

Ontological Modeling and Clustering Techniques for Service Allocation

Subjects: Computer Science, Interdisciplinary Applications

Contributor: Marcelo Karanik, Iván Bernabé-Sánchez, Alberto Fernández

In complex distributed IoT-based applications, multiple and heterogeneous IoT devices are deployed in a given environment. Those devices typically act as information input providers (e.g., sensor networks) or actuators. Edge computing arose as a solution to reduce the high demand for data traffic between IoT devices and the cloud that processes them. Moreover, several notions have been introduced to shape the gap between the cloud and edge, like fog or mist computing.

Keywords: IoT ; service allocation ; clustering ; edge computing ; knowledge graphs ; ontologies

1. Introduction

In the last few years, we have witnessed the increasing development of the Internet of Things (IoT), where physical devices form interconnected systems. In complex distributed IoT-based applications, multiple and heterogeneous IoT devices are deployed in a given environment. Those devices typically act as information input providers (e.g., sensor networks) or actuators. Edge computing arose as a solution to reduce the high demand for data traffic between IoT devices and the cloud that processes them. Moreover, several notions have been introduced to shape the gap between the cloud and edge, like fog or mist computing ^[1]. One of the current research areas is application/service orchestration in the edge–cloud continuum ^[2], that is, deciding where to offload applications based on the computing characteristics of heterogeneous edge nodes, as well as requirements such as the network load, execution time and carbon emissions ^[3]. The aforementioned technologies are essential components on which smart cities are based. This is also the case for large smart areas suffering from harsh environments, with multiple IoT devices geographically distributed. In these environments, devices may suffer frequent contingencies (e.g., loss of connectivity, low battery autonomy, etc.) due to changing weather conditions or other unexpected events. While assuring the correct functioning of complex IoT systems in controlled environments (e.g., smart buildings) is not easy, doing so in large, hard and changing environments is a real challenge. Many developments working for smart cities are also useful for such complex smart areas. However, they require additional aspects to face the mentioned issues.

2. Ontological Modeling and Clustering Techniques for Service Allocation on the Edge

One of the goals of edge computing is to bring computing resources closer to devices. Edge computing ^[4] eliminates the need to use a cloud environment for extensive computations because edge computing provides computing resources, such as memory and processors, at the edge of the network (such as the base station ^[5]), so devices can use these local computational resources and not the remote resources located in the cloud. In edge computing, resource allocation mechanisms are used to assign computing resources to tasks or services located at the edge of the network. Switching processing tasks between computational resources is not an easy task because, if it is not performed correctly, the computational resources may be exhausted or the tasks may not be able to be performed. The environments formed by mobile devices, edge computing and cloud computing make up ecosystems of computational resources where software can be deployed somewhere in the system to be executed in the most suitable place. Resource allocation mechanisms calculate the best location based on preset objectives. In this direction, several works propose resource allocation mechanisms with different objectives, such as optimizing the energy consumption, bandwidth or computation. In ^{[6][7]}, allocation tasks are made to reduce energy consumption. Goudarzi et al. ^[8] propose a model to optimize aspects of the energy consumption and execution time in the distribution of tasks between IoT devices and servers located in the fog or the cloud ^[7], unlike ^[6], which proposes a strategy based on the allocation of physical resources to minimize the energy consumption and processing time of the overall system. In ^[9], another type of problem is considered in which task

allocation is performed taking into account the energy consumption, connectivity time and amount of data to be transferred.

Optimizing the allocation of resources according to aspects of the energy savings, bandwidth or amount of data used is an important aspect in smart cities. However, the optimization in the task allocation based on computing nodes' capability is crucial for the proper functioning of the services. In this line, several works propose solutions taking into account computational aspects. For example, Pan and Li ^[9] propose to consider the computing capability of the mobile device and then determine whether the task needs to be offloaded. In that case, the algorithm transfers the tasks to the edge computing servers with the highest capacity. In ^{[10][11][12]}, virtualization mechanisms are proposed to allocate resources to satisfy task requirements. Finally, other authors propose to use clustering mechanisms to manage resources according to the task priorities ^[13] and the overload of CPU, communications and I/O operations ^[14].

To explore the solutions mentioned above, various studies have introduced novel mechanisms for managing these tasks, such as those presented in works such as ^{[10][11]}. Additionally, some works propose adapting pre-existing algorithms and applying them to the specific domain of resource and task allocation, such as ^{[12][15][16][17]}.

The solutions described have been designed for specific domains, and their application to other domains is complicated by the fact that they have been designed for a specific context. This implies that similar concepts appear in different works but are represented differently. In addition, there is a gap between the concepts that people use and the data that systems interpret. To overcome this problem, using semantic descriptions facilitates the common definition of the elements between people and computers when handling information. Semantic representation mechanisms provide a common language or structure for modeling IoT devices and service data, irrespective of the format. Typically involving a graph structure ^[18], semantic representation enables the interpretation of data beyond textual information. This would imply reducing human intervention by reducing the rate of errors introduced and increasing the speed of allocation.

Works such as ^[19] have explored the application of ontologies in cloud environments. The mOSAIC ontology ^[20] is one of the most important examples. mOSAIC offers a description detailed of cloud computing resources, and it is focused on interoperability within cloud-based systems. However, this solution is not tailored to address the IoT devices, such as sensors, actuators and gateways. Some ontologies have been specifically developed to model those devices. For instance, the Semantic Sensor Network ontology (SOSA/SSN) ^[21] is designed to describe sensor and actuator networks, detailing their capabilities, characteristics of interest and observations. SOSA/SSN is used as the core for the creation of other ontologies. Another relevant ontology is the Smart Applications REference Ontology (SAREF) ^[22], which is specifically designed to model devices and their functions. SAREF is aligned with the oneM2M base ontology ^[23], which enables syntactic and semantic interoperability between devices and external systems. This strategic alignment enhances the overall effectiveness of semantic representation in the IoT domain.

References

1. Yousefpour, A.; Fung, C.; Nguyen, T.; Kadiyala, K.; Jalali, F.; Niakanlahiji, A.; Kong, J.; Jue, J.P. All one needs to know about fog computing and related edge computing paradigms: A complete survey. *J. Syst. Archit.* 2019, 98, 289–330.
2. Ullah, A.; Dagdeviren, H.; Ariyattu, R.C.; DesLauriers, J.; Kiss, T.; Bowden, J. Micado-edge: Towards an application-level orchestrator for the cloud-to-edge computing continuum. *J. Grid Comput.* 2021, 19, 47.
3. Kimovski, D.; Matha, R.; Hammer, J.; Mehran, N.; Hellwagner, H.; Prodan, R. Cloud, Fog, or Edge: Where to Compute? *IEEE Internet Comput.* 2021, 25, 30–36.
4. Cao, K.; Liu, Y.; Meng, G.; Sun, Q. An overview on edge computing research. *IEEE Access* 2020, 8, 85714–85728.
5. Araldo, A.; Stefano, A.D.; Stefano, A.D. Resource allocation for edge computing with multiple tenant configurations. In *Proceedings of the 35th Annual ACM Symposium on Applied Computing, Virtual*, 30 March–3 April 2020; pp. 1190–1199.
6. Goudarzi, M.; Wu, H.; Palaniswami, M.; Buyya, R. An Application Placement Technique for Concurrent IoT Applications in Edge and Fog Computing Environments. *IEEE Trans. Mob. Comput.* 2021, 20, 1298–1311.
7. Ning, Z.; Hu, X.; Chen, Z.; Zhou, M.; Hu, B.; Cheng, J.; Obaidat, M.S. A cooperative quality-aware service access system for social Internet of vehicles. *IEEE Internet Things J.* 2017, 5, 2506–2517.
8. Zhang, Y.; Zhao, L.; Liang, K.; Zheng, G.; Chen, K.C. Energy Efficiency and Delay Optimization of Virtual Slicing of Fog Radio Access Network. *IEEE Internet Things J.* 2023, 10, 2297–2313.

9. Pan, M.; Li, Z. Multi-user Computation Offloading Algorithm for Mobile Edge Computing. In Proceedings of the 2021 2nd International Conference on Electronics, Communications and Information Technology (CECIT), Sanya, China, 27–29 December 2021; pp. 771–776.
10. Deepika, T.; Rao, A.N. Active resource provision in cloud computing through virtualization. In Proceedings of the 2014 IEEE International Conference on Computational Intelligence and Computing Research, Coimbatore, India, 18–20 December 2014; pp. 1–4.
11. Usman, M.J.; Samad, A.; Chizari, H.; Aliyu, A. Energy-Efficient virtual machine allocation technique using interior search algorithm for cloud datacenter. In Proceedings of the 2017 6th ICT International Student Project Conference (ICT-ISPC), Johor, Malaysia, 23–24 May 2017; pp. 1–4.
12. Wang, C.F.; Hung, W.Y.; Yang, C.S. A prediction based energy conserving resources allocation scheme for cloud computing. In Proceedings of the 2014 IEEE International Conference on Granular Computing (GrC), Noboribetsu, Japan, 22–24 October 2014; pp. 320–324.
13. Liu, X.; Yu, J.; Wang, J.; Gao, Y. Resource allocation with edge computing in IoT networks via machine learning. *IEEE Internet Things J.* 2020, 7, 3415–3426.
14. Ullah, I.; Youn, H.Y. Task classification and scheduling based on K-means clustering for edge computing. *Wirel. Pers. Commun.* 2020, 113, 2611–2624.
15. Adhikari, M.; Nandy, S.; Amgoth, T. Meta heuristic-based task deployment mechanism for load balancing in IaaS cloud. *J. Netw. Comput. Appl.* 2019, 128, 64–77.
16. Somasundaram, T.S.; Govindarajan, K. CLOUDRB: A framework for scheduling and managing High-Performance Computing (HPC) applications in science cloud. *Future Gener. Comput. Syst.* 2014, 34, 47–65.
17. Behera, I.; Sobhanayak, S. Task scheduling optimization in heterogeneous cloud computing environments: A hybrid GA-GWO approach. *J. Parallel Distrib. Comput.* 2024, 183, 104766.
18. Hogan, A.; Blomqvist, E.; Cochez, M.; d'Amato, C.; Melo, G.D.; Gutierrez, C.; Kirrane, S.; Gayo, J.E.L.; Navigli, R.; Neumaier, S.; et al. Knowledge graphs. *ACM Comput. Surv.* 2021, 54, 1–37.
19. Imam, F.T. Application of ontologies in cloud computing: The state-of-the-art. *arXiv* 2016, arXiv:1610.02333.
20. Moscato, F.; Aversa, R.; Di Martino, B.; Fortiş, T.F.; Munteanu, V. An analysis of mosaic ontology for cloud resources annotation. In Proceedings of the 2011 Federated Conference on Computer Science and Information Systems (FedCSIS), Szczecin, Poland, 18–21 September 2011; pp. 973–980.
21. Guha, R.V.; Brickley, D.; Macbeth, S. Schema. org: Evolution of structured data on the web. *Commun. ACM* 2016, 59, 44–51.
22. Daniele, L.; den Hartog, F.; Roes, J. Created in close interaction with the industry: The smart appliances reference (SAREF) ontology. In Proceedings of the Formal Ontologies Meet Industry: 7th International Workshop, FOMI 2015, Berlin, Germany, 5 August 2015; pp. 100–112.
23. Liquori, L.; Scarrone, E.; Peraldi-Frati, M.A.; Jeong, S.M.; Cimmino, A.; Castro, R.G.; Koss, J.; Khan, A.Q.; Kumar, S.; El Khatab, S. ETSI SmartM2M Technical Report 103715; Study for oneM2M; Discovery and Query Solutions Analysis & Selection; Technical Report; European Telecommunications Standard Institute: Sophia Antipolis, France, 2021.