Survey of Service Discovery Protocols in Mobile Ad Hoc Networks[1]

Technical Report 4/06

Adnan Noor Mian, Roberto Beraldi, Roberto Baldoni

Dipartimento di Informatica e Sistemistica "Antonio Ruberti"
Università degli Studi di Roma "La Sapienza"
Via Salaria, 113 – 00198
Rome, Italy

{Adnan, Roberto.Beraldi, Roberto.Baldoni}@dis.uniroma1.it

Abstract. Mobile devices are inherently scarce in resources, necessitating the need to cooperate among them for performing tasks that cannot be done alone. This cooperation is in the form of services that are offered by other devices in the network. To get benefit from the services offered by other devices, they have to be discovered. Service Discovery Protocols (SDPs) are used for this purpose. This is an important area of research in mobile computing and ubiquitous computing. In this paper twelve SDPs for multihop mobile ad hoc networks are analyzed with respect to their service discovery architectures, management of service information, search methods, service selection, methods for supporting mobility and service description techniques, Among these the most important aspect is the service discovery architecture as this affects other aspects of the service discovery. In this paper the service discovery architectures are categorized in four groups namely directory-based with overlay support architecture, directory-based without overlay support architecture, directory-less with overlay support architecture and directory-less without overlay support architecture. The management of service information and search methods mainly depend on the type of service discovery architecture. It is found that mobility support and service selection methods are independent of the SDP architecture. Also the description of services is found to be independent of the SDP architectures. Mostly the services are described using XML or the extensions of XML. At the end of the paper open issues and areas of further research are discussed.

Keywords: Mobile Ad hoc Networks, Service Discovery and Service Discovery Protocols

1 This work was supported by the Italian Ministry of Education,University, and Research (MIUR) under the "IS-MANET" project and by the NoE "RESIST" (European Commission 026764)

1 Introduction

A service can be any tangible or intangible thing that can be useful for someone. For example when a laborer works for building a house, he is giving his services for which he is paid. Similarly a teacher teaches to his students providing them knowledge that is useful for students. This act of teaching is a service provided by a teacher to his students. In our context of mobile ad hoc networks any facility provided by a device that can be useful for any other device is a service. A service in this context could be a software service like providing an implementation of some algorithm (for example, converting one audio file format to another) on a device so that when some device needs this service, it can contact that device and use it. A service can also be a hardware service like a printer that can be used by a mobile device to print a file. To get benefit from these services a device must be able to locate them in the network and also have the ability to invoke these services. Here comes the role of service discovery protocols.

In fixed and wired networks service discovery protocols simplify the interaction among users, devices and services [8]. Service discovery protocols allow devices to automatically discover network services thus making the task of network administration and configuration easy.

In wireless mobile ad hoc networks devices are free to move. The characteristic limitation of a mobile device is that it has to be small in size and weight. Such devices inherently have few and limited number of resources as compared to fixed devices. So it becomes important to utilize the resources and services available in other devices to accomplish the tasks that cannot be done alone. For example a mobile device without a printing support will require a printer to fulfill the printing task. Thus forming ad hoc network between mobile devices and getting benefit of resources available in a network require knowledge of services available by other devices and how to interact with these services. The service discovery protocols aim at these aspects. More specifically the service discovery protocols not only provide mechanisms for locating a particular service but also mechanisms to advertise a service, invoke a service, select a service if there are more than one services of the same type available and to describe a particular service so as to make its searching easy.

There is a lot of research going in the field of service discovery. Basically there are three types of networks as far as the research in service discovery is concerned. First are the wired networks, second are single hop wireless networks and third are the wireless multihop mobile ad hoc networks. The service discovery protocols suggested for one type of network are not suitable for another type of network because each network is based on different assumptions, the most important being the mobility and rate of joining and leaving of devices from the network. In the first type devices do not move at all and there is no join/leave at all or the join and leaves are few and far between. In the second type the network formed is ad hoc with very restricted mobility and having low rate of join/leave. There are one or more nodes that are fixed. But in the third type of network the devices are assumed to have unrestricted mobility and these can join or leave the system at any rate. There may be no fixed node. Due to these assumptions the problem of service discovery is very challenging in the third type of networks.

In wired networks many service discovery protocols have being suggested, some of which have gained the status of industry standard. For example Jini [17] by Sun, Universal Plug and Play (UPnP) [18] by Microsoft, Salutation [19] by IBM and Service Location Protocol (SLP) [20] by IETF. Mainly the service discovery protocols for wired networks fall into one of the three categories. Some are directory-based, like Jini that has a centralized place to store information about the services. Some are directory-less like UPnP that has a peer-to-peer (P2P) architecture and the service information is stored on each device. The third category is the hybrid of the above two. For example SLP can work in both modes, that is with or with out a directory depending on the situation.

In single hop ad hoc networks there are also some mature protocols available. For example Bluetooth SDP [22] and DEAPSpace [21]. Bluetooth's SDP is an industry standard. These protocols may follow P2P architecture like DEAPSpace or a client-service approach like Bluetooth SDP.

In multihop Mobile Ad hoc Networks (MANETs) a lot of research is going on but still it has not been mature enough to be used by industry. The main reason for this lack of mature research, in spite of lot efforts by the research community is because of the challenging issues due to the unrestricted mobility of devices. A lot of work has been done in the field of routing in MANETs. One can take advantage of this work for studying the service discovery problem. But essentially the service discovery problem is different from the routing problem. In routing we know the ID of a node, which is unique and data is only sent to that particular device, whereas in service discovery there is a service, which may not be unique and its multiple copies can reside on different devices. The task is to find that service, which best fits some given criteria. A service discovery protocol (SDP) may use an underlying routing protocol to invoke and get a reply from a particular service residing on a particular device. There are some SDPs that integrate the functionality of routing and service discovery. Thus service discovery and routing although are quite related to each other but specifically have distinct characteristics. One can take advantage of the research work going in one field for the benefit of other.

There are some good surveys of the service discovery protocols that also include SDPs for MANETs. For example the surveys done by Cho and Lee [14], Zhu and Mutka [16] and Marin-Perianu, Hartel and Scholten [15]. These surveys survey the service discovery protocols for all of the three types of networks and none of the survey go deep into surveying service discovery protocols for only the multihop mobile ad hoc networks. In this paper we have concentrated on SDPs for MANETs only. For this purpose we selected twelve SDPs given in the reference from [1] to [12] that have been referenced quite often by other authors. These are protocol by Cheng et al [1], GSD [2], Allia [3], Konark [4], Service Rings [5], Lanes [6], protocol by Kozart et al [7], Splendor [8], protocol by Varshavsky et al [9], protocol by Tyan et al [10], Field Theoretic Approach [11] and DSD [12]. We have selected six major components or aspects of any SDP and then analyzed the selected SDP with respect to these aspects. These aspects are service discovery architectures, management of service information, search methods, service selection methodologies, mobility support mechanisms and service description. From our point of view the most important aspect of any service discovery protocol is its architecture as other aspects may also depend on it.

The paper is organized as follows. In section 2 service discovery architectures are discussed. After discussing the existing two types of categorizations, another categorization is proposed that is a better representation of service discovery architectures of MANETs. It is also shown how different MANET SDPs fit in this categorization. In section 3 and section 4 the management of service information and search methods respectively are discussed. Service selection methodologies and the metrics used for algorithm based service selection are discussed in section 5. Section 6 is devoted for describing methods for dealing with mobility in MANETs. First the conditions in which a system of mobile nodes need mobility support is explained and then three methods for supporting mobility in existing SDPs are presented. Section 7 describes three different trends in service description techniques. Finally section 8 presents the conclusions and areas of future work.

2 Service discovery architectures

Broadly speaking architecture specifies the layout of any structure and how major components of that structure are connected with each other. The SDP architecture mainly depends on having or not having a directory. A directory is an entity that stores information about services available in the network. It helps in the service discovery process [14]. With respect to the directory the architectures for any service discovery protocol could be directory-based, directory-less and hybrid of the two. C.K. Toh [13] has already categorized three possible service discovery architectures in MANETs. These architectures are service coordinator based, distributed query-based and hybrid of these two.

In service coordinator based architecture, which is similar to directory-based architecture or centralized directory as called by Cho and Lee [14], certain nodes in the MANET are chosen to be Service Coordinators (SCs), a role quite similar to the Directory Agent (DA) in SLP [20] or the lookup service in Jini [17]. SCs announce their presences to the network periodically by flooding SC announcement messages. The flooding is limited to a certain number of hops, determined by the SC announcement scope parameter. Directory-based protocols include Service rings [5], Splendor [8], protocols by Kozart and Tassiulas [7] and Tyan and Mahmoud. [10].

In distributed query-based architecture, which is same as the decentralized directory or directory-less as Cho and Lee [14] has called it; there are no such Service Coordinator nodes. Instead, a client floods the service discovery request throughout its surroundings in the network. The flooding is limited by the flooding scope parameter. Directory-less approach is used far more than directory-based approach in MANETs. Examples of this approach are GSD [2], Allia [3], Konark [4], Field theoretic approach [11], protocols by Cheng and Marsic [1], Varshavsky, Reid et. al. [9] and DSD [12] by Chakraborty, Joshi et. al.

The hybrid architecture combines the above two architectures. Service providers within the announcement scope of one or more SCs will register with them their available services and access information, but the service coordinators themselves must also be ready to respond to flooded service requests. When a User

Agent unicasts a service request to its affiliated SC according with the service coordinator based architecture, the SC responds with a positive or negative service reply. However, if there is no SC in the User Agent's surroundings or if the affiliated SC returned a negative service reply, the User Agent will simply fall back to the Distributed query based or directory-less architecture. Hybrid architecture is mostly found in wired networks and there is no real example of such type of SDP architecture in MANETs.

The SDP architectures for MANETs can also be categorized by having overlay network support or not having overlay network support.. An overlay network can be explained as follows. If a node, which is a part of an ad hoc network, knows the address of another node of the same network and can communicate with it, then we say that there is an overlay link between the two nodes [5]. An overlay link does not necessarily mean that the two nodes have a direct physical or wireless connection. Two nodes can form an overlay link even if they can communicate through many intermediate nodes. An overlay network is a collection of such overlay links and the nodes they connect. The overlay can be a structured overlay network when there is some organization between the nodes forming the network or it can unstructured in which there is no organization and the node are connected randomly. In both cases there is a bootstrapping phase in which the nodes form the overlay network (structured or unstructured). Also there are special algorithms for joining and leaving of nodes from the overlay network. Note that there is a difference between unstructured overlay network and a network that does not have an overlay at all. In an unstructured overlay network firstly, a node forms overlay links to nodes, that is, it has the addresses of nodes beyond the neighbor nodes and can communicate with those nodes and secondly, these overlay links are randomly connected with the node. In case of not having an overlay network, a node has only the addresses of its neighbors and does not have addresses of nodes beyond its neighbors. Thus a node can only communicate directly to its neighbor nodes. From neighbors we mean all nodes that are in the radio range of a node.

Also note that whenever there is an overlay network support, it will always be a structured overlay and not an unstructured overlay. The reason for this is that the structured overlay network has the advantage of controlled multicast of service query or advertisement message. This controlled multicast restricts and greatly reduces the network traffic. Thus, although we pay for forming and maintaining the structure but also get an advantage of reduced network traffic. In case of having unstructured overlay networks in MANETs there is no such advantage of reduced network traffic but the cost of forming and maintaining the overlay is always there, although it may be less than the previous case. We thus have to pay but without getting any advantage. Therefore it does not make any sense to use unstructured overlay networks in MANETs. Examples of protocols forming overlay networks (which are structured) are Allia [3], Service rings [5], Lanes [6], protocols by Kozart and Tassiulas [7] and Tyan and Mahmoud [10].

The SDPs that do not form overlay, do not have a bootstrapping phase or special algorithms to maintain the overlay structure. Thus nodes may show low latency in forming a network and during join and leave operation. But on the other hand the multicast cannot be controlled. The only way to restrict the service query or advertisement message is by specifying the Time To Live (TTL) parameter of the messages. Example of protocols that do not form an overlay network are GSD [2], DSD [12], Konark [4], Splendor [8], Field theoretic approach [11] and protocols by Cheng and Marsic [1] and Varshavsky, Reid et. al. [9].

If we combine the existing ways of classifying the service discovery architectures, we get a classification that is a better representative of SDP architectures in MANETs. On this basis we categorize the SDP architectures in four categories, as given below.

(i) directory-based with overlay support architecture
(ii) directory-based without overlay support architecture
(iii) directory-less with overlay support architecture
(iv) directory-less without overlay support architecture

On the basis of above classification we give a brief overview of the SDPs in each of the category. One of the protocols in directory-based with overlay network support is Service rings [5]. It forms an overlay structure by grouping of nodes that are physically close and offer similar services. This overlay is formed on top of transport layer of ad hoc networks. The structure is called service ring. Each service ring has a designated service access point (SAP) through which the nodes within the ring can be accessed as it has all the information about the services offered within the ring. These SAPs are also connected with SAPs of other service rings thus forming a hierarchical structure. The directory information is kept in chosen edges that are dynamically selected. The protocol Lanes [6] also falls in the same category. It is inspired by Content Addressable Network (CAN) protocol, which is used for service discovery in wired peer-to-peer networks. Some nodes are grouped together to form an overlay network forming lanes of nodes. Each group is called a lane. Nodes in the same lane have the same directory replicated in each node cache. There are different lanes in a network that are loosely coupled with each other. Similarly in the protocol given by Kozart and Tassiulas [7] a subset of the network nodes forms a dominating set, also called virtual backbone. The nodes in the virtual backbones keep the directory, which stores the advertised information about other services in the network. The protocol given by Tyan and Mahmoud [10] forms cluster of mobile nodes in which each cluster has a gateway node. This gateway node is used for routing and keeping the directory.

The work in the field of directory-based without overlay network support is not much. Here we just have the protocol Splendor [8] by Zhu, Mutka et al. Even in this protocol the emphasis is on security aspect. It has four components, which are clients, services, directories and proxies. The directories are used for caching the service information and answering the client service requests. The proxies are used to authenticate the mobile services.

Similarly we have just one example of the work done in the directory-less with overlay network support architecture. Allia [3] is the only example. It follows a decentralized directory-less approach in which the nodes, which are geographically close form groups called alliances.

In the category of directory-less without overlay network support we have many SDPs as this architecture seems most obvious for the mobile ad hoc systems. For example the protocol by Cheng and Marsic [1] is directory-less P2P based on ondemand multicast routing protocol (ODMRP). GSD [2] and DSD [12] protocols are also P2P based and has a decentralized approach. Similarly an interesting approach is a Field theoretic [11] suggested by Lenders, May et. al. This approach is inspired by electric field concept and uses a simple and distributed mechanism to find the best route to the closest service instance. It is totally a decentralized protocol without any central server. Konark [4] is a completely distributed protocol and assumes IP connectivity between ad hoc nodes. Each device runs a stack of Konark application, SDP managers and registry. Another protocol that is directory-less and do not form overlay is proposed by Varshavsky, Reid et. al. [9]. It has two main components. A routing protocol independent Service Discovery Library (SDL) and Routing Layer Driver (RDL). SDL function is to store information about the service providers. RLD, which is closely coupled with the MANET routing mechanism, is used to disseminate service discovery requests and advertisements. Each node has the stack containing SDL and RLD to form a P2P networking or a directory-less architecture.

The categorization of protocols in different SDPs architectures is shown below.

Table 1. Categorization of SDPs Architectures

	Directory-based	Directory-less
Overlay	• Service Rings [5] • Lanes [6] • Kozart et al [7] • Tyan et al [10]	• Allia [3]
No overlay	• Splendor [8]	• Cheng et al [1] • GSD [2] • Konark [4] • Varshavsky et al [9] • Field Appr. [11] • DSD [12]

3 Management of service information

Service information includes any information about a service that is provided by service providers in advertisement. This information is used to describe, identify and discover a service in a network. The information may include name of the service, ID of the service, IP address of the service provider, port number of service point, protocol that server and a client may use to invoke a service [1], etc. The information about services must be stored somewhere so that other nodes can contact this node to discover a particular service. Management of service information include aspects like where to store service information, time duration for which the information will be stored, distance in number of hops the information will travel as advertised by the service provider, etc. There are different ways different SDPs manage service information. Some protocols select a particular node among a group of nodes and store this information in a directory that resides on that particular node. Some store only on their local cache. The service information, depending on a particular advertisement policy, can be made to travel far or the advertisements can just be stored on neighbors. Some protocols require the service providers to refresh the service information before being deleted from the storage place. The exact mechanism how different SDP manages the service information is explained below.

First let us describe the mechanism of management of service information in protocols that do not have a centralized directory of storing service information. For example in Cheng and Marsic [1] protocol service information is stored on every node that is interested in the service. Each service provider multicasts advertisements about the services it can provide to the ad hoc network. Each server and its possible consumer make a multicast group. Any node that is interested in a particular service or services stores the advertisement in its local service registry and may send a service awareness reply to the service provider. Once some clients send back service awareness replies, the server sends its updated services advertisements by multicasting them to only those clients. With a similar idea in GSD [2] and Field theoretic approach [11] every node caches advertisement to maximum N hops (called the advertisement diameter). The service cache in each node thus contains a list of the local as well as remote services that this node has seen through advertisements. In this way each node in its local cache, contains the description of services that are within the advertisement diameter. By restricting the advertisement hops these protocols achieves better local memory utilization and also the probability of discovering a service in its vicinity. In Allia [3] the difference is that each node advertises its services only to the immediate neighbors, that is, the advertisement diameter is just one hop. Some nodes according to a local policy cache these advertisements. These nodes form an alliance with the advertising node. Thus every node stores, in its local cache, the advertisements from nodes in its alliances. As in Allia, DSD [12] also advertises its service information to all nodes in its radio range but here each node in addition to storing the service information may also forward the advertisement to other nodes, depending on the forwarding policy. Similarly in Konark [4] service information is stored on each node. This protocol has a special structure called service registry that is present on every device. It is used to store all the local service descriptions as well as the service descriptions that a node comes across through advertisements. In Varshavsky, Reid et. al. [9] scheme there is a service discovery library (SDL) on top of routing layer on every node. SDL maintains a service table that keeps record of service information.

The following explain the mechanism of management of service information in protocols that keep a centralized directory for storing the service information. For example Service rings [5] forms groups of nodes that are physically close to each other. Each group called "service ring" has a designated node called the "service access point" which keeps record of the services present in the ring. Similarly Tyan and Mahmoud [10] scheme form groups of mobile nodes based on their location. A node is chosen as a gateway that contains the directory. Lanes [6] also form group of nodes called lanes in which each node knows it predecessor and successors. There is an anycast address of each lane in which all nodes share the same anycast address. The services advertised are sent to an anycast address of a lane. All nodes within a lane have the same directory replicated and thus have the same information stored. Kozart and Tassiulas [7] scheme is also not very different from the previous protocols as this protocol also selects a particular node for storing the service information. This protocol forms virtual backbone from some of the mobile nodes. When a particular node advertises its services, these are stored on the directories located on the virtual backbone nodes in case the service provider is itself not a backbone node. If the service provider is itself a backbone node then it registers services on the same node. Splendor [8] is a little different as in this protocol the service information is stored in special nodes called directories that are pre-assigned the task of storing service information and are not selected by the SDP.

4 Search methods

Searches can be used either to find the node having the directory or to find the services. The exact purpose of search depends on the type of storage method and the SDP architecture [15]. The search method depends on the type of network in which the search is made. Mainly there are two ways to search for information about the services available in the network.

(i) The first method is used in networks that have directories for storing service information. The directory nodes keep information about the services available in a group and clients query these directories.

(ii) The second method is used when there are no directories. Service providers advertise the services to all of the nodes. When a node is in need of a particular service, it searches its local cache for the presence of the service. If the service is not found query messages is sent to all nodes.

Service rings [5] is an example of SDPs in which searching is done using directories. It has special nodes called Service Access Points (SAP) that keep all the information about the services within the ring. When a node wants to search for a service, the query is routed through the ring structure, passing through SAPs of other rings and reaching only to those subrings that can possibly offer the service. This use of special overlay network with SAPs restricts the query flooding to only the selected nodes. Similarly in the protocol presented by Kozart and Tassiulas [7] the client forwards the service request to a virtual access point. These virtual access points, also called the virtual backbone node (VBN) broadcast or multicast the query message to all the other VBNs. Only flooding the backbone nodes instead of all nodes in the network thus reduces the overhead of broadcasting a query. In Splendor [8] directories are first discovered by sending queries by the clients or the directories themselves announce their presence periodically in the network. Clients after knowing the directory addresses, query the directories for services. In the SDP by Tyan and Mahmoud [10] the gateway of each cell provides the directory services containing information about the services of other gateways. When a client wants to search for a service, it sends a service request to its local gateway. The gateway searches its service advertisement cache and in turn gets a list of advertisements that corresponds to the service. In Lanes [6] the case is a little different. It has directories but the same directory is replicated throughout the lane overlay nodes. The service announcements are sent through a lane and service requests are sent through other lanes. These messages are sent through lanes by anycast routing.

The protocol given by Cheng and Marsic [1] does not have directories. When an appliance needs a service, it sends a query to service query multicast group. This group consists of a service provider and its possible consumers formed during the bootstrapping phase. In GSD [2] first the search is done in the local cache as it contains information about all the services within the advertisement diameter. When service is not found in the local cache then query request is selectively forwarded to a set of nodes based on semantic information. Similarly Allia [3] first checks the local cache for service information and if it is not available then active discovery is done by multicasting query request to the members of its alliance. If the service is still not available then the query request is broadcasted to other alliances in its vicinity. In Konark [4] and the protocol given by Varshavsky, Reid et. al. [9] also, the services are searched by first looking at the local cache and then, if not found, multicasting the service request to a fixed group of nodes in the network. Those nodes respond to the query message that can provide the service. In field theoretic approach [11] service advertisements are flooded through the network within a limited scope. Each node temporarily stores the advertisement and calculates the potential. When a client wants to search for a service, it forms a service query containing the service type of the desired service. This query is routed to the neighbor with a higher potential for that service, eventually reaching the service. In DSD [12] a service request based on ontology-based description is formed. The request is first matched with the services in the local cache and if services are not found, the request is selectively forwarded to other nodes based on the ontology descriptions. When the node does not have enough information to selectively forward a request, then a broadcast is made to the neighboring nodes.

5 Service Selection Methodologies

The query request from a client node to the network can result in many responses of matching services. Although there are many service discovery protocols that do not deal with the selection issue but for a service discovery protocol to be complete, handling of multiple responses of the same services should be taken care of and it should be part of the of service discovery protocol to select one of the available services for invocation. There can be different ways to select a service. For example it could be done manually or the selection procedure can be automated using some algorithm based on some criteria. The criteria or the metrics for service selection have been defined differently by different protocols. For example the lowest hop count, current load of a service provider, bandwidth available of the communication channel between the service provider and the client, velocity of the service provider are some of the criteria.

Varshavsky, Reid et. al. [9] protocol integrates the service discovery and selection feature with the underlying routing protocol. They have demonstrated that proper service selection improves the overall network performance, by localizing the network communication. The mechanism used for service selection is simple. When multiple entries in the service table match the request, the client selects with the lowest hop count.

In Tyan and Mahmoud [10] proposal mobile agents are used for the service selection. When the mobile agents receive a list of advertisements from the service discovery phase, these agents move to different nodes while selecting the services according to some criteria. For example the user can specify the mobile agent to choose services with highest rating returned or services having some index values higher than user specified index value or the user can specify the mobile agent to just return the first available service.

In Field theoretic approach [11] client selects services using two metrics, one is the network distance, that is, the number of hops and other is the capacity of service (CoS). The algorithm for service selection is distributed and does not involve direct interaction with the client.

Splendor [8] specifies that the service selection will occur at client end but does not give detail of the algorithm used for service selection and also does not tell about the selection criteria. The protocols by Kozart and Tassiulas [7] and by Cheng and Marsic [1], GSD [2], Allia [3], Konark [4], Service rings [5], Lanes [6], DSD [12] do not tell any thing about selection mechanism.

6 Mechanisms for Mobility Support

In a system of mobile ad hoc network nodes keep on moving and changing their position with respect to each other. In a system in which the all of the nodes just keep information about their own services and not of the other nodes, mobility is not a big issue as searching is done by multicasting a query message to all nodes in the system. But the limitation of such systems is that they have less number of nodes and are not scalable. For example in [1] the mobility support is implicitly provided by the multicast mechanism. Systems that:

(i) have directory nodes that keep all of the information about other services in the network or
(ii) have nodes that keep partial information about the services, for example services present in the neighbor node or
(iii) form structured overlay networks,

mobility is a real issue that has to be taken care by a SDP if the protocol has to function properly.

Mobility support implies that the information about services in the directory nodes is up-to-date under mobility. That is if a directory node is supposed to keep information about the all nodes in a group then it must have that correct information. If that node changes it position with respect to the group, the directory information should also be updated quickly. Only by this way the SDP can search the services in a timely manner. Mainly there are three different ways to support the mobility.

(i) Updating service information
   a. Event driven updating of service information
   b. Periodical updating of service information

(ii) Advertisement controls
   a. Changing the rate of advertisement
   b. Changing the diameter of advertisement

(iii) Algorithms that maintain the structure of overlay network in SDPs that form structured overlay networks

Service information can be updated mainly by two ways. One is to update the service whenever there is any event occurring. For example when there is no route available to the service provider, the service information should be updated. The other way is to update the service information on regular basis for example by periodical advertisements as done by Konark [4], Splendor [8], DSD [12], Field theoretic approach [11], protocols by Kozart and Tassiulas [7] and by Tyan and Mahmoud [10]. The protocol by Varshavsky, Reid et. al. [9] uses both methods.

Some protocols change the rate and diameter of advertisements as the mobility of node changes. If the nodes are moving faster then rate of advertisement is increased and the diameter, that is the number of hops an advertisement can travel, is reduced. This type of mechanism is done in GSD [2] and Allia [3].

Some protocols form overlay structures and can only search correctly for services if that overlay structure is maintained. Due to mobility the overlay structure may get faulty. In this case there are special algorithms that try to maintain the overlay network structure. For example Service rings [5], Lanes [6], protocols by Kozart and Tassiulas [7] and by Tyan and Mahmoud [10].A brief description of how mobility is managed in each protocol is given below.

Let us briefly describe how different protocols support mobility. The protocol by Cheng and Marsic [1] supports mobility by using multicasting for discovery of services in the networks. The mobility support is thus not explicitly provided by any special mechanism but it implicit in the multicasting the service requests [15].

In GSD [2] there are two parameters that can be adjusted for different mobility scenarios: the advertisement diameter and advertisement time interval. Advertisement diameter is the number of hops that an advertisement is expected to travel in the network and advertisement time interval is the time interval after which every nodes sends a list of services it has to all the nodes in its radio range. In high mobility scenarios, for example, the advertisement time interval can be reduced to cater for the rapidly changing vicinity. Similarly the advertisement diameter can be regulated with the dynamism of the network. In DSD [12], in addition to the mechanism discussed in GSD [2], this protocol takes care of the effects of the mobility of nodes in the following way. The services announce when they enter the network and the neighbour nodes cache this information. If the advertisement is not refreshed after a specified time the information about the service will be removed from the cache of other nodes.

Allia [3] takes care of mobility by adjusting the advertisement rates and alliance diameter based on the mobility of the nodes. Regarding the advertisement rates one of the three methods can be employed. First is simply use a constant frequency rate for advertisements. This can be used for relatively stable networks. Second method is to use Multiplicative Increase Linear Decrease (MILD) algorithm or a Binary Exponential Back-off (BEB) Mechanism to vary the advertisement frequency. The advertisement frequency would be higher for more dynamic networks and low for less dynamic networks. Third possibility is sending out and advertisement only when it receives a new advertisement. The alliance diameter is the number of hops the advertisement may propagate in the network. Any node within the diameter would be able to cache the advertisement. For highly dynamic networks small advertisement diameter is adjusted and vice versa.

In Service rings [5] the overlay network is corrected which gets faulty due to the mobility of nodes. Each ring member only knows its successor and it predecessor. RingCheck messages, initiated periodically by the appropriate Service Access Points (SAP), circle through each ring to check its consistency. Every ring member receiving the message puts its predecessor information and forwards it to its successor. If a node does not receive such a message in one of its rings for a certain time it checks for a link breakage or a partition in the network. If any of these cases is detected then an appropriate algorithm is initiated to repair the ring.

In Lanes [6] the lane structure of the overlay network is maintained by different algorithms. Each node pings its upper neighbor and receives pings from its lower neighbor to maintain the lane. If any of the pings is missing either a node is detected to be vanished or the network is detected to be partitioned. In any case there are appropriate algorithms that are initiated to build a regular overlay structure according to the lane protocol specifications. Also there are algorithms for node logging in and logging off that keep the regular overlay structure.

In the protocol by Kozart and Tassiulas [7] the service registrations are done on periodic basis. In this protocol a virtual backbone is formed. To take care of frequent topology changes due to mobility or nodes vanishing, the dominating set feature of the backbone is maintained with the help of specific algorithms.

Splendor [8] and Konark [4] store service information as a soft state. When a service advertises itself, it also announces its lifespan. Before a service expires, it has to announce again. The proxies cache the information about the mobile services. Thus regular advertisements keep the information updated.

Due to mobility some of the service providers may not be accessible and some new ones may be in range. Reselection and rediscover are two mechanisms through which the protocol given by Varshavsky, Reid et. al. [9] takes care of the mobility of nodes. In reselection the services based only on the current entries in the service table are reconsidered. The policy when reselection should occur could be different. For example one reselection policy could be that reselection should occur when there is any change in the service table. Another policy could be to reselect the services when there is no route to the server. In rediscovery the network is probed for up-to-date information about the available service providers.

In the protocol by Tyan and Mahmoud [10] mobility is supported by two mechanisms. First when a gateway node moves to another cell, it broadcast the service registry tree to the nodes in its previous cell. These nodes elect another gateway node. This gateway then starts using the service registry information. The second mechanism is by specifying time to live parameter, which is the physical clock time after which a service has to refresh its advertisement.

Field theoretic approach [11] protocol also uses periodical advertisements. The nodes can be disconnected from their neighbors due to mobility. This is determined by listening to the periodic update message from the neighbor node. If a node does not receive such a message for a long time it assumes a broken link and removes the neighbors from its table.

7 Service Description techniques

Service description is an abstraction of the facilities and characteristics of a service. The description of a service is necessary if it is to be utilized by other devices or services. The nodes in a network search for services by only looking at the descriptions of the services advertised by the service provider. A service, not properly described, may remain completely unknown to other devices in the network, thus defeating the objectives for which a service was formed. For these reasons SDPs usually describe the way services are described and the languages used for description. In MANETs SDPs we find three trends with regard to service description.

(i) Most commonly used language for service description is eXtensible Markup Language (XML) and its extensions like DAML (DARPA Markup Language) [23] and Web Ontology Language (OWL) [24]. For example GSD [2], Konark [4], Service rings [5] and DSD [12].

(ii) Some SDPs are independent of any description language. Any language or description method can be used in these protocols. For example one is free to use simple text attribute-value schemes or XML for describing services. For example Allia [3], Lanes [6] and the protocol presented by Varshavsky, Reid et. al. [9].

(iii) In SDPs the issue of description is not discussed. These protocol are usually concerned only with the searching of a service and do not go into the details of other aspects of SDP. The authers by Kozart and Tassiulas [7], Zhu, Mutka et al in Splendor [8], Lenders, May et. al. in their Field theoretic approach [11] and Cheng and Marsic. [1] do not touch the issue of description of services in their protocols.

The details of describing a service in different protocols are following. Konark [4] protocol defines an eXtensible Markup Language (XML) based service descriptions. The description file is a plain text file that has all the information about the characteristic and functions of the service. GSD [2] use DAML (DARPA Markup Language) and OIL (Ontology Interference Layer) to define ontology to describe the services in mobile ad hoc networks. DAML + OIL is based on XML and the Resource Description Framework (RDF) [25]. The semantic capabilities of DAML make it a good choice for the description of services. The service requests are also expressed in DAML that are matched with the service description during the discovery process. The services are classified into groups based on class-subclass hierarchy present in DAML. The semantic features of DAML are used to reduce the network flooding. Web Ontology Language (OWL) is used in DSD [12] to describe services. OWL is also based on XML and RDF and is used in wired networks to describe services. The semantic class-subclass hierarchy present in OWL is used to described service groups. This also helps in selectively forwarding the service request.
In Allia [3] framework services can be described using any method, for example using XML or any other alternative. During the service discovery mechanism no description mechanism is specified, thus making Allia independent of any descriptions of services. Also the protocol presented by Varshavsky, Reid et. al. [9] is independent of any service description. To make the protocol [9] independent of any service description language, a matching of service advertisements by the service providers and service requests are handled by a pluggable matching module. The approach given in Lanes [6] is also independent of the service description. Similarly Service rings [5] will work with all the descriptions that satisfy the two conditions. First, there should be a distance function that allow to compare different service descriptions and second, there should be a summarize function which should produce a single new description if it is given a set of service descriptions. For example on simple taxonomies of services, both these functions can be defined. Another example is DAML-S language.

8 Conclusions and Future Work

In this paper we have surveyed SDPs for multihop MANETs. We selected twelve SDPs for MANETs and compared these protocols with respect to six important aspects. These aspects, which we chose for evaluating the protocols, are service discovery architectures, service information storage, search methods, service selection, mobility support and service description. There are many other aspects from which any SDP (for wired or wireless) can be evaluated, for example one of these is security, but these aspects are either not important with regards to MANETs or most of the protocols at present do not discuss these aspects. We have found a clear categorization that is a better representation of SDPs in MANETs. This categorization based on the service discovery architectures is given below.

(iii) directory-based with overlay support architecture
(iv) directory-based without overlay support architecture
(v) directory-less with overlay support architecture
(vi) directory-less without overlay support architecture

Most of the SDPs are in the category (iv), which seems natural for wireless mobile ad hoc networks. In spite of lot of research work most of the protocols are still in their initial phase of research and have only been verified using simulation studies. Very few have been implemented but just using a couple of devices. We strongly feel that there is a lot of potential in category (i), the directory-based with overlay support architecture for having scalable practicable real implementation of SDPs in MANETs. The reason being that in real world there are mobile nodes with varying degrees of mobility and with varying degrees of resources. A real world mobile ad hoc network may consist of mobile phones, PDAs, laptops and even we can include desktops, which are most of the time immobile and just can leave or enter a system. We observe an inverse relation between the mobility of a device and resources it has and the services it can offer. A mobile phone although less in resources or services to offer is much more mobile as compared to a laptop which is less mobile but has much more services to offer and also have lot of resources. Normally the protocols in the category (iv) consider all nodes having very few resources and therefore propose solution that does not pose any overheads on the protocols that is, having a directory-less without overlay support architecture. But this architecture has not been successful in providing a scalable and a practicable solution. Our position is that a more practicable solution for large scalable mobile ad hoc networks is only possible with directory-based and forming some sort of overlay structure. Presence of directory decreases the latency time for service discovery and service invocation. An overlay structure is helpful for having controlled multicast, thus helping in developing scalable protocol. SDPs with directories and also having an overlay structure clearly require more resources and may not be as lightweight as SDPs in category (iv). We can get rid of these limitations if we also include nodes that have more resources like laptops and even desktops (which although are not mobile but can be included in ad hoc category as they can join and leave the system).

We found that there are not many protocols that discuss the security aspect of SDP. Any SDP if it has to be practicable cannot ignore the security aspect. This is another area of research that can be pursued in the domain of SDP for MANETs.

Mobility is an important dimension in SDP. We found that there are mainly three ways that are used to handle mobility. These are:

(i) Updating service information
(ii) Advertisement controls
(iii) Algorithms that maintain the structure of overlay network

Most of the protocols use either one of these methods. We think this is another area of research that can be probed into for finding ways to improve the mobility support by using some intelligent technique based on all these three methods and even some other method.

Service discovery is an important and an active field of research. Especially in the domain of Mobile ad hoc networks, which is also a very active field of research, the importance of service discovery protocols is even more. Still there are many open problems that need to be addressed before SDPs can be made practicable.

References

1. L. Cheng and I. Marsic. Service discovery and invocation for mobile ad hoc networked appliances, December 2000.

2. Dipan Chakraborty, Anupam Joshi, Tim Finin and Yelena Yesha. GSD: A .novel group-based service discovery protocol for MANETs. In 4rth IEEE Conference on Mobile and Wireless Communications Networks (MWCN). IEEE, September 2002.

3. Olga Vladi Ratsimor, Dipanjan Chakraborty, Anupam Joshi, Timothy Finin. Allia: Alliance-based service discovery for ad hoc environments. In ACM Workshop on Mobile Commerce WMC`02, September 2002.

4. S. Helal, N. Desai, V. Verma and C. Lee. Konark- a service discovery and delivery protocol for ad hoc networks. Proceeding of the Third IEEE Conference on Wireless Communication Networks WCNC, March 2003.

5. Michael Klein, Birgitta Konig-Ries and Philipp Obreiter. Service rings – a semantic overlay for service discovery in ad hoc networks. In DEXA Workshops, pages 180-185, 2003.

6. Michael Klein, Birgitta Konig-Ries and Philipp Obreiter. Lanes – a light weight overlay for service discovery in mobile ad hoc networks. Technical Report 2003-6, University of Karlsruhe, May 2003.

7. Ulas C. Kozart and Leandos Tassiulas. Network layer support for service discovery in mobile ad-hoc networks. Proceeding of IEEE/INFOCOM-2003, April 2003.

8. Feng Zhu, Matt Mutka and Lionel Ni. Splendor: A secure, private and location-aware service discovery protocol supporting mobile services. Proceedings of the First International Conference on Pervasive Computing and Communication PerCom`03, Pages 235-242, 2003. ACM Press.

9. Alex Varshavsky, Bradley Reid and Eyal de Lara. A cross layer approach to service discovery and selection in MANETs. January 2004.

10. Jerry Tyan and Qusay H. Mahmoud. A network layer based architecture for service discovery in mobile ad hoc networks. CCECE 2004, May 2004 IEEE.

11. Vincent Lenders, Martin May and Berhard Plattner. Service discovery in mobile ad hoc networks: A field theoretic approach. Pervasive and Mobile Computing 2005.

12. Dipanjan Chakraborty, Anupan Joshi, Yelena Yesha and Tim Finin. Toward distributed service discovery in pervasive computing environments. IEEE Transactions on Mobile Computing, February 2006.

13. Toh, C.K., "Ad Hoc Mobile Wireless Networks. Protocols and Systems", Prentice Hall PTR, New Jersey, 2002, pp. 231-242.

14. Chunglae Cho and Duckki Lee. Survey of service discovery architectures for mobile ad hoc networks. Unpublished 2005, Computer and information sciences and Engineering Department, University of Florida Gainesville, USA.

15. Raluca Marin-Perianu, Pieter Hartel and Hans Scholten. A Classification of service discovery protocols. Unpublished June 2005.

16. Feng Zhu, Matt W. Mutka and Lionel M. Ni. Service discovery in pervasive computing environments. IEEE Pervasive Computing, October 2005

17. Jini Technology Core Platform Specification, v. 2.0, Sun Microsystems, June 2003 www.sun.com/software/jini/specs/core2_0.pdf.

18. UPnP Device Architecture 1.0, UpnP Forum, Dec. 2003 www.upnp.org/resources/documents/CleanUPnPDA10120031202s.pdf

19. Salutation Architecture Specification, Salutation Consortium, 1999

20. E.Guttman et al., Service Location Protocol, v. 2, IETF RFC 2608, June 1999 www.ietf.org/rfc/rfc2608.txt

21. M. Nidd, "Service Discovery in DEAPspace," IEEE Personal Comm., August 2001, pp. 39–45.

22. Specification of the Bluetooth System, Bluetooth SIG, Feb. 2003.

23. http://www.daml.org

24. http://www.w3.org/TR/owl-features/

25. http://www.w3.org/RDF/