

# Balanceamento de carga cognitivo em redes móveis heterogêneas e auto-organizáveis

Cesar A. Sierra Franco e José Roberto B. de Marca

**Resumo**— A contínua evolução das tecnologias móveis celulares e o uso de topologias de rede com maior densificação de estações rádio-base, têm gerado novos desafios em termos de configuração e otimização do sistema. Especificamente, nos cenários denominados heterogêneos que misturam o uso de macro e pico células, o balanceamento de carga é um desses desafios dado que os terminais selecionam usualmente a estação rádio-base da qual recebem o sinal com a maior potência. Este artigo propõe um esquema de expansão dinâmica da célula baseado em uma abordagem de aprendizado automático. Simulações ao nível do sistema mostraram a efetividade desta abordagem para expandir dinamicamente a cobertura das pico células conforme as condições de tráfego, balanceando a carga e melhorando o desempenho dos usuários.

**Palavras-Chave**— Balanceamento de carga; redes móveis; redes heterogêneas; redes cognitivas;

**Abstract**— The continuous evolution of cellular communication infrastructure into dense, dynamic and heterogeneous networks has posed new challenges for system configuration as well as coverage and capacity optimization, especially in areas with unequal user traffic distribution. In a mixed macro/pico cell scenario, load balance is one of those challenges, since users typically select the base station with the highest received signal power. This paper proposes a Self-Optimizing Cell Range Expansion Scheme based on a cognitive learning approach for a mobile heterogeneous network. System level simulations show the effectiveness of this approach in dynamically expanding the Pico cell coverage according to traffic conditions, balancing traffic load and improving the network performance.

**Keywords**— Load balance; mobile networks; heterogeneous networks; cognitive networks

## I. INTRODUÇÃO

A evolução das redes móveis tem sido caracterizada pela constante introdução de novas tecnologias, dispositivos e arquiteturas de rede, com o objetivo de acompanhar o aumento no número de usuários (*User Equipments*, UE) e as mudanças nos padrões de consumo de banda pelas aplicações emergentes. Uma das principais abordagens utilizadas para melhorar a capacidade da rede é o aumento ou densificação do número de elementos na rede de acesso via rádio [1] e a implementação de cenários heterogêneos (Macro e Pico eNB, *Relay Nodes*, etc.). No entanto, o uso dessas técnicas traz consigo o aumento nos custos e a complexidade nas tarefas de operação e gerenciamento do sistema. Para enfrentar estes desafios, o NGMN (*Next Generation Mobile Networks*) introduziu o conceito de redes auto-organizáveis (*Self Organizing Networks*, SON) como uma das principais abordagens para otimizar o desempenho

da rede enquanto se procura reduzir os seus custos operacionais. A auto-organização é um conceito inspirado na natureza, onde certos sistemas biológicos apresentam um comportamento organizado, observando, aprendendo e se adaptando ao meio a fim de atingir objetivos desejados. Nos últimos anos, o 3GPP (*3rd Generation Partnership Project*) tem identificado a necessidade de introduzir as ditas funcionalidades para os seus padrões de LTE (*Long Term Evolution*) e *LTE-Advanced*, procurando fornecer à rede inteligência e adaptabilidade para automatizar funcionalidades relacionadas com a configuração, otimização e recuperação de falhas. Durante a última década, o uso de SON nas redes de acesso via rádio tem atraído o interesse da indústria de comunicações e da comunidade acadêmica, levando a várias contribuições em temas relacionados com a otimização de cobertura, capacidade, balanceamento de carga e mitigação de interferências. Em [4], Aliu e colaboradores realizaram um estudo detalhado sobre o estado da arte de SON nas redes móveis celulares, comparando os pontos positivos e as fraquezas das soluções atuais e destacando os principais desafios de pesquisa nesta temática. O presente artigo foca na relevância da funcionalidade de balanceamento de carga para lidar com a distribuição do tráfego entre macro e pico células. No contexto das redes heterogêneas, o principal problema é que devido às altas potências de transmissão das Macro eNBs (MeNB), em alguns casos não é possível transladar uma quantidade suficiente de UEs da MeNB para as Pico eNBs (PeNB). O motivo é que nos métodos tradicionais de seleção de célula, os UEs comparam as potências dos sinais de referência recebidas, escolhendo a estação base da qual é percebido o maior nível. Diversas contribuições sobre o balanceamento de carga em redes móveis têm sido apresentadas na literatura [6], sendo essas soluções resultantes de problemas definidos como: alocação de recursos; controle de acesso (*Access Control*, AC); controle dos processos de *Handover*; ou ainda de ajuste de potências de transmissão das estações rádio-base. Em [5] Simão et al. propuseram um mecanismo de balanceamento de carga em um cenário móvel heterogêneo a partir de uma estratégia de controle de acesso. O algoritmo desenvolvido naquele artigo foi baseado em uma solução de programação inteira, procurando achar a distribuição ótima dos fluxos de tráfego na rede de acesso. De forma geral a maioria das soluções desenvolvidas na atualidade tem sido projetadas a partir da geração de modelos estáticos obtidos a partir de processos analíticos ou simulações.

O presente artigo propõe um esquema de expansão dinâmico de célula baseado em uma abordagem de aprendizagem automática. A ideia é dar à rede a capacidade de aprender como ajustar dinamicamente a cobertura da pico eNB

Cesar A. Sierra Franco e José Roberto B. de Marca, Centro de estudos em telecomunicações da PUC-Rio. Rio de Janeiro, Brasil. E-mails: {cesarsierra,jrbdemarca}@cetuc.puc-rio.br

conforme as condições de tráfego na célula, em busca de um melhor desempenho da mesma. O restante do artigo está organizado da seguinte forma: a Seção II descreve os conceitos relacionados com o balanceamento de carga em redes heterogêneas, a técnica CRE e os detalhes do cenário em estudo. A Seção III apresenta a arquitetura e o algoritmo de aprendizado utilizado. Na Seção IV são avaliados os resultados obtidos, enquanto a Seção V finaliza o artigo, apresentando as conclusões e indicações sobre possíveis trabalhos futuros.

## II. BALANCEAMENTO DE CARGA EM REDES MÓVEIS HETEROGÊNEAS

O uso de redes heterogêneas através da adição de uma camada de estações base de baixa potência é útil em casos de otimização de cobertura e capacidade, além de melhoras na qualidade da comunicação dos usuários da borda da célula [2]. Nos métodos tradicionais de seleção de célula, os UE comparam as potências dos sinais de referência recebidas (*Reference Signal Received Power*, RSRP), escolhendo desta forma a estação base da qual é percebido o maior nível. Para solucionar este problema pode-se utilizar uma técnica conhecida como “*cell range expansion*” (CRE) [3], que permite aumentar a cobertura das PeNB sem necessidade de incrementar suas potências de transmissão. Isto é conseguido ao adicionar um *offset* ao RSRP durante o processo de seleção de célula nos terminais. Desta forma, os UE não vão se ligar com a célula com maior RSRP, senão com aquela que tiver maior RSRP mais seu *offset*. Através da implementação do CRE será possível associar um maior número de UE com as PeNB melhorando o balanceamento de carga. No entanto, este benefício vem com o custo associado ao incremento dos níveis de interferência para os usuários agora localizados na borda das PeNB.

### A. Cenário de estudo e simulação

Neste trabalho foi considerado um cenário LTE heterogêneo com 7 eNBs e distâncias entre eles de 500 metros. Cada site é composto por três setores com fator de reuso 1 e uma pico célula omnidirecional na borda da célula. No cenário implementado, o centro da célula e as pico-células representam a área de interesse enquanto o resto de eNBs geram interferência aos UEs avaliados. A Tabela 1 mostra os principais parâmetros e modelos de propagação utilizados na rede de acesso (*Evolved Universal Terrestrial Radio Access Network*, E-UTRAN). Estes valores foram baseados nos parâmetros definidos pela 3GPP para a simulação de sistemas heterogêneos, especificamente para o modelo 1 (*Urban Macro cell + outdoor Hotzone cells*, tabela A.2.1.2-3 em [7]).

## III. APRENDIZAGEM E SELEÇÃO DINÂMICA DO CRE *offset*

Nesta pesquisa, foi utilizada uma abordagem cognitiva para melhorar as funcionalidades SON relacionadas com o balanceamento de carga através da técnica CRE. O cenário proposto foi simulado através do simulador de código aberto NS-3 [11], o qual foi integrado com a ferramenta Matlab onde foram programados os algoritmos de aprendizado automático. A seguir, serão apresentados os detalhes referentes

TABELA I  
PARÂMETROS DE SIMULAÇÃO DO E-UTRAN

Parâmetro	Valor
frequência de portadora	2 Ghz
largura de banda	5 Mhz
potência de transmissão - MeNB	46 dBm
potência de transmissão - PeNB	30 dBm
perdas de propagação - MeNB	$128.1 + 37.6\log_{10}(R)$
perdas de propagação - PeNB	$140.7 + 36.7\log_{10}(R)$
perdas de penetração	20 dB

à metodologia utilizada e o algoritmo empregado durante a fase de treinamento da rede móvel.

### A. Arquitetura cognitiva e cooperativa em redes móveis auto-organizáveis

A Figura 1 ilustra a arquitetura genérica proposta para implementar os mecanismos de auto-organização através de processos cognitivos. A ideia principal é dotar os diversos dispositivos de redes com capacidades de aprendizado a partir de técnicas de inteligência computacional, permitindo-lhes construir modelos dinâmicos sobre os quais as funcionalidades SON são implementadas. Com esta abordagem, procura-se que o sistema aperfeiçoe no tempo seu desempenho através da experiência obtida a partir da interação com o meio (realização de medidas de desempenho locais e obtenção de informações cooperativas de nós vizinhos), construindo modelos mais flexíveis a mudanças no ambiente quando comparados com os modelos obtidos a partir de abordagens convencionais.

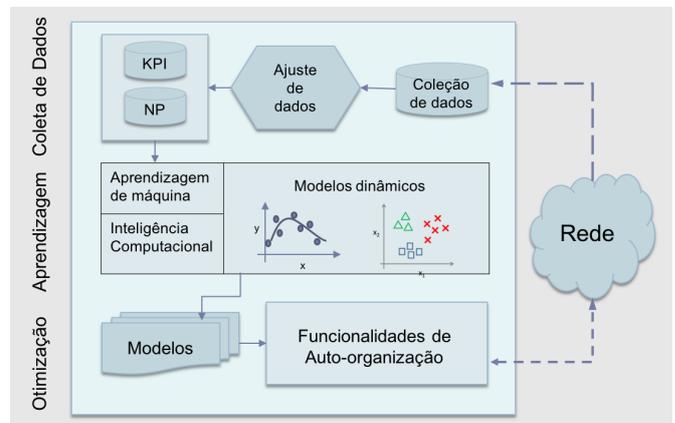


Fig. 1. Arquitetura cognitiva e cooperativa proposta para redes móveis auto-organizáveis

### B. Etapa de aprendizagem

Neste trabalho, a fase de aprendizado esteve baseada em uma técnica de aprendizagem por reforço (*Reinforcement Learning*, RL). O RL representa o problema encontrado por um agente autônomo que deve aprender como agir em um ambiente dinâmico através de interações de tentativa e erro [8]. Especificamente para o problema abordado, cada Pico-eNB foi modelada como um agente que percebe o estado do ambiente e interage com ele ao modificar seu próprio CRE

*offset* em busca de maximizar o desempenho da célula (Figura 2). Assim, o processo de decisão foi modelado como um processo de Markov (PDM) onde o agente percebe a cada instante  $t$  seu estado atual  $s$  (descrito na Seção III.C), para posteriormente escolher uma ação  $a$  do conjunto de ações possíveis  $A = \{a_1, a_2, \dots, a_j\}$ . Uma vez executada a dita ação, o agente mudará para um estado futuro  $s'$ , recebendo uma recompensa numérica  $r$  do ambiente.

O grande desafio nos problemas de aprendizado por reforço é que o agente não conta com o modelo para resolver o PDM (definido pelas funções de transição e recompensa), portanto, este deverá interagir com ambiente de forma direta para obter informações suficientes que permitam a obtenção das políticas ótimas que maximizem a utilidade cumulativa esperada  $R_k$ , definida por,

$$R_k = r_k + \gamma r_{k+1} + \gamma^2 r_{k+2} + \dots = \sum_{t=0}^{\infty} \gamma^t r_{k+t} \quad (1)$$

onde  $r_k$  corresponde à recompensa no estado inicial  $k$  do agente,  $r_{t+k}$  denota a recompensa numérica instantânea obtida como resultado de realizar uma determinada ação no instante  $t+k$  e  $\gamma \in [0, 1]$  é o fator de desconto que determina a importância relativa das recompensas futuras.

Neste trabalho, a solução de aprendizado por reforço foi realizada através de um algoritmo baseado em diferenças temporais conhecido como *Q-Learning*, o que permitiu ao agente construir de forma incremental (a cada transição  $s, a, r, s'$ ), a função  $Q(s, a)$  que indica a utilidade futura esperada que pode ser obtida desde um estado  $s$  quando se realiza a ação  $a$ . Usualmente, a função  $Q(s, a)$  pode ser representada por uma tabela que relaciona a cada estado  $s$  o resultado de aplicar cada possível ação  $a$ . No entanto, como o cenário de estudo neste trabalho apresenta um conjunto de variáveis contínuas, foi considerada uma adaptação do algoritmo *Q-Learning* conhecido como *fuzzy Q-Learning* (FQL) [9]. Esta mistura entre a lógica *fuzzy* e o *Q-learning*, nos permitiu discretizar e segmentar as variáveis contínuas utilizadas para a representação dos estados e ações. A seguir será apresentado o algoritmo específico usado para realizar o processo de aprendizado do CRE *offset*, além dos detalhes de implementação.

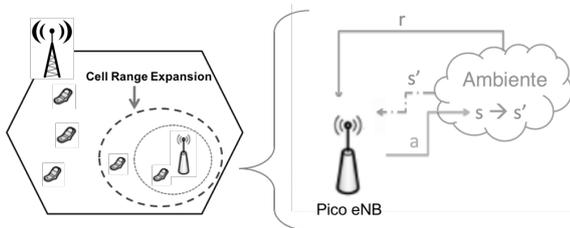


Fig. 2. PeNB como agente de aprendizagem

### C. Algoritmo fuzzy Q-Learning

Tal como mencionado na seção anterior, o algoritmo *Q-Learning* se baseia na estimação dos valores do par estado-ação da função  $Q(s, a)$ , que representam a recompensa

esperada quando realizada a ação  $a$  no estado  $s$ . O FQL é uma abordagem que permite expandir o *Q-Learning* para o ambiente da lógica *fuzzy*, onde as variáveis contínuas utilizadas para representar os estados e ações são transformadas em um número finito de variáveis linguísticas, que se encontram representadas pelas chamadas funções de pertinência *fuzzy*. Por outro lado, a função  $Q(s, a)$  será representada através de um conjunto de regras da forma:

**se**  $\vec{x}$  é  $\vec{s}_i$  **então**  
 ação =  $a_1$  com  $q(s_i, a_1)$  **ou** ação =  $a_2$  com  $q(s_i, a_2)$   
**ou ... ou** ação =  $a_j$  com  $q(s_i, a_j)$   
**fim se**

Onde  $a_j$  é a  $j$ -ésima ação possível a escolher dado que o estado percebido  $\vec{x}$  se corresponde com o estado  $\vec{s}_i$ , e  $q(s_i, a_j)$  é o valor da função  $Q$  para o par estado-ação  $(s_i, a_j)$ . O algoritmo 1 ilustra passo a passo a implementação do algoritmo FQL. Note-se que há tantos estados como regras *fuzzy* (i.e.,  $N$ ) e  $J$  ações possíveis por regra.

### Algorithm 1 Algoritmo fuzzy Q-Learning

- 1: Inicializar a tabela para armazenar os valores-q:  
 $q[i, j] = 0$  para  $1 \leq i \leq N$  e  $1 \leq j \leq J$
- 2: Observar o valor da variável de estado de entrada  $\vec{x}$ .
- 3: para cada regra ativada selecionar uma ação (exploração).  
 $a_i = \arg \max_k q(i, k)$  com probabilidade  $(1 - \varepsilon)$   
 $a_i = \text{random} \{a_k, k = 1, 2, \dots, J\}$  com probabilidade  $\varepsilon$
- 4: Calcular a ação global inferida pelo sistema de controle fuzzy  
 $a = \sum_{i=1}^N \alpha_i(s) \times a_i$   
 onde  $\alpha_i(s)$  representa o grau de ativação da regra  $i$
- 5: Aproximar a função  $Q$  a partir dos valores-q atuais e o grau de ativação de cada regra:  
 $Q(s(t), a) = \sum_{i=1}^N \alpha_i(s) \times q[i, a_i]$   
 Onde  $Q(s(t), a)$  é o valor da função  $Q$  para o estado  $s(t)$  na iteração  $t$  e para a ação  $a$ .
- 6: Realizar a ação indicada pelo controlador fuzzy e deixar o sistema evoluir ao próximo estado  $s(t+1)$
- 7: Observar o sinal de reforço (recompensa)  $r(t+1)$  e calcular o valor do novo estado definido por:  
 $V_t(s(t+1)) = \sum_{i=1}^N \alpha_i(s(t+1)) \cdot \max_k (q[i, a_k])$
- 8: Calcular a diferença entre  $Q(s(t), a)$  e  $Q(s(t+1), a)$ :  
 $\Delta Q = r(t+1) + \gamma V_t(s(t+1)) - Q(s(t), a)$   
 onde  $\gamma$  representa o fator de desconto
- 9: Atualizar a tabela de valores-q através de diferenças temporais:  
 $q[i, a_i] = q[i, a_i] + \lambda \cdot \Delta Q \cdot \alpha_i(s(t))$   
 onde  $\lambda$  representa a taxa de aprendizagem.
- 10: Repetir o processo para o estado atual desde o passo 2 até achar convergência.

### D. Definição de estados, ações e recompensas

A Figura 3 ilustra o sistema de inferência *fuzzy* que vai determinar o processo de aprendizado e que posteriormente

será o responsável por decidir o valor do CRE *offset* a utilizar conforme as condições de tráfego na célula. A seguir serão definidos os detalhes da implementação do algoritmo FQL, definindo os conjuntos de estado, ações e função de recompensa para os agentes.

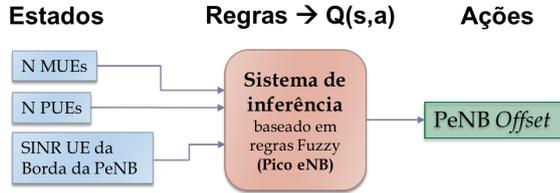


Fig. 3. Sistema de inferência fuzzy para a determinação do valor do CRE *offset* adaptativo

1) **Estados:** O estado percebido pela PeNB é definido pela tripla,  $S = [NMUE, NPUE, 5SINR]$ , onde  $NMUE$  e  $NPUE$  representam a carga da Macro e da Pico célula respectivamente em termos de número de usuários e  $5SINR$  faz referência ao Percentil 5 (5th Percentile) do SINR dos usuários localizados na pico célula, valor comumente utilizado para determinar o desempenho dos usuários da borda de uma célula. Nas Figuras 4.a e 4.b se ilustra a definição dos termos e valores escolhidos para cada variável *fuzzy* que compõe a percepção do entorno (estados) por parte da PeNB.

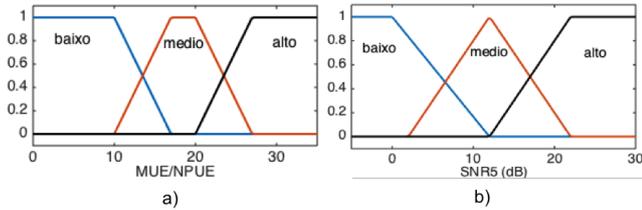


Fig. 4. Funções de pertinência para as variáveis de estado a) número de UEs na Macro/Pico eNB e b) 5SINR

2) **Ações:** O conjunto de ações possíveis  $A$  consiste no valor escolhido para o parâmetro CRE *offset* utilizado pela PeNB. Na Figura 5 pode-se observar a definição das variáveis linguísticas *fuzzy* definidas para designar o dito *offset*.

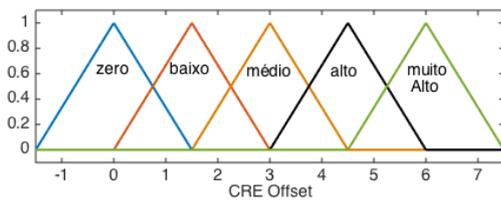


Fig. 5. Função de pertinência para a variável de saída: CRE *offset*

3) **Função de recompensa instantânea:** A recompensa no sistema FQL proposto é dado pela vazão média instantânea ( $v^t$ ) por usuário na célula. Definindo  $U$  como o número total de usuários com uma comunicação ativa na célula em um instante  $t$ ,  $v^t$  pode ser expressa como:

$$v^t = \sum_{u \in U} \frac{v_u^t}{U} \quad (2)$$

Onde  $v_u^t$  indica a vazão instantânea do usuário  $u$ .

## IV. AVALIAÇÃO DE RESULTADOS

Nesta seção é apresentada uma avaliação dos resultados da técnica proposta para o balanceamento de carga em redes móveis LTE heterogêneas, especificamente para o cenário descrito na Seção II-A. Para isto, foram realizadas simulações ao nível de sistema, empregando a ferramenta NS-3, para comparar e avaliar o desempenho da vazão média por UE na célula. A técnica proposta, baseada em processos cognitivos e de aprendizagem (FQL-*offset*), foi comparada com o esquema de seleção de célula baseado no máximo RSRP (MaxRSRP) e com a técnica de referência ABC (*Adaptive Bias Configuration*) [10] desenvolvida por Tiang e colaboradores. Esta última tem como objetivo modificar dinamicamente o CRE *offset* através de um algoritmo analítico e iterativo a partir de realimentações de desempenho do sistema, resultando num desempenho ótimo dos UEs na célula. Para cada cenário, as condições de tráfego em cada setor foram variadas alterando o número de UEs que recebiam uma transmissão UDP no enlace descendente de taxa constante de 512 kbps.

### A. Resultados da aprendizagem

Na Figura 6 pode-se observar o gráfico da superfície que representa a relação resultado do aprendizado pela PeNB, entre a carga da MeNB (em termos do número de UEs ativos) e a razão sinal-ruído experimentada pelos usuários da borda da PeNB em função do valor do CRE *offset* que produz o melhor desempenho na célula. Neste caso de aprendizado, observa-se como o PeNB identificou a possibilidade de aumentar o valor do seu CRE *offset* quando a carga do tráfego na MeNB alcançar valores altos sempre que a razão Sinal-Ruído (SINR) de seus UEs na borda atinjam valores acima de 20 dB. Por outro lado, quando o SINR dos UEs da borda da PeNB experimentem valores abaixo de 0 dB, o CRE *offset* deve ser mantido no menor valor para não prejudicar o desempenho da célula. De forma similar, a Figura 7 ilustra o gráfico de superfície (aprendido pela PeNB) para as variáveis SUE, 5SINR e CRE *offset* que produzem o melhor desempenho na célula.

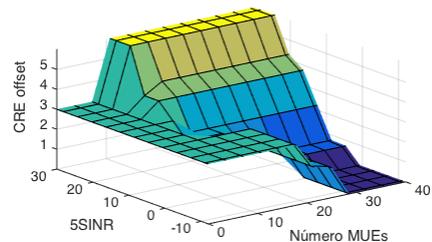


Fig. 6. Relação resultado do aprendizado, entre as variáveis MUE, 5SINR e CRE *offset* que produzem o melhor desempenho na célula (componente cooperativa)

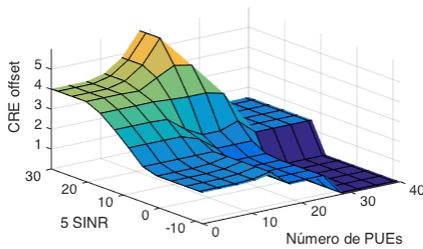


Fig. 7. Relação resultado do aprendizado, entre as variáveis SUE, 5SINR e CRE *offset* que produzem o melhor desempenho na célula (componente local)

### B. Avaliação de desempenho

A Figura 8 ilustra a relação entre a vazão média por UE e o número de UEs ativos dentro da célula (Macro + Pico), enquanto a Figura 9 ilustra a função cumulativa de probabilidade (*Cumulative Distribution Function*, CDF) obtida especificamente para para o cenário com 40 UEs ativos na célula. Em ambas figuras pode-se observar a efetividade da técnica proposta para aprender de forma autônoma os valores do CRE *offset* que melhoraram o desempenho dos UEs conforme as condições de tráfego da mesma. A seleção dinâmica do *offset* garante esse melhor desempenho já que uma maior quantidade de UEs consegue ser transferida para a PeNB (de forma controlada). Também pode-se observar (Figuras 8 e 9) que a técnica desenvolvida conseguiu igualar o desempenho ótimo conseguido pela técnica ABC usada como referência, porém com melhores resultados em termos de tempo de convergência, tal como ilustra a Figura 10.

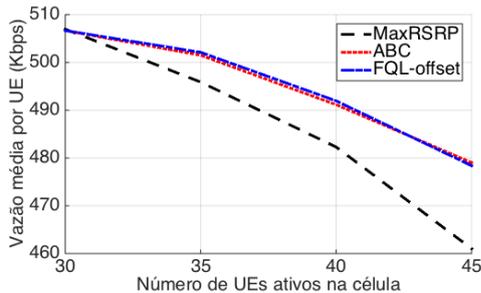


Fig. 8. Vazão média por UE Vs. número de UEs ativos na célula

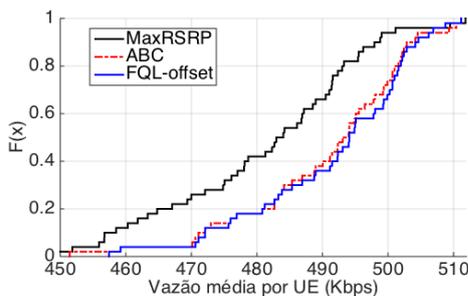


Fig. 9. CDF empírica para a vazão média por UE (40 UEs)

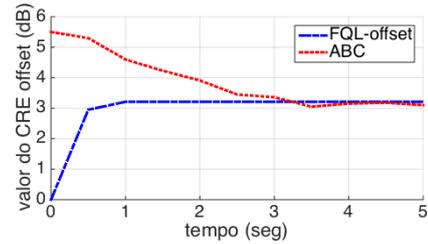


Fig. 10. Variação do valor do CRE *offset* no tempo

## V. CONCLUSÕES

Neste artigo foi apresentada uma técnica de balanceamento de carga para redes móveis heterogêneas, através de uma abordagem cognitiva e aprendizado automático. Os resultados das simulações mostraram a eficiência do mecanismo implementado ao aprender e posteriormente realizar o ajuste adequado do *offset* utilizado para modificar virtualmente a cobertura das pico células conforme as condições de tráfego observadas. Desta forma foram obtidas melhores condições de balanceamento da carga, reduzindo o congestionamento nas Macro eNB e consequentemente melhorando a vazão média das UEs. Como um próximo passo deste trabalho, pretende-se considerar cenários que incluam a transmissão de tráfego misto (i.e. voz, dados, vídeo) com restrições de qualidade de serviço diferenciadas e a utilização de esquemas de coordenação de interferências.

## AGRADECIMENTOS

Agradecemos à Faperj pelo financiamento desta pesquisa.

## REFERÊNCIAS

- [1] N. Bhushan, J. Li, D. Malladi, R. Gilmore, D. Brenner, A. Damnjanovic, R. T. Sukhvasi, C. Patel, and S. Geirhofer, "Network densification: The dominant theme for wireless evolution into 5G," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 82-89, 2014.
- [2] C. A. S. Franco and J. R. B. de Marca, "An open-source simulation of QoE for video applications in relay enhanced LTE-A networks," in 2014 11th International Symposium on Wireless Communications Systems (ISWCS), 2014, pp. 406-410.
- [3] Qualcomm. "Range expansion for efficient support of heterogeneous networks". TSG-RAN WG1, 2008.
- [4] O. G. Aliu, A. Imran, M. A. Imran, and B. Evans, "A Survey of Self Organisation in Future Cellular Networks", *IEEE Communications Surveys*, vol. 15, no. 1, pp. 336-361, 2013
- [5] S. Eduardo, A. Rodrigues, A. Mihovska, and N. R. Prasad, "Cell load balancing in heterogeneous scenarios: A 3GPP LTE case study", in *Wireless VITAE*, 2013.
- [6] A. Sharma, A. Roy, S. Ghosal, R. Chaki, and U. Bhattacharya, "Load balancing in Cellular Network: A review," in *ICCCNT*, 2012.
- [7] 3GPP, "TR 36.814: Further advancements for E-UTRA physical layer aspects (Release 9)," 2010.
- [8] L. P. Kaelbling, M. L. Littman, and A. W. Moore, "Reinforcement Learning: A Survey", *Journal of artificial intelligence research*, pp. 237-285, 1996.
- [9] P. Y. Glorionec and L. Jouffe, "Fuzzy Q-learning," *Proceedings of the Sixth IEEE International Conference on Fuzzy Systems, Barcelona, 1997*, pp. 659-662 vol.2.
- [10] P. Tian, H. Tian, J. Zhu, L. Chen, and X. She, "An adaptive bias configuration strategy for range extension in LTE-advanced heterogeneous networks," *IET International Conference on Communication Technology and Application (ICCTA 2011)*, pp. 336-340, 2011
- [11] N. Baldo, M. Miozzo, M. Requena, and J. N. Guerrero, "An open source product-oriented LTE network simulator based on ns-3," in *Proc. of MSWIM*, 2011.