2016 IEEE International Symposium on Consumer Electronics

Device discovery strategies for the IoT

Pablo Calcina Ccori
IME – USP
pcalcina@ime.usp.br

Laisa Caroline Costa De Biase
LSITEC
laisa.costa@lsitec.org.br

Marcelo Knorich Zuffo
Escola Politécnica – USP
mkzuffo@usp.br

Flávio Soares Corrêa da Silva
IME – USP
fcs@ime.usp.br

***Abstract*—Billions of devices are expected to be Internet-enabled by 2020, levering the Internet of Things (IoT). Device discovery is a problem of finding useful devices in order to accomplish a task, and it constitutes a major challenge for IoT. Once a device finds another with the desired characteristics and establishes a negotiated communication, novel services can be generated through cooperation. Several efforts to solve the discovery problem have been proposed, focusing on different aspects of the process: service description, index organization, query definition, communication protocol, etc. In the present work, we focus on the topology of a network of devices and how it influences discovery. Three network topologies are evaluated using multi-agent simulations: (1) centralized, (2) decentralized and (3) hierarchical. These topologies are assessed with respect to discovery time, server-side infrastructure, reliability, and network traffic. Moreover, the obtained results have been used to design an IoT registry infrastructure.**

I. Introduction

A growing number of physical objects is being connected to the Internet, thus generating the Internet of Things (IoT). Such objects have embedded computational power, shared information, and coordinated decisions. A related paradigm is the Swarm, based on independent, cross-niche and heterogeneous devices cooperating to execute tasks synergistically [1].

A natural approach to implement cooperation among devices is service oriented computing, encapsulation a device as a service [2]. In a scenario comprising billions of devices, a major challenge is the discovery of such services.

The *discovery of services* requires the location in enormous networks of connected devices of a service featuring the required characteristics given a scope of interest, and the capability of securing a required quality of services. Service Discovery is designed to allow consumers to know the existence of other services without human intervention [3]. It leverages novel service compositions through cooperation.

Efficient discovery of services depends on the topology of service registrations, given that many devices are resource-constrained on data storage, processing speed, and energy consumption. Moreover, the topology used for discovery affects infrastructure investments, the reliability of discovery and the response time of services, since it can require a centralized server or rely directly on each node, as well as requiring several or few messages to find a requested service. In the present work we simulate and evaluate three topologies for communication in networks of devices: (1) *centralized*, (2) *decentralized* and (3) *hierarchical*.

Centralized network of devices considers that all nodes are connected to a central node, which acts as a directory, and discovery is performed through a request to that central node. In a *decentralized* network there is no central node and each node is connected to some neighbors; discovery is performed by flooding a request message until the target node is found or the Time To Live (TTL) is reached. In a *hierarchical* network, devices are connected to some supernodes which maintain partial directories; discovery is performed by flooding a request message among the supernodes.

II. Simulation

Multi-agent-based simulations have been used to analyze the impact on discovery by changing the topology of networks of devices. Each device is represented as an agent that communicates with other devices to discover a service in the network.

In *centralized* networks there are two agent types: one representing the central node that acts as a directory and a second representing all the other agents that represent the devices connected to the central node.

In *decentralized* networks we have only one uniform agent type. The *average node degree* parameter is used to define the average number of neighbors of each node. Large average node degrees require additional storage and generate larger numbers of messages, with the advantage of featuring greater probabilities to find a path between two devices.

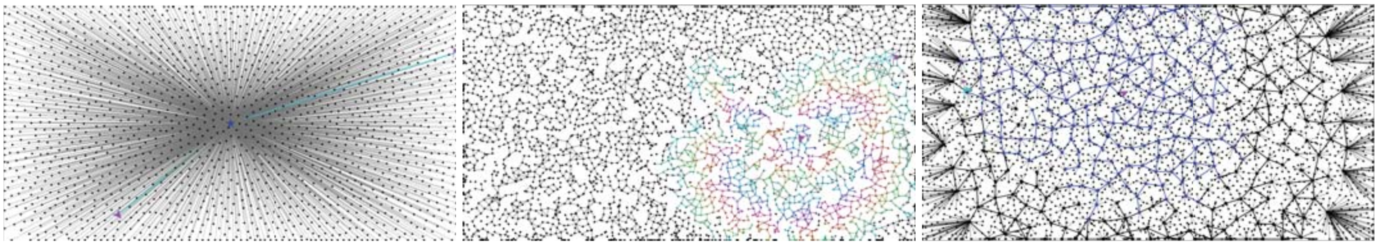
Finally, in *hierarchical* networks there are two agent types: supernodes and standard nodes, which are located randomly in the network. Every standard node is connected to the nearest supernode. The parameter *average node degree*, in this case, defines the average number of neighbors at the supernode level. The parameter *nodes failure probability* equally affects both node types. These topologies are depicted in Fig. 1.

The following parameters were used to control the simulation: (1) number of simulations (3000), (2) number of nodes (3500), (3) average node degree (2-4), (4) nodes failure probability (0-100), (5) number of failing nodes (0-500), (6) maximum TTL (Time to Live) (10-100), and (7) number of super nodes (500).

A. Evaluation criteria

We have adopted five criteria to compare these topologies [4]: (1) required storage, (2) discovery time, (3) network traffic, (4) success rate and (5) reliability. *Required storage* is used to keep the network structure information. In centralized and hierarchical networks, central nodes act as directories of sibling nodes. In decentralized networks, nodes keep the information about neighbors. *Discovery time* measures the time

978-1-5090-1549-8/16/$31.00 ©2016 IEEE



(a) Centralized network (b) Decentralized network (c) Hierarchical network

Fig. 1: Three network topologies used in the simulations

units required to find a target device. *Network traffic* represents the amount of messages issued until the target device is found or TTL is reached. *Success rate* measures the percentage of successful searches. In hierarchical and decentralized networks the target device is not always reached, due to the Time To Live (TTL) parameter in the search. *Reliability* is the measurement of success achieved in the discovery process, i.e. whether the requested service was found. Randomly chosen devices will be deactivated during the simulation in order to verify the impact in the discovery.

To better visualize the comparison we normalized the values and inverted the direction of scale for three criteria: for *required storage*, *discovery time* and *network traffic*, we have called its counterparts respectively: *storage efficiency*, *time efficiency* and *communication efficiency*.

B. Implementation

The model and simulation has been implemented using the NetLogo[1] software.

III. Results and discussion

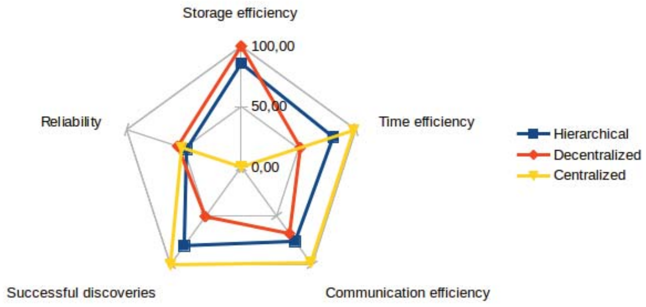


Fig. 2: Comparison of network features

Fig. 2 summarizes the simulation results. The measurements were normalized and the scale inverted, to obtain a chart in which the ideal configuration coincides with the perimeter of the polygon.

Storage efficiency is higher with decentralized and hierarchical topologies, since the directory is partitioned, so, this efficiency decreases with the number of supernodes used; centralized topology performed worst since it concentrates all the index storage in the central node. *Time efficiency* is higher with centralized topology since it requires a single request to the central directory to find any other node in the network. Hierarchical topology has a good time efficiency since it only finds in supernodes, which are considerably less than the total of nodes. Decentralized topology has the worst efficiency since it looks across a vast part of the nodes. *Communication efficiency* is better in centralized topology since it only uses one message to perform discovery; in the other two networks, a message flooding is performed. The number of *successful discoveries* was higher for the centralized network since it is guaranteed to find the node in the index when existent; for decentralized and hierarchical, it strongly depends on the *Max TTL* parameter. *Reliability* is similar for all configurations, and close to the probability of failure of individual devices; however, when varying the number of failing supernodes, reliability is significantly improved for decentralized and hierarchical networks.

IV. Conclusions

Agent-based simulation has proven to be very helpful to understand and to visualize the quality of service and device discovery considering different communication topologies for devices the in IoT. After evaluating the results, we have identified each topology strengths and weaknesses, which can be used according to each project needs. Hierarchical topology has featured comparatively good balance across the evaluated criteria when compared with Centralized and Decentralized ones.

References

[1] L. C. P. Costa, J. Rabaey, A. Wolisz, M. Rosan, and M. K. Zuffo, "Swarm os control plane: an architecture proposal for heterogeneous and organic networks," *Transactions on Consumer Electronics, IEEE*, vol. 61, no. 4, pp. 454–462, 2015.

[2] S. De, P. Barnaghi, M. Bauer, and S. Meissner, "Service modelling for the internet of things," *Federated Conference on Computer Science and Information Systems (FedCSIS)*, pp. 949–955, 2011.

[3] F. Zhu, M. W. Mutka, and L. M. Ni, "Service discovery in pervasive computing environments," *IEEE Pervasive computing*, no. 4, pp. 81–90, 2005.

[4] M. Patil and R. C. Biradar, "A survey on routing protocols in wireless sensor networks," in *Networks (ICON), 2012 18th IEEE International Conference on*. IEEE, 2012, pp. 86–91.

[1] https://ccl.northwestern.edu/netlogo/

978-1-5090-1549-8/16/$31.00 ©2016 IEEE