



Universidade Federal do Ceará
Centro de Ciências
Departamento de Ciência da Computação
Programa de Pós-Graduação em Ciência da Computação

Adyson Magalhães Maia

Um mecanismo para escalonamento de pacotes no uplink da rede LTE no contexto da
comunicação máquina-a-máquina

Fortaleza, 2013

ADYSON MAGALHÃES MAIA

**UM MECANISMO PARA O ESCALONAMENTO DE PACOTES NO
UPLINK DA REDE LTE NO CONTEXTO DA COMUNICAÇÃO
MÁQUINA-A-MÁQUINA**

Dissertação submetida à Coordenação do Curso de Pós-Graduação em Ciência da Computação da Universidade Federal do Ceará, como requisito parcial para a obtenção do grau de Mestre em Ciência da Computação.

Área de concentração: Ciência da Computação.

Orientadora: Prof. Dr. Miguel Franklin de Castro.

FORTALEZA, CEARÁ

2013

Dados Internacionais de Catalogação na Publicação
Universidade Federal do Ceará
Biblioteca Universitária
Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

M184m Maia, Adyson Magalhães.

Um mecanismo para escalonamento de pacotes no uplink da rede LTE no contexto da comunicação máquina-a-máquina / Adyson Magalhães Maia. – 2013.
66 f. : il. color.

Dissertação (mestrado) – Universidade Federal do Ceará, Centro de Ciências, Programa de Pós-Graduação em Ciência da Computação, Fortaleza, 2013.

Orientação: Prof. Dr. Miguel Franklin de Castro.

Coorientação: Prof. Dr. Dário Vieira Conceição.

1. Comunicação máquina-a-máquina. 2. Internet das coisas. 3. Escalonador de pacotes para o uplink da rede LTE. I. Título.

CDD 005

ADYSON MAGALHÃES MAIA

**UM MECANISMO PARA O ESCALONAMENTO DE PACOTES NO
UPLINK DA REDE LTE NO CONTEXTO DA COMUNICAÇÃO
MÁQUINA-A-MÁQUINA**

Dissertação submetida à Coordenação do Curso de Pós-Graduação em Ciência da Computação, da Universidade Federal do Ceará, como requisito parcial para a obtenção do grau de Mestre em Ciência da Computação. Área de concentração: Ciência da Computação.

Aprovada em: __/__/2013

BANCA EXAMINADORA

Prof. Dr. Miguel Franklin de Castro
Universidade Federal do Ceará - UFC
Orientador

Prof. Dr. Dario Vieira Conceição
EFREI
Co-orientador

Profa. Dra. Rossana Maria de Castro Andrade
Universidade Federal do Ceará - UFC

Prof. Dr. Danielo Gonçalves Gomes
Universidade Federal do Ceará - UFC

Prof. Dr. Yacine Ghamri-Doudane
Université de La Rochelle

Aos meus pais, familiares e amigos

AGRADECIMENTOS

Agradeço à minha família que sempre me apoiou e incentivou nas minhas decisões que levaram ao caminho que estou trilhando.

Aos meus grandes amigos imiscíveis, Cecilia Feitoza, Hendrio Santiago, Kleber Gomes, Mailson Magalhães, Monyke Silveira e Regina Justa, por ainda me aturarem mesmo depois deste grande tempo de amizade e é claro pelos apoios, conselhos e incentivos.

Ao meu orientador, professor Miguel Franklin, pelo apoio ao meu trabalho através de seus ensinamentos e experiências. Graças ao seu guia foi possível terminar esta árdua tarefa de fazer um mestrado.

Aos professores Dario Viera e Yacine pelo convite do estágio realizado na França na fase final do trabalho. Especialmente ao professor Dario que me recebeu, auxiliou, tanto em assuntos pessoais como acadêmicos, durante a estadia na França e também pela sua colaboração no meu trabalho.

Aos membros da Efrei Internacional, são tantos que ocorre o risco de esquecer alguém, pode ter me ajudado nessa transição cultural Brasil-França e por ter mostrado um pouco deste grande polo cultural que é Paris. Em especial à Catherine Nusbaum e ao Francois Calvarési.

Aos professores Danielo Gomes e Rossana Andrade, que compõem a banca examinadora e certamente contribuirão com este trabalho.

Também agradeço aos colegas e amigos que me acompanharam durante esta jornada, André Fonteles, Benedito, Cecilia Rebouças, Christiano Rodrigues, Ítalo Linhares, Nayane, Níbia, Paulo Artur, Rafael Lima, Rainara, Ricardo Viana e Thalisson pelo companheirismo e pelas trocas de conhecimentos.

À equipe e aos demais colegas do GReat que contribuíram de alguma forma para a conclusão deste trabalho.

Finalmente à CAPES pelo financiamento da bolsa de estudos para realização deste trabalho.

“We’re all stories, in the end. Just make it a good one, ok?!”

(The Doctor)

RESUMO

As redes celulares são esperadas para desempenhar um importante papel no futuro da Internet das Coisas (*Internet of Things* - IoT) e na comunicação Máquina-a-Máquina (*Machine-to-Machine* - M2M), com a rede LTE ganhando cada vez mais destaque tanto da academia quanto da indústria na implantação deste cenário. Contudo, melhorias na rede LTE são necessárias devido às características intrínsecas da comunicação M2M, como a grande quantidade de dispositivos e a diversidade nos requerimentos das suas aplicações, que os diferenciam da comunicação Humano-a-Humano (*Human-to-Human* - H2H) para a qual a rede foi inicialmente projetada. Dentre essas melhorias, está o aperfeiçoamento no escalonamento de pacotes no *uplink* definido no padrão para que trate o caso da comunicação M2M e atenda aos seus requerimentos sem causar grandes impactos na comunicação H2H. Este trabalho propõe um mecanismo para o escalonamento de pacotes no *uplink* da rede LTE no contexto da comunicação M2M que utiliza as informações atuais e passadas da rede e de cada dispositivo para satisfazer os requisitos de Qualidade de Serviço (QoS) dos dispositivos M2M, evitar o problema de inanição e controlar o congestionamento causado por estes dispositivos. Os resultados obtidos indicam que a proposta consegue reduzir o impacto da comunicação M2M sobre a H2H e garantir justiça na alocação de recursos. No entanto, o nível de satisfação dos requisitos de QoS dos dispositivos M2M diminui com o aumento do tráfego H2H.

Palavras-chave: Comunicação Máquina-a-Máquina. Internet das Coisas. Escalonador de Pacotes para o *Uplink* da rede LTE.

ABSTRACT

Cellular networks are expected to play an important role in the future of Internet of Things (IoT) and in the Machine-to-Machine (M2M) communication with the LTE network gaining increasing highlights for this purpose. However, improvements in the LTE network are necessary because of the intrinsic characteristics of M2M communication, such as the large number of devices and the diversity of the requirements of their applications, which differentiate from the Human-to-Human (H2H) communication for which the network was initially designed. Among these improvements, it includes improving the uplink packet scheduler of LTE standard to dealing the M2M communication and meets your requirements without serious impacts on the H2H communication. This work proposes a mechanism for the uplink packet scheduler in the context of M2M communication that uses current and past information from the network and each device to satisfy their requirements of Quality of Service (QoS), avoid the problem of starvation and control the congestion caused by these devices. The results indicate that the proposal can reduce the impact of the M2M communication over the H2H and ensure justice in resource allocation, however, the level of satisfaction of QoS requirements decreases with increasing of H2H traffic.

Keywords: Machine-to-Machine Communication. Internet of Things. Uplink Packet Scheduler for LTE network.

LISTA DE FIGURAS

| | | |
|------------|---|----|
| Figura 2.1 | Arquitetura para a comunicação M2M. Adaptado de (GALETIĆ et al., 2011) | 23 |
| Figura 2.2 | Arquitetura da rede LTE. Reproduzido de (TEKTRONIX, 2012) | 27 |
| Figura 2.3 | Pilha das camadas da interface E-UTRAN. Reproduzido de (ALI-YAHIYA, 2011) | 28 |
| Figura 2.4 | Estrutura de quadros do LTE. Adaptado de (AMBERKAR, 2013) | 29 |
| Figura 4.1 | Divisão dos recursos entre as comunicações H2H e M2M | 42 |
| Figura 5.1 | Impacto da comunicação M2M sobre a taxa de transferência da comunicação H2H | 54 |
| Figura 5.2 | Impacto da comunicação M2M sobre a satisfação de QoS da comunicação H2H | 55 |
| Figura 5.3 | Taxa de transferência do tráfego das aplicações M2M baseadas em evento | 56 |
| Figura 5.4 | Taxa de transferência do tráfego das aplicações M2M baseadas em tempo | 57 |
| Figura 5.5 | Nível de insatisfação dos requisitos de QoS das aplicações M2M baseadas em evento | 58 |
| Figura 5.6 | Nível de insatisfação dos requisitos de QoS das aplicações M2M baseadas em tempo | 59 |
| Figura 5.7 | Índice de justiça para o tráfego das aplicações M2M baseadas em evento | 60 |
| Figura 5.8 | Índice de justiça para o tráfego das aplicações M2M baseadas em tempo | 61 |

LISTA DE TABELAS

| | | |
|------------|---|----|
| Tabela 2.1 | Exemplo de aplicações para a comunicação M2M | 22 |
| Tabela 2.2 | Requisitos de QoS para as classes de Liu et al. (2011) | 25 |
| Tabela 2.3 | Classes de QoS do padrão do sistema LTE. | 30 |
| Tabela 3.1 | Comparação dos trabalhos relacionados | 35 |
| Tabela 4.1 | Requisitos de QoS para as classes de Amokrane, Ksentini e Hadjadj-Aoul (2011) | 38 |
| Tabela 4.2 | Classes de QoS para as aplicações M2M | 40 |
| Tabela 5.1 | Parâmetros dos modelos para os tráfegos H2H e M2M | 50 |
| Tabela 5.2 | Parâmetros das simulações | 51 |

LISTA DE SIGLAS

| | |
|---------|---|
| IoT | Internet of Things |
| M2M | Machine-to-Machine |
| LTE | Long Term Evolution |
| H2H | Human-to-Human |
| QoS | Qualidade de Serviço |
| H2M | Human-to-Machine |
| M2H | Machine-to-Human |
| VoIP | Voice over IP |
| RFID | Radio-Frequency Identification |
| IP | Internet Protocol |
| RSSF | Redes de Sensores Sem Fio |
| POS | Point of Sale |
| 3GPP | Third Generation Partnership Project |
| GSM | Global System for Mobile Communications |
| UMTS | Universal Mobile Telecommunication System |
| OFDMA | Orthogonal Frequency-Division Multiplexing Access |
| SC-FDMA | Single Carrier Frequency-Division Multiplexing Access |
| TDD | Time Division Duplex |
| FDD | Frequency Division Duplexing |
| ITU | International Telecommunication Union |
| LTE-A | LTE Advanced |
| RAN | Random Access Network |
| E-UTRAN | Evolved Universal Terrestrial Radio Access Network |
| CN | Core Network |
| SAE | System Architecture Evolution |
| EPC | Evolved Packet Core |
| EPS | Evolved Packet System |
| eNB | evolved-NodeB |
| UE | User Equipement |
| MAC | Media Access Control |
| RLC | Radio Link Control |
| PDCP | Packet Data Convergence Protocol |
| RRC | Radio Resource Control |
| TTI | Transmission Time Interval |
| RB | Resource Block |
| QCI | QoS Class Identifier |
| SR | Scheduling Request |

| | |
|------|---|
| BSR | Buffer Status Report |
| CSI | Channel State Information |
| SINR | Signal-to-Interference-plus-Noise Ratio |
| SRS | Sounding Reference Signal |
| TDPS | Time Domain Packet Scheduling |
| FDPS | Frequency Domain Packet Scheduling |
| MCS | Modulation and Coding Scheme |
| PF | Proportional Fair |

LISTA DE SÍMBOLOS

| | |
|----------------------------|---|
| \mathcal{U} | Conjunto com todos os dispositivos |
| \mathcal{U}_{H2H} | Conjunto com os dispositivos H2H |
| \mathcal{U}_{M2M} | Conjunto com os dispositivos M2M |
| \mathcal{RB} | Conjunto com todos os RBs disponíveis |
| $\widehat{RB}_{H2H}(u, t)$ | Demanda de RBs para o UE H2H u no TTI t |
| BS_{H2H} | Quantidade de dados no <i>buffer</i> do UE H2H u à espera de transmissão no TTI t |
| $RB_{H2H}^{avg}(u, t)$ | Quantidade média de recursos alocados para o UE H2H u até o TTI t |
| $BS_{H2H}^{avg}(u, t)$ | Quantidade média de dados no <i>buffer</i> do UE H2H u até o TTI t |
| RB_{H2H}^{min} | Demanda mínima de recursos por UE H2H |
| α | Constante da média móvel exponencial de $RB_{H2H}^{avg}(u, t)$ e $BS_{H2H}^{avg}(u, t)$ |
| $\widehat{RB}_{H2H}(t)$ | Demanda total de recursos dos UEs H2H no TTI t |
| P_{M2M}^{min} | Porcentagem mínima dos RBs para a comunicação M2M |
| RB_{M2M}^{min} | Quantidade alocada de recursos por UE M2M |
| $RB_{M2M}(t)$ | Quantidade de RBs reservada para a comunicação M2M no TTI t |
| U_{M2M}^{max} | Quantidade máxima de UEs M2M que receberão recursos no atual TTI |
| $RB_{H2H}(t)$ | Quantidade de RBs reservada para a comunicação H2H no TTI t |
| $M_{M2M}^{TD}delay$ | Função utilitária para priorizar o UE M2M u no TTI t segundo seus requisitos de QoS |
| $\Delta D(u)$ | Função que avalia o quão próximo o UE M2M u está de superar o atraso máximo tolerável no TTI t |
| $M_{M2M}^{TD}fair$ | Função utilitária que avaliar a alocação justa nos dispositivos M2M |
| $T_{M2M}^{avg}(u, t)$ | Taxa de transferência média do UE M2M u até o TTI t |
| β | Constante da média móvel exponencial de $T_{M2M}^{avg}(u, t)$ |
| $M_{M2M}^{TD}(u, t)$ | Função prioritária da fase TDPS |
| ϖ | Constante da função $M_{M2M}^{TD}(u, t)$ |
| σ | Porcentual sobre $\Delta D(u)$ para o tempo de espera máximo do UE M2M u após a requisição atual ser negada |
| $T_{M2M}(u, t)$ | Taxa de transferência espera para o UE M2M u no TTI t |
| $T_{M2M}(u, g)$ | Taxa de transferência espera para o UE M2M u no TTI t ao utilizar o conjunto de RBs g |
| $\eta(u, g)$ | Eficiência espectral entre o dispositivo u e o grupo de recursos g |
| Γ | Coefficiente da diferença entre o limite teórico e a performance real |
| $\gamma(u, g)$ | SINR entre o dispositivo u e o conjunto de RBs g |

SUMÁRIO

| | | |
|------------|--|----|
| 1 | INTRODUÇÃO | 16 |
| 1.1 | Contextualização | 16 |
| 1.2 | Objetivos | 17 |
| 1.2.1 | Objetivo geral | 17 |
| 1.2.2 | Objetivos específicos | 17 |
| 1.3 | Planos de atividades | 18 |
| 1.4 | Organização da dissertação | 18 |
| 2 | REVISÃO BIBLIOGRÁFICA | 20 |
| 2.1 | Internet das coisas | 20 |
| 2.2 | Comunicação máquina-a-máquina | 21 |
| 2.2.1 | Classificação das aplicações M2M | 24 |
| 2.3 | Rede celular LTE | 25 |
| 2.3.1 | Arquitetura da rede | 26 |
| 2.3.2 | Estrutura de quadros | 28 |
| 2.3.3 | Classes de QoS | 29 |
| 2.3.4 | Escalonamento de pacotes | 29 |
| 2.3.5 | Mensagens de controle para o escalonamento | 31 |
| 3 | TRABALHOS RELACIONADOS | 32 |
| 4 | MECANISMO DE ESCALONAMENTO PARA A COMUNICAÇÃO M2M | 36 |
| 4.1 | Formulação do problema | 36 |
| 4.2 | Classes de QoS | 37 |
| 4.3 | Cálculo da demanda por recursos | 39 |
| 4.4 | Descrição do algoritmo de escalonamento | 42 |
| 4.4.1 | Primeira fase - TDPS | 43 |
| 4.4.2 | Segunda fase - FDPS | 45 |
| 5 | EXPERIMENTOS | 47 |
| 5.1 | Soluções comparadas | 47 |
| 5.2 | Modelagem do tráfego | 48 |

| | | |
|------------|---|----|
| 5.3 | Descrição do plano de testes | 49 |
| 5.4 | Métricas de avaliação | 51 |
| 5.5 | Análise dos resultados obtidos | 52 |
| 5.5.1 | Impacto da comunicação M2M sobre a H2H | 53 |
| 5.5.2 | Taxa de transferência | 53 |
| 5.5.3 | Nível de satisfação dos requisitos de QoS | 56 |
| 5.5.4 | Justiça na alocação dos recursos | 58 |
| 6 | CONCLUSÃO E TRABALHOS FUTUROS | 62 |
| 6.1 | Conclusões | 62 |
| 6.2 | Contribuições | 62 |
| 6.3 | Trabalhos futuros | 63 |
| | REFERÊNCIAS BIBLIOGRÁFICAS | 64 |

1 INTRODUÇÃO

Esta dissertação apresenta a implementação de um mecanismo para o escalonamento de pacotes no *uplink* na rede LTE que trata a comunicação Máquina-a-Máquina além da comunicação Humano-a-Humano. Este escalonamento utiliza informações atuais do tráfego de cada dispositivo assim como informações do histórico de alocações para identificar o tipo de comunicação utilizado nos dispositivos e em suas aplicações e também para priorizar a alocação de recursos compartilhados necessários para o envio de dados.

Na seção 1.1 é descrita a contextualização onde este trabalho está inserido. Os objetivos do trabalho são apresentados na seção 1.2. As atividades que compuseram este trabalho são descritas na seção 1.3. Finalizando, a seção 1.4 expõe a organização dos capítulos subsequentes desta dissertação.

1.1 Contextualização

As evoluções das nanotecnologias, das tecnologias móveis e sem fio propiciam que uma gama maior de objetos físicos do nosso dia-a-dia (eletrodomésticos, automóveis, celulares etc) como também sensores sem fio, atuadores e máquinas industriais possam se conectar à Internet (STRATEGY, I. T. U.; UNIT, Policy, 2005). Um novo paradigma chamado Internet das Coisas (*Internet of Things* - IoT) surge deste cenário onde uma diversidade de objetos (também chamados de objetos inteligentes) estarão conectados e criando, assim, uma junção entre o mundo físico e o mundo das informações ou mundo virtual (UCKELMANN; HARRISON; MICHAHELLES, 2011).

Neste novo ambiente não existirá somente a comunicação Humano-a-Humano (*Human-to-Human* - H2H), a Humano-a-Máquina (*Human-to-Machine* - H2M), e a Máquina-a-Humano (*Machine-to-Human* - M2H), mas também a comunicação Máquina-a-Máquina (*Machine-to-Machine* - M2M), que terá um importante papel neste ambiente sendo o tipo de comunicação que mais gerará e receberá tráfego (WU et al., 2011). Esta comunicação é caracterizada pela conexão entre dois ou mais objetos inteligentes sem a necessidade direta de intervenção humana para trocar informações ou cooperar atividades.

Espera-se que as redes celulares possuam um papel fundamental na implantação da IoT, como também na comunicação M2M, destacando-se a rede LTE para esta implantação (AT-ZORI; IERA; MORABITO, 2010; MIORANDI et al., 2012). Porém, melhorias na rede LTE são necessárias devido às características intrínsecas da comunicação M2M, como a massiva quantidade de dispositivos, a diversidade dos requisitos de Qualidade de Serviço (QoS) e o tráfego principal no sentido do *uplink*, que diferenciam da comunicação H2H para a qual a rede foi inicialmente projetada (CHEN; WANG, 2010). Dentre essas melhorias, há o aperfeiçoamento no escalonamento de pacotes no *uplink* definido no padrão para que trate o caso da comunicação M2M e atenda aos seus requerimentos sem causar grandes impactos na comunicação H2H.

A transmissão de dados via rádio entre os dispositivos e a rede LTE é feita através

de canais compartilhados que oferecem uma quantidade limitada de recursos de rádio necessários para a transmissão. Estes recursos são alocados para uso exclusivo, durante um intervalo de tempo, para os dispositivos que necessitam transmitir dados. Desta forma, é papel do escalonador de pacotes, tanto do *downlink* como do *uplink*, decidir, com base em parâmetros que refletem a condição do sistema, quais dispositivos vão receber quantos e quais recursos de rádio com a finalidade de satisfazer as necessidades de cada dispositivo e também de maximizar um objetivo global do sistema.

Dentro deste contexto, este trabalho aborda a problemática da alocação de recursos compartilhados de rádio na rede LTE, levando em consideração tanto a comunicação M2M e a H2H, propondo um mecanismo de escalonamento de pacotes para o *uplink*. O foco desta dissertação no *uplink* é devido a característica de que o principal tráfego dos dispositivos da comunicação M2M ser neste sentido pois a geração de dados é maior que a recepção deles (WU et al., 2011). Os objetivos deste trabalho serão discutidos na próxima seção.

1.2 Objetivos

A seguir estão apresentados os principais objetivos do trabalho desenvolvido nesta dissertação. Eles estão divididos em objetivo geral e objetivos específicos.

1.2.1 Objetivo geral

O objetivo geral deste trabalho é desenvolver um mecanismo para o escalonamento de pacotes no *uplink* da rede LTE que utilize informações atuais e passadas da condição da rede e dos dispositivos para satisfazer os requerimentos de QoS das aplicações que utilizam a comunicação M2M; evitar o problema de inanição, ou seja, garantir que todos os dispositivos recebam uma quantidade justa de recursos ao longo do tempo; e, além disso, evitar causar grandes impactos na comunicação H2H ao implantar a comunicação M2M na rede.

1.2.2 Objetivos específicos

Para alcançar o objetivo geral proposto, objetivos específicos foram definidos. Estes subobjetivos são descritos a seguir:

- Estabelecer os requisitos que devem ser atendidos pela solução que será proposta, considerando as características da comunicação M2M e da H2H.
- Propor um mecanismo de escalonamento de pacotes no *uplink* que atenda os requisitos estabelecidos.
- Implementar e analisar os resultados para verificar a validade da solução proposta.

1.3 Planos de atividades

1. **Revisão bibliográfica.** Inicialmente foi realizado um levantamento bibliográfico sobre os conceitos de Internet das Coisas e comunicação Máquina-a-Máquina. Também foi realizado um estudo sobre a arquitetura e os protocolos da rede LTE. Além disso, foi realizada uma revisão bibliográfica sobre as dificuldades e os problemas da implantação da comunicação M2M na rede LTE. Por fim, um levantamento bibliográfico sobre os algoritmos de escalonamento de pacotes no *uplink* que levam ou não em consideração a comunicação M2M foi feito.
2. **Definição do problema e especificação da solução proposta.** Utilizando o estudo das soluções atuais existentes na literatura, esta atividade foi realizada para definir o problema para o qual se propõe a solução. O estudo de algoritmos de escalonamento que não consideram a comunicação M2M também foi realizado para especificar a solução proposta.
3. **Implementação e execução dos experimentos.** A solução proposta e os experimentos foram implementados e simulados no simulador de redes NS-3 (NS-3, 2013), utilizando a linguagem C++. Além da solução proposta, foram implementados também: o escalonamento **Proporcional Fair** (LEE et al., 2009) como solução base que não trata distintamente a comunicação M2M da H2H e uma solução escolhida para comparação, o segundo algoritmo apresentado em (LIOUMPAS; ALEXIOU, 2011).
4. **Análise dos resultados obtidos.** Após a realização dos experimentos, foram coletados e analisados os dados provenientes das simulações. As análises e as discussões dos resultados obtidos são apresentadas no capítulo 5.

1.4 Organização da dissertação

O conteúdo restante desta dissertação está organizado nos capítulos subsequentes, da seguinte forma:

- **Capítulo 2.** Apresenta a revisão bibliográfica realizada, destacando as definições, principais características e aplicações da comunicação M2M e da IoT. É também abordada uma visão geral da arquitetura da rede LTE e suas características. Além disso, o escalonamento de pacotes também é apresentado. Por fim, os problemas da introdução da comunicação M2M na rede LTE são discutidos.
- **Capítulo 3.** Neste capítulo, são apresentados os trabalhos relacionados com a proposta desta dissertação. A descrição das soluções assim como uma análise crítica delas também são abordadas.
- **Capítulo 4.** Neste capítulo é apresentada a definição do problema para o qual este trabalho propõe solução. Além disso, duas abordagens para estender as classes de QoS da

rede LTE para suportar a comunicação M2M são propostas. Para finalizar, este capítulo descreve os detalhes do mecanismo para o escalonamento de pacotes proposto nesta dissertação.

- **Capítulo 5.** Este capítulo define os experimentos realizados. Por fim, os resultados obtidos e uma análise sobre estes resultados são discutidos neste capítulo.
- **Capítulo 6.** Neste capítulo estão as considerações finais deste trabalho após a análise dos resultados obtidos. Este capítulo também aponta as principais contribuições deste trabalho. Finalmente, as perspectivas de trabalhos futuros são apresentadas.

2 REVISÃO BIBLIOGRÁFICA

Neste capítulo é apresentada a revisão bibliográfica dos assuntos necessários para a compreensão deste trabalho. Na seção 2.1, os principais conceitos e definições da Internet das Coisas são apresentados. A seção 2.2 apresenta a definição, as principais características, as aplicações e suas classificações da comunicação M2M. Uma visão geral das características e da arquitetura da rede LTE é descrita na seção 2.3. A seção apresenta ainda o escalonamento de pacotes da rede.

2.1 Internet das coisas

A IoT é um novo paradigma que vem ganhando destaque no cenário das telecomunicações (ATZORI; IERA; MORABITO, 2010). A ideia básica sobre este conceito é que uma grande variedade de objetos do nosso dia-a-dia (eletrodomésticos, automóveis, celulares etc) assim como sensores, atuadores, máquinas industriais, entre outros, possam ser identificados e interligados à Internet para trocar informações e tomar decisões para atingir objetivos em comuns (ATZORI; IERA; MORABITO, 2010). Portanto, cria-se uma junção entre o mundo físico e o mundo das informações ou mundo virtual (UCKELMANN; HARRISON; MICHAHELLES, 2011), onde há conexão em qualquer momento, em qualquer lugar para qualquer um ou qualquer coisa (também chamado de objeto inteligente) (STRATEGY, I. T. U.; UNIT, Policy, 2005).

Embora o conceito básico sobre a IoT seja facilmente visualizado, não há na literatura um consenso na definição da IoT. Segundo Atzori, Iera e Morabito (2010), as definições e as pesquisas para a IoT podem ser classificadas de acordo com o ponto de vista na qual elas abordam o termo "Internet das Coisas". Assim, a IoT pode ser compreendida como um paradigma computacional formado pela sobreposição dos seguintes pontos de vista:

- **Visão orientada aos objetos.** Segundo Bassi e Horn (2008), a IoT envolve objetos que possuem identidades e personalidades virtuais que operam em espaços inteligentes e utilizam interfaces inteligentes para se conectar e se comunicar dentro de contextos sociais, ambientais e de usuário. Os trabalhos que focam nesta visão buscam a integração e identificação dos objetos inteligentes em um arcabouço comum (ATZORI; IERA; MORABITO, 2010).
- **Visão orientada à Internet.** Além da visão orientada aos objetos, Bassi e Horn (2008) define a IoT como uma rede mundial de objetos interligados, endereçáveis e se comunicando através de protocolos padronizados de comunicação. Os trabalhos que focam neste ponto de vista procuram criar modelos e técnicas para a interoperabilidade dos dispositivos em rede (ATZORI; IERA; MORABITO, 2010).
- **Visão orientada à semântica.** Devido ao grande número de objetos inteligentes que estão conectados à Internet, os trabalhos focados neste ponto de vista apresentam propostas

para a apresentação, o armazenamento, a interconexão, a pesquisa e a organização das informações geradas por estes objetos (ATZORI; IERA; MORABITO, 2010).

Baseando-se nas previsões que bilhões para trilhões de objetos inteligentes estarão conectados à Internet em um futuro próximo (WU et al., 2011), então, a principal forma de comunicação da Internet atual será alterada. Ou seja, passará de uma comunicação direcionada aos humanos (comunicações H2H, H2M e M2H) para uma comunicação direcionada aos objetos ou também chamada de comunicação M2M. No restante deste trabalho, nós referenciaremos a comunicação direcionada aos humanos somente como comunicação H2H e estão inclusos neste tipo de comunicação serviços tais como: navegação web, *e-mail*, transmissão/reprodução de conteúdo multimídia, voz sobre IP (Voice over IP - VoIP). Desta maneira, a IoT também pode ser vista como a composição e a integração de diversos serviços das comunicações H2H e M2M (ALAM; NIELSEN; PRASAD, 2013).

Um dos desafios da IoT é fornecer conectividade para a grande quantidade de objetos inteligentes. Neste contexto, as redes celulares, mais especificamente a rede LTE, emergem como soluções promissoras para o provisionamento desta conectividade (ATZORI; IERA; MORABITO, 2010; MIORANDI et al., 2012; GOTSIS; LIOUMPAS; ALEXIOU, 2012). Entre os principais fatores para o uso da rede LTE na IoT, destacamos: o baixo custo de implantação das aplicações ao utilizar uma rede já disseminada e padronizada, a alta capacidade da rede, a flexibilidade nas configurações da rede, a interoperabilidade entre outros padrões de rede e o grande alcance da rede com conectividade tanto em ambientes externos como internos. Contudo, como será visto no decorrer desta dissertação, ainda são necessários melhorias na rede LTE para o suporte ao grande volume de objetos conectados à rede.

Nas próximas seções, abordaremos com mais detalhes a comunicação M2M (seção 2.2) e a rede LTE (seção 2.3).

2.2 Comunicação máquina-a-máquina

A comunicação M2M é um novo modelo de negócio originado das tecnologias de telemetria, utilizado para transmissão e coleta automática de dados de uma fonte remota através de um meio físico, como uma rede cabeada ou sem fio ou um híbrido deles, sem ou com pouca intervenção humana (GALETIĆ et al., 2011). A telemetria e a comunicação M2M são baseadas na transmissão de dados, mas enquanto a primeira utiliza meios dedicados ou privados (ex.: sinais de rádio, linhas telefônicas dedicadas) a comunicação M2M utiliza as redes públicas existentes e já bem implantadas para a transmissão (GALETIĆ et al., 2011). A razão para utilizar a infraestruturas dessas redes, principalmente as redes sem fio e celulares, é que a comunicação M2M beneficia-se de alguns de seus aspectos, tais como: a larga cobertura da rede, o uso de dispositivos móveis, a comunicação autenticada e segura, e a facilidade na implantação de baixo custo em ambientes industriais (BOSWARTHICK; ELLOUMI; HERSENT, 2012).

A comunicação M2M pode ser empregada para melhorar os processos existentes em diferentes setores da indústria. Algumas das proeminentes áreas e suas aplicações onde a comunicação M2M pode ser aplicada são apresentadas em (BOSWARTHICK; ELLOUMI;

HERSENT, 2012) e visualizadas na Tabela 2.1.

Geralmente, as aplicações M2M utilizam um grande número de dispositivos de baixo custo, com várias limitações de recurso e podem empregar uma variedade de tecnologias, tais como: sensores, atuadores, identificação por radiofrequência (*radio-frequency identification* - RFID), bluetooth, Wi-Fi (BOSWARTHICK; ELLOUMI; HERSENT, 2012). As limitações de recursos incluem uma carga limitada de energia provida por uma bateria e uma limitada quantidade de processamento e de armazenamento.

A função dos dispositivos das aplicações M2M é monitorar algum aspecto (temperatura, umidade, velocidade, posição, batimento cardíaco, etc.) do ambiente implantado (AMOKRANE; KSENTINI; HADJADJ-AOUL, 2011). Os dados capturados pelo monitoramento são transportados em mensagens (geralmente pacotes IP) por uma rede para um servidor (programa de software). No servidor, os dados são processados e traduzidos em informações úteis (detecção de ameaças, armazenamento de dados, etc.). Finalmente, respostas podem ser enviadas para os dispositivos.

Tabela 2.1: Exemplo de aplicações para a comunicação M2M

| Área | Aplicações |
|---------------------|---|
| Assistência médica | monitoramento remoto de pacientes; assistência domiciliar; rastreamento de recursos hospitalares. |
| Transporte | gerenciamento de frota; provisionamento de conteúdo multimídia; prevenção de roubos e furtos; manutenção veicular; rastreamento de veículos e objetos; cobrança automática de pedágio. |
| Utilidades públicas | monitoramento, provisionamento e cobrança das utilidades públicas tais como: eletricidade, água, gás, etc. |

Baseado no funcionamento das aplicações M2M, uma simples e genérica arquitetura para a comunicação M2M pode ser definida como está ilustrado na Figura 2.1 (AMOKRANE; KSENTINI; HADJADJ-AOUL, 2011). Esta arquitetura é composta de três partes principais:

- **Domínio do dispositivo.** Formado pelos dispositivos e também pelos *gateways*. Os *gateways* podem ser utilizados para interconectar os dispositivos, como também, para redirecionar as mensagens para a rede. Consequentemente, os dispositivos podem conectar-se à rede diretamente ou pelo intermédio de um *gateway*. Além disso, poderá haver interação direta entre os dispositivos através de uma rede de área pessoal.
- **Domínio da rede.** É a infraestrutura da rede responsável pelo transporte de mensagens entre os dispositivos/*gateways* e o servidor.
- **Domínio da aplicação.** Composto pelo servidor e pelos clientes que utilizam os serviços oferecidos pela aplicação.

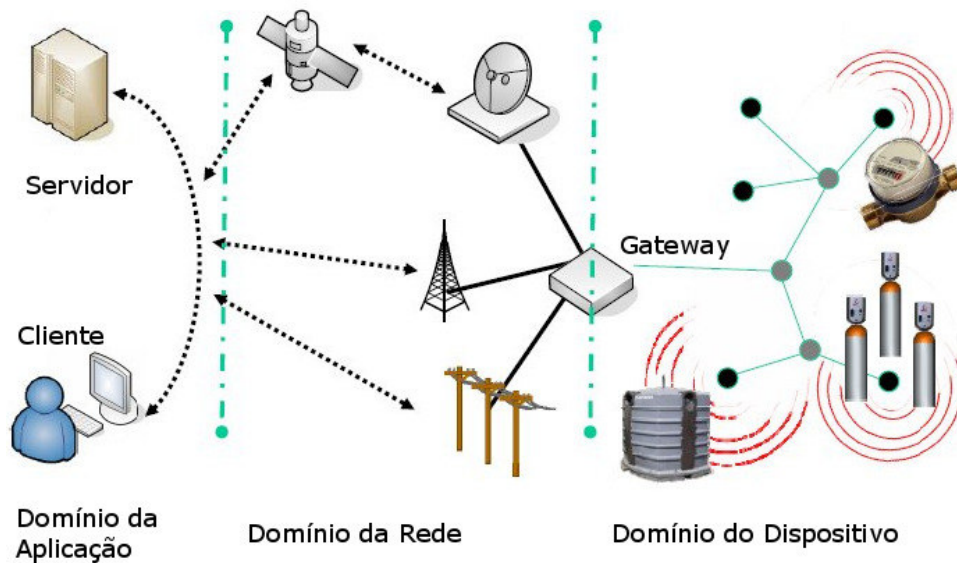


Figura 2.1: Arquitetura para a comunicação M2M. Adaptado de (GALETIĆ et al., 2011)

As aplicações M2M e H2H possuem algumas características em comuns, tais como: conexão segura, mobilidade, comutação de pacotes (LIEN; CHEN; LIN, 2011). Entretanto, como as aplicações M2M desempenham suas funções sem ou com pouca intervenção humana, várias outras características as diferenciam da H2H. Além disso, a diversidade de aplicações M2M também implica em uma grande variedade de características que elas podem possuir. As principais características das aplicações M2M que as diferenciam da H2H são (3GPP, 2012b):

- **Grande número de dispositivos.** As aplicações tipicamente requerem uma grande quantidade de dispositivos implantados em uma mesma área.
- **Baixo consumo de energia.** Os dispositivos podem requerer baixíssimo consumo de energia.
- **Baixa mobilidade.** Os dispositivos possuem baixa ou nenhuma mobilidade.
- **Tempo controlado.** A transmissão de dados é feita somente em intervalos de tempo predefinidos.
- **Tolerância.** Atrasos no envio e recebimento de dados são toleráveis, assim como, falhas são aceitáveis e previstas. Contudo, algumas aplicações requerem transmissões confiáveis e em tempo real.
- **Transmissão de dados de pequeno porte.** Apenas pequenas quantidades de dados são transmitidas por vez.
- **Tráfego principal no sentido ascendente** Os dispositivos geram e enviam mais mensagens para o servidor do que recebem dele. Ou seja, o principal tráfego é no sentido ascendente (do dispositivo para o servidor) ou *uplink*.
- **Transmissão esporádica.** Longo período entre a transmissão consecutiva de dois dados.

- **Agrupamento.** Os dispositivos podem se agrupar. Isto é interessante para questões como cobrança, endereçamento e gerenciamento dos dispositivos de um grupo.

2.2.1 Classificação das aplicações M2M

Em (AMOKRANE; KSENTINI; HADJADJ-AOUL, 2011), os autores baseiam-se nas Redes de Sensores Sem Fio (RSSF) para classificar as aplicações M2M segundo seu tráfego e sua origem. As três classes definidas são:

- **Baseada em Evento (*Event Driven*).** O dispositivo alerta o servidor quando um evento ocorre no ambiente monitorado. Por exemplo, mensagens são enviadas para o servidor quando é detectada a presença de intrusos em aplicações de segurança. As aplicações desta classe normalmente requerem transmissão confiável e em tempo real.
- **Baseada em Consulta (*Query Driven*).** As aplicações interativas estão incluídas nessa classe. Geralmente, o cliente envia consultas para o servidor. O servidor, então, envia a consulta para os dispositivos responsáveis pelo monitoramento do ambiente. Finalmente, os dispositivos enviam a resposta da consulta para o servidor. Outro passo que poderá existir nessas aplicações é que o servidor pode enviar ações que devem ser executadas pelos dispositivos (atuadores) de acordo com as respostas recebidas, criando assim, o laço monitorar-computar-atuar. Por exemplo, uma aplicação que mantém a temperatura de uma sala de acordo com a preferência do usuário está inclusa nesta classe. Esta aplicação comanda o aquecedor ou o ar-condicionado para aumentar, diminuir ou manter a temperatura dependendo da preferência do usuário e da temperatura atual.
- **Baseada em Tempo (*Time Driven*).** A maioria das aplicações M2M estão inclusas nesta classe. Nestas aplicações, os dispositivos enviam seus dados coletados para o servidor em intervalos regulares de tempo (que podem variar de poucos milissegundos até vários minutos ou horas). O servidor é, geralmente, responsável pelo resto do tratamento dos dados.

O trabalho (LIU et al., 2011) também utiliza o tráfego para classificar as aplicações. Porém, outros parâmetros são utilizados na classificação: o conteúdo da mensagem transferida, o tamanho da mensagem, o modo de transmissão e a prioridade de transmissão. As cinco classes são descritas abaixo:

- ***Streaming Móvel.*** Os dispositivos transmitem conteúdos do tipo *streaming* (áudio, vídeo, etc.) para o servidor. Por esta razão, a transmissão é contínua e de alta intensidade. Entretanto, para evitar congestionamentos, a transmissão possui baixa prioridade e possui tolerância a erros.
- ***Smart Metering.*** Nesta classe, os dispositivos enviam uma grande quantidade de dados (vários megabytes) ao serem requisitados pelo servidor. Desta forma, o modo de transmissão é do tipo requisição-resposta. A prioridade de transmissão é baixa para evitar congestionamentos.

- **Monitoramento regular.** É a classe largamente utilizada no mercado e que engloba uma maior variedade de aplicações. Esta classe é caracterizada pelas aplicações cujo os seus dispositivos transmitem periodicamente uma pequena quantidade de dados. Além disso, as aplicações são tolerantes a falhas.
- **Alerta de Emergência.** Mensagens de alerta são enviadas ao detectar situações de emergência. O tipo do conteúdo da mensagem de alerta pode ser *streaming* ou dados normais (não *streaming*) o que afeta o tamanho da mensagem enviada. Devido a urgência, a transmissão precisa ser confiável, em tempo real e ter alta prioridade. Como essa situação de emergência é inesperada, o modo de transmissão é do tipo rajada.
- **Ponto de Venda Móvel (*Mobile Point of Sale - POS*).** Essa classe engloba as aplicações iterativas voltadas à venda e pagamento. Pagamento instantâneo em pedágios inteligentes e em máquinas de venda são exemplos dessas aplicações. Para este tipo de aplicação é recomendado a transmissão confiável e em tempo real, porém a prioridade é baixa. O modo de transmissão é em rajadas devido ao comportamento inesperado das vendas e pagamentos.

De acordo com a análise das classes discutidas acima, os requisitos de QoS das aplicações podem ser definidos (LIU et al., 2011). Estes requisitos são especificados, conforme ilustrados na Tabela 2.2, segundo as propriedades das transmissões das classes (tempo real, exatidão, prioridade).

Tabela 2.2: Requisitos de QoS para as classes de Liu et al. (2011)

| | Tempo real | Exatidão | Prioridade |
|-----------------------|-------------------|-----------------|-------------------|
| Streaming móvel | Alto | Baixa | Baixa |
| Smart metering | Baixo | Alta | Baixa |
| Monitoramento regular | Baixo | Baixa | Baixa |
| Alerta de emergência. | Alto | Alta | Alta |
| POS móvel | Alto | Alta | Baixa |

2.3 Rede celular LTE

O *Long Term Evolution* (LTE) é um padrão criado pela *Third Generation Partnership Project* (3GPP) para a comunicação sem fio e uma evolução dos padrões GSM/UMTS sendo compatível com os mesmos. A rede LTE adota a tendência das redes móveis de se basearem no padrão IP, o protocolo principal da Internet. Desta maneira, o tráfego de voz é suportado principalmente através da tecnologia de VoIP. Portanto, a rede LTE utiliza a comutação de pacotes ao invés da comutação de circuito no tráfego de voz típica das redes celulares de terceira geração. Os objetivos da rede LTE são aumentar a capacidade e a velocidade da redes sem fio utilizando novas tecnologias, além de simplificar arquitetura da rede e reduzir a latência (3GPP, 2013a).

A rede oferece um rico conjunto de características com uma grande flexibilidade em termos de opções de implantação e potenciais ofertas de serviços. Algumas das características mais importantes da rede são (3GPP, 2013a):

- Sua arquitetura é baseada totalmente sobre o IP e a comutação de pacotes.
- Suporte para interoperação e co-existência com os padrões antigos (GSM/UMTS, CDMA2000, etc.).
- Utiliza a tecnologia de acessos múltiplos *Orthogonal Frequency-Division Multiplexing Access* (OFDMA) no *downlink* para obter um alto pico da taxa de transferência em uma largura de banda de alto espectro (até 300 Mb/s com 20 MHz de largura de banda).
- Utiliza a tecnologia de acessos múltiplos *Single Carrier Frequency-Division Multiplexing Access* (SC-FDMA) no *uplink* com taxa de transferência de até 75 Mb/s com a largura de banda de 20 Mhz.
- Suporta a duplexação por divisão de tempo (Time Division Duplex - TDD) e a por divisão de frequência (Frequency Division Duplexing - FDD).
- A rede pode operar em alocações de espectro de diferentes tamanhos para o *downlink* e *uplink*, tais como: 1.4 MHz, 3 MHz, 5 MHz, 10 MHz, 15 MHz, and 20 MHz.
- Suporte a garantias de QoS.

Para garantir que a LTE cumpra os requerimentos das tecnologias de quarta geração das redes celulares definidos pela *International Telecommunication Union* (ITU) (ITU-R, 2008), como o pico da taxa de dados no *downlink* de até 1 Gb/s, a 3GPP estendeu as capacidades da rede a partir da sua versão 10. Por esta razão, a rede LTE também pode ser referida como *LTE Advanced* (LTE-A) desde esta versão. Nesta dissertação, não diferenciaremos a rede LTE da LTE-A.

2.3.1 Arquitetura da rede

A arquitetura geral da rede é composta de duas sub-redes como ilustrado na Figura 2.2: a rede de acesso de rádio (RAN - *Random Access Network*) chamada de *Evolved Universal Terrestrial Radio Access Network* (E-UTRAN) e a rede de núcleo (CN - *Core Network*) denominada *System Architecture Evolution* (SAE) ou também chamada de *Evolved Packet Core* (EPC). A combinação dessas duas sub-redes é nomeada de *Evolved Packet System* (EPS).

A rede EPC realiza o roteamento das informações transmitidas e as funções relacionadas sobre este roteamento, tais como: gerenciamento de mobilidade, tratamento de QoS, política de controle dos fluxos dos dados do usuário, interconexão com redes externas.

A E-UTRAN é caracterizada por uma rede de estações base, também chamadas de *evolved-NodeB* (eNB), que interagem com os dispositivos nomeados de *User Equipment* (UE).

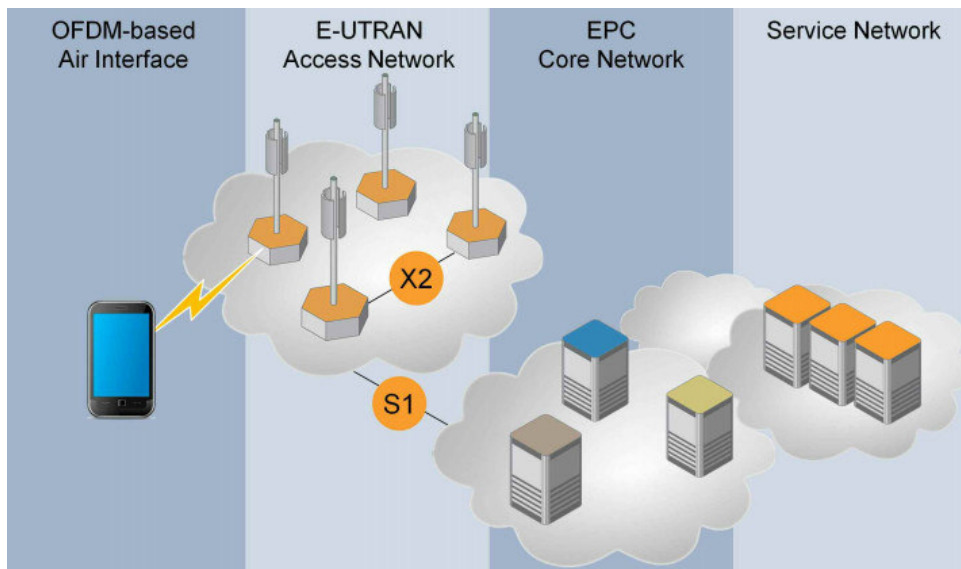


Figura 2.2: Arquitetura da rede LTE. Reproduzido de (TEKTRONIX, 2012)

As eNBs estão conectadas entre elas e com a rede de núcleo. A estação base é responsável por várias funções, entre elas: o gerenciamento dos recursos de rádio; a compressão do cabeçalho IP; a criptografia de dado; o escalonamento e alocação de recursos de rádio tanto do *downlink* como do *uplink*; e a coordenação de *handover* entre as eNBs vizinhas. Como ilustrado na Figura 2.3, as camadas da pilha de protocolos da rede E-UTRAN são:

- **Física.** Carrega todas as informações dos canais de transporte da camada MAC sobre a interface aérea. Esta camada também exerce as funções de adaptação do *link*, de controle de potência, de pesquisa de células (para fins de sincronização e *handover*) e de prover algumas medições para a camada RRC.
- **MAC.** A camada de controle de acesso ao meio (MAC - *Media Access Control*) oferece um conjunto de canais lógicos para a camada RLC que são multiplexados nos canais de transporte da camada física. A camada MAC também gerencia a correção de erros, lida com a priorização dos canais lógicos para o mesmo UE e o escalonamento de recursos de rádio entre os dispositivos (UEs).
- **RLC.** A camada *Radio Link Control* (RLC) executa as tarefas de correção de erros, segmentação/concatenação de dados, re-ordenação para entrega ordenada e detecção de duplicatas.
- **PDCP.** A camada *Packet Data Convergence Protocol* (PDCP) é principalmente responsável pela compressão/descompressão do cabeçalho dos pacotes IP, por assegurar a entrega ordenada de dados e também por realizar a criptografia dos pacotes.
- **RRC.** A camada *Radio Resource Control* (RRC) cuida do sistema de difusão de informações, da paginação, do estabelecimento e da liberação das conexões RRC, do gerenciamento das chaves de segurança, do *handover* e das funções de mobilidade e de gerenciamento de QoS.

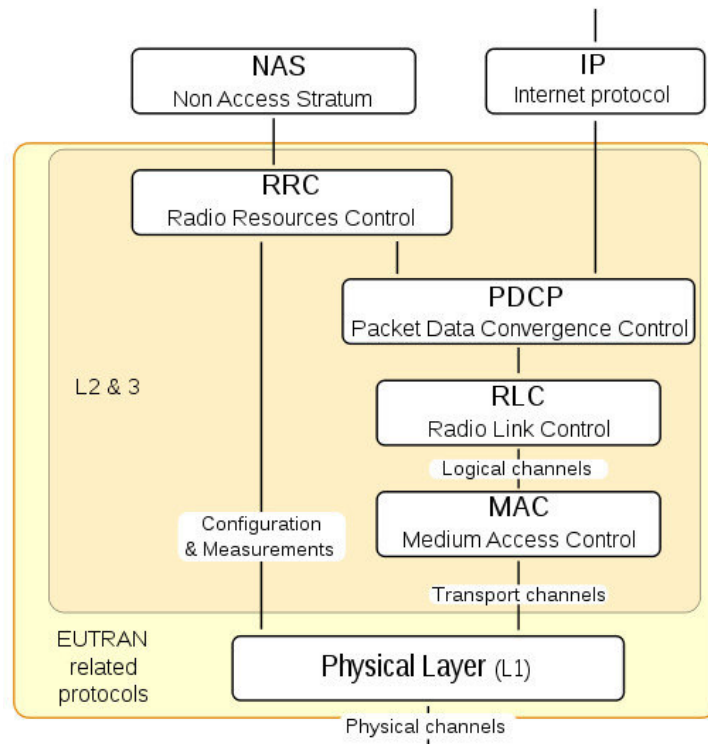


Figura 2.3: Pilha das camadas da interface E-UTRAN. Reproduzido de (ALI-YAHIYA, 2011)

2.3.2 Estrutura de quadros

A camada física suporta dois esquemas de acesso como previamente descritos: o OFDMA para o *downlink* e o SC-FDMA para o *uplink*. Ambos os esquemas utilizam a mesma estrutura de quadros. Esta estrutura é dividida em dois domínios (tempo e frequência) e representada na forma de eixos na Figura 2.4.

No domínio do tempo, cada quadro tem duração de 10 ms e é formado por 10 sub-quadros de 1 ms cada. Um sub-quadro constitui um TTI (*Transmission Time Interval*) e é formado por dois *slots* cada um com 0,5 ms de duração. Finalmente, um *slot* é formado por 7 símbolos.

No domínio da frequência, a largura de banda total utilizada pela rede (1,4, 3, 5, 10, 15 ou 20 MHz) é dividida em sub-canais com 12 sub-portadoras de 15 KHz, totalizando 180 KHz.

O bloco de 7 símbolos X 12 sub-portadoras (Tempo X Frequência) constitui a unidade básica de recurso de rádio chamada de bloco de recurso (*Resource Block - RB*). A quantidade total de blocos de recursos (RBs) depende da largura de banda utilizada cujo os valores são definidos no padrão do sistema LTE. Por exemplo, 25 RBs para 5 MHz e 50 RBs para 10 MHz.

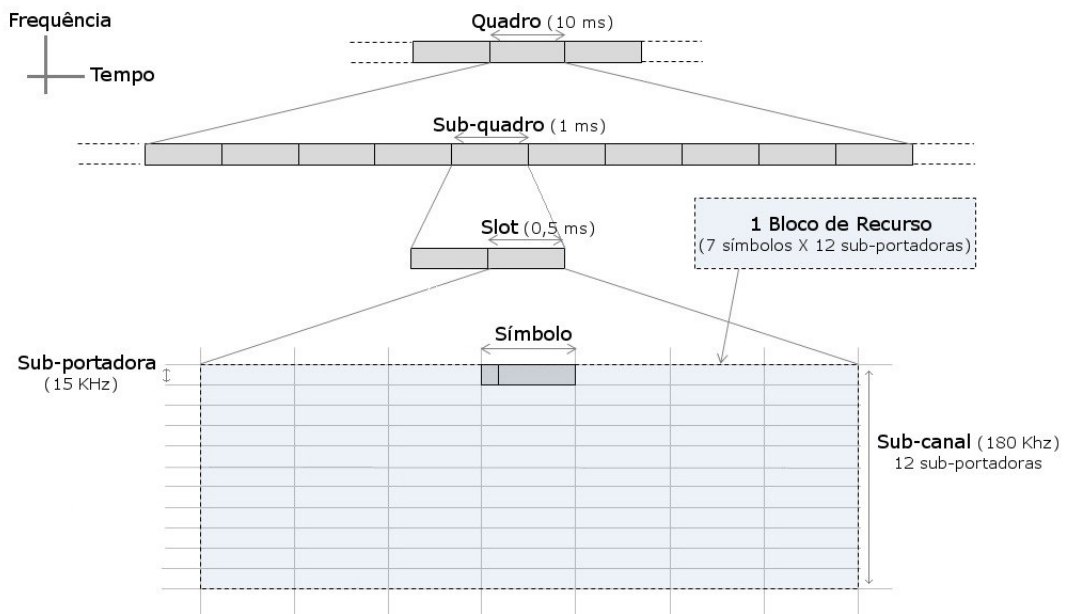


Figura 2.4: Estrutura de quadros do LTE. Adaptado de (AMBERKAR, 2013)

2.3.3 Classes de QoS

O padrão LTE suporta garantias de QoS para os fluxos de tráfego. Estes tráfegos são categorizados segundo seus requisitos de QoS em classes que refletem os serviços e as aplicações suportadas pela rede (*e.g.*, VoIP, *streaming* de vídeo ou áudio e web) (3GPP, 2013b).

Como apresentado na Tabela 2.3, as classes de QoS são identificáveis através do QCI (*QoS Class Identifier*). Estas classes são caracterizadas pelo conjunto dos seguintes parâmetros: (i) o atraso máximo tolerável entre o dispositivo e o *gateway* de saída da rede; (ii) a taxa máxima de perda de pacotes; (iii) se a taxa de bits é garantida (*Guaranteed Bitrate* - GBR) ou não (Non-GBR); e (iv) a prioridade com que os tráfegos devem ser manipulados na rede, sendo que quanto menor o valor maior a prioridade.

Ao longo desta dissertação, os dispositivos das aplicações que utilizam a comunicação M2M serão nomeados de UE M2M, Consequentemente, UE H2H será utilizado para o caso da comunicação H2H. Quando o tipo da comunicação não for especificado, o termo UE será utilizado para referenciar-se a qualquer dispositivo, não importando o tipo de comunicação associado. Além disso, para ter simplicidade sem perder generalidade, consideraremos que cada UE possui um único tipo de tráfego. Desta forma, poderemos utilizar as expressões que são relacionados ao tráfego diretamente ao UE. Por exemplo, o QCI do UE refere-se ao QCI associado ao único tipo de tráfego gerado pelo UE.

2.3.4 Escalonamento de pacotes

O escalonamento de pacotes refere-se a atividade de alocação de recursos de rádio (RBs) necessários para a transmissão de dados, para os UEs. Esta atividade é realizada a cada TTI (1 ms) pelos escalonadores de pacotes do *downlink* e do *uplink* localizados na camada

Tabela 2.3: Classes de QoS do padrão do sistema LTE.

| QCI | GBR | Prioridade | Atraso Máximo | Perda de Pacotes | Aplicações |
|-----|-----|------------|---------------|------------------|---|
| 1 | Sim | 2 | 100 ms | 10^{-2} | VoIP |
| 2 | Sim | 4 | 150 ms | 10^{-3} | Live streaming |
| 3 | Sim | 3 | 50 ms | 10^{-3} | Jogos em tempo real |
| 4 | Sim | 5 | 300 ms | 10^{-6} | Buffered streaming |
| 5 | Não | 1 | 100 ms | 10^{-6} | Sinalização IMS |
| 6 | Não | 6 | 300 ms | 10^{-6} | Baseados em TCP (web, email, ftp, chat, etc.) |
| 7 | Não | 7 | 100 ms | 10^{-3} | Vídeo |
| 8 | Não | 8 | 300 ms | 10^{-6} | Baseados em TCP (web, email, ftp, chat, etc.) |
| 9 | Não | 9 | 300 ms | 10^{-6} | Baseados em TCP (web, email, ftp, chat, etc.) |

MAC da eNB (estação base). Desta forma, o escalonador do *downlink* tem a função de escolher quais RBs serão utilizados para cada UE no próximo período de transmissão (próximo TTI) na transmissão no sentido descendente (da eNB para o UE). Similarmente, o escalonador do *uplink* decide quais RBs serão utilizados para os UEs na transmissão no sentido ascendente (do UE para a eNB).

Uma importante característica que difere os dois escalonadores, é que no *uplink*, todos os RBs alocados para um único UE devem ser contínuos em relação à frequência. Contudo, o escalonador do *downlink* não possui esta restrição.

Os escalonadores tomam a decisão de quais recursos serão alocados para quais UEs com a finalidade de satisfazer as necessidades de recursos dos UEs, assim como, maximizar um objetivo global do sistema. Ressaltando que como a 3GPP define apenas o protocolo de troca de mensagens do escalonamento para o padrão do sistema LTE, cabe ao fornecedor da rede decidir qual o objetivo pretendido e, conseqüentemente, qual solução para o escalonamento utilizar.

Várias soluções para o escalonamento podem ser encontradas na literatura (KWAN; LEUNG, 2010). Os principais objetivos dessas soluções são obter o melhor desempenho das seguintes medidas ou uma combinação delas:

- **Taxa de transferência.** Garantir o uso eficiente da largura de banda para maximizar a quantidade de dados transmitidos. Os trabalhos (CALABRESE et al., 2008), (WANG; KONISHI, 2010) e (LIMA et al., 2012) propõem soluções com o objetivo de maximizar a taxa de transferência.
- **Justiça.** Assegurar alocação justa de recursos entre os dispositivos. Conseqüentemente, evitar o problema de inanição, *i.e.*, garantir que todos os UEs irão receber recursos ao longo do tempo. Exemplos de escalonadores que possuem este objetivo são apresentados

em (AHMED; MOHAMED; SHAKEEL, 2010) e (LEE et al., 2009).

- **Satisfação de QoS.** Maximizar o provimento de QoS para os UEs. Os trabalhos (ALI; ZEESHAN, 2012), (ANAS et al., 2008) e (MARWAT et al., 2012) propõem soluções para satisfazer o requisito de QoS de atraso máximo tolerável. Os trabalhos (MEHTA; KHA-KUREL; KARANDIKAR, 2012) e (PHAM; VU; HWANG, 2011) possuem o objetivo de diminuir as perdas de pacotes.
- **Utilização de energia.** Minimizar o consumo energético dos dispositivos na transmissão de dados. Exemplos de soluções com este objetivo podem ser encontradas em (SOK-MEN; GIRICI, 2010) e (GHANDOUR; FRIKHA; TABBANE, 2011).

2.3.5 Mensagens de controle para o escalonamento

Como discutido anteriormente, o padrão LTE define apenas para o escalonamento o protocolo de troca de mensagens entre a eNB e os UEs que estão conectados à esta eNB. Além disso, como o tráfego principal gerado pela comunicação M2M é no sentido ascendente (do dispositivo para o servidor remoto/eNB), este trabalho foca no escalonamento de pacotes do *uplink*. Diante disso, esta subseção aborda as principais mensagens de controle utilizadas no escalonamento do *uplink*.

O mecanismo do escalonamento de pacotes é acionado pelo UE ao enviar para a eNB uma requisição de escalonamento (*Scheduling Request* -SR). Esta mensagem de controle informa à eNB que o UE possui uma quantidade incerta de dados para transmitir. Para que a eNB saiba a quantidade de dados que cada UE pretende transmitir, os dispositivos precisam enviar à estação base este valor. O relatório sobre o estado do buffer (*Buffer Status Report* -BSR) informa a quantidade de dados armazenados no *buffer* à espera de serem transmitidos. O tamanho do dado informado na mensagem de controle do BSR é indexado em intervalos predefinidos pelo padrão do sistema LTE.

Outra informação que pode ser utilizada no escalonamento é a qualidade no canal de transmissão entre o UE e a eNB. Para o *uplink*, a informação do estado do canal (*Channel State Information* - CSI) é medida pela relação sinal-ruído mais interferência (*Signal-to-Interference-plus-Noise Ratio* - SINR). Para fazer esta mediação, o UE envia à eNB o sinal de referência SRS (*Sounding Reference Signal*) em relação a toda largura de banda ou uma parte dela. Ou seja, o SRS é utilizado para calcular a qualidade da transmissão entre o UE e todos os RBs disponíveis ou entre um subconjunto deles.

Após o escalonamento dos recursos ser realizado, o resultado dele é informado para os dispositivos que receberam os recursos através da mensagem de controle de recurso concedido (*scheduling grant*). Esta mensagem contém todas as informações necessárias para a transmissão, como os recursos alocados e o esquema de modulação a ser utilizado.

3 TRABALHOS RELACIONADOS

Várias soluções sobre o escalonamento de pacotes no *uplink* da rede LTE podem ser encontradas na literatura (KWAN; LEUNG, 2010) no contexto do tráfego presente atualmente na rede. Contudo, estas soluções, geralmente, não são adequadas para a comunicação M2M por causa da suposição da pouca quantidade de serviços e requisitos de QoS existentes da H2H (*e.g.*, voz, vídeo e web) que diferencia da ampla variedade de aplicações M2M e de seus requisitos de QoS (LIOUMPAS; ALEXIOU, 2011). Outro fator importante da não adequação é que essas soluções não tratam do congestionamento causado pelo uso dos recursos compartilhados entre a grande quantidade de dispositivos M2M implantados em uma mesma área.

Neste capítulo, serão discutidas algumas das soluções que consideram a comunicação M2M presente na rede. Estas soluções serão comparadas segundo seus objetivos e parâmetros utilizados para cumpri-los.

Em (LIEN; CHEN; LIN, 2011), os autores propõem um escalonador baseado em grupos no qual os dispositivos (UEs M2M) se agrupam em classes, criadas dinamicamente, segundo seus requisitos de QoS. Os requisitos de QoS dos dispositivos em uma classe são definidos por dois parâmetros: (i) taxa de chegada de pacotes e (ii) o máximo *jitter* tolerável. A classe com maior taxa de chegada de pacotes possui maior prioridade sendo que seus dispositivos só podem requisitar recursos em intervalos inversamente proporcionais a esta taxa. Desta maneira, esta solução utiliza o padrão do tráfego para garantir seus requisitos de QoS. Entretanto, em cenários onde há uma grande quantidade de dispositivos, o escalonamento não é justo para os dispositivos nas classes com baixas prioridades. Ademais, a solução não utiliza as informações da qualidade do canal entre o dispositivo e a estação base que influencia na taxa de transferência.

Duas soluções são apresentadas em (LIOUMPAS; ALEXIOU, 2011) com o objetivo de garantir os requisitos de QoS. Para isto, dois parâmetros são utilizados: (i) a qualidade do canal entre o dispositivo e a estação base, e (ii) o atraso máximo tolerável como métrica de QoS. As duas soluções se diferenciam pelo peso dos dois parâmetros na tomada de decisão. Na primeira, a qualidade do canal tem maior peso na priorização, enquanto na segunda solução, o maior peso é dos UEs M2M menos toleráveis ao atraso. Outra característica das soluções apresentadas em (LIOUMPAS; ALEXIOU, 2011) é que cada dispositivo envia para a rede seus valores para os parâmetros de QoS. Contudo, as duas soluções não são justas para os UEs M2M que possuem as piores condições no parâmetro mais relevante. Por exemplo, no cenário onde há uma grande quantidade de dispositivos, a primeira solução não é justa na alocação de recursos para os UEs M2M que possuem uma baixa qualidade do canal e a segunda solução não é justa para os UEs M2M mais toleráveis ao atraso.

Assim como nas RSSF, quando um dispositivo captura um evento e faz uma requisição por recursos para transmitir os dados capturados desse evento, pode ocorrer um aumento da probabilidade de que os outros dispositivos, que estão próximos dele e que também monitoram o mesmo evento, também requisitem recursos para transmitir os dados do mesmo evento. Brown e Khan (2013) utiliza esta característica para alocar recursos para os UEs M2M antes

que eles façam a requisição por recursos, diminuindo, assim, o atraso da transmissão. Para isto, a solução tenta prever quais dos UEs M2M vizinhos e também relacionados ao dispositivo que fez a requisição de recursos diretamente à rede, irão também fazer requisições de recursos nos próximos instantes. Porém, previsões erradas, que desperdiçam recursos, agravam ainda mais a situação de escassez de recursos existente ao introduzir a comunicação M2M na rede.

Em um cenário real, ambas as comunicações H2H e M2M estarão presentes na rede. Neste cenário, a escassez de recursos de rádio e, conseqüentemente, o congestionamento da rede será um grande problema. A principal causa deste problema é que embora cada dispositivo M2M transmita pouca quantidade de dados resultando numa menor necessidade de recursos de rádio, a grande quantidade dos UEs M2M requisitando recursos leva à uma escassez destes. Em vista disso, a comunicação H2H será gravemente afetada caso o escalonamento não trate este problema. Considerando isto, nenhuma das soluções apresentadas acima tratam do impacto da comunicação M2M sobre a H2H.

A solução apresentada em (ZHENQI et al., 2013) controla este impacto ao dividir o escalonamento em duas partes. Na primeira parte, o escalonamento é feito para os UEs H2H com o objetivo de satisfazer os requisitos de QoS e garantir justiça na alocação dos recursos. Na segunda parte, o escalonamento é feito para os UEs M2M com os recursos restantes após a primeira parte e com os mesmos objetivos do primeiro. Entretanto, poderá ocorrer que após a primeira parte do escalonamento não haja mais recursos disponíveis para os UEs M2M em cenários onde há um grande tráfego H2H. Causando desta maneira, o problema de inanição para os UEs M2M, mesmo que o escalonamento seja justo na alocação de recursos entre os UEs M2M. Além disso, esta solução considera apenas as aplicações baseadas em tempo, de forma que um grande nível de não satisfação dos requisitos de QoS seja aceitável devido à tolerância a atrasos dessas aplicações.

O trabalho (GOTSIS; LIOUMPAS; ALEXIOU, 2013) propõe modificações em (LIEN; CHEN; LIN, 2011) de forma que o impacto da comunicação M2M na H2H seja controlado ao reservar uma quantidade constante de recursos para a comunicação M2M. Além disso, para classificar o tráfego nas classes de QoS, o processo de Poisson, aplicado na modelagem da transmissão em rajadas das aplicações baseadas em evento como proposto em (3GPP, 2012a), é utilizado ao invés do intervalo regular e constante das transmissões das aplicações baseadas em tempo. Outra diferença entre as soluções propostas em (GOTSIS; LIOUMPAS; ALEXIOU, 2013) e em (LIEN; CHEN; LIN, 2011), é que a primeira utiliza uma quantidade predeterminada de classes, enquanto na segunda as classes são criadas dinamicamente. No entanto, a não garantia de justiça e o não aproveitamento da informação da qualidade do canal ainda é existente em (GOTSIS; LIOUMPAS; ALEXIOU, 2013). Também a separação constante dos recursos para as comunicações M2M e H2H causa um mau aproveitamento deles. Por exemplo, em situações onde somente os dispositivos H2H requisitam recursos.

Em (ZHOU et al., 2012) complementada em (ZHOU; NIKAEIN; KNOPP, 2013), o escalonamento proposto nestes trabalhos também reserva uma quantidade de recursos para os dispositivos M2M. Porém, o acesso a estes recursos é baseado em contenção. Ou seja, ao invés de requisitar recursos à rede, os dispositivos escolhem aleatoriamente os recursos, dentre aqueles reservados, e os utilizam sem informar à qualquer entidade esta escolha. Por conseguinte,

colisões podem existir quando pelo menos dois dispositivos tentam transmitir recursos comuns entre eles. A probabilidade desta colisões pode ser bem alta quando há uma grande quantidade de dispositivos, acarretando deste modo, uma baixa performance do sistema.

A Tabela 3 resume a comparação entre os trabalhos relacionados. Observa-se nesta tabela que como somente o mecanismo de predição é proposto em (BROWN; KHAN, 2013), então os objetivos do escalonamento não podem ser totalmente especificado. Além disso, como a quantidade de recursos utilizados por cada dispositivo é feita de forma independente no escalonamento com contenção proposto em (ZHOU et al., 2012), logo não há alocação igualitária dos recursos entre os dispositivos. Portanto, esta solução pode ser considerada injusta.

Com base nas deficiências das soluções anteriormente discutidas, este trabalho propõe uma novo mecanismo para o escalonamento de pacotes que controla o impacto da comunicação M2M na H2H, garante justiça na alocação de recursos e tenta maximizar a satisfação dos requisitos de QoS. A Tabela 3 também apresenta a comparação com o mecanismo proposto neste trabalho. Mais detalhes sobre este mecanismo serão abordados nos próximos capítulos.

| | Objetivos para a M2M | | | Características principais | Pontos fracos |
|-----------------------------------|----------------------|---------|-----------------------------------|---|--|
| | QoS | Justiça | Controle do Impacto da M2M na H2H | | |
| (LIEN; CHEN; LIN, 2011) | Sim | Não | Não | Agrupamento dos dispositivos | Considere apenas os dispositivos baseados em tempo |
| (LIOUMPAS; ALEXIOU, 2011) | Sim | Não | Não | Utiliza os valores exatos dos parâmetros de QoS | Inanição nos dispositivos com baixas prioridades |
| (BROWN; KHAN, 2013) | Sim | - | Não | Predição das requisições | Uso ineficiente dos recursos em predições erradas |
| (ZHENQI et al., 2013) | Sim | Sim | Sim | Escalonamento separado para as comunicações H2H e M2M | Inanição na M2M no alto tráfego H2H |
| (GOTSIS; LIOUMPAS; ALEXIOU, 2013) | Sim | Não | Sim | Agrupamento dos dispositivos | Uso ineficiente dos recursos ao reservar uma quantidade fixa dos RBs para a M2M |
| (ZHOU et al., 2012) | Sim | Não | Sim | Baseado em contenção | Alta colisão leva ao uso ineficiente dos recursos |
| Mecanismo Proposto | Sim | Sim | Sim | Garante justiça e o Uso eficiente dos recursos | Insatisfação de QoS para a M2M em alto tráfego H2H (mais detalhes no capítulo 5) |

Tabela 3.1: Comparação dos trabalhos relacionados

4 MECANISMO DE ESCALONAMENTO PARA A COMUNICAÇÃO M2M

Esta dissertação aborda a problemática da alocação de recursos de rádio no *uplink* na rede LTE para os dispositivos das aplicações M2M. Neste capítulo, serão abordados os detalhes do mecanismo proposto sobre o problema citado.

As seções a seguir definem o problema da alocação de recursos (seção 4.1); a classificação do tráfego M2M com base nos seus requisitos de QoS (seção 4.2); a divisão de recursos entre as comunicações H2H e M2M segundo as suas demandas (seção 4.3); e por fim, o algoritmo de escalonamento proposto neste trabalho (seção 4.4).

4.1 Formulação do problema

O problema da alocação de recursos no escalonamento de pacotes pode ser definido como um problema de otimização cujo objetivo é maximizar alguma performance desejada do sistema. Alguns dos objetivos foram citados na seção 2.3.4. Esta performance do sistema é expressa pela medida do desempenho da escolha dos recursos alocados para cada dispositivo.

Seja $\mathcal{U} = \{1, \dots, U\}$ o conjunto dos dispositivos (UEs) ativos que requisitaram recursos no TTI t , tal que $\mathcal{U}_{H2H} \cup \mathcal{U}_{M2M} = \mathcal{U}$, onde \mathcal{U}_{H2H} , \mathcal{U}_{M2M} são os conjuntos dos UEs H2H e M2M respectivamente. Considere B como a largura de banda utilizada e com R recursos (RBs) indexados pelo conjunto $\mathcal{RB} = \{1, \dots, R\}$. Assumindo isto, a alocação de recursos tem como objetivo maximizar a função utilitária M_{sys} que mede a performance do sistema como ilustrado nas equações (4.1a), (4.1b), (4.1c).

$$\max M_{sys} \quad (4.1a)$$

$$M_{sys} = \sum_{u \in \mathcal{U}} M_u \quad (4.1b)$$

$$M_u = \sum_{r \in \mathcal{RB}} M_{u,r} A_{u,r} \quad (4.1c)$$

Onde M_u é a função utilitária para o UE u e $M_{u,r}$ é a métrica que avalia a performance da alocação do RB r para o UE u com o intuito de alcançar o objetivo do escalonamento. $A_{u,r} = \{0, 1\}$ é uma função no qual o valor é 1 se o RB r for alocado para o UE u e 0 caso contrário. Além disso, as seguintes restrições são necessárias:

$$\sum_{r=1}^R A_{u,r} \leq 1; \forall u \in \mathcal{U} \quad (4.2)$$

$$\begin{aligned} \forall u \in \mathcal{U} : \\ \forall r (r \in \{r' | A_{u,r'} = 1\}) \implies r \in \{i, i+1, \dots, i+l\} \\ \text{para algum } i, l | 1 \leq i \leq i+l \leq R \end{aligned} \quad (4.3)$$

A primeira restrição assegura que cada RB só pode ser alocado para no máximo um dispositivo. A restrição imposta pela Equação (4.3) assegura que os recursos alocados para um dispositivo sejam contínuos como exigido na escalonamento do *uplink*.

Um simples algoritmo para o escalonamento do *uplink* pode ser definido a partir das definições citadas acima. Este algoritmo faz uma busca exaustiva nas combinações entre os RBs e os UEs respeitando as restrições supracitadas. Assumindo R RBs disponíveis e U UEs ativos, se k dos U receberão recursos, então há $\binom{R+k}{2k}$ opções de k subconjuntos com RBs contínuos (REN et al., 2013). Como existem $\frac{U!}{(U-k)!}$ opções de escolher os k UEs, logo o tamanho do espaço de busca para este caso é igual à $\binom{R+k}{2k} \frac{U!}{(U-k)!}$. Portanto, o tamanho total do espaço de busca é igual a:

$$\sum_{k=0}^U \binom{R+k}{2k} \frac{U!}{(U-k)!} \quad (4.4)$$

Esta abordagem pode ser inviável considerando uma grande quantidade de UEs, como no caso da comunicação M2M, onde o escalonamento possui um tempo máximo baixo de 1 TTI (1 ms) para ser concluído. Portanto, heurísticas podem ser empregadas para reduzir este espaço de busca e encontrar soluções sub-ótimas. Exemplo de heurísticas utilizadas na literatura são: estratégia gulosa (WONG; OTERI; MCCOY, 2009; REN et al., 2013), programação linear (WANG; KONISHI, 2010), teoria dos jogos (YAACOUB; DAWY, 2009), *first maximum expansion* (TEMINO et al., 2008; LIOUMPAS; ALEXIOU, 2011).

Uma vez definido o problema, é necessário especificar os objetivos do mecanismo de escalonamento proposto neste trabalho. O primeiro objetivo é reduzir o impacto da comunicação M2M sobre a H2H. Para medir este impacto a métrica da taxa de transferência é utilizada. Neste mecanismo, o impacto é controlado ao separar os RBs para os dois tipos de comunicação (H2H e M2M). Criando assim, dois subconjuntos de RBs consecutivos, cada um de uso exclusivo para cada tipo de comunicação. Para obter uma melhor eficiência no uso dos recursos, o tamanho de cada subconjunto é variável ao depender da demanda por recursos dos UEs. Na seção 4.3, os detalhes sobre esta divisão de recursos são apresentados.

Outro objetivo do escalonamento proposto nesta dissertação é maximizar o nível de satisfação dos requisitos de QoS dos UEs M2M. O requisito de QoS avaliado é o atraso máximo permitido. Para cumprir este objetivo, os UEs M2M são classificados pelo seus requisitos de QoS. Na seção 4.2, estes requisitos, seus possíveis valores e a classificação são definidos.

O terceiro e último objetivo é evitar o problema de inanição na alocação de recursos para os UEs M2M. Para isto, uma alocação igualitária e justa é utilizada. As funções utilitárias que avaliam o desempenho de cumprir os objetivos da solução proposta desta dissertação, além do algoritmo para o escalonamento são apresentadas na seção 4.4.

4.2 Classes de QoS

Neste trabalho, utilizamos a classificação das aplicações M2M propostas em (AMOKRANE; KSENTINI; HADJADJ-AOUL, 2011). Porém, para utilizar os requisitos de QoS de-

finidos em (LIU et al., 2011), apresentados na Tabela 2.2, nas classes propostas por Amokrane, Ksentini e Hadjadj-Aoul (2011), um relacionamento entre esses dois trabalhos precisa ser realizado como será proposto a seguir. As classes baseadas em evento e em tempo são diretamente relacionadas com as classes alerta de emergência e monitoramento regular respectivamente. As classes *smart metering* e POS móvel são relacionadas à classe baseada em consulta, pois as aplicações interativas com o modo de transmissão do tipo requisição-resposta estão inclusas nessas classes. A classe *streaming* móvel não possui relacionamento por causa da sua transmissão contínua que não se compatibiliza com nenhuma das classes apresentadas em (AMOKRANE; KSENTINI; HADJADJ-AOUL, 2011).

Para definir os requisitos de QoS das classes propostas em (AMOKRANE; KSENTINI; HADJADJ-AOUL, 2011), os seguintes parâmetros serão utilizados: (i) a tolerância a atrasos, (ii) a tolerância a perdas de pacotes e (iii) a prioridade. Considerando que as transmissões em tempo real são poucos tolerantes a atrasos e que as transmissões confiáveis possuem baixas taxas de perda de pacotes, então os valores para os parâmetros dos requisitos de QoS são diretamente especificados da relação entre as duas classificações supracitadas. A Tabela 4.1 informa estes valores.

Tabela 4.1: Requisitos de QoS para as classes de Amokrane, Ksentini e Hadjadj-Aoul (2011)

| | Tolerância a atrasos | Tolerância a perdas de pacotes | Prioridade |
|---------------------|-----------------------------|---------------------------------------|-------------------|
| Baseado em Consulta | Baixo/Alta | Baixa | Baixa |
| Baseado em Tempo | Alta | Alta | Baixa |
| Baseado em Evento | Baixa | Baixa | Alta |

A classificação do tráfego das aplicações M2M nas classes de QoS do padrão LTE pode ser feita ao avaliar os requisitos de QoS das classes das aplicações M2M definidos acima. As aplicações da classe baseada em consulta podem ser relacionadas com os QCI 8 ou 9, pois estas aplicações utilizam o modo de transmissão requisição-resposta típica das aplicações baseadas no protocolo TCP e também possuem baixa prioridade. As aplicações da classe baseada em evento requerem alta prioridade, baixa tolerância a atrasos e a perdas de pacotes. Desta forma, o único relacionamento possível desta categoria é com o QCI 5, contudo, esta classe é utilizada exclusivamente para mensagens de controle da rede. Nenhum relacionamento pode ser realizada entre a categoria das aplicações baseadas no tempo e as classes de QoS do LTE em consequência das propriedades de transmissão em intervalos regulares e da tolerância a atrasos e a perdas de pacotes destas aplicações. Ou seja, nenhuma classe de QoS do padrão se adequa à grande variação do atraso máximo tolerável (de poucos milissegundos até vários minutos) com alta taxa de perda de pacotes e baixa prioridade.

Apesar de que nenhuma classe proposta por Amokrane, Ksentini e Hadjadj-Aoul (2011) é relacionada com a classe *Streaming* Móvel, esta pode ser relacionada com o QCI 7 por possuir baixa tolerância a atrasos, ser tolerante a perdas de pacotes e possuir baixa prioridade.

Apoiando-se no relacionamento supracitado, somente as aplicações baseadas em evento e em tempo serão consideradas aplicações M2M no restante desta dissertação. Isto se deve aos fatos de que as aplicações baseadas em consulta podem ser classificadas em classes de QoS já existentes para o padrão LTE e também de que uma grande parte das aplicações M2M

estão contidas nas outras duas classes (AMOKRANE; KSENTINI; HADJADJ-AOUL, 2011).

Este trabalho propõe duas abordagens para estender as classes de QoS definidas no padrão do sistema LTE para suportar os requisitos de QoS das aplicações M2M. As duas propostas possuem como vantagem a identificação dos dispositivos (M2M ou H2H) e de suas aplicações através da classe de QoS no qual o seu tráfego está relacionado. As abordagens propostas são as seguintes:

1. Na primeira abordagem, duas classes são adicionadas ao padrão. Uma classe será utilizada para as aplicações M2M baseadas em evento e a outra será utilizada para as aplicações baseadas no tempo. Entretanto, o valor do parâmetro de atraso máximo permitido não será especificado nesta última classe e cabe a cada dispositivo das aplicações baseadas em tempo enviar para a rede (estação base) o seu próprio valor para este parâmetro.
2. Na segunda abordagem, $1 + N$ classes são adicionadas ao padrão. Como na abordagem anterior, uma classe será utilizada para as aplicações baseadas em evento, no entanto, N classes são utilizadas para as aplicações baseadas no tempo. Estas N classes possuirão como diferencial entre elas o atraso máximo tolerável. Assim sendo, um UE M2M que transmite seus dados a cada x intervalo de tempo escolherá a classe com o maior atraso máximo permitido igual ou inferior a x .

Na Tabela 4.2 é ilustrado um exemplo das abordagens propostas. A primeira abordagem utiliza somente as duas primeiras classes (QCI 10 e 11) e a segunda conta com todas as classes ilustradas. Os valores dos parâmetros para a classe das aplicações baseadas em evento foram escolhidas para atender os seus requisitos de QoS utilizando valores das classes já existentes no padrão. Além disso, a classe com QCI 10 possui a maior prioridade após a classe utilizada para mensagens de controle (QCI 5). O atraso máximo tolerável foi incrementado para cada classe subsequente das aplicações baseadas no tempo. A diferença do atraso máximo entre duas classes consecutivas aumenta conforme as aplicações sejam mais toleráveis ao atraso. Desta maneira, haverá uma menor granularidade na classificação das aplicações menos toleráveis ao atraso que também possuem maior prioridade.

Nos próximos capítulos e seções, quando as classes de aplicações M2M forem citadas, elas serão referentes à classificação proposta em (AMOKRANE; KSENTINI; HADJADJ-AOUL, 2011). Exceto em situações com citação direta do trabalho no qual a classificação é proposta.

4.3 Cálculo da demanda por recursos

Um dos objetivos deste trabalho é controlar o impacto da comunicação M2M na H2H. Para isto, o escalonador irá reservar RBs para o tráfego dos UEs H2H baseando-se nas demandas por recursos de cada UE H2H. Porém, a quantidade atual de dados no *buffer* de transmissão no *uplink* do UE não é relacionada diretamente com a quantidade de recursos requeridos, pois a taxa de transferência é influenciada, além da quantidade de RBs, pela qualidade do canal

Tabela 4.2: Classes de QoS para as aplicações M2M

| QCI | GBR | Prioridade | Atraso Máximo | Perda de Pacotes | Aplicações |
|-----|-----|------------|---------------|------------------|-------------------|
| 10 | Não | 2 | 50 ms | 10^{-6} | Baseada em Evento |
| 11 | Não | 10 | 50 ms | 10^{-2} | Baseada em Tempo |
| 12 | Não | 11 | 100 ms | 10^{-2} | Baseada em Tempo |
| 13 | Não | 12 | 150 ms | 10^{-2} | Baseada em Tempo |
| 14 | Não | 13 | 300 ms | 10^{-2} | Baseada em Tempo |
| 15 | Não | 14 | 400 ms | 10^{-2} | Baseada em Tempo |
| 16 | Não | 15 | 500 ms | 10^{-2} | Baseada em Tempo |
| 17 | Não | 16 | 1 s | 10^{-2} | Baseada em Tempo |
| 18 | Não | 17 | 5 s | 10^{-2} | Baseada em Tempo |
| 19 | Não | 18 | 15 s | 10^{-2} | Baseada em Tempo |
| 20 | Não | 19 | 30 s | 10^{-2} | Baseada em Tempo |
| 21 | Não | 20 | 1 min | 10^{-2} | Baseada em Tempo |

entre o UE e cada RB alocado. Como nesta fase inicial os RBs ainda não foram alocados, a quantidade de recursos requeridos por UE H2H precisa ser calculada.

Para o cálculo da demanda atual de recurso, este trabalho propõe o uso da relação da quantidade de recursos por tamanho de dado presenciada pelos UE H2H ao longo do tempo. Portanto, o cálculo da demanda de recursos de um UE H2H é realizado utilizando o histórico de alocação de recursos para este dispositivo como mostra a equação a seguir:

$$\widehat{RB}_{H2H}(u,t) = \max \left(RB_{H2H}^{min}, \frac{BS_{H2H}(u,t) \times RB_{H2H}^{avg}(u,t-1)}{BS_{H2H}^{avg}(u,t-1)} \right) \quad (4.5)$$

onde $\widehat{RB}_{H2H}(u,t)$ é a demanda atual de RBs do UE H2H u no TTI t , BS_{H2H} é o tamanho atual dos dados no *buffer*, $RB_{H2H}^{avg}(u,t)$ é a quantidade média de RBs alocados para o dispositivo e $BS_{H2H}^{avg}(u,t)$ é a média do tamanho dos dados no *buffer* de transmissão. Além disso, cada UE H2H possui um limite mínimo $RB_{H2H}^{min} > 0$ de recursos requeridos.

As funções RB_{H2H}^{avg} , BS_{H2H}^{avg} são calculadas utilizando uma média móvel exponencial como mostrado nas equações (4.6) e (4.7) respectivamente, onde α é uma constante medida em TTIs.

$$RB_{H2H}^{avg}(u,t) = \frac{1}{\alpha} RB_{H2H}(u,t) + \left(1 - \frac{1}{\alpha} \right) RB_{H2H}^{avg}(u,t-1) \quad (4.6)$$

$$BS_{H2H}^{avg}(u,t) = \frac{1}{\alpha} BS_{H2H}(u,t) + \left(1 - \frac{1}{\alpha} \right) BS_{H2H}^{avg}(u,t-1) \quad (4.7)$$

A demanda total de recursos de todos os dispositivos H2H no atual TTI t ($\widehat{RB}_{H2H}(t)$) é calculada como a soma das demandas de cada dispositivo tendo um limite superior igual à quantidade total de recursos disponíveis $|\mathcal{RB}|$ como mostrado a seguir:

$$\widehat{RB}_{H2H}(t) = \min \left(|\mathcal{RB}|, \sum_{u \in \mathcal{U}_{H2H}} \widehat{RB}_{H2H}(u,t) \right) \quad (4.8)$$

Entretanto, os UEs H2H podem requerer todos os recursos disponíveis causando inanição nos UEs M2M. Assim sendo, uma porcentagem mínima de recursos ($0 \leq p_{M2M}^{min} \leq 1$) para os UEs M2M é garantida com objetivo de evitar o problema de inanição. Outra importante característica da solução proposta neste trabalho é que como os dispositivos M2M transmitem pouca quantidade de dados por vez (menos de 1000 bits) (3GPP, 2012a), estes dispositivos irão receber uma quantidade fixa e pequena de recursos RB_{M2M}^{min} .

Ao alocar uma quantidade fixa de recursos por UE M2M, a mensagem de controle BSR que reporta a quantidade de dados a espera de transmissão é desnecessária. Assim, esta mensagem de controle não precisa ser enviada, diminuindo a quantidade de mensagens trafegadas na rede e, conseqüentemente, reduzindo o congestionamento.

Com a garantia de uma quantidade mínima de recursos reservada para a comunicação M2M ($p_{M2M}^{min} \times |\mathcal{RB}|$) e que cada UE M2M receberá uma quantidade fixa de recursos (RB_{M2M}^{min}), então pelo menos $(p_{M2M}^{min} \times |\mathcal{RB}|) / RB_{M2M}^{min}$ dispositivos M2M irão receber recursos. Esta quantidade de UEs M2M selecionada pode aumentar caso a demanda de recursos da comunicação H2H seja baixa. A equação (4.9) formulariza esta discussão, onde o segundo argumento da função max calcula a quantidade de UEs M2M que seriam selecionados caso a demanda da comunicação H2H fosse totalmente atendida.

$$U_{M2M}^{max}(t) = \max \left(\left\lfloor \frac{p_{M2M}^{min} \times |\mathcal{RB}|}{RB_{M2M}^{min}} \right\rfloor, \left\lfloor \frac{|\mathcal{RB}| - \widehat{RB}_{H2H}(t)}{RB_{M2M}^{min}} \right\rfloor \right) \quad (4.9)$$

Como mostra a equação (4.10a), a quantidade de recursos reservada para a comunicação M2M ($RB_{M2M}(t)$) é calculada considerando que $|\mathcal{U}_{M2M}|$ dispositivos M2M requisitaram recursos e que no máximo U_{M2M}^{max} dessas requisições serão atendidas (uma requisição por dispositivo). Assim sendo, a quantidade de RBs reservada para a comunicação H2H ($RB_{H2H}(t)$) é igual a quantidade de recursos restantes após reservar os RBs para a comunicação M2M como mostra a equação (4.10b).

$$RB_{M2M}(t) = RB_{M2M}^{min} \times \min(|\mathcal{U}_{M2M}|, U_{M2M}^{max}(t)) \quad (4.10a)$$

$$RB_{H2H}(t) = |\mathcal{RB}| - RB_{M2M}(t) \quad (4.10b)$$

Depois da demanda por recursos de cada tipo de comunicação calculada, os RBs disponíveis são separados em dois subconjuntos com RBs contínuos, em relação à frequência, um para cada tipo de comunicação. O tamanho do subconjunto, referindo-se a quantidade de RBs, é determinado pela quantidade de recursos solicitados pelo tipo de comunicação associado a este subconjunto. A Figura 4.1 ilustra esta divisão.

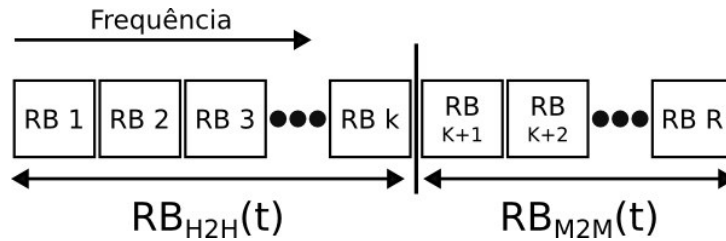


Figura 4.1: Divisão dos recursos entre as comunicações H2H e M2M

4.4 Descrição do algoritmo de escalonamento

Com as demandas de recursos para os dispositivos calculadas, os recursos são divididos em dois grupos consecutivos em relação à frequência, um para cada tipo de comunicação. Deste modo, as alocações de RBs para as comunicações H2H e M2M podem ser tratadas separadamente. Portanto, o mecanismo proposto neste trabalho focará somente no escalonamento para os UEs M2M, beneficiando-se do uso de qualquer solução existente na literatura para o escalonamento dos dispositivos da comunicação H2H.

Baseando-se no modelo especificado em (POKHARIYAL et al., 2007), a modelagem do algoritmo para o escalonamento pode ser dividida em duas fases.

1. **TDPS.** A primeira fase constitui o escalonamento de pacotes no domínio do tempo (*Time Domain Packet Scheduling* - TDPS) que executa a priorização dos dispositivos que fizeram a requisição de recursos. Depois, são escolhidos os dispositivos que possuem maiores prioridades. Os dispositivos não selecionados nesta fase terão seus pedidos por recursos negados.
2. **FDPS.** O escalonamento de pacotes no domínio da frequência (*Frequency Domain Packet Scheduling* - FDPS) constitui a segunda fase no modelo do algoritmo de escalonamento. Esta fase executa a alocação de recursos em si, ao escolher quais recursos serão utilizados pelos dispositivos selecionados na primeira fase. Como discutido na seção 4.1, esta escolha tem o objetivo de maximizar alguma performance desejada do sistema. Após a alocação ser realizada, o resultado desta é enviado para os dispositivos.

Ao dividir o escalonamento em duas fases, a função utilitária que avalia o desempenho do escalonamento para alcançar o objetivo global do sistema (seção 4.1) também é separada nos domínios relacionados às duas fases. A equação (4.11) exemplifica esta discussão, onde $M^{TD}(u, t)$ é métrica da primeira fase para o UE u no TTI t e $M^{FD}(u, r, t)$ é a métrica da segunda fase entre o UE u e o RB r no mesmo instante de tempo t .

$$M_u \rightarrow M^{TD}(u, t), M^{FD}(u, r, t) \quad (4.11)$$

4.4.1 Primeira fase - TDPS

Após a separação dos recursos entre os tipos de comunicação (H2H e M2M), RB_{M2M} recursos estão disponíveis para serem alocados nos UEs M2M, além disso, cada UE M2M poderá receber uma quantidade fixa de recursos igual a RB_{M2M}^{min} . Portanto, no máximo RB_{M2M}/RB_{M2M}^{min} dispositivos M2M poderão receber recursos. Caso esta quantidade seja inferior à quantidade total de UEs M2M, a requisição de recursos deverá ser negada para um conjunto destes. Assim sendo, esta fase escolhe os UEs M2M que receberão os RBs disponíveis para a comunicação M2M.

No mecanismo proposto neste presente trabalho, a decisão de quais UEs M2M irão receber os recursos é feita de tal forma que os dispositivos M2M que receberam menos RBs ao longo do tempo e que também estejam mais próximos de superar o atraso máximo tolerável tenham maiores prioridades na alocação de recursos. Portanto, nesta fase, é garantida a justiça na alocação de recursos. Outro objetivo é a maximização da satisfação dos requisitos de QoS medidos pela métrica de atraso máximo tolerável. Para cumprir estes objetivos, a função utilitária desta fase é composta por duas outras funções utilitárias, uma para cada objetivo.

A primeira função utilitária é a função $M_{M2M}^{TD}delay$ que prioriza os dispositivos segundo seus requisitos de QoS. Esta priorização é realizada por intermédio da função $\Delta D(u)$ que avalia o quão próximo o UE M2M u está de superar o atraso máximo tolerável no TTI t . Os UEs M2M que possuem valores para ΔD mais próximos de zero estão mais próximos de não satisfazer os requisitos de QoS. Desta forma, a função $M_{M2M}^{TD}delay$ foi definida, como observado em (4.12), para que os seus valores estejam contidos no intervalo fechado $[0, 1]$, onde quanto maior o seu valor, maior será a prioridade.

$$M_{M2M}^{TD}delay(u,t) = \begin{cases} 1 & \text{Se } (\max_{n \in \mathcal{U}_{M2M}} \Delta D(n)) = 0 \\ 1 - \frac{\Delta D(u)}{\max_{n \in \mathcal{U}_{M2M}} \Delta D(n)} & \text{Caso contrário} \end{cases} \quad (4.12)$$

$\Delta D(u)$ é calculada de duas formas. Se a última requisição por recursos do UE M2M u foi atendida ou nenhuma requisição foi feita, então $\Delta D(u)$ é igual ao atraso máximo tolerável. Caso contrário, $\Delta D(u)$ é igual a diferença entre o atraso máximo e o tempo esperado desde a última requisição negada ou zero caso esta diferença seja negativa. O atraso máximo tolerável pode ser definido utilizando uma das duas abordagens propostas na seção 4.2 para classificar o tráfego M2M.

A segunda função utilitária $M_{M2M}^{TD}fair$ prioriza os UEs M2M que receberam poucos recursos ao longo do tempo. A média da taxa de transferência $T_{M2M}^{avg}(u,t)$ é utilizada para avaliar esta quantidade de recursos recebidos. Para obter os valores de $M_{M2M}^{TD}fair$ no intervalo $[0, 1]$, os valores de $T_{M2M}^{avg}(u,t)$ são divididos pela maior média da taxa de transferência dentre os UEs M2M. Como pode ser observado na equação (4.13), quanto menor for o valor desta divisão, maior será a prioridade.

$$M_{M2M}^{TD}fair(u,t) = \begin{cases} 1 & \text{Se } (\max_{n \in \mathcal{U}_{M2M}} T_{M2M}^{avg}(n,t-1)) = 0 \\ 1 - \frac{T_{M2M}^{avg}(u,t-1)}{\max_{n \in \mathcal{U}_{M2M}} T_{M2M}^{avg}(n,t-1)} & \text{Caso contrário} \end{cases} \quad (4.13)$$

$T_{M2M}^{avg}(u,t)$ é calculado através de uma média móvel exponencial da taxa de transferência do UE M2M u até o TTI t como mostrado na equação (4.14), onde β é uma constante medida em TTIs e $T_{M2M}(u,t)$ é a taxa de transferência esperada para o dispositivo no TTI t cujo cálculo será mostrado na próxima subseção.

$$T_{M2M}^{avg}(u,t) = \frac{1}{\beta} T_{M2M}(u,t) + \left(1 - \frac{1}{\beta}\right) T_{M2M}^{avg}(u,t-1) \quad (4.14)$$

Como discutido anteriormente, a função prioritária $M_{M2M}^{TD}(u,t)$ desta fase é definida pela composição das funções $M_{M2M}^{TD}delay(u,t)$ e $M_{M2M}^{TD}fair(u,t)$, como mostra a equação (4.15). A constante ϖ ($0 \leq \varpi \leq 1$) indica o peso da priorização dos dois objetivos pretendidos desta fase: (i) garantir justiça nas alocações de recursos através de $M_{M2M}^{TD}fair(u,t)$; e (ii) satisfazer os requisitos de QoS através de $M_{M2M}^{TD}delay(u,t)$.

$$M_{M2M}^{TD}(u,t) = (1 - \varpi) M_{M2M}^{TD}fair(u,t) + \varpi M_{M2M}^{TD}delay(u,t) \quad (4.15)$$

Com a função de priorização definida ($M^{TD}(u,t) = M_{M2M}^{TD}(u,t)$), os passos desta fase são os seguintes:

1. Cacule a prioridade $M_{M2M}^{TD}(u,t)$ para todos os UEs M2M que requisitaram recursos no TTI t .
2. Ordene os UEs M2M de forma descendente segundo sua prioridade em uma lista L^{TD} .
3. Escolha os RB_{M2M}/RB_{M2M}^{min} primeiros UEs M2M de L^{TD} para a próxima fase e os remova de L^{TD} .
4. As requisições são negadas para os UEs M2M restantes em L^{TD} . Além disso, cada UE M2M u em L^{TD} só poderá fazer uma nova requisição após um período de tempo escolhido aleatoriamente no intervalo $[1, \sigma \Delta D(u)]$ ms. A constante $0 < \sigma \leq 1$ indica o percentual do tempo $\Delta D(u)$ que o dispositivo esperará. Portanto, esta constante é estabelecida de tal forma que após a espera, o atraso não tenha ultrapassado, provavelmente, o máximo tolerável.

O quarto passo é executado com a finalidade de espalhar as possíveis próximas requisições no decorrer do tempo sem que o atraso máximo tolerável seja ultrapassado. Assim, esta fase também tenta reduzir um possível congestionamento na rede causado pela grande quantidade de dispositivos requisitando recursos ao mesmo tempo.

4.4.2 Segunda fase - FDPS

Seja $n = RB_{M2M}/RB_{M2M}^{min}$. Logo, n grupos de recursos consecutivos em relação à frequência são criados nesta fase, um para cada dispositivo selecionado pela fase anterior. A decisão de qual grupo será alocado para qual dispositivo é feita com objetivo de aumentar a eficiência do uso dos recursos, *i.e.*, aumentar a taxa de transferência dos dispositivos. O nível de satisfação dos requisitos de QoS também é afetado nesta fase, dado que uma maior taxa de transferência, é necessário menos tempo para transmitir os dados e, por conseguinte, uma redução no atraso.

Considerando que há n grupos e n dispositivos, então, o espaço de busca total dos possíveis mapeamentos entre os grupos e os dispositivos é igual a $n!$ No pior caso onde a largura de banda é de 20 MHz com 100 RBs e com 1 RB por UE., este espaço é igual a 100!

Para reduzir este espaço de busca, uma simples abordagem gulosa é utilizada. Os passos desta abordagem são:

1. Para cada grupo de RBs g e UE M2M u , calcule a taxa de transferência ($T_{M2M}(u, g)$) caso o dispositivo utilize este grupo de recursos.
2. Salve os resultados calculados no passo anterior na forma do par $\langle u, g \rangle$ em uma lista L^{FD} ordenada de forma decendente em relação à taxa de transferência dos pares.
3. Se $L^{FD} \neq \emptyset$, então, remova desta lista o primeiro elemento $\langle u', g' \rangle$. Caso contrário, a fase é finalizada.
4. Aloque os RBs do grupo g' ao UE M2M u' somente se o dispositivo ainda não recebeu recursos.
5. Repita o passo 2.

A complexidade no pior caso para esta etapa é igual a $\Theta(n^2)$. Portanto, o uso da abordagem gulosa proposta neste trabalho é uma solução viável, visto que o escalonamento possui a restrição de tempo de 1 TTI (1 ms) para ser executada.

A taxa de transferência $T_{M2M}(u, t)$ esperada para o UE u no TTI t é igual à taxa de transferência $T_{M2M}(u, g)$ para o grupo de recursos g selecionado. O cálculo de $T_{M2M}(u, g)$ pode ser realizado utilizando o limite teórico de Shannon para a capacidade máxima do sistema (LIM

et al., 2006; PIRO; BALDO; MIOZZO, 2011), como mostrado a seguir:

$$T_{M2M}(u, g) = \frac{B}{R} \eta(u, g) \quad (4.16a)$$

$$\eta(u, g) = \log_2 \left(1 + \frac{\gamma(u, g)}{\Gamma} \right) \quad (4.16b)$$

$$\Gamma = \frac{-\ln(5 \times 10^{-5})}{1,5} \quad (4.16c)$$

$$\gamma(u, g) = \left[\frac{1}{\frac{1}{|g|} \sum_{r \in g} \frac{SINR(u, r)}{SINR(u, r) + 1}} \right]^{-1} \quad (4.16d)$$

onde, B é a largura de banda do sistema, R é quantidade de RBs disponíveis para esta largura e $\eta(u, g)$ é a eficiência espectral entre o dispositivo u e o grupo de recursos g . O coeficiente Γ modela a diferença entre o limite teórico e a performance real do esquema de modulação e codificação (*Modulation and Coding Scheme - MCS*) (PIRO; BALDO; MIOZZO, 2011). A função $\gamma(u, g)$ mede a relação sinal-ruído mais interferência (SINR) entre o UE u e o grupo g utilizando o equalizador MMSE (*Minimum Mean Squared Error*) (LIM et al., 2006). Por fim, $SINR(u, r)$ é o valor da relação SINR entre o dispositivo e o RB r .

A relação SINR é umas das métricas utilizadas no padrão do sistema LTE para medir a qualidade do canal cujo valor depende de vários fatores, como a distância do dispositivo à estação base e a potência utilizada para a transmissão. O cálculo do SINR está fora do escopo desta dissertação, mas o seu cálculo pode ser encontrado em (PIRO; BALDO; MIOZZO, 2011).

5 EXPERIMENTOS

Após definir o mecanismo proposto nesta dissertação no capítulo 4, a próxima etapa consiste na implementação deste mecanismo e de outras duas soluções existentes na literatura. Após os experimentos executados, os seus resultados obtidos são analisados. Desta forma, os experimentos e seus parâmetros são especificados neste capítulo, assim como, a discussão dos resultados obtidos é apresentada.

Na seção 5.1, os escalonamentos utilizados nos experimentos são apresentados. A seção 5.2 apresenta o modelo para os tráfegos das comunicações H2H e M2M. A descrição do plano de testes para a realização dos experimentos, do mesmo modo, as especificações dos parâmetros utilizados nas simulações são abordadas na seção 5.3. Na seção 5.5, os resultados obtidos são apresentados, analisados e discutidos.

5.1 Soluções comparadas

Para a avaliação do mecanismo proposto nesta dissertação com as duas abordagens de classes de QoS apresentadas no capítulo 4, estas foram comparadas com duas soluções existentes na literatura. A primeira solução é o escalonador *Proportional Fair* (PF), que é uma das soluções mais utilizadas e pesquisadas na literatura para o escalonamento justo de recursos (LEE et al., 2009). No PF, os tipos de comunicações (H2H e M2M) não são tratados distintamente. Para fins de comparação, o PF foi utilizado para o escalonamento da comunicação H2H no mecanismo proposto neste trabalho. Como discutido no capítulo 3, dois algoritmos são propostos em (LIOUMPAS; ALEXIOU, 2011) sendo o segundo algoritmo escolhido para os experimentos visto que o mesmo apresenta melhor desempenho na satisfação do QoS que o primeiro algoritmo.

O escalonador de pacotes PF tem os objetivos de (i) garantir justiça proporcional entre os UEs e (ii) maximizar a taxa de transferência dos UEs. Para isto, a função utilitária $M_{PF}(u, t)$ é definida (equação (5.1)) como a razão entre a taxa de transferência esperada $T(u, t)$ e a taxa de transferência média $T^{avg}(u, t)$ do UE u no TTI t . As funções $T^{avg}(u, t)$ e $T(u, t)$ são definidas de forma equivalente às mostradas nas equações (4.15) e (4.16) respectivamente. O algoritmo 1, reproduzido em pseudocódigo, apresenta uma visão geral da abordagem gulosa, baseada em (WONG; OTERI; MCCOY, 2009), para o escalonador PF utilizado nos experimentos.

$$M_{PF}(u, t) = \frac{T(u, t)}{T^{avg}(u, t)} \quad (5.1)$$

Os valores para dois parâmetros de entrada do segundo algoritmo proposto em (LIOUMPAS; ALEXIOU, 2011) precisam ser definidos para que este algoritmo seja comparado aos outros escalonadores utilizados nos experimentos. Primeiramente, os RBs disponíveis para o escalonamento dos UEs M2M precisam ser definidos. Como a comunicação H2H não é abordada em (LIOUMPAS; ALEXIOU, 2011), apesar dela ter maior prioridade em relação à M2M,

Algoritmo 1: Pseudocódigo para o algoritmo do *Provincial Fair*

Entrada: Conjunto \mathcal{U} dos UEs e \mathcal{RB} dos RBs

Saída: Mapeamento dos RBs para os UEs

```

1 Crie o conjunto  $\mathcal{G}$  com os índices dos possíveis subgrupos de RBs contínuos;
2 para cada  $u \in \mathcal{U}$  faça
3   |  $\mathcal{RB}_u \leftarrow \emptyset$ ;
4   | para cada  $g \in \mathcal{G}$  faça
5   |   | Calcule a métrica  $M_{PF}(u, g)$ , equação (5.1);
6   |   | fim
7   | fim
8 enquanto  $\mathcal{RB} \neq \emptyset$  e  $\mathcal{U} \neq \emptyset$  faça
9   | Encontre o  $u \in \mathcal{U}$  e  $g \in \mathcal{G}$  que possuam o maior valor para  $M_{PF}(u, g)$ ;
10  |  $\mathcal{RB}_u \leftarrow \mathcal{RB}_u \cup \{r\}$ ;
11  |  $\mathcal{G} \leftarrow \mathcal{G} \setminus \{g\}$ ;
    | //  $\mathcal{RB}_g$  é o conjunto dos RBs de  $g$ 
12  |  $\mathcal{RB} \leftarrow \mathcal{RB} \setminus \{\mathcal{RB}_g\}$ ;
13  | para cada  $g' \in \mathcal{G}$  faça
14  |   | se  $\mathcal{RB}_g \cap \mathcal{RB}_{g'} \neq \emptyset$  então
15  |   |   |  $\mathcal{G} \leftarrow \mathcal{G} \setminus \{g'\}$ ;
16  |   |   | fim
17  |   | fim
18  | fim
19 se  $\mathcal{RB} \neq \emptyset$  então
20  | Distribua aleatoriamente os RBs restantes para os UEs respeitando a
    | restrição de contiguidade;
21 fim

```

então a divisão de recursos entre H2H e M2M descrita na seção 4.3 foi utilizada para valorar o parâmetro dos recursos disponíveis para o escalonamento no segundo algoritmo apresentado em (LIOUMPAS; ALEXIOU, 2011). Finalmente, a quantidade de recursos requerida por cada dispositivo M2M também precisa ser determinada. Para isto, os recursos foram divididos igualmente entre os dispositivos, mas com um limite inferior igual à RB_{M2M}^{min} . Este limite é definido, seção 5.3, com base na característica de pouca transmissão de dados da comunicação M2M, logo cada UE M2M poderá requerer uma pequena quantidade de recursos. O algoritmo 2 apresenta em pseudocódigo o segundo algoritmo proposto em (LIOUMPAS; ALEXIOU, 2011), onde as linhas 1 à 4 referem-se a especificação dos valores para os parâmetros de entrada discutidos anteriormente. Ademais, o escalonador PF também é utilizado para o escalonamento dos dispositivos H2H.

5.2 Modelagem do tráfego

As aplicações H2H utilizadas nas simulações foram VoIP, vídeo e FTP. As modelagens dos tráfegos destas aplicações foram baseadas nos trabalhos (SALAH et al., 2011) (VoIP

Algoritmo 2: Pseudocódigo do algoritmo proposto em (LIOUMPAS; ALEXIOU, 2011)

Entrada: Conjunto \mathcal{U}_{H2H} dos UEs H2H, \mathcal{U}_{M2M} dos UEs M2M e $\mathcal{R}\mathcal{B}$ dos RBs

Saída: Mapeamento dos RBs para os UEs

```

1 Defina  $\mathcal{R}\mathcal{B}_{H2H}$  para os RBs da H2H como especificado na seção 4.3 ;
2 Defina  $\mathcal{R}\mathcal{B}_{M2M}$  para os RBs da M2M como especificado na seção 4.3;
3 Execute o Algoritmo 1 para as entradas  $\mathcal{U}_{H2H}$ ,  $\mathcal{R}\mathcal{B}_{H2H}$ ;
4  $maxRB \leftarrow \min \left( \left\lfloor \frac{|\mathcal{R}\mathcal{B}_{M2M}|}{|\mathcal{U}_{M2M}|} \right\rfloor, RB_{M2M}^{min} \right)$ ;
5 para cada  $u \in \mathcal{U}_{M2M}$  faça
6   |  $\mathcal{R}\mathcal{B}_u \leftarrow \emptyset$ ;
7 fim
8 enquanto  $\mathcal{R}\mathcal{B}_{M2M} \neq \emptyset$  e  $\mathcal{U}_{M2M} \neq \emptyset$  faça
9   | Encontre  $u \in \mathcal{U}_{M2M}$  que possua o menor atraso máximo tolerável  $d_u$ ;
10  | Encontre  $r \in \mathcal{R}\mathcal{B}_{M2M}$  com o menor valor SINR para o UE  $u$  e adjacente
    | à algum RB de  $\mathcal{R}\mathcal{B}_u$ ;
11  |  $\mathcal{R}\mathcal{B}_u \leftarrow \mathcal{R}\mathcal{B}_u \cup \{r\}$ ;
12  |  $\mathcal{R}\mathcal{B}_{M2M} \leftarrow \mathcal{R}\mathcal{B}_{M2M} \setminus \{r\}$ ;
13  | se  $|\mathcal{R}\mathcal{B}_u| \geq maxRB$  então
14  |   |  $\mathcal{U}_{M2M} \leftarrow \mathcal{U}_{M2M} \setminus \{u\}$ ;
15  |   fim
16 fim
17 se  $\mathcal{R}\mathcal{B}_{M2M} \neq \emptyset$  então
18  | Distribua aleatoriamente os RBs restantes para os UEs M2M respeitando
    | a restrição de contiguidade;
19 fim

```

e FTP) e (POTSCH et al., 2013) (Vídeo).

As categorias baseadas em evento e em tempo foram utilizadas para simular a comunicação M2M. A transmissão em rajadas das aplicações M2M baseadas em eventos foram modeladas através do processo de Poisson com a taxa média igual à 0,02 pacotes por TTI (GOTSIS; LIOUMPAS; ALEXIOU, 2013). O intervalo de transmissão de cada dispositivo das aplicações baseadas em tempo foi distribuído uniformemente entre 50 ms e 550 ms. Ambas as categorias de aplicações M2M possuem o tamanho do pacote de 125 bytes (3GPP, 2012a).

A Tabela 5.1 resume os parâmetros da modelagem dos tráfegos das aplicações H2H e M2M.

5.3 Descrição do plano de testes

Os escalonadores discutidos na seção 5.1 foram implementados, bem como os experimentos foram realizados, no simulador de redes NS-3 (NS-3, 2013) com suporte à rede LTE através do projeto LENA (BALDO et al., 2011).

Tabela 5.1: Parâmetros dos modelos para os tráfegos H2H e M2M

| | Tamanho do pacote | Tempo entre transmissões | QCI |
|--------------------|-------------------|--|-----|
| Tráfego H2H | | | |
| Vídeo | 1200 bytes | 75 ms | 2 |
| VoIP | 40 bytes | 20 ms | 1 |
| FTP | 256 bytes | 16,625 ms | 8 |
| Tráfego M2M | | | |
| Baseado em Evento | 125 bytes | Poisson, média de 50 ms (0,02 pacotes/TTI) | |
| Baseado em Tempo | 125 bytes | Distribuição uniforme [50, 550] ms | |

Para analisar o impacto da comunicação M2M na H2H três casos foram estabelecidos para esta avaliação. Em cada caso, os outros objetivos (garantir justiça e satisfazer os requisitos de QoS) do mecanismo proposto nesta dissertação também foram avaliados. Estes três casos refletem diferentes cargas dos tráfegos das aplicações H2H, que são:

1. **Baixo tráfego.** No primeiro caso, um baixo tráfego é gerado ao utilizar 6 UEs H2H divididos igualmente entre os três tipos de tráfego H2H simulados (Vídeo, VoIP, FTP). Para simplificação, em todos os casos, os dispositivos geram somente um tipo de tráfego.
2. **Médio tráfego.** No segundo caso, o tráfego H2H é reduzido para valores intermediários em relação aos outros casos. Assim como no primeiro caso, os dispositivos são divididos igualmente para os diferentes tipos de tráfego H2H simulados, porém este caso emprega 15 UEs H2H.
3. **Alto tráfego.** No último caso, um alto tráfego H2H é gerado ao usar 30 UEs H2H, dez para cada tipo de tráfego H2H.

Nas simulações dos três casos, a quantidade de dispositivos M2M foi variada e dividida entre as classes baseada em evento e em tempo de tal forma que 1/3 dos UEs foram para a primeira classe e 2/3 para a segunda. Esta relação foi escolhida visto que a classe baseada em tempo contém grande parte das aplicações M2M como discutido no capítulo 2. Todos os UEs foram distribuídos uniformemente em torno de uma única eNB. A largura de banda utilizada foi de 5 MHz, 25 RBs disponíveis por TTI (1 ms), com a finalidade de simular um ambiente com escassez de recursos mesmo utilizando poucos dispositivos. O tempo de simulação (3 s) foi definido para conter uma quantidade razoável de escalonamentos (3000). Para obter resultados confiáveis como discutido em (JAIN, 1991), 30 simulações, gerando amostras aleatórias e independentes, foram executadas para cada caso e foi utilizado o intervalo de confiança de 95% para analisar os resultados obtidos.

Com relação aos parâmetros do mecanismo proposto, a Tabela 4.2 foi utilizada para classificar o tráfego M2M segundo as duas abordagens já discutidas na seção 4.2. Além disso, na fase inicial da implementação dos experimentos, foram realizadas simulações com o objetivo de enriquecer a solução proposta nesta dissertação e definir os seus parâmetros. Deste modo, os valores das constantes das equações (4.6), (4.7) e (4.14) foram estabelecidos de tal forma que as decisões mais antigas ainda influenciem nos cálculos dessas funções. As quantidades mínimas

Tabela 5.2: Parâmetros das simulações

| Parâmetros Gerais | |
|----------------------------------|---|
| Tempo de simulação | 3s (3000 TTI) |
| Amostras | 30 por caso |
| Distribuição dos UEs | fixo e uniforme no radio 1400 |
| Número de eNB | 1 |
| Largura de banda | 5 MHz (25 RBs) |
| Número de UEs M2M | 0, 50, 100, 150, 200, 250; 1/3 Baseado em Evento e 2/3 Baseado em Tempo |
| Número de UEs H2H | 30, 15, 6; 1/3 Vídeo, 1/3 VoIP, 1/3 FTP |
| Parâmetros do Mecanismo Proposto | |
| α, β | 99 |
| $RB_{H2H}^{min}, RB_{M2M}^{min}$ | 3 |
| P_{M2M}^{min} | 0,48 |
| σ | 0,1 |
| ϖ | 0,72 |

de recursos (RB_{H2H}^{min} e RB_{M2M}^{min}) requisitadas por UE M2M foram escolhidas para que seja possível transmitir todos os dados no *buffer* com apenas esta quantidade caso o UE M2M apresente uma boa qualidade no canal e a quantidade de dados seja pequena como nas aplicações M2M. Além disto, a porcentagem mínima de recursos para os dispositivos M2M foi definida para que haja pelo menos 4 ($25 \times P_{m2m}^{min} / RB_{M2M}^{min}$) UEs M2M recebendo recursos por TTI. Na equação (4.15), o valor da constante ϖ foi definido para que a satisfação dos requisitos de QoS tenha maior peso do que a justiça na priorização das requisições. O tempo de espera após a requisição ser negada foi estabelecida como no máximo 10% ($\sigma = 0,1$) do valor de ΔD .

Por fim, a Tabela 5.2 resume os parâmetros das simulações. Parâmetros não incluídos nesta tabela possuem os valores padrões do simulador NS-3 (NS-3, 2013) na sua versão 3.18.

5.4 Métricas de avaliação

Para avaliar o desempenho dos objetivos do mecanismo para o escalonamento de pacotes proposto nesta dissertação, as seguintes métricas foram utilizadas:

- **Taxa de transferência.** Esta métrica foi utilizada para avaliar o impacto da comunicação M2M sobre a H2H. A taxa de transferência do dispositivo foi calculada como a quantidade de RBs em bits alocados para o dispositivo durante a simulação dividido pelo tempo de simulação. A transformação da quantidade de RBs alocados em bits é feita pelo simulador NS-3 ao mapear a qualidade do canal (a eficiência espectral da equação (4.16b)) e o esquema de modulação e codificação (MCS) para o tamanho do bloco de transmissão em bits como definido pelos trabalhos (PIRO; BALDO; MIOZZO, 2011) e (ZHOU; BALDO; MIOZZO, 2013).
- **Nível de insatisfação dos requisitos de QoS.** Esta métrica é calculada como a razão entre

a quantidade de pacotes perdidos mais a quantidade de pacotes cujo o atraso ultrapassou o máximo tolerável associado ao seu tráfego pela quantidade total de pacotes gerados pelo dispositivo. O atraso é calculado pelo intervalo de tempo entre o pacote chegar no *buffer* de transmissão até ele sair da rede LTE pelo *gateway* de saída. Desta forma, quanto maior o valor da razão menor será o nível de satisfação. Além de avaliar o nível de insatisfação dos requisitos de QoS, esta métrica também foi utilizada para avaliar o impacto da comunicação M2M na H2H.

- **Índice de justiça de Jain.** O índice de justiça proposto por Jain (1991) avalia o quão igualitária é a distribuição dos recursos. A equação (5.2) apresenta o cálculo deste índice, onde $T(u)$ é a taxa de transferência do UE u descrita acima. O valores dessa métrica podem ser compreendidos da seguinte maneira. Para exemplificar o que o valor deste índice indica, considere que $I = k/n$ e que exista no total n dispositivos. Isto significa que k dispositivos receberam recursos igualmente e os outros $n - k$ dispositivos não receberam recursos durante o intervalo de tempo analisado. Portanto, quanto mais próximo o valor do índice estiver de 1, mais justo o resultado será. De modo consequente, índices mais próximos de zero indicam que mais dispositivos sofreram de inanição.

$$I = \frac{(\sum_{u \in \mathcal{U}} T(u))^2}{|\mathcal{U}| \times \sum_{u \in \mathcal{U}} T(u)^2} \quad (5.2)$$

Para cada amostra, a média entre todos os dispositivos, do mesmo tipo de comunicação, foi feita para as duas primeiras métricas citadas acima. O índice de justiça é calculado diretamente da equação (5.2). Deste modo, ao analisar os resultados obtidos por cada caso é feita a média dos valores das amostras com o intervalo de confiança com nível de confiança igual à 95%.

5.5 Análise dos resultados obtidos

Nesta seção, os resultados obtidos após a implementação dos três casos de teste, definidos na seção 5.3, são apresentados em formas de gráficos. Também os resultados são analisados para verificar o impacto da comunicação M2M sobre a H2H (subseção 5.5.1), a taxa de transferência (subseção 5.5.2), o nível de satisfação dos requisitos de QoS (subseção 5.5.3) e, por fim, o índice de justiça (subseção 5.5.4).

Como discutido na seção 5.1, duas abordagens foram propostas neste trabalho para classificar o tráfego M2M segundo seus requisitos de QoS. Ao longo desta seção, o uso da primeira abordagem pelo escalonador proposto será referido como “Proposta V1” e ao utilizar a segunda abordagem, o escalonador proposto será referido como “Proposta V2”. As modificações também descritas na seção 5.1 para o segundo escalonador proposto em (LIOUMPAS; ALEXIOU, 2011) serão designadas de “Lioumpas 2”. Por fim, O escalonador *Proporcional Fair* será identificado como PF.

5.5.1 Impacto da comunicação M2M sobre a H2H

Conforme apresentado na seção 5.4, o impacto da comunicação M2M sobre a H2H será avaliado pelas métricas taxa de transferência e nível de satisfação dos requisitos de QoS. Ressaltando que como os escalonadores Proposta V1, Proposta V2 e Lioumpas 2 utilizam a mesma sub-rotina para separar os recursos entre as comunicações H2H e M2M, então somente os escalonadores Proposta V1 e PF são visualizados nos gráficos desta subseção.

A Figura 5.1 apresenta os valores médios da taxa de transferência dos UEs H2H em relação à variação no número de UEs M2M, onde as sub-figuras 5.1(a), 5.1(b) e 5.1(c) são referentes aos casos com baixo, médio e alto tráfego H2H do plano de testes respectivamente. A quantidade de dispositivos mostrada no eixo x da Figura 5.1, assim como nas figuras seguintes desta seção, é referente à quantidade total de dispositivos (UEs H2H + UEs M2M).

Observa-se que, para todos os casos, o escalonador PF apresenta uma drástica queda na taxa de transferência com o aumento do número de UEs M2M. O escalonador Proposta V1 também apresenta quedas nas suas taxas, contudo, a sua diminuição é menos acentuada. Este melhor desempenho apresentado pelo Proposta V1 em relação ao PF é decorrente da limitação da quantidade de UEs M2M escalonados em um TTI com presença de tráfego H2H como discorrido na seção 4.3.

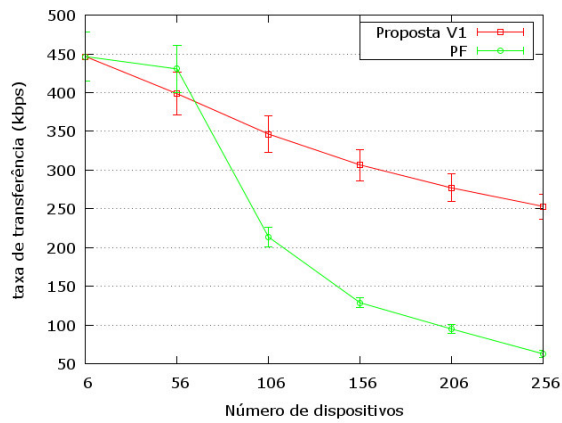
A Figura 5.2 apresenta os valores médios dos percentuais de pacotes do tráfego H2H que ultrapassam, durante a transmissão, o atraso máximo tolerável para os três casos do plano de testes. Assim como na Figura 5.1 sobre a taxa de transferência, as sub-figuras 5.2(a), 5.2(b) e 5.2(c) são referentes aos casos com baixo, médio e alto tráfego H2H do plano de testes respectivamente.

Em todos os casos ilustrados na Figura 5.2, o nível de insatisfação dos requisitos de QoS para o escalonador PF sofre um grande aumento com o acréscimo dos UEs M2M. O PF apresenta uma aumento nas suas medidas de mais 80% em todos os três casos. O escalonador Proposta V1 satisfaz totalmente os requisitos de QoS quando o tráfego H2H é baixo (6 UEs H2H) ou moderado (15 UEs H2H) e apresenta, no final da simulação, uma aumento em torno de 11% do nível de não satisfação dos requisitos de QoS quando 30 UEs H2H são empregados. Estes resultados estão relacionados diretamente com o impacto na taxa de transferência supracitados, *i.e.*, com a diminuição da taxa de transferência, há um aumento de tempo na entrega dos pacotes (atraso).

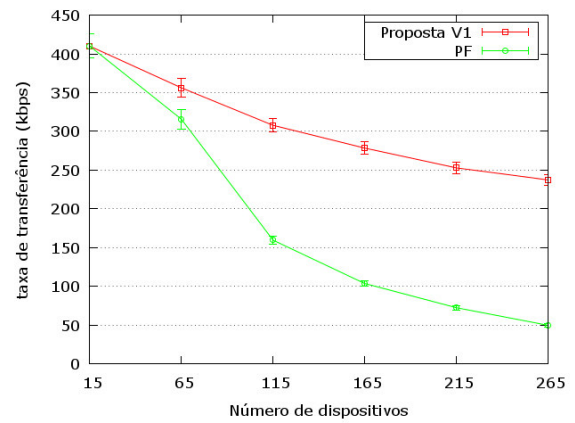
Conclui-se ao analisar os resultados das figuras 5.1 e 5.2 que a abordagem utilizada no escalonamento Proposta V1, a separação dos recursos para as comunicações H2H e M2M, apresenta melhor controle do impacto da comunicação M2M sobre a H2H do que o escalonamento PF que não distingue estes tipos de comunicações.

5.5.2 Taxa de transferência

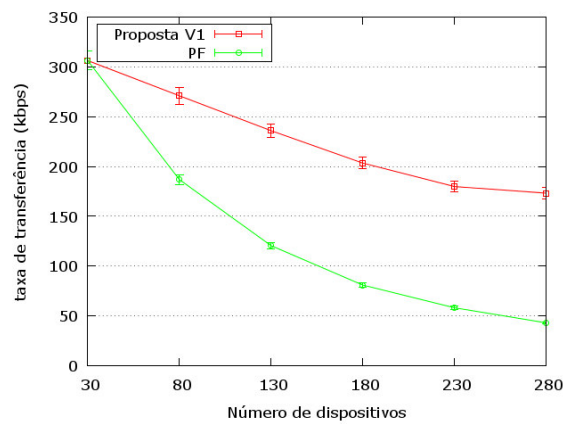
Nesta subseção, os resultados obtidos da taxa de transferência dos dispositivos M2M são analisados. A Figura 5.3 apresenta os valores médios da taxa de transferência dos



(a) Tráfego H2H com 6 UEs H2H



(b) Tráfego H2H com 15 UEs H2H



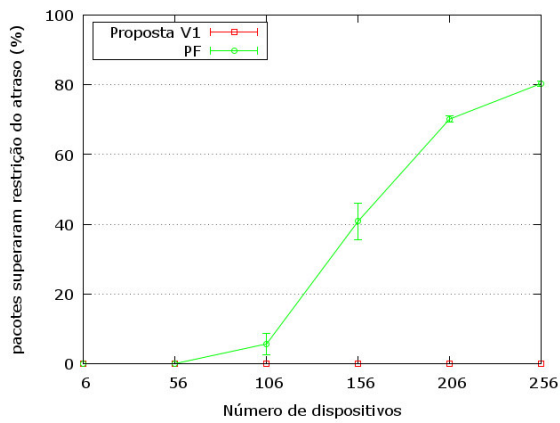
(c) Tráfego H2H com 30 UEs H2H

Figura 5.1: Impacto da comunicação M2M sobre a taxa de transferência da comunicação H2H

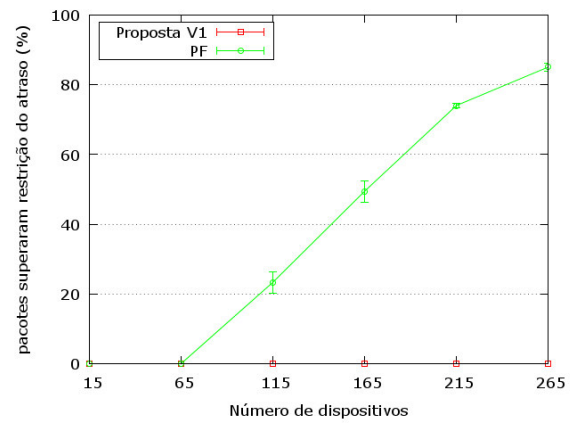
tráfegos M2M das aplicações baseadas em evento, onde as sub-figuras 5.3(a), 5.3(b) e 5.3(c) são referentes aos casos de testes com 6, 15 e 30 UEs H2H respectivamente. Similarmente, a Figura 5.4 apresenta estes valores para as aplicações M2M baseadas em tempo.

Verifica-se nas Figuras 5.3 e 5.4 que os escalonamentos Proposta V1 e Proposta V2 possuem valores com poucas variações em todos os casos. Somente no caso com alto tráfego H2H (30 UEs H2H) que estes escalonamentos apresentam uma redução de cerca de 16% da taxa de transferência no final da simulação para as aplicações baseadas em evento. Esta queda está associada com a garantia de justiça que faz com que os UEs M2M das aplicações baseadas em tempo recebam recursos mesmo tendo, geralmente, menores prioridades na satisfação dos requisitos de QoS por serem mais toleráveis ao atraso. Estas prioridades mais baixas são decorrentes dos parâmetros de simulações utilizados como discutidos na seção 5.3. Assim sendo, ao utilizar 280 dispositivos dos quais são 30 UEs H2H, esta garantia de justiça começa a afetar a taxa de transferência do tráfego das aplicações baseadas em tempo. Além disso, para todas as situações, a taxa de transferência do tráfego M2M apresenta valores baixos devido à pouca quantidade de recursos que são alocados por UE M2M.

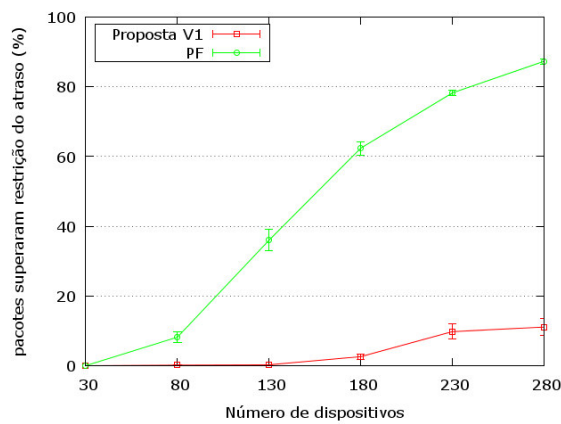
Para o escalonador Lioumpas 2, é observado nos casos onde o tráfego H2H não é alto (6 e 15 UEs H2H) que os valores das taxas de transferência diminuem até obterem valores



(a) Tráfego H2H com 6 UEs H2H



(b) Tráfego H2H com 15 UEs H2H



(c) Tráfego H2H com 30 UEs H2H

Figura 5.2: Impacto da comunicação M2M sobre a satisfação de QoS da comunicação H2H

próximos dos escalonadores Proposta V1 e V2. Isto é explicado pelo fato de que para calcular a demanda de recursos dos UEs M2M, os recursos são divididos igualmente entre os dispositivos com limite inferior igual à quantidade de recursos alocados por UEs M2M nos escalonadores Proposta V1 e V2. Para o caso do alto tráfego H2H, ocorre a situação contrária aos escalonadores Proposta V1 e V2, pois há uma redução de aproximadamente 35% da taxa de transferência do tráfego das aplicações baseadas no tempo. Isto ocorre, pois, ao contrário dos escalonadores Proposta V1 e V2, o Lioumpas 2 não possui garantias de justiça para os dispositivos com menores prioridades.

O escalonador PF apresentou redução em todas as situações analisadas. Esta redução é causada pela tentativa de alocação justa e igualitária entre os UEs, desta forma, a quantidade de RBs por UE é diminuída com o aumento do número de dispositivos. Apesar de a taxa ser reduzida, os valores da taxa de transferência são, na maioria das situações, maiores do que os valores apresentados pelos outros escalonadores. Isto se deve à não limitação dos recursos para a comunicação M2M.

Com essas observações, pode-se concluir que ao controlar o impacto da comunicação M2M sobre a H2H através da restrição da quantidade de recursos para a comunicação M2M, a taxa de transferência do tráfego M2M é afetada. Contudo, esta baixa taxa é aceitável,

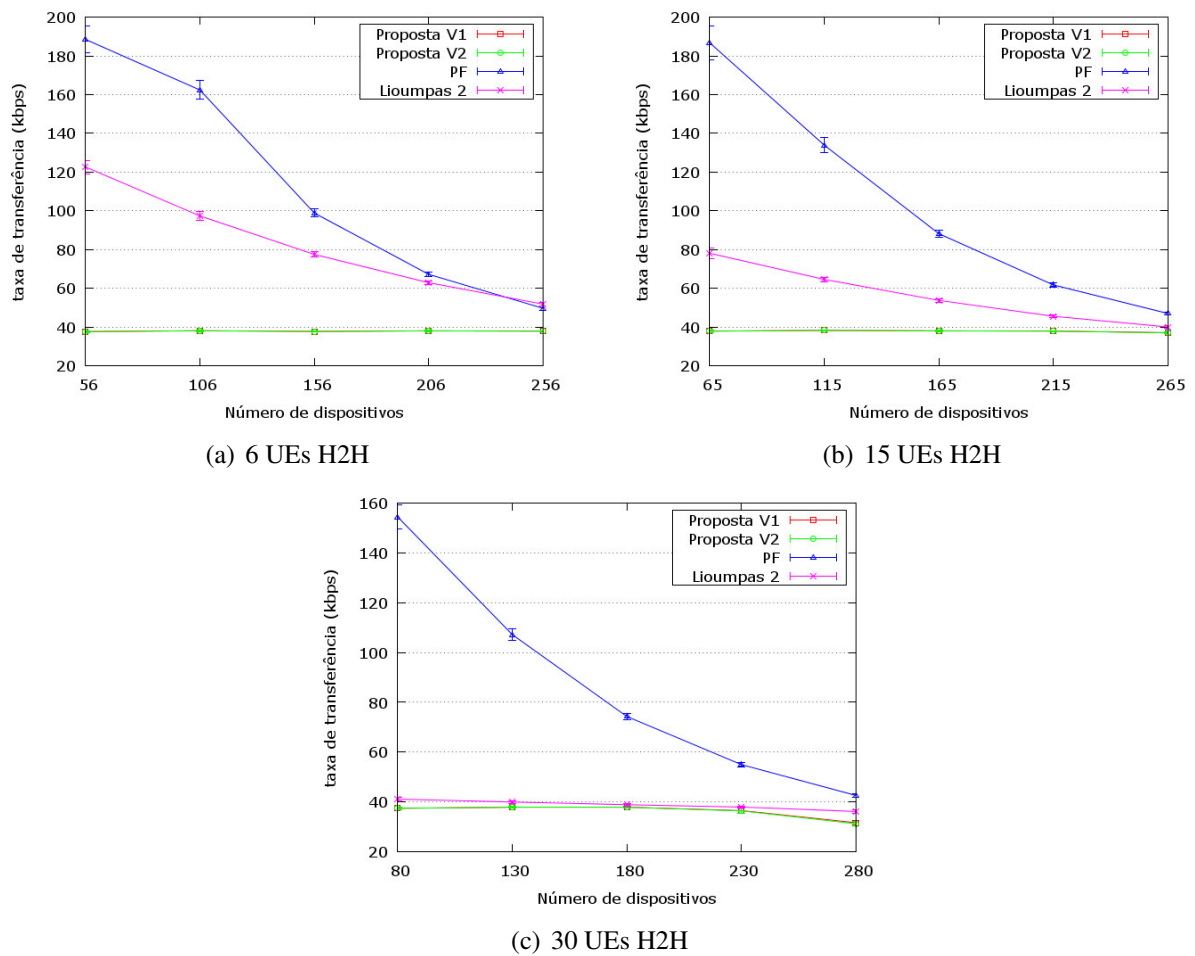


Figura 5.3: Taxa de transferência do tráfego das aplicações M2M baseadas em evento

pois poucos dados são transmitidos por cada dispositivo M2M.

5.5.3 Nível de satisfação dos requisitos de QoS

A Figura 5.5 apresenta os valores médios dos percentuais de pacotes do tráfego das aplicações baseadas em evento que ultrapassaram o atraso máximo tolerável durante a sua transmissão. As sub-figuras 5.5(a), 5.5(b) e 5.5(c) são referentes aos resultados desta métrica nos casos com baixo, médio e alto tráfego H2H respectivamente. De modo similar, a Figura 5.6 apresenta os resultados obtidos para o tráfego das aplicações baseadas em tempo.

Os escalonadores Proposta V1 e V2 apresentam, na Figura 5.5, resultados próximos à zero para os casos com baixo e médio tráfego H2H. Ao utilizar 15 UEs H2H (Figura 5.5(b)), há um aumento de 2% da insatisfação dos requisitos de QoS com 265 dispositivos. Contudo, no caso com 30 UEs H2H, uma grande alta é constatada, chegando a um aumento de quase 40% com 280 dispositivos. Este aumento está associado, como discutido na subseção 5.5.2, com a garantia de que mesmo os UEs das aplicação baseadas em tempo com menores prioridades, em relação a satisfação de QoS, recebam recursos.

O escalonador Lioumpas 2 apresenta, em todos os casos, valores dos percentuais

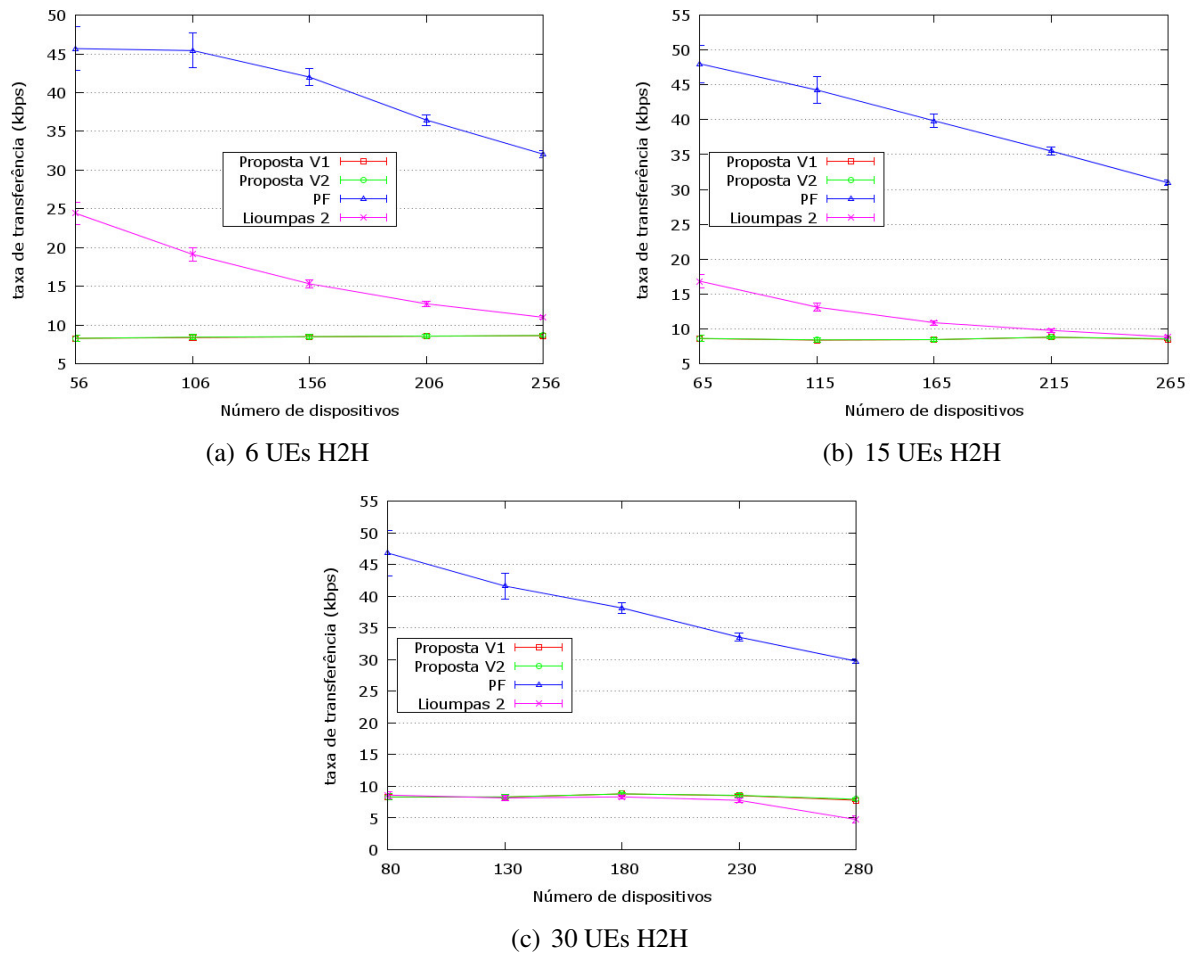


Figura 5.4: Taxa de transferência do tráfego das aplicações M2M baseadas em tempo

de pacotes que não satisfizeram os requisitos de QoS inferiores a 5% para o tráfego das aplicações baseadas em evento. Este escalonador apresenta estes baixos valores, pois as aplicações baseadas em evento possuem maior prioridade por serem poucos tolerantes ao atraso.

Para o tráfego das aplicações baseadas em tempo, o Lioumpas 2 apresenta crescimentos mais acentuados com o aumento do tráfego H2H. Esta situação ocorre pela combinação de dois motivos. O primeiro é que com o aumento do tráfego H2H que possui maior prioridade na alocação de recursos, menos recursos são disponíveis para os UEs M2M. O segundo motivo é que as aplicações baseadas em tempo possuem menores prioridades no escalonador por serem mais tolerantes a atraso. Para os escalonadores Proposta V1 e V2, o crescimento é menor que o do Lioumpas 2 porque o segundo motivo não é completamente válido para esses dois primeiros escalonadores. Isto se deve ao fato de que a garantia de justiça é também levada em consideração na priorização, disponibilizando assim, recursos para os tráfegos das aplicações baseadas em tempo.

Observa-se que em todos os casos do plano de teste, o escalonador PF apresenta um grande aumento da insatisfação do requisitos de QoS tanto para as aplicações baseadas em evento como para as baseadas em tempo. Este aumento na insatisfação é causado pela diminuição da taxa de transferência que foi discutida anteriormente na subseção 5.5.2.

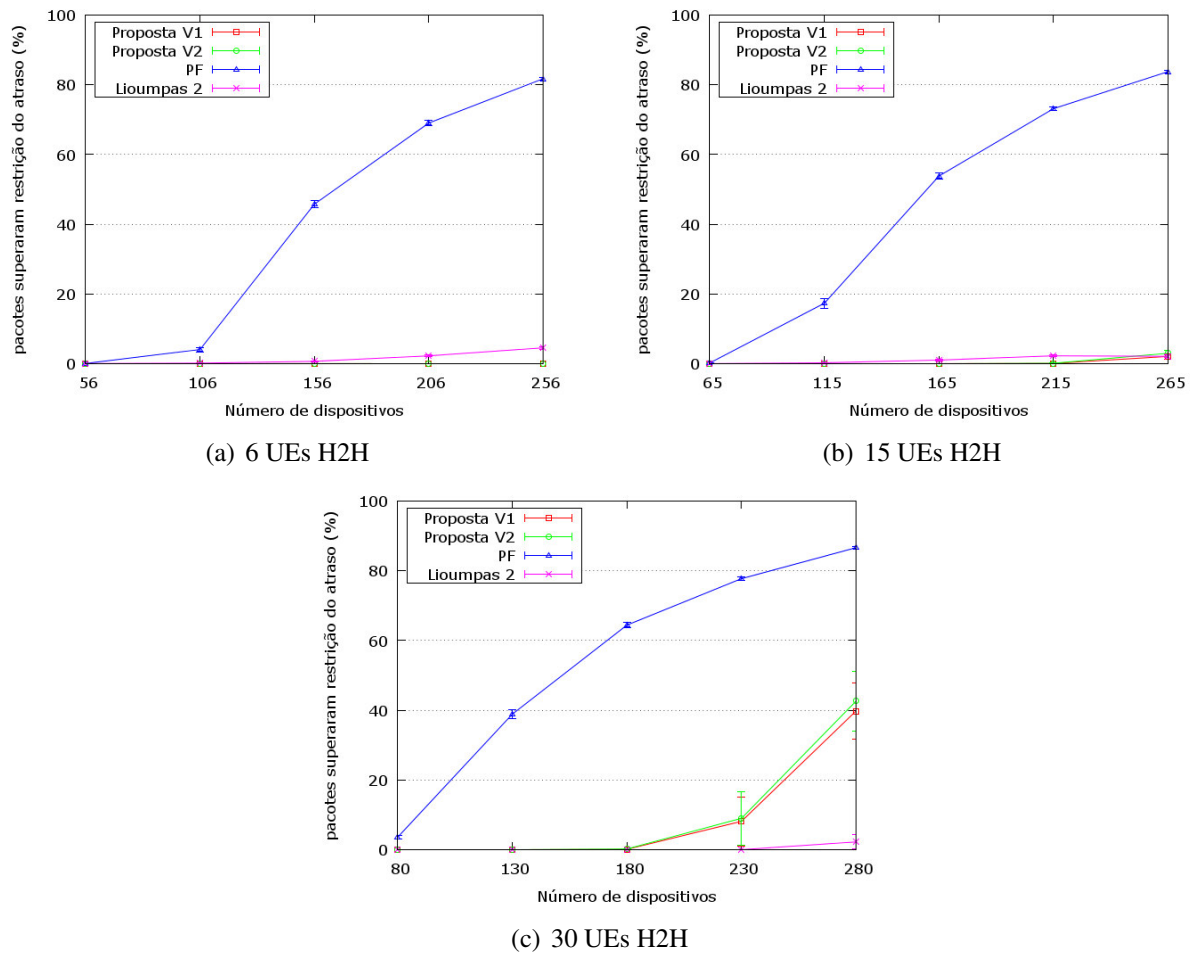


Figura 5.5: Nível de insatisfação dos requisitos de QoS das aplicações M2M baseadas em evento

Conclui-se desta análise que os escalonadores Proposta V1 e V2 apresentam resultados iguais ou inferiores que os do Lioumpas 2 para a porcentagem dos pacotes que não satisfizeram os requisitos de QoS quando o tráfego H2H não é alto. No entanto, estes dois escalonadores apresentam resultados melhores que o do Lioumpas 2 sobre o nível de satisfação dos requisitos de QoS dos tráfegos das aplicações baseadas em tempo. Contudo, ao garantir justiça na alocação dos recursos pelos escalonadores Proposta V1 e V2, a satisfação dos requisitos de QoS das aplicações baseadas em evento é afetada na situação de alto tráfego H2H.

5.5.4 Justiça na alocação dos recursos

Para avaliação da justiça na alocação dos recursos entre os tráfegos, o índice de justiça intra-classe é utilizado. Este índice é calculado pela equação (5.2) ao considerar apenas os tráfegos da mesma classe de aplicação M2M. Desta forma, é verificada a alocação igualitária dos recursos entre os tráfegos de uma mesma classe.

As Figuras 5.7 e 5.8 apresentam os valores médios do índice de justiça para o tráfego das aplicações baseadas em evento e em tempo respectivamente. Cada uma dessas figuras é composta por três sub-figuras que são relacionadas à quantidade do tráfego H2H presente nas

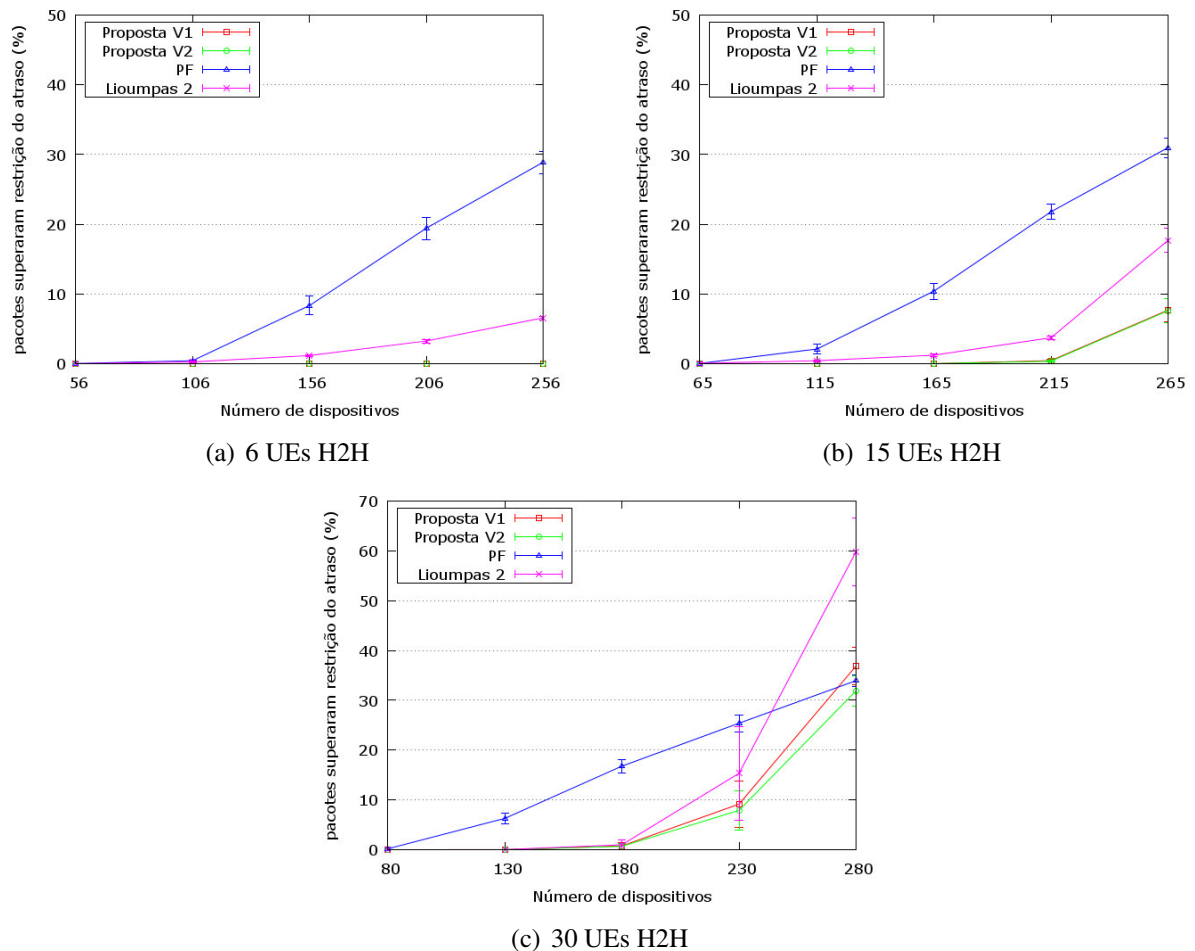


Figura 5.6: Nível de insatisfação dos requisitos de QoS das aplicações M2M baseadas em tempo simulações.

O escalonador PF possui poucas variações dos índices para as aplicações baseadas em evento com o aumento da quantidade de dispositivos M2M. Além disso, para as aplicações baseadas em tempo, há um aumento do índice nos três casos de testes. Este crescimento está relacionado com o aumento da quantidade de tráfegos com maiores prioridades do escalonador. Os tráfegos das aplicações baseadas em tempo têm maior prioridade no PF por causa da característica de transmissão infrequente dessas aplicações, conseqüentemente, estes tráfegos possuem baixas taxas de transferências.

Há também aumento do índice de justiça das aplicações baseadas em evento para o escalonador Lioumpas 2 nos casos com 6 e 15 UEs H2H. Assim como no caso do PF, o aumento do índice está relacionado com o aumento dos tráfegos com maiores prioridades, mas tem, aqui neste caso, com o aumento dos UEs M2M das aplicações baseadas em evento. Porém, no caso com alto tráfego H2H, a combinação de escassez de recursos com a grande demanda por estes recursos começa a afetar os índices de justiça dos tráfegos prioritários.

Percebe-se para o escalonador Lioumpas 2 e para as aplicações baseadas em tempo, pouca variação nos índices de justiça para os casos com pouco ou médio tráfego H2H. Contudo, com 30 UEs H2H, é verificada uma redução de mais de 50% do valor índice, chegando a um

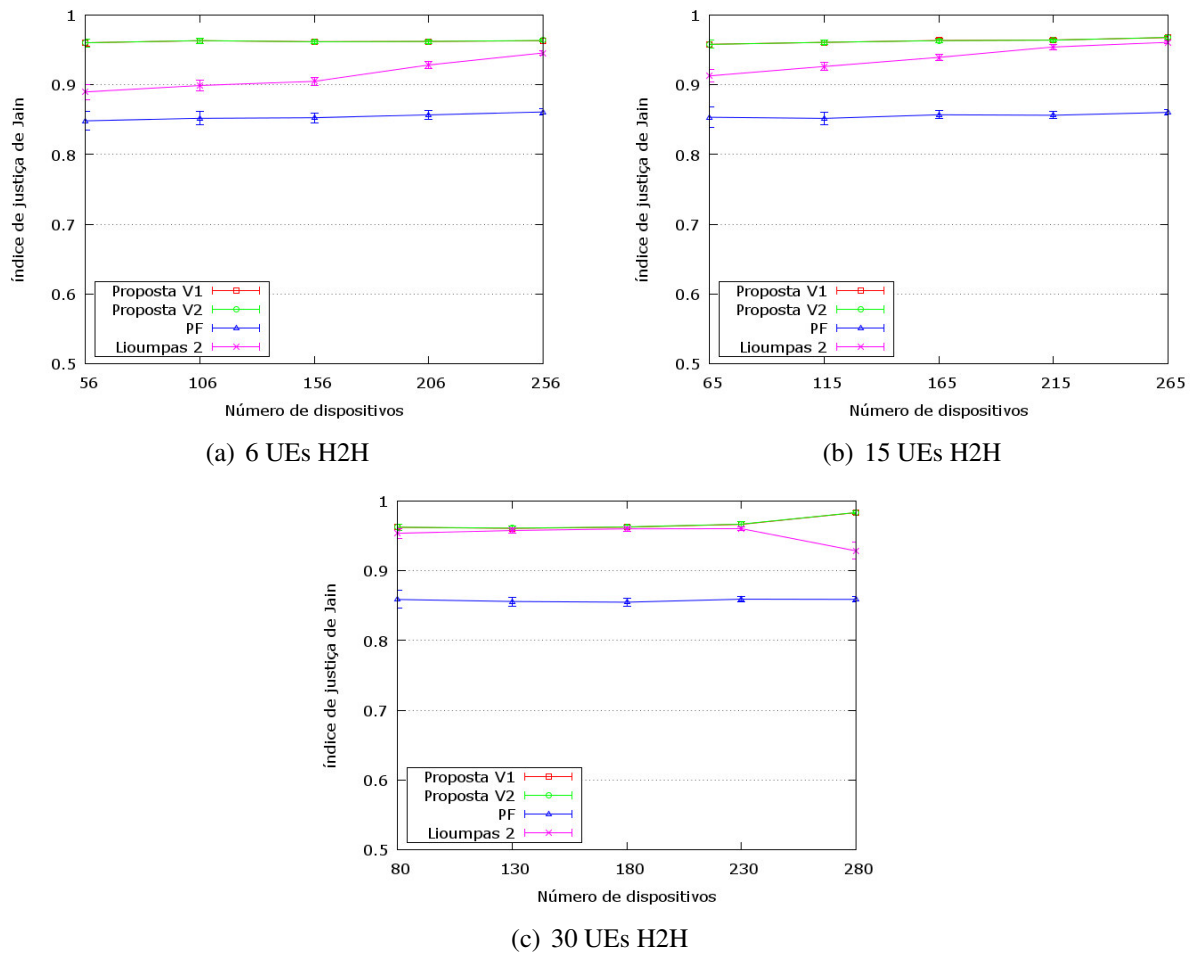
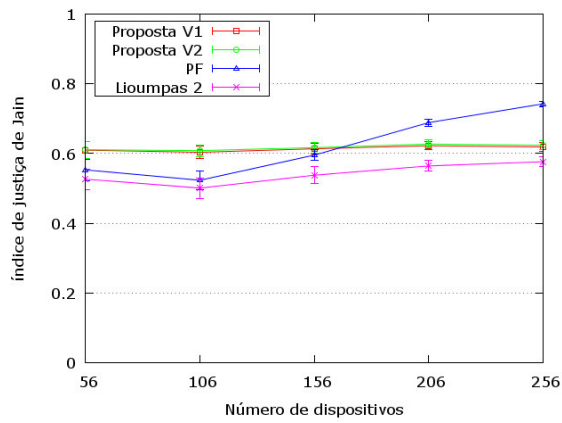


Figura 5.7: Índice de justiça para o tráfego das aplicações M2M baseadas em evento

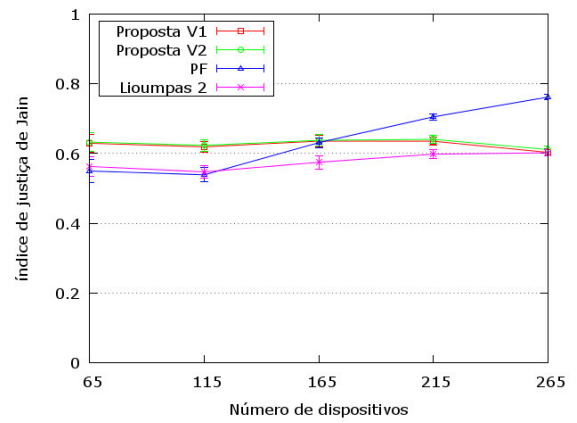
baixo índice de menos de 0,3 quando 280 dispositivos são simulados. Este baixo valor do índice é causado pela baixa prioridade desse tipo de aplicação. Portanto, há uma grande possibilidade de o problema de inanição ocorrer nesta situação.

Os escalonadores Proposta V1 e V2 apresentam poucas variações nos índices de justiça na maioria das situações para ambas as classes de aplicações M2M (baseado em evento e em tempo). A exceção é para o caso com alto tráfego H2H (30 UEs H2H), onde há um aumento do índice para as aplicações baseadas em evento e uma diminuição deste índice para as baseadas em tempo. Isto é explicado pelo fato de que em uma situação de escassez de recursos, não há recursos suficientes para compensar a maior prioridade das aplicações baseadas em evento na alocação justa dos recursos entre todos os UEs M2M.

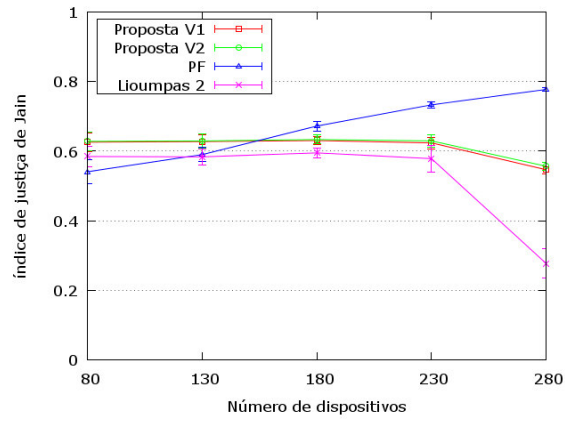
Após essas análises, conclui-se que o escalonador Proposta V1 e Proposta V2 apresentam os melhores índices de justiça para as aplicações baseadas em evento e bons índices para as aplicações baseadas em tempo. Além disto, verifica-se que o escalonador Lioumpas 2 poderá apresentar o problema de inanição para as aplicações baseadas em tempo em situações com escassez de recursos e com grande demanda por estes recursos.



(a) 6 UEs H2H



(b) 15 UEs H2H



(c) 30 UEs H2H

Figura 5.8: Índice de justiça para o tráfego das aplicações M2M baseadas em tempo

6 CONCLUSÃO E TRABALHOS FUTUROS

Esta dissertação propôs um mecanismo para o escalonador de pacotes no *uplink* da rede LTE para tratar a comunicação M2M utilizando informações do histórico de alocações de recursos, qualidade do canal e os requisitos de QoS dos dispositivos para (i) controlar o impacto da comunicação M2M sobre a H2H, (ii) evitar o problema de inanição com a alocação justa dos recursos e (iii) satisfazer as garantias de QoS. Desta forma, a seção 6.1 expõe as conclusões inferidas a partir dos resultados obtidos e discutidos no capítulo 5. Na seção 6.2 são apresentadas as principais contribuições deste trabalho. Finalmente, na seção 6.3, são discutidos os possíveis trabalhos futuros oriundos desta dissertação.

6.1 Conclusões

A partir das análises realizadas dos resultados obtidos pelas simulações, foi observado que a abordagem da separação de recursos entre a comunicação H2H e a M2M, de tal forma que a quantidade de recursos alocados para a M2M seja controlada, é uma alternativa viável para controlar o impacto da comunicação M2M sobre a H2H. Além disso, o escalonador proposto evita o problema de inanição ao garantir alocação justa dos recursos. Contudo, em um ambiente com uma grande quantidade de tráfego H2H e para cumprir os objetivos (i) e (ii) supracitados, a satisfação dos requisitos de QoS das aplicações baseadas em evento é comprometida.

Ademais, os resultados mostraram que a abordagem de estender as classes de QoS do padrão do LTE para adicionar mais n classes para as aplicações M2M baseadas em tempo possui resultados similares aos da abordagem de adicionar mais uma mensagem de controle no padrão do LTE para o envio dos valores dos parâmetros de QoS dos UEs. Por conseguinte, isto evita a geração de mais tráfego de controle em um ambiente congestionado e com escassez de recursos. É importante apontar que essas abordagens são as únicas modificações para o padrão LTE propostas nesta dissertação. Vale ressaltar que a abordagem de adicionar $1 + n$ classes apresenta uma simples modificação no padrão.

6.2 Contribuições

A principal contribuição desta dissertação é o mecanismo de alocação de recursos no *uplink* da rede LTE para a comunicação M2M. Contribuições secundárias surgem das funcionalidades deste mecanismo, entre elas:

- Identificação do tipo de comunicação (M2M e H2H) através dos requisitos de QoS do tráfego.
- Controle do impacto nas principais comunicações atuais da rede LTE (H2H) ao introduzir a futura comunicação M2M na rede.

- O algoritmo simples e viável do escalonamento de pacotes da comunicação M2M para satisfazer os requisitos de QoS do tráfego M2M e evitar o problema de inanição na alocação dos recursos.

Outra contribuição secundária são os códigos das implementações dos experimentos em C++ para o simulador de rede NS-3 que são disponibilizados aos pesquisadores interessados em evoluir este trabalho. Os códigos podem ser encontrados no endereço <http://bit.ly/1gPIV6d>.

6.3 Trabalhos futuros

No desenvolvimento desta dissertação, algumas questões foram negligenciadas em função do tempo disponível para sua implementação. Desta forma, alguns pontos que podem ser trabalhados para melhorar o desempenho da solução proposta nesta dissertação são listados a seguir:

- Analisar o desempenho do mecanismo proposto neste trabalho para cenários com dispositivos M2M que transmitem em intervalos regulares maiores que o limite superior utilizado nos experimentos (550 ms).
- Implementar uma heurística mais elaborada para a escolha dos recursos alocados para os dispositivos na segunda fase do algoritmo de escalonamento (seção 4.4).
- Realizar um refinamento na parametrização das constantes da solução proposta para obter uma melhor performance, do que a que foi apresentada, dos objetivos discutidos principalmente para as aplicações M2M baseadas em evento.
- Aperfeiçoar os parâmetros das classes de QoS para o tráfego M2M para se aproximar mais das características das aplicações M2M reais.

Outros trabalhos futuros interessantes que podem ser citados são (i) a utilização da característica de agrupamento dos dispositivos das aplicações M2M no escalonamento e (ii) a adaptação das constantes do mecanismo, proposto nesta dissertação, segundo o estado do sistema.

REFERÊNCIAS BIBLIOGRÁFICAS

- 3GPP. *Analysis on traffic model and characteristics for MTC and text proposal*. [S.l.], February 2012. Disponível em: <<http://www.3gpp.org>>.
- 3GPP. *Service requirements for Machine-Type Communications*. [S.l.], March 2012.
- 3GPP. *LTE*. December 2013. Disponível em: <<http://www.3gpp.org/LTE>>. Acesso em: Dezembro 12, 2013.
- 3GPP. *Policy and charging control architecture*. [S.l.], Setembro 2013.
- AHMED, I.; MOHAMED, A.; SHAKEEL, I. On the group proportional fairness of frequency domain resource allocation in l-sc-fdma based lte uplink. In: *GLOBECOM Workshops (GC Wkshps), 2010 IEEE*. [S.l.: s.n.], 2010. p. 1312–1317.
- ALAM, M.; NIELSEN, R.; PRASAD, N. The evolution of m2m into iot. In: *Communications and Networking (BlackSeaCom), 2013 First International Black Sea Conference on*. [S.l.: s.n.], 2013. p. 112–115.
- ALI, S.; ZEESHAN, M. A utility based resource allocation scheme with delay scheduler for lte service-class support. In: *Wireless Communications and Networking Conference (WCNC), 2012 IEEE*. [S.l.: s.n.], 2012. p. 1450–1455. ISSN 1525-3511.
- ALI-YAHIYA, T. *Understanding LTE and its Performance*. 1st. ed. [S.l.]: Springer Science+Business Media, 2011. ISSN 978-1-4419-6456-4.
- AMBERKAR, S. *LTE - basic PHY structure*. Dezembro 2013. Disponível em: <<http://samiramberkar-tutorials.blogspot.com.br/2011/07/lte-20.html>>. Acesso em: Dezembro 9, 2013.
- AMOKRANE, A.; KSENTINI, A.; HADJADJ-AOUL, Y. *Congestion Control in the context of Machine Type Communications in Long Term Evolution networks*. Dissertação (Mestrado) — ENS Cachan Bretagne, 2011.
- ANAS, M. et al. Combined admission control and scheduling for qos differentiation in lte uplink. In: *Vehicular Technology Conference, 2008. VTC 2008-Fall. IEEE 68th*. [S.l.: s.n.], 2008. p. 1–5. ISSN 1090-3038.
- ATZORI, L.; IERA, A.; MORABITO, G. The internet of things: A survey. *Computer Networks*, v. 54, n. 15, p. 2787 – 2805, 2010. ISSN 1389-1286. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S1389128610001568>>.
- BALDO, N. et al. An open source product-oriented lte network simulator based on ns-3. In: *Proceedings of the 14th ACM International Conference on Modeling, Analysis and Simulation of Wireless and Mobile Systems*. New York, NY, USA: ACM, 2011. (MSWiM '11), p. 293–298. ISBN 978-1-4503-0898-4. Disponível em: <<http://doi.acm.org/10.1145/2068897.2068948>>.
- BASSI, A.; HORN, G. Internet of things in 2020: A roadmap for the future. *European Commission: Information Society and Media*, 2008.

BOSWARTHICK, D.; ELLOUMI, O.; HERSENT, O. *M2m communications: a systems approach*. [S.l.]: Wiley. com, 2012.

BROWN, J.; KHAN, J. Predictive resource allocation in the lte uplink for event based m2m applications. In: *Communications Workshops (ICC), 2013 IEEE International Conference on*. [S.l.: s.n.], 2013. p. 95–100.

CALABRESE, F. et al. Search-tree based uplink channel aware packet scheduling for utran lte. In: *Vehicular Technology Conference, 2008. VTC Spring 2008. IEEE*. [S.l.: s.n.], 2008. p. 1949–1953. ISSN 1550-2252.

CHEN, Y.; WANG, W. Machine-to-machine communication in lte-a. In: *Vehicular Technology Conference Fall (VTC 2010-Fall), 2010 IEEE 72nd*. [S.l.: s.n.], 2010. p. 1–4. ISSN 1090-3038.

GALETIĆ, V. et al. Basic principles of machine-to-machine communication and its impact on telecommunications industry. In: *MIPRO, 2011 Proceedings of the 34th International Convention*. [S.l.: s.n.], 2011. p. 380–385.

GHANDOUR, F.; FRIKHA, M.; TABBANE, S. A fair and power saving uplink scheduling scheme for 3gpp lte systems. In: *Network of the Future (NOF), 2011 International Conference on the*. [S.l.: s.n.], 2011. p. 6–9.

GOTSIS, A.; LIOUMPAS, A.; ALEXIOU, A. M2m scheduling over lte: Challenges and new perspectives. *Vehicular Technology Magazine, IEEE*, v. 7, n. 3, p. 34–39, Sept 2012. ISSN 1556-6072.

GOTSIS, A. G.; LIOUMPAS, A. S.; ALEXIOU, A. Analytical modelling and performance evaluation of realistic time-controlled m2m scheduling over lte cellular networks. *Transactions on Emerging Telecommunications Technologies*, Wiley Online Library, 2013.

ITU-R. Technical Report, *Requirements Related to Technical Performance for IMT-Advanced Radio Interface(s)*. 2008.

JAIN, R. *The art of computer systems performance analysis*. [S.l.]: John Wiley & Sons Chichester, 1991.

KWAN, R.; LEUNG, C. A survey of scheduling and interference mitigation in lte. *JECE*, Hindawi Publishing Corp., New York, NY, United States, v. 2010, p. 1:1–1:10, jan. 2010. ISSN 2090-0147. Disponível em: <<http://dx.doi.org/10.1155/2010/273486>>.

LEE, S.-B. et al. Proportional fair frequency-domain packet scheduling for 3gpp lte uplink. In: *INFOCOM 2009, IEEE*. [S.l.: s.n.], 2009. p. 2611–2615. ISSN 0743-166X.

LIEN, S.-Y.; CHEN, K.-C.; LIN, Y. Toward ubiquitous massive accesses in 3gpp machine-to-machine communications. *Communications Magazine, IEEE*, v. 49, n. 4, p. 66–74, 2011. ISSN 0163-6804.

LIM, J. et al. Channel-dependent scheduling of uplink single carrier fdma systems. In: *Vehicular Technology Conference, 2006. VTC-2006 Fall. 2006 IEEE 64th*. [S.l.: s.n.], 2006. p. 1–5.

- LIMA, F. R. M. et al. Maximizing spectral efficiency with acceptable service provision in multiple antennas scenarios. In: *European Wireless, 2012. EW. 18th European Wireless Conference*. [S.l.: s.n.], 2012. p. 1–8.
- LIOUMPAS, A.; ALEXIOU, A. Uplink scheduling for machine-to-machine communications in lte-based cellular systems. In: *GLOBECOM Workshops (GC Wkshps), 2011 IEEE*. [S.l.: s.n.], 2011. p. 353–357.
- LIU, R. et al. M2m-oriented qos categorization in cellular network. In: *Wireless Communications, Networking and Mobile Computing (WiCOM), 2011 7th International Conference on*. [S.l.: s.n.], 2011. p. 1–5. ISSN 2161-9646.
- MARWAT, S. et al. Design and performance analysis of bandwidth and qos aware lte uplink scheduler in heterogeneous traffic environment. In: *Wireless Communications and Mobile Computing Conference (IWCMC), 2012 8th International*. [S.l.: s.n.], 2012. p. 499–504.
- MEHTA, M.; KHAKUREL, S.; KARANDIKAR, A. Buffer-based channel dependent uplink scheduling in relay-assisted lte networks. In: *Wireless Communications and Networking Conference (WCNC), 2012 IEEE*. [S.l.: s.n.], 2012. p. 1777–1781. ISSN 1525-3511.
- MIORANDI, D. et al. Internet of things: Vision, applications and research challenges. *Ad Hoc Networks*, v. 10, n. 7, p. 1497 – 1516, 2012. ISSN 1570-8705. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S1570870512000674>>.
- NS-3. *The network simulator NS-3*. Dezembro 2013. Disponível em: <<http://www.nsnam.org>>. Acesso em: Dezembro 12, 2013.
- PHAM, H.; VU, X. N.; HWANG, S.-H. Service class-aided scheduling for lte. In: *Advanced Communication Technology (ICACT), 2011 13th International Conference on*. [S.l.: s.n.], 2011. p. 39–43. ISSN 1738-9445.
- PIRO, G.; BALDO, N.; MIOZZO, M. An lte module for the ns-3 network simulator. In: *Proceedings of the 4th International ICST Conference on Simulation Tools and Techniques*. ICST, Brussels, Belgium, Belgium: ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), 2011. (SIMUTools '11), p. 415–422. ISBN 978-1-936968-00-8. Disponível em: <<http://dl.acm.org/citation.cfm?id=2151054.2151129>>.
- POKHARIYAL, A. et al. Harq aware frequency domain packet scheduler with different degrees of fairness for the utran long term evolution. In: *Vehicular Technology Conference, 2007. VTC2007-Spring. IEEE 65th*. [S.l.: s.n.], 2007. p. 2761–2765. ISSN 1550-2252.
- POTSCH, T. et al. Influence of future m2m communication on the lte system. In: *Wireless and Mobile Networking Conference (WMNC), 2013 6th Joint IFIP*. [S.l.: s.n.], 2013. p. 1–4.
- REN, F. et al. Frequency domain packet scheduling with stability analysis for 3gpp lte uplink. *Mobile Computing, IEEE Transactions on*, v. 12, n. 12, p. 2412–2426, 2013. ISSN 1536-1233.
- SALAH, M. et al. Evaluating uplink schedulers in lte in mixed traffic environments. In: *Communications (ICC), 2011 IEEE International Conference on*. [S.l.: s.n.], 2011. p. 1–5. ISSN 1550-3607.
- SOKMEN, F.; GIRICI, T. Uplink resource allocation algorithms for single-carrier fdma systems. In: *Wireless Conference (EW), 2010 European*. [S.l.: s.n.], 2010. p. 339–345.

- STRATEGY, I. T. U.; UNIT, Policy. *Itu internet reports 2005: The internet of things. Geneva: International Telecommunication Union (ITU), 2005.*
- TEKTRONIX. *LTE Networks: Evolution and Technology Overview*. June 2012. Disponível em: <www.tektronixcommunications.com/LTE>. Acesso em: June 26, 2012.
- TEMINO, L. Ruiz de et al. Channel-aware scheduling algorithms for sc-fdma in lte uplink. In: *Personal, Indoor and Mobile Radio Communications, 2008. PIMRC 2008. IEEE 19th International Symposium on*. [S.l.: s.n.], 2008. p. 1–6.
- UCKELMANN, D.; HARRISON, M.; MICHAHELLES, F. An architectural approach towards the future internet of things. In: UCKELMANN, D.; HARRISON, M.; MICHAHELLES, F. (Ed.). *Architecting the Internet of Things*. [S.l.]: Springer Berlin Heidelberg, 2011. p. 1–24. ISBN 978-3-642-19156-5.
- WANG, X.; KONISHI, S. Optimization formulation of packet scheduling problem in lte uplink. In: *Vehicular Technology Conference (VTC 2010-Spring), 2010 IEEE 71st*. [S.l.: s.n.], 2010. p. 1–5. ISSN 1550-2252.
- WONG, I.; OTERI, O.; MCCOY, W. Optimal resource allocation in uplink sc-fdma systems. *Wireless Communications, IEEE Transactions on*, v. 8, n. 5, p. 2161–2165, 2009. ISSN 1536-1276.
- WU, G. et al. M2m: From mobile to embedded internet. *Communications Magazine, IEEE*, v. 49, n. 4, p. 36–43, 2011. ISSN 0163-6804.
- YAACOUB, E.; DAWY, Z. A game theoretical formulation for proportional fairness in lte uplink scheduling. In: *Wireless Communications and Networking Conference, 2009. WCNC 2009. IEEE*. [S.l.: s.n.], 2009. p. 1–5. ISSN 1525-3511.
- ZHENQI, S. et al. Research on uplink scheduling algorithm of massive m2m and h2h services in lte. In: *Information and Communications Technologies (IETICT 2013), IET International Conference on*. [S.l.: s.n.], 2013. p. 365–369.
- ZHOU, D.; BALDO, N.; MIOZZO, M. Implementation and validation of lte downlink schedulers for ns-3. In: *Proceedings of the 6th International ICST Conference on Simulation Tools and Techniques*. ICST, Brussels, Belgium, Belgium: ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), 2013. (SIMUTools '13).
- ZHOU, K.; NIKAEIN, N.; KNOPP, R. Dynamic resource allocation for machine-type communications in lte/lte-a with contention-based access. In: *Wireless Communications and Networking Conference (WCNC), 2013 IEEE*. [S.l.: s.n.], 2013. p. 256–261. ISSN 1525-3511.
- ZHOU, K. et al. Contention based access for machine-type communications over lte. In: *Vehicular Technology Conference (VTC Spring), 2012 IEEE 75th*. [S.l.: s.n.], 2012. p. 1–5. ISSN 1550-2252.