capítulo. As Seções 4.1 e 4.2 descrevem as avaliações realizadas para medir as influências das medições fora das vias urbanas e da densidade de ocupação urbana sobre a técnica avaliada, respectivamente.

Como mencionado no capítulo anterior, o gerador de *grids* informa o número de células geradas em cada um dos *grids*. A Tabela 4.1 apresenta esses números para *grids* regulares e irregulares com todas as resoluções inteiras de 1 a 20 metros. A relação entre o número de células no *grid* regular $N_{regular}$ e a resolução r pode ser descrita pela Equação 4.1, onde Δx e Δy correspondem aos comprimentos da região de cobertura nos eixos x e y , respectivamente. Já a relação entre o número de células no *grid* irregular $N_{irregular}$ e a resolução r , nos cenários simulados neste trabalho, pode ser descrita pela Equação 4.2, onde n_x e n_y correspondem aos números de vias urbanas paralelas aos eixos x e y , respectivamente. As constantes c_1 e c_2 foram definidas para simplificar as expressões.

$$N_{regular} = \frac{\Delta x}{r} \times \frac{\Delta y}{r} = \frac{\Delta x \Delta y}{r^2} = \frac{c_1}{r^2} \quad (4.1)$$

$$N_{irregular} = \frac{n_x \Delta x}{r} + \frac{n_y \Delta y}{r} = \frac{n_x \Delta x + n_y \Delta y}{r} = \frac{c_2}{r} \quad (4.2)$$

Além do número de células em cada *grid*, a Tabela 4.1 contém a taxa de redução (TR) do número de células nos *grids* irregulares em relação aos regulares para cada resolução, definida pela Equação 4.3. Podemos encontrar a relação entre TR e a resolução r , substituindo $N_{regular}$ e $N_{irregular}$ na Equação 4.3 pelas Equações 4.1 e 4.2. Esse resultado é apresentado na Equação 4.4. A constante c também foi definida para simplificar a expressão. Por fim, observamos que TR decai linearmente com o aumento da resolução r .

$$TR = \left(1 - \frac{N_{irregular}}{N_{regular}}\right) \times 100 \quad (4.3)$$

$$TR = \left(1 - \frac{c_2}{c_1} r\right) \times 100 = (1 - cr) \times 100 \quad (4.4)$$

Essa relação entre TR e a resolução r nos permite concluir que o uso do modelo de *grid* irregular em substituição ao regular é ainda mais vantajoso quando se utiliza baixas resoluções r , em termos da redução da complexidade de processamento e memória. Desta forma, o modelo

Tabela 4.1 Número de células e taxa de redução em cada um dos *grids* gerados

Resolução (m)	Regular (células)	Irregular (células)	TR (%)
1	2006136	29091	98.55
2	502579	14559	97.10
3	224000	9720	95.66
4	126000	7290	94.21
5	80700	5835	92.77
6	56475	4881	91.36
7	41495	4185	89.91
8	31772	3663	88.47
9	25200	3258	87.07
10	20385	2931	85.62
11	16851	2667	84.17
12	14238	2451	82.79
13	12180	2271	81.35
14	10476	2103	79.93
15	9191	1971	78.56
16	8075	1845	77.15
17	7120	1734	75.65
18	6460	1650	74.46
19	5760	1560	72.92
20	5168	1476	71.44

de *grid* irregular viabiliza a aplicação de técnicas de localização baseadas em *fingerprint* de RF sobre *grids* com baixa resolução r , em busca de uma melhor precisão na estimativa de localização. Como evidência disso, a partir da Tabela 4.1, podemos observar que os *grids* irregulares com resoluções de 1 m , 3 m e 5 m tem aproximadamente o mesmo número de células que os *grids* regulares com resoluções de 8 m , 14 m e 19 m , respectivamente.

O estudo realizado em [1] aponta que os *grids* irregulares obtiveram um desempenho melhor do que os regulares em relação ao erro médio, para todas as resoluções testadas. No entanto, os autores afirmam que a vantagem dos *grids* irregulares sobre os regulares é mais marcante quando se considera a quantidade de células necessárias para se chegar a um erro médio desejado. Por essa razão, o número de células geradas em cada *grid* será importante também nas avaliações do erro médio da técnica de localização com cada um dos modelos de *grid*. Essas avaliações serão apresentadas nas próximas seções deste capítulo.

4.1 Influência de Medições Fora das Vias Urbanas

A técnica de avaliação utilizada neste trabalho foi a mesma utilizada em [1], que corresponde a um *k-fold cross validation* com $k = 10$. Como métrica de avaliação, é calculado o erro médio μ para cada *grid* avaliado. Para cada cenário simulado, foram avaliados os *grids* regulares e irregulares com todas as resoluções r inteiras de 5 m a 20 m . Os *grids* com valores de r menores que 5 m não foram avaliados por conta do alto custo computacional das etapas *offline* e *online* da técnica de localização ao utilizar os *grids* regulares com essas resoluções. O alto número de células nesses *grids*, mostrados na Tabela 4.1, evidencia a alta complexidade de tempo e memória nas fases de treino e teste dessa técnica. Restringindo-se as avaliações a essas resoluções mencionadas, foi possível avaliar a técnica de localização sobre uma quantidade maior de cenários simulados.

Conforme mencionado no capítulo anterior, foram simulados cenários com diferentes proporções de usuários localizados fora das ruas e avenidas, com o objetivo de avaliar o impacto dessas medições sobre o desempenho da técnica de localização avaliada neste trabalho. Essas medições foram utilizadas tanto na fase de treino quanto na fase de teste. Considerando que a técnica proposta em [1] foi idealizada para cenários urbanos densamente ocupados, a avaliação realizada nesta seção contempla apenas simulações com ambiente do tipo urbano. Para esse ambiente, foram avaliados cenários com as seguintes frações de medições realizadas nas vias urbanas: 100%, 90%, 80%, 70% e 60%. Os erros médios por resolução para esses cenários são apresentados nas Figuras 4.1, 4.2, 4.3, 4.4 e 4.5, respectivamente.

No cenário com 100% das medições realizadas nas vias urbanas, pode-se observar na Figura 4.1 que o modelo de *grid* irregular obteve menor erro médio para $r \geq 14$ m . Em contraposição, para $r < 14$ m , o modelo de *grid* regular apresentou menor erro médio. Uma possível explicação para esse comportamento é o fato de que as células do *grid* irregular são sempre posicionadas no centro das ruas e avenidas, cuja largura nos cenários simulados é sempre de 25 m . Deste modo, as medições realizadas longe do centro (local que corresponderia às calçadas), sempre terão um erro mínimo associado a elas. Já os *grids* regulares com valores menores para a resolução r contém mais células dentro das ruas e avenidas, diminuindo o erro associado a essas medições. No entanto, as diferenças de erro entre os dois modelos de *grid* são de apenas

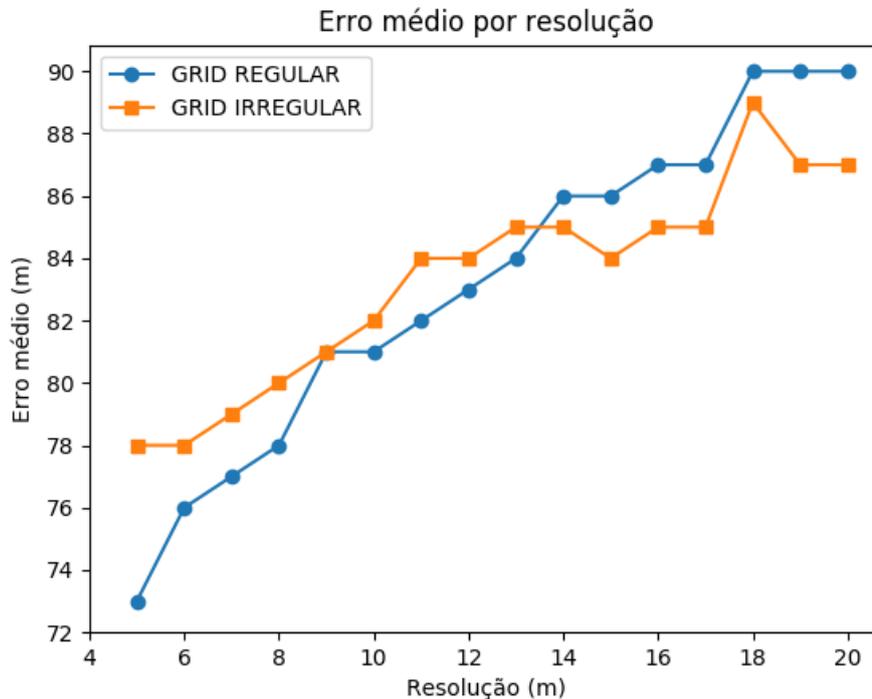


Figura 4.1 Erro médio por resolução num cenário urbano com 100% de medições em vias urbanas.

alguns metros e, por isso, não são tão significativas, especialmente para uma técnica voltada para localização *outdoor*. Logo, sendo os erros médios similares, o uso dos *grids* irregulares se justifica pela redução no número de células para cada resolução apresentada na Tabela 4.1. Esses resultados mostram que o modelo de *grid* irregular proporciona reduções significativas na complexidade de processamento e memória da técnica de localização sem reduzir significativamente a precisão da mesma.

Uma forma mais justa de se comparar os dois modelos de *grid* em relação à precisão da localização foi proposta em [1] e consiste em comparar a precisão de dois *grids* com aproximadamente o mesmo número de células, em vez de selecionar dois *grids* com a mesma resolução. A ideia desta proposta é identificar qual modelo de *grid* obtém melhor precisão, com aproximadamente o mesmo custo computacional. O *grid* irregular com resolução $r = 5\text{ m}$ tem 5835 células e apresenta um erro médio de 78 m, enquanto o *grid* regular com resolução $r = 19\text{ m}$ tem 5760 células e apresenta um erro médio de 90 m. Comparações análogas a essa podem ser feitas para diferentes resoluções e nos permite concluir que os *grids* irregulares apresentam melhor precisão na localização quando comparados a *grids* regulares com aproximadamente o mesmo número de células.

Os erros médios para o cenário com 90% das medições realizadas nas vias urbanas, mostrado na Figura 4.2, apresenta comportamento semelhante ao cenário anterior, onde o modelo de *grid* irregular obteve menor erro para valores altos de r e o modelo de *grid* regular apresentou menor erro para valores baixos de r . Esse comportamento tem a mesma explicação

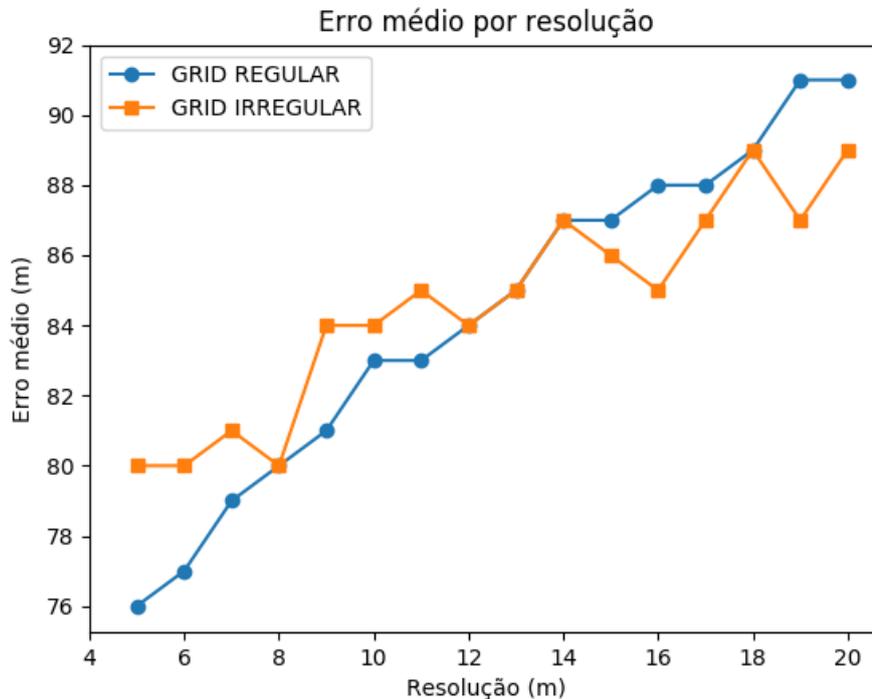


Figura 4.2 Erro médio por resolução num cenário urbano com 90% de medições em vias urbanas.

apresentada para o cenário anterior.

Os erros médios continuam sendo muito próximos, indicando que os 10% de medições realizadas fora das ruas e avenidas tiveram pouca influência sobre a precisão das técnicas. Por fim, observamos que o *grid* irregular com resolução $r = 5\text{ m}$ apresenta um erro médio de 80 m , enquanto o *grid* regular com resolução $r = 19\text{ m}$ apresenta um erro médio de 91 m , mostrando que os *grids* irregulares ainda apresentam melhor precisão na localização quando comparados a *grids* regulares com aproximadamente o mesmo número de células.

Um comportamento diferente dos anteriores pode ser observado nos erros médios do cenário com 80% das medições realizadas nas vias urbanas, como mostrado na Figura 4.3. Neste cenário, os *grids* regulares apresentam erro médio menor ou igual para todas as resoluções testadas, o que indica que os 20% de medições realizadas fora das ruas e avenidas já impactam negativamente de forma significativa a precisão do modelo de *grid* irregular. Para valores pequenos de r , a diferença entre os erros médios já é de 6 a 7 metros.

No entanto, nesse cenário, os *grids* irregulares ainda apresentam melhor precisão na localização quando comparados a *grids* regulares com aproximadamente o mesmo número de células. Como evidência disso, comparamos novamente o *grid* irregular com resolução $r = 5\text{ m}$, que tem erro médio igual a 86 m , com o *grid* regular com resolução $r = 19\text{ m}$, que apresenta um erro médio de 93 m .

Por fim, nos cenários com 70% e 60% das medições realizadas nas vias urbanas, os *grids* regulares apresentam erro médio inferior para todas as resoluções testadas, conforme mostrado

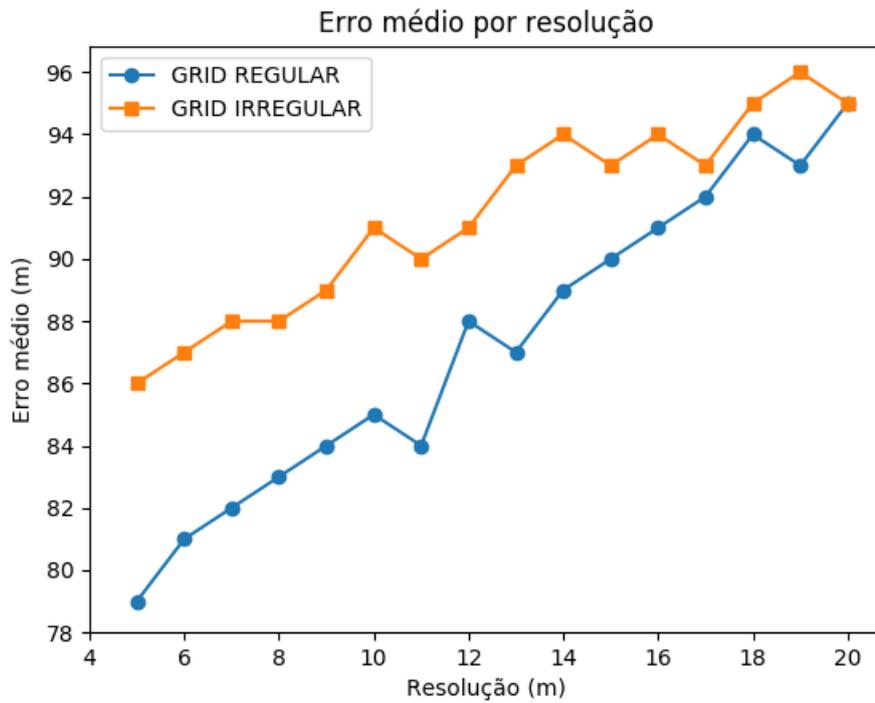


Figura 4.3 Erro médio por resolução num cenário urbano com 80% de medições em vias urbanas.

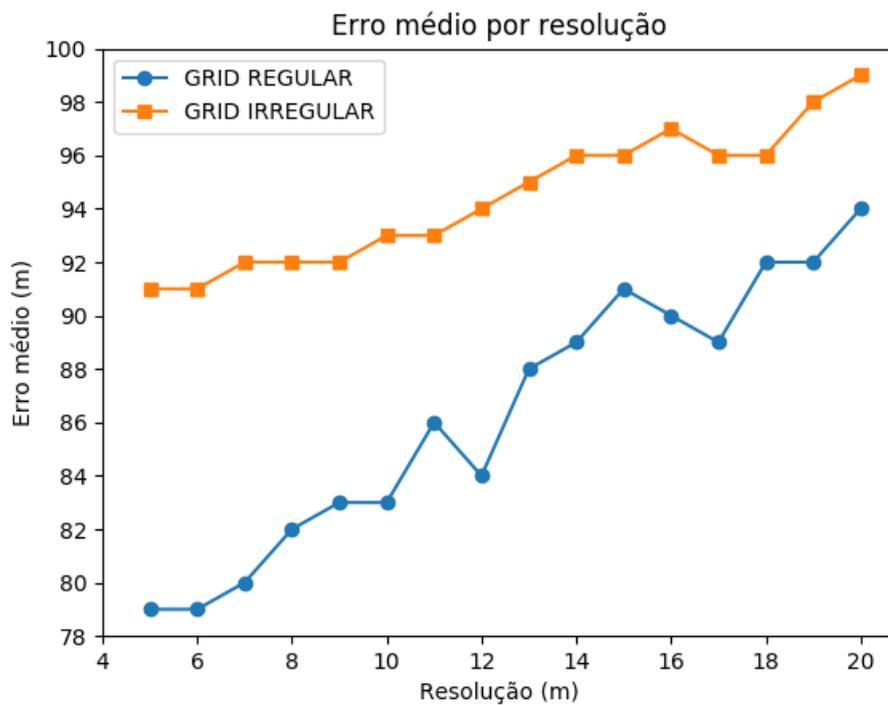


Figura 4.4 Erro médio por resolução num cenário urbano com 70% de medições em vias urbanas.

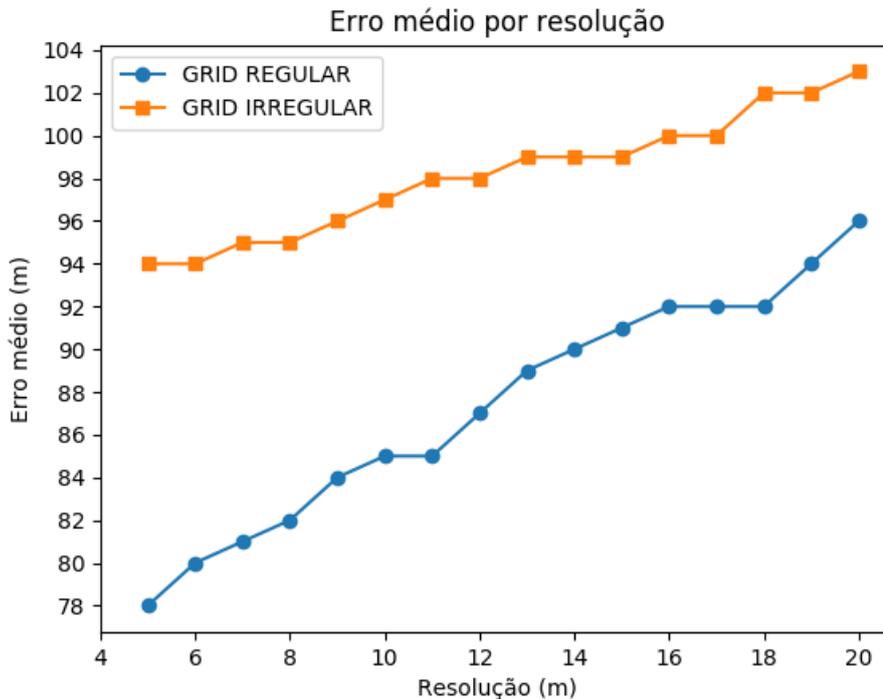


Figura 4.5 Erro médio por resolução num cenário urbano com 60% de medições em vias urbanas.

nas Figuras 4.4 e 4.5. Logo, as frações de 30% e 40% de medições realizadas fora das ruas e avenidas impactam ainda mais a precisão do modelo de *grid* irregular. Para valores pequenos de r , a diferença entre os erros médios chegam a valores entre 12 e 15 m. Além disso, para esses dois cenários, os *grids* irregulares já não apresentam mais uma precisão melhor na localização quando comparados a *grids* regulares com aproximadamente o mesmo número de células.

Portando, os experimentos realizados indicaram que, em cenários com 100%, 90% ou 80% de medições em ruas ou avenidas, os *grids* irregulares apresentam um menor erro médio de localização quando comparados a *grids* regulares com aproximadamente o mesmo número de células. Em contrapartida, esses experimentos apontam que, em cenários com 70% ou 60% de medições nas vias urbanas, os *grids* irregulares já não melhoram o erro médio de localização em relação a *grids* regulares com aproximadamente o mesmo número de células.

É importante ressaltar que o modelo de *grid* proposto em [1] foi idealizado para problemas de localização *outdoor* e que a avaliação descrita nesta seção tem como um de seus objetivos testar até que ponto esse modelo pode ser utilizado de forma genérica para localização *outdoor* e *indoor*. A redução no desempenho desse modelo de *grid* com a inserção de medições fora das vias urbanas já era esperada, uma vez que as posições estimadas para essas medições serão, naturalmente, localizadas em alguma rua ou avenida da região de cobertura. Além disso, em ambientes *indoor*, é comum que as técnicas de *fingerprinting* enfrentem grandes problemas com difração, reflexão e dispersão na propagação dos sinais de RF recebidos [12]. No entanto, esse estudo permitiu estimar a magnitude da redução no desempenho, além de indicar aproxi-

madamente a proporção limite de medições fora das vias urbanas que pode ser utilizada com esse modelo.

4.2 Influência da Densidade de Ocupação Urbana

A técnica proposta em [1] foi idealizada para cenários urbanos densamente ocupados. No entanto, nesta seção, é analisado se essa técnica pode ser aplicada também a outros tipos de ambiente. Uma vez que o ambiente do tipo urbano já foi contemplado na seção anterior, essa seção se concentra em ambientes dos tipos suburbano e rural.

4.2.1 Ambiente Suburbano

Primeiramente, analisaremos o cenário suburbano com 100% de medições nas das vias urbanas, para o qual foram obtidos os erros médios exibidos na Figura 4.6. Assim como no cenário urbano com mesma proporção de medições nessas vias, a diferença de erro entre os dois modelos de *grid* é pequena quando comparada com a magnitude do erro. No entanto, no cenário suburbano, os erros dos *grids* irregulares são quase sempre maiores do que os dos *grids* regulares de mesma resolução.

Analogamente ao cenário urbano, nesse cenário suburbano os *grids* irregulares apresentam melhor precisão na localização quando comparados a *grids* regulares com aproximadamente o mesmo número de células. Isso pode ser mostrado comparando novamente o *grid* irregular com resolução $r = 5\text{ m}$, que tem erro médio igual a 81 m , com o *grid* regular com resolução $r = 19\text{ m}$, que apresenta um erro médio de 95 m .

Em seguida, completamos nossa análise de ambientes suburbanos considerando cenários com 80% e 60% de medições nas ruas e avenidas, para os quais foram obtidos os erros médios apresentados nas Figuras 4.7 e 4.8. Os resultados encontrados são semelhantes aos apresentados para ambientes urbanos com as mesmas frações de medições nessas vias, nos quais o aumento da proporção de medições fora dessas vias deteriora gradativamente a precisão da técnica de localização. No cenário com 80% de medições nas vias urbanas, os *grids* irregulares ainda apresentam melhor precisão na localização quando comparados a *grids* regulares com aproximadamente o mesmo número de células. No cenário com 60% de medições nessas vias, isso já não é mais verdade, de modo que o modelo de *grid* irregular já não apresenta mais vantagens sobre o modelo de *grid* regular.

Além disso, pode se observar que os erros médios encontrados para ambos os modelos de *grid* no ambiente suburbano são, em média, 5 m maiores que os erros médios para o ambiente urbano. No entanto, a média das diferenças de erro entre os dois modelos de *grid* é aproximadamente a mesma nos ambientes urbanos e suburbanos. Logo, podemos concluir que nos cenários suburbanos simulados neste trabalho, a técnica de localização baseada em *fingerprint* de RF perdeu precisão em relação aos cenários urbanos simulados, independente do modelo de *grid* usado.

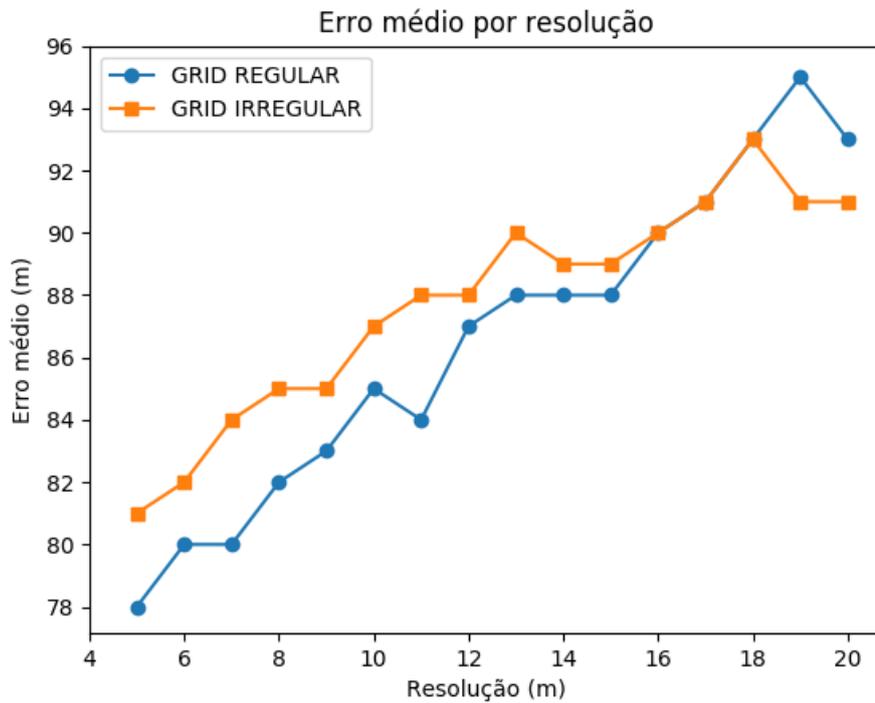


Figura 4.6 Erro médio por resolução num cenário suburbano com 100% de medições em vias urbanas.

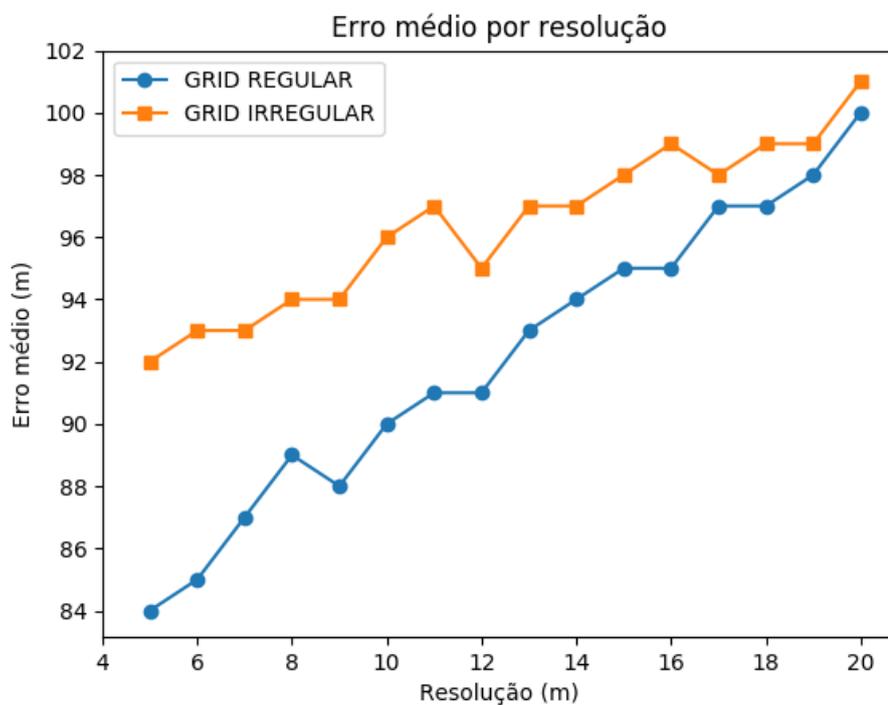


Figura 4.7 Erro médio por resolução num cenário suburbano com 80% de medições em vias urbanas.

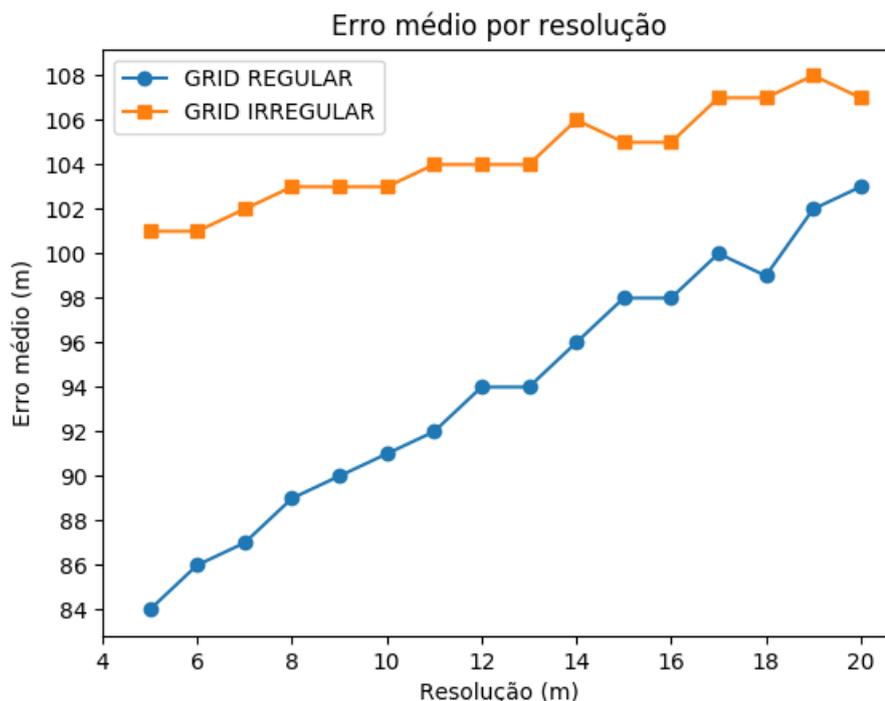


Figura 4.8 Erro médio por resolução num cenário suburbano com 60% de medições em vias urbanas.

4.2.2 Ambiente Rural

Seguindo a mesma ordem de avaliação, analisaremos primeiro o cenário rural com 100% de medições nas vias urbanas, para o qual foram obtidos os erros médios mostrados na Figura 4.9. Novamente, assim como nos ambientes urbanos e suburbanos com mesma porcentagem de medições nessas vias, a diferença de erro entre os dois modelos de *grid* não é significativa quando comparada com a magnitude do erro.

Assim como nos outros dois tipos de ambiente, nesse cenário rural os *grids* irregulares tem melhor precisão na localização quando comparados a *grids* regulares com aproximadamente o mesmo número de células. Tal propriedade pode ser constatada comparando novamente o *grid* irregular com resolução $r = 5 m$, que tem erro médio igual a $81 m$, com o *grid* regular com resolução $r = 19 m$, que apresenta um erro médio de $94 m$.

Por fim, analisamos cenários rurais com 80% e 60% de medições nas vias urbanas, para os quais foram obtidos os erros médios apresentados nas Figuras 4.10 e 4.11. A queda de precisão da técnica com o aumento da porcentagem de usuários fora dessas vias segue o mesmo padrão dos ambientes urbanos e suburbanos. Ou seja, os *grids* irregulares ainda apresentam vantagem sobre os regulares para uma proporção de 80% de usuários em vias urbanas, porém já não apresentam mais tal vantagem para uma proporção de 60%. Vale ressaltar que essa vantagem é observada quando comparados dois *grids* com aproximadamente o mesmo número de células.

No ambiente rural, os erros médios encontrados para ambos os modelos de *grid* em um cenário com 100% de medições nas vias urbanas são, em geral, maiores do que os encontrados

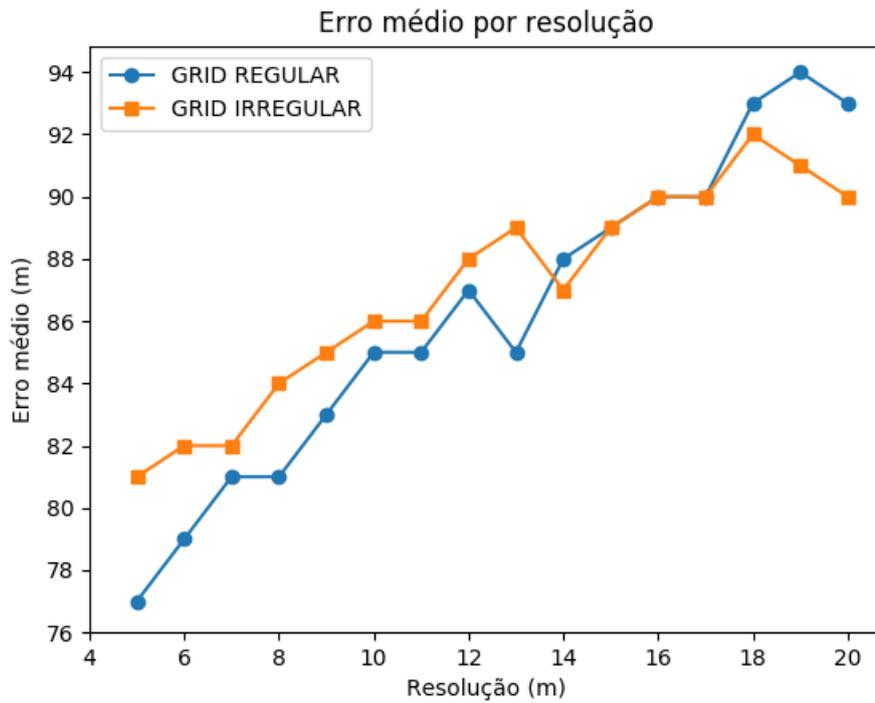


Figura 4.9 Erro médio por resolução num cenário rural com 100% de medições em vias urbanas.

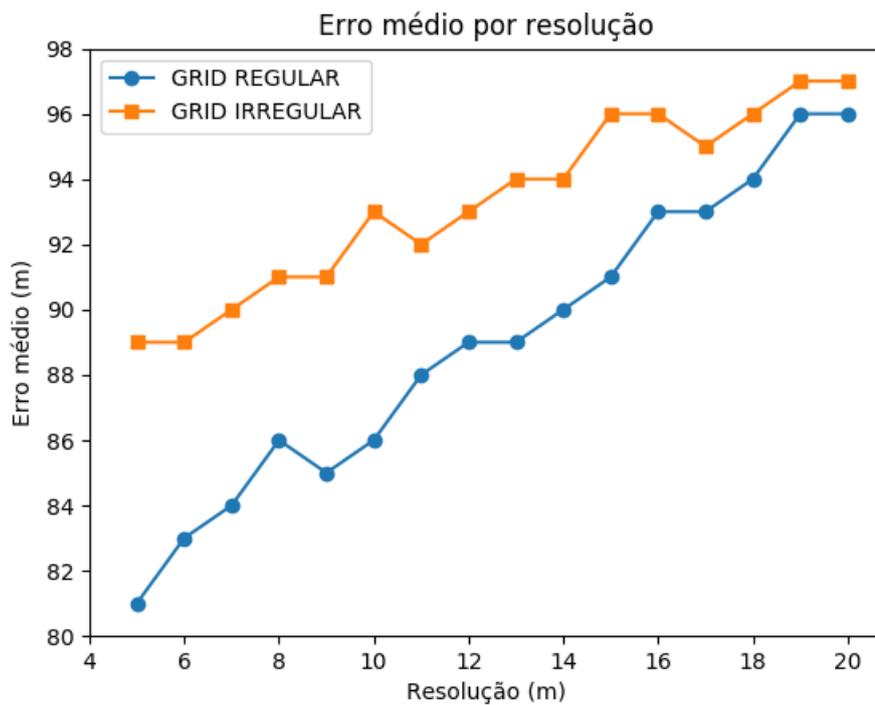


Figura 4.10 Erro médio por resolução num cenário rural com 80% de medições em vias urbanas.

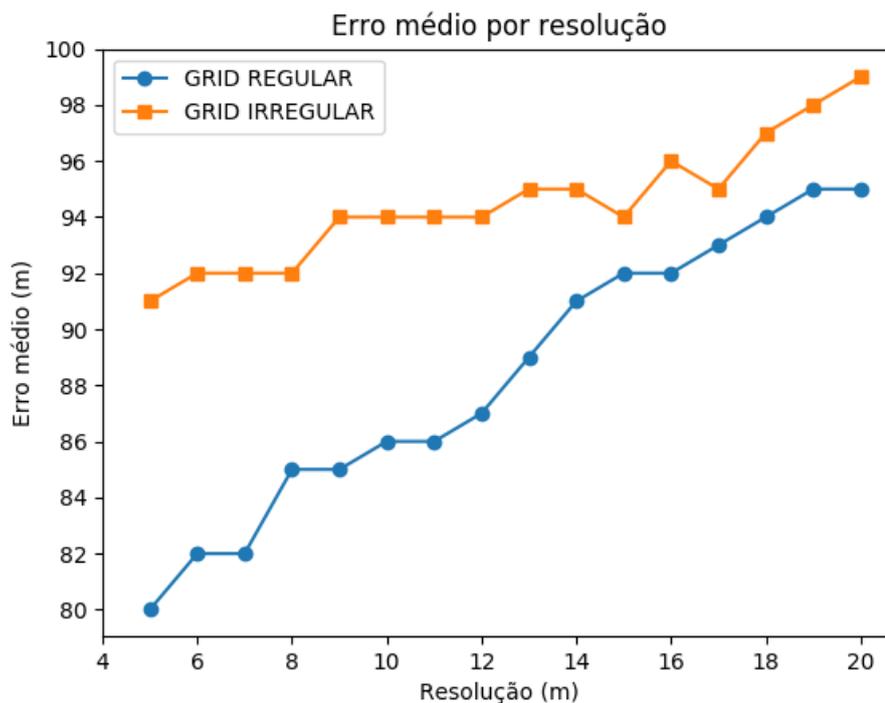


Figura 4.11 Erro médio por resolução num cenário rural com 60% de medições em vias urbanas.

para o ambiente urbano. No entanto, em um cenário com 60% de medições nas vias urbanas, os erros médios para algumas resoluções são maiores em ambientes urbanos do que em ambientes rurais. Logo, para os cenários rurais simulados neste trabalho, não foi encontrado um padrão de aumento ou redução na precisão da técnica de localização baseada em *fingerprint* de RF em relação aos cenários urbanos simulados. Portanto, com os resultados obtidos nesses experimentos, não foi possível chegar a uma conclusão sobre a relação entre os desempenhos dessa técnica em ambientes rurais e urbanos.

Conclusão e Trabalhos Futuros

5.1 Contribuições

Neste trabalho, foram estudadas técnicas de localização baseadas em *fingerprint* de RF, bem como diferentes modelos de *grid* utilizados pelas mesmas. A fim de avaliar um desses modelos de *grid*, proposto por Bittencourt, Urbano e Cunha [1], foram modelados sinais de RF utilizando o módulo LENA LTE do simulador ns-3. Essas simulações incluem estações base, dispositivos móveis e edifícios, além de modelar diferentes aspectos da propagação de sinais de RF, como perda de caminho, sombreamento e desvanecimento.

Com o objetivo de avaliar o desempenho da técnica em questão quando uma fração das medições de treino e de teste é realizada fora das vias urbanas, foram simulados cenários com diferentes porcentagens de medições nessas vias. Comparando-se *grids* regulares e irregulares com aproximadamente o mesmo número de células, os resultados mostram que os *grids* irregulares são mais precisos que os regulares em cenários cuja porcentagem de medições em vias urbanas corresponde a 80% ou mais. Utilizando o mesmo tipo de comparação, pode-se concluir que o modelo de *grid* irregular tem precisão menor ou igual à do modelo de *grid* regular em cenários nos quais essa porcentagem corresponde a 70% ou menos. Isso indica que o limiar para a porcentagem de medições em vias urbanas encontra-se entre 70% e 80%.

Outra avaliação realizada neste trabalho corresponde ao impacto da densidade de ocupação urbana no desempenho dessa técnica. Para essa análise, foram simulados três tipos de ambientes: urbanos, suburbanos e rurais. Os resultados encontrados para ambientes urbanos foram apresentados no parágrafo anterior. Nos ambientes suburbanos e rurais, foram encontradas as mesmas relações entre os erros médios dos modelos de *grid* regular e irregular encontradas nos ambientes urbanos, para todas as porcentagens de medições em vias urbanas avaliadas. Isso nos permite concluir que os ambientes suburbanos e rurais não afetam o desempenho dos *grids* irregulares sobre os regulares. Além disso, nos ambientes suburbanos, é observado um aumento no erro médio de ambos os modelos de *grid* em relação aos ambientes urbanos. Isso nos leva a concluir que a técnica de localização baseada em *fingerprint* de RF tem um desempenho pior nesses ambientes do que nos ambientes urbanos, independente do modelo de *grid* utilizado.

5.2 Trabalhos Futuros

Uma possível extensão deste trabalho consiste em avaliar separadamente o erro médio da estimativa de localização para medições *indoor* e *outdoor*. Esta avaliação possibilitaria estimar a precisão desta técnica em cada um dos dois tipos de ambientes, o que seria interessante para

aplicações voltadas para apenas um destes tipos.

Outra possível extensão consiste em avaliar cenários nos quais todas as medições de treino são localizadas nas vias urbanas, porém as medições de teste são localizadas nas vias urbanas e fora das mesmas. Essa avaliação corresponde a um cenário no qual o algoritmo é treinado com dados de um *drive-test* realizado apenas em ruas e avenidas, porém o mesmo seria utilizado por uma aplicação voltada para todos os pontos da região de cobertura.

Por fim, outro trabalho futuro consiste em adicionar o tempo de execução como um dos resultados da avaliação. Dessa forma, seria possível comparar a precisão de *grids* regulares e irregulares com aproximadamente o mesmo tempo de execução.

Referências Bibliográficas

- [1] G. P. Bittencourt, A. A. F. Urbano, and D. C. Cunha, “A proposal of an RF fingerprint-based outdoor localization technique using irregular grid maps,” *IEEE WCNC*, Apr. 2018.
- [2] Q. D. Vo and P. De, “A survey of fingerprint-based outdoor localization,” *IEEE Communications Surveys & Tutorials*, vol. 18, no. 1, pp. 491–506, 2016.
- [3] I. Guvenc and C.-C. Chong, “A survey on TOA based wireless localization and NLOS mitigation techniques,” *IEEE Communications Surveys & Tutorials*, vol. 11, no. 3, pp. 107–124, 2009.
- [4] Y. Han, H. Ma, and L. Zhang, “An efficient RF fingerprint positioning algorithm based on uneven grid layout,” *ICWMMN2015 Proceedings*, pp. 250–254, Nov. 2015.
- [5] *The LENA ns-3 LTE Module Documentation Release v8*. <http://networks.cttc.es/mobile-networks/software-tools/lena>, Jan. 2014.
- [6] R. Mondal, J. Turkka, T. Ristaniemi, and T. Henttonen, “Performance evaluation of MDT assisted LTE RF fingerprint framework,” *Seventh International Conference on Mobile Computing and Ubiquitous Networking (ICMU)*, pp. 33–37, 2014.
- [7] J.-H. Kim, K. S. Min, and W.-Y. Yeo, “A design of irregular grid map for large-scale Wi-Fi LAN fingerprint positioning systems,” *The Scientific World Journal*, Sept. 2014.
- [8] *ns-3 Tutorial Release ns-3.27*. <https://www.nsnam.org>, Oct. 2017.
- [9] 3GPP TR 36.814, *3rd Generation Partnership Project; Technical Specification Group Radio Access Network; Evolved Universal Terrestrial Radio Access (E-UTRA); Further advancements for E-UTRA physical layer aspects*, Mar. 2010. Release 9.
- [10] 3GPP R4-092042, *Simulation assumptions and parameters for FDD HeNB RF requirements*, May 2009. Alcatel-Lucent, picoChip Designs, Vodafone.
- [11] 3GPP TS 36.104, *Evolved Universal Terrestrial Radio Access (E-UTRA); Base Station (BS) radio transmission and reception*, July 2010. Version 9.4.0, Release 9.
- [12] H. Liu, H. Darabi, P. Banerjee, and J. Liu, “Survey of wireless indoor positioning techniques and systems,” *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 37, pp. 1067–1080, Nov. 2007.