On the Scalability of Internet Gateway Discovery Algorithms for Ad hoc Networks

Mona Ghassemian, Vasilis Friderikos, A. Hamid Aghvami
Centre for Telecommunications Research, King's College London, UK

***Abstract*— The aim of this work is to evaluate the scalability of methods applied to interconnect ad hoc networks to the Internet. We describe some of the solutions proposed for Internet connectivity in ad hoc networks. We define scalability space, absolute, relative and weak relative scalability terms. The scalability comparison of these mechanisms is presented by means of analytical modelling with respect to different parameters such as number of mobile nodes, rate of link changes and rate of traffic sessions per each mobile node. To optimise the total amount of overhead generated by the discovery protocols, we propose a feedback control algorithm to adapt period and transmission range of gateway advertisements.**

INTRODUCTION

The main appeal of a mobile ad hoc network is its peer-to-peer communication capability in a dynamic environment without any form of infrastructure support. However, the recent widespread success of IEEE 802.11 WLAN technology in the consumer, enterprise and service provider markets demonstrates that Internet connectivity remains the primary application driver and so infrastructure support remains a key part of most wireless mobile networks.

In a multihop radio access network (MRAN), some infrastructure nodes, called internet gateways (IGWs), offer internet connection to mobile nodes (MNs) reside in an ad hoc network. We address the problem where the nodes in an ad hoc network have to discover at least one of these IGWs:

Reactive discovery mechanisms: A MN broadcasts a message throughout the ad hoc network, soliciting for a connection to the Internet. An IGW receiving this message will reply to the MN offering its IP prefix address [1].

Proactive discovery mechanisms: All IGWs periodically broadcast their IP prefix address throughout the ad hoc network. When the MN is connected to an IGW and receives an advertisement that offers a better path (e.g. with lower number of hops), it will optimise the existing path [1].

Hybrid discovery mechanisms: All IGWs periodically broadcast their IP prefix address within a proactive zone. A MN out of one of the proactive zones should solicit for a connection [2].

The aim of this work is to present a scalability evaluation of methods to interconnect ad hoc networks to the Internet. Section I gives an overview of related works to measure ad hoc protocol overhead bounds and define scalability space term. Modelling assumptions necessary for overhead analysis are described and derived in Section II. Section III of this paper presents the analysis of the total overhead generated by running each of the Internet connectivity mechanisms. The overhead bound analysis gives us the ability to check the scalability of the discovery mechanisms and compare their degree of scalability. We investigate absolute and weak relative scalability factors of the discovery mechanisms with respect to (w.r.t) different parameters in Section IV. The results show that the performance of the gateway discovery mechanisms varies in different scenarios. Therefore we propose an adaptive gateway discovery algorithm by optimising the frequency of periodical advertisements and the proactive zone of IGWs in the hybrid IGW discovery protocol in Section V. Conclusions are presented in Section VI.

I. RELATED WORKS

In this section, we point out the most related works in ad hoc protocol overhead and scalability analysis.

Our main work references [3] and [4] model a number of ad hoc routing protocols. In [3], a novel framework is presented for the study of scalability in ad hoc networks. Using this framework, the first asymptotic analysis is provided w.r.t network size, mobility and traffic for each fundamental class of ad hoc routing algorithms. Beside this paper, a report from INRIA also covers analytical model of proactive and reactive ad hoc routing protocols [5].

Most of the works on the overhead analysis focus on ad hoc networks stand alone. The only publication in hand related to theoretical analysis of ad hoc networks connected to the Internet is discussed in [6]. It is presented that in a wireless network, a relatively sparse fixed infrastructure is well suited for very populated regions, whereas pure ad hoc networking can be used for areas with a relatively low density of nodes. However, if in some regions the density is sub-critical, only cellular network can offer an acceptable connectivity.

As many research papers present the scalability of a mechanism w.r.t a given parameter in a particular environment [7], it is necessary to say that a discovery mechanism is scalable w.r.t a triple (environment, parameter, metric) space. "Independent parameters" are parameters that can be freely varied, such as number of MNs, traffic load and mobility rate. "Primary metrics" are performance quantities that are observed in the network, such as throughput and delay in the network. Finally, "environmental parameters" are parameters that define the operational conditions of the network, such as: network characteristics, node characteristics, traffic pattern, routing, MAC, PHY layer being used.

II. Modelling preliminaries

This section presents the model assumptions and definition of notations employed in our analysis. The total amount of overhead generated in the ad hoc networks is a combination of protocol dependent and protocol independent overhead. Protocol dependent overhead depends on the protocol parameter settings, such as period between routing updates in a proactive routing protocol. However a protocol independent overhead is due to topology configurations of the ad hoc network, such as the topology size and positioning of the MNs in the network.

The amount of overhead generated by the protocol dependent section depends on several independent parameters that models network, traffic and protocol features, as described below:

Network model parameters are number of MNs (n) in the ad hoc network, number of IGWs (m), average path length in hops ($\overline{L}$), maximum expected path length (L_{max}), the average rate of link changes (v) as a result of mobility, power failure,... for each active node[1], area covered by n nodes (A), network average density (σ) and neighbouring degree (d) (the degree is the number of neighbours of a node).

Traffic model parameters are the average rate of new sessions generated by each node (λ) and the average traffic rate for each node (r). It is also important to know the number of active sessions per node (s), which depends on the session duration and the traffic distribution, to calculate the expected number of nodes engaged with already started sessions.

Protocol dependent parameters are the average signalling packet size (l_{RREQ}, l_{RREP}, l_{RERR}, l_P) in bits, the rate of periodical IGW advertisement (IGW-ADV) messages (p), and the number of hops the signalling packets are forwarded (k), i.e. TTL of the packets. The model assumptions and definitions employed in our analysis are as follows:

a.1 We assume that as the network size increases, the average neighbouring degree d remains constant [3]. Imposing a fixed neighbouring degree in a network is desirable, because allowing the density to increase without bound jeopardizes the achievable network throughput. This is as a result of applying a power control mechanism that reduces the transmission range of each node. On the absence of a power control mechanism, the number of neighbours to maintain connectivity in an ad hoc network is proved to be $\Theta(\log n)$ in [8].

a.2 Let A be the area covered by n nodes of the network, an $\sigma = \frac{n}{A}$ be the network average density. Then, the expected average number of nodes inside an area A_1 is approximately σA_1 [3]. For example, it is expected that half the area covered by the network contains approximately one half of the nodes in the network.

a.3 The number of nodes that are at distance of k or less hops away from a source node S increases as [3]:

$$n_k = \Theta(dk^2) \tag{1}$$

Therefore number of hops within distance k hops of the node that broadcast a message can be modelled by n_k.

a.4 One important factor in our analysis is the source to destination node distribution. In the presented analysis, all source nodes send packets to IGWs. Therefore, the destination nodes are not equally probable. This feature makes the analysis of IGW discovery different from ad hoc routing protocols. In the reactive discovery approach, discovery of at least one of the m IGWs is desired. However in a reactive routing protocol, all the n nodes in the network can be the destination nodes. In proactive discovery approach, only m number of IGWs broadcast their advertisements through the network, while in a proactive routing protocol, all nodes broadcasts their information periodically.

To find an IGW reachability condition for nodes in an ad hoc network, we define a "reachability zone" for each IGW. A reachability zone is the area within which source MNs can connect to the Internet through an IGW. The concept of reachability does not necessarily show the connectivity. Once a node is connected, it is reachable, but when it is reachable it is not necessarily connected as there may not be any on going session, due to policy, security or QoS requirements.

Assume IGW_i^S be a set of IGWs that can be discovered by node S, considering that each IGW covers an average n_k nodes (assumption a.3) and that each node is covered directly or in a multihop manner by at least one IGW. This yields that adding up the number of nodes covered by all IGWs results in a number greater than n.

$$(\sum IGW_s^i)n_k \geq n$$

Therefore to maintain the reachability:

$$m = \Omega(\frac{n}{n_k}) \tag{2}$$

a.5 The path length (in hops) among nodes ($\overline{L}$) in a connected subset of (n) nodes both increase as $\Theta(\sqrt{n})$[2].

Assuming that IGWs are all located in the centre of topology and each IGW can be reached by $\hat{n} = \frac{n}{m}$ nodes. Thus average path length can be defined as $\overline{L} = \Theta(\sqrt{\frac{n}{m}})$.

The question is how many hops connects a source node S within the reachability zone of an IGW in worse case (L_{max}). In the proactive discovery approach, IGW-ADV messages broadcast to $\hat{n}$ nodes with $TTL = \hat{n}$. Therefore the maximum number of hops ($\sqrt{\hat{n}}$) is the total number of hops that may forward a packet to reach to an IGW. In the reactive approach we consider expanding ring search mechanism with TTL set to $k_i \in \{k_1, k_2, .., k_{Threshold}, \hat{n}\}$. As we do not consider either the intermediate nodes' reply on behalf of IGWs, or the local route repair mechanism, the maximum path length is bounded to k_i. For the hybrid discovery approach, the path length depends on the size of the proactive and reactive zones. If we consider IGWs broadcasts their IGW-ADV with $TTL = k_{pro}$ for MNs

[1] An active node is either a source, intermediate or destination of an ongoing traffic

[2] See [9] for a tighter bound analysis of path length in ad hoc networks.

reside in the proactive zone, then $L_{max} = k_{pro}$. For MNs outside this zone, the path length depends on their distance to the border nodes in the proactive zone and $L_{max} = k_{re} + k_{pro}$ where $k_{re} \leq \sqrt{\hat{n}} - k_{pro}$.
To summerise:

$$L_{max} = \begin{cases} \sqrt{\hat{n}} & \text{Proactive} \\ \{k_1, k_2, .., k_{Threshold}, \hat{n}\} & \text{Reactive} \\ k_{re} + k_{pro} & \text{Hybrid} \end{cases} \quad (3)$$

a.6 Average traffic injected per node depends on the rate of each session (r), number of simultaneous active routes per node (s) and rate of new sessions initiated per node (λ). Thus the average traffic injected per node is:

$$\overline{r} = \frac{1}{n} \sum_{i=1}^{\lambda} r_i s_i \quad (4)$$

III. Overhead definition and analysis

Measuring only the control packets generated by a discovery mechanism does not provide enough information regarding the total amount of overhead since there are other factors, such as route sub-optimality that may become more relevant as traffic or network size increases. For example, the discovery mechanism that produces less control overhead may be forming longer paths than necessary. This may not be an issue at low traffic rate, but as the traffic rate increases the extra hops may be comparable to, or even greater than, the control overhead. There is an inherent trade-off between control overhead and sub-optimal routing cost.

The different sources of overhead that contribute to the total overhead may be grouped and expressed in terms of route discovery and maintenance, and sub-optimal routing overheads. The total overhead is a combination of all of them. In this section, we present the analysis of overhead for reactive, proactive and hybrid discovery approaches.

A. Overhead in the reactive discovery mechanism

The reactive overhead of a protocol is the amount of bandwidth consumed to build paths from a source to a destination after a traffic flow to that destination has been generated at the source. The reactive overhead is a function of both traffic and topology changes. We describe all means of overhead in the reactive discovery mechanism as follows:

1) Route discovery overhead: C_{RD} presents the maximum amount of signalling overhead generated due to solicitations for discovering a route towards the Internet for each new session per node per second (λ). Each RREQ message broadcast to the entire network with $TTL = k$. Therefore route request messages are retransmitted for all the nodes that are $k-1$ hops away from the source nodes S. According to assumption a.3, number of hops within distance of k hops of a nodes (n_k) increases on average as $\Theta(d(k-1)^2)) = \Theta(dk^2)$.

If $k < \overline{L}$ where $\overline{L}$ is the distance between source node S and an IGW, the expanding ring search technique is applied, and this will be the only reason for overhead when the route request messages do not reach to any IGWs.

If $k \geq \overline{L}$ and equation 2 are fulfilled, IGWs that receive the route request messages send a RREP message back to the source node S that generates additional overhead in addition to the RREQ messages. However these messages will unicast to the source MNs. Assuming that the IGWs receives all the route request messages, the total amount of RREP overhead is equal to $\overline{L}\hat{n}\lambda m l_{RREP}$ (where l_{RREP} is the average size of reply packets to the route request in bits). RREP messages unicast to the source nodes, considering that source nodes are at the farthest distance from a reachable IGW, the amount of overhead will be forwarded over $L_{max} = \min(\sqrt{\hat{n}}, k_i)$ hops (where $i = \{k_1, k_2, .., k_{Threshold}, \hat{n}\}$ is the expanding ring size) that results in $L_{max}\hat{n}\lambda m l_{RREP}$ bits of overhead. Thus, total amount of IGW discovery process (C_{RD}) is:

$$C_{RD} = \Theta(\hat{n}\lambda(n_{k_i} l_{RREQ} + \min(\sqrt{\hat{n}}, k_i) m l_{RREP})) \quad (5)$$

where $\hat{n}\lambda$ is the number of new sessions, and k_i is the ring size that the RREQ message is reached by an IGW.

$$n_{k_i} = \sum_{k_i \in \{k_1, k_2, .., k_{Threshold}, \hat{n}\}} \Theta(d{k_i}^2),$$

$$i = \begin{cases} \{k_1\} & k_1 \geq \overline{L} \\ \{k_1, k_2\} & k_1 < \overline{L} \leq k_2 \\ \{k_1, k_2, .., k_{Threshold}, \hat{n}\} & k_{Threshold} < \overline{L} \end{cases}$$

If a MN broadcasts a request message per second and only one IGW receives it, there will be only one reply message per second sent back. However if more than one IGW receives the broadcast message, the maximum of m reply messages will be sent back per second. This results in maximum of $\hat{n}\lambda m$ unicast messages forwarded via $\min(\sqrt{\hat{n}}, k_i)$ hops.

2) Route maintenance overhead: In mobile ad hoc networks, paths are established not only due to new flows but also due to number of link failures in an already active path. The link failures can result from relative mobility of active nodes, channel degradation, or node power failure. Note that a link break is harmless when no route uses that link. The overhead generated by re-establishing an existing path is almost the same as establishing a new path with different rate in link changes (v) of active sessions per node (s). With the same logic for C_{RD}, the bound for the route maintenance overhead C_{RM} that is generated as the result of active link breaks is:

$$C_{RM} = \Theta(vs\hat{n}(n_{k_i} l_{RREQ} + \min(\sqrt{\hat{n}}, k_i) m l_{RREP})) \quad (6)$$

3) Sub-optimal routing overhead: The reactive discovery approach maintains the same path to the same IGW as long as it is possible, regardless of the route optimality. Therefore, when the optimal path length between S and its serving IGW is $\overline{L}$, the length of an suboptimum path can be l ($l > \overline{L}$). Since an average $\overline{r}s\hat{n}$ data packets are generated every second, the additional bandwidth required for transmission of all packets is $(l - \overline{L})\overline{r}s\hat{n}$ (bps). For the case of no local route repair and no routing cache capability in intermediate nodes, maximum number of sub-optimal hops is $\min(\sqrt{\hat{n}}, k_i) - 1$. Hence,

$$C_{RS} = O(\hat{n}\bar{r}s\min(\sqrt{\hat{n}}, k_i)) \quad (7)$$

Thus, suboptimal routing overhead increases with session duration and decreases with mobility that results in active link changes. In our analysis, we do not consider optimisation of a rediscovered path as a result of failure of an active route for simplicity. By adding up the three equations (5), (6) and (7), the total amount of overhead (C_R) in the reactive approach can be achieved by:

$$C_R = C_{RD} + C_{RM} + C_{RS} \quad (8)$$

B. Overhead in the proactive discovery mechanism

The proactive overhead of a protocol is the amount of bandwidth consumed by the protocol in order to propagate route information before it is needed.

1) Discovery routing overhead: The overhead in the proactive approach depends on the number of IGWs (m), the hop limit of the periodical advertisements (k), the period between updates (T_p) and the size of the periodical messages (l_P).

As the rate of IGW periodical updates increases ($p = \frac{1}{T_p}s^{-1}$), the overhead flooding to the ad hoc fringe also increases. So on average mp periodical messages are generated at any given second. Each periodical message is retransmitted at least once per each node for $\hat{n}$ times. This induces an overhead of $l_P\hat{n}$ bits. Thus the periodical overhead per second is:

$$C_{PD} = \Theta(\hat{n}mpl_P) \quad (9)$$

2) Sub-optimal routing overhead: Although higher frequency of IGW-ADVs increases the overhead in C_P, it optimises routes more frequently that consequently minimises the sub-optimal routing overhead. Length of suboptimal route within $\frac{1}{p}$ seconds depends on the applied mobility pattern ($\hat{n}$ presents an upper bound for this value). Therefore the sub-optimal overhead is:

$$C_{PS} = O(\frac{\hat{n}\bar{r}l_P}{p}) \quad (10)$$

Thus the total amount of overhead (C_P) in the proactive approach can be achieved:

$$C_P = C_{PD} + C_{PS} \quad (11)$$

C. Overhead in the hybrid discovery mechanism

The overhead in the hybrid discovery approach depends on the location of the MN. MNs within the proactive zone receive IGW-ADV messages transmitted with $TTL = k_{pro}$. MNs outside this zone should broadcast a RREQ message. The closed MN in the border of the proactive zone will reply back with its IGW information.

1) Hybrid discovery overhead in the proactive zone: If the MN resides in the proactive zone, then the total overhead will be equal to the total overhead of the proactive approach (C_P) considering TTL of the IGW-ADVs as k_{pro} instead of $\hat{n}$.

$$C_{HD}^{pro} = C_{PD} \mid_{\hat{n}=n_k}$$
$$C_{HS}^{pro} = C_{PS} \mid_{\hat{n}=n_k}$$

where $n_k = \Theta(dk_{pro}^2)$.

2) Hybrid discovery overhead in the reactive zone: If the MN resides outside the proactive zone, the total overhead will be aggregation of the overhead in the proactive and the reactive zones. Therefore number of the MNs outside this zone is $\hat{n}_{re} = 1 - n_k \mid_{k=k_{pro}}$. Considering a uniform traffic distribution, the probability that a source node resides in the reactive zone is $\alpha = 1 - \frac{n_k|_{k=k_{pro}}}{\hat{n}}$.

MNs broadcast a request message with $TTL = k_{re}$, where $k_{re} \leq \sqrt{\hat{n}} - k_{pro}$. Therefore for the case of expanding ring search, the maximum value to forward the RREQ message to be reached by a MN in the proactive zone will be $\sqrt{\hat{n}} - k_{pro}$. Hence, $k_i^{hy} = \{k_1, k_2, .., \hat{n}_{re}\}$.

$$C_{HD}^{re} = C_{RD} \mid_{k=k_i^{hy}}$$
$$C_{HM}^{re} = C_{RM} \mid_{k=k_i^{hy}}$$
$$C_{HS}^{re} = C_{RS} \mid_{\hat{n}=\hat{n}_{re}}$$

The total amount of the hybrid discovery approach is:

$$C_H = \alpha(C_{HD} + C_{HM} + C_{HS})^{re} + (1-\alpha)(C_{HD} + C_{HS})^{pro} \quad (12)$$

When $k_{pro} = \hat{n}$ and therefore $\alpha = 0$ the analysis is for a proactive approach.

IV. Scalability Analysis

The importance of being able to determine the scalability of a method lies in the fact that different methods may scale differently. Performance bounds of overhead, derived in Section III, gives us the ability to check the scalability of the discovery mechanisms and compare their degree of scalability.

A. Absolute scalability

A mechanism is termed absolutely scalable w.r.t a given scalability space, if the efficiency of the network does not vanish as the independent parameter tends to infinity. Efficiency of a network is said to vanish as an independent parameter tends to infinity if any of the primary metrics becomes arbitrarily large. It should be noted that although in some cases the efficiency does not vanish, it is possible that the primary metrics are finite, but very large [7].

To evaluate whether or not a method is absolutely scalable, precise definition of the primary metric, e.g. overhead, is essential. The total amount of bandwidth consumption in a network is equal to a protocol dependent factor (total amount of the overhead consumed by running the IGW discovery approach) plus an independent protocol factor (minimum traffic load).

The *minimum traffic load* of a network is the minimum amount of bandwidth required to forward packets over the optimal paths available, assuming all the nodes have instantaneous a priori full topology information of the network. This definition is independent of the routing protocol applied in the network; therefore it does not include the protocol overhead [3].

In order to quantify scalability, we use the concept of minimum traffic load defined in [3] for the network scalability factor as follows:

$$\Psi_P \triangleq \lim_{P\to\infty} \frac{\log T_r(P)}{\log P} \quad (13)$$

where $T_r(P_1, P_2, ...)$ the minimum traffic load experienced by a network under independent parameters $P_1, P_2, ...$.

To quantify a discovery mechanism's scalability, the respective scalability factor is defined based on the total overhead concept, as:

$$\rho_P^X \triangleq \lim_{P\to\infty} \frac{\log C_X(P)}{\log P} \quad (14)$$

where ρ_P^X is the respective IGW discovery scalability factor, C_X is the total amount of overhead generated by the IGW discovery protocol and X is the IGW discovery and P is the independent parameter.

A mechanism C_X is absolutely scalable if and only if, when $P \longrightarrow \infty$, the total amount of overhead generated to route and maintain the traffic is less than the number of bits a network can simultaneously transmit in a second, i.e. the capacity of the network [3].

Thus, the rate of bandwidth consumption in an scalable network caused by protocol dependent overhead should increase slower than the rate of protocol independent section:

$$\rho_P^X \leq \Psi_P \quad (15)$$

ρ_P^C shows the rate of overhead generated due to increase of P and Ψ_P shows the rate of increase in the capacity network w.r.t the independent parameter[3].

The network scalability factor (Ψ_P) is a number that asymptotically relates the increase in network load to the different network parameters. For the class of mobile ad hoc networks under assumptions in subsection II, each node ($\hat{n}$) generates r bits per seconds that must be retransmitted $\overline{L}$ times (hops). Thus, each node induces a load of $\overline{r}\overline{L}$ that after adding all the nodes results in a $T_r(r, v, \hat{n}) = \frac{\overline{r}\overline{L}n}{m}$. As we assumed neither caching nor route repair mechanisms, the model for minimum traffic load considering equations 3 and 4 will be simplified to:

$$T_r(r, v, \hat{n}) = \frac{k}{m} \sum_{i=1}^{\lambda} r_i s_i \quad (16)$$

Therefore, according to equation (13), $\Psi_v = 0$, $\Psi_r = 1$, $\Psi_p = 0$, $\Psi_\lambda = 1$, $\Psi_s = 1$, $\Psi_k = 1$, $\Psi_n = 0$, and $\Psi_m = -1$.

[3]This is the reason for the logarithmical equations (13) and (14) to express the rate of the entities w.r.t a specific parameter

Thus, for the class of network under study, according to equation (14) a mechanism X is said to be absolutely scalable w.r.t the number of nodes, mobility rate, traffic rate, new session generation rate, rate of periodical updates and the range of advertisement transmission if and only if $\rho_n^X \leq 0$, $\rho_m^X \leq -1$, $\rho_v^X \leq 0$, $\rho_r^X \leq 1$, $\rho_\lambda^X \leq 1$, $\rho_k^X \leq 1$, $\rho_k^X \leq 1$, $\rho_s^X \leq 1$, and $\rho_p^X \leq 0$ respectively.

By replacing equations (8), (9), (12) in equation (15), absolute scalability results can be derived. A summary of absolute scalability of all the discovery mechanisms are shown in table I. "+" and "-" show that the approach is absolutely scalable and absolutely unscalable respectively.

As there is no periodical advertisement from IGWs in reactive mechanism, it is absolutely scalable w.r.t the periodical update rate and number of IGWs. The proactive discovery is absolutely scalable w.r.t the rate of new session generation, duration of traffic sessions and number of link failures. The hybrid approach is absolutely scalable w.r.t to the parameters that both zones are absolutely scalable w.r.t them. Therefore all the discovery mechanisms are absolutely scalable w.r.t the rate of data traffic and the new session generation rate

B. *Relative scalability*

Although a network using a method may not be absolutely scalable [7] w.r.t particular independent parameters, one method may be more scalable than another. Even when for a metric both mechanisms are absolutely scalable, it is of the operator's interest to verify the cost of each mechanism to provide a service and how many services can be provided in by each mechanism. The entire scalability vectors should be considered, but for simplicity we evaluate the total overhead component only. We aim to compare the scalability of the discovery mechanisms w.r.t different independent parameters in this part of the paper. A discovery mechanism α is more scalable than discovery mechanism β w.r.t P, if

$$\psi_P \triangleq \lim_{P\to\infty} \frac{f(\alpha)}{f(\beta)} \longrightarrow 0$$

where P is an independent parameter and $f(\alpha)$ and $f(\beta)$ show the rate of growth of the measuring metric (i.e. overhead in this work) of the discovery mechanisms α and β respectively. If relative scalability factor (ψ_p) is a positive constant, then α and β are equally scalable.

However it may be possible that a parameter grows large in an ad hoc network, but cannot grow arbitrarily large. For instance, the traffic rate r in the network cannot increase above the capacity of the links or TTL of signalling messages can be set to a maximum value of n. Setting this value to larger value than n will result to unnecessary overhead. Therefore, it might be impossible/impractical for a parameter to grow larger than some value, B_{max}. In this case, we consider only an interval $[B_{min}, B_{max}]$, where B_{min} is the minimum value of the parameter and B_{max} is its maximum value.

As defined in [7], a method α is termed *more weakly scalable* than another method β w.r.t (E, P, M) triple, if the rate of growth of the metric measuring α, $f(\alpha)$, is slower than that of metric measuring β, $f(\beta)$.

Therefore, by the fundamental law of calculus, the discovery mechanism α is *more weakly scalable* than the discovery mechanism β w.r.t the metric M and the parameter P over the range $[B_{min}, B_{max}]$, if

$$f(\alpha \mid_{B_{max}}) - f(\alpha \mid_{B_{min}}) < f(\beta \mid_{B_{max}}) - f(\beta \mid_{B_{min}})$$

Similarly, for the equality case, then α and β are *equally scalable* w.r.t the metric M and the parameter P over the range $[B_{min}, B_{max}]$. In table I "Pro", "Re", "Hy" and "Eq" present that the proactive, the reactive, the hybrid approach, or all approaches equally scalable respectively.

Absolute	λ	s	v	r	p	k	n	m
Reactive	+	-	-	+	+	-	-	-
Proactive	+	+	+	+	-	+	-	-
Hybrid	+	-	-	-	-	-	-	
Weak / Relative	λ	s	v	r	p	k	n	m
Re vs. Pro	Pro	Pro	Pro	Eq	Re	Pro	Pro	Eq
Hy vs. Pro	Pro	Pro	Pro	Eq	Eq	Pro	Pro	Eq
Hy vs. Re	Hy	Hy	Hy	Eq	Re	Eq	Hy	Eq

TABLE I

SUMMERY OF ABSOLUTE AND RELATIVE SCALABILITY COMPARISON.

We discuss some of the results shown in table I. As the number of IGWs increases the multihop routing changes to single hop routing. Therefore the performance of all the discovery mechanisms improves. In the proactive discovery mechanism, number of nodes sending unnecessary periodical advertisements increases. Whereas in reactive approach more number of IGWs send back replies to every RREQ message and causes high amount of overhead. The hybrid approach follows a combination of these additional overheads.

The proactive approach is more scalable w.r.t (multihop routing, s, total overhead). There are two justifications for higher amount of overhead in the reactive approach: First, due to the fact that if nodes handle voice communications, it is likely to have a small number of new sessions per node due to limited wireless link capacity. Secondly, for longer sessions, it is more likely to have broken sessions.

V. OUTLOOK: AN ADAPTIVE DISCOVERY MECHANISM

In this section, we propose to adapt the value of IGW-ADV lifetime k_{pro} to optimise the reachability of nodes in the hybrid discovery approach. This may introduce more retransmission of signalling overhead but minimises the number of IGWs (i.e. the cost of set up and maintenance of IGWs will be reduced). The proactive zones can be adapted according to the topological information of active nodes. Once an IGW receives a number of RREQ packets, it will get an idea of how many and how far the nodes outside its coverage area are that demand for Internet connectivity. The IGW can proceed to increase its coverage range by increasing the TTL of the IGW-ADV messages. The amount of this increase can be made in steps of few hops or by looking at the hop count of RREQ messages. To reduce the coverage range, an IGW monitors the received data packet hop count value. If this value is less than the current TTL of IGW-ADV messages, then the IGW reduces its coverage range to the maximum number of hop count of the data it has received within the monitoring window duration.

We also propose to optimise the frequency of IGW-ADVs (p) to adapt the routing accuracy in this mechanism. The packet delivery ratio can increase by advertising more frequently for a high traffic load. This on the hand can inject high amount of overhead once it is not required. Therefore we propose to adapt the frequency of IGW-ADVs according to the traffic load received by each IGW. As the cost of signalling overhead is relative to successfully delivered data, for higher data rates increasing the rate of IGW-ADVs do not affect the overhead as much as in a light traffic scenario.

For these purposes we apply a feedback controller algorithm in [10] to the hybrid IGW discovery mechanism. Traffic load and active node hop count are the input parameters for optimisation decision of frequency and TTL of the IGW-ADVs.

VI. CONCLUSIONS

In this paper, we have considered Internet connectivity of ad hoc networks via internet gateways. We compared proactive and reactive internet gateway discovery approaches analytically. We defined the scalability space and relative scalability terms. We modelled total overhead generated by each discovery mechanism and analysed and compared their scalability degree. Finally, we propose to adapt the frequency and TTL of the gateway advertisements to adapt the discovery mechanism according to the traffic load and distance of the active nodes to the gateways.

ACKNOWLEDGMENT

The authors would like to thank Dr Cesar Santivanez and Dr Christian Bettstetter for their useful comments to enhance this paper.

REFERENCES

[1] R. Wakikawa, J. T. Malinen, C. E. Perkins, A. Nilsson, and A. J. Tuominen, "Global connectivity for ipv6 mobile ad hoc networks," Internet Draft, draftwakikawamanetglobalv602.txt, 2002.

[2] P. Ratanchandani and R. Kravets, "A hybrid approach to internet connectivity for mobile ad hoc networks," *Proceedings of IEEE WCNC*, vol. 3, pp. 1522–1527, 2003.

[3] C. A. Santivanez, B. McDonald, I. Stavrakakis, and R. Ramanathan, "On the scalability of ad hoc routing protocols," *IEEE INFOCOM*, 2002.

[4] C. A. Santivanez, "A framework for multi-mode routing in wireless ad hoc networks: theoretical and practical aspects of scalability and dynamic adaptation to varying network size, traffic and mobility patterns," Ph.D. dissertation, Northeastern University, Boston, MA, 2001.

[5] P. Jacquet and L. Viennot, "Overhead in mobile ad hoc network protocols," Institut National de Recherche en Informatique et en Automatique (INRIA), France, Tech. Rep., 2000.

[6] O. Dousse, P. Thiran, and M. Hasler, "Connectivity in ad hoc and hybrid networks," *IEEE Infocom*, 2002.

[7] O. Arpacioglu, T. Small, and Z. J. Haas, "Notes on scalability of wireless ad hoc networks," *Internet Draft*, 2003.

[8] F. Xue and P. Kumar, "The number of neighbors needed for connectivity of wireless networks," *Journal of Wireless Networks*, 2004.

[9] L. Dmitri and S. JeanMarc, "Path lengths in ad hoc networks," *International Workshop on Wireless Ad hoc Networks (IWWAN)*, 2004.

[10] M. Ghassemian, V. Friderikos, and A. H. Aghvami, "A generic feedback controller algorithm to improve performance of proactive ad hoc mechanisms," *IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks*, 2005.