

O Impacto da Clusterização na Eficiência Energética de Redes de Comunicação entre Máquinas

W. S. Souza e C. H. M. de Lima

Resumo—Nesse trabalho, estudamos o desempenho de sistemas hierárquicos com implantação massiva de máquinas usando estratégias de clusterização baseadas em *Machine Learning*. Simulações computacionais utilizando a abordagem de Monte Carlo são empregadas para obter as figuras de desempenho com relação à latência e eficiência energética do sistema. Nos cenários em estudo, o canal rádio móvel incorpora o efeito degradante do desvanecimento de larga escala, e o sistema opera baseado em estratégias de múltiplo acesso com sensoriamento de portadora. Os resultados numéricos mostram que a implantação da técnica *K-Means* aumentou em cerca de 3dB a eficiência energética dos sistemas em estudo.

Palavras-Chave—M2M, URC, backoff, contenção, eficiência energética, Machine Learning, latência, CSMA/CA.

Abstract—In this work, we assess the system performance of large scale deployments of machine type communication employing clustering strategies based on Machine Learning. Performance figures evaluating the end-to-end latency and energy efficiency were obtained through computational simulations following the Monte Carlo approach. The scenarios under study are degraded by large scale fading, and the system operates based on multiple access protocols with carrier sensing. Our numerical results show that the clustering strategy *K-Means* increases the energy efficiency of systems requiring URC in 3dB.

Keywords—M2M, URC, backoff, content, energy efficiency, machine learning, latency, CSMA/CA.

I. INTRODUÇÃO

As comunicações entre máquinas (M2M, em inglês *machine-to-machine*) estão cada vez mais presentes em nosso cotidiano, viabilizando em qualquer lugar e a qualquer momento o acesso a serviços essenciais, como transações bancárias, aprendizagem, saúde e entretenimento. Além de aplicações essenciais como as descritas, a Internet das Coisas (IoT, em inglês *Internet of Things*) e a comunicação entre máquinas terão aplicações em dispositivos domésticos, sistemas de estacionamento, transporte público, segurança pública e sistemas industriais [1]. Para viabilizar um grande número de dispositivos conectados, a próxima geração de redes de comunicações móveis requer grandes mudanças [2].

É previsto que 50 bilhões de dispositivos estejam conectados em 2020, um número dez vezes maior do que os valores de 2010 [3]. Além disso, mais da metade da população estará sob a cobertura de sistemas LTE (*Long Term Evolution*). Nessas condições, será necessário um aumento na capacidade de tráfego de até mil vezes com relação à capacidade atual.

W. S. Souza e C. H. M. de Lima, Universidade Estadual Paulista (UNESP), Câmpus de São João da Boa Vista-SP, Brasil, E-mails: weler-son.s@hotmail.com, carlos.lima@unesp.br. Este trabalho foi financiado pela FAPESP (2017/18351-5). Os autores agradecem ao suporte do CNPq através do projeto Universal n.432925/2016-3

Estudos identificaram as redes 5G como capazes de habilitar a visão de uma Sociedade Conectada (em inglês, *Networked Society*) [4]–[6] onde todas as pessoas e todos dispositivos estão conectados e recebem informação a qualquer momento.

A ideia de ter uma Sociedade Conectada é muito útil, porém, para que se torne realidade, um desafio técnico é imposto. O grande desafio para termos uma vida onde tudo que usamos, desde eletrodomésticos até os nossos carros, estejam conectados é a necessidade de uma capacidade de tráfego muito superior aos níveis atuais. Para obter esses resultados, vários modelos de implantação foram sugeridos e, além disso, diversos fatores limitadores foram identificados. Três exemplos de sistemas identificados foram o IEEE 802.11ah (tecnologia WiFi com baixo consumo), redes com baixo consumo de energia em espectro não licenciado, e comunicação entre máquinas em redes celulares. Dos exemplos citados, o mais atrativo é utilizar as redes celulares por já permitirem muitos dispositivos conectados e ter ampla cobertura. O desafio de usar as redes celulares é viabilizar a coexistência de sistemas de comunicação sem fio orientados a humanos e orientados às máquinas [7].

No cenário atual, os sistemas são na sua maior parte orientados a humanos e projetados para assegurar boa conectividade na maior parte do tempo, porém, em locais com pouca cobertura, alta interferência ou sobrecarga, a taxa de transferência é quase nula. Já sistemas orientados às máquinas priorizam a confiabilidade e continuidade da comunicação, geralmente em taxas de transmissão moderadas. Os pacotes dos sistemas atuais são divididos em uma grande parte de dados (em inglês, *payload*) e uma pequena parte de informações de controle (em inglês, *metadata*). Já na Comunicação entre Máquinas (MTC, em inglês *Machine Type Communication*) os pacotes são curtos e o tamanho do *Payload* e do *Metadata* são equivalentes [8].

As aplicações que utilizam esses sistemas orientados às máquinas exigem latência pequena e confiabilidade alta, logo, operam em regime de Comunicação Ultra Confiável (URC, em inglês *Ultra Reliable Communication*). URC é um modo de operação que não existe atualmente e representa a necessidade de um alto nível de qualidade de serviço por quase todo o tempo de conexão [9]. A Comunicação Ultra Confiável entre máquinas tem aplicações como automação industrial, *Cloud Computing* e comunicação entre veículos [10], [11].

A. Apresentação do problema e organização do trabalho

A comunicação massiva entre máquinas (MTC, em inglês *Machine Type Communication*) aumenta a latência dos nós que estão ativos em dada região e compartilham o enlace de comunicação. Se vários nós tentam transmitir pacotes

simultaneamente, eles entram em contenção para uso do canal e o algoritmo de tempo de *backoff* exponencial os divide em *slots* de tempo para diminuir as colisões. Quando existe a primeira colisão, os nós ativos com pacotes a serem transmitidos recebem um tempo aleatório que devem aguardar até que tentem transmitir novamente. Além de organizar a contenção, o algoritmo também auxilia na eficiência energética do sistema pois os nós só tentam transmissão após o tempo que foi designado a eles.

Ainda assim, dispositivos que operam em regime URC podem experimentar latência alta uma vez que podem esperar muito tempo até que possam utilizar o canal. Para que esses consigam atender os requisitos de latência, uma estratégia de *backoff* que leva em consideração a prioridade do dispositivo é utilizada e esses tem menor latência [12], [13].

Foram implementadas duas técnicas de clusterização a fim de comparar os resultados numéricos, a primeira foi a técnica de clusterização em setores já utilizada em [13]. Já a segunda foi a técnica de clusterização *K-Means* baseada em *Machine Learning*. A performance das duas técnicas foi avaliada em cenários de comunicação massiva hierárquica entre máquinas pelos tempos de transmissão fim-a-fim dos nós e a eficiência energética de cada nó.

Este trabalho está dividido em 4 seções. A Seção II descreve o modelo de propagação utilizado nas simulações computacionais, bem como a estratégia utilizada para diminuir a latência dos nós e então, apresenta as técnicas de clusterização utilizadas para implementar a hierarquia de comunicação entre máquinas. Os resultados das simulações computacionais são apresentados e discutidos na Seção III e por fim, na Seção IV, são apresentadas as observações finais e conclusões sobre o trabalho realizado.

II. MODELO DE SISTEMA

Inicialmente, apresentaremos o modelo de propagação do enlace de rádio utilizado, em seguida apresentaremos a estratégia que foi utilizada para diminuir a latência dos nós e atender os requisitos do regime URC. Finalmente, serão apresentadas as técnicas de clusterização dos nós que serão objetos de estudo desse trabalho.

A. Modelo de propagação e implantação da rede

Em nosso modelo, o enlace de rádio é degradado pela perda de percurso e sombreamento [14], [15]. A perda de percurso tem comportamento linear em função do logaritmo da distância e quanto maior a distância entre o transmissor e o receptor, maior a perda de percurso. Já o sombreamento tem distribuição gaussiana. A potência recebida pelo receptor de interesse a uma distancia d de um transmissor é dada por

$$p_{rx} = p_{tx} d^{-\alpha} x_{\sigma}, \quad (1)$$

em que p_{tx} é a potência transmitida em *Watts*, α é o expoente de perda de percurso e x_{σ} representa o sombreamento Log-Normal (LN). Os valores de expoente de perda de percurso α e desvio padrão de sombreamento σ são encontrados em [15] e dependem do ambiente de propagação. O número de pontos

em uma região arbitraria de \mathbb{R}^2 segue distribuição de *Poisson* [16], logo o numero de dispositivos que são dispostos na área da rede aleatoriamente é dado dessa maneira.

B. Controle de acesso ao meio

Nesse trabalho, os dispositivos compartilham o enlace de rádio e utilizam o método de múltiplo acesso CSMA/CA onde os dispositivos com dados a transmitir acessam o canal rádio móvel independentemente. Uma vez que o canal está ocupado, atribui-se um tempo de *backoff* para que o dispositivo tente transmitir o pacote novamente. O algoritmo de *backoff* exponencial segue os padrões IEEE 802.11 e 802.15.4 e o tempo que é atribuído aos dispositivos após cada colisão é dado por

$$\Delta = \lfloor 2^i \times \text{Uniform}(0, 1) \rfloor \times \delta, \quad (2)$$

em que $\text{Uniform}(0, 1)$ é um número aleatório com distribuição uniforme entre 0 e 1, $\lfloor \cdot \rfloor$ trunca o número para o menor inteiro, δ é o slot de tempo de *backoff* definido pelo sistema e i é um numero não negativo. O índice i representa o número da tentativa de transmissão, é iniciado em i_{\min} e acrescido de 1 a cada tentativa. Se o valor de i_{\max} for atingido, o pacote é perdido. Além disso, se o t_{\max} (tempo máximo de espera para que o pacote chegue até seu destino final) for atingido, o pacote também é descartado.

A latência para qualquer dispositivo ainda pode ser alta, um dispositivo que opera em URC e requer latência baixa pode ficar no processo de contenção durante um período muito longo ou ainda não conseguir transmitir. A alternativa utilizada para diminuir a latência desses dispositivos é designando prioridade a eles para que sempre tenham tempo de transmissão fim-a-fim menor que os outros. O critério para atribuição dessa prioridade pode ser o tamanho dos pacotes enviados, a importância da aplicação de determinado dispositivo, a relevância dos dados a serem transmitidos, entre outros.

Nesse trabalho, nosso objetivo é avaliar a eficiência do algoritmo com prioridade, logo, separaremos os nós entre os que não tem prioridade e os que tem sem utilizar um critério específico para designação de prioridade. O tempo de *backoff* utilizando essa regra é dado por

$$\Delta' = \left\lfloor \frac{2^i \times \text{Uniform}(0, 1)}{fp} \right\rfloor \times \delta, \quad (3)$$

em que fp é o fator de prioridade.

C. Hierarquia do sistema de comunicação M2M e técnicas de clusterização

Foram implementadas redes de comunicação entre máquinas hierárquicas nas simulações computacionais usando a abordagem de Monte Carlo. Em nosso trabalho, os nós dispostos na área da rede são divididos em três *clusters* e cada um tem um nó denominado *cluster head* localizado estrategicamente para aumentar a eficiência energética do sistema. A função desse é receber os pacotes de todos os nós de seu *cluster* e enviar para o *cluster head* subsequente até que o pacote chegue no *cluster head* do terceiro *cluster*, denominado como *sink node*. A Fig. 1 ilustra uma rede hierárquica com três *clusters*.

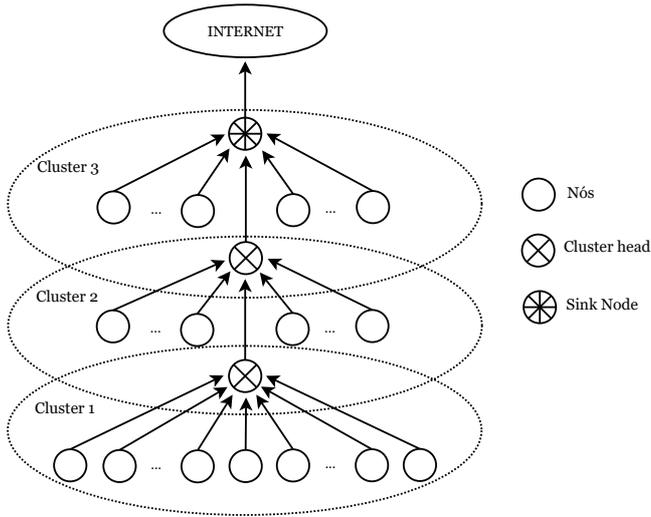


Fig. 1. Hierarquia da Rede de Comunicação entre Máquinas.

Utilizaremos duas estratégias para dividir os nós em *clusters*: o método de clusterização em setores e o *K-Means*.

1) *Clusterização em setores*: O método de clusterização em setores, já utilizado em [13], é um método idealizado onde a área de implantação da rede é dividida em três faixas retangulares verticais idênticas e cada uma corresponde a um *cluster*. O *cluster head* de cada *cluster* é o nó que tem a menor distância até o centro de cada setor e são escolhidos deste modo para que a diferença das distâncias entre os nós e o *cluster head* seja a menor possível. A Fig. 2 mostra um cenário clusterizado em setores, as setas demonstram o trajeto feito por um pacote saindo de um nó localizado no *cluster 1* em direção ao *sink node*.

2) *K-Means*: Esse método agrupa os nós tentando separá-los em n *clusters* de variância igual, minimizando a soma de mínimos quadrados dentro do *cluster* [17]. Nesse trabalho os cenários de implantação da rede contam com 3 *clusters*.

O algoritmo *K-Means* divide um conjunto de N amostras no espaço \mathcal{A} , área de implantação da rede, em K *clusters*, cada um descrito pela média μ_j das amostras do *cluster*, denominada centroide do *cluster*. Os centroides do *cluster* não necessariamente são nós que foram dispostos na área. O algoritmo visa escolher centroides que minimizem o critério de soma dos mínimos quadrados dentro do *cluster*:

$$\sum_{i=0}^n \min_{\mu_j \in C} (||\alpha_j - \mu_i||^2), \quad (4)$$

em que α_j é a posição de cada nó e μ_i é a posição do centroide. *K-Means* é geralmente chamado de algoritmo de *Lloyd* e adota três etapas. A primeira consiste em fazer a escolha dos centróides iniciais, o algoritmo escolhe k amostras do espaço X . Após isso, o algoritmo faz um *loop* entre atribuir cada amostra ao seu centróide mais próximo e criar novos centróides tomando o valor médio de todas as amostras atribuídas a cada centróide anterior. O algoritmo repete essas duas etapas até que a diferença entre o antigos e os novos centróides seja menor que um dado valor.

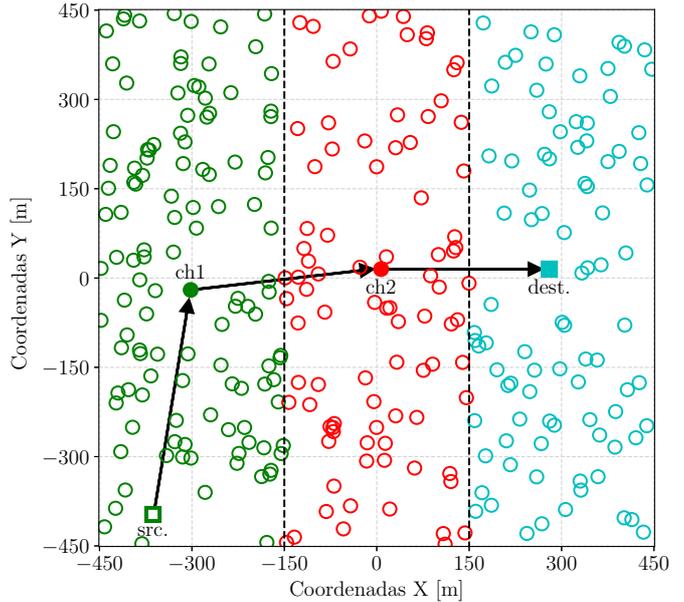
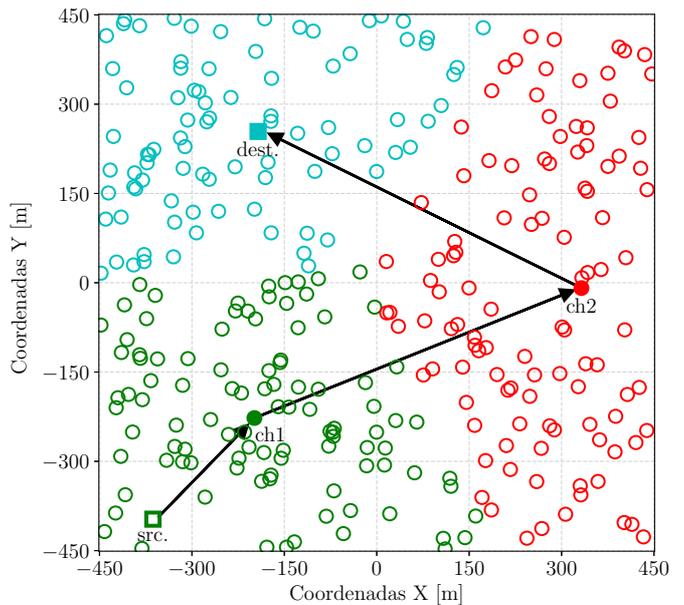


Fig. 2. Cenário clusterizado em setores.


 Fig. 3. Cenário clusterizado por *K-Means*.

A Fig. 3 mostra um cenário clusterizado utilizando esse método, assim como no método anterior, as setas demonstram a transmissão de um pacote do *cluster 1* até o *sink node*.

III. AVALIAÇÃO DE DESEMPENHO

O impacto de utilização do método *K-Means* baseado em *Machine Learning* na clusterização de redes hierárquicas de comunicação entre máquinas (M2M) será avaliado em termos de latência e eficiência energética da rede. Para isso, os cenários foram construídos e simulados utilizando a abordagem de Monte Carlo para obter as métricas em 10000 realizações (snapshots) de cada cenário. Os parâmetros utilizados nas simulações são apresentados na Tabela I.

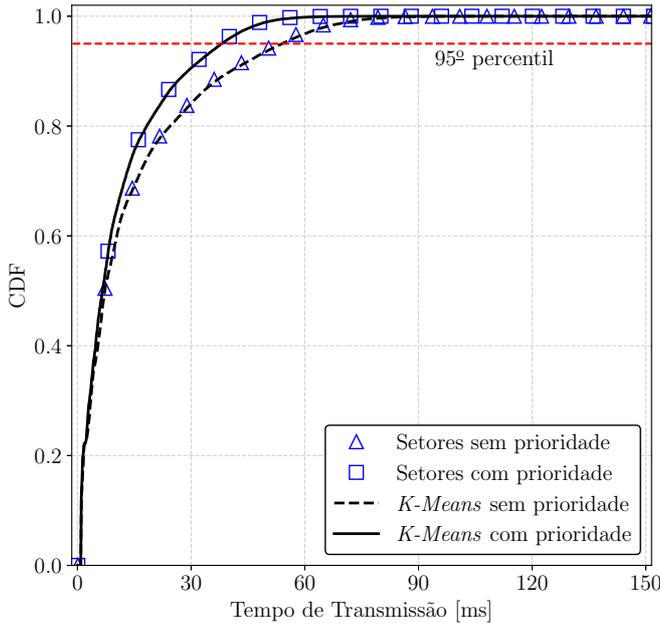


Fig. 4. CDF dos tempos de transmissão fim-a-fim.

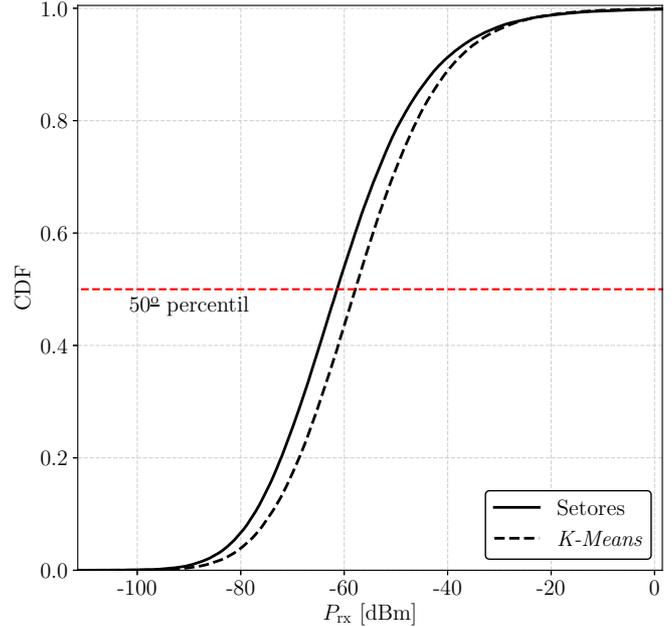
 TABELA I
 PARÂMETROS UTILIZADOS NAS SIMULAÇÕES.

Parâmetro	Valor
Número médio de dispositivos por cenário	60
Potência transmitida pelos nós	2 W (33 dBm)
Expoente de perda de percurso (α)	4
Desvio padrão de sombreamento (σ)	10 dB
Fator de prioridade (fp)	1 ou 2
Tempo de transmissão	1ms
$backoff_slot_time$	100 μ s
i_{min}	2
i_{max}	10

O número médio de dispositivos por cenário foi definido de acordo com o tamanho da área em que os cenários serão implantados e simulados. Os valores do expoente de perda de percurso e do desvio padrão foram escolhidos assumindo a implantação da rede em uma área urbana. Habitualmente, em comunicações de dispositivos, um *frame* tem 10ms e é dividido em 10 *slots* de 1ms que corresponde ao tempo de transmissão de um pacote. O $backoff_slot_time$, i_{min} e o i_{max} são parâmetros encontrados em [12].

Inicialmente, a latência será avaliada. Os tempos de transmissão fim-a-fim foram avaliados para cenários clusterizados pelos dois métodos estudados utilizando a estratégia de *backoff* com e sem prioridade.

A Fig. 4 mostra a função de distribuição acumulada (em inglês, Cumulative Distribution Function) dos tempos de transmissão fim-a-fim da rede clusterizada em setores e pelo algoritmo *K-Means* considerando a regra de *backoff* com e sem prioridade. A curva tracejada mostra a CDF do sistema clusterizado pelo método *K-Means* sem utilizar designação de prioridades e a curva contínua utiliza o mesmo método de clusterização mas, nessa, é utilizada a designação de prioridades aos nós.


 Fig. 5. CDF das potências recebidas pelos *cluster-heads*

A latência dos dispositivos no sistema que atribui prioridades é consideravelmente menor, apresentando uma redução na latência de até 15 ms (cerca de um *frame* e meio) no 95º percentil de probabilidade acumulada. Os triângulos e quadrados mostram a CDF do sistema clusterizado em setores sem e com designação de prioridade, respectivamente.

Os resultados dos dois métodos coincidiram pois a diferença de distância entre os nós e os *cluster heads*, obtida pela escolha do método de clusterização, tem influência desprezível no tempo total fim-a-fim uma vez que nesses casos o tempo de propagação do sinal é muito menor que o tempo de espera na fila. Sendo assim, o método de designar prioridades a dispositivos que operam em URC, além de ser eficaz para sistemas clusterizados em setores, também é eficaz para diminuir a latência de sistemas que utilizam a técnica de *K-Means*.

A escolha do método de clusterização não proporciona diferença efetiva na latência dos dispositivos, porém, o efeito aparece em outra métrica de grande importância em comunicações entre máquinas, os níveis de potência recebidos pelos *cluster heads*. A Fig. 5 mostra a CDF desses níveis. Vemos que os níveis de potências recebidos pelos *cluster heads* no cenário clusterizado pelo método *K-Means* são maiores que os recebidos no clusterizado em setores, em torno de 3 dB de ganho.

IV. CONCLUSÕES

Neste trabalho, discutiu-se métodos para clusterizar os nós em sistemas com implantação massiva de comunicação entre máquinas a fim de que o sistema opere de maneira hierárquica e em regime URC.

O desempenho do sistema é avaliado em cenários clusterizados em setores e cenários clusterizados utilizando o algoritmo *K-Means* baseado em *Machine Learning*. Observou-se que a latência no sistema clusterizado pelo método *K-Means* tem

valores similares aos encontrados no sistema clusterizado em setores com uma redução na latência de até 15 milissegundos utilizando a estratégia de *backoff* com prioridade. Por outro lado, analisando a potência recebida pelos *cluster heads* em cada um dos cenários, nota-se que os níveis de potência recebidos no cenário clusterizado pelo método *K-Means* é cerca de 3 dB maior. Dessa maneira, embora não exista ganho em relação a latência ao utilizar o método de clusterização baseado em *Machine Learning*, existe um ganho considerável na eficiência energética.

REFERÊNCIAS

- [1] Cisco, “The internet of everything,” www.cisco.com/web/about/ac79/docs/innov/IoE.pdf, acesso em: 13 de fevereiro de 2018.
- [2] “The Metis2020 project - laying the foundation of 5G,” www.metis2020.com, acesso em 13 de fevereiro de 2018.
- [3] A. Osseiran, F. Boccardi, and V. Braun, “Scenarios for 5G mobile and wireless communications: the vision of the METIS project,” vol. 52, no. 5, pp. 26–35, May 2014.
- [4] Ericsson, “The networked society is here,” acesso em: 14 de fevereiro de 2018.
- [5] J. vanDijk, *The Network Society*, 3rd ed. SAGE Publications, 2012.
- [6] M. Castells, *The Rise of the Network Society – The Information Age: Economy, Society, and Culture*, 2nd ed. Wiley–Blackwell, 2009.
- [7] H. Shariatmadari, R. Ratasuk, S. Iraj, A. Laya, T. Taleb, R. Jäntti, and A. Ghosh, “Machine-type communications: current status and future perspectives toward 5G systems,” *IEEE Communications Magazine*, vol. 53, no. 9, pp. 10–17, 2015.
- [8] T. M. Cover and J. A. Thomas, Eds., *Elements of Information Theory*, 2nd ed. Wiley-Interscience, 2006.
- [9] P. Popovski, “Ultra-reliable communication in 5G wireless systems,” in *5G for Ubiquitous Connectivity (5GU), 2014 1st International Conference on*. IEEE, 2014, pp. 146–151.
- [10] C.-X. Wang, F. Haider, X. Gao, X.-H. You, Y. Yang, D. Yuan, H. Aggoune, H. Haas, S. Fletcher, and E. Hepsaydir, “Cellular architecture and key technologies for 5G wireless communication networks,” *IEEE Communications Magazine*, vol. 52, no. 2, pp. 122–130, February 2014.
- [11] F. Boccardi, R. W. Heath, A. Lozano, T. L. Marzetta, and P. Popovski, “Five disruptive technology directions for 5g,” *IEEE Communications Magazine*, vol. 52, no. 2, pp. 74–80, February 2014.
- [12] I. Park, D. Kim, and D. Har, “MAC achieving low latency and energy efficiency in hierarchical M2M networks with clustered nodes,” *IEEE Sensors Journal*, vol. 15, no. 3, pp. 1657–1661, 2015.
- [13] W. S. Souza and C. H. M. Lima, “Estratégia de backoff com prioridade em sistemas hierárquicos de comunicação entre máquinas,” *Anais do XXXV Simpósio Brasileiro De Telecomunicações E Processamento De Sinais*, pp. 776–777, 2017.
- [14] A. Goldsmith, *Wireless communications*. Cambridge university press, 2005.
- [15] T. S. Rappaport, *Comunicações sem fio: princípios e práticas*. Pearson Prentice Hall, 2009.
- [16] J. F. C. Kingman, *Poisson processes*. Wiley Online Library, 1993.
- [17] D. Arthur and S. Vassilvitskii, “k-means++: The advantages of careful seeding,” in *Proceedings of the eighteenth annual ACM-SIAM symposium on Discrete algorithms*. Society for Industrial and Applied Mathematics, 2007, pp. 1027–1035.