Information Sciences 376 (2017) 95–109

Contents lists available at ScienceDirect

Information Sciences

journal homepage: www.elsevier.com/locate/ins

Bayesian resource discovery in infrastructure-less networks

Francesco Palmieri

Department of Computer Science, University of Salerno, Via Giovanni Paolo II 132, I-84084 Fisciano (SA), Italy

ARTICLE INFO

Article history:
Received 2 August 2014
Revised 2 October 2016
Accepted 9 October 2016
Available online 11 October 2016

Keywords:
Resource discovery
Ad-hoc networks
Bayesian inference
Probabilistic broadcast
Future Internet

ABSTRACT

Resource discovery is one of the most important infrastructure-level facilities enabling the success of modern ad-hoc mobile communication networks, integrating services and capabilities provided by heterogeneous objects into the Future Internet environment. However, the highly dynamic and infrastructure-less nature of these organizations, where nodes continuously join and leave the network or change their attachment connections by moving between different coverage areas, makes resource discovery an extremely challenging tasks. In lack of specific knowledge about the availability of resources and their location over the network, flooding-based or pure probabilistic exploration approaches are the only feasible options to support search/discovery operations. By considering the high communication cost and the incomplete coverage problems respectively characterizing the aforementioned approaches we propose a novel adaptive random-walk search strategy, for resource discovery in ad-hoc networks, structured according to a selective stochastic query/response scheme where the exploration process is driven by posterior probability, using Bayesian inference and relying on the history of past discovery operations. This strategy, by also taking advantage of the scale-free properties characterizing the aforementioned ad-hoc organizations, is able to significantly contain the broadcast traffic without compromising the overall success of search operations by seamlessly accommodating to dynamic changes in resource location and network topology.

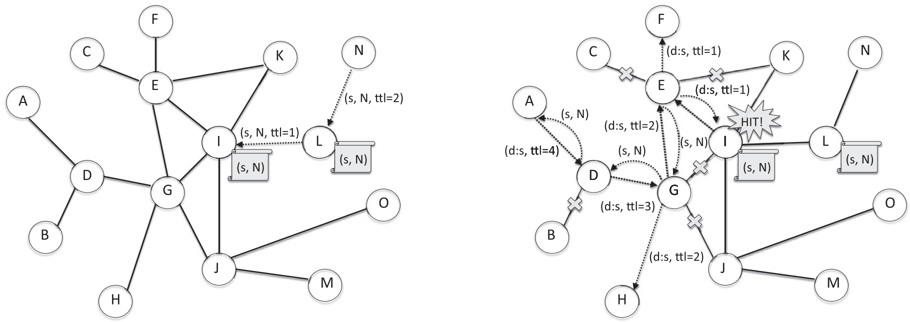
1. Introduction

Resource discovery is an essential element for the operations of infrastructure-less networks characterizing many future Internet scenarios, where many kind of mobile cooperating entities, without the need of specific infrastructural coverage elements such as access points or base stations, establish ad-hoc connections to communicate and lack of any prior knowledge about the other nodes and the resources made available by them.

A resource can identify any kind of capability or service provided through the network, such as a specific information, eventually acquired by a sensor device or a shared data repository, a publish/subscribe service as well as a runtime or storage element or any kind of networked object/equipment/actuator offering its capabilities to other nodes. The underlying wireless communication infrastructure allows the seamless integration of these heterogeneous resources and their provider entities into a coherent architecture, mapped on the global Internet according to the emerging IoT (Internet of Things) paradigm. These resources actively participate in business and everyday's life under the control of network-based applications that trigger their actions and take advantage of their services with or without the need of human contribution.

E-mail addresses: fp@di.unisa.it, fpalmieri@unisa.it

http://dx.doi.org/10.1016/j.ins.2016.10.021



(a) N advertises s, ATTL=2. (b) A searches for s, STTL=4.

Fig. 1. A simple advertisement and discovery scenario.

In principle, resource discovery is a rather simple task. Any new object joining the network has to announce all the services and facilities it provides, by making available information about its location and attributes through some kind of distributed directory service. Any other object/node or application on the same network that needs to access these services must issue an explicit request/query, by specifying the attributes describing the desired facility, to the same distributed directory, obtaining as a response the information or list of information (in presence of multiple provider objects) necessary to locate and access their target. However, the success of such scheme implies some knowledge about the entities providing directory services and hence an implicit hierarchical organization emerging around these predefined nodes. Unfortunately, the highly dynamic nature of infrastructure-less ad-hoc networks, where we cannot rely on any kind of organization with some predefined nodes assuming the role of directory servers, makes resource discovery a really challenging task. In such a complex scenario, the only viable strategy to successfully locate specific resources, also in the worst case, is propagating query messages throughout the whole network until at least a target providing the needed resource or service is reached. Such target, if it exists, replies to the query with its resource attributes so that the discovery process may be completed. Per-node caching mechanisms can be used to reduce the amount of query traffic by taking advantage of the history of past queries, so that some already traversed nodes, accumulating resource information in their cache may respond on behalf of the target resource providers, by dynamically assuming the role of directory servers. Clearly, the efficacy of such mechanisms depends on the query dissemination policy, which in turn, being tightly coupled with the underlying network dynamics, may hardly condition the success of the whole discovery framework.

The unpredictable mobility of nodes together with the lack of a fixed topology makes *flooding* the most used option for network-wide broadcast dissemination of queries for resources in infrastructure-less organizations [8,15,20], where each participating object continuously establishes multiple ephemeral ad-hoc connections with the other ones, in order to gain network access and exchange information with its neighbors. Unfortunately, flooding originates a non-negligible amount of redundant traffic that may rapidly exhaust critical communication resources such as radio channels or the energy needed for transmission as well as cause contention and conflicts, interference, and consequently, information loss. Hence, redundant broadcast containment will become a first-class objective for the performance and effectiveness of all the discovery mechanisms implemented on dynamic ad-hoc organizations [29]. Unfortunately, reducing the amount of broadcasts propagated throughout the network consequently reduces the degree of reliability and hence the chances of success of the above mechanisms. Thus, we have to consider a balance between the introduced level of redundancy and the needed service reliability. Probabilistic query broadcasting techniques [11,25], especially adaptive ones, if driven by proper heuristics towards successful solutions, may be really effective for selective broadcast containment in all the flooding-based network exploration mechanisms.

Accordingly, in order to cope with the above performance problems we propose an adaptive active search strategy, based on a step-by-step selective stochastic query/response process where an originating node relays its query to a subset of its neighbors chosen according to a Bayesian probabilistic decision model working on posterior probabilities estimated on the history of previous searches, and aiming at maximizing the success likelihood of each selection step. This model enables the search process in taking more informed decisions when choosing suitable query forwarders and thus significantly improves the search coverage and success rate. Each neighbor when receiving the query iterates the same Bayesian forwarding process, by taking another step, independent from the previous ones, towards its neighborhood, and so on, thus generating multiple parallel random walks throughout the network, until the required resource information is found within the cache of a visited node (see Algorithm 1(b) lines 12–20 and Fig. 1(b)). This strategy avoids the query traffic explosion associated to flooding and thus greatly reduces the communication overhead. The resulting framework results to be fully distributed and inherently parallel in its nature, and capable of working in lack of any kind of centralized control, predefined roles and stable communication links. It provides a good degree of resiliency to the whole discovery process, that can easily and seamlessly adapt its behavior to any on-the-fly change in resource location and query distribution. Furthermore, the unique