On the Impact of Neighborhood Discovery on Geographical Routing in Wireless Sensor Networks

Eun Kyung Lee, John Paul Varkey, and Dario Pompili
Department of Electrical and Computer Engineering
Rutgers University, New Brunswick, New Jersey
eunkyung_lee@cac.rutgers.edu, johnpaul_varkey@cac.rutgers.edu, pompili@ece.rutgers.edu

***Abstract*—In Wireless Sensor Networks (WSNs), sensing nodes operate in dynamic environments resulting in neighboring nodes being discovered or lost at any moment causing the network topology to change constantly. Hence, routing schemes especially geographical ones (which use node positions to route data packets) require periodic exchange of control packets to discover neighboring nodes. Even though it is intuitive that the overhead caused by their periodic broadcasts may affect the end-to-end performance of the routing scheme, previous works have not thoroughly studied the impact of transmission power and frequency of control packets in static as well as mobile environments. Hence, based on our study, Distributed Neighborhood Discovery Protocol (DNDP) is proposed that can make online decisions to find the best transmit power and frequency for sending discovery packets so to minimize the effect on routing.**

***Index Terms*—Wireless Sensor Networks, Geographical Routing, Neighborhood Discovery Protocol, Mobility.**

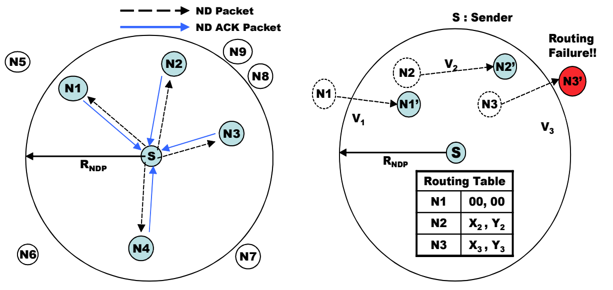
I. INTRODUCTION

Wireless Sensor Networks (WSNs) – comprising of ad hoc sensor nodes with computation and communication capabilities – are deployed to perform collaborative monitoring tasks over a physical terrain. Geographical routing, which relies on location information to forward data packets, is one of the widely used routing techniques in WSNs. In geographical routing, sensor nodes need to possess knowledge about their neighborhood to route packets efficiently towards the destination. To discover the neighborhood information and, hence, the network topology, Neighborhood Discovery (ND), which involves control message exchanges, is essential. However, *Neighborhood Discovery Protocols (NDPs)* that *periodically* exchange control message using *maximum power* will degrade the performance of the routing scheme itself by causing interference and congestion in the network. Also, the nodes may end up losing energy rapidly thereby dying out and changing the network topology. The situation is complicated further in mobile sensor networks, where the acquired information gets out of date rapidly.

The networking research community has investigated transmission power control techniques that help reduce energy consumption of wireless sensor nodes, reduce interference, and solve *exposed terminal problem* in wireless channel. However, they can worsen the *hidden terminal problem* resulting in packet losses and, hence, energy-costly retransmissions if power margin is not set properly. The exposed terminal problem occurs when a node is prevented from sending packets to other nodes due to a neighboring transmitter. The hidden terminal problem occurs when a node A is visible from the node B, but not from other nodes communicating with node B due to the different communication range. Power control techniques rely heavily on theoretical propagation models that are often unrealistic as they cannot capture the uncertainty in the wireless channel. Moreover, the benefit of using power control depends on many factors that cannot be considered offline: 1) topology, which depends on channel conditions, 2) mobility, 3) traffic, and 4) end-to-end (e2e) metric the application is interested in. Hence, online decisions have to be taken by probing the network.

In [1], the authors study how extensive the knowledge of the network topology at each node should be so that energy-efficient geographical routing decisions can be taken. However, the approach is based entirely on the distance between nodes instead of transmission power and, hence, is not accurate. In [2], the authors analyze the impact of neighbor sensing on the performance of Optimized Link State Routing (OLSR) protocol that uses information from the IEEE 802.11 MAC for ad hoc networks. However, their study is just based on the impact of frequency at which control packets are transmitted and it does not consider the effect of NDP transmission power. In [3], the authors propose a scalable NDP for infrastructure wireless mesh networks. However, the authors show only the performance gain of their algorithm in terms of localized metrics compared to OLSR and they do not study the e2e performance of their scheme in terms of packet delivery ratio and total energy consumption of the network.

In this paper, firstly, we conduct a detailed study on the impact of the neighborhood discovery process on geographical routing schemes; then, based on the observations, we propose a Distributed Neighborhood Discovery Protocol (DNDP) that uses probe packets to find the suitable power and frequency for neighborhood discovery. DNDP finds these optimal points based on our study on neighborhood discovery taking into account mobility of the nodes as well. We propose two approaches to find the best frequency and transmit power to be used for the discovery process; “Random Search” and “Selective Search”. We also use Brute-Force Reinforcement Learning (BFRL) algorithm to adjust these two parameters on-the-fly. In “Random Search”, a node randomly selects the transmit power and frequency, whereas in “Selective Search” a node chooses the power and frequency from a set of powers



(a) (b)

Fig. 1. (a) Vanilla Neighborhood Discovery Protocol; (b) Routing failure when nodes are mobile.

and frequencies based on the feedback received from previous probing. We assume that for geographical routing all the nodes are aware of their own and the destination node's position.

The remainder of the paper is organized as follows: in Sect. II, we study the impact of NDP on geographical routing schemes; in Sect. III, we propose DNDP; in Sect. IV, we evaluate the performance of the proposed solution; finally, in Sect. V, we draw the main conclusions.

II. STUDY ON THE NETWORKING IMPACT OF NEIGHBORHOOD DISCOVERY

In this section, we discuss the impact of periodic exchange of ND packets on the performance of geographical routing schemes. We start with a simple NDP called Vanilla NDP as shown in Fig. 1(a). Node S periodically sends ND packets at a certain power level with certain range (R_{NDP}) and gathers information through ACK packets (ND ACK) from its neighbors; then it decides the best next hop to route data packets based on their information. However, this neighborhood discovery process has some associated tradeoffs. The higher the transmission power used for sending ND packets, the greater the size of the neighborhood due to ACKs and, hence, the higher interference leading to more energy consumption in the network. In addition, the higher the frequency of exchange of ND packets, the more the traffic injected leading to network congestion. Last, but not least, when nodes are mobile, the faster the nodes move, the faster the perceived topology changes.

In mobile scenarios, routing failure rate is higher than immobile scenarios. For example, with reference to Fig. 1(b) where nodes are mobile, routing would fail as node $N3$ goes out of the range of sender S (R_{NDP}) while S is sending the data packet. It can also happen that sender S may route the packet to $N2$, which is still within R_{NDP} even after it has moved ($N2'$), but the other node $N1$, which is not yet in the routing table of S, has moved to previous position of node $N2$. The faster the nodes move, the lesser the possibility that the sender can capture the mobility of the nodes, thus necessitating the need for frequent update of the routing table to avoid (or limit) routing failures. Therefore, selecting optimal NDP parameters is crucial for selecting the best next hop.

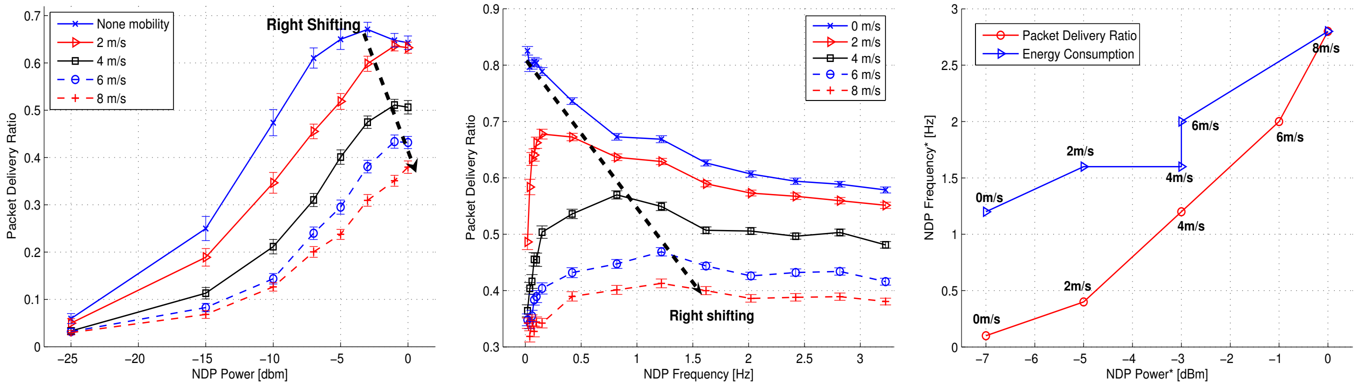
In this section, we present the impact of NDP through simulations by exploring the following questions: how big should the neighborhood be?, and how often should a node update the neighborhood information? To answer these questions, we conducted simulations by varying different NDP parameters (power and frequency) and by considering different routing schemes. Simulations were done using TOSSIM 2.x, TinyOS simulator and the radio propagation model used is described in [4]. To evaluate e2e performance, we use *packet delivery ratio* and *energy consumption of entire network per received bit* as metrics.

The whole data traffic was directed towards a single sink node at a rate of 2 Hz in terrain of area $100\text{x}100\ \mathrm{m}^2$ using one of the four different routing schemes - Most advance [5], Channel aware [6], Compass [7], and Energy aware (which selects the node that has the maximum available energy). Based on extensive simulations, we found that all the aforementioned *routing schemes are affected in a similar way by the NDP parameters chosen*. Hence, we opt for Energy Aware for our later studies.

Using all the above mentioned routing schemes, first we explore the question *'how big should the neighborhood be?'* This question is related to the transmission power used for discovery. The e2e packet delivery ratio and energy consumption for different power levels and different speeds of mobility are shown in Fig. 2(a) for a network of 36 nodes deployed in a uniform-random manner. To analyze the effect of note mobility, we use random waypoint model with various node velocities (2, 4, 6, and $8\ \mathrm{m/s}$) and assume that all the nodes are moving except the sink when ND packets transmission frequency is fixed at 1 Hz. The right shift of the peak of plots with velocity indicates that, as the nodes move faster, higher NDP power is needed to compensate for mobility (and avoid routing failures as in Fig. 1(b)). Hence, *as velocity increases, higher NDP power is required to achieve higher packet delivery ratio.*

We also explore the question of *'how often should a node update the neighborhood information?'* For the static case, Fig. 2(b) (with $0\ \mathrm{m/s}$) shows that the packet delivery ratio decreases with the increase in NDP frequency as frequent transmission of control packets results in network congestion. The plots of different velocities (2, 4, 6, and $8\ \mathrm{m/s}$) shows that an increase in the velocity of the nodes results in a decrease in the packet delivery ratio when ND packets transmission power is fixed at $0\ \mathrm{dBm}$. The right shift in the peak of plot in Fig. 2(b) indicates that when nodes are moving faster at fixed NDP power, higher NDP frequency is needed to capture the mobility of the nodes. Hence, *the higher the velocity of the nodes, the higher the NDP frequency is required in order to achieve a high packet delivery ratio.*

We summarize our evaluation in terms of best NDP frequency and best power for various node velocities in Fig. 2(c). This plot shows that best power and frequency linearly increase with node velocity for both the metrics (packet delivery ratio and energy consumption). Hence, it can be inferred that, as the nodes move faster, we need higher NDP



(a) (b) (c)

Fig. 2. (a) Packet delivery ratio vs. NDP power with mobility at NDP frequency, 1 Hz for 36 nodes; (b) Packet delivery ratio vs. NDP frequency with 0dBm NDP power for various velocities; (c) Optimal NDP frequency vs. optimal power.

power and frequency to achieve higher packet delivery ratio and to reduce energy consumption per received bit. Therefore, *it is evident that using maximum power and fixed frequency for NDP may not be the best solution for a reliable traffic delivery*. Figure 2(c) also indicates that for a fixed node velocity and fixed NDP power, the optimal NDP frequency for reduced energy consumption is higher than optimal NDP frequency for higher packet delivery ratio.

III. PROPOSED SOLUTION

In light of the simulation- and experiment-based study done to investigate the impact of NDP on routing as described in Sect. II, we propose a Distributed Neighborhood Discovery Protocol (DNDP) that dynamically probes the network to find the suitable power and frequency for discovering the neighbors so to reduce NDP's impact on routing performance. The benefit of DNDP is that nodes can dynamically adapt to topology changes that may result from various factors such as channel variation, mobility of the nodes, and traffic congestion by dynamically adjusting the transmission power and frequency of discovery packets.

For DNDP, we use an energy model that takes into account transmission power, transmission frequency, and type of routing. According to our energy model, the topology *information cost* for a generic node i (i.e., cost to acquire information about the neighbors) is defined as,

$$C_i^{INF}(p_i, f_i, L_A) = \Bigg[p_i \cdot \frac{L_D}{B} + E_{elec}^{tran} + N_i(p_i) \cdot E_{elec}^{rec} + \sum_{m\epsilon\zeta_i(p_i)} \left(p_{mi} \cdot \frac{L_A}{B} + E_{elec}^{tran}\right) + N_i(p_i) \cdot E_{elec}^{rec}\Bigg] \cdot f_i, \tag{1}$$

where p_i and f_i are NDP power [dBm] and NDP frequency [Hz], respectively; L_D and L_A are the length of ND and ND ACK packets [bit]; $E_{elec}^{tran}/E_{elec}^{rec}$ is the energy spent by the radio to transmit/receive a packet [Joule]; B is the data rate [bps]; $N_i(p_i)$ is the number of neighbors perceived by node i with transmission power p_i; and finally $\zeta_i(p_i)$ is set of node indices in the transmission range of node i.

The expression $p_i \cdot \frac{L_D}{B} + E_{elec}^{tran}$ represents the energy needed for node i to transmit ND packet to all nodes in its power range p_i, whereas $N_i(p_i)$ nodes in the transmission range of node i spend E_{elec}^{rec} each to receive ND packet. After receiving ND packet from node i, each of the $N_i(p_i)$ nodes transmit ND ACK packet with all the information required for the routing scheme to node i. The energy spent by each of these nodes to send ND ACK packet with power p_{mi} (depending on the distance between transmitter m and i) is represented by $p_{mi} \cdot \frac{L_A}{B} + E_{elec}^{tran}$. Moreover, node i spends $N_i(p_i) \cdot E_{elec}^{rec}$ to receive the ND ACK packet from each of the $N_i(p_i)$ nodes. By adding all these components and multiplying by NDP frequency f_i we obtain the expression for information cost for node i, C_i^{INF}.

The *communication cost* for node i can be defined as,

$$C_i^{COM}(\mathbf{R}) = \sum_{(s,d)\epsilon\Pi_i(\underline{R})} p_i + 2 \cdot P_{elec}, \tag{2}$$

with

$$\Pi_i(\mathbf{R}) = \{(s,d) \quad s.t. \; x_{ij}^{sd} = 1 \; for \; at \; least \; one \; j\}, \tag{3}$$

where $x_{ij}^{sd} = 1$ iff the link between node i and j is part of the path between source s and destination d, and matrix $\mathbf{R}$ describes the NDP power and frequency of all the nodes in the network.

Set $\Pi_i(\mathbf{R})$ contains all source-destination (s,d) pairs whose path includes node i as a relay node as well as those for which node i is the source. The component $2 \cdot P_{elec}$ represents the power required by the circuit to transmit and receive a packet. Thus, $C_i^{COM}(\mathbf{R})$ represents the power expenditure for all the communications node i is involved in. Hence, the information cost of each node depends on its NDP power p_i, NDP frequency f_i and also to an extent on the routing scheme used (The length of ND ACK packet L_A varies according to type of information needed for routing). This is the rationale behind our proposed solution. The total cost of node i can be computed as,

$$C_i^{TOT}(\mathbf{R}) = C_i^{INF}(p_i, f_i) + C_i^{COM}(\mathbf{R}). \qquad (4)$$

DNDP is executed at every node i that is either a source or a relay node. We indicate as K_i the set of connections where i has an active role. Periodically, each active node selects a certain power and frequency to probe the neighborhood, different from the current NDP power and NDP frequency from a discrete set of possible transmit powers and frequencies. We refer to the selected probe power as $\hat{p}$ and probe frequency as $\hat{f}$ and current NDP power as p^* and current NDP frequency as f^*. For each connection $k \epsilon K_i$, node i selects the next hop $l_i(d^k, \hat{p})$, where d^k is the destination node of the connection $k \epsilon K_i$. The node calculates

$$C_i^{TOT}(\hat{p}, \hat{f}) = C_i^{INF}(\hat{p}, \hat{f}) + \sum_{k \epsilon K_i} c_i^k(\hat{p}, \hat{f}), \qquad (5)$$

where $c_i^k(\hat{p}, \hat{f})$ is the cost of the transmissions along the path from node i to the destination node of the connection k, with probe power $\hat{p}$ and probe frequency$\hat{f}$. This accounts for the cost of transmitting data from the node itself to all the destinations, plus the cost of information associated to the probe power $\hat{p}$ and probe frequency$\hat{f}$. Once $C_i^{TOT}(\hat{p}, \hat{f})$ is calculated, if the $C_i^{TOT}(\hat{p}, \hat{f}) < C_i^{TOT}(p^*, f^*)$, then the value of NDP power and NDP frequency are adjusted accordingly such that $(p^* = \hat{p})$ and $(f^* = \hat{f})$.

Based on how to select transmission power and frequency for probing the network, we propose two approaches, “Random Search” and “Selective Search”. In “Random Search” a node randomly selects the NDP transmit power and frequency for probing such that it is not same as current transmit power and frequency used for NDP. Conversely, in “Selective Search” a node chooses power and frequency from a set of powers and frequencies based on the feedback received from previous probing. The starting set of powers and frequencies can be chosen offline based on the summary of simulation-based study (as in Fig. 2(c)), which shows the best NDP transmission power and frequency for various node velocities. If the node’s average velocity is found to be $4\ \mathrm{m/s}$, then the cutoff NDP power and frequency are $-3\ \mathrm{dBm}$ and $1.22\ \mathrm{Hz}$ based on the “packet delivery ratio” metric and, thus, the node can choose any power greater then $-3\ \mathrm{dBm}$ and any frequency greater then $1.22\ \mathrm{Hz}$ to maximize network reliability, shortening convergence time of the algorithm.

However, as employing an offline simulation-based study cannot capture all the scenarios (i.e., node density, traffic pattern, and area of deployment), a simple BFRL algorithm can be adapted to improve our “Selective Search” solution. The aim of BFRL is to maximize the cumulative reward $RW_i(p^*, f^*, v)$ (minimize the total cost) of node i on-the-fly given the different parameters where p^* is the current power, f^* is the current frequency, and v is the node velocity. Suggested BFRL paradigm assumes an approximated reward function based on our simulation-based study in Fig. 2(c). This function is adjusted toward improving the possibility of choosing the best set of NDP power and frequency based on the updated reward function.

For either of the approaches, a probe packet has six fields. The first two data fields contain the geographical coordinates of the source and the destination. The third field contains the *cumulative communication cost* and the next two fields contain the probe power $\hat{p}$ and probe frequency $\hat{f}$. The last field is a 1-bit flag, which is equal to 1 if the packet is on a forward path towards the destination or equal to 0 if it is on the reverse path. When the probe packet is created, the *cumulative communication cost* is initialized to 0 and is incremented hop-by-hop by adding *incremental communication cost*, i.e., the communication cost to reach the next hop as the packet proceeds in the forward path.

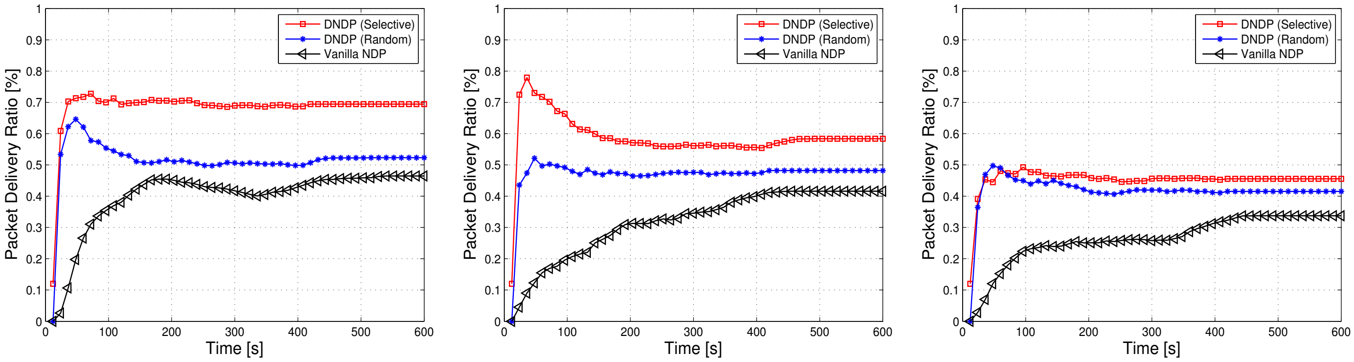
Algorithm 1 DNDP (Random and Selective Search)

```
if (Random Search) then
    randomly select $\hat{p} \neq p^*$ and $\hat{f} \neq f^*$
else if (Selective Search) then
    select $\hat{p} \epsilon \{P_0, P_1, ..., P_N\}$ and $\hat{f} \epsilon \{f_0, f_1, ..., f_N\}$
end if
for each $k \epsilon K_i$ do
    $i \rightarrow l_i(d^k, \hat{p})$
end for
wait for return packets
$C_i^{TOT}(\hat{p}, \hat{f}) = C_i^{INF}(\hat{p}, \hat{f}) + \sum_{k \epsilon K_i} c_i^k(\hat{p}, \hat{f})$
if $(C_i^{TOT}(\hat{p}, \hat{f}) \cdot RW_i(p^*, f^*, v) \leq C_i^{TOT}(p^*, f^*))$ then
    $p^* = \hat{p}$ and $f^* = \hat{f}$
    update $RW_i$
end if
```

Once a certain probe power $\hat{p}$ and probe frequency $\hat{f}$ are chosen for each connection in K_i, node i sends a probe packet to the relevant next hop and waits for its return. When a node receives a probe packet on the forward path, it looks into a cost record table to check if it already knows the incremental communication cost needed to reach this destination. If it does, there is no need to forward the probe packet to the destination. The probe packet is sent back with the updated information and the path bit is set to reverse. If it does not, the packet is forwarded to the next hop towards the destination to evaluate the communication cost. The packet is forwarded until a node with information for that destination or the destination itself is reached. When a node has gathered all the cost information associated the probe power $\hat{p}$ and probe frequency $\hat{f}$, it calculates the cost associated to them as in (5). Algorithm 1 describes the two search approaches and operations performed by a node that executes DNDP. Eventually, the NDP power, NDP frequency, and the reward function are updated only if the the total cost for the last N_{probe} values multiplied by the reward function is lower than the cost of the current NDP power and frequency.

IV. PERFORMANCE EVALUATION

We implemented the distributed neighborhood discovery protocol (DNDP) described in Sect. III and evaluated its performance with the two proposed approaches - “Selective” and “Random” against Vanilla NDP. The simulation were performed in a TOSSOM, the TinyOS simulator. We present



(a) velocity 0 m/s

(b) velocity 4 m/s

(c) velocity 8 m/s

Fig. 3. Cumulative Packet Delivery Ratio vs. Time for various velocities, 16 nodes deployment.

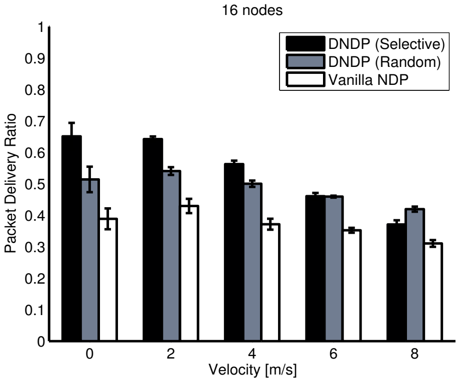


Fig. 4. **Simulation results** - Packet Delivery Ratio vs. Velocity.

simulation results for uniform random deployment of 16 nodes in a terrain of area $50\text{x}50m^2$. For mobility, we considered different node velocities (4 and 8 $\mathrm{m/s}$). In order to reach stability, we set the N_{probe} value to 3. We considered the following set of power values (0, -1, -3, -5, -7, -10) dBm and frequency values of (0.02, 0.42, 0.82, 1.22, 1.62, 2.02, 2.42, 2.82, 3.22) Hz for our simulations. For "Selective" approach, the initial set of powers and frequencies were decided based on the cut-off frequencies for various node velocities as shown in Fig. 2(c). The data traffic rate was set to 1 Hz.

Packet delivery ratio was chosen as the performance metric. Figure 3 shows the cumulative "packet delivery ratio" with respect to "time" (for a single run) with 16 nodes deployment. To assess the performance of each protocol in terms of packet delivery ratio we stopped the data traffic after 400 s of simulations but the NDP and probing continued. From Fig. 3, it is evident that "Selective" DNDP performs better than both "Random" DNDP and Vanilla NDP as it has higher cumulative packet delivery ratio. In addition, with increase in velocity of nodes (Figs. 3(a), 3(b), and 3(c)) the cumulative packet delivery ratio drops slightly. We run several simulations for all the flavors of NDP discussed with 95 % confidence intervals. Figure 4 shows the cumulative packet delivery ratio for various velocities after several runs for 16 nodes. The results indicate clearly that DNDP with "Selective" approach outperforms both DNDP with "Random" approach and Vanilla NDP.

V. CONCLUSIONS

We analyzed the impact of the neighborhood discovery protocol on geographical routing schemes by observing the effect of transmission power and frequency of neighborhood discovery probes on the packet delivery ratio and energy consumption. The analysis was carried out in both static and mobile environments. Based on the simulation- and experimental-studies, we proposed a Distributed Neighborhood Discovery Protocol (DNDP) that can make online decisions in a distributed manner to find the best transmit power and frequency for transmitting probe packets. We also explored the use of Brute-Force Reinforcement Learning (BFRL) algorithm to learn how to adjust the aforementioned parameters on the fly. We proposed two flavors of our DNDP - "Selective" and "Random". Our simulations showed that "Selective" DNDP performs better than "Random" DNDP and Vanilla NDP.

REFERENCES

[1] T. Melodia, D. Pompili, and F. Akyildiz, "On the Interdependence of Distributed Topology Control and Geographical Routing in Ad Hoc and Sensor Networks," *IEEE Journal on Selected Areas in Communications*, vol. 23, no. 3, pp. 520–532, Mar. 2005.

[2] M. Voorhaen and C. Blondia, "Analyzing the impact of neighbor sensing on the performance of the OLSR protocol," in *In Proc. of International Symposium of Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks (WiOpt)*, Boston, MA.

[3] G. Elias, M. Novaes, G. Cavalcanti, and D. Porto, "A Scalable Neighborhood Discovery Protocol for Infrastructure Wireless Mesh Networks," in *Proc. of International Conference on Advances in Mesh Networks (MESH)*, Athens, Greece, Jun. 2009.

[4] P. Levis, N. Lee, M. Welsh, and D. Culler, "Tossim: accurate and scalable simulation of entire tinyos applications," in *Proc. of the International conference on Embedded Networked Sensor Systems (SenSys)*, Los Angeles, CA, 2003.

[5] T.-C. Hou and V. Li, "Transmission Range Control in Multihop Packet Radio Networks," *IEEE Trans. Commun.*, vol. 34, no. 1, pp. 38–44, Jan. 1986.

[6] D. Pompili, T. Melodia, and I. F. Akyildiz, "Routing Algorithms for Delay-insensitive and Delay-sensitive Applications in Underwater Sensor Networks," in *Proc. of International Conference on Mobile Computing and Networking (MobiCom)*, Los Angeles, CA, Sep. 2006.

[7] G. G. Finn, "Routing and Addressing Problems in Large Metropolitanscale Internetworks," Information Sciences Institute, University of Southern California, Tech. Rep. ISI/RR-87-180, Mar. 1987.