

Análise da Alocação de Recursos em *Smart Cities* e *Fog Computing* abordando Múltiplos Períodos e Serviços

Mayron C. O. Moreira^{1,2}, Samuel M. A. Araújo³ e Geraldo R. Mateus²

¹Departamento de Ciência da Computação
Universidade Federal de Lavras (UFLA) – Lavras, MG - Brasil

²Departamento de Ciência da Computação
Universidade Federal de Minas Gerais (UFMG) – Belo Horizonte, MG - Brasil

³Departamento de Tecnologia em Engenharia Civil, Computação,
Automação, Telemática e Humanidades
Universidade Federal de São João Del-Rei (UFSJ) – Ouro Branco, MG - Brasil

mayron.moreira@ufla.br, sabreu@ufs.j.edu.br, mateus@dcc.ufmg.br

Abstract. *The Internet of Things (IoT) and Fog Computing (FC) technologies, when integrated, enable the development of smart urban systems and efficient processing of large volumes of data. This study presents a mathematical model for integrating these technologies and allocating resources to serve multiple network services in Smart Cities environments. In this context, multi-period service provision is justified in planning and scheduling the response to IoT sensor demands, minimizing fog and cloud nodes activation costs. The computational results show the limitations of an exact approach to solve the problem, and show future avenues concerning solution methods improvement based on heuristics.*

Resumo. *As tecnologias provenientes de Internet das Coisas (IoT) e Fog Computing (FC), se integradas, possibilitam o desenvolvimento de sistemas urbanos inteligentes e o processamento eficaz de grandes volumes de dados. Este trabalho apresenta um modelo matemático para a integração dessas tecnologias e a alocação de recursos para atender múltiplos serviços de rede em ambientes de Cidades Inteligentes. Neste contexto, o atendimento em múltiplos períodos se justifica no planejamento e programação do atendimento às demandas dos sensores de IoT, minimizando os custos de ativação de nós das camadas fog e cloud. Os resultados computacionais mostram as limitações de uma abordagem exata para a resolução do problema, e abrem discussões sobre o aprimoramento de métodos de solução baseados em heurísticas.*

1. Introdução

A IoT é amplamente utilizada em contextos tais como Cidades Inteligentes, Agricultura de Precisão e Planejamento da Logística [Zhou et al. 2020, Queiroz et al. 2022]. Trata-se de uma rede de dispositivos que, incorporados a *hardware* e *software*, conectam e trocam informações com outros dispositivos pela internet. Seu crescimento é impulsionado pelos avanços nas redes, que visam atender à mobilidade dos usuários e objetos. Para satisfazer os requisitos de dispositivos de IoT e aplicativos em tempo real, sensíveis à latência, surge

um paradigma de computação adicional à *cloud computing*, denominado *fog computing* [Bonomi et al. 2012, Srirama 2024]. Os nós da camada *fog* possuem, em geral, poder de processamento e memória inferiores aos nós da *cloud*. No entanto, seu tempo de resposta às demandas de dispositivos é inferior ao tempo gasto na *cloud*, o que impulsiona sua aplicabilidade em ambientes cada vez mais conectados. Este trabalho aborda o problema de alocação de recursos em redes baseadas em tecnologias *fog* para atender demandas provenientes de sensores de IoT em um contexto de *Smart Cities*, com múltiplos serviços e múltiplos períodos.

Considera-se que as demandas que surgem em um determinado período podem ser atendidas ou não pelos servidores da rede, não havendo a possibilidade da mesma ser atendida em algum período subsequente. As contribuições deste estudo são: (i) definição formal e proposição de um modelo linear inteiro para o problema de localização e alocação de recursos em redes hierarquicamente divididas em sensores, camada *fog* e camada *cloud*; (ii) estudo preliminar do potencial que um método exato possui para a resolução do problema proposto. Na sequência, a Seção 2 apresenta o modelo matemático proposto. Resultados preliminares são discutidos na Seção 3. Por fim, a Seção 4 apresenta as principais conclusões do estudo e aponta possibilidades de trabalhos futuros.

2. Modelo matemático

Seja S o conjunto de sensores, F o conjunto de nós *fog* e C o conjunto de nós *cloud*. Considera-se que os sensores podem demandar tipos serviços pertencentes ao conjunto R , em um horizonte de planejamento de T períodos. Define-se $F^s \subseteq F$ como o conjunto de nós *fog* que cobrem o sensor s . Seja $F^f \subseteq F$ o conjunto de nós *fog* conectados ao nó *fog* f , e $C^f \subseteq C$ como o conjunto de nós *cloud* que podem atender o nó *fog* $f \in F$. A largura de banda (Mb/s) demandada pelo sensor s , referente ao tipo de requisição r é denotada por bw_s^r , enquanto bw_{sf} indica a largura de banda (em Mb/s) disponível para a comunicação *wireless* do sensor s com o nó *fog* $f \in F^s$. Denota-se por bw_{fi} a largura de banda (em Mb/s) disponível para a comunicação cabeada do nó *fog* f com o nó $i \in F \cup C$. As capacidades de processamento e memória de nós *fog* e nós *cloud* são, respectivamente: \mathcal{D}_i (CPUs) e \mathcal{D}_i^m (Mb), $\forall i \in F \cup C$. Estabelece-se que o custo, em US\$, de utilização de um nó $i \in F \cup C$ é dado por α_i .

A modelagem considera que cada sensor s envia um único serviço de um tipo. Em cenários os quais um sensor possui múltiplas requisições de vários tipos, basta fragmentar cada requisição de cada tipo em uma única demanda, e abstraí-las com se fossem vários sensores. Assim, seja \mathcal{S} o conjunto de nós sensores artificiais, e $\mathcal{S}_t \subseteq \mathcal{S}$ o conjunto de nós artificiais a serem atendidos no período $t \in T$. Define-se $\text{dem}(s)$, $\text{mem}(s)$ e $\text{bw}(s)$ como funções que retornam os parâmetros relativos à demanda de processamento, demanda de memória e largura de banda requerida pelo sensor artificial $s \in \mathcal{S}$, respectivamente. O conjunto $\mathcal{F}^s \subseteq F$ representa os nós *fog* que cobrem o sensor $s \in \mathcal{S}$. Os parâmetros relativos à largura de banda dos enlaces da camada IoT e *fog* são redefinidos por meio de $\text{bw}(s, f)$, em que $s \in \mathcal{S}$ e $f \in \mathcal{F}^s$. A demanda do sensor $s \in \mathcal{S}$ pode ser atendida por um nó *fog*, ou por um nó *cloud*. Toda designação deve passar pela camada *fog*. Considere a variável $x_{sf}^t \in \{0, 1\}$ igual a 1 se no período $t \in T$, o sensor $s \in \mathcal{S}_t$ envia requisição para o nó *fog* $f \in \mathcal{F}^s$. A variável de fluxo $y_{sfi}^t \in \{0, 1\}$ assume o valor 1 se no período $t \in T$, o fluxo do nó $s \in \mathcal{S}_t$ é enviado para o nó *fog* $f \in F$ que envia para outro nó $i \in F^f \cup C^f$. Note que se $y_{sff}^t = 1$, então o nó *fog* $f \in F$ processa a demanda

do sensor $s \in \mathcal{S}_t$. A variável real $w_s^t \in \{0, 1\}$ é igual a 1 se o atendimento à demanda do sensor $s \in \mathcal{S}_t$ no período $t \in T$ é realizado com êxito. Por fim, seja $z_i \in \{0, 1\}$ igual a 1 se o nó $i \in F \cup C$ está ativo ou não.

$$\text{minimize } \sum_{i \in F \cup C} \alpha_i z_i + \mathcal{M} \sum_{t \in T} \sum_{s \in \mathcal{S}_t} (1 - w_s^t) \quad (1)$$

sujeito a:

$$\sum_{f \in \mathcal{F}^s} x_{sf}^t = w_s^t, \quad \forall t \in T, \forall s \in \mathcal{S}_t. \quad (2)$$

$$\sum_{f \in F} y_{sff}^t + \sum_{f \in F} \sum_{c \in C^f} y_{sfc}^t = w_s^t, \quad \forall t \in T, \forall s \in \mathcal{S}_t. \quad (3)$$

$$x_{sf}^t + \sum_{j \in F | f \in F^j} y_{s fj}^t = \sum_{j \in F^f} y_{s fj}^t + y_{sff}^t + y_{sfc}^t, \quad \forall t \in T, \forall s \in \mathcal{S}_t, \forall f \in F, \forall c \in C^f. \quad (4)$$

$$x_{sf}^t \leq z_f, \quad \forall t \in T, \forall s \in \mathcal{S}_t, \forall f \in \mathcal{F}^s. \quad (5)$$

$$y_{s fj}^t \leq z_f, \quad \forall t \in T, \forall s \in \mathcal{S}_t, \forall f \in F, \forall j \in F^f. \quad (6)$$

$$y_{sfc}^t \leq z_c, \quad \forall t \in T, \forall s \in \mathcal{S}_t, \forall f \in F, \forall c \in C^f. \quad (7)$$

$$\sum_{s \in \mathcal{S}_t} \text{proc}(s) y_{sff}^t \leq \mathcal{D}_f z_f, \quad \forall t \in T, \forall f \in F. \quad (8)$$

$$\sum_{s \in \mathcal{S}_t} \text{mem}(s) y_{sff}^t \leq \mathcal{D}_f^m z_f, \quad \forall t \in T, \forall f \in F. \quad (9)$$

$$\sum_{f \in F^c} \sum_{s \in \mathcal{S}_t} \text{proc}(s) y_{sfc}^t \leq \mathcal{D}_c z_c, \quad \forall t \in T, \forall c \in C. \quad (10)$$

$$\sum_{f \in F^c} \sum_{s \in \mathcal{S}_t} \text{mem}(s) y_{sfc}^t \leq \mathcal{D}_c^m z_c, \quad \forall t \in T, \forall c \in C. \quad (11)$$

$$\text{bw}(s) x_{sf}^t \leq \text{bw}(s, f), \quad \forall t \in T, \forall s \in \mathcal{S}_t, \forall f \in \mathcal{F}^s. \quad (12)$$

$$\sum_{s \in \mathcal{S}_t} \text{bw}(s) y_{sfi}^t \leq \text{bw}_{fi}, \quad \forall t \in T, \forall f \in F, \forall i \in F^f \cup C^f. \quad (13)$$

$$y_{s fj}^t + y_{s jf}^t \leq 1, \quad \forall t \in T, \forall s \in \mathcal{S}_t, \forall f \in F, \forall j \in F^f. \quad (14)$$

$$x_{sf}^t \in \{0, 1\}, \quad \forall t \in T, \forall s \in \mathcal{S}_t, \forall f \in \mathcal{F}^s. \quad (15)$$

$$y_{sfi}^t \in \{0, 1\}, \quad \forall t \in T, \forall s \in \mathcal{S}_t, \forall i \in F^f \cup C^f. \quad (16)$$

$$w_s^t \in [0, 1], \quad \forall t \in T, \forall s \in \mathcal{S}_t. \quad (17)$$

$$z_i \in \{0, 1\}, \quad \forall i \in F \cup C. \quad (18)$$

A função objetivo (1) minimiza o custo de instalação dos nós *fog* e *cloud*, além de maximizar o número de serviços realizadas através da penalização $\mathcal{M} = \max_{c \in C} \alpha_c$. As Restrições (2) e (3) estabelecem o atendimento de uma demanda do sensor s ao seu recebimento por um nó *fog* $f \in \mathcal{F}^s$ e seu processamento na camada *fog* ou *cloud*. As Restrições (4) estabelecem a conservação de fluxo. O processo de ativação de nós *fog* e *cloud* é definido pelas Restrições (5)-(7), enquanto as capacidades de processamento e memória da camada *fog* e *cloud* são garantidas pelas Restrições (8)-(11). A limitação do tráfego dada pela largura de banda dos enlaces é apresentada pelas Restrições (12)-(13).

As Restrições (14) indicam a quebra de simetria necessária ao roteamento entre nós *fog*. O domínio das variáveis de decisão é definido pelas Restrições (15)-(18).

3. Resultados computacionais

Esta seção apresenta experimentos preliminares para verificar a escalabilidade do modelo matemático ao considerar cenários de pequeno/médio porte, compostos por até 300 sensores. Três instâncias foram geradas a partir de adaptações dos dados inspirados em localizações da cidade de Antuérpia (Bélgica), obtidos em [Santos et al. 2021]. A topologia da rede estabelece que um sensor s pode enviar fluxo para um conjunto de nós *fog* (\mathcal{F}^s), um nó *fog* $f \in F$ pode receber fluxo de sensores e de outros nós *fog*, e um nó da *cloud* $c \in C$ pode receber fluxo de qualquer nó *fog* a ele conectado. O esquema para a geração de instâncias detalhado pode ser visto em [Moreira et al. 2024]. O modelo foi implementado utilizando a linguagem C++ e o resolvidor comercial Gurobi 11.0, com 1 *thread* e tempo limite de 300s. Os testes foram realizados em um computador Intel(R) Core(TM) i7-4770, 3.40 GHz, 8Gb RAM, sistema operacional Ubuntu 22.04.4 LTS.

As três instâncias geradas possuem serviços relacionados à captação de imagem, mensuração da qualidade do ar e mensuração de volume de lixo em lixeiras inteligentes. A quantidade de períodos considerados foi igual a 10, e os valores de alcançabilidade de sensores a nós *fog*, assim como a largura de banda para a comunicação de ambas as camadas, seguiram o padrão 4G. A simulação de conexão entre nós *fog* adotou valores de largura de banda entre 1 e 3 Gb/s, enquanto a conexão com a *cloud* adotou o valor de 50 Gb/s. Os serviços que compõem cada período foram gerados utilizando distribuição Poisson. A instância “*id0*” possui $|S| = 100$, $|F| = 10$ e $|C| = 1.157$. A solução ótima foi encontrada em 21s, com atendimento de todos os serviços demandados pelos sensores de cada período, com uma taxa de utilização de 2 de nós *fog* e 1 nós *cloud*. A instância “*id1*” possui $|S| = 200$, $|F| = 20$ e $|C| = 3.997$. Uma solução factível com *gap* de otimalidade igual a 3, 44% foi encontrada, utilizando 5 nós *fog* e 2 nós *cloud*. Considerando a instância “*id2*”, composta por $|S| = 300$, $|F| = 30$ e $|C| = 9.380$, o modelo não obteve sequer uma solução viável em 300s. Nesse caso, o Gurobi interrompeu sua execução pela demanda por memória exceder seu limite. Isso indica que para a resolução desse problema com uma maior quantidade de sensores e serviços, métodos exatos não são escolhas mais apropriadas.

4. Conclusões

Este estudo aborda o problema de alocação de recursos e localização de nós *fog* e nós *cloud* em redes de sensores com multi-serviços e multi-períodos. Um modelo matemático linear inteiro foi proposto, e testes computacionais preliminares foram realizados para analisar a escalabilidade da resolução do modelo em instâncias de pequeno/médio porte. Os experimentos indicaram a ineficácia de uma metodologia exata para a resolução das três instâncias compostas por 300 sensores. A baixa taxa de utilização de nós *fog* pode ser explicada pelo fato das redes geradas serem esparsas. Como trabalhos futuros, pretende-se: (i) propor heurísticas e metaheurísticas para a obtenção de soluções em cenários com redes mais densas; (ii) desenvolver algoritmos de aprendizado estatístico para que, com base no histórico das informações dos serviços atendidos em múltiplos períodos, o sistema possam prever a alocação apropriada de recursos para requisições dinâmicas em ambientes *online*; (iii) aprimorar o gerador de instâncias.

Referências

- Bonomi, F., Milito, R., Zhu, J., and Addepalli, S. (2012). Fog computing and its role in the internet of things. In *Proceedings of the first edition of the MCC workshop on Mobile cloud computing*, pages 13–15.
- Moreira, M. C. O., Araújo, S. M. A., and Mateus, G. R. (2024). Sensor, fog and cloud generator: new test instances based on the city of antwerp. <https://shorturl.at/tAKY9>. Accessed on March 22, 2024.
- Queiroz, T. A. d., Canali, C., Iori, M., and Lancellotti, R. (2022). *An Optimization View to the Design of Edge Computing Infrastructures for IoT Applications*, pages 1–30. Springer International Publishing, Cham.
- Santos, J., Wauters, T., Volckaert, B., and De Turck, F. (2021). Towards end-to-end resource provisioning in fog computing over low power wide area networks. *Journal of Network and Computer Applications*, 175:102915.
- Srirama, S. N. (2024). A decade of research in fog computing: Relevance, challenges, and future directions. *Software: Practice and Experience*, 54(1):3–23.
- Zhou, Y., Yu, F. R., Chen, J., and Kuo, Y. (2020). Cyber-physical-social systems: A state-of-the-art survey, challenges and opportunities. *IEEE Communications Surveys and Tutorials*, 22:389–425.