

Caracterização da Latência de Borda sob Efeito de Mobilidade a Partir de Dados Reais

Luiz Felipe Cantanhede Cristino¹, Pedro Henrique Cruz Caminha¹

¹GTA/DEL-POLI/PEE-COPPE
Universidade Federal do Rio de Janeiro (UFRJ)

{cristino, cruz}@gta.ufrj.br

Resumo. *A Computação de Borda de Múltiplos Acessos (Multi-Access Edge Computing – MEC) é uma solução promissora de computação de borda que permite que dispositivos limitados ofereçam aos usuários aplicações de alto poder computacional utilizando baixa latência. No entanto, o sistema MEC possui limitações quanto à mobilidade de usuários, pois a latência aumenta à medida que os usuários se afastam dos servidores. Assim, este estudo coleta mais de 15.000 amostras de latência da comunicação entre dispositivos em uma mesma rede celular, avaliando o impacto da mobilidade em um possível sistema MEC. As observações mostram que a latência é fortemente impactada pela qualidade do sinal e pela distância entre dispositivos.*

Abstract. *Multi-Access Edge Computing (MEC) is a promising edge computing solution, enabling limited devices to provide applications with high computational power and low latency to users. However, MEC faces challenges, particularly in scenarios where user mobility can disrupt the proximity between users and the physical resources serving them. In a MEC system, low latency is achieved through the proximity of users to their serving physical resources. This study collects 15.000 latency measurements from communication between devices within a cellular network, assessing the impact of user mobility on a potential MEC system. We show that distance and signal quality play a major role in latency.*

1. Introdução

A proliferação de aplicações exigentes em termos de poder computacional impõe desafios a dispositivos com limitações computacionais e energéticas. Em um cenário típico, uma aplicação exigente executada em um dispositivo limitado utiliza a rede para realizar requisições a um poder computacional externo, mais robusto. Assim, as tarefas mais intensas são realizadas pelo poder computacional externo, num procedimento chamado de *offloading* [Islam et al. 2021]. Algumas aplicações podem ter baixa tolerância à latência, como jogos [Robaina and Fiorese 2023], por exemplo. Nesses casos, a Computação em Nuvem não é adequada para oferecer o poder computacional, sendo indicada a computação de borda (*Edge Computing*). A borda oferece recursos computacionais geograficamente e topologicamente próximos aos usuários, oferecendo baixa latência. A disponibilidade de recursos e a baixa latência são cruciais para oferecer uma Qualidade de Experiência (*Quality of Experience – QoE*) adequada [Ren et al. 2019].

O padrão *Multi-Access Edge Computing* (MEC), desenvolvido pelo Instituto Europeu de Padrões de Telecomunicações (*European Telecommunications Standards Institute* – ETSI), padroniza a oferta de recursos de computação de borda pelas operadoras de redes móveis (*Mobile Network Operators* – MNOs). Nesse contexto, os principais dispositivos limitados são móveis, denominados Equipamentos de Usuário (*User Equipment* – UE). Pelo padrão MEC, as MNOs posicionam servidores MEC em estações rádio base (*Base Stations* – BSs) ou em outros pontos da rede. Esses servidores MEC atendem a solicitações de aplicações nos UEs. No entanto, a implementação do MEC enfrenta desafios para otimizar a alocação de recursos em cenários com múltiplos UEs [Chen and Hao 2018]. A mobilidade dos UEs pode afastá-los dos recursos de borda originalmente designados, causando um aumento na latência e degradando a QoE. Portanto, estratégias de alocação devem considerar a mobilidade como fator fundamental para que UEs sejam atendidos por recursos suficientemente próximos [Luo et al. 2021]. Assim, faz-se necessária a criação de conjuntos de dados que relacionem atraso e mobilidade em uma mesma rede móvel.

Este trabalho apresenta um estudo da latência entre um possível servidor MEC e um UE a partir de dados reais. Para tal, são instanciados um UE servidor e um UE cliente dentro da rede de uma mesma operadora e são obtidas as latências nas requisições entre eles. Assim, é possível estimar o atraso entre esses dois dispositivos internos a uma rede MNO, assumindo que existe uma correlação entre a distância entre os dois dispositivos e o número de saltos na topologia da rede. Nos experimentos, o UE cliente se movimenta e o UE servidor permanece estático, a fim de capturar a influência da mobilidade na latência entre ambos. O protocolo IPv6 é utilizado para evitar interceptação por parte da MNO. Além dos valores de latência, os conjuntos de dados contêm informações sobre as condições do canal (e.g., potência do sinal), localização do UE (i.e., coordenadas do GPS), identificação de células conectadas (Cell ID) e outros valores. O software desenvolvido para a captura é disponibilizado. Finalmente, é estabelecida uma relação entre a latência e a distância entre os dois UEs. Assim, é possível sumarizar três contribuições principais:

- Produção de traços de latência entre dispositivos dentro de uma mesma MNO;
- Disponibilização do software para captura de dados;
- Caracterização da latência entre dois dispositivos na rede de uma mesma MNO, emulando um servidor MEC e um UE.

Diversos artigos na literatura abordam a mobilidade de usuários enquanto requisitam o tráfego de dados de uma rede móvel. Assim, são capturados traços de mobilidade que extraem as condições da rede enquanto o UE conectado à rede móvel se movimenta e troca informações com servidores externos à operadora [Raca et al. 2020, Xiao et al. 2014]. Também há trabalhos que caracterizam a latência entre UEs e borda de maneira teórica [Ko et al. 2018]. No entanto, não são encontrados na literatura traços públicos de tráfego dentro de uma mesma rede celular. Assim, este artigo cria o primeiro conjunto de dados público capaz de emular o tráfego entre um servidor MEC e um UE, permitindo a criação de estratégias eficientes de alocação e de *offloading*.

As seções subsequentes deste trabalho estão organizadas da seguinte forma: na Seção 2 são apresentados os trabalhos relacionados. Na Seção 3 é discutida a variação e o aumento da latência à medida que um dispositivo se afasta de um servidor de borda. A metodologia da coleta de dados e os dados a serem coletados são definidos na Seção 4.

Uma interpretação dos resultados gerados é apresentada na análise da Seção 5. Finalmente, a conclusão é apresentada na Seção 6.

2. Trabalhos Relacionados

Estudos têm produzido conjuntos de dados sobre a capacidade da rede frente a requisições contínuas enquanto UEs se movem. Esses estudos visam avaliar a latência e a variação no comportamento de conexões em redes móveis ou até mesmo testar possíveis topologias de MNOs. Raca *et al.* disponibilizam um conjunto de dados de produção para avaliar a variação das condições do canal em redes parcialmente 5G [Raca et al. 2020]. No estudo, o UE realiza o *download* de um arquivo ou assiste a um *streaming* de vídeo oriundos de um servidor na nuvem. Os experimentos são realizados de maneira estática e com a movimentação de um automóvel. A partir de um aplicativo público para mensurar informações da rede, é capturada a latência em ambos os cenários. Raca *et al* encontraram um aumento médio de 15 ms na latência quando o dispositivo estava em movimento. O mesmo grupo de pesquisa possui um trabalho similar, utilizando a mesma metodologia para produzir um conjunto de dados com redes 4G [Raca et al. 2018].

Xiao *et al.*, utilizando o protocolo TCP em redes 4G, propõem um conjunto de dados a partir de um aplicativo que se conecta com servidores locais para a análise do desempenho de uma rede celular chinesa sob os efeitos de alta velocidade. Para medir as condições da rede, o estudo emprega um trem-bala, um automóvel e um ponto estacionário, que têm velocidades de até 300 km/h, até 100 km/h e 0 km/h, respectivamente. Xiao *et al.* observam que no trem a latência medida chega a ser mais de 150 ms, enquanto no ponto estacionário e no automóvel, a latência é de 30 ms, em média [Xiao et al. 2014]. Baseando-se em conjuntos de dados disponibilizados publicamente, Bastos *et al.* realizam diversas simulações de topologias de redes MEC para melhor entregar vídeos sob demanda. É constatado que a instanciação de componentes em servidores MEC próximos aos usuários é capaz de melhorar a QoE, mas o movimento dos UEs não é levado em consideração para a análise desse comportamento [Bastos et al. 2019].

Também é possível encontrar na literatura análises teóricas sobre a latência entre dispositivo e borda. O trabalho de Ko *et al.*, por exemplo, realiza uma análise teórica da latência experimentada por usuários de MEC [Ko et al. 2018]. O trabalho mostra que a análise de latência é importante para o projeto de serviços MEC, mas não utiliza dados diretamente empíricos para obter seus resultados. Gonçalves *et al.* fazem uma análise profunda sobre o impacto da mobilidade em servidores de borda a partir de simulações, mostrando o quanto a predição de movimento pode reduzir a latência e o número de migrações, embora seu erro tenha um grande custo [Gonçalves et al. 2019].

É possível encontrar na literatura análises e contribuições a respeito da mobilidade sem considerar a existência do padrão MEC. Mehmeti e Porta analisam um conjunto de dados público de uma operadora móvel irlandesa. Os autores demonstram que, no caso de um UE em movimento dentro de um carro, o tempo que o UE fica conectado à mesma BS segue a distribuição de Pareto [Mehmeti and Porta 2021]. Bouchelaghem *et al.* produziram um conjunto de dados que contém comportamentos da rotina de um usuário detectado a cada cinco minutos em uma região de alcance de 500 m de cada BS. Este conjunto de dados foi gerado utilizando traços de mobilidade de um conjunto de dados público, permitindo predizer a localização de um usuário baseado em seu histórico de

movimento, o que auxilia na criação de estratégias para um sistema MEC melhor servir um usuário em seu trajeto [Bouchelaghem et al. 2022].

Os trabalhos apresentados trazem com excelência um melhor entendimento do impacto da mobilidade em redes móveis e a importância de seu estudo para o avanço da rede 5G, concentrando-se na comunicação entre um dispositivo local e um servidor em outra rede. No entanto, nenhum trabalho aborda a comunicação entre dispositivos que permanecem dentro da mesma rede celular, permitindo avaliar o impacto da mobilidade em possíveis sistemas MEC estabelecidos dentro de uma mesma MNO. A literatura carece de conjuntos de dados que capturem e que consigam relacionar essas grandezas em um mesmo experimento. Assim, este trabalho inova ao trazer publicamente dados práticos e relevantes para a criação de estratégias de alocação que valorizem a mobilidade dos usuários.

3. Latência entre Dispositivo e Borda

Ao realizar o *offloading* de alguma tarefa, uma aplicação executada por um UE realiza uma requisição a uma aplicação executada por um servidor para que execute tal tarefa e que retorne o resultado. Caso o UE esteja utilizando um serviço de nuvem, o servidor que executa a aplicação pode estar distante do UE - tanto geograficamente quanto topologicamente. Assim, existe uma latência na resposta à requisição. Essa latência é relativa ao tempo de transmissão da requisição ao servidor, do processamento de uma resposta e da transmissão da resposta. Caso o servidor de nuvem esteja distante, o tempo de transmissão da requisição e de transmissão da resposta pode ser elevado. A união desses atrasos pode inviabilizar aplicações sensíveis à latência.

3.1. Multi-Access Edge Computing

A fim de mitigar o problema da latência, a borda oferece poder computacional próximo aos UEs. Assim, o tempo de transmissão das requisições e das respostas é reduzido. Uma das formas possíveis de implementar a borda é o MEC. Nele, operadoras espalham servidores MEC (*MEC Hosts*) por suas redes de acesso. Esses servidores são gerenciados por um sistema MEC (*MEC System*). Os UEs, por sua vez, executam aplicações locais denominadas aplicações UE (*UE applications*), que interagem diretamente com os usuários finais. As aplicações MEC (*MEC applications*) são contrapartes das aplicações UE, sendo capazes de responder a requisições de *offloading* oriundas das aplicações UE.

Quando um UE deseja realizar o *offloading* de alguma tarefa, ele solicita ao sistema MEC que instancie a aplicação MEC correspondente. O sistema MEC verifica as permissões do usuário e da aplicação UE e, caso não haja impedimentos, escolhe um servidor MEC para executar a aplicação MEC e responder aos pedidos de *offloading*. A escolha de qual servidor MEC deve executar a aplicação MEC não é trivial e depende da política de alocação da MNO [Djigal et al. 2022]. A estratégia de alocação pode levar em consideração apenas a latência para o UE, mas também pode levar em consideração a utilização da rede, a disponibilidade de recursos, a mobilidade de usuários ou outros fatores [Cruz et al. 2022, Luo et al. 2021].

3.2. Latência entre dispositivo e borda sob mobilidade

A Figura 1¹ ilustra a situação na qual um UE executa uma aplicação UE enquanto está se movendo dentro de um automóvel. O sistema MEC instancia a aplicação MEC correspondente no servidor MEC mais próximo. Porém, o movimento do automóvel poderá fazer com que o UE se afaste do servidor MEC inicial e se aproxime mais de outro servidor MEC. Tal situação altera qual o servidor MEC ótimo do ponto de vista da latência.

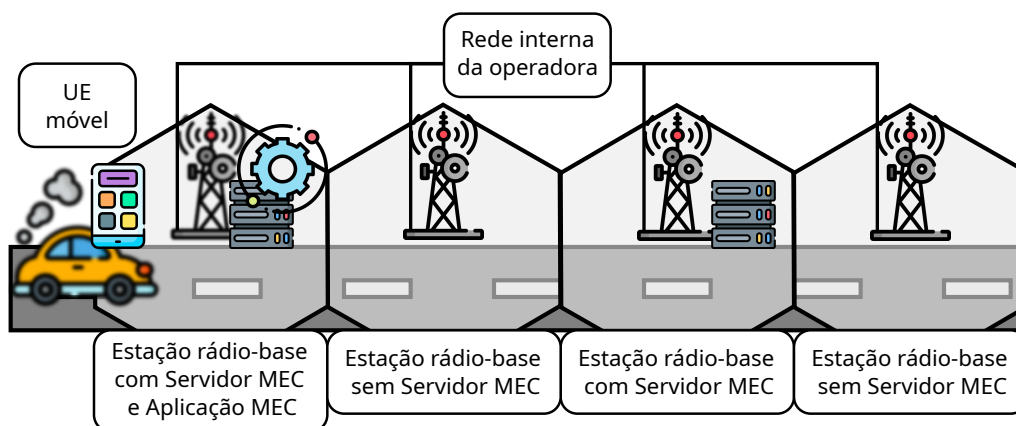


Figura 1. Movimentação de um UE com relação a um servidor MEC.

O atraso entre o UE e o servidor MEC depende também da topologia da rede interna da MNO. As topologias de MNOs são variadas, contendo suas redes de acesso e de núcleo, podendo abranger serviços variados. Frequentemente, essas topologias são tratadas pelas MNOs como segredos comerciais, a fim de evitar prejuízos pela concorrência ou por atacantes. No entanto, não é necessário conhecer precisamente a topologia das redes internas das MNOs para estimar o atraso entre dois dispositivos internos a uma rede. Porém, é necessário que um dos dispositivos faça uma requisição e o outro responda. Isso determina a importância de os experimentos serem realizados com dois UEs na mesma rede. Assim, mesmo que ambas as conexões sejam sem fio, é possível emular uma comunicação entre o servidor MEC e um usuário, embora ter acesso diretamente a um servidor MEC em uma BS fosse o cenário ideal para menos interferências. As redes das MNOs são protegidas por *firewalls* e NATs que impedem que dispositivos internos a essas redes recebam requisições. A Seção 4 detalha o procedimento de coleta de dados.

4. Procedimento de Coleta de Dados

A coleta experimental dos conjuntos de dados se dá a partir do desenvolvimento de duas aplicações para dispositivos Android [Google 2024]. Os códigos estão disponíveis em um repositório público². Os testes iniciais demonstram a impossibilidade de realização de *ping* ou requisições HTTP entre dispositivos da MNO escolhida utilizando o protocolo IPv4. Assim, para capturar os dados em uma mesma rede móvel, fez-se necessário a utilização do protocolo IPv6. Este trabalho apresenta o cliente WolfClient e o servidor WolfServer, duas aplicações que trocam requisições HTTP [Fielding et al. 2022]. As aplicações são instaladas em dois UEs conectados a uma mesma rede celular.

¹Figura com ícones de Freepik, de Flaticon.com.

²<https://github.com/GTA-UFRJ/WolfLatency>

A aplicação WolfClient realiza uma requisição GET a um servidor que esteja executando a aplicação WolfServer, a partir do endereço IPv6 do servidor. A aplicação WolfServer, então, retorna uma simples página como resposta. Ao receber a resposta do servidor, a aplicação WolfClient calcula a latência gerada a partir do tempo de ida e volta (*Round Trip Time* – RTT) dessa requisição. A aplicação WolfClient também coleta métricas a respeito da camada física, utilizando bibliotecas próprias do Android. Assim, as aplicações WolfClient e WolfServer emulam a latência observada por um UE ao se conectar a um servidor MEC. Após a obtenção das métricas, a aplicação WolfClient utiliza POST para enviar uma linha de informações à aplicação WolfServer, contendo os dados já gerados em sua requisição. A aplicação WolfServer cria um arquivo de backup com os dados coletados. Não são coletadas métricas de camada física e latência do POST, a fim de que essas requisições não interfiram nos resultados. Assim, a contribuição técnica desses softwares permitirá que novos cenários sejam estudados e uma precisão ainda maior das decisões de alocação sejam alcançadas em trabalhos futuros.

4.1. Dispositivos utilizados

Nos experimentos, são utilizados dois dispositivos UE. Um dos dispositivos executa o WolfServer e é mantido estático, a fim de simular um possível servidor MEC. O outro dispositivo executa o WolfClient e se move, se afastando para gerar dados referentes à latência e à variação das condições de conexão.

Nome do dispositivo	Modelo	Função
UE_1	Galaxy A54 5G SMA546E/DS	Servidor
UE_2	Galaxy Tab A9+ 5G SM-X216B	Cliente

Tabela 1. Dispositivos utilizados nos experimentos.

A Tabela 1 sumariza os dispositivos utilizados. Ambos possuem conectividade 5G e são equipados com cartões SIM com planos de dados válidos. Os planos são oferecidos por uma das três maiores MNOs do Brasil, segundo o ranking de fatias de mercado de operadoras da Anatel [Anatel 2021].

4.2. Métricas obtidas

De maneira geral, as aplicações trocam mensagens e coletam a latência da comunicação. Além da latência, também são coletadas outras informações, que estão sumarizadas na Tabela 2 e detalhadas a seguir. São coletadas informações referentes à geolocalização do UE (latitude e longitude), à data e instante quando a requisição é feita (Timestamp), ao meio de transporte utilizado (p. ex.: a pé, trem ou automóvel). Também ao tipo de rede móvel utilizada (5G, 4G, 3G ou 2G), à potência e o nível do sinal, medida em decibel-milliwatts (dbm), ao Código de Área de Rastreamento (*Tracking Area Code* – TAC) e Código de Localização de Área (*Localization Area Code* – LAC), ao Código Móvel do País (*Mobile Country Code* – MCC), ao Código da Operadora (*Mobile Network Code* – MNC), à identificação da célula (*Cell Identification* – Cell ID). Adicionalmente, são coletadas outras informações referentes à qualidade do sinal, como Qualidade do Sinal de Referência (*Reference Signal Received Quality* – RSRQ), Relação Sinal/Ruído (*Reference Signal Signal to Noise Ratio* – RSRNR) e também a identificação do canal (*New Radio - Absolute Radio Frequency Channel Number* – NRARFCN).

Campo	Definição	Campo	Definição
Transporte	Meio de transporte utilizado	MNC	Código da operadora do país
Timestamp	Data e hora da requisição	CellId	Identificação da célula conectada
Latência	Valor da latência em ms	TAC/LAC	Código de área da célula
Latitude	Coordenada de Latitude do GPS	Rede_Móvel	Conectado a 2G, 3G, 4G ou 5G
Longitude	Coordenada de Longitude do GPS	RSRQ	Qualidade do sinal recebido
Sinal_dbm	Potência do sinal obtido, em dbm	RSSNR	Quantia de ruído na comunicação
Sinal_nível	Tradução do Sinal_dbm em níveis	NRARFCN	Número do canal de rádio
MCC	Código móvel do país	-	-

Tabela 2. Grandezas capturadas pelos aplicativos.

Os dados de identificação reúnem informações para identificar a localização do UE no espaço através de suas coordenadas de latitude e longitude do Sistema de Posicionamento Global (GPS), que estabelece um parâmetro usado de maneira similar para identificar a localização na qual o acesso à rede está sendo feito. O código MCC endereça o país e o código MNC identifica a MNO. Assim sendo, ambos MCC e MNC são os mesmos ao longo de todos os experimentos, mas estão presentes para comparação com a realização de experimentos em outras áreas. Os códigos de localização de área (TAC/LAC), no entanto, servem para descrever qual das regiões dessa MNO o usuário está presente, permitindo endereçar diretamente a célula de estação rádio base que o usuário está conectado (CellId) [Han et al. 2019].

Os dados de qualidade do sinal reúnem dados para mensurar as condições de rede enfrentadas pelo usuário, como a informação referente ao indicador de potência de sinal recebido (Sinal_dbm), que depende do tipo de rede móvel experimentada pelo usuário. O Sinal_dbm pode significar Received Signal Strength Indicator (RSSI) para 2G, Received Signal Code Power (RSCP) para 3G e Reference Signal Received Power (RSRP) para 4G e 5G. Esses sinais são medidos em dbm e são transcritos para o mostrador de nível de sinal do usuário. Normalmente o mostrador é apresentado na forma de barras na interface e codificados em níveis de zero a quatro (Sinal_nível). De maneira a complementar à potência do sinal, há indicadores que dizem a qualidade do sinal que está sendo recebido (RSRQ) e o quanto de ruído ou está sendo recebido junto a esse sinal na comunicação entre as duas pontas (RSSNR) [Simpson and Sun 2018]. O número absoluto de canal de radiofrequência das redes 5G-NR (NRARFCN), no entanto, é utilizado para identificar um canal específico. Assim, a partir dos dados captados em diversas funcionalidades do dispositivo móvel, espera-se não apenas contribuir com o estudo da mobilidade, mas também disponibilizar informações relevantes para outros estudos referentes a redes móveis.

4.3. Captura de dados em movimento

Os experimentos são realizados no estado do Rio de Janeiro e são divididos em três etapas: um Trajeto Curto, um Trajeto Médio e um Trajeto Longo. Nos três experimentos, o servidor permanece estático em um espaço aberto ou fechado, enquanto o cliente se move em um automóvel. O mapa referente às posições dos UEs e aos caminhos percorridos pode ser encontrado na Figura 4.

A Tabela 3 resume as configurações de cada experimento. A coluna Duração exhibe a duração de cada experimento, no formato *horas:minutos:segundos*. A coluna Número de amostras indica quantas amostras são obtidas em cada experimento. A coluna

Latência média mostra a latência média e seu intervalo de confiança a 95%. Finalmente, as colunas Distância mínima e Distância máxima exibem, respectivamente, a menor e a maior distância registrada entre os UEs ao longo de cada experimento.

O Trajeto Curto é realizado com ambos os dispositivos posicionados na Cidade Universitária da Universidade Federal do Rio de Janeiro (UFRJ). O dispositivo UE_2 recebe a aplicação cliente e o dispositivo UE_1 recebe a aplicação servidor. A distância entre o cliente e o servidor varia entre 0,02 km e 2,01 km. O dispositivo UE_1 permaneceu em espaço aberto no estacionamento do Bloco H do Centro de Tecnologia (CT), enquanto o dispositivo UE_2 percorreu um trajeto interno à Ilha do Fundão. No experimento do Trajeto Médio, o dispositivo UE_1 recebe a aplicação servidor e o dispositivo UE_2 recebe a aplicação cliente. O servidor é mantido estático em um espaço aberto do campus da universidade em frente ao restaurante universitário CT1, enquanto o cliente se locomove até o bairro de Vila Isabel, na mesma região da cidade, viajando uma distância de 0,02 km até 6,41 km do servidor. O Trajeto Longo, no entanto, estabelece uma conexão entre cliente no dispositivo UE_2 que inicia seu trajeto no bairro de Vila Isabel e viaja até Petrópolis, enquanto o servidor no dispositivo UE_1 permanece estático em espaço fechado no bairro de Santa Cruz, iniciando o trajeto em uma distância de cerca de 39,52 km. O cliente se afasta até outra cidade, de forma que a distância entre ambos chega a 67,34 km. É esperado que esses experimentos sejam significativos para um grande número de cenários.

Para iniciar os experimentos, o endereço IPv6 do servidor e o meio de transporte utilizados são introduzidos manualmente na aplicação cliente. O cliente realiza uma requisição HTTP por segundo ao servidor, exceto nos casos nos quais a resposta do servidor demora mais do que um segundo. A cada requisição, o cliente atualiza suas informações de localização e de condições de rede. A cada requisição recebida do cliente, o servidor atualiza seu próprio conjunto de dados com a sua contínua localização e seus respectivos dados de rede, estabelecendo uma diferença do estado de conexão entre as duas pontas da comunicação. Os dados capturados pelo cliente também são armazenados em um arquivo de backup presente no servidor para evitar que os dados do cliente sejam perdidos. Os dados estão disponíveis em um repositório público².

Experimento	Duração	Número de amostras	Latência média	Distância mínima	Distância máxima
Trajeto Curto	00:12:54 h	769	449.5 ±16,0 ms	0,02 km	2,01 km
Trajeto Médio	00:19:37 h	1.157	323.9 ±40,9 ms	0,02 km	6,41 km
Trajeto Longo	02:53:04 h	13.521	465.3 ±17,8 ms	39,52 km	67,34 km

Tabela 3. Cenários dos experimentos realizados.

A captura presente nos mapas da Figura 4 apresenta duas regiões percorridas nas quais não são capturadas informações referentes a qualquer tipo de rede móvel. Essas áreas não possuem cobertura pela MNO para nenhum tipo de dispositivo móvel.

5. Análise do Conjunto de Dados

A Seção 4 apresenta a metodologia de captura de traços de rede sob mobilidade, culminando na formação de um conjunto de dados. Esta seção apresenta e avalia os resultados experimentais, buscando entender o compromisso entre proximidade dos UEs com recursos de borda que os servem e atraso na comunicação.

5.1. Compromisso entre latência e distância

A Figura 2 apresenta a medida da latência para as diferentes distâncias avaliadas nos experimentos. A latência é apresentada em escala logarítmica. A linha em preto é a reta ajustada pelo método dos mínimos quadrados. Seu coeficiente angular é 3.11 ms/km, o que mostra o crescimento da latência com a distância. Entretanto, o coeficiente de determinação R^2 da reta ajustada é de 0,0032, o que indica que há outros fatores influenciando a latência, além da distância. É possível observar que há atrasos que destoam ao longo do trajeto, embora uma série de atrasos se acumulem em cima de uma faixa que cresce lentamente com a distância. Dessa forma, são observados dois grupos de atrasos que apresentam comportamentos diferentes: um grupo contido e um grupo com atrasos destoantes. Esses valores destoantes se acentuam conforme os dispositivos se afastam. Em todos os experimentos são verificados atrasos de mais de 1 s, podendo chegar até a 15 s, valores que podem prejudicar a QoE para diversas aplicações, como jogos ou operação de equipamentos à distância. Assim como destacado em [Hassebo et al. 2018], estes valores inviabilizam aplicações críticas, pois muito se exige uma latência inferior a 10s.

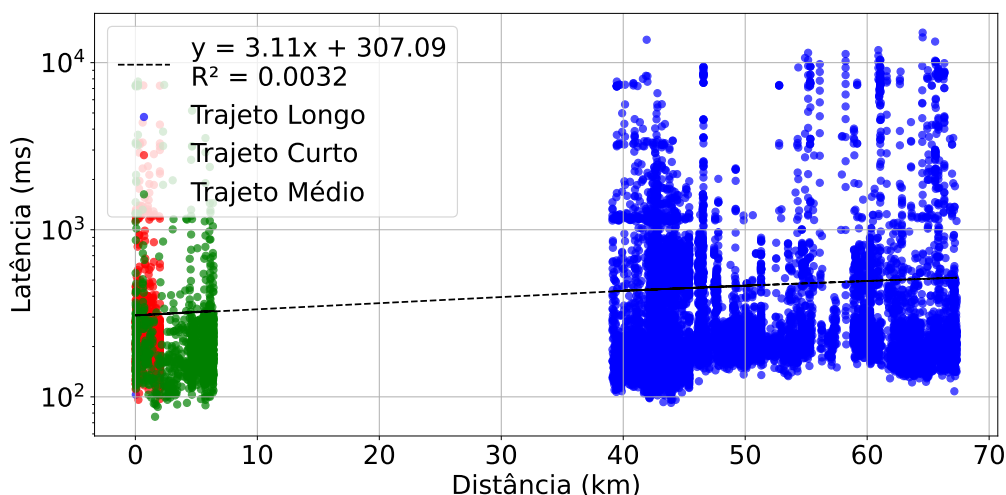


Figura 2. Latência em função da distância entre servidor e cliente.

O *jitter* é obtido utilizando o método descrito no RFC 4689 [Poretsky et al. 2006] e está exposto na Figura 3. A sua análise demonstra o impacto da mobilidade na transmissão de pacotes. Os altos valores de Jitter influenciam diretamente no atraso de voz, imagens quebradas e transferências de dados no geral.

Observando os mapas comparativos, ilustrados na Figura 4, torna-se perceptível que as cores mais claras responsáveis por marcar os locais com maiores valores de latência encontrados, aparecem com maior frequência à medida que o cliente se locomove no mapa e se distancia do servidor. A Figura 4d ilustra os pontos com latência maior do que 3000 ms, justamente, a fim de destacar os valores expressivos geograficamente.

Ao longo dos trajetos são observados um total de 335 células desde o início da conexão estabelecida pelo cliente, trocando entre as tecnologias de 3G, 4G e 5G. Contudo, foi observado que as conexões 5G presentes no trajeto percorrido pelo cliente tratava-se de conexões baseadas nas BSs da tecnologia 4G, utilizando a transmissão de sinal

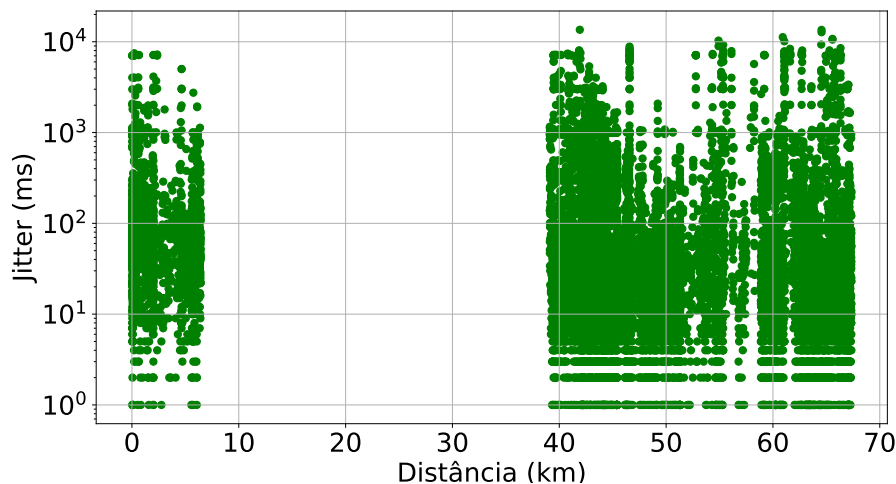


Figura 3. Jitter experimentado ao longo dos três trajetos.

consolidada da geração passada. Essa adaptação da rede 5G é conhecida como NR-NSA (New-Radio Non-Standalone) e leva o dispositivo a reconhecer a rede como 4G com sobreposição de dados da 5G [Narayanan et al. 2021].

5.2. Avaliação de latência mínima por janela

A Figura 5 apresenta a função de distribuição acumulada empírica (*Empirical Cumulative Distribution Function* – CDF empírica) das latências presentes em cada experimento. Em cada experimento há uma região de inflexão, situada entre 300 ms e 500 ms. Observa-se que o Trajeto Curto possui proporcionalmente latências maiores. Também pode-se observar, ainda que sutilmente, que o Trajeto Longo é o primeiro a apresentar valores maiores de latência para a quantidade de amostras coletadas, denotando a maior presença de valores destoantes. Uma hipótese possível é que a distribuição possua alguns atrasos menores, causados pela simples transmissão de dados e outros atrasos maiores, causados pela retransmissão de dados.

A Figura 2 apresenta valores de latência que destoam, sendo muitas vezes maiores do que valores obtidos a distâncias próximas. Essa observação está consoante com as observações relativas à Figura 5. Este trabalho considera que a menor latência em uma determinada janela de tempo representa, com alta probabilidade, a latência de propagação e encaminhamento entre os dois dispositivos, sem retransmissões. Para estimar a latência de propagação entre os dois dispositivos, são utilizadas as amostras da Figura 2. Assim, é obtida uma lista de tuplas (distância, latência), ordenadas por distância. As tuplas são agrupadas de cinco em cinco e a tupla de menor latência é selecionada. Assim, um subconjunto é criado, contendo as menores latências entre cada cinco amostras próximas. Por motivos de simplicidade, essa métrica é denominada de *latência mínima*.

A Figura 6 ilustra a latência mínima em função da distância. A Figura 6a ilustra o ajuste de reta para as latências mínimas de todas as amostras obtidas. A Figura 6b ilustra o ajuste de reta para as latências mínimas obtidas nos experimentos do Trajeto Médio e do Trajeto Longo. Ainda é possível observar grande variância entre as amostras. Também é possível observar que o experimento do Trajeto Curto possui valores menos condizentes

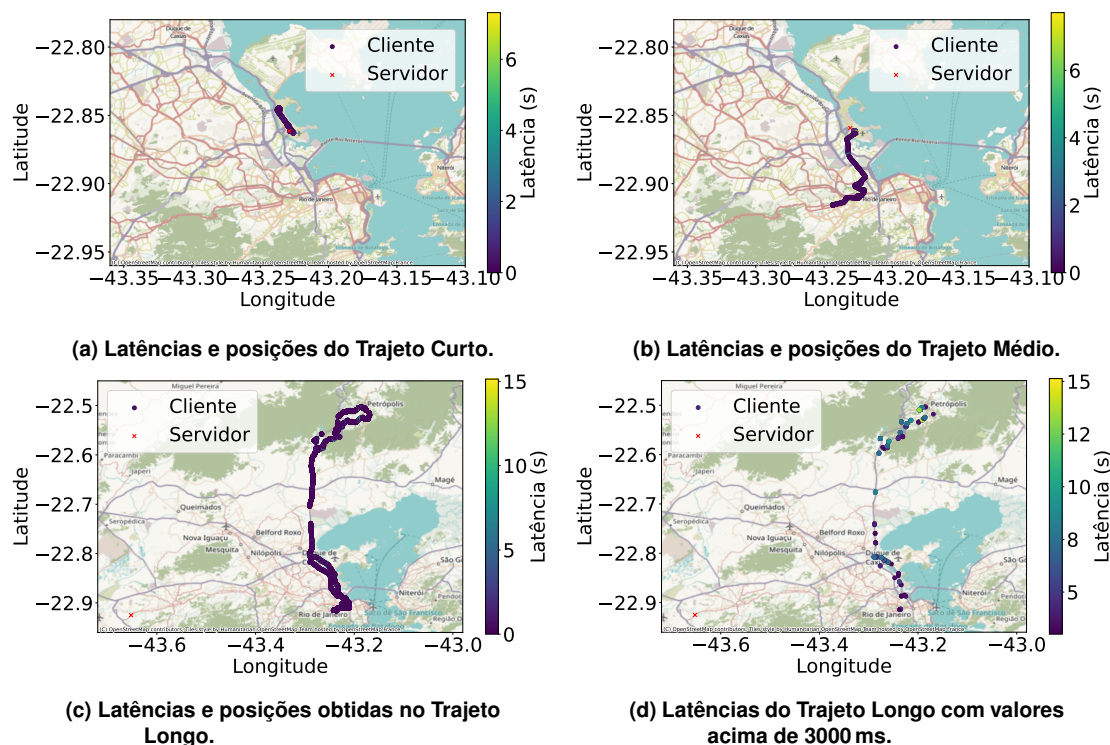


Figura 4. Mapas de latência para os diferentes experimentos.

com os valores encontrados nos outros dois experimentos.

5.3. Influência das condições de rede

Ao relacionar as condições de rede enfrentadas pelo cliente, é possível observar forte influência da potência de sinal na latência mínima e na latência destoante. A Figura 7a apresenta a proporção de amostras para cada nível de sinal. É possível notar que o sinal se manteve com maior qualidade durante a maior parte dos experimentos. A Figura 7b apresenta um diagrama de caixas para a latência em função de cada nível de sinal, com a latência em escala logarítmica. Ela mostra que o espalhamento dos valores de latência é maior quando o nível do sinal é menor. Pode-se concluir que a latência sofre forte influência do nível do sinal.

6. Conclusão

A Computação de Borda de Múltiplos Acessos (Multi-Access Edge Computing – MEC) é um padrão do Instituto Europeu de Padrões de Telecomunicações (*European Telecommunications Standards Institute* – ETSI) para que operadoras de redes móveis (*Mobile Network Operators* – MNOs) ofereçam recursos de computação de borda para seus usuários. O MEC oferece poder computacional de servidores MEC (*MEC hosts*) para que equipamentos de usuários (*User Equipments* – UEs) possam executar o *offloading* de tarefas. Adicionalmente, o MEC confia na proximidade entre os UEs e os servidores MEC para garantir baixa latência. Assim, a mobilidade de UEs pode prejudicar o potencial do MEC de oferecer baixas latências à medida que um UE móvel se afasta do servidor MEC.

Este trabalho buscou compreender a relação entre a distância e os atrasos entre UE e servidor MEC. Para tal, foram coletados dados experimentais a partir de dois UEs

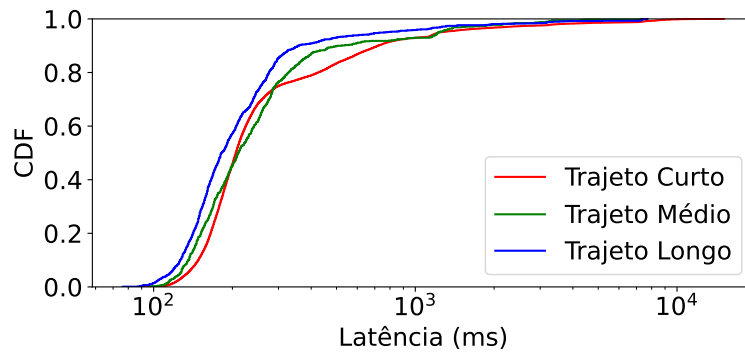


Figura 5. Função de probabilidade acumulada da latência entre os trajetos.

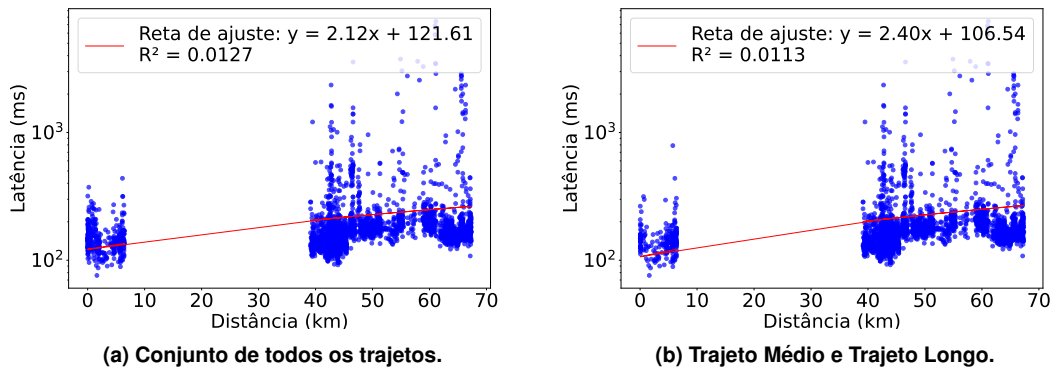


Figura 6. Latências mínimas por janelas de cinco amostras.

em movimento. Um dos UEs atua como servidor enquanto o outro atua como cliente. Dados foram coletados a partir da movimentação do cliente, com mais de 15.000 amostras de latência coletadas. Além das amostras de latência, também foram coletados dados relativos à qualidade do sinal. São apresentados resultados da relação entre a distância e a latência, assim como a influência do nível do sinal para a latência.

Como trabalhos futuros, espera-se utilizar o conjunto de dados gerado com os resultados dos traços de mobilidade em projetos de simulação, auxiliando na criação de estratégias de alocação a partir do comportamento móvel de UEs em uma rede real. Assim, devem ser iniciados novos trabalhos para estudar o impacto da mobilidade em cenários de um sistema MEC, emulando o comportamento esperado e as melhores estratégias de lidar com múltiplos usuários.

Referências

- Anatel (2021). Ranking das operadoras da agência nacional de telecomunicações. <https://informacoes.anatel.gov.br/paineis/acessos/ranking>. Acessada em: 28 de Janeiro de 2024.
- Bastos, I., Moraes, I., Thi-Mai-Trang, N., and Pujolle, G. (2019). Modelo e avaliação da recuperação de conteúdos através de funções de rede virtuais na arquitetura de computação na borda em redes móveis. In *Anais do XXXVII Simpósio Brasileiro de*

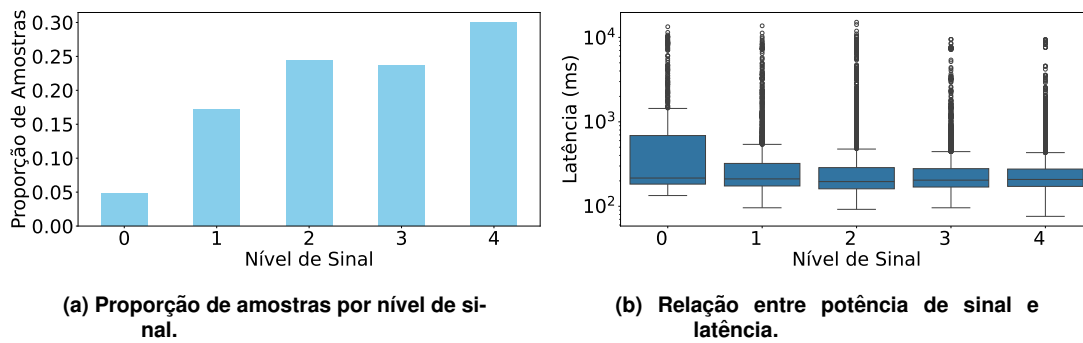


Figura 7. Relação entre a proporção e a latência gerada por nível de sinal.

Redes de Computadores e Sistemas Distribuídos, pages 749–762, Porto Alegre, RS, Brasil. SBC.

Bouchelaghem, S., Boudjelaba, H., Omar, M., and Amad, M. (2022). User mobility dataset for 5g networks based on gps geolocation. In *2022 IEEE 27th International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD)*, pages 128–133.

Chen, M. and Hao, Y. (2018). Task offloading for mobile edge computing in software defined ultra-dense network. *IEEE Journal on Selected Areas in Communications*, 36(3):587–597.

Cruz, P., Achir, N., and Viana, A. C. (2022). On the edge of the deployment: A survey on multi-access edge computing. *ACM Comput. Surv.*, 55(5).

Djigal, H., Xu, J., Liu, L., and Zhang, Y. (2022). Machine and deep learning for resource allocation in multi-access edge computing: A survey. *IEEE Communications Surveys & Tutorials*.

Fielding, R. T., Nottingham, M., and Reschke, J. (2022). Http semantics. *Internet Engineering Task Force, Request for Comments RFC*, 9110.

Gonçalves, D., Bittencourt, L. F., and Madeira, E. (2019). Análise da predição de mobilidade na migração de aplicações em computação em névoa. In *Anais do XXXVII Simpósio Brasileiro de Redes de Computadores e Sistemas Distribuídos*, pages 580–593, Porto Alegre, RS, Brasil. SBC.

Google (2024). Android api reference. <https://developer.android.com/reference>. Acessada em: 28 de Janeiro de 2024.

Han, S., Kang, T., and Seo, J. (2019). Smartphone application to estimate distances from lte base stations based on received signal strength measurements. In *2019 34th International Technical Conference on Circuits/Systems, Computers and Communications (ITC-CSCC)*, pages 1–3.

Hassebo, A., Obaidat, M., and Ali, M. A. (2018). Commercial 4g lte cellular networks for supporting emerging iot applications. In *2018 Advances in Science and Engineering Technology International Conferences (ASET)*, pages 1–6.

Islam, A., Debnath, A., Ghose, M., and Chakraborty, S. (2021). A survey on task offloading in multi-access edge computing. *Journal of Systems Architecture*, 118:102225.

- Ko, S.-W., Han, K., and Huang, K. (2018). Wireless networks for mobile edge computing: Spatial modeling and latency analysis. *IEEE Transactions on Wireless Communications*, 17(8):5225–5240.
- Luo, Q., Hu, S., Li, C., Li, G., and Shi, W. (2021). Resource scheduling in edge computing: A survey. *IEEE Communications Surveys & Tutorials*, 23(4):2131–2165.
- Mehmeti, F. and Porta, T. F. L. (2021). Analyzing a 5g dataset and modeling metrics of interest. In *2021 17th International Conference on Mobility, Sensing and Networking (MSN)*, pages 81–88.
- Narayanan, A., Zhang, X., Zhu, R., Hassan, A., Jin, S., Zhu, X., Zhang, X., Rybkin, D., Yang, Z., Mao, Z. M., Qian, F., and Zhang, Z.-L. (2021). A variegated look at 5g in the wild: performance, power, and qoe implications. In *Proceedings of the 2021 ACM SIGCOMM 2021 Conference, SIGCOMM '21*, page 610–625, New York, NY, USA. Association for Computing Machinery.
- Poretzky, S., Erramilli, S., Perser, J., and Khurana, S. (2006). Terminology for Benchmarking Network-layer Traffic Control Mechanisms. RFC 4689.
- Raca, D., Leahy, D., Sreenan, C. J., and Quinlan, J. J. (2020). Beyond throughput, the next generation: a 5g dataset with channel and context metrics. In *Proceedings of the 11th ACM Multimedia Systems Conference, MMSys '20*, page 303–308, New York, NY, USA. Association for Computing Machinery.
- Raca, D., Quinlan, J. J., Zahran, A. H., and Sreenan, C. J. (2018). Beyond throughput: A 4g lte dataset with channel and context metrics. In *Proceedings of the 9th ACM multimedia systems conference*, pages 460–465.
- Ren, J., Yu, G., He, Y., and Li, G. Y. (2019). Collaborative cloud and edge computing for latency minimization. *IEEE Transactions on Vehicular Technology*, 68(5):5031–5044.
- Robaina, G. P. and Fiorese, A. (2023). Gaming on the edge: Uma arquitetura de computação na borda para jogos em dispositivos móveis. In *Anais do XLI Simpósio Brasileiro de Redes de Computadores e Sistemas Distribuídos*, pages 574–587. SBC.
- Simpson, O. and Sun, Y. (2018). Lte rsrp, rsrq, rsnr and local topography profile data for rf propagation planning and network optimization in an urban propagation environment. *Data in Brief*, 21:1724–1737.
- Xiao, Q., Xu, K., Wang, D., Li, L., and Zhong, Y. (2014). Tcp performance over mobile networks in high-speed mobility scenarios. In *2014 IEEE 22nd International Conference on Network Protocols*, pages 281–286.