Abstract— An ever-growing demand for coverage extension, higher data-rates and improved connectivity has motivated the interconnection of mobile ad-hoc networks to fixed IP networks, to form a Hybrid Wireless Network (HWN). The convergence of ad-hoc and infrastructure network is also attractive in real-world scenarios due to its practicality and usefulness. In HWNs, gateways are generally used as an interconnection point for the integrated network. In order to improve the overall performance of HWN, an efficient gateway discovery and route maintenance scheme is indispensable. In this paper, we present a novel adaptive gateway discovery scheme called ADD (Adaptive Distributed gateway Discovery), in which the adaptation is done in a fully distributed manner. Simulation results show that the distributed adaptation capabilities of our proposed scheme facilitate outperforming the existing gateway discovery mechanisms in terms of control overhead, packet delivery and delay.

I. INTRODUCTION

In Hybrid Wireless Networks (HWNs), we aim at the integration and coexistence of two wireless paradigms i.e. infrastructure-based (cellular/data) and ad-hoc (spontaneous), in order to offer higher flexibility and improved performance. HWNs not only preserve the benefits of conventional infrastructure-based networks, where the service infrastructure is constituted of fixed entities (gateway), but it also incorporates the flexibility of ad-hoc networks, where wireless transmission through mobile stations in multihops is allowed.

As shown in Figure 1, a hybrid wireless network is a mix of infrastructure and mobile ad-hoc networks. Gateway (GW) facilitates the Mobile Nodes (MN) to connect with the core IP network. MN, which is not in the direct coverage of GW, discovers the route to GW in multi-hop fashion i.e. passing through other MNs in order to connect to the infrastructure network (Internet, for example). Each set of nodes covered by a GW form an ad-hoc network. In this paper, we frequently refer to GW as a means to connect to the Internet, although the GW is flexible enough to offer a large range of other services as well.

For the mobile nodes to connect to the Internet, a gateway discovery mechanism is required in order for the nodes to discover the route towards the gateway. As compared to infrastructure-based networks, where the gateway is always at a single-hop, the problem of gateway discovery in hybrid networks is challenging and node mobility, scalability of ad-hoc networks etc. makes it even more complex [1].

The intended contributions of this paper are twofold. First, we describe and discuss the operation of most well-known approaches towards gateway discovery. Second, and the most important, we present an adaptive gateway discovery scheme called ADD (Adaptive Distributed gateway Discovery), in which the adaptation is done in a fully distributed manner. To the best of our knowledge, this scheme is a first step towards the distributed adaptation of gateway discovery in HWNs, which results into significant performance gain, as validated by simulation results.

The remainder of this paper is organized as follows. In Section II, we present the state of the art achievements in gateway discovery mechanisms and also implicitly highlight our motivation for a new adaptive distributed gateway discovery scheme. After that we propose our novel scheme called Adaptive Distributed gateway Discovery (ADD) in Section III. Furthermore, simulation-based performance evaluation of ADD scheme is presented in Section IV. Finally, we conclude the paper in Section V.

II. MOTIVATION AND RELATED WORKS

Gateway discovery is one of the major components in order to realize a Hybrid Wireless Network. Gateways are generally the specialized nodes, which lie in the ad-hoc network and also have connectivity with the fixed network, such as Internet. With the help of gateway discovery mechanisms, nodes in the mobile ad-hoc network discover the route towards the gateway for different reasons, for instance, to communicate with a node in the fixed network or to access the Internet or Intranet.

In the existing literature, different proposals for gateway discovery have used either reactive or proactive mechanisms.
[1]. In proactive approaches, the gateway periodically sends the Gateway Advertisements (GWADV) messages, which are flooded throughout the entire ad-hoc network. On the other hand, in reactive approaches, the nodes which require connectivity to the gateway broadcast the Gateway Solicitation (GWSOL) messages. These solicitation messages are responded by the gateways.

In the recent studies, hybrid schemes have also been proposed. In hybrid gateway discovery, part of the nodes in ad-hoc network discovers the gateway proactively, whereas the rest do it reactively [2]. As shown in Figure 2, all the nodes which are located in two hops (i.e. \(TTL=2\)) of the gateways periodically receive the GWADV from the gateway whereas the other nodes reactively discover the gateway. Although, in certain cases, a hybrid scheme outperforms the reactive and proactive mechanisms [4]; the definition of a hybrid scheme is quite vague i.e. defining the optimised \(TTL\) value. Therefore, the hybrid scheme should incorporate a certain level of adaptation in order to dynamically respond to the network changes [5]. Moreover, in a medium to large scale network, the overhead to maintain the routes towards the gateway is much more expensive then initially discovering the route. To this end, an adaptive gateway discovery is a simple but powerful concept, which pledges to offer an efficient and cost-effective gateway discovery and maintenance scheme. In [4], the gateway is initially discovered reactively, furthermore for route maintenance the scope of the gateway advertisements i.e. \(TTL\), is adapted according to the active sources in the network. By simply looking at the IP header of the data packets, the gateway keeps track of the number of hops at which each of its active sources is located. The gateway adaptively selects the \(TTL\) value such that all the active sources receive the periodic advertisements from the gateway. This latter approach is referred as Maximum Source Coverage (MSC). Figure 3 presents a hybrid networking scenario with the MSC scheme for gateway discovery. Here, nodes \(M\) and \(N\) are data sources which are active and sending data to some node in the fixed network (Internet). Initially, the active sources reactively discover the routes towards the gateway and initiate the data transfer. By analyzing the data packet, the gateway keeps track of the number of hops of each of these data sources are far from the gateway, for instance, in this example \(M\) is at 3 hops and \(N\) is at 2 hops. As shown in Figure 3, since source \(M\) is farthest away from the gateway i.e. 3 hops, the gateway sets \(TTL=3\) for GWADV in order to cover all the active sources. Now, in spite of sending periodic advertisements to all the nodes in the network, the gateway only covers the active sources which are located within 3-hop radius.

MSC may suffer from suboptimal performance in certain networking scenarios such as when a small number of sources are at a large number of hops from the gateway. In all the adaptive gateway discovery schemes presented in the literature [5-7], the gateway stores the information of active connections i.e. the address of data sources and number of hops, these sources are away from the gateway. To this end, the adaptation is done centrally at the gateway. In a more dynamic network such as HWN, where the network conditions change drastically, the central adaptation may result into far from optimal performance. It is quite possible that one part of the network has considerably large number of active sources than the other. Therefore, one adaptive \(TTL\) value may be useful for a certain sub-network but not for the whole network. Inspired by “think-global and act-local” principle, distributed adaptation should be considered, where the decisions are taken locally and in a totally independent fashion, according to the local-view of the network with the objective of achieving global performance.

III. A NEW SCHEME: ADAPTIVE DISTRIBUTED GATEWAY DISCOVERY (ADD) FOR HWNS

Adaptive Distributed gateway Discovery (ADD) algorithm is a fully distributed approach towards gateway discovery in HWNs. The ADD algorithm is inspired by a hypothesis that the gateway advertisements should only be targeted to those nodes, which are looking for the gateway (such as active data sources); and other nodes should not be hampered with the periodic gateway advertisements. However, unlike other adaptive discovery mechanisms where the adaptation is performed centrally at the gateway and a single \(TTL\) value is used for the whole network, we advocate a totally distributed approach in
our ADD scheme. In addition to the support of gateways, ADD also involves the participation of common nodes (called intermediate nodes) which are used as relays to forward the data packets from the source towards the gateway, for an adaptive gateway discovery. In other words, intermediate nodes act as multicast group members in order to forward the gateway advertisements only to the active sources in a multicast fashion. Therefore, the TTL value for GWADV is adapted at each hop in the network towards the active data source in a distributed manner, in order to cover all the active data sources with minimum annoyance for the other nodes in the HWN.

In the ADD algorithm, initially, the potential source node reactively discovers the gateway. Once the route towards the gateway is known, the source node starts sending the data packets. By simply looking at the IP header of the data packets, the gateway keeps track of the number of hops at which each of its active sources is located. Based on this information, a source table is constructed at the gateway. Apart from this, each intermediate node between the source and the gateway also captures the following information from the IP header of the data packet: the number of hops the source is away from the intermediate node and the source address. With the help of this information, the corresponding source node is marked as "active" in the routing table of the intermediate node. As a result of data packet analysis, those nodes, which are relaying the data packets learn that they are intermediate nodes for some given source in the network. This information is further leveraged in the gateway rediscovery and route maintenance scheme.

During the periodic gateway advertisement, initially, the gateways send the GWADV message with TTL=1. Then, on the reception of the GWADV message, each node verifies whether it is an intermediate node. If the node finds that it is already relaying data of an active source towards the gateway, it learns that it is an intermediate node. In case if it is an intermediate node, it forwards the GWADV message further with TTL=1, otherwise it does nothing. Following this procedure, all the intermediate nodes broadcast the GWADV message. This distributed gateway discovery adaptation ensures that all the nodes (active sources) which are actively sending data towards the gateway and are actually willing to maintain the route towards the gateway, receives the GWADV message, traversing hop-by-hop through all the intermediate nodes. Other nodes in the network, which are not interested in knowing the route towards the gateway, are not hindered with GWADV messages; therefore the control overhead is significantly decreased. This per hop distributed adaptation for gateway discovery and route maintenance is the essence of our proposed scheme.

Since in our algorithm, GWADV is broadcasted with TTL=1 by the intermediate nodes, all the one-hop neighbours of the intermediate nodes (either active sources or not) receives the advertisement messages. This later results into the formation of an active region comprised of all the nodes between the gateway and the active source. We formally define an active region as a logical sub-network which has at least one active source, and where all the nodes know the route towards the gateway. Apparently, the GWADV messages received by nodes in an active region, which are neither active source nor intermediate node, seems "excessive overhead" because these nodes are not interested in localizing the gateways. However, in case of high mobility, these nodes may help the active sources in order to adapt to the continuously changing network topology. As the active source moves, the intermediate nodes also change which eventually results into the movement of the active region. To this end, an active source always remains intact with the active region, and the active region adapts itself according to the movement of the active source. During the dynamic changes in the network, if an active source looses a valid route to the gateway, it can query any of the active region nodes to learn an updated route towards the gateway. This characteristic of ADD helps the active sources to quickly learn the updated route towards the gateway in a dynamic network, without generating an excessive overhead.

The ADD scheme is further explained with the help of rectangular lattice as shown in Figure 4. We assume that the nodes are uniformly distributed in a rectangular lattice covering a certain area. Each vertex of the lattice is a possible location for a node, however only one node can be at a concrete vertex. A "red" node in the middle of the rectangular lattice represents the gateway, whereas "black" nodes represent the active sources and "white" nodes are the intermediate nodes for the given active sources. The established (active) routes between the gateway and the active sources are shown in "solid" lines. These routes are established initially, with the help of reactive gateway discovery. Once the data flow is established, the data packets are analysed and intermediate nodes are marked. For the gateway rediscovery and route maintenance process, the gateway broadcasts the GWADV within its one-hop neighbours. Furthermore, only the intermediate nodes broadcast the GWADV within its own one-hop neighbourhood. By doing so, not only all the active sources receive the GWADV message but also the active regions are formed. In the lattice shown in Figure 4, the shaded regions represent the active regions formed by the intermediate nodes between the active sources and the gateway. The concept of active region introduced in ADD scheme can be very useful in case of high mobility, network partitioning and for load balancing.
IV. IMPLEMENTATION AND PERFORMANCE EVALUATION

In this section, we evaluate our proposed ADD scheme. For this performance evaluation, we conduct extensive simulation under a variety of WiFi-based multi-hop hybrid networking scenarios.

A. Simulation Methodology

It is difficult to capture the details of gateway discovery schemes in an analytical model, since the real-time behavior of the multi-hop hybrid network can not be accurately quantified. For that reasons we evaluate and analyze the performance of ADD in the 802.11-based hybrid network, using the ns-2 simulator snapshot ns-2.29. We use locally modified and extended version of ad-hoc routing protocol AODV called AODV+ [9], which implements the interconnection between a MANET and the Internet. AODV+ already provides the hybrid gateway discovery scheme. We further extend the AODV+ with our proposed adaptive distributed scheme i.e. ADD and a centralized adaptive gateway discovery scheme such as MSC. The radio channel capacity for each mobile node is 2 Mbps, using the IEEE 802.11b DCF MAC layer and a communication range of 250 m. In addition, there are two gateways; located at the coordinates (100, 250) and (700, 250) respectively. In the hybrid and MSC approaches, both gateways use TTL=2 for gateway advertisements as it is recommended in [2].

For our simulated scenarios, we use a random waypoint model. During initial phase of simulation, the nodes are static for pause-time seconds. Then, they pick up a random destination inside the simulation area and start moving to the destination at a speed uniformly distributed between 0 and 10m/s (maximum-speed). In the simulated scenarios, some mobile nodes are randomly selected as active sources, which communicate with some fixed node located across the Internet. The traffic under consideration is Constant Bit Rate (CBR) traffic. Each of these randomly selected CBR sources start sending data at a uniformly distributed time within the first 10 seconds of the simulation and lasts until the end of simulations (i.e. 300 secs). Each of the sources generates 512 bytes data packets at a rate of 5 packets per second (20 Kb/s). The simulation model parameters employed in our study are summarised in Table I.

B. Performance Metrics and Evaluation Model

The performance evaluation of reactive, proactive, hybrid and adaptive gateway discovery schemes has been extensively simulated and evaluated in the past. The lessons learnt from these studies [4] [10] motivated us to design a novel adaptive network/gateway discovery scheme, where the adaptation is performed in a totally distributed and decentralized manner. In order to evaluate and assess the effectiveness of the different gateway discovery mechanisms compared to our ADD scheme, the performance metrics which have been evaluated are packet delivery fraction, average end-to-end delay, and normalized routing load.

To the best of our knowledge, ADD is a first step towards the distributed adaptation for gateway discovery and route maintenance in HWNs; therefore we could not compare it with any other distributed adaptation approach. To this end, we compare ADD with a hybrid gateway discovery scheme and with a centralized adaptation scheme such as the MSC approach.

<table>
<thead>
<tr>
<th>Simulation / Scenario</th>
<th>MAC / Traffic / Gateway</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simulation Time</td>
<td>300s</td>
</tr>
<tr>
<td>Mobile Nodes</td>
<td>10 – 90</td>
</tr>
<tr>
<td>Traffic Sources</td>
<td>5 – 45</td>
</tr>
<tr>
<td>Mobility Model</td>
<td>Random way</td>
</tr>
<tr>
<td>Maximum Speed</td>
<td>5 m/s</td>
</tr>
</tbody>
</table>

C. Simulation Results

The first simulation scenario measures the performance of the scheme in terms of variable mobility patterns (i.e. changing pause-time). Seven different pause-times were used: 0, 15, 30, 60, 120, 240 and 300 seconds. A pause-time of 0 seconds corresponds to a continuous motion whereas a pause-time of 300 seconds corresponds to a static scenario. This behaviour is repeated several times for the whole duration of the simulation. The first simulated scenario consists of 15 mobile nodes randomly distributed over an area of 800x500 m², out of which 5 nodes are randomly selected as active data sources. The performance results for the first simulation scenario are shown in Figure 5. The simulation results show that our proposed approach is able to offer a Packet Delivery Fraction (PDF) as high as the other approaches with a much lower overhead compared to hybrid and adaptive gateway discovery i.e. MSC approaches. In contrast, our approach experiences a slightly higher delay as compared to the other more proactive approaches, as in Figure 5.

The high packet delivery fraction of hybrid and MSC is attributed to their proactive way of advertising the gateway information throughout the network. In contrast, ADD only serves the gateway information to the active sources i.e. nodes in communication with the gateway. In Figure 5, we can intuitively see that in terms of normalized load, ADD consistently generate very less overhead and outperforms the hybrid and MSC gateway discovery mechanisms. This more than two times gain compared to MSC and more than four times gain compared to hybrid is mainly attributed to the distributed adaptation mechanism of ADD which permits to periodically inform only the active sources about the route towards the gateway. Predictably, load performance of hybrid, MSC and ADD schemes experience a steep improvement with increasing pause-time. In contrast, since both of the hybrid and MSC schemes are inherently based on the proactive gateway advertisement principles, in case of high mobility, the nodes quickly learn the updated routes towards the gateway at the expense of high overhead. Whereas, ADD broadcasts the gateway advertisement only towards the active sources, thereby the learning of new routes is comparatively slower in case of high link breaks and low node-density (i.e. 15 nodes). However, this slight increase in delay can be reduced with the help of active regions concept introduced by ADD, in considerable node-density scenario. In case of high mobility with significant node-density (reasonable assumption in
HWNs, the active sources may learn the routes towards the gateway by simply asking the neighbouring nodes within the active region. The low node-density (as studied in the first simulation scenario) influences the stability of the active regions in ADD.

To this end, the second simulation scenario assesses the performance of network discovery schemes in case of high node-density and node-mobility. The mobility model is configured at pause-time of 2 seconds and maximum-speed of 10 m/sec. The simulation area is extended to 1300x800 m², while the number of nodes in the given area is uniformly increased from 20 to 90 nodes. Moreover, half of the given total simulated nodes are randomly selected as active data sources.

The simulation results reported in Figure 6, show considerable overall gain in terms of normalized load, delay and PDF for ADD scheme under high mobility and node-density. As it can be seen from the Figure 6, ADD predictably outperforms in normalized load. In contrast, before reaching a density of 45 nodes, ADD scheme offers a slightly low PDF and high delay as compared to other simulated schemes, however just after 45 nodes, there is a noticeable improvement. Starting from a node density of 45 at high-mobility, ADD outperforms MSC and hybrid schemes in terms of delay and packet delivery fraction. This latter result validates our hypothesis that the concept of active regions introduced by ADD is very useful towards overall performance optimization in high mobility multi-hop hybrid networking scenarios. In summary, ADD significantly generates less control overhead than other existing schemes in any network scenario. Moreover, the active regions formed during distributed gateway discovery helps ADD to also outperform in terms of delay and successful packet delivery in case of significant mobility and network density scenarios.

V. CONCLUSION

HWNs are emerging as a promising new technology, benefiting from both fixed and ad-hoc networks technologies whilst alleviating some critical problems in these networks. HWNs seem attractive in opening new business opportunity for network operators and service providers, enabling commercial services provision with broad coverage on one hand and seamless mobility for mobile clients with improved overall QoS on the other hand. In this paper, we exposed the motivation and importance of deploying HWNs, while discussing the existing mechanisms of gateway discovery in the hybrid networks ranging from reactive, proactive and hybrid discovery to adaptive discovery schemes. Furthermore, we presented the ADD mechanism which is, according to the best of our knowledge, the first attempt towards distributed adaptation for discovering gateways in HWNs.

Being intelligent and capable to cope up with the ever changing conditions of the network, our distributed adaptive gateway discovery schemes outperforms the existing gateway discovery mechanisms and provides an excellent platform decreasing gateway re-discovery and route maintenance overhead, as validated by the simulations. The results show that ADD is able to offer a packet delivery fraction comparable to the other schemes, but with 3 to 4 times lower control overhead. Moreover, in case of high mobility and reasonable node density scenario, ADD also exhibits gains in terms of delay and packet delivery fraction. We are currently investigating the active region stability and group mobility scenarios, which we aim to report in our future works.

REFERENCES

Figure 5. PDF, Normalized Load and Delay with variable mobility-pattern (i.e. pause-time)

Figure 6. PDF, Normalized Load and Delay with increasing node-density, variable traffic pattern and high-mobility