Multicast Service Discovery Profiles for Deployment within Dynamic Edge Networks

Joseph Macker, Justin Dean
Naval Research Lab
Washington DC, USA
joseph.macker@nrl.navy.mil
jdean@itd.nrl.navy.mil

Ian Taylor, Andrew Harrison
School of Computer Science
Cardiff University, Cardiff, UK
Ian.J.Taylor@cs.cardiff.ac.uk
and A.B.Harrison@cs.cardiff.ac.uk

Abstract—Deploying survivable group-oriented applications and enterprise services across network infrastructures incorporating the use of mobile ad hoc edge networks is of interest in military communication scenarios. This requires a more flexible approach to service discovery than conventional solutions typically provide. In this paper, we discuss our service discovery design extensions leveraging improved multicast capabilities in multi-hop, wireless networks. The design also supports multiple discovery profiles, through the use of flexible timing parameters, caching and operating modes. We also provide an overview of our working software prototype and methodology used in examining scenario-based system performance. Our experiments investigate various service provider and consumer distributions across a set of mobile ad-hoc network scenarios. Our experimental scenarios involve node mobility and varying degrees of temporal connectivity. We illustrate the differences in success rates, time delays and network overhead of several discovery profiles. Results indicate that a one-size-fits all profile is not practical and that a flexible means of deploying specific discovery profiles for different applications is needed to optimize performance.

I. INTRODUCTION

There has been extensive research work related to dynamic wireless networking, including the often cited case of mobile ad hoc networks (MANETs) [1]–[3]. Significant DoD, academic, standards, and industry focus has been placed on the development and testing of routing protocol variants for MANET environments. Much of this work focuses on the development of unicast routing capabilities and to a lesser extent on multicast routing and forwarding [4]–[6]. There are now a myriad of MANET routing solutions to choose from and applied deployment experience has improved (e.g., operational community networks [7]). However, far less research has been focused on upper layer distributed network services and collaborative applications operating within such networks. Here we address a related technology gap by examining enhanced network service discovery methods for use within tactical edge networks. Our goal is to develop and identify decentralized and mobility tolerant service discovery solutions that can operate effectively within tactical edge networks. We are also interested in supporting interconnection between fixed and mobile network architectures as depicted in Figure 1.

Use cases cover a broad range of possibilities, not only in our focus area for tactical edge military applications, but also in related areas, such as medical emergency scenarios and disaster response. Mobile network scenarios range in infrastructure types from highly autonomous MANET operations to the hybrid use of unidirectional satellite links and cellular systems. A general design challenge is to operate a distributed network service layer above a physical network architecture that is subject to dynamics, as shown in Figure 2. Since service discovery in mobile edge networks require additional design considerations [8], its mechanism benefits from a variety of operational modalities (unicast, multicast, reactive, i.e. solicited service advertisement publishing, proactive, i.e. unsolicited service advertisement publishing) and configuration flexibility (e.g. retransmission timers, data cache settings).

In the application space, there are a number of motivational examples illustrating the need for service discovery within MANET networks. Dynamic situational awareness and collaboration, through effective discovery and sharing of your operational environment and related services, is a key motivational concept. As an example, mobile platforms and assets may dynamically publish local services, discover and maintain other relevant mission and application services, and gain access to other collaborative services. Collaborative applications may include: chat, VoIP, command and control for providing orders, alarms, reports; data services, such as in-theatre video, real-time sensors (e.g. weather, sensor data, maps), access to caches of historical data, logistical support; and composable application and workflow support that benefits
Dynamic Wireless Network: dynamic routing, transport, mobility, environment

Churn consumers producers

Services Overlay: failures, collaboration groups, replicated services

from on-the-fly collection and management of services and data. The middleware used to exchange messages may also vary, depending on the nature of the application (e.g. XMPP [9] for collaborative systems and WS-Notification [10] for publish/subscribe systems.)

In this paper, we consider two example notional use cases that require different combinations of service providers and consumers:

* **Hierarchical Command and Control, Single Provider Service** – involves multiple consumers and typically one provider of information, rising to several depending on the specific application. Here, notional force order updates (data) are issued by a single point and consumed by many participants. Consumers may need to discover and maintain access to the command and control service or service components, as the network topology changes. This service type can also represent other in-theatre dynamic services such as real-time video feeds making them more useful in distributed environments.

* **Peer Group Situational Awareness and Collaborative Services** – an example of multiple consumers and multiple providers. The collaboration could involve service components: database exchanges, cataloging of service locations (DA), tracking information, real-time data streams, sensor data sharing, or even locations of media files. For our experiment presented in Section IV, we focus on showing one consumer interacting with multiple providers within a situational awareness use case, and observe how discovery profiles impact the results.

These two examples provide two diverse cases and are not intended to be all inclusive but indicate examples where reactive or proactive discovery modes might be appropriate. For example, in cases where service consumers far outweigh service providers (e.g. our command and control use case), we show that proactive or mixed mode discovery offers a more robust approach. When there is more availability of service providers (e.g. our situational awareness use case), we show that a reactive approach performs equally as well with potentially better response times. Collectively, these experiments demonstrate that a single discovery approach with fixed parameters is not effective across all possible combinations of network mobility and usage scenario. We also vary as independent variables the degree of connectivity and mobility of these two use cases to illustrate how parameter choice and discovery mode influences performance.

The rest of the paper is organized as follows. The next section provides an overview of the enhancements on existing discovery solutions that we focus on in this research. Section III describes the design and implementation of the resulting discovery system that encompasses these enhancements. Section IV shows results for the two deployment scenarios, performed for multiple levels of mobility and connectivity. In Section V we summarize our findings and state future work to be undertaken in the area.

II. DISCOVERY ADVANCEMENTS OVER PREVIOUS WORK

A combined solution of MANET-appropriate multicast forwarding and multicast-enhanced service discovery can improve wireless resource usage, system survivability, and mobile collaborative interaction for a myriad of applications. Our work has resulted in the service discovery software toolkit shown in Figure 3 that can deployed across standard IP multicast capable networks. We call our prototype the Independent Network Discovery Interface (INDI)

1. INDI means data, information, indication or pointer in Latin

2. SMF Draft is at https://tools.ietf.org/html/draft-ietf-manet-smf
discovery messaging within the control plane mechanism [16]. We do not suggest a particular approach to the message forwarding here, as there are many valid options and INDI supports standard IP multicast traffic but can also be adapted for middleware encapsulation. Rather, we focus on examining the use of such techniques, whether encapsulated or not, to enable the potential for more robust distributed querying, response, and proactive notification.

III. INDI OVERVIEW

The present INDI implementation is toolkit for network service discovery and notification. As illustrated in Figure 3 INDI provides several layers focusing on different levels of the discovery stack. Our motivation for developing INDI is to achieve a modular research platform for investigating lightweight, scalable and flexible service discovery paradigms within a variety of network architectures, focusing primarily on its use in dynamic, mobile environments. To this end, INDI supports profiles that can contain one of three modes of discovery in order to address different network dynamics: Reactive, Proactive and Opportunistic Caching.

To support these three modes of operation, INDI leverages some design elements from SLPv2 [14]. INDI uses SLP’s core roles: User Agent (UA) — a service consumer; Service Agent (SA) — a service provider; and Directory Agent (DA) — a cache/proxy for service advertisements. In the INDI design, each role (UA, SA, DA) is configured using one of the three operational modes. Each mode is constrained by a few important parameters: timeouts — a list of exponentially increasing wait intervals; required results – the number of successful requests the UA requires; and maximum retries – the number of retries to attempt per request. These parameters define fault tolerance behavior beyond any provided by the underlying network stack, which provide resilience to the dynamically changing connectivity and mobility patterns.

Shown in the lower part of Figure 3 is a pluggable interface for supporting multiple unicast and multicast protocol stacks. In this paper we use the SMF protocol to provide multicast forwarding. No additional reliable transport protocol was used here but we are considering protocols such as NORM [17] for potential reliable multicast and unicast improvements.

A. Reactive Mode

The reactive mode provides a more conventional means of service discovery. A consumer dispatches a service request using a query (containing attributes or a template for matching purposes) and providers respond with a service advert, which contains service endpoint identification and any additional metadata related to that service. In brokered systems, the service contact might be a centralized registry which may differ from the service endpoint location. Many discovery subsystems e.g. Jxta [18] and Jini [19], use a combination of unicast to contact directories and multicast to search multicast groups for peers that have services on the network. For this paper, we focus on the MANET case so there is a greater interest in the decentralized multicast discovery mechanisms.

Figure 4 illustrates this mode for a MANET. An INDI UA performs a query in a standard fashion for a decentralized network, i.e., using multicast, and then performs a retransmission backoff algorithm defined by the timeouts parameter. For these experiments the initial wait time is set to 100 milliseconds (set to around twice a typical response delay in our test environment) and exponential retransmission backoff is performed beyond this using a maximum retries setting of 5. This mechanism is analogous to the convergence algorithm used by SLPv2 and supports the inclusion of “previous responders” in the query, allowing providers to know whether a consumer has already received a response from them. In addition, INDI supports a retry field representing the times the consumer has attempted a request within a single retransmission session. This allows providers to not re-send a response to a particular retry even if the consumer never received their response was therefore unable to insert them as a previous responder.

In the experiments, a service discovery event is not restricted to the process of finding a matching service advertisement. The additional step of actually attempting to connect the service provider described in the service advert is performed. The outcome of the connect process encompasses a total discovery process from discovery through the connect stage. This feature is important in dynamic environments where service adverts are more likely to become stale due to dynamics.

Therefore, if a matching service advert is received, a connect thread is initiated. The connection process in our experimental model uses the same exponential backoff algorithm as the querying process. It determines how many providers it can connect to at each timeout and attempts as many as it can. In our simulations connection interval is set to 100 milliseconds.
Hence, if the amount of time allocated to the connection thread is 400 milliseconds, then up to four providers can be connected to in that time period. Of course the actual number of providers to attempt is further bounded by the actual number of appropriate service adverts known. The results described in Section IV show successful retries that result in an actual service connection. Corresponding discovery rates are also shown for comparison.

B. Proactive Discovery

In proactive discovery the providers, SAs and DAs, perform a service push by periodically sending out their service adverts using multicast. These adverts are delivered to all consumers subscribed to the service discovery multicast group, as illustrated in Figure 5. The effectiveness of the service push periodicity is dependent on the dynamic nature of the network; more frequent pushes may deal with increased dynamics at the expense of increased overhead. INDI UAs operating in the proactive mode use an internal cache, which caches service adverts received from producers.

When a UA invokes a service request, it first checks its local cache for valid matching service adverts. If matching services are found it attempts to connect to a service (discovery time 0), otherwise it sleeps and waits for the next interval before checking again. This continues until the maximum number of retries are exhausted or enough service connections are successful. The primary difference between reactive and proactive mode in terms of consumer behavior is that consumers do not issue service requests during the convergence process in proactive mode, instead they wait to receive adverts from providers.

C. Opportunistic Caching or Hybrid Discovery

The distributed opportunistic caching scheme is based on the concept that if a consumer queries for a service, other consumers may be interested in that service and can opportunistically cache them for future use. Opportunistic caching UAs will automatically cache all adverts overhead on the network. When an opportunistic caching UA requires a service, it will first check its local cache for valid matching service adverts. Like proactive mode, if there are matching adverts available, it will attempt to connect to the appropriate end point. Otherwise, it will revert to reactive mode and begin the reactive multicast service request process. The main difference between opportunistic caching and reactive is that using opportunistic caching, SAs and DAs will reply with multicast service adverts, as opposed to a unicast adverts. This allows any other opportunistic caching UA to potentially overhear and cache that information for possible later recall.

IV. SIMULATIONS FOR APPLICATION SCENARIOS

To capture the effects of two scenarios identified in Section I, we chose a testbed of 50 mobile nodes and performed simulations for both scenarios with: 1 provider and 25 consumers (Command and Control Scenario) and 25 providers and 1 consumer (Situational Awareness Scenario). In the simulations, all service providers are equal, nodes never host both consumers and providers (for modeling purposes to force non-local discovery) and the selection of services and consumers is consistent across scenarios regardless of motion scenarios and service discovery settings. We tested the following INDI discovery profiles:

- Opp Cache: opportunistic caching mode where each service expires after 9 seconds.
- Opportunistic Cache 54 Sec: opportunistic caching with each service lifetime set to 54 seconds.
- Proactive Cache 3 Sec: proactive mode where each node multicasts their service onto the network every 3 seconds.
- Proactive Cache 18 Sec: proactive mode where the adverts are propagated every 18 seconds.
- Reactive: Reactive mode

For Proactive, the service lifetime is set to three times the periodicity of the service pushes onto the network. Therefore, the opportunistic caching and proactive schemes employ the same service lifetime for comparison; that is, 9 and 54 seconds.

Aside from the two core scenarios and modes, we also wanted to explore how different mobility and connectivity profiles would impact the results of the experiment. For mobility, we chose a modified Random Walk motion model which independently adjusts speed and directional heading (+/- 1m/s and +/- 15 degrees), with limited scope (0-5 m/s), for each node at every selected time interval (1 second). This approach maintains a more uniform distribution of nodes within in a bounded grid space over often-used random motion techniques like the random vector waypoint model [20], [21]. For “connectivity”, we used a normalized measure of network connectivity to categorize different random mobility scenarios. A time snapshot of the mobile network was taken at regular intervals, represented by undirected graph $G$. We defined...
mobile network connectivity \((N_c)\) metric as the average expected fraction of the network reachable from a randomly selected node, measured across all topological time intervals, as follows:

\[
N_c = \frac{1}{\sum_{n=1}^{N} |C_n|/N^2} \sum_{n=1}^{N} |C_n|/N^2
\]

where \(|C_n|\) is the order of the connected component in \(G\) containing node \(n\), \(T\) is the number of uniform time intervals within the mobility scenario \(300\) in our case, and \(N\) is the number of nodes or vertices, \(V\), within \(G\). A \(N_c\) value of 1.0 (100%) is a fully connected network at all times. Lesser values indicate partial network connectivity for at least some time portion of the simulation. Using the \(N_c\) metric and we generated ten separate random mobility scenarios with resulting connectivity coefficients of 60, 70, 80, 90, and 100, resulting in 50 mobility profiles per scenario. Mobility profiles with higher connectivity coefficients were generated using smaller mobility areas resulting in higher network densities than those with lower coefficients. This allowed settings for Random Walk node mobility generation to be constant across all 50 mobility profiles, resulting in similar average link change rates \((1.5\) to \(2.0\) percent per second) for each profile.

INDI used the Agentl [22], [23] toolkit to execute the unmodified Java INDI code within the ns2 environment [24] (i.e. we ran actual code rather than discretizing our approach).

To generate the scenario files and analyze the results, we developed a software analysis package which defines the test environments, combines the appropriate mobility scripts, populates the service agents, and establishes other parameter settings. Consumer discovery events were generated randomly using a series of event times according to a Poisson distribution with a mean interval of 6 seconds.

![Fig. 7. Legend for 8, 9 & 10](image)

**Fig. 7.** Legend for 8, 9 & 10

Figure 8 shows the coupled “successful discovery and connect” results for the 1 provider and 25 consumer simulation scenario (1-25, Command And Control). It can be seen in this case that generally the proactive-based protocols outperform reactive for discovery, both in terms of success rate and overall responsiveness (delay). However, in Figure 9 which shows only the discovery rates, it can be seen that the underlying multicast algorithms perform far better across the board. The decrease in success rates between Figures 8 and 9 is far more pronounced for the proactive-based protocols. In the case of long expiration intervals, valid discovered service advert endpoints have a higher failure rate due to network dynamics, resulting in a larger difference between discovery and success rates. For discovery and connect, the reactive protocol has the lowest success rates, while the proactive protocol with the longer \((18\) second) expiration interval comes in a close second. Both the opportunistic caching using a \(9\) second lifetime and proactive using 3 second intervals perform well with respect to success rate in our scenarios, showing their ability to respond well to the dynamics of the network. Although, the proactive protocol performs well across our test scenarios, opportunistic caching proves a competitive combination when considering both success rate and delay metrics.

![Fig. 8. Discovery/Connect Success with Delay (1 provider, 25 consumers)](image)

**Fig. 8.** Discovery/Connect Success with Delay (1 provider, 25 consumers)

<table>
<thead>
<tr>
<th>Connectivity</th>
<th>60</th>
<th>70</th>
<th>80</th>
<th>90</th>
<th>100</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reactive Mode</td>
<td>649</td>
<td>796</td>
<td>952</td>
<td>1.1K</td>
<td>1.2K</td>
</tr>
<tr>
<td>Proactive Mode</td>
<td>3K</td>
<td>3K</td>
<td>3K</td>
<td>3K</td>
<td>3K</td>
</tr>
<tr>
<td>OppCach Mode</td>
<td>295</td>
<td>386</td>
<td>535</td>
<td>599</td>
<td>717</td>
</tr>
</tbody>
</table>

For the 25 providers and 1 consumer (25-1, Collaborative) scenario, shown in Figure 10, there is little significant difference between the discovery schemes. This is somewhat intuitive since we are achieving additional robustness through an abstracted model of service replication. The provider locations are so distributed that the probability of finding one of them in a network containing 25 is high even in the more fragmented cases. However, the reactive protocol here offers the best balance between success rate and responsiveness with delay values almost half of the other protocols. However, when message overhead (combined multicast and unicast) is considered:

then opportunistic caching proves to be an excellent compromise between proactive and reactive protocols for this scenario. The results from Figure 10 show that not only does it provide comparable performance, but it also generates far less overhead than the other two protocols. Opportunistic caching generates around 40% less messages than a reactive protocol and 76% less messages than the competing proactive protocol.
With respect to providing further fault tolerance to the underlying multicast algorithm through the use of INDI “retries” in the convergence gains, we can see the following gains:

<table>
<thead>
<tr>
<th>Type</th>
<th>Retries (success rate)</th>
<th>No Retries (success rate)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-25</td>
<td>67.23%</td>
<td>54.9%</td>
</tr>
<tr>
<td>25-1</td>
<td>95.3%</td>
<td>93.0%</td>
</tr>
</tbody>
</table>

For low numbers of distributed providers, it can be clearly seen that there is a significant gain in successful discovery rates by using retries. For, 1-25 there is a 13% performance gain and for 25-1 there is a 2% gain. Our findings also show more significant gains are realized when applying retries in scenarios of lower connectivity degree. In these cases, the network forwarding is less robust and the network mobility may also become a positive factor in discovery success of servers that may be initially unreachable.

V. CONCLUSION AND FUTURE WORK

Supporting a distributed network service paradigm is vital to hiding the physical network structure and configuration from end users. Such a capability is essential in dynamic, self-organizing wireless networks in which network nodes may join, leave, fail, or move within a given architecture. Also addresses may change and services may migrate. This paper has discussed extensions to existing service discovery protocol frameworks that we feel address many issues for effective operation within wireless, multi-hop mobile network architectures. We have shown that discovery flexibility, in terms of operating modes and parameter settings, is vital for appropriately operating service discovery in different deployment profiles, degrees of connectivity, and mobility patterns. Particularly, early findings indicate that proactive profiles are a better fit when demand far exceeds supply and that an opportunistic caching scheme proves both efficient in terms of success rate and message overhead for scenarios where supply far exceeds demand.

Beyond our initial design and findings there is much that remains to be explored. One area involves modeling additional system heterogeneity in terms of service types and network resources beyond our present models. It is also presently not well understood if or how reliable transport protocols, directory agents, or service proxying techniques provide useful performance gains in MANET type service discovery architectures. Through future studies we hope to document more about the architectural tradeoffs and design options to aid future system designs. Finally, we are also migrating INDI to support interfaces to other service discovery frameworks (e.g., WS-Notify, mDNS-SD) to support interoperable system capabilities.

ACKNOWLEDGMENT

INDI has been developed by the Networks and Communication Systems Branch of the Information Technology Division (ITD) at NRL. On-going modifications to the core system are being funded by the Office of Naval Research (ONR) Service Oriented Networks Operating in MANET (SONOMA) project. The authors would like to thank Brian Adamson for his advice and for the use of his SMF implementation.

REFERENCES