



UNIVERSIDADE FEDERAL DO AMAZONAS
INSTITUTO DE COMPUTAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM INFORMÁTICA

Impacto do consumo de energia no repasse de mensagens em redes oportunistas

GILMARA DO MONTE DOS SANTOS

Manaus - Amazonas

Dezembro de 2021

GILMARA DO MONTE DOS SANTOS

Impacto do consumo de energia no repasse de mensagens em redes oportunistas

Dissertação apresentada ao Programa de Pós-Graduação em Informática do Instituto de Computação da Universidade Federal do Amazonas como requisito parcial para obtenção do título de Mestre em Informática.

Orientador: Edjair de Souza Mota, Dr.-Ing.

Manaus - Amazonas

Dezembro de 2021

Ficha Catalográfica

Ficha catalográfica elaborada automaticamente de acordo com os dados fornecidos pelo(a) autor(a).

S237i Santos, Gilmara do Monte dos
Impacto do consumo de energia no repasse de mensagens em
redes oportunistas / Gilmara do Monte dos Santos . 2021
102 f.: il. color; 31 cm.

Orientador: Edjair de Souza Mota
Dissertação (Mestrado em Informática) - Universidade Federal do
Amazonas.

1. Redes oportunistas. 2. Economia de energia. 3. Clusterização.
4. Repasse de mensagens. I. Mota, Edjair de Souza. II.
Universidade Federal do Amazonas III. Título



FOLHA DE APROVAÇÃO

"Impacto do consumo de energia no repasse de mensagens em redes oportunistas"

GILMARA DO MONTE DOS SANTOS

Dissertação de Mestrado defendida e aprovada pela banca examinadora constituída pelos Professores:

Prof. Edjair de Souza Mota - PRESIDENTE

Prof. Alexandre Passito de Queiroz - MEMBRO EXTERNO

Prof. Celso Barbosa Carvalho - MEMBRO EXTERNO

Manaus, 01 de Dezembro de 2021

*Esta dissertação é dedicada à minha família meu apoio
em todos os momentos e a todos que contribuem com a educação.*

Agradecimentos

Agradeço a Deus por toda a oportunidade de aprimoramento.

Agradeço ao meu orientador, professor Edjair Mota, pelos esclarecimentos, ajuda, pelo incentivo que tornou possível a realização desse trabalho. Agradeço ao meu noivo Maurício pelo incentivo e compreensão. Agradeço principalmente a toda minha família, que sempre acreditou e torceu por mim.

Agradeço aos colegas do Grupo de Redes e Comunicação Multimídia (GRCM) pelo incentivo, pela companhia, pelos momentos descontraídos e pelas contribuições prestadas para melhoria do trabalho realizado, em especial ao Diogo Soares por contribuir com meu crescimento acadêmico. E ao meu amigo Ronaldo Nascimento pelo incentivo.

Resumo

Nas Redes Oportunistas (OppNets), a mobilidade e o contato entre os nós são usados para criar oportunidades de comunicação e trocar mensagens e informações. Uma premissa básica para um melhor desempenho dessas redes é a colaboração dos nós durante a comunicação. No entanto, devido a fatores de restrição de energia, os nós podem eventualmente deixar de colaborar com as trocas de mensagens. Neste trabalho, propomos um mecanismo de roteamento chamado eGPDMI para melhorar a probabilidade de entrega da mensagem enquanto otimiza o consumo de energia dos nós. Ao contrário de outros algoritmos propostos na literatura OppNets, o eGPDMI agrupa nós por nível de energia e interesses de nós usando técnicas de agrupamento. Nossa principal suposição é que reter mensagens nos nós com os níveis de energia mais altos pode melhorar o desempenho da rede, superando assim o problema de desconexão dos nós devido à falta de vontade de cooperar devido aos baixos valores de energia. Através da aplicação de questionários e experimentos de planejamento fatorial, caracterizamos os impactos dos níveis de energia em OppNets. Além disso, aplicamos a avaliação de desempenho do mecanismo eGPDMI em termos de eficácia usando a mobilidade de cenários do mundo real. Os resultados mostram que nosso mecanismo reduz efetivamente a degradação da probabilidade de entrega quando aumenta o nível de energia mínimo usados para os nós cooperarem com a comunicação.

Palavras-Chave: Redes Oportunistas, Economia de Energia, Clusterização, Repasse de Mensagens.

Abstract

In Opportunistic Networks (OppNets), mobility of and contact between nodes are used to create communication opportunities and exchange messages and information. A basic premise for a better performance of these networks is a collaboration of the nodes during communication. However, due to energy restriction factors, nodes may eventually fail to collaborate with message exchanges. In this work, we propose a routing mechanism called eGPDMI to improve message probability of delivery while optimizing nodes' energy consumption. Unlike other algorithms proposed in OppNets literature, eGPDMI groups nodes by energy level and nodes interests using clustering techniques. Our major assumption is that retaining messages in nodes with the highest energy levels can improve network performance, thus overcoming the problem of nodes' disconnection due to unwillingness to cooperate due to low energy values. Through questionnaire application and factorial design experiments, we characterize the impacts of energy levels in OppNets. Further, we apply performance evaluation of the eGPDMI mechanism in terms of effectiveness using mobility from real-world scenarios. The results show that our mechanism effectively reduces the degradation of the probability of delivery when the minimum energy level used for nodes to cooperate with communication increases.

Key Words - Opportunistic Networks, Energy-saving, Clustering, Forwarding of Messages.

Lista de Figuras

Figura 1: Fases da Operação TCP (adaptado de Oliveira et al., 2007)	20
Figura 2: Internet Interplanetária (CanalTech, 2018)	21
Figura 3: Comunidade Ribeirinha (Google Imagens, 2021)	22
Figura 4: Encaminhamento de Mensagens em Redes DTN	23
Figura 5: Camada de Agregação Bundle	24
Figura 6: Contato previsível de uma DTN em zona rural (<i>adaptado de Melo, 2011</i>)	26
Figura 7: Comunicação Oportunista (adaptado de Pelusi et al., 2006)	27
Figura 8: Clusters em duas dimensões (Jain e Dubes, 1988)	35
Figura 9: Algoritmo K-means (Santos et. al., 2015)	39
Figura 10: Execução do algoritmo K-means (adaptado de Linden, 2009)	40
Figura 11: Ilustração do algoritmo Bubble Rap (adaptado de Crowcroft et. al., 2011)	43
Figura 12: Componentes Int-Tree (adaptado de Xia et. al., 2016)	44
Figura 13: Funcionamento do Algoritmo ECF de encaminhamento baseado	47
Figura 14: Energia social gerada pela colisão de nós, distribuída e compartilhada (Li et al., 2017)	48
Figura 15: Fluxograma dos trabalhos relacionados por tópicos	50
Figura 16: Fluxograma Geral de processos do eGPDMI	56
Figura 17: Formação de Grupos por níveis de energia e Interesses Comuns	57
Figura 18: Impacto da densidade do fator cenário e do limite da bateria	66
Figura 19: Impacto da densidade do fator cenário e do limite da bateria	68
Figura 20: Overview of the ONE (Keränen et al., 2009)	70
Figura 21: Tela de Execução The ONE - GUI do Simulador (O autor, 2021)	70
Figura 22: Trace Reality – Grafo de formação de contatos (Neves, 2018)	73
Figura 23: Trace Infocom5 - Grafo de formação de contatos (O autor, 2021)	73

Lista de Gráficos

Gráfico 1: Probabilidade de entrega no cenário de mobilidade Infocom	81
Gráfico 2: Atraso médio de entrega no cenário de mobilidade Infocom	83
Gráfico 3: Sobrecarga no cenário de mobilidade Infocom	84
Gráfico 4: Distribuição de dados das médias de taxa de entrega - Infocom	84
Gráfico 5: Probabilidade de entrega no cenário de mobilidade Reality	87
Gráfico 6: Atraso médio de entrega no cenário de mobilidade Reality	88
Gráfico 7: Sobrecarga no cenário de mobilidade Reality.....	89
Gráfico 8:Distribuição de dados das médias de taxa de entrega – Reality	89

Lista de Tabelas

Tabela 1: Exemplos de Traces de Contato disponíveis.....	29
Tabela 2: Passo a passo K-means (Linden, 2009).....	40
Tabela 3: Resumo dos Trabalhos Relacionados.....	51
Tabela 4: Nível de energia e questionário OppNets.....	61
Tabela 5: Fatores de experimentos.....	63
Tabela 6: Interações entre fatores com métrica de resultado.....	63
Tabela 7: Avaliação fatorial epidemic.....	65
Tabela 8: Interações entre fatores com métrica de resultado.....	66
Tabela 9: Avaliação fatorial da proposta.....	67
Tabela 10: Parâmetros do conjunto de dados de mobilidade.....	76
Tabela 11: Parâmetros de Energia.....	77
Tabela 12: Formação dos Grupos de energia e interesse.....	77
Tabela 13: Síntese dos cálculos da análise de variância.....	79
Tabela 14: Análise de Variância (Taxas Entrega) – Anova – Infocom.....	85
Tabela 15: Teste de Tukey (Taxas Entrega) – diferença absoluta das médias – Infocom.....	86
Tabela 16: Análise de variância (Taxa Entrega) - Anova - Reality.....	90
Tabela 17: Teste de Tukey (Taxas Entrega) – diferença absoluta das médias - Reality.....	91

Sumário

AGRADECIMENTOS	IV
RESUMO	V
ABSTRACT	VI
LISTA DE FIGURAS	VII
LISTA DE GRÁFICOS.....	VIII
LISTA DE TABELAS.....	IX
INTRODUÇÃO	12
1.1. DEFINIÇÃO DO PROBLEMA.....	15
1.2. OBJETIVOS	16
1.3. MOTIVAÇÃO	17
1.4. CONTRIBUIÇÕES DESTE TRABALHO	17
1.5. ORGANIZAÇÃO DO TRABALHO	18
FUNDAMENTOS.....	19
2.1 REDES TOLERANTES A ATRASOS E DESCONEXÕES.....	19
2.2 ARQUITETURA DTN.....	23
2.3 TIPOS DE CONTATO EM UMA REDE DTN	25
2.4 MÉTRICAS UTILIZADAS EM UMA REDE DTN	26
2.5 REDES OPORTUNISTAS.....	26
2.6 MODELOS DE MOBILIDADE SINTÉTICOS.....	28
2.7 TRACES DE MOBILIDADE	29
2.8 ROTEAMENTO EM REDES OPORTUNISTAS	30
2.9 GRAFOS SOCIAIS E FORMAÇÃO DE GRUPOS EM OPPNETS.....	31
2.10 EGOÍSMO EM REDES OPORTUNISTAS	32
2.11 CONSUMO DE ENERGIA EM OPPNETS	33
2.12 FONTES DE CONSUMO DE ENERGIA	33
2.13 APRENDIZADO DE MÁQUINA	34
2.13.1 TÉCNICAS DE AGRUPAMENTO	34
2.13.2 ALGORITMO ESPERANÇA-MAXIMIZAÇÃO (EM).....	37
2.13.3 ALGORITMO K-MEANS.....	39
2.14 CONSIDERAÇÕES FINAIS.....	41
TRABALHOS CORRELATOS.....	42
3.1 ESTADO DA ARTE	42
3.2 SÍNTESE DOS RESULTADOS.....	50
3.3 CONSIDERAÇÕES FINAIS.....	54
MÉTODO.....	55
4.1 MECANISMO PROPOSTO (EGPDMI).....	55

4.1.1	ALGORITMO DE REPASSE DE MENSAGENS ENTRE OS NÓS	58
4.2	MÉTODO DE AVALIAÇÃO.....	59
4.3	CONSIDERAÇÕES FINAIS.....	60
PROJETO DE EXPERIMENTOS.....		61
5.1	PROCESSO DE PLANEJAMENTO DO PROJETO DE EXPERIMENTOS.....	61
5.2	CONSIDERAÇÕES FINAIS	68
IMPLEMENTAÇÃO E AVALIAÇÃO		69
6.1	THE OPPORTUNISTIC NETWORK ENVIRONMENT SIMULATOR (THE ONE)	69
6.2	MODIFICAÇÕES NO <i>THE ONE SIMULATOR</i>	70
6.3	CONJUNTO DE DADOS DE MOBILIDADE UTILIZADOS	72
6.4	ALGORITMO PROPHET	73
6.5	ALGORITMO GPDMI (<i>GROUPING AND DISSEMINATION OF MESSAGES BY INTERESTS</i>)	75
6.6	CARACTERIZAÇÃO DO CENÁRIO	76
6.7	MÉTODO ESTATÍSTICO	78
6.7.1	ANOVA.....	78
6.7.2	TESTE TUKEY	80
6.8	DISCUSSÃO DOS RESULTADOS	80
6.8.1	RESULTADOS DO CENÁRIO INFOCOM	81
6.8.1.1	TAXA DE ENTREGA DE MENSAGENS	81
6.8.1.2	ATRASO MÉDIO DE ENTREGA	82
6.8.1.3	SOBRECARGA.....	83
6.8.1.4	ANÁLISE DE VARIÂNCIA - PROBABILIDADE DE ENTREGA DE MENSAGENS	84
6.8.1.5	TESTE DE TUKEY – PROBABILIDADE DE ENTREGA (INFOCOM).....	85
6.8.2	RESULTADOS DO CENÁRIO REALITY	87
6.8.2.1	TAXA DE ENTREGA DE MENSAGENS	87
6.8.2.2	ATRASO MÉDIO DE ENTREGA	87
6.8.2.3	SOBRECARGA.....	88
6.8.2.4	ANÁLISE DE VARIÂNCIA - PROBABILIDADE DE ENTREGA DE MENSAGENS	89
6.8.2.5	TESTE DE TUKEY - PROBABILIDADE DE ENTREGA (REALITY).....	90
6.9	CONSIDERAÇÕES FINAIS	92
CONCLUSÕES		93
BIBLIOGRAFIA.....		95

Capítulo 1

Introdução

Apesar do desenvolvimento da Internet, tanto em termos de tecnologia e cobertura, ainda existem áreas remotas e cenários onde a conectividade é difícil de alcançar. Em situações em que há longos períodos de interrupção, longos atrasos e falta de conectividade entre diferentes nós de uma determinada rede, as tecnologias de comunicação sem fio podem oferecer soluções adequadas para estes ambientes desafiadores, onde não há infraestrutura de rede fixa ou onde existem altas taxas de atrasos.

Uma solução possível é o uso de redes tolerantes a atrasos e desconexões (*Delay and Disruption Tolerant Networks* - DTN). Outras terminologias usadas para nomear essas redes são redes oportunistas (*Opportunistic Networks* - OppNet) ou redes desafiadoras [Preshiya e Suriyakala, 2015]. As redes do tipo DTN ou OppNet são compostas por um conjunto de dispositivos móveis (nós) utilizados por indivíduos que se comunicam através de mensagens de forma assíncrona quando se encontram na mesma área de cobertura [Soares, 2018].

O suporte de comunicação nesses cenários depende principalmente do contato entre dispositivos móveis, definida como a proximidade entre eles suficiente para permitir a troca direta de mensagens. No instante desse evento, um dos dispositivos decide se aceita receber e carregar consigo mensagens destinadas a outros dispositivos [Oliveira, 2007].

Importante destacar que a mobilidade de nós influencia a quantidade e a qualidade das oportunidades de contato, sendo a principal premissa para o sucesso do roteamento e entrega de mensagens entre os membros da rede. Isso acontece porque o tempo de envio da mensagem pode variar desde minutos, horas ou dias, dependendo da densidade da população na rede e da distância entre os indivíduos comunicantes.

Além disso, essas redes são uma alternativa interessante para compartilhamento de dados entre nós próximos, pois suportam comunicação de baixo custo entre dispositivos independentemente de sua localização, podendo serem

utilizadas para múltiplas aplicações como compartilhamento de arquivos [D. M. Reis et al., 2018] ou áudio / vídeo, bate-papo, entre outros [Bridgefy, 2020].

A disseminação de mensagens neste tipo de rede ocorre segundo uma abordagem conhecida como armazena-carrega-encaminha. O nó remetente armazena a mensagem recebida de uma maneira persistente para tentar repassá-la futuramente quando houver novos contatos com outros nós intermediários, isto é, nós que não sejam o destino da mensagem. Entretanto, esse mecanismo exige espaço para armazenamento (*buffer*) e energia suficiente para a troca das mensagens.

Sabe-se, que o consumo de energia é um fator crítico na análise do desempenho de sistemas computacionais. Estudos feitos por Feeney e Nilsson (2001) e Wang et al. (2007), indicam que o subsistema de comunicação é o maior responsável pelo consumo de energia de dispositivos utilizados em comunicações sem fio. Além disso, o consumo de energia pode ter impactos diferentes na avaliação de desempenho em OppNets. Enquanto dispositivos móveis com bateria totalmente descarregada não podem ser usados para transmissão de dados, dispositivos móveis com baixo nível de bateria, não podem ser encorajados a cooperar em comunicações oportunísticas [B. Jedari et al., 2019].

Este consumo ocorre principalmente durante a transmissão de dados, seguido pelo processo de descoberta de novos dispositivos na rede. Tal fato é agravado em OppNets, pois estas redes são geralmente caracterizadas por cenários que demandam a utilização de dispositivos com pouca capacidade de armazenamento de energia em suas baterias e grande mobilidade [Azevedo, 2010].

Depois que um nó é desconectado, não é possível prever quando ele estará disponível novamente causando impacto no sucesso do roteamento e na taxa de entrega de mensagens. Faz-se necessário então maximizar a disponibilidade dos nós [Silva et al., 2012]. O que torna interessante pesquisas na área com o objetivo de oferecer possíveis soluções para problemas causados na operacionalidade das OppNets provenientes do impacto causado pelo consumo de energia durante eventos de encaminhamento de mensagens.

Na literatura, existem diversas abordagens e propostas de algoritmos que visam aumentar as taxas de entrega em OppNets, dentre as mais recentes está o uso de características sociais que tem ajudado a modelar novas estratégias de repasse de mensagens [Zhu et al., 2013]. O algoritmo *Friendship and Selfishness Forwarding (FSF)*, por exemplo, verifica o egoísmo dos nós durante as tentativas de repasse das

mensagens, considerando que mesmo tendo uma forte relação social com o nó de destino, o nó pode recusar o recebimento da mensagem por ser um indivíduo egoísta ou por possuir restrições de recursos naquele momento [Souza, 2019].

Outra abordagem consiste em um protocolo de disseminação de mensagens em OppNets que utiliza o interesse dos nós pela mensagem gerada pelo nó de origem, como critério para o repasse de mensagens através de grupos formados por interesses comuns [Neves, 2018]. O conhecimento de características sociais e comunitárias pode ajudar o protocolo na escolha dos melhores nós para retransmitir mensagens para seus respectivos destinos e, portanto, melhorar suas chances de entrega [Bulut et al., 2009].

Entretanto, o consumo de energia precisa ser levado em consideração no projeto de novos algoritmos de encaminhamento de mensagens em OppNets para que esses novos algoritmos possam aumentar a eficácia e a eficiência nas taxas de entrega de mensagens. Em um ambiente real pode-se não obter os resultados esperados das taxas de entrega de mensagem se o consumo de energia desses dispositivos móveis não for levado em consideração.

Para responder essas hipóteses aplicamos um questionário no campus da UFAM, com alunos de diversos cursos para avaliar o comportamento de usuários reais, no encaminhamento de mensagens em OppNets quando estão usando dispositivos com níveis de energia da bateria críticos. O detalhamento sobre o resultado do questionário será apresentado no Capítulo 5.

Este trabalho propõe um mecanismo para otimizar o encaminhamento de mensagens, adicionando o consumo de energia dos nós como fator para tomada de decisão. Isto faz com que os nós evitem o repasse de mensagens através de nós que podem se tornar egoístas ou inoperantes, devido ao pouco recurso de energia disponível. Optou-se por técnicas de aprendizagem de máquina para a classificação e agrupamento de nós conforme seu nível de energia, alinhando o consumo de energia dos nós com fatores sociais em sua classificação.

Neste contexto, foi idealizado, projetado e implementado um mecanismo de decisão de encaminhamento, que leva em consideração dispositivos com níveis de energia semelhantes, agrupados de tal forma que atendam aos critérios de repasse utilizando níveis de energia considerados aceitáveis para encaminhamento da mensagem, assim evitando o consumo de recursos dos nós e o comprometimento da

rede. Tendo como premissa que o padrão de uso da bateria depende do comportamento do seu usuário.

O novo mecanismo chamado ϵ GPDMI foi desenvolvido customizando o protocolo oportunista GPDMI [Neves, 2018], que possui como base social de roteamento, grupos formados pelo interesse social dos nós. Essa hipótese é baseada no fato de que em OppNets, os nós tendem a cooperar na comunicação se o conteúdo da mensagem é de interesse deles [F. Xia et al., 2016].

Os mecanismos de formação de grupos de interesse foram preservados no protocolo sendo então adicionados novos contextos para decisão de roteamento, como, agrupamentos definidos através da classificação por níveis de energia, assim, modificando a estrutura geral do algoritmo e implementando novas modificações que formaram o ϵ GPDMI. As motivações acerca desta percepção são profundamente estudadas e a proposta é comparada a outros mecanismos de roteamento apresentados na literatura para avaliação de desempenho.

1.1. Definição do Problema

Em OppNets, a capacidade de energia dos nós é um recurso crítico. Os dispositivos que compõem a rede, são equipados com baterias que tem duração limitada. Isso impacta na conectividade entre os nós dessa rede e, conseqüentemente, nas taxas de entrega de mensagens e no tempo médio entre o envio e o recebimento das mensagens na rede.

Naturalmente, os nós móveis deste tipo de rede possuem restrições devido ao consumo de energia. A simples busca por nós vizinhos consome bastante recurso. Isso implica que cada nó precisa deslocar-se mais até entrar na área de cobertura de um outro nó móvel também em deslocamento [Yoon e Haas, 2012].

Também consome energia o repasse de uma mensagem para um outro nó. Redes oportunistas dependem da disponibilidade de seus nós para executar o roteamento de mensagens de um nó de origem para um nó de destino e isto consiste em utilizar tal recurso da forma mais eficiente possível [Azevedo, 2010]

Então, o problema consiste em nos perguntarmos se é possível reduzir as desconexões dos nós e aumentar a taxa de entrega de mensagens em OppNets móveis através da otimização do consumo de energia dos nós durante o repasse de mensagens, a fim de que os nós da rede não se desconectem devido ao esgotamento

de energia de suas baterias ou se tornem nós naturalmente egoístas por causa de recursos críticos de energia.

Uma proposta promissora é agrupar nós móveis conforme os níveis de energia de suas baterias. Repassar as mensagens aos nós de acordo com o grupo de classificação do seu nível de energia, possuindo como mecanismo de decisão os níveis considerados bons para repasse.

Pode ser uma estratégia interessante a fim de reduzir a desconexão de nós da rede, devido ao esgotamento da energia de suas baterias e otimizar o encaminhamento de mensagens. Aliar essa proposta ao interesse social dos nós por tópicos de mensagens que é explorado no algoritmo GPDMI [Neves, 2018] pode potencializar as entregas de mensagens na rede além de equilibrar o consumo de energia dos nós. O mecanismo considera o comportamento dos usuários móveis com o consumo de energia de suas baterias.

1.2. Objetivos

Esta dissertação tem como objetivo geral contribuir para aperfeiçoar o desempenho e a quantidade de mensagens entregues em cenários de redes oportunistas (OppNets), utilizando a otimização do consumo de energia durante o encaminhamento de mensagens.

A partir desse objetivo, considerando as lacunas identificadas na literatura sobre mecanismos para economia de consumo de energia em redes oportunistas móveis (OppNets), o presente trabalho tem como objetivos específicos:

1. Projetar um mecanismo para o roteamento de mensagens que visem o gerenciamento de recursos dos nós da rede, com foco principal no problema do consumo de energia dos nós em OppNets móveis.
2. Modelar a rede adicionando níveis de energia de bateria para restrições de participação dos nós.
3. Aperfeiçoar o algoritmo de roteamento oportunista GPDMI [Neves, 2018], associando o agrupamento social dos nós por interesses comuns com a formação de grupos por níveis de energia.
4. Verificar a eficiência do algoritmo proposto nas condições de recurso de energia conforme limites de carga de energia diferentes.

5. Analisar o algoritmo proposto comparando-o com algoritmos existentes na literatura.

1.3. Motivação

O gerenciamento e a forma de utilização de recursos dos nós no momento do repasse de mensagens são grandes alvos na busca por maneiras cada vez mais eficientes de potencializar o repasse de mensagens e o aumento na eficiência de entrega de mensagens, tornando-se um dos principais desafios neste tipo de rede.

Embora existam pesquisas focadas no gerenciamento de recursos dos dispositivos em OppNets, elas estão, em sua maioria, dissociadas dos algoritmos para roteamento de mensagens, principalmente quando se trata de recursos de energia. É necessário pensar em cenários reais, onde ambas são importantes. Neste contexto, podemos encontrar aspectos importantes para uma análise mais completa do desempenho da rede em cenários mais realistas.

Além disso, é possível contribuir significativamente para a utilização desta tecnologia no mundo real, aumentando o interesse de pesquisadores sobre a convergência de soluções voltadas cada vez mais para atender e focar na análise de desempenho de demandas e questões mais reais. Buscando considerar tais questões, o foco deste trabalho é associar os aspectos de consumo de recursos de energia dos dispositivos móveis em OppNets.

1.4. Contribuições deste Trabalho

As contribuições deste trabalho são:

1. Um mecanismo de roteamento que utilize a formação de grupos de nós formados por diferentes níveis de energia (considerando o comportamento do usuário com sua bateria) e interesses sociais para serem usados como fator de decisão durante o repasse de mensagens, potencializando o repasse através de interesses comuns e evitando nós naturalmente egoístas devido o esgotamento do recurso crítico de energia dos dispositivos.
2. Apresentar uma análise do desempenho do algoritmo proposto com outras abordagens presentes na literatura.

3. Indicar novas abordagens e campos de pesquisa com o objetivo de melhorar a disseminação de mensagens e a maximização energética dos nós em redes oportunistas.
4. Publicação do artigo “*An Energy-Saving Forwarding Mechanism Based on Clustering for Opportunistic Networks*”, no *journal Sensors* em novembro de 2021¹.

1.5. Organização do Trabalho

Este documento está organizado da seguinte forma. No capítulo 2 são apresentados os fundamentos, conceitos e definições teóricas necessárias para o entendimento dos métodos adotados. No capítulo 3 é apresentado uma síntese dos trabalhos relacionados. O mecanismo proposto é descrito no capítulo 4. O capítulo 5 apresenta um projeto de experimentos, enquanto o capítulo 6 detém-se na implementação e na avaliação dos resultados do experimento. Por fim, no capítulo 7 são discutidas as conclusões e as possíveis extensões futuras deste trabalho.

¹ <https://doi.org/10.3390/s21227427>

Fundamentos

Neste capítulo, são apresentados os conceitos necessários para o desenvolvimento do trabalho. Tais conceitos incluem uma visão geral sobre DTN e OppNets, aspectos fundamentais, utilização de características sociais em projeto de protocolos, consumo de energia e técnicas de aprendizado de máquina aplicadas neste trabalho.

2.1 Redes Tolerantes a Atrasos e Desconexões

As tecnologias de rede tradicionais têm como base de sua arquitetura a pilha de protocolos TCP/IP e baseiam-se em premissas básicas para o seu funcionamento tais como a existência de uma rota fim-a-fim, baixa perda de pacotes e baixa latência.

A conectividade proporcionada por essa arquitetura permitiu a expansão da Internet de maneira global, contudo determinados ambientes de rede não possuem as características necessárias para operações de comunicação fim-a-fim, consequentemente a rede terá falhas, atrasos e desconexões frequentes que causam a baixa qualidade das conexões, tornando-a assim inadequada e pouco robusta.

Aspectos que são comuns a todas as DTNs dificultam as operações do protocolo TCP (*Transmission Control Protocol*), um protocolo orientado à conexão que busca garantir a confiabilidade na entrega de dados fim-a-fim. Seu funcionamento é baseado em três fases, a saber: estabelecimento de conexão, transferência de dados e desconexão [Oliveira et al., 2007]. Estas fases são exemplificadas na Figura 1.

Quando um nó transmissor envia dados para um nó receptor e não ocorrem falhas, o nó receptor envia um reconhecimento positivo para o transmissor (ACK, do Inglês *acknowledgment*), indicando sucesso na transmissão [Oliveira et al., 2007]. Quando há falhas na transmissão, a mensagem é enviada novamente. O mesmo ocorre quando ou a mensagem ou o reconhecimento demoram mais que o tempo definido previamente para alcançarem seus destinos. A retransmissão gera maior latência na comunicação. Na Internet convencional, há um limite máximo para essa latência e, quando ultrapassado, a transmissão é abortada.

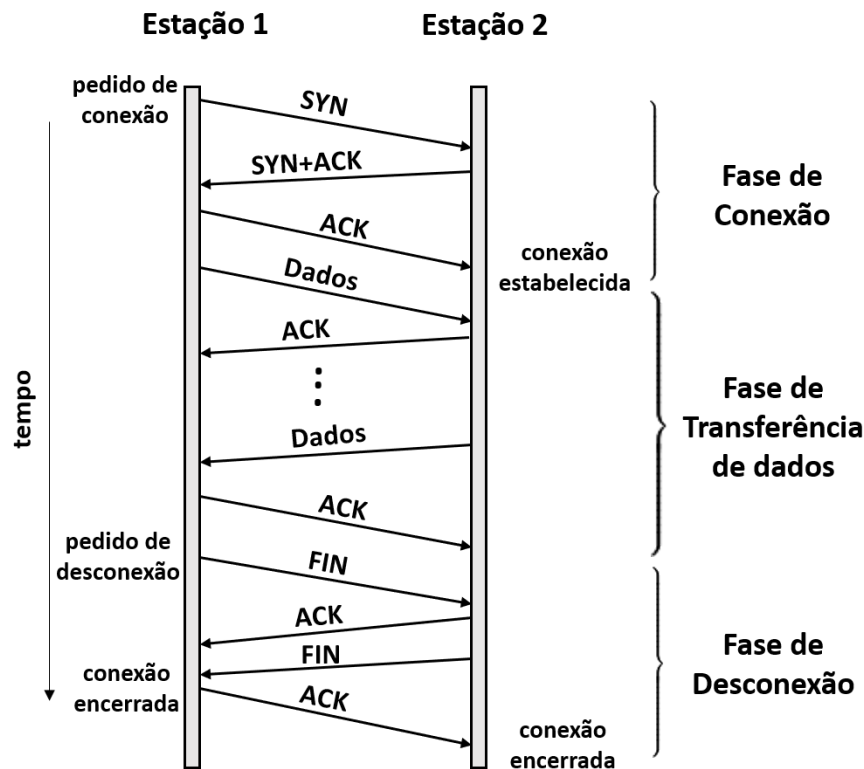


Figura 1: Fases da Operação TCP (adaptado de Oliveira et al., 2007)

Dentre as situações em que os protocolos convencionais da Internet falham, estão as redes Ad hoc móveis (MANETs, do Inglês Mobile Ad Hoc Networks), redes de sensores sem fio, comunicações espaciais e redes que atendem regiões carentes de infraestrutura de energia elétrica e telecomunicações [Demmer et al., 2004].

A seguir, temos exemplos de ambientes reais onde esses protocolos convencionais de Internet falham e onde não existe infraestrutura mínima para utilização. A Figura 2 é uma foto retirada do site CanalTech² sobre um modelo de comunicação interplanetária da NASA usando DTN.

A comunicação interplanetária pode permitir que naves e satélites se comuniquem com a Terra propagando informações do espaço. Este modelo de comunicação apresenta grandes desafios por causa da distância extrema entre os planetas, como os dados são transmitidos através de milhões de milhas de distância, sofrendo grandes atrasos e interrupções, podem ocorrer perdas de dados significativos.

Para utilizar DTN neste cenário, é importante que a rede consiga ser capaz de resolver enormes atrasos e pacotes de dados perdidos, devido aos ambientes

² <https://canaltech.com.br/espaco/nasa-lanca-as-bases-de-uma-internet-interplanetaria-118034/>

espaciais de alta latência. De fato, a arquitetura DTN é capaz de atender a estas questões, pois ele armazena automaticamente os dados transmitidos em cada nó até que o próximo fique disponível, podendo transmitir dados para um satélite em órbita, a uma distância menor e este repassando para outros satélites planetários até a informação chegar na terra.

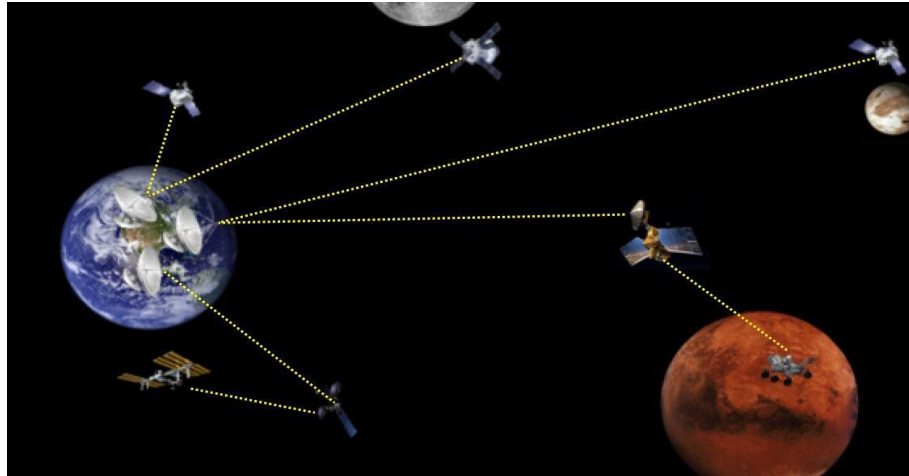


Figura 2: Internet Interplanetária (CanalTech, 2018)

A Figura 3 mostra a foto de uma comunidade ribeirinha. Segundo notícia publicada em 2020, no Amazonas, 97% das comunidades ribeirinhas não tem acesso à Internet [Portal Amazônia, 2020].

Isso ocorre por diversos fatores, como a geolocalização dos municípios e comunidades ribeirinhas, que geralmente são áreas bastante extensas com difícil acesso e vasta extensão territorial entre as comunidades separadas pelos rios e florestas.

As grandes distâncias geram um alto custo logístico, além dos fenômenos naturais como seca e cheia nas regiões, o que gera impactos, pois muitas comunidades não têm rede elétrica e as que possui, também sofrem com problemas de telecomunicação como lentidão e falhas constantes na comunicação.

Levar tecnologia de comunicação para essas áreas remotas se torna um grande desafio devido a exigência de uma infraestrutura mínima que seja capaz de atender aos requisitos dos protocolos de internet que exigem um caminho fim a fim para comunicação.



Figura 3: Comunidade Ribeirinha (Google Imagens³, 2021)

As Redes Tolerantes a Atrasos e Desconexões (*Disruption and Delay Tolerant Networks - DTNs*) se propõem a mitigar características como longos atrasos e frequentes desconexões que impactam diretamente o funcionamento dos protocolos que atuam na pilha TCP/IP [Oliveira et al, 2007].

Para superar os problemas associados a conectividade intermitente, as redes DTN usam o princípio armazena-carrega-e-envia (*Store-Carry-and-Forward*) para o encaminhamento de mensagens [Neto, 2012]. Este princípio, diferentemente das redes baseadas em comutação de pacotes, consiste em agrupar blocos de dados, gerados pela aplicação do usuário, em uma única mensagem agregada (*bundle*) que é movida entre os dispositivos de armazenamento dos nós [Warthman, 2003]. A técnica é ilustrada na Figura 4.

O nó de origem (*source node*) possui armazenado em sua memória persistente uma mensagem agregada endereçada ao nó de destino (*destination node*). Ao mover-se, carrega consigo a mensagem agregada e, quando surge uma oportunidade de contato com um nó intermediário, o nó de origem encaminha a mensagem agregada para o nó intermediário. O nó intermediário, por sua vez, armazena em sua memória persistente a mensagem agregada. Este procedimento se repete com outros nós intermediários até que a mensagem agregada seja entregue ao nó de destino.

De acordo ainda com a arquitetura implementada na camada de agregação (*bundle layer*) de redes DTN, a transmissão de dados e mensagens acontece através de um mecanismo denominado transferência de custódia (TC), onde a camada de agregação utiliza um temporizador para transmitir e retransmitir mensagens.

³ <https://images.app.goo.gl/r7W5S8Mq8iksiBrRA>

Observando o exemplo abaixo, quando o nó de origem (source node) envia a mensagem para o nó intermediário **A**, ele solicita uma transferência de custódia (TC). Se a camada de agregação do nó **A** aceitar o pedido, uma resposta é retornada através do ACK para o nó de origem. Isso tudo acontece dentro de um tempo determinado pelo temporizador. Uma vez que este tempo acaba o nó de origem tenta retransmitir novamente, caso o nó intermediário **A** não tenha aceitado as mensagens.

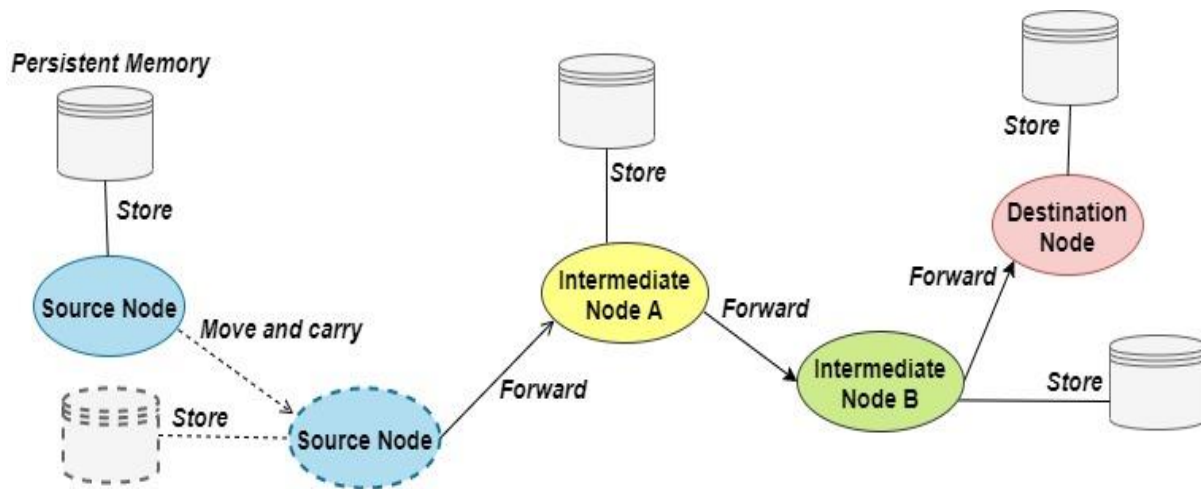


Figura 4: Encaminhamento de Mensagens em Redes DTN
(adaptado de Taylor e Francis Group, 2019).

Neste caso, em situações sem conexão ou com atraso, o armazenamento persistente possibilita o armazenamento integral da mensagem que se deseja enviar, ou seja, primeiramente o nó recebe totalmente a mensagem, armazena persistentemente para depois enviar para o destino ou para nós intermediários.

2.2 Arquitetura DTN

A arquitetura DTN é então definida no documento RFC 4838, que descreve como um conjunto de nós se organiza para armazenar e encaminhar mensagens em ambientes sujeitos a atrasos longos ou com frequentes desconexões [Cerf et al., 2007].

Este mecanismo torna-se possível em DTN com a criação de uma sobrecamada na pilha TCP/IP, denominada de Camada de Agregação (*Bundle Layer*) [Souza, 2013].

Esta sobrecamada foi colocada estrategicamente entre a camada de aplicação e a camada de transporte, com o objetivo de manter a interoperabilidade de uma DTN com qualquer tipo de rede que utilize a pilha de protocolos TCP/IP, conforme ilustra a Figura 5.

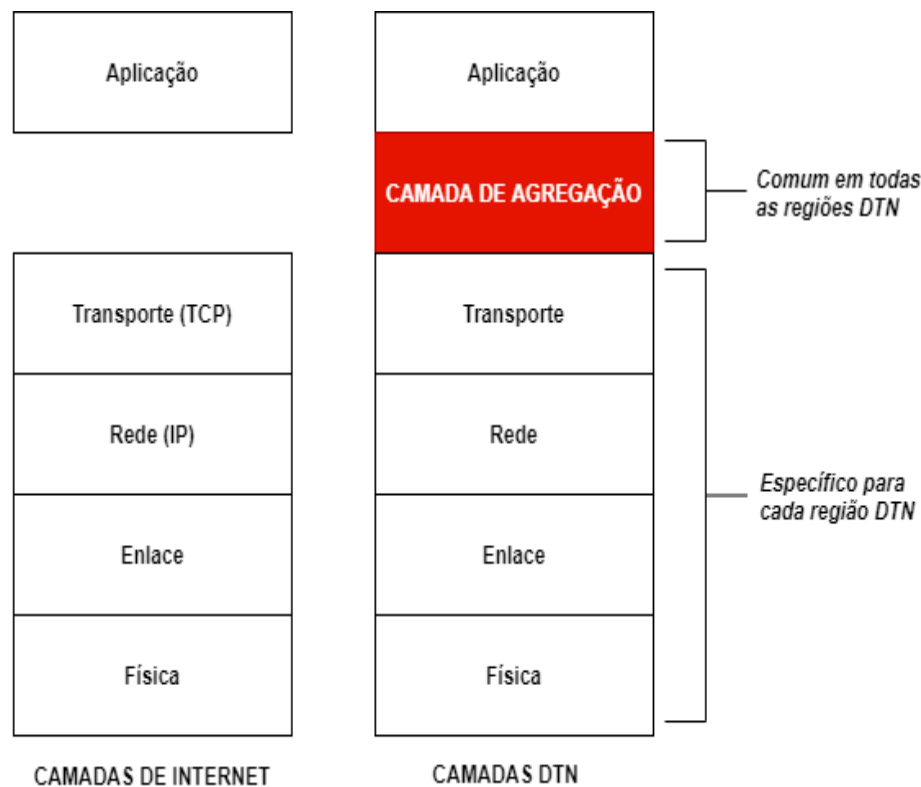


Figura 5: Camada de Agregação Bundle

A camada de agregação permite o armazenamento persistente através do protocolo de agregação, possibilitando que as mensagens sejam divididas em tamanhos variáveis chamadas de unidades de dados da aplicação (*Application Data Units – ADUs*) e transformadas em uma ou mais unidades de protocolo (*Protocol Data Units - PDUs*) chamadas agregados (*Bundles*) [Neves, 2018].

Essas mensagens são transformadas pela camada de agregação em mensagens agregadas, ficando disponíveis para armazenamento (*storage*) e encaminhamento (*forwarding*). Esses agregados são transmitidos entre os nós da rede quando acontece o fenômeno de contato entre eles, podendo ser repassados ou não dependendo de vários fatores.

É importante destacar que todas as informações necessárias para uma conexão e troca de dados são agregadas e enviadas uma única vez evitando diversas trocas de mensagens que são realizadas em uma transferência de dados padrão de uma rede TCP/IP [Cerf et al., 2007].

2.3 Tipos de Contato em uma rede DTN

O mecanismo de transmissão de dados, em uma rede DTN, é sustentado pela existência do contato e nem sempre os nós que fazem parte desta rede são alcançáveis e estão disponíveis, isso ocorre devido a mobilidade dos nós da rede ou restrições de seus recursos (pouca bateria dos dispositivos, pouco espaço de memória). Por esse motivo, o contato entre os nós é extremamente importante neste tipo de rede. Os dados armazenados são transferidos durante o tempo de duração do contato. [Neto, 2012].

Um contato corresponde a uma ocasião favorável à troca de dados entre dois nós. Tal evento ocorre quando os nós estão no raio de alcance um do outro, possibilitando que os dados armazenados, dependendo da estratégia de roteamento sejam transferidos [Xu et al, 2009].

Os tipos de contatos podem ser:

- **Persistente** – são contatos sempre disponíveis. Por exemplo, uma conexão de internet cabeada [Neto, 2012].
- **Em demanda** – são contatos que, para ocorrerem, necessitam que uma ação seja tomada [Souza, 2013]. Por exemplo, conexões bluetooth de dispositivos móveis que dependem da ação do usuário para serem conectados.
- **Agendado (Programado)** – ocorrem em espaços de tempo definidos e com duração conhecida [Neto, 2012], onde o horário e a duração do contato são definidos previamente entre os nós que precisam se comunicar, isso exige uma sincronização no tempo para a troca de mensagens ocorrer com sucesso, outros autores também denominam este contato de programado [Oliveira, 2008]. Redes interplanetárias são exemplos deste tipo de contato.
- **Oportunista** – são contatos que ocorrem à medida que os nós se movimentam e entram no mesmo raio de ação, surgindo assim a oportunidade de troca de informações [Souza, 2013].
- **Previsível** – Podem ter sua ocorrência prevista através do histórico de contatos anteriores [Neto, 2012]. Onde os nós podem fazer previsões sobre o horário e a duração dos contatos com base nos históricos previamente realizados [Oliveira, 2008].

Podemos exemplificar através da Figura 6, uma rede em uma área rural, onde um ônibus funciona como intermediário no transporte de informações. Este é um exemplo de contato previsível, baseado nos históricos anteriores, teríamos o horário que o ônibus chega ao ponto e o tempo que fica parado [Melo, 2011].

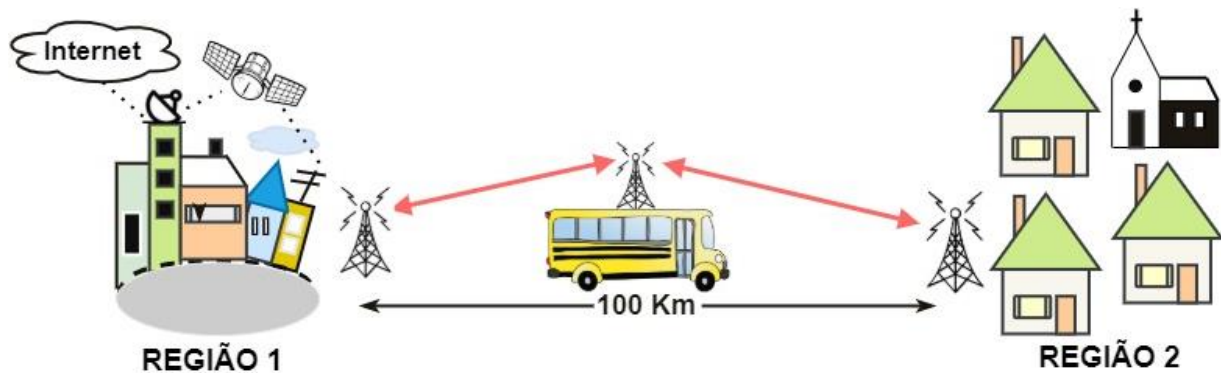


Figura 6: Contato previsível de uma DTN em zona rural (adaptado de Melo, 2011)

2.4 Métricas utilizadas em uma rede DTN

Para avaliar o desempenho de uma rede DTN, são necessárias a utilização de métricas como medida de desempenho, eficiência e a eficácia da rede [Keränen, Ott e Kärkkäinen, 2009]. A seguir são apresentadas as métricas mais utilizadas conforme os autores:

- **Tempo de contato** - Mede o tempo médio de contato entre os nós da rede.
- **Tempo entre contatos** - Mede o tempo médio entre os contatos dos nós da rede.
- **Taxa de entrega** - Razão entre o número de mensagens recebidas e o número de mensagens enviadas.
- **Atraso de Entrega** - Intervalo de tempo médio entre o envio e o recebimento das mensagens na rede.
- **Ocupação da rede** - Porcentagem média de ocupação de espaço de armazenamento dos nós da rede.

2.5 Redes Oportunistas

O termo redes oportunistas tem sido usado para diferentes tecnologias. Definimos redes oportunistas como o conjunto de aplicativos e serviços executados em dispositivos de usuários finais (por exemplo, *smartphones*, *tablets* e dispositivos

digitais semelhantes) que usam oportunidades de comunicação direta para trocar informações entre si.

Dessa forma, todos os dispositivos na área podem ser facilmente acessados, sem a necessidade de um aplicativo especial ou assinaturas pagas de dados móveis. Além disso, se a infraestrutura de comunicação, como 3G ou 4G, estiver indisponível ou com falhas, as mensagens ainda serão trocadas. O termo “oportunista” refere-se ao fato de que as oportunidades de comunicação com outros dispositivos são usadas à medida que elas acontecem [Dede et al., 2018].

O conceito de redes oportunistas segue a abordagem de redes tolerantes a atrasos e desconexões (DTNs), onde não há necessidade de conhecimento da topologia, e o encaminhamento é feito oportunisticamente ao longo dos contatos esporádicos que ocorrem ao passar do tempo [Fall, 2003]. Em suma, as redes oportunistas são consideradas uma evolução das MANETs, e uma subcategoria de redes tolerantes a atrasos e desconexões (DTNs) herdando todas as suas características [Rolim, 2016].

As redes oportunistas são altamente dinâmicas, compostas por dispositivos móveis e fixos, tiram vantagem dos contatos oportunistas ao longo do tempo para trocar informações carregadas pelos usuários [Pond et al., 2013], elas exploram a mobilidade e os contatos oportunistas entre os nós como forma de proporcionar a entrega de dados [Soares, 2016].

O objetivo é se beneficiar da mobilidade desses dispositivos como apresentado, na Figura 7 para realizar a troca de mensagens conforme as oportunidades de contato surgem, assumindo não haver conectividade fim-a-fim entre o nó gerador de mensagem e o nó destino da mensagem.

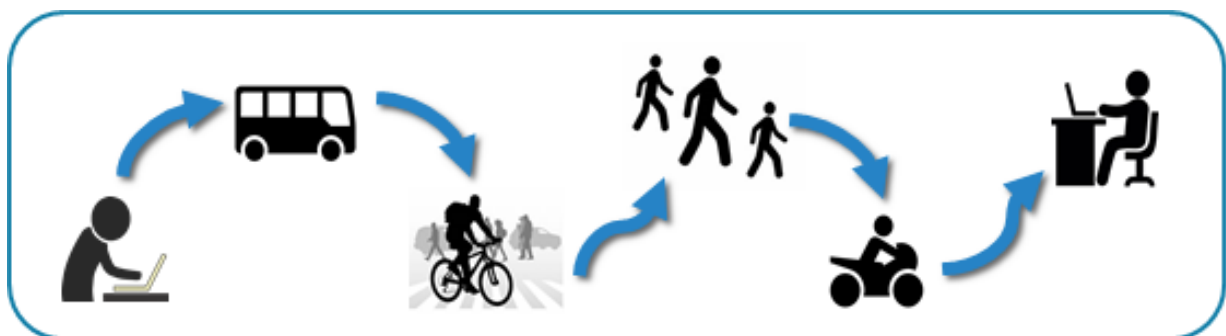


Figura 7: Comunicação Oportunista (adaptado de Pelusi et al., 2006).

As redes oportunistas utilizam todos os seus nós como mulas de dados, uma vez que não é possível antecipar quais serão os nós disponíveis para o roteamento

das mensagens, ou seja, quando não é possível saber antecipadamente o próximo salto entre o remetente e o receptor [Palazzo et al., 2011].

A movimentação dos nós pode ser representada por um modelo matemático de mobilidade, ou através de um arquivo de registros chamado de rastro (*trace*, em inglês) contendo informações de deslocamento dos nós, coletados em um experimento real.

Segundo [Musolesi et al., 2009], os modelos de mobilidade podem ser classificados em: **modelos de mobilidade sintéticos**, que se baseiam em formulações matemáticas como equações analíticas ou eventos estocásticos e os **modelos baseados em *traces*** (descrição de contatos entre os nós da rede), por exemplo *traces* baseados na captura de dados reais de movimentação humana.

2.6 Modelos de mobilidade sintéticos

▪ Modelo *Random Waypoint*

Como modelo de mobilidade sintéticos e dentre vários modelos um dos modelos mais utilizados na literatura é *Random Waypoint (RWP)* suas principais características são a escolha aleatória da posição, velocidade e tempo de pausa a partir de uma distribuição de probabilidade uniforme, entre os intervalos de mudança de direção do movimento dos nós [Yoon et al. 2003], então dentro de uma área de interesse $N \times M$, os nós se movimentam ao longo desta superfície até o próximo *waypoint* com velocidade aleatória, repetindo todo o processo novamente [Soares, 2018].

▪ Modelos baseados em comunidade

Segundo Neves (2018), comunidade é um conceito importante tanto na ecologia quanto na sociologia. Na ecologia, uma comunidade é um conjunto de duas ou mais populações de diferentes espécies que ocupam a mesma área geográfica. Na sociologia, a comunidade é definida como um grupo de pessoas que interagem vivendo em uma localização [Neves, 2018].

Intuitivamente, uma comunidade é um grupo de entidades sociais, dentro de uma grande rede social, que se associam umas às outras por laços sociais. Contudo, transformar essa noção intuitiva em uma definição formal não é tão simples. Como não existe um consenso na própria definição de comunidade, o conjunto de algoritmos de detecção de comunidades proposto na literatura também é muito variado [Cruz, 2013].

Dessa forma vamos exemplificar utilizando o trabalho de [Spyropoulos et al., 2006], que apresenta um modelo baseado em comunidade representada por células geográficas onde cada nó é associado a uma célula chamada de região local preferida pelo nó, essa região é definida pelo movimento aleatório do nó dentro deste local (célula/comunidade).

2.7 Traces de mobilidade

Pode-se definir traces em redes oportunistas (OppNets), como uma coletânea de dados sobre contato ou sobre localização que são extraídos a partir de experimentos em ambientes reais. Sendo provenientes de duas possíveis variantes: traces de localização e traces de contato [Soares, 2018].

Os traces de localização são extraídos a partir da localização e medições geográficas de sistemas de GPS (*Global Positioning System*), enquanto traces de contato representam medidas de proximidade entre objetos da rede com maior similaridade entre si, geralmente são obtidos durante eventos em diferentes ambientes como seminários, conferências, locais urbanos, campus universitários entre outros, utilizando o escaneamento de contatos através das tecnologias de *Wi-fi Direct* ou *Bluetooth* [Soares, 2018].

Sua utilização em redes oportunistas é importante para os estudos realizados sobre desempenho de novos algoritmos, além do mais, suas informações são mais precisas, o que torna os experimentos mais realistas, pois apresenta padrões de comportamento, interação e mobilidade humana dos participantes do local de coleta de dados.

Na Tabela 1 é apresentada, como exemplo, algumas informações sobre traces disponíveis:

Tabela 1: Exemplos de Traces de Contato disponíveis

Trace	Dispositivo	# de nós	Duração (dias)	Descrição
Cambridge	iMotes	54	11	Universidade de Cambridge – Campus universitário.
Reality	Smartphone	97	246	Massachusetts Institute of Technology – MIT - Campus universitário.
Sassy	T-mote	27	79	Universidade de St. Andrews - Campus universitário.

2.8 Roteamento em redes oportunistas

As desconexões e alta taxa de rotatividade dos nós são características normais em redes oportunistas. Por isso o roteamento é um desafio nesse ambiente. A taxonomia dos métodos de roteamento utilizada por [Nguyen and Giordano, 2008] separa os algoritmos em classes, como mostrado a seguir:

- **Roteamento com contexto obvio:** protocolos de roteamento que utilizam técnicas de *flooding*⁴ (inundação da rede por meio da duplicação de mensagens) para encaminhamento de mensagens.
- **Roteamento baseado na mobilidade dos nós:** informações sobre o movimento dos nós são utilizados para determinar os nós por onde as mensagens serão encaminhadas.
- **Roteamento baseado nas informações sociais:** os contextos sociais dos portadores de dispositivos são recuperados e utilizados para determinar os nós por onde as mensagens serão encaminhadas.

Em uma rede OppNet, como os nós são dinâmicos, ela é construída a partir dos instantes de contato imprevisíveis e curtos, à medida que os nós entram e saem da conectividade [Palazzo et al., 2011]. Portanto, todos os nós podem ser mulas de dados, que visam encontrar o próximo nó móvel que aceite armazenar a mensagem e a conduza em direção ao destino.

Aspectos sociais têm sido amplamente utilizados por muitos protocolos de roteamento propostos pela literatura, como protocolos de transmissão de dados que utilizam comunidades e similaridades e protocolos que utilizam o egoísmo em OppNets.

Um membro de uma determinada comunidade provavelmente interage de maneira quase constante com outro membro da mesma comunidade do que com um membro da população escolhido aleatoriamente. Portanto, as comunidades refletem naturalmente o relacionamento social entre as pessoas, ou seja, dispositivos dentro de uma mesma comunidade tem maiores chances de se encontrar [Neves, 2018].

⁴ Utilizando as técnicas de *flooding*, um nó envia uma mensagem para todos os outros nos vizinhos a ele, cada nó recebe e repassa a mesma mensagem para todos os seus vizinhos, exceto, para aquele que lhe enviou a mensagem.

Através deste conhecimento muitas abordagens foram desenvolvidas na literatura utilizando estruturas comunitárias para escolher os melhores nós para retransmissão de mensagens até seu destino, melhorando a chance de entrega.

A probabilidade de dois nós serem conectados é maior quando eles têm um vizinho em comum, então utilizando o conceito de similaridade que é uma medida do grau de semelhança, pode-se medir o número de vizinhos entre indivíduos em redes sociais.

Porém, se o contato entre os vizinhos dos nós são improváveis, significando que não existem vizinhos em comum e nenhuma interação entre eles, espera-se que a difusão demore mais tempo do que quando a similaridade é alta, com vizinhos em comum e maior interação. Existem outras maneiras de similaridades com vizinhos comuns, tais como, a semelhança de interesses e a semelhança nos mesmos locais dos utilizadores [Hidi, 2006].

Existem protocolos que foram desenhados para utilizar roteamentos oportunistas baseados nessas duas características sociais, comunidade ou grupos de interesse e similaridades. Já o aspecto social de egoísmo em redes oportunistas ocorre por diversos motivos, principalmente como forma de economizar os recursos (*buffer*, energia etc.), muitas vezes limitados, dos nós da rede.

Para compreender esse efeito, podemos nos voltar a sociologia para rever os mecanismos por quais as pessoas decidem com quem interagir. Essas interações sociais requerem recursos que são limitados, como as pessoas tendem a se relacionar com outras que possuem os mesmos gostos sociais, então podemos dizer que os seres humanos escolhem cuidadosamente com quem compartilhar seus recursos, maximizando seus próprios recursos individuais [Pujol et al., 2009].

De um modo geral, as pessoas tendem a interagir mais com pessoas que compartilham mais gostos sociais que outras [Soares, 2018] ou pode ocorrer também de se tornarem naturalmente egoístas, simplesmente pelo fato de não disponibilizar de recurso suficiente naquele momento, como por exemplo, energia.

2.9 Grafos sociais e formação de grupos em OppNets

Considerando-se uma rede onde os nós são constituídos por dispositivos móveis carregados pelos usuários, podemos gerar um grafo social da rede e com base nos relacionamentos sociais de cada usuário determinar a melhor rota para um envio oportunista de informação até um nó destino.

Já foram propostos uma diversidade de protocolos que variam de acordo com o encaminhamento empregado. Usar algum teor social como fator de decisão no roteamento oportunista tem grandes vantagens na entrega de dados de forma eficiente [Moreira et al., 2013]. Algumas estratégias com caráter social são: utilização de laços sociais como a força da amizade e a formação de comunidades ou grupos formados por nós que possuem o mesmo interesse social.

Podemos definir que uma comunidade é um conjunto de duas ou mais populações de diferentes espécies que ocupam a mesma área geográfica. Na sociologia, a comunidade é definida como um grupo de pessoas que interagem vivendo na mesma localidade [Neves, 2018]. As comunidades refletem naturalmente o relacionamento social entre as pessoas.

A formação de grupos de interesse resulta de maneira geral da busca de objetivos comuns por parte de indivíduos que compartilham interesses afins e que, se beneficiariam atuando coletivamente [Olson, 1965].

O “interesse comum” sobre algum assunto é um dos fatores que levam pessoas a formarem grupos. Por exemplo, a união de pessoas que possuem interesse em fazer o mesmo tour turístico e gostariam de trocar informações com outras que já o fizeram ou que ainda pretendem fazê-lo [Albarelllo, 2005].

2.10 Egoísmo em Redes Oportunistas

No mecanismo comum da comunicação em OppNets, os nós servem como encaminhadores de dados para outros nós ao carregar mensagens de forma intermediária. Isto requer uma comunicação totalmente cooperativa. Contudo, em aplicações reais, os nós podem ser dispositivos controlados por pessoas reais e estas pessoas podem não querer encaminhar determinadas mensagens [Soares, 2018].

A razão é que interações sociais requerem recursos que são limitados, então os seres humanos escolhem cuidadosamente com quem compartilhar seus recursos, maximizando seus próprios recursos individuais [Pujol et al., 2009]. O egoísmo pode ocorrer por diversos motivos, uma delas como forma de economizar os recursos limitados, dos nós da rede tais como energia, consumo de memórias etc. [Soares, 2018].

É considerado como socialmente egoísta o indivíduo que participa parcialmente das ações de transmissão e retransmissão de mensagens na rede.

Esses nós desejam apenas armazenar e carregar mensagens destinadas a outros nós com quem ele tem uma relação social forte [Zhu et al., 2013].

Podemos dividir o egoísmo relacional dos nós de duas maneiras. Se um nó demanda por economizar algum recurso e ele não coopera com a comunicação nem mesmo dentro do mesmo grupo social, denominamos egoísmo individual (*individual selfish*). Podemos denominar egoísmo social (*social selfish*), quando um nó coopera mais com usuários de um mesmo grupo social que ele, mas não auxilia no repasse com outros de fora do seu grupo social [Soares, 2018].

2.11 Consumo de energia em OppNets

Um dos problemas que se destaca em OppNets é o consumo de energia. O consumo de energia ocorre principalmente no processo de descoberta de dispositivos vizinhos e a economia eficiente deste recurso é um dos principais desafios. [Azevedo, 2010].

A energia é uma fonte preciosa em aparelhos móveis, por isso o padrão 802.11 provê capacidade de gerenciamento de energia, permitindo que os nós minimizem o tempo de suas funções de percepção, transmissão e recebimento, e outros circuitos necessários para o seu funcionamento [Kurose e Ross, 2013].

2.12 Fontes de Consumo de energia

O intervalo de tempo entre o encontro de nós é geralmente maior do que as durações de contato, o que mostra que os nós passam a maior parte do tempo em busca de nós vizinhos em vez de enviar ou receber informações.

Todas as operações em dispositivos de usuários finais precisam de energia e o comportamento de ambas as partes consumidoras de energia (processamento, visualização, comunicação, por um lado, e as baterias, por outro), é complexo.

Essa complexidade resultou no desenvolvimento de vários modelos que caracterizam e enfatizam diferentes aspectos do manuseio do uso de energia para utilização em pesquisas sobre impacto de consumo energético [Dede et al., 2018].

Considerando a funcionalidade básica, os modelos disponíveis podem ser classificados da seguinte forma:

- **Modelo de bateria** – deve idealmente imitar as características comportamentais das baterias do mundo real, como capacidade, tensão e corrente máximas, carregamento e descarregamento, dependência de temperatura, vida útil etc.

- **Modelos de gerador de energia** – simulam componentes como estações de recarga, painéis solares e componentes similares. Eles são muitas vezes combinados com o modelo de bateria.
- **Modelos de consumo de energia ou despesa de energia** – refletem o uso de energia pelos componentes do dispositivo e/ou atividades. Nestes modelos, o consumo de energia dos dispositivos simulados é implementado como requisitos de energia conforme estados ou eventos tais como ocioso, transmitindo, recebendo ou processando as operações dos dispositivos de rede.

2.13 Aprendizado de máquina

Aprendizado de máquina (*machine learning*) é uma subárea da Inteligência Artificial (IA) pode ser definida como qualquer mudança em um sistema que melhore o seu desempenho na segunda vez que ele repetir a mesma tarefa, ou outra tarefa da mesma população [Simon, 1983].

Pode ser classificado como aprendizado supervisionado e não supervisionado. Se as informações que estão contidas nas bases de dados estiverem rotuladas com sua classe correspondente, pode-se utilizar algoritmos de aprendizado supervisionado; estes induzem padrões a partir de dados. Mas no caso, se estes dados não estiverem rotulados, se utilizam algoritmos de aprendizado não supervisionado [Sanches, 2003].

Neste trabalho, usamos os algoritmos de aprendizagem de máquina não supervisionado *EM* (*Expectation-Maximization*) e *K-means*.

2.13.1 Técnicas de agrupamento

Existem diversas abordagens ou técnicas utilizadas dentro da aprendizagem de máquina dentre elas a chamada técnica de agrupamento ou clusterização (*Clustering*). *Clustering* é um método de aprendizado não supervisionado, sendo comum sua utilização em análise de dados estatísticos. As técnicas de clusterização estão divididas em hierárquica e particionada.

O agrupamento hierárquico está subdividido em aglomerativa e divisível, que criam *clusters* gradualmente a cada iteração [Sharma et al., 2012]. Um cluster é um grupo ou conjunto de dados formado por amostras que possuem algumas

características e comportamentos que são comuns ou que possuem semelhanças entre si.

A ideia básica é que elementos que compoñham um mesmo cluster devem apresentar alta similaridade (sejam elementos bem parecidos, que seguem um padrão similar), mas devem ser muito dissimilares de objetos de outros clusters [Puc Rio, 2010].

Os algoritmos de agrupamentos descobrem *clusters* diretamente por deslocamento iterativo de pontos entre subconjuntos ou identifica agrupamentos em áreas altamente povoadas com dados. Frequentemente, são definidas por alguma métrica de similaridade e dissimilaridade entre membros de um mesmo cluster interno ou separação entre clusters diferentes [Sharma et al., 2012].



Figura 8: Clusters em duas dimensões (Jain e Dubes, 1988)

As técnicas de agrupamento são aplicadas quando não existem classes previamente definidas, ou seja, não existe uma base contendo o histórico de informações que pode ser utilizada durante uma coleta de informações. Sendo assim, o uso dessa técnica é feito quando as amostras devem ser divididas em grupos com bases em suas características [Soares, 2008].

O *Clustering* consiste em agrupar exemplos de tal forma que estes pertencentes a um mesmo cluster sejam mais similares entre si, de acordo com alguma medida de similaridade pré-definida, torna-se necessário e importante então definir o conceito de similaridade utilizado nesse processo [Sanches, 2003].

A similaridade pode ser calculada de diversas maneiras, mas três métodos destacam-se em algoritmos de *clustering*:

1. medidas de distância,
2. medidas de correlação e
3. medidas de associação.

Temos então que as técnicas de agrupamento podem assumir os relacionamentos entre os objetos de determinados conjuntos podem ser representados por uma matriz contendo medidas de dissimilaridades ou de proximidade entre cada par de objetos.

Cada entrada p_{ij} na matriz corresponde em um valor numérico que demonstra a proximidade entre os objetos i e j . Algumas métricas calculam a similaridade, outras calculam a dissimilaridade, mas em essência elas são idênticas. Todos os coeficientes de dissimilaridade são funções $d: \Gamma \times \Gamma \Rightarrow \mathfrak{R}$, onde Γ denota o conjunto de objetos pertencentes ao conjunto em questão [Linden, 2009].

$$\Gamma = \begin{bmatrix} x_{ll} & \cdots & x_{lf} & \cdots & x_{lp} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ x_{il} & \cdots & x_{if} & \cdots & x_{ip} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ x_{nl} & \cdots & x_{nf} & \cdots & x_{np} \end{bmatrix}$$

Essas funções realizam a transformação da matriz, então a matriz de distância é dada pela entrada $d(i,j)$ que é a distância entre os elementos i e j .

$$d = \begin{bmatrix} 0 & & & & \\ d(2,1) & 0 & & & \\ d(3,1) & d(3,2) & 0 & & \\ \vdots & \vdots & \vdots & & \\ d(n,1) & d(n,2) & \cdots & \cdots & 0 \end{bmatrix}$$

Os métodos de similaridade por distância mais comuns são:

1. Distância de Manhattan,
2. Distância Euclidiana e
3. Similaridade de Cossenos

A medida distância de Manhattan tem esse nome porque em várias cidades é praticamente impossível estabelecer uma rota entre dois pontos através de uma reta, devido ao fato de as cidades serem subdivididas em quadras, e estas não estarem vazias, mas sim preenchidas com prédios. Esta medida é definida pela função abaixo, que é simplesmente a soma das diferenças absolutas das dimensões [Sanches, 2003]. A distância dada entre dois pontos $x = (x_1, x_2, \dots, x_n)$ e $y = (y_1, y_2, \dots, y_n)$ é dada pela equação 2.1:

$$d(x, y) = \sum_{i=1}^n |x_i - y_i| \quad (2.1)$$

Onde x_i e y_i são as i coordenadas de cada objeto e x e y são objetos de uma massa de dados [Neves, 2018].

A distância euclidiana é a distância mais citada e utilizada na literatura, sendo expressa pela distância geométrica euclidiana entre os exemplos em um espaço multidimensional [Sanches, 2003].

Sendo representada pela distância dada entre dois pontos $x = (x_1, x_2, \dots, x_n)$ e $y = (y_1, y_2, \dots, y_n)$ através da equação 2.2:

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (2.2)$$

Temos ainda a similaridade de cossenos, onde se analisa a distância angular entre dois pontos a partir da origem, ou seja, é a medida do ângulo entre dois vetores de n dimensões.

Cada objeto assim é representado por um vetor e a medida do ângulo entre eles representa o grau de similaridade dos dois objetos. Esta similaridade é dada pela equação 2.3, onde A e B são objetos representados por vetores [Neves, 2018].

$$CosSim(A, B) = \frac{A \cdot B}{\|A\| \|B\|} \quad (2.3)$$

2.13.2 Algoritmo Esperança-Maximização (EM)

As técnicas de agrupamento são amplamente utilizadas através de algoritmos que automatizam a clusterização dos dados. Vários métodos e algoritmos vem sendo desenvolvidos com o intuito de dinamizar a aplicação das técnicas de agrupamento de dados em âmbitos sociais e econômicos [Palma, 2018]. Para o trabalho usaremos os algoritmos de agrupamento EM e K-means.

O algoritmo EM (*Expectation-Maximization*), assim como a maioria dos algoritmos que pertencem à classe das técnicas probabilísticas, baseia-se em um modelo estatístico chamado "misturas finitas". Uma mistura finita é um conjunto de k distribuições de probabilidades, representando k clusters, que governam os valores dos atributos dos exemplos naquele cluster.

Em outras palavras, cada distribuição determina a probabilidade de um determinado exemplo possuir um certo conjunto de valores de atributos dado que esse exemplo pertence a esse cluster [Sanches, 2003].

O método usado para clusterização consiste em encontrar estimativas de máxima verossimilhança para um conjunto de parâmetros, quando se tem variáveis não observáveis. Essas variáveis não-observáveis são calculadas pela combinação linear das variáveis originais [McLachlan e Krishnan, 2007].

O algoritmo EM (*Expectation-Maximization*) executa um cálculo iterativo de estimadores de máxima verossimilhança quando as observações podem ser vistas como dados incompletos (Picard, 2007). Seu objetivo é estimar o número de populações (*clusters*) na amostra, onde cada cluster representa uma distribuição de probabilidade, resumindo estima a média amostral e sua variância.

O EM associa cada objeto a um cluster de acordo com um peso (probabilidade de distribuição), representando sua probabilidade de membresia. Começa com uma estimativa inicial de um vetor de parâmetros, e de maneira iterativa reavalia os objetos com relação à mistura de distribuições produzida pelo vetor de parâmetros, os objetos reavaliados são usados para atualizar a estimativa dos parâmetros. A cada objeto é associado uma probabilidade de pertencer a um cluster [Neves, 2018].

O algoritmo possui dois passos de interação distintos para refinar a formação dos clusters, onde:

1. No passo E (*Expectation*), a verossimilhança do conjunto de dados como um todo é calculada, associando cada objeto x_i ao cluster C_k , onde i é o índice do objeto da amostra e k é o índice do cluster encontrado no conjunto de amostras,

$$P(x_i \in C_k) = p \frac{C_k}{x_i} = \frac{p(C_k) p(x_i/C_k)}{p(x_i)}, \quad (2.4)$$

onde $p(x_i/C_k) = N(m_k, E_k(x_i))$ segue uma distribuição normal de probabilidade com média m_k e valor esperado de E_k .

2. No passo M (*Maximization*), os valores desconhecidos são recalculados por meio da maximização da função do passo anterior, usando então as probabilidades estimadas no passo anterior para reestimar temos:

$$m_k = \frac{1}{n} \sum_{i=1}^n \frac{x_i p(x_i \in C_k)}{\sum_j p(x_i \in C_j)} \quad (2.5)$$

A execução desses passos continua até que algum critério de convergência seja satisfeito.

2.13.3 Algoritmo K-means

Diferentemente do agrupamento hierárquico, no qual o agrupamento funde ou particiona os grupos a cada iteração, no K-means o agrupamento é particionado em exatos K grupos, não havendo qualquer hierarquia entre eles. Assim, no k-means, o objetivo traçado é encontrar um agrupamento tal que os grupos sejam tão homogêneos quanto possível e os objetos em cada grupo estejam bem distintos de objetos em outros grupos [Jain e Dubes, 1988].

O *k-means* posiciona de forma aleatória K pontos centrais, chamados centroides, dentro do espaço $\mathbb{R}^c$, onde c representa a dimensão utilizada e definida pelo número de atributos c utilizados nas amostras. Cada amostra a é associada ao grupo mais próximo através da distância entre a e o centroide $C_i \forall i \in K$ [Soares, 2018].

Assim, cada centroide é recalculado, a cada nova amostra que entra no grupo. Essa iteração do *k-means* termina quando o processo de atualização encerra após obter a posição ideal dos centroides, significando assim que nenhuma nova associação pode ser feita, e os K grupos formados representam as similaridades mais próximas quanto possíveis [Soares, 2018].

Após escolhidos os centros, cada dado é assinalado como pertencente à classe cujo centro está mais próximo. Após todos os dados serem processados, a posição média de cada classe é então calculada e usada como o novo centro, e os dados processados novamente.

Como mencionado anteriormente esse processo é repetido até que não mais ocorram mudanças nos centros encontrados ou um número limite de iterações definido pelo usuário seja atingido [Santos et al, 2015]. O algoritmo na Figura 9 demonstra o funcionamento do algoritmo de clusterização *K-means*.

Algoritmo 1 Algoritmo de aglomeração *K-means*

1. Inicialize K centros ($w_1, \dots, w_k$), representam as classes;
2. Cada classe C_j é associada a um conjunto w_j ;
3. Para cada dado de entrada x , assinale-o como pertencente à classe C_j caso w_j seja o centro mais próximo a este dado;
4. Para cada classe C_j , atualize os centros w_j , através da medida de todas as amostras de dados assinaladas como pertencentes a cada classe C_j ;
5. Calcular a função de erro E

$$E = \sum_{j=1}^k \sum_{x \in C_j} \|x - w_j\|^2;$$

6. Repetir passos 3, 4 e 5 até que o erro E não mude significativamente.
-

Figura 9: Algoritmo K-means (Santos et. al., 2015)

Além da formalização do algoritmo k-means apresentado na Figura 9, ele pode ser facilmente compreendido com a descrição feita seguindo os passos citados por Linden (2009) na Tabela abaixo:

Tabela 2: Passo a passo K-means (Linden, 2009)

- | |
|---|
| <ol style="list-style-type: none">1. Escolher K distintos valores para os centros dos grupos (possivelmente de forma aleatória)2. Associar cada ponto ao centro mais próximo (maior similaridade)3. Recalcular o centro de cada grupo4. Repetir os passos 2 a 3 até não haver mais mudanças significativas entre os pontos |
|---|

Segundo descreveu Ricardo Linden (2009), este algoritmo é extremamente veloz, geralmente convergindo em poucas iterações para uma configuração estável, na qual nenhum elemento está designado para um cluster cujo centro não lhe seja o mais próximo.

O algoritmo K-means é largamente aceito como uma técnica "padrão" de detecção automática de classes pela análise de dados. Trata-se de um método baseado em heurística, que possui um grande poder de modelagem. O princípio de aglomeração consiste na determinação do modelo da distribuição dos exemplos no espaço atribuído [Santos et al, 2015].

Podemos observar um exemplo de execução do algoritmo K-means na Figura 10, onde,

(a) cada elemento foi designado para um dos três grupos aleatoriamente e os centroides (representados pelos círculos maiores) de cada grupo foram calculados.

(b) Os elementos foram designados agora para os grupos cujos centroides lhe estão mais próximos e

(c) Os centroides foram recalculados e já estão em sua forma final, caso não estivesse continuariam o procedimento (a) e (b) respectivamente até atingir sua forma final [Linden, 2009].

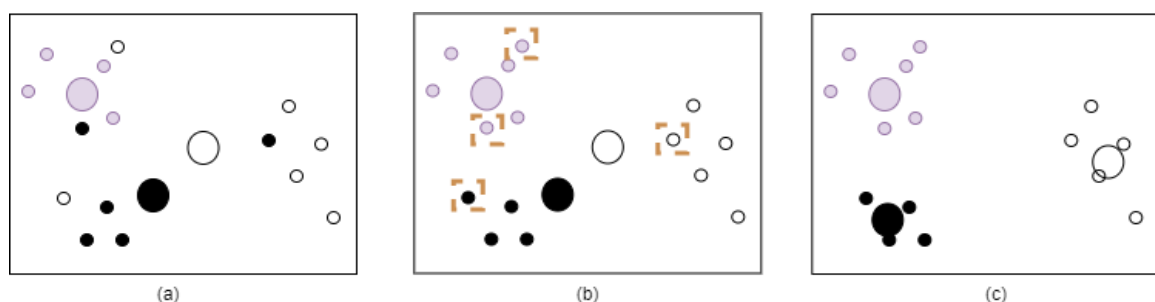


Figura 10: Execução do algoritmo K-means (adaptado de Linden, 2009).

2.14 Considerações finais

Este capítulo descreveu os principais conceitos para a compreensão deste trabalho. Assim, nas Seções 2.1 a 2.13 foram introduzidos aspectos fundamentais sobre redes tolerantes a atrasos e desconexões.

Na Seção 2.5 foram brevemente apresentadas as características das redes oportunistas e na Seção 2.6 os modelos de mobilidade e roteamento em OppNets, dentre as quais este trabalho irá utilizar o contexto de roteamento baseado em características sociais além disso, foi feita uma explanação na Seção 2.7 sobre traces de mobilidade, na seção 2.9 sobre a formação de grafos sociais e na seção 2.10 abordamos sobre aspectos do egoísmo em redes oportunistas.

Nas seções 2.11 e 2.12 consumo de energia em redes OppNet e modelos de fontes de consumo de energia utilizados em pesquisas, na Seção 2.13 fizemos uma apresentação sobre aprendizagem de máquina e técnicas de agrupamento além da apresentação formal do algoritmo Esperança-Maximização (EM) e K-means utilizados neste trabalho.

Trabalhos Correlatos

A literatura acadêmica compreende diversos trabalhos sobre o encaminhamento de mensagens em OppNets, dentre os quais se utilizam vários métodos de roteamento diferentes, iremos nos restringir aos trabalhos com abordagens de roteamento com aspectos sociais e investigação sobre recursos de energia. Este capítulo apresenta e discute os principais trabalhos existentes na literatura e que se relacionam com a proposta desta dissertação.

3.1 Estado da Arte

Ao analisar os cenários para os quais as OppNets são projetadas, pode-se ver que a mobilidade dos nós tem uma grande influência na quantidade e na qualidade das oportunidades de contato. Assim, o roteamento de mensagens e a mobilidade de nós têm um impacto determinante no consumo de energia. Já mencionamos que em ambientes DTN típicos como por exemplo áreas de difícil acesso como comunidades ribeirinhas, o acesso a uma fonte de alimentação externa é limitado. Portanto, nos algoritmos projetados para roteamento em OppNets, o consumo de energia deve ser cuidadosamente considerado como uma importante medida de avaliação [Silva et al., 2012].

Os mecanismos tradicionais de roteamento não são mais adequados e recentes abordagens usam as relações sociais entre usuários móveis como um critério para o processo de roteamento. O protocolo *Bubble Rap Forwarding*, por exemplo, é considerado o primeiro algoritmo de roteamento para OppNets baseado em características sociais [Souza, 2019].

Basicamente, este algoritmo toma as decisões de roteamento usando os conceitos de comunidade e centralidade. Nesse algoritmo, quando uma oportunidade de contato ocorre, mensagens são encaminhadas primeiramente para aqueles nós com alta centralidade e/ou membros da mesma comunidade a que pertence o destinatário da mensagem. Para realizar os cálculos necessários para detecção das comunidades e centralidades dos nós da rede, os autores utilizaram técnicas de

grafos como k-clique e um algoritmo centralizado de pesos das arestas do grafo que representa a rede [Crowcroft et al., 2011].

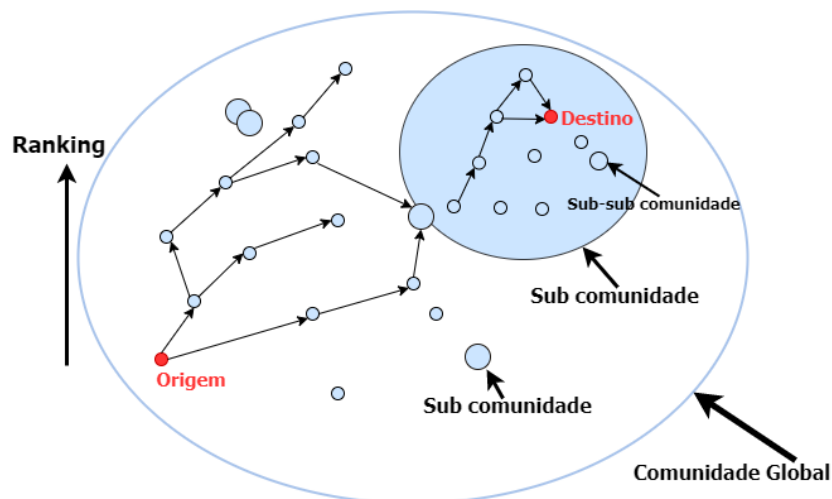


Figura 11: Ilustração do algoritmo Bubble Rap (adaptado de Crowcroft et. al., 2011)

Nos modelos mais recentes que utilizam características sociais para roteamento destacamos alguns trabalhos a seguir como o algoritmo “*Friendship and Selfishness Forwarding*” (FSF), que toma suas decisões de roteamento considerando os relacionamentos sociais e o egoísmo dos nós [Souza, 2019].

Para escolher nós intermediários para uma mensagem, FSF usa os laços sociais entre os nós como critério. A motivação para isso é que quanto mais forte for o laço social entre os nós, maior será a probabilidade deles se encontrarem no futuro. O algoritmo FSF divide os relacionamentos sociais entre os nós em três grupos: amigos, conhecidos e desconhecidos [Souza, 2019].

Outra abordagem utilizada por Eric Neves (2018) propõe um algoritmo para disseminação de mensagens usando os interesses dos nós da rede como fator social, formando diferentes grupos de repasse de mensagens usando clusterizadores para organizar os grupos por níveis de interesse direto e indireto sendo a mensagem então encaminhada entre os grupos até o nó de destino, a ideia é de que características sociais podem influenciar diretamente na estratégia de disseminação adotada, podendo ser usadas para agrupamento e repasse das mensagens até seus grupos e nós destinatários, contribuindo de forma direta para a economia de recursos da rede e sendo eficiente nas entregas das mensagens aos seus destinatários [Neves, 2018].

Int-Three é um algoritmo que utiliza o uso de características de interesses para disseminação de mensagens, isto ocorre segundo Hidi (2006), porque as

peessoas com interesses semelhantes se encontram com frequência e compartilham mais dados umas com as outras.

O trabalho de Xia et al. (2016) apresenta como base a exploração de duas características sociais, estrutura de comunidade e interesse dos usuários. É construído a Interest-Three, estrutura de comunidade baseada no interesse de cada usuário que é atualizada dinamicamente a cada contato, onde é recalculado a densidade da comunidade e o laço social entre os usuários.

Para a eleição de um bom nó repassador a Int-Three decide se o nó intermediário é adequado para ser um repassador de acordo com os critérios na estratégia de encaminhamento [Xia et al., 2016].



Figura 12: Componentes Int-Tree (adaptado de Xia et. al., 2016)

Outro protocolo de roteamento é o denominado *SocialCast*, este possui suporte a publicações e assinaturas de conteúdos em redes DTN. Esse protocolo leva em conta o interesse dos nós receptores para rotear a informação através dos laços sociais entre os nós e, conseqüentemente, prever sua mobilidade. A disseminação dos dados sociais e de interesse dos nós é usada para uma melhor abordagem do repasse e armazenamento das mensagens pelos nós na rede [Costa et al. 2008].

Os autores Yoneki et al. (2007) utilizam a detecção de comunidades distribuídas a partir do rastreamento e propõe uma camada de distribuição de comunicação chamada *Socio-Aware Overlay*, que controla publicações e assinaturas dos nós da rede. Os nós com maior centralidade possuem uma maior visibilidade para outros nós na rede, por esse motivo é criada uma camada com os nós mais centrais de cada comunidade [Yoneki et al., 2007].

Dada a importância do impacto do consumo de energia na vida operacional de redes DTN, Denis Silva et al. (2012) projeta e implementa um módulo responsável pelo consumo de energia no simulador *The ONE* para analisar o consumo de energia dos nós de rede em protocolos de roteamento propostos para DTN [Silva et al., 2012].

O modelo considera o consumo de energia na transmissão e recepção de mensagens e na descoberta de nós vizinhos. Ele tem uma opção de recarga da bateria, que torna o ambiente de simulação mais realista para o estudo do impacto do consumo de energia nos protocolos de roteamento.

O módulo de energia recebe como parâmetro a quantidade de energia gasta durante a busca por outros nós, a transmissão e o recebimento de mensagens em unidades por segundo para cada operação, ou seja, a energia total gasta é a função de sua duração. Neste módulo o nó tem a opção de recarregar sua bateria. Durante a simulação, os autores reduziram a carga de bateria do nó para cada operação realizada.

Sendo assim, a proposta permite ao usuário definir a capacidade de bateria, quantidade de energia consumida por segundo para suas operações e o intervalo de tempo entre as recargas de bateria.

Os resultados apontam que quando o consumo de energia é considerado, os parâmetros operacionais dos protocolos mudam substancialmente. O módulo projetado por Denis Silva (2012) permite fazer a avaliação de energia para qualquer novo protocolo implementado neste ambiente de simulação permitindo a lógica interna do protocolo tomar decisões para economizar energia [Silva et al., 2012].

Recentemente Salah Loudari e Nabil Benamar (2019), fizeram um estudo sobre a avaliação e comparação do consumo de energia em protocolos de roteamento levando em consideração modelos de mobilidade versus consumo de energia e nós egoístas, chegando à conclusão que os protocolos de roteamento em redes oportunistas se concentram em diminuir o atraso médio, a sobrecarga e aumentar a taxa de entrega de mensagens, para isso aumentam o número de cópias de pacotes na rede, portanto, mais energia é consumida pelos nós deixando o ambiente mais hostil, pois o consumo de energia desempenha um papel crucial na modificação do comportamento dos nós na rede porque os nós com pouca energia na bateria podem se tornar egoístas e não cooperativos.

Os autores utilizaram os protocolos SPRAY e WAIT, DLife, SPRAY e FOCUS ambos apresentaram resultados melhores em comparação com o PROPHET,

EPIDEMIC e PROPHET v2 em termos de consumo de energia. Os experimentos foram feitos em dois ambientes distintos: cooperativo e egoísta. As simulações foram realizadas em três diferentes modelos de movimento: passeio aleatório, ponto de passagem aleatório e modelo de cluster [Loudari e Benamar, 2019].

Em se tratando de consumo de energia, existem trabalhos que dão ênfase em protocolos de roteamento que utilizam a centralidade dos nós. Sendo assim a centralidade de um nó é usada como critério de roteamento, os nós centrais (chamados *hubs*) são nós conectados à maioria dos outros nós sendo caracterizados através da análise do histórico de encontros, resultados desse tipo de abordagem mostram que os esquemas de encaminhamento que tendem a usar apenas os nós mais preferidos causam cargas injustas no sistema.

Uma abordagem publicada por Qaisar Ayub e Sulma Rashid (2020) propõe um protocolo baseado em detecção de nós inativos com eficiência energética, trata-se de um método para detectar nós em estado morto ou inativo na rede e fornecer uma política de gerenciamento de buffer eficiente com o objetivo de evitar congestionamento na rede. A arquitetura consiste em um módulo padrão de conectividade, um módulo para detecção de nós inativos, um módulo de reconhecimento de nós inativos e um módulo de tratamento de mensagens [Ayub e Rashid, 2020].

O protocolo é focado no controle de transmissão das mensagens que se destinam a nós mortos, seu método determina um nó morto, estudando a conectividade entre os nós e aplica uma política de buffer para eliminar as mensagens destinadas a nós mortos [Ayub e Rashid, 2020].

Seguindo nessa linha de utilização de centralidade de nós, o trabalho proposto por Socievole et al. (2015), busca usar os benefícios (algoritmos totalmente distribuídos etc.) desse tipo de abordagem centralizada em conjunto com mecanismos para economia de consumo de energia, para tal, projetaram um novo protocolo de roteamento oportunista chamado *Energy-Aware Centrality-Based Forwarding (ECF)*. Durante um encontro entre dois nós, cada nó calcula seu valor de centralidade levando em consideração seu nível de energia atual. O protocolo modulariza a centralidade de um nó com seu nível de energia atual [Socievole et al., 2015].

Os autores em sua proposta levam em consideração quatro níveis de energia baseados em observações empíricas: (1) acima de 50% (2) entre 40% e 50%, (3)

entre 30% e 40% e (4) abaixo de 30% do seu valor energético inicial. Para o nível de energia (1), assumem que o nó ainda tem energia suficiente para usar a versão não consciente de sua centralidade.

Ao contrário, abaixo deste limite, o valor de centralidade do nó é decrementado usando uma regra exponencial para os níveis de energia (2) e (3), chegando a zero com o nível de energia (4). Cada nó então examina as mensagens que está transportando e as encaminha somente se o nó encontrado for o destino ou tiver um valor de centralidade maior [Socievole et al., 2015].

Dessa maneira evita-se transmissões em nós mais centrais, mas com energia limitada, conforme pode ser observado na Figura 13 abaixo.

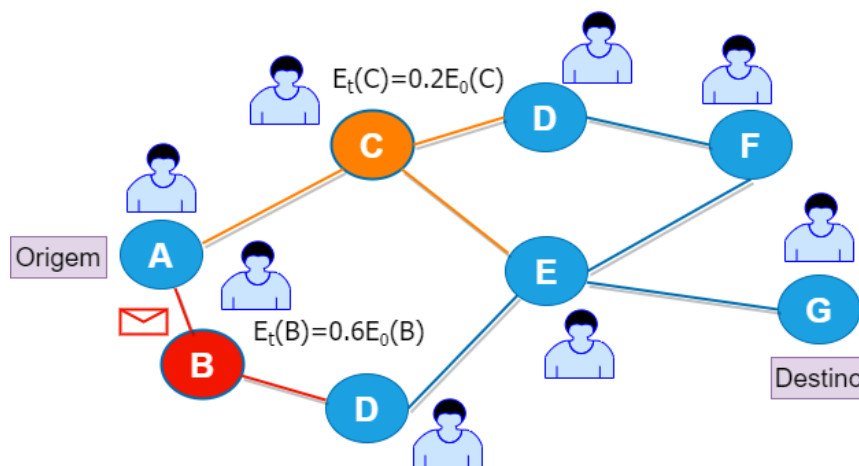


Figura 13: Funcionamento do Algoritmo *ECF* de encaminhamento baseado na centralidade dos nós e na consciência de energia (adaptado de Socievole et al. 2015).

Ao enviar uma mensagem do nó de origem A para o nó de destino G, A pode escolher entre os nós B e C como possíveis próximos-saltos no tempo t . Mesmo se C tiver uma centralidade maior (os três elos de contato em laranja), o *Energy-Aware Centrality-Based Forwarding (ECF)* escolhe B como o próximo salto, tendo uma centralidade mais baixa (os dois elos de contato em vermelho) e um nível de energia mais alto (60% de sua energia inicial [Socievole et al., 2015]).

Uma outra abordagem proposta por Fan Li et al. (2017) é a criação de um roteamento baseado em energia social chamado SEBAR, onde uma nova métrica chamada energia social é incluída para quantificar a capacidade de um nó encaminhar pacotes para outros nós utilizando leis gerais de física de partículas. Esse algoritmo trabalha correlacionando essa nova métrica social com métodos de detecção de comunidades, ou seja, a métrica energia social é calculada com base na

estrutura da comunidade e compartilhada por membros da comunidade [Li et al., 2017].

Na física de partículas, a energia pode ser gerada quando duas partículas colidem. Da mesma forma, quando dois nós móveis v_k e v_l se encontram, ocorre uma colisão entre eles e essa colisão gera certa quantidade de energia social, então essa energia gerada será distribuída uniformemente para esses dois nós. Cada nó contribui com parte dessa energia gerada para as comunidades a que pertence e detém a parte restante da energia, o objetivo é espalhar a energia social para um alcance mais amplo. O funcionamento do SEBAR é exemplificado na Figura 14.

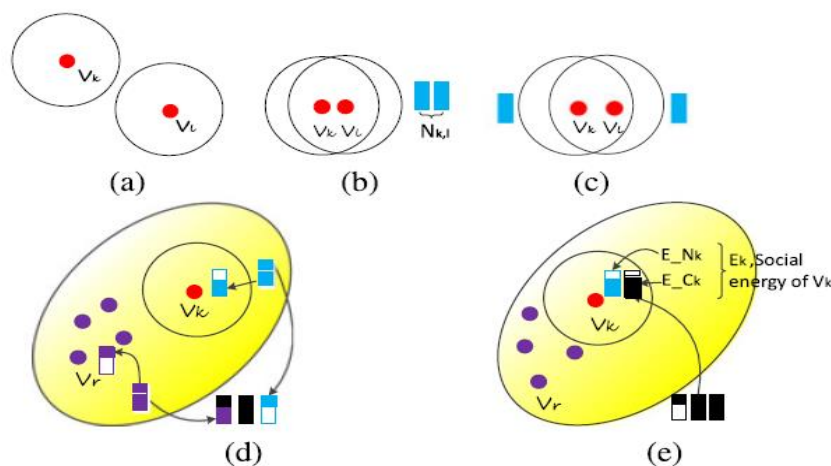


Figura 14: Energia social gerada pela colisão de nós, distribuída e compartilhada (Li et al., 2017).

Na Figura a letra (a) representa dois nós v_k e v_l . O círculo centrado no nó é o seu alcance de transmissão. (b) v_k encontra com v_l , e essa colisão gera $N_{k,l}$ unidades de energia. (c) $N_{k,l}$ unidades de energia são divididas igualmente por v_k e v_l . (d) v_k contribui com percentual de energia gerada pelos encontros para sua comunidade e reserva $(1 - p)$ percentual para si mesma.

A comunidade de v_k recebe energia de todos os seus nós membros, como v_k e v_r . (e) A comunidade realoca sua energia para seus membros com base em suas centralidades comunitárias. O bloco azul é a energia obtida pelos encontros diretos do nó de v_k , o bloco preto é a energia realocada pelas comunidades do nó v_k [Li et al., 2017].

O algoritmo CASASR proposto por Sepehr Keykhaie e Maryam Rostaei (2017) é um protocolo de roteamento com características sociais que utiliza o processo de Poisson para encontrar a probabilidade de contato entre os membros da comunidade do nó de destino, o protocolo utiliza uma função para selecionar um nó de retransmissão com laços sociais mais fortes e maior probabilidade encaminhar a

mensagem para o destino, este algoritmo também propõe um mecanismo de controle de congestionamento para evitar nós congestionados na rede, outra propriedade do algoritmo é o gerenciamento de energia que é feito através de um controle que ajusta o nível de energia disponível para retransmissão de mensagens com base em informações sociais e individuais do nó [Keykhaie e Rostaei, 2017].

Em outro trabalho Jiagao Wu et al. (2016) propõem um algoritmo otimizado de limites de cópia com eficiência energética modelado para cenários de roteamento em comunidades com egoísmo social, o protocolo tem objetivo de melhorar o desempenho de roteamento epidêmicos. Utiliza equações diferenciais ordinárias, sua otimização baseia-se no método complexo de Box⁵ sob DTN social [Wu et al., 2016].

Utilizando outra perspectiva, Cristian Chilipirea et al. (2013) propõem uma customização do algoritmo de roteamento baseado em rede social *BUBBLE Rap* com o objetivo de aumentar o tempo de vida útil da rede. O protocolo proposto combina encaminhamento socialmente consciente com a otimização do consumo de energia, os autores denominaram o novo protocolo com o nome “*Energy-Aware BUBBLE Rap*” [Chilipirea et al., 2013].

No algoritmo “*BUBBLE Rap*” original, quando dois nós se encontram, eles usam uma função de utilidade para decidir por cada mensagem para encaminhá-la através do nó vizinho. A decisão inclui as classificações global e local dos dois nós que se encontram. Se um nó tiver uma classificação mais alta, a probabilidade de receber mais pacotes aumenta, já que sua centralidade é maior.

Na abordagem “*Energy-Aware BUBBLE Rap*” a função de permitir que ela também diminua foi estendida se o nó vizinho tiver recursos de energia insuficientes para suportar a transferência de mensagens. À medida que a energia diminui, a probabilidade de que esse nó seja um portador bem-sucedido também diminui.

Para tornar a decisão de roteamento consciente, foi introduzido o custo de usar recursos específicos (modelo de bateria), com as seguintes propriedades: a função de utilidade diminui com o aumento do consumo de energia, portanto um nó não aceitará mensagens se a bateria estiver esgotada. Dessa forma, a decisão de roteamento deve ser capaz de reagir rapidamente a mudanças nas condições de energia.

⁵ O método complexo de Box é um algoritmo de pesquisa direta amplamente aplicado, desenvolvido para solucionar problemas de otimização não-linear restritos e originado do método simplex.

Com essas propriedades os autores detectaram que existe um limite para esta função de utilidade que é o nível de energia de um nó. Sob essa suposição o processo de roteamento com reconhecimento de energia foi modelado de maneira semelhante a um processo evolucionário em um circuito RLC elétrico.

Por sua vez o trabalho feito por Nelson Machado (2013) propõe um algoritmo chamado SOCLEER (*Social-based Energy-Efficient Routing Protocol*) que em vez de enviar uma mensagem imediatamente para o nó conectado, que possui maior centralidade do que o nó de origem como ocorre no Bubble rap, isso passar a ser feito durante o tempo t , quando o gráfico de contato é formado, a centralidade é calculada e armazenada numa matriz de centralidades globais e locais, a partir daí essa matriz é classificada de forma decrescente e, em seguida, é chamado um método implementado para desenhar uma interface de 1 a 10 de n nós com a maior centralidade, dentro de uma faixa existente na matriz, para então encaminhar a mensagem adiante.

Essa redistribuição da carga de tráfego de mensagens alivia as ocorrências dos nós preferidos gerando uma economia de energia dos nós devido a essa redistribuição [Machado, 2013].

3.2 Síntese dos Resultados

Dividimos os trabalhos em tópicos conforme o fluxograma apresentado na Figura 15. A Tabela 3 resume os aspectos fundamentais dos trabalhos relacionados neste capítulo, apresentando autor, algoritmo proposto e/ou arquitetura, estando separado por tipo de apresentação, tais quais, alguns algoritmos relevantes que utilizam roteamento social, experimentos envolvendo impacto do consumo em OppNets e propostas de algoritmos com roteamento social que incluem o consumo de energia no seu projeto de algoritmo.

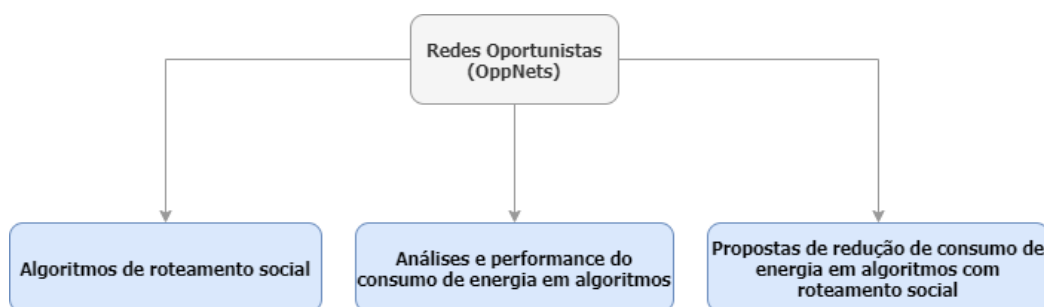


Figura 15: Fluxograma dos trabalhos relacionados por tópicos

Tabela 3: Resumo dos Trabalhos Relacionados.

Algoritmos de roteamento social	
Friendship and Selfishness Forwarding - FSF (Souza, 2019)	Este algoritmo classifica os laços sociais entre o destinatário da mensagem e o nó candidato a receber a mensagem (relay), posteriormente, utilizando funções logísticas, o algoritmo proposto avalia o egoísmo do nó relay para considerar os casos em que o nó receptor é egoísta seja porque seu dispositivo está com limitações de recursos, ou porque ele é racionalmente egoísta.
Grouping protocol and dissemination of messages by interests - GPDM (Neves, 2018)	Protocolo de disseminação de mensagens em OppNets, usando os interesses dos nós da rede como fator social, para a formação de níveis de repasse de mensagens.
Int-Three (Xia et al., 2016)	Utiliza características de interesse dos nós para a formação de comunidades através do cálculo da densidade da comunidade e o laço social entre os nós para eleição de nós classificados como bons repassadores. Tem o objetivo de aumentar a eficiência do repasse e a taxa de entrega de mensagens.
Bubble Rap (Crowcroft et. Al., 2011)	Utiliza a estrutura social contida na mobilidade dos nós, histórico de contatos, para a formação de comunidades entre eles, eleição de hubs de comunicação e repasse das mensagens. É considerado o primeiro algoritmo de roteamento em OppNets baseado em características sociais.
SocialCast (Costa et al. 2008)	Utiliza o interesse dos nós receptores para repassar as mensagens e restringir o acesso a mensagem com o objetivo de aumentar a eficiência de repasse e a diminuição do overhead do buffer dos nós da rede.
Socio-Aware Overlay (2007)	Utiliza a estrutura de um sistema de distribuição de conteúdo, que é formado através da detecção de comunidades, onde os nós mais centrais formam uma camada chamada Socio-Aware Overlay.
Análises e performance do consumo de energia em algoritmos de roteamento	
Efeitos do egoísmo no consumo de Energia em redes oportunistas uma avaliação (Loudari e Benamar, 2019)	O trabalho tem como enfoque principal a avaliação de performance de protocolos em ambientes com nós egoístas, tendo como principal avaliação o impacto do egoísmo no consumo de energia relacionado aos protocolos de roteamento em OppNets.
Arquitetura de Simulação (Silva et al., 2012)	Módulo de Energia proposto na arquitetura do Simulador <i>The ONE</i> para analisar o consumo de energia de um nó de rede em protocolos de roteamento propostos para DTN, o modelo tem uma opção de recarga da bateria, que torna o ambiente de simulação mais realista para o estudo do impacto do consumo de energia.
Propostas de redução de consumo de energia em algoritmos com roteamento baseado em aspecto social	
Energy Efficient Inactive Node Detection Based Routing Protocol for Delay Tolerant Network. (Qaisar Ayub e Sulma Rashid, 2020)	Propõe um protocolo de roteamento baseado em detecção de nó inativo com eficiência energética para redes tolerantes a atrasos (<i>IND</i>). O protocolo utiliza a mobilidade dos nós para detectar nós inativos. Além disso, uma política de gerenciamento de buffer inteligente foi projetada para detectar e descartar as mensagens destinadas a nós mortos.
SEBAR (Fan Li et al., 2017)	Propõe a métrica energia social que é calculada com base na estrutura da comunidade e compartilhada por membros da comunidade. Nesse modelo baseado na lei da física de partículas os nós ao se encontrarem geram uma colisão capaz de gerar uma certa quantidade de energia que é compartilhada conforme a centralidade de sua comunidade.
CASASR (Keykhaie et al., 2017)	Protocolo de conscientização social e egoísmo consciente, que usando características sociais e consciência de nós egoístas para decidir se deve realizar retransmissão de mensagens selecionando a melhor forma de retransmissão e o comportamento egoísta social e individual dos nós para ajustar o nível de energia disponível para a retransmissão de mensagens.
Energy-Efficient Routing in Multi-Community DTN with Social Selfishness Considerations (Wu et al., 2016)	Os autores propõem um algoritmo para encontrar o limite ideal de cópias em várias comunidades utilizando equações diferenciais ordinárias visando melhoria na eficiência energética em cenários de redes com comunidades e nós egoístas.
ECF (Socievole et al., 2015)	Faz a comparação da centralidade do nó durante um encontro entre dois nós, calculando seu valor de centralidade levando em consideração seu nível de energia atual. Dessa forma o valor da centralidade do nó vai caindo exponencialmente caso o nível de energia esteja baixo evitando transmissões em nós mais centrais, mas com energia limitada.
Energy-Aware BUBBLE Rap (Chilipirea et al., 2013)	Decisão de roteamento baseada nas mudanças nas condições de energia, para tornar a decisão de roteamento consciente, foi introduzido o custo de usar recursos específicos (modelo de bateria), a função de utilidade diminui com o aumento do consumo de energia, portanto um nó não aceitará mensagens se a bateria estiver esgotada.
SOCLEER: Social-based Energy-Efficient Routing Protocol (Machado Jr; Nelson, 2013)	Proposta de protocolo de roteamento oportunista com base social, com objetivo de aliviar a frequente repetição de nós preferidos para repasse de mensagens, distribuindo a carga de participação destes nós no envio de mensagens com outros de características sociais semelhantes, bem como reduzir seu consumo de energia.

Embora existam diversas abordagens utilizadas na comunidade acadêmica voltadas para a arquitetura DTN, ainda são apresentadas questões em aberto para serem exploradas através de pesquisas, destacam-se os problemas com o roteamento de mensagens, gerenciamento de buffer e gerenciamento de energia.

Na literatura, há uma grande quantidade de propostas de protocolos de roteamento. Em muitos deles, aspectos sociais são usados para construir esses protocolos. Como redes oportunistas são compostas por humanos, é razoável afirmar que os modelos de redes sociais são uma alternativa melhor para a construção de protocolos de roteamento social [Zhu et al., 2013].

É possível também identificar na literatura protocolos que enfatizam a importância de características sociais no projeto de seus algoritmos. Esses algoritmos utilizam características como amizade, egoísmo, centralidade, popularidade, similaridade, comunidades entre outras. Novas políticas de encaminhamento de mensagens podem ser projetadas estendendo-se aos nós da rede as características de comportamento pertencentes aos usuários que carregam dispositivos na rede, adotando esses parâmetros para o repasse de mensagens [Lee et al., 2009].

Muitos protocolos de roteamento foram projetados para aumentar as métricas de desempenho usando características sociais no cerne de seus protocolos. Embora essas estratégias possam trazer melhorias, percebemos que ainda faltam estudos que analisem a relação entre o encaminhamento de mensagens e o consumo de energia.

Os eventos como descoberta de nós vizinhos e a troca de mensagens na rede pode aumentar o consumo de energia e diminuir o desempenho geral da rede devido aos nós ficarem sem recursos energéticos. Assim, à medida que os recursos dos dispositivos são utilizados, a capacidade de energia diminui, pois os dispositivos que compõem a rede são equipados com bateria de energia que tem uma duração limitada.

Como mencionado anteriormente, restrições de recursos são aspectos importantes em redes oportunistas, pois são responsáveis por manter ou não a interoperabilidade da rede, e devem ser levados em consideração durante a fase de projeto de novos protocolos. No entanto, a revisão bibliográfica mostrou que são poucos os estudos mais aprofundados que consideram o consumo de energia e seus possíveis impactos no repasse de mensagens em redes oportunistas, uma vez que

esse fator pode influenciar consideravelmente as taxas de entrega de mensagens e o desempenho da rede.

Em redes oportunistas, é essencial que os nós atuem como colaboradores na comunicação, cada nó carregando uma réplica da mensagem original aumenta a probabilidade de entrega ao nó destinatário, mas este paradigma pode consumir rapidamente os recursos dos nós da rede incentivando que tenham um comportamento egoísta [Soares, 2018].

Uma proposta promissora utilizando roteamento social foi desenvolvido por Neves (2018), para a disseminação de mensagens em redes oportunistas. A proposta utiliza o conceito social de formação de comunidades ou grupos sociais formados por pessoas que interagem vivendo em uma determinada localização.

Partindo do pressuposto que é provável que membros da mesma comunidade interajam mais com membros da própria comunidade do que com uma pessoa escolhida aleatoriamente. Sendo assim as comunidades podem refletir o relacionamento social entre as pessoas e isso é explorado para retransmitir mensagens e melhorar a chance de entregas [Neves, 2018].

O algoritmo usa o interesse que os membros da comunidade possuem por determinados assuntos então utiliza técnicas de aprendizagem de máquina para agrupar nós por grupos de interesse utilizando essas características sociais previamente estabelecidas. E dessa maneira, as mensagens são encaminhadas através dos grupos formados até chegar ao nó destino uma vez que grupos com interesses comuns tendem a se encontrar e trocar mensagens mais vezes.

No entanto, embora o protocolo GPDMI (*Grouping and dissemination of messages by interests*) tenha conseguido alcançar resultados promissores, é importante ressaltar que o autor não avaliou o protocolo com relação ao consumo de energia dos nós e dependendo das limitações de seus recursos, incluindo o nível de energia, o nó pode tornar-se egoísta.

3.3 Considerações finais

Neste capítulo descrevemos e analisamos os trabalhos relacionados observando os métodos e técnicas utilizados no roteamento social dos nós em OppNets, assim como experimentos relevantes sobre o impacto do consumo de energia em ambientes distintos bem como as propostas de gerenciamento de energia voltadas para os cenários de mobilidade e roteamento com características sociais em OppNets.

Embora alguns dos trabalhos relacionados utilizem mecanismos para reduzir o consumo de energia em OppNets, como é o caso dos trabalhos do Socievole et al. (2015), Machado (2013), Ayub e Rashid (2020), Fan Li et al (2017) que são promissores, nenhum deles leva em consideração estudar o comportamento de usuários reais e a forma como utilizam a bateria de seus dispositivos móveis.

Aprender sobre esse comportamento e usar esse histórico pode ajudar a desenhar soluções melhores que atenuem os impactos causados pelo consumo de energia nas métricas da rede. Grande parte dos trabalhos citados utiliza como mecanismo o ajuste dos níveis de energia dos nós durante a retransmissão na rede associado a centralidade ou popularidade desses nós.

Além disso nenhum dos trabalhos encontrados utiliza técnicas de aprendizagem de máquina para o agrupamento de nós por nível de energia levando em consideração o comportamento dos usuários na utilização de suas baterias, técnicas de aprendizagem de máquina são bastante utilizadas em problemas onde a natureza dos dados muda constantemente, encontrar padrões nesses dados pode auxiliar a usar novas abordagens e a ter predições mais assertivas.

Método

Neste capítulo, descrevemos o mecanismo desenvolvido como abordagem para otimizar os impactos do consumo de energia durante repasse de mensagens em OppNets. Identificamos os principais desafios relacionados a este trabalho, seguido por uma descrição detalhada do modelo utilizado neste trabalho e suas técnicas.

4.1 Mecanismo Proposto (eGPDMI)

O mecanismo proposto neste trabalho tem como premissa evitar o desperdício de recursos da rede ao tentar transmitir mensagens de nós com pouco recurso de energia disponível, com forte tendência a se tornarem nós egoístas, levando em consideração estudos realizados que indicam que as limitações por recursos de energia tendem a tornar os nós egoístas e comprometer as taxas de entrega neste tipo de rede e a sua interoperabilidade⁶.

Neste sentido, este trabalho propõe a otimização do consumo de energia através do agrupamento dos nós da rede por níveis de energia, com o objetivo de evitar desperdício de recurso durante tentativas de repasse de mensagens na rede. É feita uma customização no algoritmo GPDMI proposto por Neves (2018), incluindo a classificação dos níveis de energia dos nós e o seu agrupamento, além de condições específicas para verificação do nível de carga de energia dos nós durante o repasse de mensagens.

A proposta se baseia em repassar mensagens através de grupos formados por níveis de energia e interesses comuns e a validação dos limites (*threshold*) de carga de energia durante o repasse das mensagens. A definição dos limites de carga de energia, utilizados para tomada de decisão de repasse de mensagens, possui característica simplesmente empírica, podendo assumir diversas variações, onde, $E = [cargaMinE, cargaMaxE]$, $E \subset [0, 100]$.

Assim, por exemplo, se $E = [50, 100]$, nossa estratégia constrói um cluster onde cada nó dentro do cluster tem nível de energia contido entre os intervalos de 50% a

⁶ É a capacidade de um sistema se comunicar utilizando linguagens e protocolos comuns.

100%. Esta condição é utilizada como fator para decidir não encaminhar uma mensagem para um nó que não está neste cluster.

Na Figura 16, apresentamos detalhadamente os processos executados pelo mecanismo *eGPDMI*, as formas geométricas que estão em amarelo representam as customizações que foram feitas no mecanismo original desde a modelagem de atributos, passando pelos processos de agrupamento e repasse de mensagens pelos grupos. As formas geométricas em azul são os processos que já existiam antes da customização.

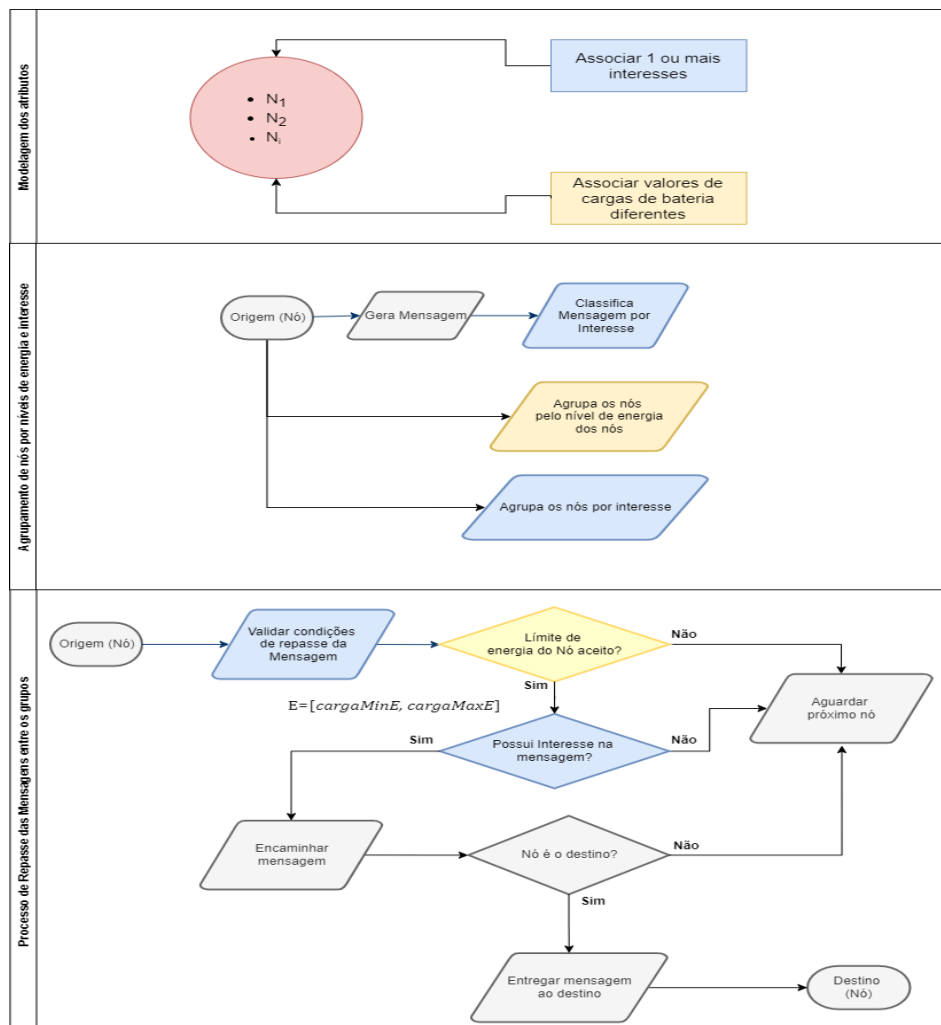


Figura 16: Fluxograma Geral de processos do *eGPDMI*

As seguintes etapas são processadas pelo algoritmo:

- 1) **Na modelagem dos atributos:** Classificação do conjunto de i nós dado por $N = \{N_1, N_2, \dots, N_i\}$, onde, é associado para cada $N = \{i\}$ um ou mais interesses do conjunto formado pelas categorias de interesse $I = \{I_1, I_2, \dots, I_i\}$.

- 2) **Na geração da mensagem:** classificação da mensagem conforme categorias de interesses, por exemplo, $I = \{\text{Arquitetura de redes, Serviços, Topologia de redes}\}$, onde, a mensagem $M1 = I[2]$ significa que a categoria de interesse da mensagem M1 é “Serviços”.

Os processos executados na etapa 1 e 2 já existiam no algoritmo original GPDMI. As etapas seguintes foram inseridas no mecanismo (*e*GPDMI).

- 3) **Na modelagem dos atributos:** Classificação do conjunto de i nós dado por $N = \{N_1, N_2, \dots, N_i\}$, onde, é atribuído de maneira pseudoaleatória um valor de energia que pode variar entre 0% e 100% de sua carga de bateria $E = \{E_1, E_2, \dots, E_i\}$.
- 4) **No processo de agrupamento:** Execução do agrupamento dos nós conforme os níveis de energia que foram atribuídos e execução dos agrupamentos de nós conforme as suas categorias de interesse formando assim grupos definidos por níveis de energia e interesse. Esse agrupamento é feito utilizando as técnicas de clusterização dos algoritmos E-M e K-Means abordados neste trabalho.
- 5) **No processo de encaminhamento da mensagem:** Execução do repasse de mensagens na rede verificando o nível de energia dos nós através do intervalo pré-definido dado por $E = [\text{cargaMinE}, \text{cargaMaxE}]$ associado a verificação do interesse do nó pela mensagem. A Figura 17 representa os grupos de energia e interesse formados através do processo de agrupamento dos nós.

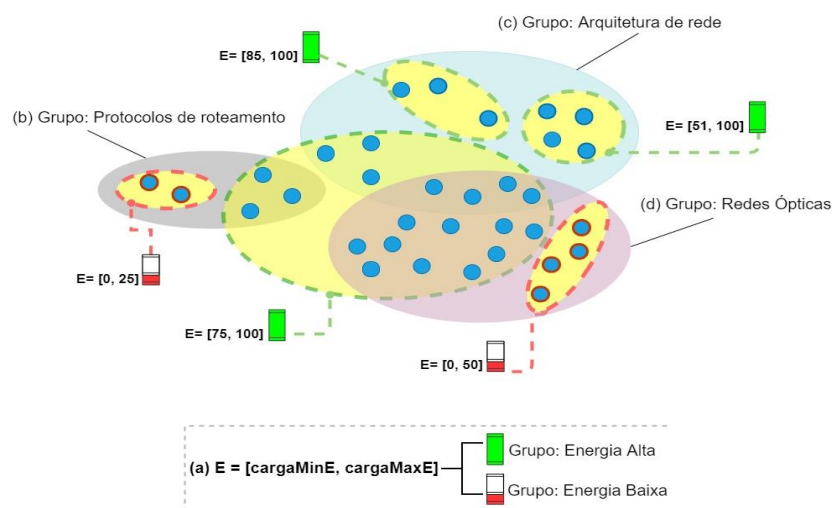


Figura 17: Formação de Grupos por níveis de energia e Interesses Comuns

Podemos observar na Figura 17 o agrupamento dos nós por níveis de energia e tópicos de interesse, os nós que estão na área de cor amarela tracejada representam os grupos de nós formados conforme a variação de carga de energia definida pelo intervalo $E=[cargaMinE, cargaMaxE]$ sendo agrupados conforme a letra (a) como níveis de energia consideradas “Alta” e “Baixa” para repasse de mensagens.

Os demais agrupamentos formados pelas letras (b), (c) e (d) ilustram os grupos de interesse desses nós. Dessa forma conseguimos visualizar aqueles nós que pertencem a grupos que possuem restrições de recurso de energia em suas baterias e por isso podem se tornar nós naturalmente egoístas.

4.1.1 Algoritmo de repasse de mensagens entre os nós

Como citado na seção anterior, o mecanismo proposto encaminha as mensagens utilizando como decisão a variação de intervalo de carga de energia dos nós definida por $E=[cargaMinE, cargaMaxE]$, e na sequência analisa os interesses dos nós pela mensagem. A seguir, apresenta-se o pseudocódigo da política de encaminhamento de mensagens da proposta.

Algorithm 1: Algoritmo de decisão para repasse de mensagem	
1:	procedure GroupRouter(ClusterEnergyLevels[], ClusterInterest[], NoA, NoB)
2:	for $i=0$ to ClusterEnergyLevels.size() do
3:	<i>// Verifica se o nó pertence ao grupo de threshold de carga de energia definido pelo intervalo $E=[cargaMinE, cargaMaxE]$</i>
4:	if NoB == ClusterEnergyLevels[i] then
5:	for $i=0$ to ClusterInterest.size() do
6:	if NoB == ClusterInterest[i] then
7:	forwardMessage(NoB.buffer, NoA.message)
10:	end if
11:	end for
12:	end if
13:	end for
14:	finishContact (NoA, NoB)
14:	end procedure

Na linha 1, o algoritmo recebe o conjunto de nós que foram agrupados por níveis de energia e agrupados por interesses. Estes estão representados pelos vetores *ClusterEnergyLevels* e *ClusterInterest* respectivamente, temos ainda o nó de

origem ou encaminhador da mensagem representado pelo **Nó(A)** e o nó para o qual se pretende passar a mensagem, ou seja, o receptor **Nó(B)**.

O algoritmo percorre o vetor *ClusterEnergyLevels* e verifica se o nível de carga de energia do **Nó(B)** está no intervalo de limite (*threshold*) estipulado como condição para o repasse de mensagens, se esta validação for verdadeira o algoritmo verifica se o nó possui interesse na mensagem.

Se as duas validações forem atendidas o **Nó(A)** repassa a mensagem para o **Nó(B)**, estas execuções ocorrem entre as linhas 2 e 7 do pseudocódigo. Se uma das validações não for verdadeira a mensagem continua com o **Nó(A)** que encerra a conexão.

4.2 Método de Avaliação

Para avaliação de desempenho dos resultados da abordagem proposta utilizaremos as métricas de rede que denotam as equações descritas abaixo [Machado, 2013]:

- Taxa de entrega - definida como a razão entre a quantidade de mensagens entregues ao destinatário (*QME*) e a quantidade de mensagens criadas na origem (*QMC*). Conforme a equação 4.1 abaixo:

$$TaxaDeEntrega = \frac{QME}{QMC} \quad (4.1)$$

- Atraso médio de entrega - é a medição do intervalo de tempo médio entre o evento de envio e o recebimento das mensagens na rede. Conforme a equação 4.2.

$$AtrasoDeEntrega = \Delta t(M_{tn} - M_{t0}) \quad (4.2)$$

Onde:

M_{tn} = Instante de tempo de entrega da mensagem

M_{t0} = Instante de tempo de envio da mensagem

- Sobrecarga - é a diferença entre o número de mensagens repassadas (*QMRep*) e as mensagens entregues ao destino (*QMEnt*) dividido pela quantidade de mensagens entregues.

$$Sobrecarga = \frac{QMRep - QMEnt}{QMEnt} \quad (4.3)$$

Os pontos dos gráficos gerados para avaliação de desempenho estão acompanhados por intervalos de confiança de 95%, calculados conforme a equação 4.4.

$$IntervaloDeConfiança = \bar{y} \pm 1,96 \left(\frac{\sigma}{\sqrt{n}} \right) \quad (4.4)$$

Onde:

$\bar{y}$ é a média das respostas y_i após n repetições e σ corresponde ao desvio padrão das amostras, conforme a equação 4.5 e a constante 1,96 indica que 95% da área sob a curva de distribuição normal das amostras está dentro de 1,96 desvios-padrões da média.

$$\sigma = \sqrt{\sum_{i=1}^n \frac{(y_i - \bar{y})^2}{n - 1}} \quad (4.5)$$

4.3 Considerações finais

Neste capítulo foi descrito o mecanismo proposto e as métricas que serão utilizadas para análise de desempenho dos resultados. Na seção 4.1 foi apresentado o mecanismo proposto que chamamos de ϵ GPDMI, na seção 4.2 as métricas de avaliação de desempenho que serão usadas.

Projeto de experimentos

Neste capítulo, descrevemos o planejamento de um projeto de experimentos sobre o impacto do nível de energia em redes oportunistas. Identificamos através de pesquisa, dados importantes para avaliar sobre a relação entre a cooperação dos nós na rede e o nível de bateria de seus dispositivos.

5.1 Processo de planejamento do projeto de experimentos

As métricas de rede podem degradar se um nó deixar as operações de rede, pois há menos nós para armazenar as mensagens. Para testar esta hipótese, conduzimos uma pesquisa para avaliar duas questões: (i) o comportamento de usuários reais ao participarem de OppNets com um nível de energia crítico; (ii) a partir de qual nível de energia os usuários consideram participar da rede apenas se favorecendo (não aceitando mensagens para outros usuários).

Para responder a essas questões, foi aplicado um questionário no campus da UFAM no final do segundo semestre de 2019 com 351 participantes de diversos cursos. A Tabela 4 descreve as duas questões principais abordadas para testar nossa hipótese:

Tabela 4: Nível de energia e questionário OppNets

Questões	Alternativas
(i) O que você faz se alguém lhe enviar uma solicitação de contato em rede oportunista enquanto o nível de bateria está crítico?	a) aceitar o pedido. b) não aceita o pedido. c) espera a bateria recarregar
(ii) Que nível de bateria do seu dispositivo faria você sair de uma comunicação oportunista?	a) abaixo de 10% b) entre 10% e 20% ... j) entre 91% e 100%

Os resultados do questionário mostram que 77,16% dos participantes não aceitariam o pedido de participação ou aguardariam o recarregamento das baterias. Os resultados colaboram com a hipótese nula de que os nós podem deixar as operações de armazenamento de dados na rede para encaminhá-los posteriormente.

Analizamos o comportamento de cooperação usando a segunda questão como um guia. Descobrimos que 68,10% dos participantes não cooperariam mais na

comunicação se a bateria fosse inferior a 30%. E apenas 19,65% dos participantes cooperariam na comunicação se seu nível de energia estiver acima de 50%.

Esse resultado reforça a ideia de que nós com baixa energia não cooperariam com o encaminhamento da rede e essa análise nos permiti escolher um projeto apropriado de experimentos definindo um conjunto adequado de fatores de investigação para avaliar os impactos da cooperação dos nós com base no nível de energia de seus dispositivos.

Uma avaliação de desempenho deve compreender dois componentes principais: o projeto de experimentos, que se refere ao processo de planejamento dos experimentos, para que os dados possam ser coletados de forma viável para análise estatística e a análise estatística real dos dados [M. W. Totaro e D. Perkins, 2005].

Para abordar o componente de projeto de experimentos, aplicamos o projeto fatorial, técnica cujo projeto consiste em dois ou mais fatores como entrada, e valores específicos atribuídos a cada fator, chamados de níveis. Em resumo, cada fator representa os parâmetros da rede que influenciam significativamente as métricas de desempenho. Além disso, os projetos fatoriais podem responder como as interações entre os parâmetros podem afetar as métricas de desempenho, classificando os fatores em ordem de impacto.

A abordagem básica para o projeto fatorial é selecionar um conjunto de k parâmetros e determinar dois níveis opostos para cada parâmetro -1 e 1 , representando os níveis inferior e superior de cada fator. Todas as combinações entre parâmetros são utilizadas em simulações, para um total de 2^k experimentos. As variáveis codificadas são utilizadas em um modelo de regressão não linear construído para representar os efeitos dos fatores e de todas suas possíveis interações na variável resposta, dada pela equação 5.1:

$$y = q_0 + q_A^{x_A} + q_B^{x_B} + q_C^{x_C} + q_D^{x_D} + q_{AB}^{x_A x_B} + q_{AC}^{x_A x_C} + \dots + q_{BC}^{x_B x_C} + \dots + q_{ABCD}^{x_A x_B x_C x_D} \quad (5.1)$$

Como estamos interessados na probabilidade de entrega, essa será nossa variável resposta y , definimos os fatores que estão apresentados na Tabela 5 como: densidade do cenário, limite de bateria, tamanho do buffer e geração de mensagens. O nível da bateria ou nível de energia é a energia restante no dispositivo do usuário. Esta energia pode ser usada para aplicações do usuário e, adicionalmente, para comunicação entre dispositivos próximos.

No entanto, quando esse recurso é baixo, assumimos que os nós da hipótese não podem cooperar na comunicação oportunística. Definimos a densidade do cenário como uma métrica para analisar como o cenário pode impactar nas simulações.

Tabela 5: Fatores de experimentos

Fator	Valor mínimo (-1)	Valor máximo (+1)
x1 - Densidade do cenário	Reality	Infocom
x2 - Limite de Bateria	75-100	0-100
x3 - Tamanho do Buffer	10M	100M
x4 - Geração de Mensagem	600, 3600	100, 1000

Determinamos duas faixas de bateria, cujo valor representa em qual nível da bateria o nó continua a cooperar na rede, transportando mensagens de outros nós. No valor mínimo (entre 75% e 100%), os nós cooperam se seus níveis de bateria estiverem nesta faixa (baixa cooperação) e, por outro lado, no valor máximo (0% a 100%), os nós carregam mensagens de outros nós independente do seu nível de energia disponível (alta cooperação).

Como o cenário precisa ser semelhante aos cenários do mundo real, escolhemos um cenário esparsa (Reality) e um cenário denso (Infocom) com informações extraídas de experimentos reais. Para análise, investigamos o efeito na métrica de entrega, cujo valor está na faixa [0, 1]. Os fatores da tabela foram analisados em 16 experimentos. Na execução do experimento, utilizamos um algoritmo de roteamento epidêmico, e posteriormente para comparação foi utilizada o mecanismo proposto neste trabalho, a variável resposta y usada foi a probabilidade de entrega na rede.

Os resultados para a métrica usando o roteamento Epidemic é apresentado na Tabela 6. Estes resultados apresentam a investigação do efeito da faixa do nível da bateria sobre a probabilidade de entrega, enquanto a Tabela 7 descreve o percentual de variação apontado pelo planejamento fatorial.

Tabela 6: Interações entre fatores com métrica de resultado da probabilidade de entrega usando epidemic

#	x1	x2	x3	x4	y
1	-1	-1	-1	-1	0.0246
2	-1	-1	-1	1	0.2308
3	-1	-1	1	-1	0.4757
4	-1	-1	1	1	0.856
5	-1	1	-1	-1	0.035
6	-1	1	-1	1	0.0682
7	-1	1	1	-1	0.5025
8	-1	1	1	1	0.8496

9	1	-1	-1	-1	0.0382
10	1	-1	-1	1	0.2308
11	1	-1	1	-1	0.4641
12	1	-1	1	1	0.8596
13	1	1	-1	-1	0.0259
14	1	1	-1	1	0.2393
15	1	1	1	-1	0.4719
16	1	1	1	1	0.9029

Dessa forma temos como exemplo:

$$y_1 = q_0 - q_A - q_B - q_C - q_D + \dots + q_{ABD} + q_{BCD} - q_{ABCD}$$

$$y_3 = q_0 - q_A - q_B + q_C - q_D - \dots - q_{ABD} - q_{BCD} + q_{ABCD}$$

$$y_8 = q_0 - q_A + q_B + q_C + q_D + \dots - q_{ABD} + q_{BCD} + q_{ABCD}$$

$$y_{16} = q_0 + q_A + q_B + q_C + q_D + \dots + q_{ABD} + q_{BCD} + q_{ABCD}$$

Depois resolve-se o sistema de equações para calcular os coeficientes em função das respostas, como nos exemplos:

$$q_0 = \frac{1}{16} (y_1 + y_2 + y_3 + y_4 + y_5 + y_6 + \dots + y_{11} + y_{12} + y_{13} + y_{14} + y_{15} + y_{16})$$

$$q_A = \frac{1}{16} (y_1 - y_2 - y_3 - y_4 - y_5 - y_6 - \dots + y_{11} + y_{12} + y_{13} + y_{14} + y_{15} + y_{16})$$

$$q_B = \frac{1}{16} (y_1 + y_2 + y_3 + y_4 + y_5 + y_6 + \dots + y_{11} + y_{12} + y_{13} + y_{14} + y_{15} + y_{16})$$

⋮

$$q_{ABD} = \frac{1}{16} (y_1 - y_2 - y_3 + y_4 + y_5 \dots - y_{11} - y_{12} + y_{13} + y_{14} + y_{15} + y_{16})$$

Para calcular o percentual da variação utiliza-se a soma dos quadrados total (SS_T), a soma dos quadrados de todas as respostas dos experimentos (SS_Y) e as somas de quadrados que atribuem porções de variação a fatores ($SS_A, SS_B, \dots$), e a interações ($SS_{AB}, SS_{AC}, SS_{AD}, SS_{ABC}, \dots$).

A importância de um fator é medida pela proporção da variação total na resposta que é explicado pelo fator, conforme o valor da variação resposta. A variação de amostra de y pode ser calculada da seguinte forma, como representado na equação 5.2.

$$SS_y = \frac{\sum_{i=1}^n (y_i - \bar{y})^2}{n - 1} \quad (5.2)$$

Onde $\bar{y}$ é a média das respostas de todos os experimentos, o numerador do lado direito da equação acima é chamado de variação total de y , então temos a soma dos quadrados total, conforme equação 5.3.

$$SS_T = \sum_{i=1}^n (y_i - \bar{y})^2 \quad (5.3)$$

A equação 5.4 descreve SS_x que podem ser atribuídas aos fatores e interações.

$$SS_x = 2^k \cdot (q_x)^2 \quad (5.4)$$

Temos então a equação 5.5 que calcula qual o percentual da variação da probabilidade de entrega está atribuído a cada um dos fatores:

$$\text{Porção de variação} = \frac{SS_x}{SS_T} \quad (5.5)$$

Embora o resultado do impacto da cooperação tenha sido estudado anteriormente por Gupta et al. (2011), este é o primeiro estudo que aplicou o planejamento fatorial para avaliar o impacto do nível de energia na probabilidade de entrega em comparação com outros parâmetros do OppNets.

O fator de intervalo da bateria tem um impacto mais significativo na probabilidade de entrega. O cenário e o fator conjunto composto pela autonomia da bateria e pelo cenário explicam 18% e 3% da variação. Assim, a quantidade de energia disponível do nó necessária para cooperar é o fator mais significativo que merece uma análise mais aprofundada.

Tabela 7: Avaliação fatorial epidemic

Porção de Variação	
msgGer	0,14%
tamanhoBuffer	0,03%
tamanhoBuffer, msgGer	0,12%
Intervalo da Bateria	76,96%
intervaloBateria, msgGer	0,10%
intervaloBateria, tamanhoBuffer	0,20%
intervaloBateria, tamanhoBuffer, msgGer	0,05%
Cenário	18,47%
cenário, msgGen	0,27%
cenário, tamanhoBuffer	0,09%
cenário, tamanhoBuffer, msgGer	0,26%
Cenário, intervaloBateria	3,15%
cenário, intervaloBateria, msgGer	0,02%
cenário, intervaloBateria, tamanhoBuffer	0,09%
cenário, [...] msgGer	0,06%

A Figura 18 apresenta a correlação entre o limite da bateria e a densidade do cenário em relação à métrica de probabilidade de entrega.

Quando a densidade do cenário é maior (mais contatos) e a faixa da bateria está no intervalo $E = [0,100]$, cooperação total, a probabilidade de entrega é maior do que no cenário esparsos ou a faixa da bateria está no intervalo de $E = [75,100]$.

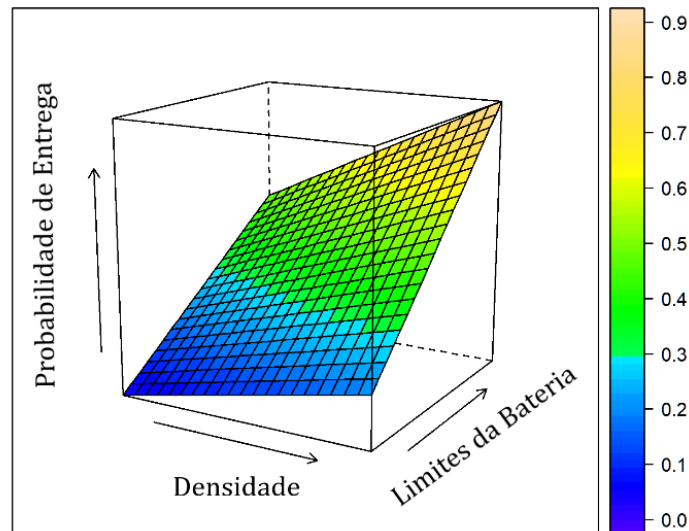


Figura 18: Impacto da densidade do fator cenário e do limite da bateria na probabilidade de entrega usando o roteamento epidêmico

Agora executamos o experimento usando o algoritmo de roteamento do nosso mecanismo (ϵ GPDMI), e assim, avaliar a diferença entre os efeitos dos fatores na variável resposta y em ambos os protocolos. Os resultados para a métrica usando nossa proposta é apresentado na Tabela 8. Da mesma forma que o projeto fatorial anterior, estes resultados apresentam a investigação do efeito da faixa do nível da bateria sobre a probabilidade de entrega.

Tabela 8: Interações entre fatores com métrica de resultado da probabilidade de entrega usando nossa proposta

#	x1	x2	x3	x4	y
1	-1	-1	-1	-1	0.3398
2	-1	-1	-1	1	0.3338
3	-1	-1	1	-1	0.3711
4	-1	-1	1	1	0.2839
5	-1	1	-1	-1	0.3325
6	-1	1	-1	1	0.2986
7	-1	1	1	-1	0.3157
8	-1	1	1	1	0.3171
9	1	-1	-1	-1	0.8657
10	1	-1	-1	1	0.8484

11	1	-1	1	-1	0.8731
12	1	-1	1	1	0.8709
13	1	1	-1	-1	0.7388
14	1	1	-1	1	0.7070
15	1	1	1	-1	0.7313
16	1	1	1	1	0.6578

A Tabela 9 descreve o percentual de variação apontado pelo planejamento fatorial da proposta. Através do resultado podemos verificar que diferente do projeto anterior o cenário virou a variável mais influente da avaliação com 93% da variação.

Comparando as duas avaliações quando a faixa da bateria está no intervalo $E = [75,100]$ que é o pior caso o nosso mecanismo conseguiu superar razoavelmente a taxa de entrega de mensagens. Uma vez que a proposta melhorou a taxa de entrega para o limite de energia ele deixou de influenciar na entrega.

No entanto, analisando o resultado do intervalo $E = [0,100]$ a proposta perde para a avaliação fatorial do Epidemic.

Tabela 9: Avaliação fatorial da proposta

Porção de Variação	
msgGer	0,43%
tamanhoBuffer	0,01%
tamanhoBuffer, msgGer	0,04%
Intervalo da Bateria	3,237%
intervaloBateria, msgGer	0,0043%
intervaloBateria, tamanhoBuffer	0,03%
intervaloBateria, tamanhoBuffer, msgGer	0,02%
Cenário	93,66%
cenário, msgGer	0,00%
cenário, tamanhoBuffer	0,00%
cenário, tamanhoBuffer, msgGer	0,00%
Cenário, intervaloBateria	2,13%
cenário, intervaloBateria, msgGer	0,15%
cenário, intervaloBateria, tamanhoBuffer	0,08%
cenário, [...],msgGer	0,21%

A Figura 19 apresenta a correlação entre o limite da bateria e a densidade do cenário em relação à métrica de probabilidade de entrega.

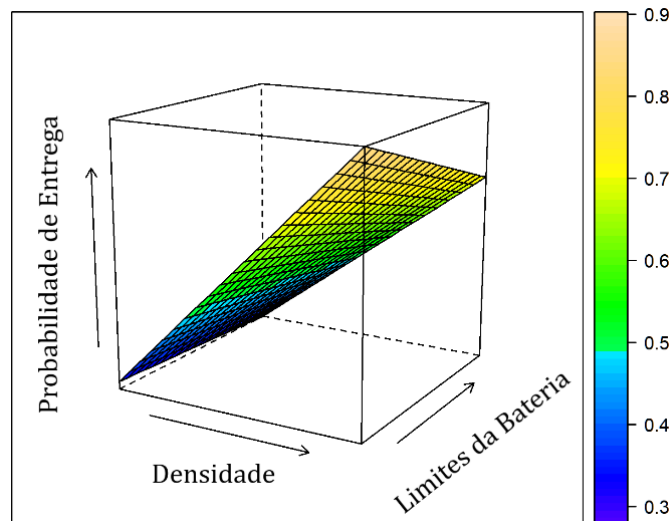


Figura 19: Impacto da densidade do fator cenário e do limite da bateria na probabilidade de entrega usando o roteamento proposto

5.2 Considerações Finais

Apresentamos os resultados de um projeto de experimentos fatorial, desde a pesquisa feita na UFAM sobre o comportamento de usuários de dispositivos móveis e recursos de energia em OppNets, até o planejamento e definição dos fatores de rede que foram investigados no projeto de experimentos. Esses resultados contribuíram para a investigação do efeito da carga do nível da bateria sobre a probabilidade de entrega, e para definirmos novas faixas de limite para serem investigadas.

Implementação e Avaliação

Neste capítulo, demonstramos os resultados experimentais e a avaliação de desempenho do mecanismo proposto neste trabalho. A técnica escolhida para avaliação do mecanismo foi a simulação, uma forma de imitar um processo ou operação do mundo real, utilizando a mobilidade extraída de ambientes reais para aumentar a similaridade com ambientes OppNets reais. A técnica de simulação é amplamente utilizada na literatura para avaliação de propostas em cenários de redes DTN, devido à grande complexidade de realização de experimentos *in loco*, contudo, a utilização de mobilidade coletada em ambientes reais, amplia a proximidade da técnica com cenários mais realísticos.

6.1 The Opportunistic Network Environment Simulator (The ONE)

O simulador *The ONE*⁷ é um ambiente de simulação *open source*, desenvolvido em Java, que permite simular os principais aspectos conceituados de uma rede DTN, como transferência de custódia, roteamento oportunista dos nós, protocolos de aplicação, consumo de energia além de oferecer suporte a modelos de mobilidade extraídos de ambientes reais (*traces*) [Keränen et al., 2009].

O simulador permite a modificação do código fonte e a parametrização de configuração dos nós da rede, o que nos permite montar diversos cenários. Como já mencionado anteriormente, possui suporte a modelos de mobilidade extraídos de ambientes reais.

A Figura 20 apresenta todo o processo de interação do simulador *The ONE*, para tornar as simulações complexas de DTN mais viáveis e compreensíveis, ele combina modelagem de movimentos, simulação de roteamento com visualização e geração de relatórios que são disponíveis através da GUI do simulador [Keränen et al., 2009].

⁷ Para fazer a utilização do simulador *The ONE* é necessário baixar um ficheiro do projeto no site <http://www.netlab.tkk.fi/tutkimus/dtn/theone/>

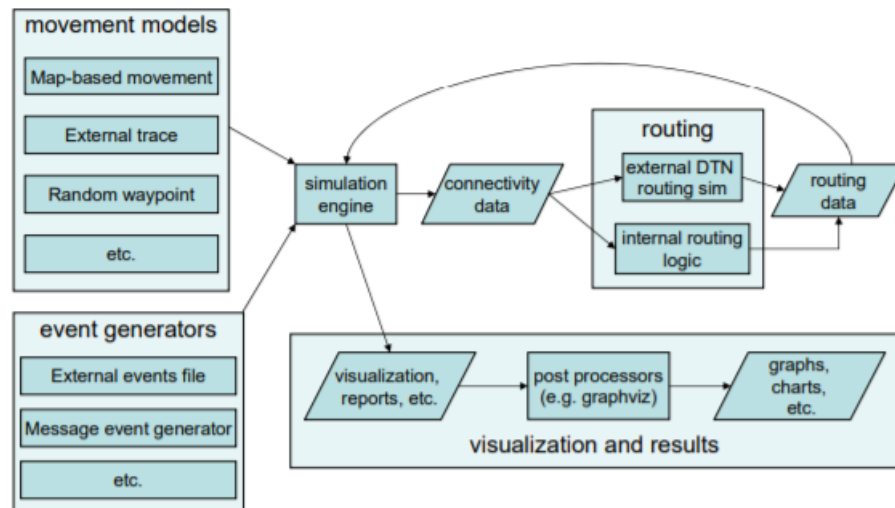


Figura 20: Overview of the ONE (Keränen et al., 2009)

Para criação de uma simulação é necessário realizar configurações básicas de alguns elementos para a construção do cenário proposto, são eles: configuração de grupos de nós, configuração de eventos, especificação de protocolos (epidemic, PROPHET, Spray and Wait e etc.).

Após a configuração destes elementos. O simulador The ONE fica apto para ser executado conforme demonstra a Figura 21 com a GUI do simulador em execução.

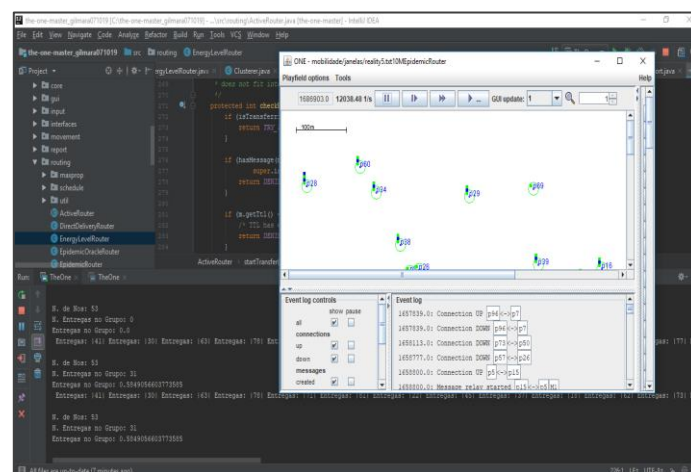


Figura 21: Tela de Execução The ONE - GUI do Simulador (O autor, 2021)

6.2 Modificações no The ONE Simulator

Para implementar a abordagem proposta foi necessário integrar algoritmos contidos na ferramenta WEKA com o simulador The ONE. A ferramenta WEKA possui diversos algoritmos de aprendizagem de máquina, inteligência artificial e mineração de dados, entre eles o algoritmo EM (*Expectation-Maximization*), que utiliza técnicas de

aprendizagem de máquina não supervisionado e o algoritmo de particionamento K-means. A integração ocorreu através da biblioteca “weka.jar”⁸ que foi importada no simulador permitindo a utilização destes algoritmos [Hall et al., 2009].

Os algoritmos de particionamento descobrem clusters diretamente por deslocamento iterativo de pontos entre subconjuntos [McLachlan e T. Krishnan, 2007]. O algoritmo Esperança – Maximização (EM) é um método que consiste em encontrar estimativas de número de populações na amostra, estimadas através de parâmetros que descrevem uma probabilidade de distribuição, clusterizando os dados estimando uma média amostral e sua variação.

Estes algoritmos analisam os padrões dos dados, no nosso caso, os níveis de energia e interesse dos nós para gerar novas distribuições e agrupamentos de dados. Algumas classes básicas do simulador foram modificadas e classes novas também foram desenvolvidas para montar o cenário dos experimentos.

A seguir as classes básicas do simulador que foram modificadas e as novas classes criadas:

- **Classes existentes que foram modificadas:**

MessageCreateEvent.java: esta classe foi modificada com o objetivo de gerar uma mensagem nova, executar as funções de classificação dos nós com níveis de energia e interesse na mensagem. Outra modificação realizada, foi a chamada do método de clusterização dos nós para o agrupamento dos mesmos por níveis de energia interesses.

ActiveRouter.java: esta classe foi utilizada para verificar a ocorrência de detecção do contato entre os nós e se ocorrem as condições necessárias para o repasse e recebimento de mensagens, incluindo as validações de energia e interesse dos nós. A modificação realizada, permite o repasse entre a localização do nó emissor e receptor dentro dos agrupamentos gerados na rede.

SimCenário.java: esta classe também pertence ao core do simulador the ONE e foi modificada para gerar valores pseudoaleatórios de carga de bateria e repassá-las aos nós da rede.

⁸ Biblioteca que permite utilizar algoritmos da ferramenta Weka que é um software livre para mineração de dados desenvolvido em Java, disponível em: <https://sourceforge.net/projects/weka/>.

DTNHost.java: é uma classe que faz parte do core do simulador the ONE e foi alterada com o objetivo de receber os valores de carga de bateria dos nós.

- **Classes Novas:**

EnergyLevel.java: esta classe foi criada para gerar o agrupamento dos nós por nível de energia, aqui ficam implementadas as solicitações para gerar os agrupamentos dos nós por níveis de energia, esta classe recebe os dados sobre os nós e seus níveis de energia e executa a clusterização.

InterestLevel.java: esta classe foi criada para gerar o agrupamento dos nós por interesse pela mensagem, aqui ficam implementadas as solicitações para gerar os agrupamentos dos nós conforme o interesse deles.

Clusterer.java: possui a implementação dos algoritmos de clusters contidos na library *weka.jar*. A biblioteca possui dois métodos distintos para cada algoritmo de cluster.

Os métodos são *evaluateInstance* e *getMeansClusterElements*, o primeiro método é utilizado para avaliar as informações que se assemelham dentro de um determinado conjunto de dados, processando todos os cálculos de acordo com o método de clusterização. O *getMeansClusterElements* retorna as instâncias ou grupos onde os nós foram inseridos.

6.3 Conjunto de dados de mobilidade utilizados

Para realização das simulações feitas neste trabalho, optou-se pela utilização de dois conjuntos de dados referentes à mobilidade humana chamados de Reality e Haggle-Infocom5.

As simulações em redes oportunistas requerem a utilização de dados de mobilidade carregados através de usuários e seus dispositivos em ambientes reais. Assim, neste trabalho os traces utilizados foram respectivamente Reality e Infocom5.

O trace **Reality**⁹ é derivado de um experimento realizado no período de 2004 a 2005 no Massachusetts Institute of Technology - MIT, utilizando um grupo de 100 usuários durante 246 dias. Foram disponibilizadas para a comunidade científica

⁹ Nathan Eagle, Alex (Sandy) Pentland, CRAWDAD dataset mit/reality (v. 2005-07-01), downloaded from <https://crawdad.org/mit/reality/20050701>, <https://doi.org/10.15783/C71S31>, Jul 2005.

informações de comunicação, proximidade, localização e informações de atividades destes usuários.



Figura 22: Trace Reality – Grafo de formação de contatos (Neves, 2018)

O trace **Haggle-infocom5**¹⁰ é um arquivo que contém informações de um experimento realizado durante uma conferência científica que aconteceu em Miami em 2005, na qual os participantes, tiveram seus contatos e mobilidades extraídos durante um período de 3 dias. O experimento contou com 41 participantes entre professores e estudantes.

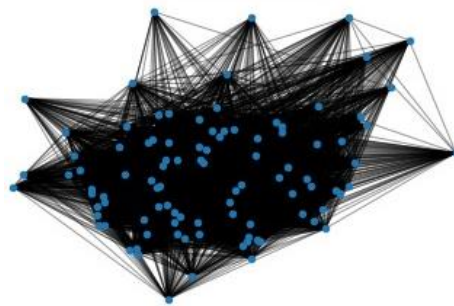


Figura 23: Trace Infocom5 - Grafo de formação de contatos (O autor, 2021)

Os algoritmos de roteamento utilizados nos experimentos e na validação da proposta foram os protocolos PROpHET e GPDMI, são baseados em replicações de mensagens.

6.4 Algoritmo PROpHET

O algoritmo de roteamento PROpHET (*Probabilistic Routing Protocol using History of Encounters and Transitivity*) foi proposto por Doria e Scheln (2003), é um dos

¹⁰ James Scott, Richard Gass, Jon Crowcroft, Pan Hui, Christophe Diot, Augustin Chaintreau, CRAWDAD dataset cambridge/haggle (v. 2009-05-29), traceset: imote, downloaded from <https://crawdad.org/cambridge/haggle/20090529/imote>.

principais algoritmos probabilísticos utilizados em DTN, esse protocolo de roteamento probabilístico utiliza históricos de encontros e transitividade.

Da mesma forma em que acontece no algoritmo de roteamento epidêmico, em que os nós em contato trocam suas listas com informações de cada nó, acontece também no algoritmo PROpHET. Porém o PROpHET possui uma informação adicional, tratando-se de uma probabilidade de cada nó entregar mensagens para um outro nó destino conhecido, assim tem-se: $P(a,b)$. O valor de $P(a,b)$ aumenta sempre que a e b se encontram.

Essa probabilidade é diretamente proporcional ao tempo transcorrido, então se a e b deixam de se encontrar frequentemente, $P(a,b)$ diminui. Esse tempo é controlado por uma constante k denominada constante de envelhecimento, que corresponde ao número de unidades de tempo transcorridas desde a última vez que a métrica foi atualizada [Luiz et al, 2011].

Então se um nó visita várias vezes uma determinada localidade, existe uma chance muito grande de visitar novamente tal localidade futuramente. Baseando-se nessa informação os autores criaram uma métrica de probabilidade chamada *delivery predictability*, ela quem indica a probabilidade de um nó entregar uma mensagem para outro nó destino [Souza, 2013]. Essa probabilidade é calculada dessa forma:

sendo P_{init} em $(0,1)$ uma constante de inicialização e y em $(0,1)$ uma constante de tempo.

$$P_{(a,b)} = \left\{ \begin{array}{ll} P_{(a,b)old} + (1 - P_{(a,b)old})P_{init} & \text{se } a \text{ e } b \text{ se encontram} \\ P_{(a,b)old} * y^k & \text{se } a \text{ e } b \text{ não se encontram por um } k \text{ tempo} \end{array} \right\} \quad (6.1)$$

Quando um nó recebe a lista do vizinho, ele calcula a probabilidade de entrega para cada uma das mensagens que ainda não possui. Em seguida, para cada mensagem, o nó compara a probabilidade indicada na sua lista com a probabilidade indicada na lista recebida do vizinho. Essa comparação é realizada para verificar qual dos dois nós possui a maior probabilidade de entrega [Luiz et al, 2011].

Esse algoritmo é bastante utilizado na avaliação de novos protocolos podendo ser utilizando em ambientes baseados em comunidades ou ambientes completamente aleatórios.

6.5 Algoritmo GPDMI (*Grouping and dissemination of messages by interests*)

O algoritmo de roteamento GPDMI foi proposto por Neves (2018). Sua função é utilizar o interesse social dos nós para disseminar mensagens em redes oportunistas. A proposta utiliza o interesse social dos nós para formar grupos que possuem interesses em comum por tipos de assuntos. O algoritmo tem objetivo de potencializar os repasses das mensagens e aumentar a eficiência das entregas na rede.

Para formar os grupos de interesses o autor utiliza a combinação de duas técnicas de agrupamento que são os algoritmos de clusterização Expectation-Maximization (EM) e K-means, o processo acontece da seguinte maneira:

1) É feita uma modelagem nos atributos, onde:

- a) Dado o conjunto $N = \{ N_1, N_2, \dots, N_i \}$, para cada nó i , pode-se associar um ou mais interesses.
- b) Cada interesse pode assumir um valor binário, onde 0 significa que um nó não possui o interesse e 1 significa que o nó possui aquele interesse. Então, o nó (N_1) possui interesses, como $N_1 < I_1, I_2, \dots, I_k >$. Assim, $N_1 < 0, 1 >$ representa que o nó N_1 está associado a dois interesses, I_1 e I_2 , mas só possui o interesse I_2 .

2) Na sequência é aplicada a técnica para o agrupamento dos nós, onde:

a) São formados níveis de agrupamentos utilizando os algoritmos EM e K-means.

b) O algoritmo K-means é utilizado para gerar os grupos de nós que possuem um único interesse, que pode ser o interesse da mensagem enviada na rede. Ou seja, são grupos de nós com interesse direto na mensagem gerada. Formado por nós que possuem um único elemento em seu conjunto de interesses.

c) O algoritmo EM é usado para gerar o agrupamento de nós que possuem vários interesses diferentes em seu conjunto de interesses, sendo um deles o interesse classificado na mensagem repassada na rede. Leva em consideração, além do conjunto de interesse dos nós destinatários, o conjunto de interesse dos nós vizinhos, podendo ser maior que um. O autor classifica como conjunto de nós com interesses indiretos na mensagem.

Existe uma interação maior dos nós da rede que possuem mais de um interesse em comum, por esse motivo o repasse das mensagens do nó de origem até o grupo de interesse do nó destinatário acontece através dos grupos que possuem interesse indireto na mensagem até que se chegue a um nó pertencente ao grupo de interesse direto do destinatário da mensagem, que é um grupo específico.

6.6 Caracterização do Cenário

Para avaliação de desempenho, simulamos nossos cenários usando o simulador The ONE. Foram utilizados dois conjuntos de dados referentes à mobilidade humana chamados de Reality e Haggle-Infocom5.

Os parâmetros de simulação são apresentados na Tabela 10. Cada cenário foi simulado 10 vezes, sendo os valores medidos para cada nó os valores médios de todas as execuções de simulação.

Calcular o número apropriado de repetições em cada instância, mostra como determinar este número para uma precisão e confiança desejadas, dada a média da distribuição e desvio padrão [Jain, 1991]. Seguindo o teorema do limite central que afirma em teorias de probabilidade que quando o tamanho da amostra aumenta, a distribuição amostral da sua média aproxima-se da distribuição normal [Heumann, 2016]. Os parâmetros de energia usados na simulação estão na Tabela 11.

Tabela 10: Parâmetros do conjunto de dados de mobilidade

Conjunto de dados experimentais	Reality	Infocom5
Tipo de dispositivo	Phone	iMote
Tipo de rede (<i>Interface</i>)	Bluetooth	Bluetooth
Duração dos traces (dias)	246	3
Número de encontros (<i>aprox.</i>)	54667	22459
Número de nós	97	41

Cada nó da rede recebeu de maneira pseudoaleatória ou pseudorrandômica valores de energia entre 0 e 100 que correspondem aos percentuais de carga de bateria de dispositivos móveis.

Na Tabela 11 foram definidos os limites de energia usados como critério para o repasse de mensagens e para observação em cada um dos cenários abaixo apresentados.

Tabela 11: Parâmetros de Energia

Parâmetros	Definições
Capacidade da Bateria (%)	$E = [0, 100]$
	$E = [25, 100]$
	$E = [50, 100]$
	$E = [75, 100]$
Roteamento	PROphet / GPDMI / Proposta
Tempo de simulação	$\approx 3,1$ dias (274884s) Infocom ≈ 10 dias (864000s) Reality
Tamanho de Buffer	10M

Sobre a clusterização dos grupos formados por energia e tópicos de interesses comuns foram usados a classificação da Tabela 12. Para definição dos grupos de energia assumimos que os valores que estão dentro dos limites variantes da capacidade da bateria na Tabela 11 retornarão grupos classificados como energia alta e o que estiver fora do limite como energia baixa, por exemplo, no caso $E = [50, 100]$ os nós que tiverem com carga de bateria menor que 50% estarão nos grupos de nós com energia baixa. No caso $E = [0, 100]$ consideraremos todos os níveis de energia para investigar o desempenho dos algoritmos.

A definição dos grupos de interesses utilizou como base um conjunto de assuntos retirados do trace Infocom5 em Barcelona no ano de 2006. Este evento também realizou coleta de dados de mobilidade dos usuários móveis participantes.

Na Tabela 12, descrevemos como os grupos foram criados. No total, usamos um conjunto de 30 tópicos de interesse diferentes, sendo este o número de interesses utilizado. Assim, $k = 30$ na técnica de agrupamento k-means.

Tabela 12: Formação dos Grupos de energia e interesse

Classificação	Energia	Interesse
Grupos	Grupo 1 – Energia Alta Grupo 2 – Energia Baixa	1. PowerControl 2. RoutingProtocols 3. Multicast/anycast 4. [...]
Roteamento	Proposta	Proposta / GPDMI

6.7 Método estatístico

A estatística experimental é a ciência que tem como objetivo estudar experimentos, englobando etapas como o planejamento, execução, coleta e análise dos dados experimentais e interpretação dos resultados obtidos [Costa Neto, 2002].

Neste trabalho aplicaremos o teste de análise de variância que é uma técnica estatística que visa, fundamentalmente, verificar se existe uma diferença significativa entre as médias e se os fatores exercem influência em alguma variável. Uma análise de variância permite que vários grupos sejam comparados a um só tempo, utilizando variáveis contínuas.

6.7.1 Anova

O método ANOVA fornece o embasamento teórico para que as médias dos tratamentos sejam comparados dentro de uma margem de erro aceitável. Este erro é denominado de nível “ α (alfa)” de probabilidade, podendo ser aplicado nos valores de 1%, 5% ou 10% [Bussab, 2013].

Para nossos experimentos adotaremos o nível de significância de 5%, onde, $\alpha = 0,05$ de modo que possamos verificar as hipóteses $H_0: \mu_1 = \mu_2 = \dots = \mu_c$ e H_1 : nem todas as médias são iguais.

Então as hipóteses adequadas ao experimento são:

H_0 – Hipótese Nula: a média de todas as populações são iguais.

H_1 – Hipótese Alternativa: nem todas as médias populacionais são iguais, ou seja, pelo menos uma média é diferente das demais.

A fórmula matemática utilizada para avaliar a variação do objeto estudado que se deseja medir é:

$$SQTotal = \sum_{i=1}^I \sum_{j=1}^J y_{ij}^2 - C, \text{ em que } C = \frac{(\sum_{i=1}^I \sum_{j=1}^J y_{ij})^2}{IJ} \quad (6.2)$$

$$SQTrat = \sum_{i=1}^I \frac{y_i^2}{j} - C \quad (6.3)$$

Onde $SQTotal$ é a soma total dos quadrados, y_{ij} é o valor na unidade j do tratamento i . E a soma dos quadrados dos resíduos ($SQRes$) é obtida pela diferença:

$$SQRes = SQTotal - SQTrat \quad (6.4)$$

Onde $SQRes$ ou variação dentro dos grupos (SQD) é chamada de variação de **Dentro**, é a função das diferenças existentes entre as repetições de um mesmo tratamento. E $SQTrat$ ou variação entre os grupos (SQE), também chamada variação **Entre**, é a variação existente entre diferentes tratamentos [Morettin et al., 2017].

O cálculo para os quadrados médios dos tratamentos ($QMTrat$) e dos resíduos ($QMRes$) podem seguir a fórmula:

$$QMTrat = \frac{SQTrat}{(I - 1)} \quad e \quad QMRes = \frac{SQRes}{(I(J - 1))} \quad (6.5)$$

Organizando as informações calculadas para a análise de variância temos a Tabela a seguir:

Tabela 13: Síntese dos cálculos da análise de variância

Causas de variação	Graus de liberdade	Soma de quadrados	Quadrados Médios	F calculado
Tratamentos	I-J	SQTrat	QMTrat	
Resíduos	I (J-1)	SQRes	QMRes	QMTrat / QMRes
Total	IJ-1	SQTotal		

As hipóteses no método ANOVAM são validadas comparando $F_{calculado}$ com o $F_{tabelado}$. De modo que a hipótese H_0 é rejeitada se $F_{calculado} > F_{tabelado}$ ou ainda $p\text{-valor} < \alpha$ (*alfa*) que é o valor de significância, dessa forma, existem evidências de diferença significativa entre pelo menos um par de médias de tratamentos. Caso contrário, não rejeitamos a hipótese H_0 , ou seja, não há evidências de diferença significativa entre os tratamentos.

Quando se utiliza mais de dois tratamentos e existe uma diferença significativa entre as amostras na análise de variância não podemos indicar qual é o melhor tratamento então é importante realizar os testes de comparação de média, como complemento para o estudo da análise de variância.

Sendo assim prevalecendo a hipótese alternativa H_1 após análise de variância, é assumido que existe diferenças significativas entre as amostras. Então é necessário identificar de quais amostras realmente existem diferenças. Para isso podemos aplicar o teste de comparação de Tukey.

6.7.2 Teste Tukey

O teste de Tukey baseia-se na diferença mínima significativa (dms). O cálculo estatístico do teste é dado da seguinte forma:

$$\Delta = dms = q_{k,gl,\alpha} \sqrt{\frac{QMRes}{r}}, \quad dms = q \cdot \sqrt{\frac{QMRes}{r}} \quad (6.6)$$

Onde, q é a amplitude total estudentizada, dado pela divisão entre a amplitude e o desvio padrão s , k é o número de tratamentos, gl são graus de liberdade, α é o nível de significância, $QMRes$ é o quadrado médio do resíduo, e r é o número de repetições. Se o contraste for maior que Δ , então as médias se diferem ao nível de significância ou $\Delta < \alpha$.

6.8 Discussão dos Resultados

Em cada sequência de simulação efetuada, utilizou-se para cada resultado coletado um intervalo de confiança de 95% para os parâmetros estimados. Para cada simulação executada, a métrica de interesse foi a probabilidade de entrega da mensagem, atraso médio e sobrecarga.

As simulações foram realizadas variando-se o limite de energia dos nós (faixa de bateria) conforme os parâmetros definidos na Tabela 11 para os algoritmos de roteamento PROPHET, GPDMI e ϵ GPDMI.

Para o mecanismo proposto ϵ GPDMI e o algoritmo GPDMI que utilizam repasses através de grupos ou comunidades além da carga de energia dos nós individuais foram utilizados os parâmetros definidos na Tabela 12 para a formação de seus respectivos grupos.

A seguir são apresentados os resultados para as métricas avaliadas neste trabalho em cada um dos conjuntos de dados utilizado. Os resultados são apresentados nas seções 6.8.1 para os resultados do cenário de mobilidade humana Infocom e 6.8.2 para os resultados do cenário de mobilidade humana Reality.

6.8.1 Resultados do cenário Infocom

6.8.1.1 Taxa de entrega de mensagens

O Gráfico 1 mostra a probabilidade de entrega em cada faixa de bateria para cada algoritmo de roteamento testado.

Ao aplicar a camada de clustering de nível de energia sobre os demais fatores de decisão no algoritmo de roteamento da nossa proposta, pretendeu-se manter as mensagens vivas e transportadas pelos nós de maior nível de energia, mitigando assim a degradação da taxa de entrega.

Os resultados confirmam essa intuição, uma vez que o ϵ GPDMI atingiu as melhores taxas de entrega quando os nós se tornaram menos cooperativos para economizar energia.

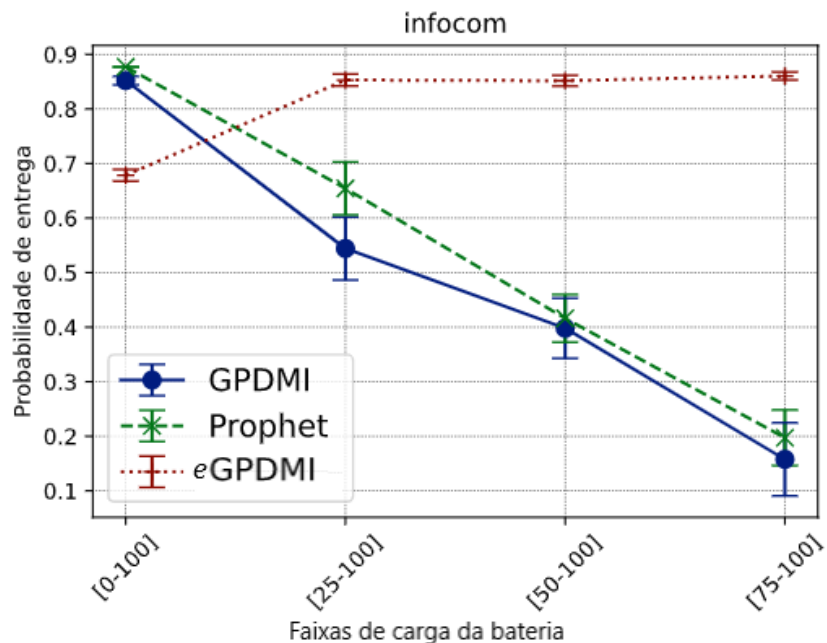


Gráfico 1: Probabilidade de entrega no cenário de mobilidade Infocom

Enquanto no caso base $E = [0,100]$, onde todos os nós cooperam, independentemente do nível de bateria nossa proposta teve uma taxa de entrega de mensagens de aproximadamente 70%, os demais algoritmos de roteamento tiveram uma taxa de entrega de cerca de 85%.

Uma justificativa razoável para este resultado refere-se ao fato de o ϵ GPDMI agrupar os nós em dois grupos, níveis de energia e interesses. Como não houve restrição de nível de energia, foram considerados apenas os interesses. Assim, para

PRoPHET e GPDMI, o encaminhamento foi maior por não haver restrição de energia. Assim, o número de mensagens replicadas na rede foi maior quando comparado ao ϵ GPDMI, como pode ser visto no Gráfico 3. Mesmo com um número menor de mensagens replicadas (142,85% menor), o ϵ GPDMI atingiu uma probabilidade de entrega apenas 19,35% menor que os outros dois protocolos.

Já no caso $E = [75, 100]$, quando os nós com menos de 75% de nível de bateria não cooperam, nossa proposta teve cerca de 80% de taxa de entrega de mensagens e os demais algoritmos de roteamento tiveram a taxa de entrega abaixo de 25%.

Isso ocorre porque esses algoritmos decidem encaminhar para nós com maior probabilidade de ter novos contatos que não são necessariamente os nós com os níveis de bateria mais altos.

6.8.1.2 Atraso médio de entrega

É importante destacar que o ϵ GPDMI teve um melhor desempenho (aprox. 104,76%) na probabilidade de entrega de mensagens quando comparado ao GPDMI e ao PRoPHET. No resultado do atraso médio na entrega de mensagens que é mostrado no Gráfico 2, o ϵ GPDMI teve melhor desempenho no caso $E = [75, 100]$.

Portanto, para verificar se havia diferença significativa entre as médias, aplicou-se o teste t, cujo resultado demonstra que tanto o ϵ GPDMI quanto o GPDMI não apresentaram diferença estatisticamente significativa ($p\text{-valor} = 0,3904$).

Da mesma forma, os resultados comparando ϵ GPDMI e PRoPHET ($p\text{-valor} = 0,3505$) não alcançaram condições suficientes para rejeitar a hipótese nula.

Em resumo, como mostrado no Gráfico 1, a probabilidade de entrega foi maior para o caso $E = [75, 100]$ sem aumentar substancialmente o atraso médio de entrega de ϵ GPDMI em relação aos outros protocolos de roteamento, como pode ser visto no Gráfico 2.

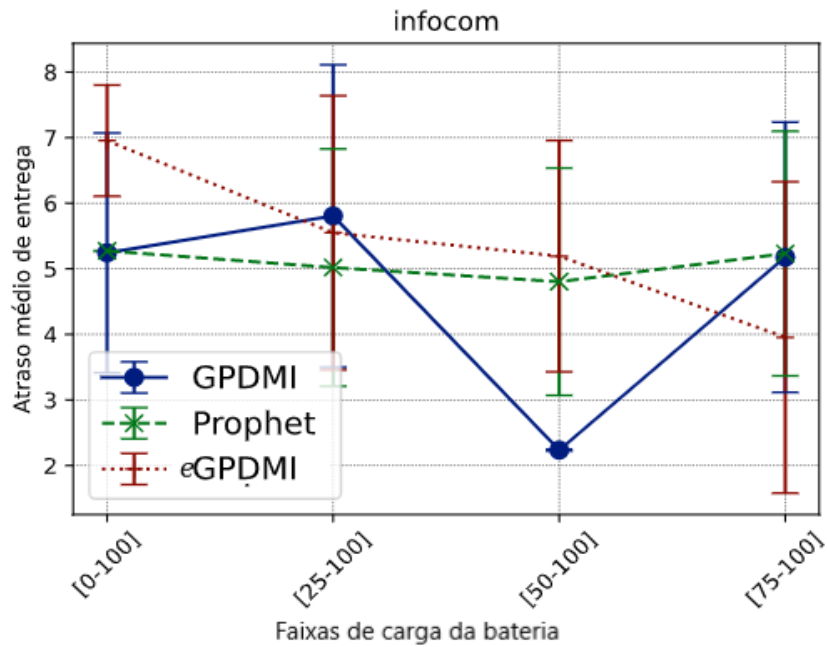


Gráfico 2: Atraso médio de entrega no cenário de mobilidade Infocom

6.8.1.3 Sobrecarga

Para medir a sobrecarga, usamos a equação 4.3 descrita anteriormente no método de avaliação, da seguinte forma:

$$Sobrecarga = \frac{QMRep - QMEnt}{QMEnt} \quad (4.3)$$

Fazendo uma análise entre as taxas de entrega para os casos em que $E=[25,100]$, $E=[50,100]$ e $E=[75,100]$, é possível notar que o ϵ GPDMI encaminha mais mensagens que o GPDMI e o PROpHET, sendo assim, é provável que a mensagem fique mais tempo viva na rede.

Esse resultado vem do fato de que as mensagens são retransmitidas principalmente entre os grupos de nós com o nível de energia mais alto, e que possuem os mesmos interesses, consequentemente isso aumenta o tempo de vida da mensagem na rede.

Conforme mostrado no Gráfico 3, a sobrecarga (*overhead*) do ϵ GPDMI é significativamente menor no caso $E= [0,100]$. Os resultados mostram que as mensagens foram retransmitidas principalmente para os nós com o nível de energia mais alto, mesmo se os nós cooperaram com um nível de bateria baixo.

Além disso, neste cenário, o ϵ GPDMI encaminhou menos mensagens do que os outros algoritmos, ou seja, a taxa de entrega é menor. Por outro lado, quando o

limite de carga da bateria para encaminhar aumenta, notamos que ϵ GPDMI retransmite mais mensagens do que GPDMI e PROphet. Esses resultados implicam dizer que o ϵ GPDMI pode manter as mensagens ativas em rede por mais tempo do que o GPDMI e o PROphet.

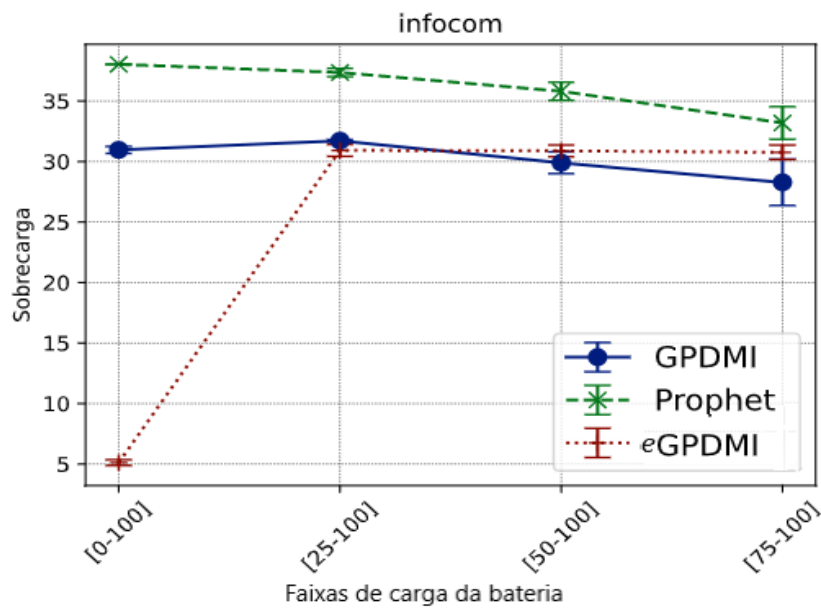


Gráfico 3: Sobrecarga no cenário de mobilidade Infocom

6.8.1.4 Análise de variância - Probabilidade de entrega de mensagens

Ao aplicar o método estatístico ANOVA nas médias de probabilidade de entrega, nossos resultados mostram que há uma diferença significativa para os limites de bateria utilizados, conforme demonstramos na Tabela 14.

O Gráfico 4 é um demonstrativo da distribuição de dados sobre as médias de probabilidade de entrega dos algoritmos de roteamento.

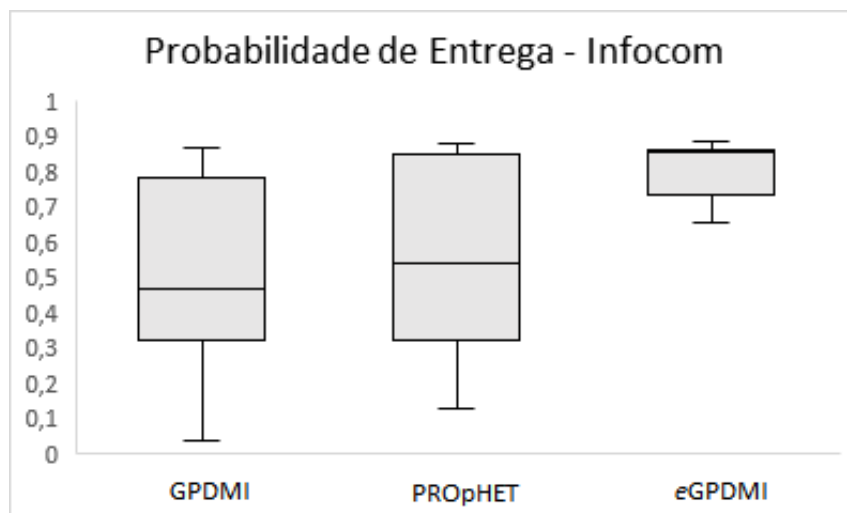


Gráfico 4: Distribuição de dados das médias de taxa de entrega - Infocom

Analisando os resultados da tabela ANOVA, observa-se pelo teste F, usando 5% de significância que, $F_{\text{calculado}} > F_{\text{tabelado}}$, ou seja, existe pelo menos um algoritmo de roteamento com performance de entrega de mensagens diferente entre os limites de bateria estudados.

Uma vez que os valores de $F_{\text{calculado}}$ são respectivamente $F_{\text{calculado}} = 409,58$ e $F_{\text{calculado}} = 295,17$ maiores que F_{tabelado} , tanto para os limites de bateria quanto algoritmos de roteamento, podemos dizer que ambos suportam a diferença de desempenho nas taxas de entrega, entre os algoritmos de roteamento descritos anteriormente no Gráfico 1.

Dessa forma, rejeita-se a hipótese nula H_0 , onde todas as médias são iguais. Para avaliar onde essa variação está ocorrendo, aplicamos o teste de Tukey.

Tabela 14: Análise de Variância (Taxas Entrega) – Anova – Infocom

Variação	<i>SQ</i>	<i>gl</i>	<i>MQ</i>	<i>F (calculado)</i>	<i>valor-P</i>	<i>F (tabelado)</i>
Limites de Bateria	2,62350745	3	0,874502483	295,1720982	$7,154413 \times 10^{-52}$	2,688691468
Algoritmos	2,426933741	2	1,213466871	409,583242	$3,787614 \times 10^{-51}$	3,080386863
Bateria x Algoritmos	2,733700332	6	0,455616722	153,7849764	$1,544337 \times 10^{-50}$	2,183656883
Dentro (resíduos)	0,319970176	108	0,002962687			
Total	8,104111699	119				

6.8.1.5 Teste de Tukey – Probabilidade de entrega (Infocom)

Para usar o teste de Tukey no cenário Infocom, foi aplicado a equação 6.6 descrita anteriormente na subseção método estatístico. Uma vez aplicado a equação sobre os valores encontrados no cálculo anova e nas demais variáveis, temos o seguinte resultado $DMS = 0,058$.

O DMS (Diferença mínima significativa) é usado para identificar por meio de testes estatísticos quais algoritmos são iguais ou diferentes nas performances de entrega conforme as faixas de energia testadas. Quando a diferença absoluta das médias (Módulo da média) calculado é maior que o valor do DMS, dizemos que as médias das taxas de entrega dos algoritmos são estatisticamente diferentes.

Tabela 15: Teste de Tukey (Taxas Entrega) – diferença absoluta das médias – Infocom

Faixas 25-100	Média 1	Média 2	Módulo das Médias
PROpHET x GPDMI	0,65432	0,5442	0,11012
eGPDMI x GPDMI	0,85301	0,5442	0,30881
eGPDMI x PROpHET	0,85301	0,65432	0,19869
Faixas 50-100	Média 1	Média 2	Módulo das Médias
PROpHET x GPDMI	0,41614	0,398	0,01814
eGPDMI x GPDMI	0,85175	0,398	0,45375
eGPDMI x PROpHET	0,85175	0,41614	0,43561
Faixas 75-100	Média 1	Média 2	Módulo das Médias
PROpHET x GPDMI	0,19725	0,15777	0,03948
eGPDMI x GPDMI	0,8604	0,15777	0,70261
eGPDMI x PROpHET	0,8604	0,19725	0,66313
Faixas 0-100	Média 1	Média 2	Módulo das Médias
PROpHET x GPDMI	0,8771	0,85198	0,02512
eGPDMI x GPDMI	0,6783	0,85198	0,17364
eGPDMI x PROpHET	0,6783	0,8771	0,19876

Podemos concluir a partir dos resultados, que existem diferenças significativas no desempenho médio das taxas de entrega entre todos os algoritmos de roteamento para o caso em que o intervalo da bateria é $E = [25,100]$, com a nossa proposta mantendo os maiores resultados nas taxas com aproximadamente 80% de entrega, comparando com PROpHET e GPDMI. Para as demais faixas de bateria, não encontramos nenhuma diferença significativa no desempenho médio entre os algoritmos de roteamento PROpHET e GPDMI.

Nas faixas de energia onde $E = [50,100]$ e $E = [75,100]$, existem diferenças significativas entre o desempenho da Proposta e dos algoritmos GPDMI e PROpHET, com taxas de aproximadamente 80% e 85% de entrega para o eGPDMI, enquanto GPDMI e PROpHET ficaram na faixa de 25% para as cooperações de energia na faixa $E = [75,100]$ e 40% de desempenho na entrega quando $E = [50,100]$.

Por fim, o teste de Tukey também apresentou diferenças entre a performance nas taxas de entrega entre todos os algoritmos da faixa $E = [0,100]$. Neste caso PROpHET e GPDMI entregaram aproximadamente 85% das mensagens na rede e o eGPDMI não ultrapassou os 70%.

Os resultados do teste de Tukey na tabela 15 suportam os já apresentados no Gráfico 1 sobre probabilidade de entrega. Sendo possível comparar estatisticamente as diferenças entre as médias.

6.8.2 Resultados do cenário Reality

6.8.2.1 Taxa de entrega de mensagens

O cenário Reality é mais esparsos que o cenário Infocom. Assim, simulamos apenas alguns dias de todo o rastreamento, conforme citado na Tabela 11. Portanto, espera-se que a probabilidade de entrega seja menor. No Gráfico 5, apresentamos o valor da probabilidade de entrega para todos os algoritmos de roteamento testados.

O resultado mostra que *e*GPDMI pode permanecer estável, mesmo quando o limite de energia da bateria diminui. Além disso, quando o limite do nível de energia da bateria diminui, o GPDMI diminui a taxa de entrega de 32% para menos de 5% e o PROPHET vai de cerca de 48% para menos de 5%. Enquanto, *e*GPDMI vai de 32% para 30%.

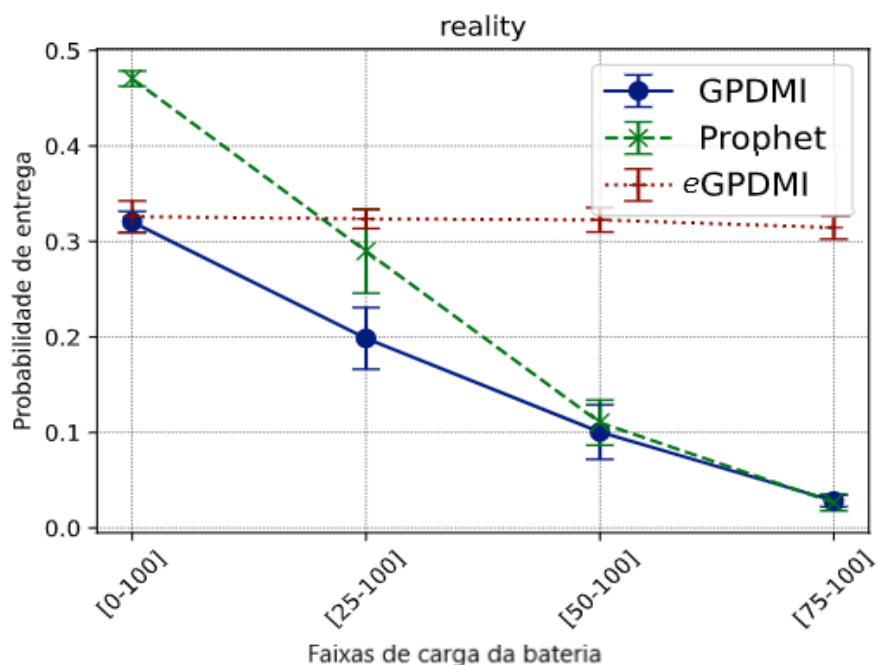


Gráfico 5: Probabilidade de entrega no cenário de mobilidade Reality

6.8.2.2 Atraso médio de entrega

Comparando os dados do Gráfico 5, vemos que o *e*GPDMI teve um desempenho de aproximadamente 142,85% melhor na probabilidade de entrega em relação ao GPDMI

e PRoPHET. Além disso, observamos que o *e*GPDMI teve um pior desempenho de atraso médio na entrega, em torno de 13% de diferença (Gráfico 6).

Porém, analisando os resultados do atraso médio de entrega entre *e*GPDMI e GPDMI, o desvio padrão foi de 18,53 e 21,82, respectivamente. Esses valores indicaram que não houve diferença significativa entre os resultados. Assim, um teste *t* não pareado foi aplicado para encontrar o valor de $p = 0,3533$, mantendo-se a hipótese nula H_0 prevalente, média do *e*GPDMI = média do GPDMI. É fundamental destacar que o mesmo teste foi aplicado ao *e*GPDMI e ao PRoPHET para obter um valor de $p > 0,05$.

É importante enfatizar que o desempenho do atraso médio de entrega é estatisticamente semelhante para GPDMI e PRoPHET. Além disso, o atraso médio de entrega foi próximo entre o *e*GPDMI e os outros algoritmos de roteamento em todos os casos de intervalos de bateria.

Essa relação ocorre pelo fato de ser um trace mais esperso, pois para entregar mais mensagens o tempo de entrega também aumenta.

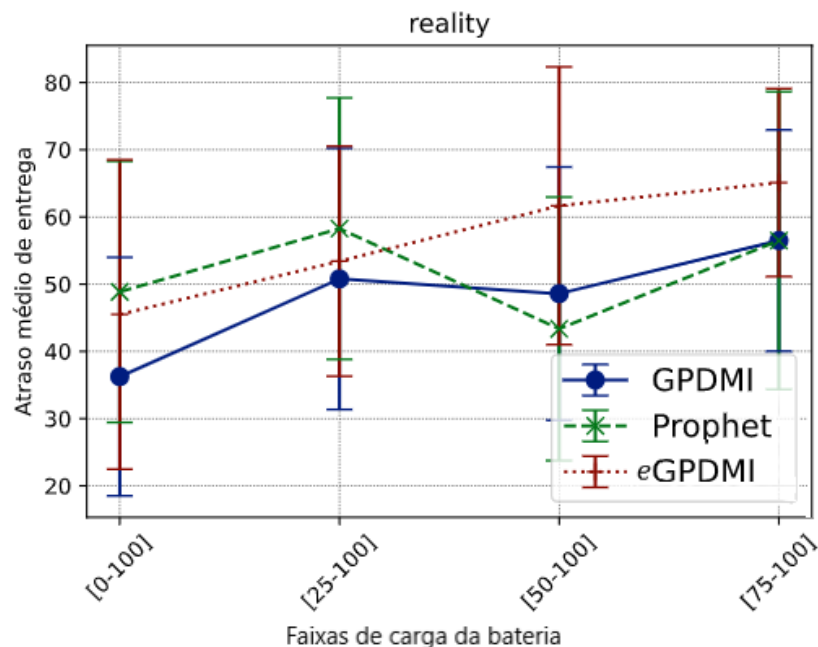


Gráfico 6: Atraso médio de entrega no cenário de mobilidade Reality

6.8.2.3 Sobrecarga

No Gráfico 7, mostramos que, para os casos em que a faixa de bateria para os nós cooperarem é maior que 25%, a sobrecarga (*overhead*) é menor que 100 para todos os algoritmos de roteamento. Por outro lado, no caso de $E = [0,100]$, descobrimos que

a sobrecarga para o protocolo de roteamento PROpHET é cerca de 10 vezes maior do que o GPDMI e o ϵ GPDMI.

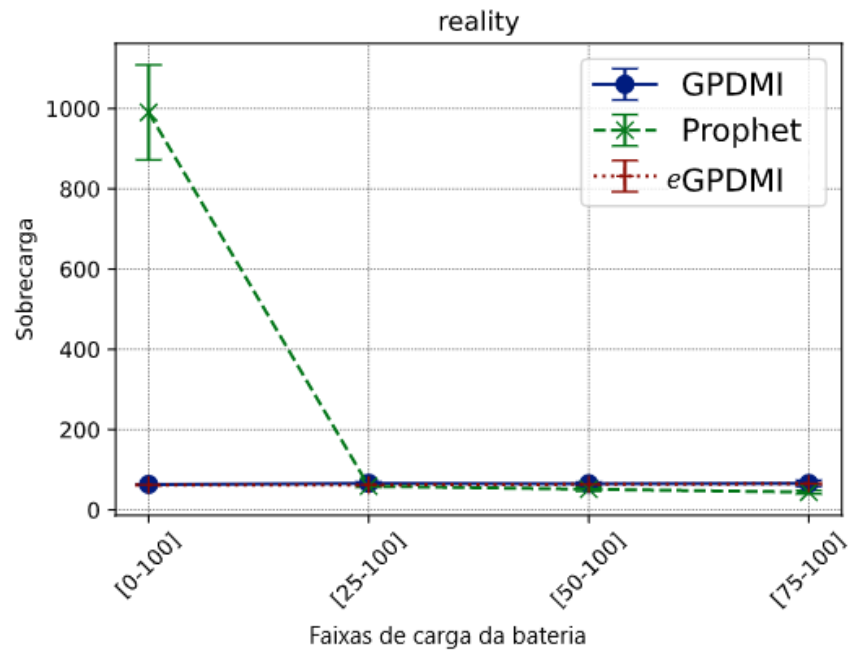


Gráfico 7: Sobrecarga no cenário de mobilidade Reality

A hipótese para este resultado é que no PROpHET, as probabilidades de encaminhamento demoram mais para se ajustar, uma vez que o Reality é um cenário esparso. Esses resultados demonstram ainda que o ϵ GPDMI atenua a sobrecarga da rede, enquanto leva à estabilização da probabilidade de entrega em ambos os cenários.

6.8.2.4 Análise de variância - Probabilidade de entrega de mensagens

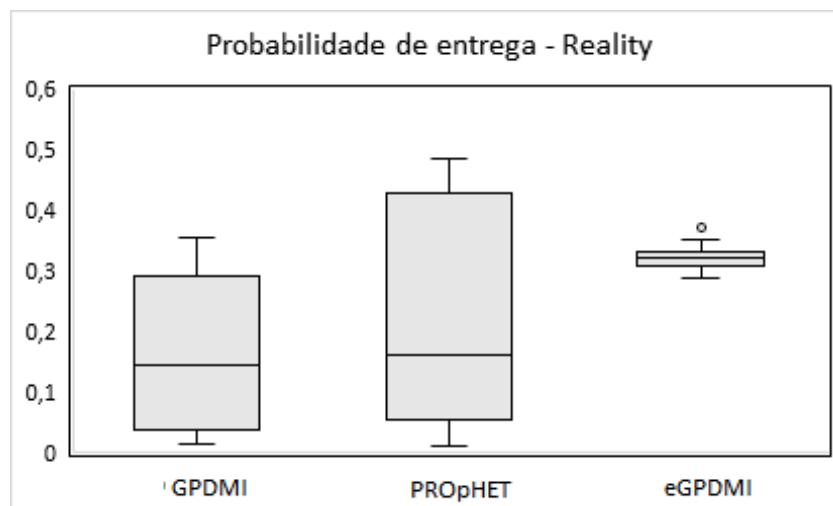


Gráfico 8: Distribuição de dados das médias de taxa de entrega – Reality

Os resultados para a análise de variância (ANOVA) no cenário Reality estão descritos na Tabela 16. Conclui-se que existe pelo menos um algoritmo de roteamento com diferenças relevantes no desempenho da probabilidade de entrega nas faixas dos limites de bateria.

Uma vez que, fazendo o teste F, usando 5% de significância temos $F_{\text{calculado}} > F_{\text{tabelado}}$. Sendo $F_{\text{calculado}} = 414,05$ para os limites de bateria e $F_{\text{calculado}} = 299,04$ para os algoritmos de roteamento, valores respectivamente maiores que o F_{tabelado} de cada um.

Portanto, é razoável declarar o impacto de cada intervalo de bateria na probabilidade de entrega. Assim como a análise feita no cenário Infocom, rejeita-se a hipótese nula H_0 , onde todas as médias são iguais. Iremos então, aplicar o teste de Tukey e calcular estatisticamente as variações.

Tabela 16: Análise de variância (Taxa Entrega) - Anova - Reality

Variação	<i>SQ</i>	<i>gl</i>	<i>MQ</i>	<i>F (calculado)</i>	<i>valor-P</i>	<i>F (tabelado)</i>
Limites de Bateria	1,077451696	3	0,359150565	414,0508078	$4,656932 \times 10^{-59}$	2,688691468
Algoritmos	0,518781447	2	0,259390724	299,0415415	$9,261242 \times 10^{-45}$	3,080386863
Bateria x Algoritmos	0,574260028	6	0,095710005	110,3403658	$9,81607 \times 10^{-44}$	2,183656883
Dentro (resíduos)	0,093679955	108	0,000867407			
Total	0,07988524	119				

6.8.2.5 Teste de Tukey - Probabilidade de entrega (Reality)

Para realizar o teste de Tukey calculamos seu DMS = 0,032. Após a análise do impacto entre os resultados de probabilidade das taxas de entrega dos algoritmos de roteamento, concluímos que, existem diferenças nas médias de desempenho entre todos os algoritmos de roteamento na faixa de bateria $E = [25, 100]$.

É demonstrado na Tabela 17 no cálculo da diferença absoluta das médias (módulo das médias) esse resultado sustenta as taxas de entrega no Gráfico 5, onde o algoritmo GPDMI tem 20%, PROpHET 30% e ϵ GPDMI 32% de probabilidade de entrega.

Na faixa de bateria $E = [0, 100]$, percebemos que não há diferença significativa entre o desempenho do algoritmo Proposto em relação ao algoritmo GPDMI. No entanto, existe uma diferença mínima significativa, entre o ϵ GPDMI com 32% e o algoritmo PROpHET com 48% das taxas de entrega.

A relação entre o algoritmo PROpHET com 32% em relação ao GPDMI também com 48%, demonstra que as duas últimas comparações mantem uma média de entregas semelhante.

Nos limites onde $E = [50,100]$ e $E = [75,100]$, nenhuma diferença significativa foi encontrada entre a comparação do algoritmo PROpHET e GPDMI.

Entre a comparação dos demais algoritmos em ambos os limites foram encontradas diferenças, mantendo-se o eGPDMI com melhores resultados comparando com GPDMI e PROpHET, onde a proposta teve cerca de 32% e 30% do desempenho de entregas sobre 10% e menos de 5% dos demais algoritmos.

Em resumo a Proposta obteve estatisticamente desempenho diferente em quase todos os limites de bateria, exceto, comparado ao GPDMI na faixa $E = [0,100]$.

Tabela 17: Teste de Tukey (Taxas Entrega) – diferença absoluta das médias - Reality

Faixas 25-100	Média 1	Média 2	Módulo das Médias
PROpHET x GPDMI	0,28978	0,19852	0,09126
eGPDMI x GPDMI	0,32365	0,19852	0,12513
eGPDMI x PROpHET	0,32365	0,28978	0,03387
Faixas 50-100	Média 1	Média 2	Módulo das Médias
PROpHET x GPDMI	0,11035	0,10047	0,00988
eGPDMI x GPDMI	0,32269	0,10047	0,22222
eGPDMI x PROpHET	0,32269	0,11035	0,21234
Faixas 75-100	Média 1	Média 2	Módulo das Médias
PROpHET x GPDMI	0,02669	0,02825	0,00156
eGPDMI x GPDMI	0,31454	0,02825	0,28629
eGPDMI x PROpHET	0,31454	0,02669	0,28785
Faixas 0-100	Média 1	Média 2	Módulo das Médias
PROpHET x GPDMI	0,47052	0,32056	0,14996
eGPDMI x GPDMI	0,32605	0,32056	0,00549
eGPDMI x PROpHET	0,32605	0,47052	0,14447

6.9 Considerações Finais

A partir da análise de todas as métricas, taxa de entrega, tempo de atraso, e sobrecarga nos dois cenários, pode-se concluir que o ϵ GPDMI obteve melhor desempenho no cenário Infocom onde a densidade de contatos é maior e o tempo de entrega das mensagens era mais limitado ou seja menor tempo de simulação do que no cenário Reality.

No cenário Reality que tem menor densidade de contatos e maior tempo para entrega das mensagens, o ϵ GPDMI obteve uma probabilidade de entrega constante no desempenho da Taxa de Entrega de Mensagens.

Observando a taxa de entrega e correlacionando com a sobrecarga em ambos os cenários, a sobrecarga se manteve bem menor em todas as faixas de energia testadas. Podemos concluir que o ϵ GPDMI possui melhor desempenho, que os demais protocolos em cenários onde a densidade dos contatos é maior e o tempo das entregas das mensagens menor.

Conclusões

Neste trabalho, apresentamos um novo algoritmo de roteamento que agrupa nós dentro de dois clusters: o cluster de nós de maior nível de energia e os clusters com base no interesse de perfil de cada nó.

Realizamos um planejamento fatorial para confirmar a hipótese de que a faixa de energia usada para os nós cooperarem na comunicação tem um impacto negativo significativo na probabilidade de entrega de mensagens.

Executamos uma avaliação de desempenho usando os traces de mobilidade Infocom e Reality, e comparamos os resultados do *e*GPDMI com os algoritmos de roteamento PROPHET e GPDMI. Sendo analisadas as métricas de taxa de entrega, atraso médio e custo de entrega.

Os resultados mostraram que o *e*GPDMI pode mitigar a degradação da probabilidade de entrega quando os nós não estão dispostos a cooperar devido ao seu nível de energia. Nossa hipótese é que isso acontece devido a formação dos grupos de energia que é composto por nós com níveis de energia semelhantes dentro de uma faixa de bateria considerada pelos nós alta para cooperação na rede.

Além disso, existe uma verificação que é feita para validar se a carga de energia dos nós vizinhos está dentro do limite especificado, o que torna muito mais viável que eles se mantenham colaborando com a rede na entrega de mensagens, por não possuírem restrições de energia, isso poderia fazê-los tornarem-se egoístas ou inoperantes. Outro fator importante é que a proposta forma grupos de energia aliado aos grupos de interesse dos nós, que por si só já contribui com o aumento das taxas de entrega, com o diferencial de considerar os níveis de bateria além dos interesses sociais.

O *e*GPDMI teve um desempenho melhor do que o algoritmo PROPHET e GPDMI em cenários de restrição de nível de energia elevada, especialmente quando a faixa de bateria para cooperação de nós é $E = [75, 100]$. Isso ocorre porque esses algoritmos decidem encaminhar para nós com maior probabilidade de ter novos contatos que não são necessariamente os nós com os níveis de bateria mais altos.

O ϵ GPDMI também manteve a maior parte dos níveis de bateria testados estáveis enquanto os outros algoritmos demonstraram ter uma diminuição em sua taxa de entrega.

Notamos também que para o cenário Infocom, que possui maior densidade de contatos entre os nós e menor tempo de entrega de mensagens, tivemos como resultado um melhor desempenho nas taxas de entrega de mensagens, para cada faixa de energia testada. Nesse cenário os nós se encontram de maneira mais restrita por um curto tempo, ou seja, interação mais entre si do que com outros grupos.

No cenário com menor densidade de contatos e maior tempo para entrega das mensagens (Reality), ϵ GPDMI se manteve constante no desempenho da taxa de entrega de mensagens entre os limites de energia testados. Ou seja, comparando com os demais algoritmos a proposta manteve um equilíbrio na performance de entrega nos cenários de diferentes níveis de energia.

Observamos que o ϵ GPDMI não superou os valores de sobrecarga quando comparada ao roteamento GPDMI e PROpHET em nenhum dos cenários. Os resultados em ambos os cenários Infocom e Reality são atribuídas ao fato, que as mensagens são retransmitidas entre os nós com maior nível de energia, podendo ter estabilizado a sobrecarga na rede.

O trabalho atual tem uma ampla gama de utilizações em OppNets. Uma vez que faltam estudos que considerem o dispositivo de bateria como um dos principais parâmetros, é relevante projetar protocolos e outros algoritmos dedicados a equilibrar o uso de energia, a cooperação dos nós e a probabilidade de entrega de mensagens.

Como trabalho futuro, pretendemos coletar dados reais de mobilidade humana com histórico sobre o consumo de energia e percentuais de bateria dos dispositivos móveis dos usuários com o objetivo de utilizar nos experimentos para avaliação da proposta. Além disso, utilizar esses dados coletados para treinar algoritmos de aprendizagem de máquina para aprender sobre a energia dos nós. Outros fatores adicionais também devem ser considerados, como tempo de recarga da bateria, variação de tamanho de *buffer*, TTL das mensagens e duração dos contatos.

Bibliografia

- Albarelo, Angela Brígida (2005). Combined optimal control of activation and transmission in delay-tolerant networks. *IEEE/ACM Transactions on Networking (TON)*.
- Azevedo, T. S. (2010). Ajuste Adaptativo do sinal de Transmissão de mensagens de descoberta para consumo eficiente de energia em Redes Tolerantes a Atrasos e Interrupções. Tese de doutorado, Universidade Federal do Rio de Janeiro.
- Ayub, Qaisar & Rashid, Sulma (2020). Energy Efficient Inactive Node Detection Based Routing Protocol for Delay Tolerant Network. *Wireless Pers Commun* 116, 227–248.
- B. Jedari, F. Xia, & Z. Ning (2018). “A survey on human-centric communications in non-cooperative wireless relay networks. *IEEE Communications Surveys Tutorials*, vol. 20, no. 2, pp. 914–944, Second quarter 2018.
- B. Jedari, F. Xia, H. Chen, S. K. Das, A. Tolba, & A.-M. Zafer (2019). “Asocial-based watchdog system to detect selfish nodes in opportunistic mobile networks, ”*Future Generation Computer Systems*, vol. 92, pp.777–788.
- Bridgefy (2020).”<https://bridgefy.me/>, Acesso em: 2020-12-10.
- Bussab, W. O. (2013), *ESTATÍSTICA BÁSICA*. São Paulo: Saraiva.
- C. Lee, D. Chang, Y. Shim, N. Choi, T. Kwon, and Y. Choi (2009)., “Regional token based routing for DTNs,” in *Information Networking, ICOIN 2009*. International Conference on. IEEE, 2009, pp. 1 – 5. 40.
- Costa Neto, P. L.O. (2002). *Estatística*. 2ª edição. São Paulo: Editora Edgar Blucher.
- Costa, Paolo; Mascolo, Cecília et al. (2008). Socially-Aware Routing for Publish-Subscribe in Delay-Tolerant Mobile Ad Hoc Networks.
- Chilipirea, Cristian; Petre, Andreea-Cristina; Dobre, Ciprian (2013). Energy-Aware Social-based Routing in Opportunistic Networks. 27th International Conference on Advanced Information Networking and Applications Workshops.

- Cerf, V.; Burleigh, S.; Hooke, A.; Torgerson, L.; Durst, R.; Scott, K.; Fall, K. & Weiss, H. (2007). Delay-tolerant networking architecture, Tech. Rep. RFC4838.
- Cruz, Igor; Assakura, Roberto (2013). Avaliação de Desempenho de Protocolos de Roteamento de Redes DTN Oportunísticas. Trabalho de conclusão de curso, Universidade Federal do Rio de Janeiro.
- D. M. Reis, T. S. Lins, J. M. S. Nogueira, and V. F. Mota (2018). "Sharefile: Sharing content through device-to-device communication," in IEEE Symposium on Computers and Communications (ISCC). IEEE, pp. 00 736–00 741.
- Dede, Jens; Forster, Anna; Hernández-Orallo, Enrique; Herrera-Tapia, Jorge et al. (2018). Simulating Opportunistic Networks: Survey and Future Directions. IEEE Communications Surveys & Tutorials.
- E. Bulut, Z. Wang, & B. K. Szymanski (2009), "Impact of social networks on delay tolerant routing" in Global Telecommunications Conference, 2009. GLOBECOM 2009. IEEE.
- E. Yoneki, P. Hui, S. Chan, & J. Crowcroft (2007), "A socio-aware overlay for publish/subscribe communication in delay tolerant networks," in Proceedings of the 10th ACM Symposium on Modeling, analysis, and simulation of wireless and mobile systems. ACM, pp. 225 – 234.
- F. Xia, Q. Yang, J. Li, J. Cao, L. Liu, & A. M. Ahmed (2016), "Data dissemination using interest-tree in socially aware networking," Computer networks, vol. 91, pp. 495 – 507.
- Feeney, L. M., Nilsson, M. (2001). Investigating the energy consumption of a wireless network interface in an ad hoc networking environment. In: INFOCOM 2001. Twentieth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE, v. 3, pp. 15481557 vol.3, August 2001.
- Forrest Warthman (2003). Delay-Tolerant Networks: A Tutorial. DTN Research Group Internet Draft.
- Keykhaie, Sepehr & Rostaei, Maryam. (2017). Congestion- and selfishness-aware social routing in delay tolerant networks. 439-444.

- K. Fall (2003). "A delay-tolerant network architecture for challenged Internets", ACM SIGCOMM.
- G. McLachlan and T. Krishnan. (2007). The EM algorithm and extensions. John Wiley & Sons.
- Hall, M.; Frank, E.; Holmes, G.; Pfahringer, B.; Reutemann, P. & Witten, I. H. (2009). The weka data mining software: An update. SIGKDD Explor. Newsl., 11(1):10-18.
- Heumann, Christian.; Schomaker, Michael and Shalabh. (2016). Introduction to Statistics and Data Analysis. Springer.
- Hui, P.; Crowcroft, J. & Yoneki, E. (2011). Bubble rap: Social-based forwarding in delay-tolerant networks. Mobile Computing, IEEE Transactions on:1576-1589.
- Jain, R. (1991). The Art of Computer Systems Performance Analysis. John Wiley & Sons.
- Jain, A., Dubes, R. (1988). Algorithms for Clustering Data. Prentice-Hall, Englewood Cliffs, NJ.
- Jiagao Wu, Y. Zhu, L. Liu, B. Yu & J. Pan. (2016). "Energy-Efficient Routing in Multi-Community DTN with Social Selfishness Considerations," IEEE Global Communications Conference (GLOBECOM).
- Keränen, A.; Ott, J. & Kärkkäinen, T. (2009). The ONE Simulator for DTN Protocol Evaluation. Em SIMUTools '09: Proceedings of the 2nd International Conference on Simulation Tools and Techniques, New York, NY, USA. ICST.
- Kurose, James F.; Ross, Keith W. (2013). Redes de Computadores e a Internet: Uma abordagem Top Down. 6. ed. São Paulo: Pearson Addison Wesley.
- Li, Fan; Jiang, Hong; Li, Hanshang (2017). SEBAR: Social-Energy-Based Routing for Mobile Social Delay-Tolerant Networks. IEEE Wireless Communications and Networking Conference (WCNC).
- Linden, Ricardo (2009). Técnicas de agrupamento, Revista de Sistemas da Informação da FSMA, n.4, pp 18-36, Faculdade Salesiana Maria Auxiliadora, Rio de Janeiro.
- Loudari, Salah & Benamar, Nabil. (2019). Effects of Selfishness on the Energy Consumption in Opportunistic Networks: A Performance Assessment.

- Luiz H. A. Correia, Daniel F. Macedo, Michel A. S. (2011). Ribeiro, Tales Heimfarth. AntRoP - Protocolo de Roteamento Bio-inspirado em Colônia de Formiga Tolerante a Falhas e Desconexões aplicado às Redes Emergenciais, XVI Workshop de Gerência e Operação de Redes e Serviços.
- Machado, Nelson. Júnior (2013). SOCLEER: uma proposta de disseminação de dados para redes oportunistas com redistribuição de carga em nós preferidos. Dissertação (Mestrado em Informática) - Universidade Federal do Estado do Rio de Janeiro - UNIRIO, pág. 106.
- M. W. Totaro & D. D. Perkins (2005). "Using statistical design of experiments for analyzing mobile ad hoc networks," in Proceedings of the 8th ACM international symposium on Modeling, analysis and simulation of wireless and mobile systems, pp. 159–168.
- M. Musolesi and C. Mascolo, Mobility models for systems evaluation. Springer, 2009. 15, 25.
- M. Melo (2011), "Implementação e Avaliação de um modelo para trocas de mensagens em DTNs utilizando Tráfego Aéreo". Dissertação de Mestrado, URFJ.
- Michael Demmer, Eric Brewer, Kevin Fall, Melissa Ho, & Robin Patra (2004). Implementing delay tolerant networking. Science.
- Moreira, W.; Mendes, P (2013), " Social-Aware Opportunistic Routing: The New Trend," Routing in Opportunistic Networks, Springer New York, pp.27-68.
- Neves, Eric Vieira das (2018). Agrupamento em dois Níveis para Disseminação de Mensagens em Redes Sociais Móveis Tolerantes a Atrasos e Desconexões. Dissertação de mestrado, Universidade Federal do Amazonas.
- Neto, J. B. P. (2012). Um Modelo para Previsão do Volume de Contato em Redes Tolerantes a Atrasos e Desconexões: Uma Abordagem Quantitativa. Tese de Doutorado, Universidade Federal do Amazonas.
- Nguyen, H.A.; Giordano, S. (2008). Spatiotemporal routing algorithm in opportunistic networks. In: Word of Wireless, Mobile and Multimedia Networks. WoWMoM 2008 International Symposium.

- N. Sharma, A. Bajpai, and M. R. Litoriya (2012). "Comparison the various clustering algorithms of weka tools," facilities, vol. 4, no. 7, 2012.
- Oliveira, C. T.; Moreira, M. D.; Rubinstein, M. G.; Costa, L. H. M. & Duarte, O. C. M. (2007). Redes tolerantes a atrasos e desconexões. SBRC Simpósio Brasileiro de Redes de Computadores e Sistemas Distribuídos.
- Oliveira, C. T. d. (2008). Uma Proposta de Roteamento Probabilístico para Redes Tolerantes a Atrasos e Desconexões. Tese de doutorado, Universidade Federal do Rio de Janeiro.
- Olson, M. The logic of collective action (1965). Public goods and the theory of groups. [S.I.], Harvard University Press.
- P. A. Morettin & W. O. Bussab (2017). Estatística Básica. Editora Saraiva.
- Palazzo, Sérgio; Campbell, Andrew T.; Amorim, Marcelo Dias (2011). Opportunistic and Delay-Tolerant Networks. EURASIP Journal on Wireless Communications and Networking.
- Palmas, Luann Farias (2018). Agrupamento de dados: K-médias. Universidade Federal do Recôncavo da Bahia, Cruz das Almas.
- Pelusi, L., Passarella, A., Conti, M.(2006) "Opportunistic Networking: Data Forwarding in Disconnected Mobile Ad hoc Networks". IEEE Communications Magazine, v. 44, n.11, pp. 134-141.
- Preshiya, D. Jyothi; Suriyakala, C. D. (2015). A Survey of Factors Influencing Network Lifetime in Delay Tolerant Network. Indian Journal of Science and Technology.
- Picard, F., (2007). An introduction to mixture models, Laboratoire Estatistique et Génome, UMR CNRS 8071 – INRA 1152 – Univ. d'Evry, France
- Pond, Bryan Larry; Oliveira, Antônio C. Júnior; Moreira, Waldir (2013). Aspectos sociais como suporte ao roteamento em redes oportunistas. XI CEEL – Conferência de estudos em engenharia elétrica, Universidade Federal de Uberlândia.
- Portal Amazônia (2020). "Pesquisa mostra que 97% de escolas rurais ribeirinhas não tem acesso a internet, no Amazonas". Disponível em:

- <<https://portalamazonia.com/noticias/cidades/pesquisa-mostra-que-97-de-escolas-rurais-ribeirinhas-nao-tem-acesso-a-internet-no-amazonas>>. Acesso em: 3 de maio de 2021.
- Puc Rio (2010). Clusterização de dados. Disponível em: <https://www.maxwell.vrac.puc-rio.br/24787/24787_5.PDF>. Acesso em: 10 de abril de 2021.
- Pujol, J. M.; Lopez Toledo, A. & Rodriguez, P. (2009). Fair Routing in Delay Tolerant Networks. Em IEEE INFOCOM 2009 - The 28th Conference on Computer Communications, pp. 837--845. IEEE. ISSN 0743-166X.
- Rolim, C. O (2016). Uso de sensibilidade à situação em redes oportunistas para intensificar a comunicação de dados em aplicações de sensoriamento urbano. Master's thesis, Universidade Federal do Rio Grande do Sul. Instituto de Informática. Programa de Pós-Graduação em Computação, Porto Alegre.
- Santos Teixeira, Luís Guilherme; Souza Frazão, Marcos Antônio (2015). Avaliação do uso de agrupamento semi supervisionado na segmentação do disco óptico para fins de detecção do glaucoma. Revista de Sistemas e Computação, Salvador, v.5. pp 50-59.
- Sanches, Marcelo Kaminski (2003). Aprendizado de máquina semi-supervisionado: proposta de um algoritmo para rotular exemplos a partir de poucos exemplos rotulados. Dissertação de mestrado, USP: São Carlos.
- S. Hidi (2006). "Interest: A unique motivational variable," Educational research review, vol. 1, no. 2, pp. 69 – 82.
- S. Gupta, C. Nagpal, & C. Singla (2011). "Impact of selfish node concentration in manets," *International Journal of Wireless & Mobile Networks (IJWMN)* Vol, vol. 3, pp. 29–37.
- Silva, Denis Rodrigues; Costa, Antônio; Macedo, Joaquim (2012). Energy Impact Analysis on DTN Routing Protocols. Extreme Conference on Communication and Computing.
- Simon, H. A. (1983). Why should machines learn? Em Machine learning, pp. 25--37. Springer.

- Soares Moreira, Diogo (2018). Um modelo de reputação com classificação via agrupamento para detecção de nós egoístas em redes oportunistas. Dissertação de mestrado, Universidade Federal do Amazonas.
- Soares Luiz, Claiton (2016). Sobre a influência de nós egoístas no descarregamento de tráfego com redes oportunistas. XVI Simpósio Brasileiro em Segurança da Informação e de Sistemas Computacionais, SBSeg 2016.
- Souza, Camilo Batista de (2013). Inserindo aspectos sociais na gerência de buffer para redes tolerantes a atrasos e desconexões. Dissertação de mestrado, Universidade Federal do Amazonas.
- Souza, Camilo Batista de (2019). Disseminação de mensagens em redes oportunistas baseada em relações sociais e aprendizagem de máquina. Tese de doutorado, Universidade Federal do Amazonas.
- Socievole, Annalisa; Rango, Floriano de (2015). Energy-aware centrality for information forwarding in mobile social opportunistic networks. International Wireless Communications and Mobile Computing Conference (IWCMC).
- Spyropoulos, T.; Psounis, K. & Raghavendra, C. S. (2006). Performance analysis of mobility-assisted routing. Em Proceedings of the seventh ACM international symposium on Mobile ad hoc networking and computing - MobiHoc '06, p. 49, New York, New York, USA. ACM Press.
- Taylor e Francis Grouping (2019). Redes tolerantes a atrasos e interrupções: interplanetárias e ligadas à Terra – arquitetura, protocolos e aplicações.
- Wang, W., Srinivasan, V., Motani, M. (2007) Adaptive contact probing mechanisms for delay tolerant applications. In: MobiCom '07: Proceedings of the 13th annual ACM international conference on Mobile computing and networking, pp. 230241, New York, NY, USA, February 2007.
- Y. Zhu, B. Xu, X. Shi, & Y. Wang, (2013). “A survey of social-based routing in delay tolerant networks: Positive and negative social effects,” IEEE Communications Surveys & Tutorials, vol. 15, no. 1, pp. 387 – 401.

- Yoon, S. K. and Haas, Z. J. (2012). Energy-aware routing protocol for delay tolerant networks. In: Vasilakos, A.; Zhang, Y.; Spyropoulos, T. V[Eds.]. Delay Tolerant Networks - Protocols and Applications. New York: CRC Press. p. 69–100.
- Yoon, Jungkeun; Liu, Mingyan; Nobre, Brian (2003). Random Waypoint Considered Harmful. IEEE INFOCOM 2003.
- Xu, K.; Hui, P.; K.Li, V. O.; Crowcroft, J.; Latora, V. & Lio, P. (2009). Impact of altruism on opportunistic communications. In Proceedings of the International conference on Ubiquitous and future networks, ICUFN09, pp. 153-158, Piscataway NJ, USA.