Neighbor Discovery in Low-Duty-Cycle Wireless Sensor Networks with Multipacket Reception

Guobao Sun, Fan Wu*, and Guihai Chen
Shanghai Key Laboratory of Scalable Computing and Systems
Department of Computer Science and Engineering
Shanghai Jiao Tong University, China
sunguobao@sjtu.edu.cn, {fwu, gchen}@cs.sjtu.edu.cn

***Abstract*—Neighbor Discovery (ND) plays an important role in the initialization phase of wireless sensor networks. In real deployments, sensor nodes may not always be awake due to limited power supply, which forms low-duty-cycle networks. Existing researches on the problem of ND in low-duty-cycle networks are all based on the assumption that a receiver can receive only one packet successfully at a time. k-Multipacket Reception (MPR) techniques (i. e., k ($k \geq 2$) packets can be successfully received at a time.) have shown their significance in improving packet transmission. However, how can MPR benefit the problem of ND is still unknown. In this paper, we the first to discuss the problem of ND in low-duty-cycle networks with MPR. Specifically, we first present an ALOHA-like protocol, and show that the expected time to discover all $n-1$ neighbors is $O(\frac{n \log n \log \log n}{k})$ by reducing the problem to a generalized form of the classic *K Coupon Collector's Problem*. Second, we show that when there is a feedback mechanism to inform a node whether its transmission is successful or not, ND can be finished in time $O(\frac{n \log \log n}{k})$. Third, we point out that lacking of knowledge of n results in a factor of two slowdown in two protocols above. Finally, we evaluate the ND protocols introduced in this paper, and compare their performance with the analysis results.**

***Keywords*-Wireless Sensor Networks, Low-Duty-Cycle, Neighbor Discovery, Multipakcet Reception**

I. Introduction

Wireless Sensor Networks (WSNs) have drawn a lot of researchers' interests because of their wide range of applications. In many cases, sensor nodes are deployed without the support of pre-existing base infrastructures, and they need to form a network through their own cooperation. *Neighbor Discovery* (ND) is a family of protocols designed to find nodes' one-hop neighbors, and is the first step in the initialization of WSNs. The information acquired through neighbor discovery protocols is extremely useful for further operations such as media access and routing.

Existing protocols for ND can be classified into three categories: deterministic protocols [1], multi-user detection-based protocols [2–4], and randomized protocols [5–7, 9–13]. Deterministic protocols usually use leaders to schedule all nodes' transmissions, and multi-user detection-based protocols identify neighbors by their pre-defined signatures. Compared with the first two categories, randomized protocols are more commonly used to conduct ND. In randomized protocols, the nodes broadcast discovery messages in randomly chosen time slots to reduce the possibility of the collision from the other nodes.

Usually the problem of ND is discussed in a synchronous system, e.g., Birthday Protocols [5]. In birthday protocols each node independently chooses to transmit during each slot by probability p and to receive by probability $1-p$. By reducing the analysis of birthday protocols to the classical *Coupon Collector's Problem*, Vasudevan *et al.* [7] discussed the time complexity of birthday protocols. Many subsequent protocols are based on birthday protocols [7, 9, 10, 13]. For example, due to the development of Code Division Multiple Access (CDMA) and Multiple-Input and Multiple-Output (MIMO), several protocols adopt the fact that nodes can receive more than one packet simultaneously, i.e., Multipacket Reception (MPR), instead of the traditional assumption of Single Packet Reception (SPR) [9, 13]. Figure 1 gives an example about how the MPR technique can help to accelerate the process of ND.

Furthermore, we notice that most existing ND protocols are based on the assumption that nodes are always awake during the ND process. This is unrealistic in WSNs due to the limited power supply. In WSNs, nodes are typically working with a certain duty-cycle (transmitting, receiving, and dormancy) to reduce the energy consumption. You *et al.* [12] discussed the issue of ND process with low-duty-cycle nodes and derived the upper bound of expected time of ND under the SPR model. Jeon *et al.* [14] discussed the issue of physical-layer signal processing to achieve MPR but the low-duty-cycle scenario was not covered.

The transition from SPR to MPR in low-duty-cycle WSNs is not trivial, because nodes act completely different from the SPR scenario. First, in traditional ALOHA-like protocols (e. g., birthday protocols), the optimal transmission probability can be easily determined to be $1/n$, where n is the clique

This work was supported in part by the State Key Development Program for Basic Research of China (Grant No. 2012CB316201), in part by China NSF grant 61073152, 61170236, 61272443, and 61133006, and in part by Shanghai Science and Technology fund 12PJ1404900. The opinions, findings, conclusions, and recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of the funding agencies or the government.

* F. Wu is the corresponding author.

1521-9097/12 $26.00 © 2012 IEEE
DOI 10.1109/ICPADS.2012.73

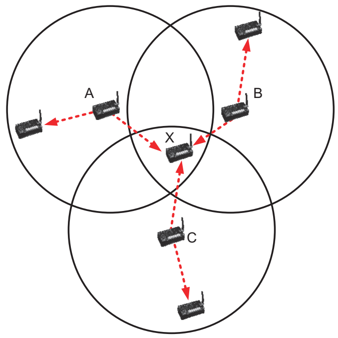

Figure 1. An example of how the ND is conducted in MPR networks. Node A, B, and C are broadcasting their discovery messages simultaneously. Node X is in the coverage of all the three nodes. If they are in a SPR network, collision will occur at X. However if they are in a 3-MPR network, X will successfully receive three nodes' discovery messages simultaneously.

size [5]. However, it is difficult to derive a closed form for the optimal transmission probability in k-MPR[1] networks. Second, previous researches with SPR model are all based on the assumption that once a node has transmitted its discovery message without collision, it will certainly be discovered by all the other nodes in a clique, which does not hold in low-duty-cycle WSNs. The reason is twofold. On one hand, it is almost impossible for all nodes to be awake at a certain time instant in low-duty-cycle networks because many nodes may be dormant. On the other hand, even if all nodes happen to be awake, it is still not enough for a node A to transmit its discovery message only once to let all other nodes find it, due to the reason that there may be more than one node, say $m(1 < m \leq k)$ nodes (including A), transmitting simultaneously. Since the radios on sensors nodes are half duplex, $n-m$ nodes can discover A successfully , while $m-1$ nodes cannot, because they are transmitting.

In this paper, we study the problem of ND in low-duty-cycle WSNs with k-MPR radios, and conduct in-depth performance analysis on ALOHA-like ND protocols with various extensions. The contributions of this paper are listed as follows:

- First, to the best of our knowledge, we are the first to consider the problem of ND using MPR radios in low-duty-cycle WSNs. We show that MPR can significantly accelerate the ND process, and thus the duration of ND in low-duty-cycle networks can be tremendously shortened. We study the ALOHA-like protocol in k-MPR networks and prove that the expected time needed is $O(\frac{n \log n \log \log n}{k})$, where n is the clique size, by reducing the problem to a generalized form of K *Coupon Collector's Problem* [17].
- Furthermore, when a feedback mechanism is introduced into the system, we prove that it provides a $\log n$ improvement over the ALOHA-like protocol.
- Finally, we extend our protocols to the case where the clique size n is unknown and show that it results in a factor of two slowdown.

[1] k-MPR means that a receiver can successfully receive at most $k (k \geq 2)$ packets simultaneously.

The rest of this paper is organized as follows. In Section II, we describe the model and assumptions under which we present our discussion. In Section III, we analyze the performance of ALOHA-like protocol in low-duty-cycle WSNs. In Section IV, the case when a feedback mechanism is introduced into the system. In Section V, we validate the theoretical results by simulation. We present related works in Section VI and the paper concludes with our future work in Section VII.

II. Network Model and Assumptions

In this section, we present our network model and assumptions under which we discuss the issue of ND. They are widely adopted by many previous works [5, 7, 9, 12, 13]. These assumptions are as follows:

- Each node has a locally unique identifier (e.g. the MAC address, the location).
- Time is identically slotted and nodes are synchronized on slot boundaries.
- n nodes are deployed in a clique. For simplicity we label them as $\{1, 2, \ldots, n\}$.
- We assume that n is known to all nodes in the clique.
- All nodes have the same transmission range and use omnidirectional antennas.
- Nodes are in a k-MPR ($k \geq 2$) network, which indicates that there is a collision in the clique if and only if there are more than k nodes transmitting simultaneously in a slot. This capabilities can be achived by MIMO or CDMA. To simplify our discussion, we neglect some real implementation issues of MPR, and assume that all sending nodes' packets can be successfully decoded at receivers' side as long as the number of transmitters is no more than k.
- Nodes are half duplex, i.e., nodes can either transmit or listen in a slot but not both at the same time.

III. ALOHA-like Protocol

In this section, we present the ALOHA-like protocol and analyze its performance based on the model and assumptions described in Section II. Recall that there are n nodes and they are in a clique. Furthermore, the case where n is unknown to nodes is discussed in Subsection III-C.

A. Protocol Description

In the ALOHA-like protocol, we assume that each node independently chooses how to act in a time slot. Precisely speaking, each node chooses to be awake by probability p_w and to be dormant by probability $1-p_w$. Dormant nodes do not transmit or receive and awake nodes will independently choose to transmit by probability p_t and receive by probability p_l. In total, there are three states for a node in each slot: transmitting, receiving, and dormancy. The corresponding probabilities are $p_w p_t$, $p_w p_l$ and $1-p_w$, respectively.

If we assume a node transmits by probability p in a slot, we can also know that it is in receiving state by probability $p_w - p$ and is dormant by probability $1-p_w$.

We can determine the probability of a *successful slot*, i.e., there is at least one node transmitting in the slot, and no collision occurs. A theorem is as follows:

Theorem 1. *For a given slot in k-MPR networks, the probability that no collision occurs in the slot is given by*

$$p_s = \sum_{i=1}^{k} \binom{n}{i} p^i (1-p)^{n-i}. \tag{1}$$

Proof: In k-MPR networks, a collision occurs if and only if more than k nodes transmit simultaneously. Since all nodes independently choose to transmit by probability p, the probability that i nodes transmit in a slot is

$$p_i = \binom{n}{i} p^i (1-p)^{n-i}.$$

Therefore, the theorem holds. ∎

B. Performance Analysis

In this subsection, before the analysis of the ALOHA-like protocol in k-MPR low-duty-cycle WSNs, we first introduce two lemmas. One is about the generalized form of *K Coupon Collector's Problem*, and the other is about the estimation of p_s.

Lemma 1. *There are n different coupons and a collector randomly chooses $k(k<n)$ distinct coupons (with replacement) in a run. Denote the number of expected runs to get all n coupons picked out and each coupon is picked out at least m times as $D^k_{m,n}$. Then*

$$D^k_{m,n} = O\left(\frac{nm \log\log n + n \log n}{k}\right). \tag{2}$$

Due to the space limitation we omit the analysis and proof of this theorem and it can be found in [17]. The k in k-MPR and the K in K Coupon Collecor's Problem is different. Although at most k nodes can successfully transmit simultaneously, there may not be exact k nodes transmitting, while the K Coupon Collecor's Problem requires that each time there are exact K coupons that are picked out.

Lemma 2. *Let $p = \frac{k-1}{n}$ and the following inequality holds:*

$$p_s > \frac{1}{2} - \frac{1}{e}.$$

Proof: Define X_i^t as a binary indicator random variable of the event "node i transmits in slot t". Then the expression of p_s in time slot t can be rewritten as follows

$$p_s = Pr[1 \le \sum_i X_i^t \le k].$$

It is obvious to see that $\sum_i X_i^t$ follows a Binomial distribution, and its mean is $np = k-1$. Due to the reason that the mean and the median are at most $\ln 2$ apart [18], the median is in $[k-1-\ln 2, k-1+\ln 2]$. Since $k-1+\ln 2 < k$, we can see that

$$Pr[\sum_i X_i^t \le k] > \frac{1}{2}.$$

Hence,

$$\begin{aligned} p_s &= Pr[\sum_i X_i^t \le k] - Pr[\sum_i X_i^t = 0] \\ &> \frac{1}{2} - (1-p)^n > \frac{1}{2} - (1-\frac{1}{n})^n \\ &> \frac{1}{2} - \frac{1}{e}, \end{aligned} \tag{3}$$

where the last inequality comes from the known inequality $(1-\frac{1}{n})^n < \frac{1}{e}$ for $\forall n \in \mathbb{N}$. ∎

In Section I, we have mentioned that it is not enough for a node to transmit successfully without collisions only once to let all other nodes discover it in low-duty-cycle WSNs. The following theorem shows how many successful transmissions are needed for a node.

Theorem 2. *If a node A transmits its discovery message $3\log_L n$ times without collisions where $L = 1/(1-p_w+p)$, then A is discovered by its $n-1$ neighbors with high probability.*

Proof: Since in these $3\log_L n$ slots no collision occurs, a node B will discover A if and only if B is in receiving state in at least one of these $3\log_L n$ slots. Hence, the probability that B does not discover A successfully is given by

$$p_B = (1-p_w+p)^{3\log_L n}.$$

Hence,

$$p_B = (\frac{1}{L})^{3\log_L n} = \frac{1}{n^3}.$$

We denote the event "there is at least one node that does not discover A after $3\log_L n$ slots" as ϵ, and we can determine $P(\epsilon)$ according to *Union Bound*

$$P(\epsilon) \le np_B = \frac{1}{n^2}.$$

We can see that $P(\epsilon) \to 0$ as $n \to \infty$. ∎

We are now ready to analyze the performance of the ALOHA-like protocol and point out its time complexity.

Theorem 3. *Let T be the time needed to discover all n nodes by using the ALOHA-like protocol. The expected value of T is given by*

$$E[T] = O\left(\frac{n \log n \log\log n}{k}\right). \tag{4}$$

Proof: In each time slot, there are $O(k)$ nodes intending to transmit and each node need to transmit $O(\log n)$ times without collisions to let all other nodes discover it. According to Lemma 1, we know that if every slot is successful the expected time needed is given by

$$W = O(\frac{n \log n \log\log n}{k}).$$

Note that T is a Pascal random variable with the parameter p_s and W, therefore we get

$$E[T] = \frac{W}{p_s} < \frac{W}{\frac{1}{2} - \frac{1}{e}} = O(\frac{n \log n \log\log n}{k}).$$

■

Now we have proven the time complexity of the ALOHA-like protocol and this is the generalization of previous works. By using lemmas and theorems proposed in this paper, we can derive many results of the performance of the ALOHA-like protocol in various scenarios.

- By setting $k = 1$ and $m = 1$, we get $O(n \log n)$ which is the time complexity of birthday protocols [5, 7] which is designed for SPR networks and nodes are always awake.
- By setting $m = 3$, we get $O(\frac{n \log n}{k})$ which is proven in [9, 14]. It is aimed at MPR networks whose nodes are keeping awake.
- By setting $k = 1$ and $m = 3 \log n$, we get $O(n \log n \log\log n)$ which is proposed in [12]. In this SPR network nodes are not always awake and have a duty cycle $p_w = 1/2$.

C. Unknown Number of Neighbors

In previous discussion, we all assume that the clique size n is known to all nodes in the clique. In this subsection, we will discuss how the protocol works if n is unknown to all nodes.

We use a standard method [7, 9, 13] to handle this situation. The basic idea is to divide the whole ND process into phases. In phase i, each node runs the protocol as if there are 2^i neighbors to discover. This phase lasts $O(\frac{2^i \log 2^i \log\log 2^i}{k})$ slots. Consequently, in the $\lceil \log_2 n \rceil$-th phase, each node will run the protocol as if there are n neighbors and this phase will last $O(\frac{n \log n \log\log n}{k})$ slots. This is just what we have derived before and the ND process can be terminated after this phase.

Now the expected time needed is

$$\sum_{i=1}^{\lceil \log_2 n \rceil} O(\frac{2^i \log 2^i \log\log 2^i}{k})$$
$$< \sum_{i=1}^{\lceil \log_2 n \rceil} O(\frac{2^i \log n \log\log n}{k})$$
$$= O(\frac{2n \log n \log\log n}{k}).$$

Hence, we observe that the lack of knowledge of n results in a factor of two slowdown.

IV. Adaptive ALOHA-like Protocol

In this section, we discuss the problem of ND in low-duty-cycle k-MPR WSNs when a transmitting node knows whether its transmission is successful or not. The feasibility and design of such kind of feedback mechanisms has been discussed in [7, 10].

In contrast to Section III, we divide a time slot into two sub-slots. Nodes independently choose to transmit or receive in the first sub-slot. For a receiving node it will check if there is a collision in the first sub-slot, i.e., more than k nodes are simultaneously transmitting [10], and broadcast a signal in the second sub-slot if a collision occurs. As a transmitting node, it keeps listening in the second sub-slot. If it hears a signal it knows that its transmission in this slot was failed; otherwise, it knows that the transmission was successful. We note that when all nodes are transmitting in the same slot, all of them will think their transmissions are successful, because they cannot receive feedback signals, which nullifies the feedback mechanism. Fortunately, the probability of this event is $(\frac{k-1}{n})^n$, which tends to be 0 as $n \to \infty$. Therefore it is reasonable to ignore it.

The main idea of our design with the feedback mechanism is similar with [9]. We refer to a node that only receives and sleeps as *passive* node, otherwise *active* node. At the beginning of the ND process all nodes are set to be active. Time is divided into phases. In phase i there are n_i nodes to be discovered; p is set to be $(k-1)/n_i$ and this phase lasts $W_i = \Theta(\frac{n_i \log\log n_i}{k})$ slots. At the end of a phase, all nodes that have successfully transmitted their discovery messages at least $3 \log_L n$ times will turn passive. We will prove that with a proper W_i, at least half of the nodes will turn passive at the end of this phase. Then in next phase, the remaining active nodes transmit by a higher probability. This process continues until there are at most $n/\ln n$ nodes that are active. Then we run the ALOHA-like protocol without feedback mechanisms as in Section III.

According to the scheme described above, the following inequality holds:

$$n_i \le \frac{n}{2^{i-1}},$$

where n_i is the number of active nodes in phase i. Hence, the total time needed is given by

$$\sum_{i=1}^{\log_2 \ln n} O(\frac{n \log\log \frac{n}{2^{i-1}}}{k 2^{i-1}}) + O(\frac{\frac{n}{\ln n} \log \frac{n}{\ln n} \log\log \frac{n}{\ln n}}{k}).$$

We can get from the equation above that the total time needed is $O(\frac{n \log\log n}{k})$, which has a factor of $\log n$ speedup compared with the one in Section III.

Because every phase runs independently and identically except that n_i and p_i are different, we will consider only

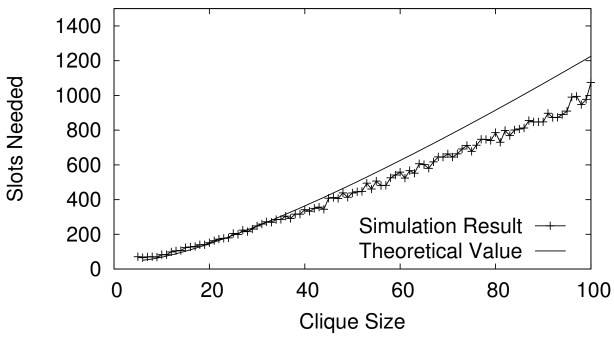

Figure 2. Validation of the ALOHA-like Protocol

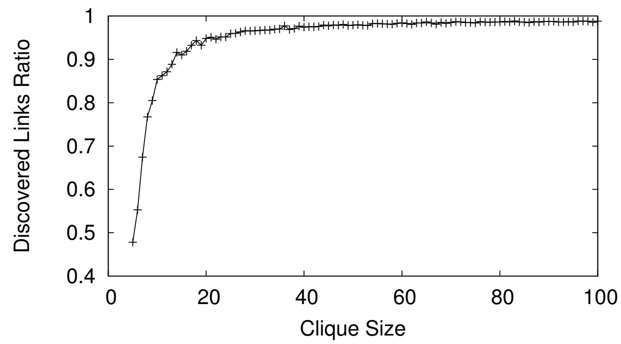

Figure 3. The Discovery Ratio of the Adaptive ALOHA-like Protocol

one phase and prove that with proper n and p, at least $n/2$ nodes will turn passive at the end of the phase.

Theorem 4. *Let S denote the set of passive nodes at the end of the phase. Let $p = (k-1)/n$ and $W = \frac{\eta n \log\log n}{k-1}$ where η satisfies the condition*

$$\frac{(\eta \log\log n - 3\log_L n + \frac{1}{2})^2}{32\eta} > \frac{k}{k-1}.$$

Then for $\forall k \geq 2$

$$Pr[|S| < \frac{n}{2}] < e^{-\frac{n}{k \log\log n}}. \quad (5)$$

Proof: Define the variable Y_i^t for node i in slot t as follows:

$$Y_i^t = \begin{cases} X_i^t & \text{if } t \text{ is a successful slot,} \\ 0 & \text{otherwise.} \end{cases}$$

By Lemma 2 and the definition of Y_i^t we can get

$$Pr[Y_i^t = 1] = Pr[X_i^t = 1] \cdot Pr[\sum_i X_i^t \leq k] \geq \frac{p}{2}.$$

Let $K_i^t = \min\{3\log_L n, \sum_{t'<t} Y_i^{t'}\}$ and $S^t = \{i | K_i^t = 3\log_L n\}$. We have

$$K^t \triangleq \sum_i K_i^t \leq |S^t| \cdot (3\log_L n) + (n - |S^t|) \cdot (3\log_L n - 1).$$

If $|S^t| < n/2$ we get

$$K^t < n(3\log_L n - \frac{1}{2}).$$

Hence,

$$Pr[|S^W| < \frac{n}{2}] \leq Pr[K^W < n(3\log_L n - \frac{1}{2})] \quad (6)$$

We define $Z^t \triangleq K^t - K^{t-1}$ if $|S^{t-1}| \leq n/2$. (If $|S^{t-1}| > n/2$ the conclusion holds obviously.) Note that $Z \triangleq \sum_{t=1}^{W} Z^t = K^W$. Therefore according to Equation (6),

$$Pr[|S^W| < \frac{n}{2}] \leq Pr[Z < n(3\log_L n - \frac{1}{2})].$$

On the other hand, we have

$$E[Z^t] = \sum_{i \notin S^{t-1}} Y_i^t \geq (n - |S^{t-1}|) \cdot \frac{p}{2}$$
$$\geq \frac{n}{2} \cdot \frac{k-1}{2n} = \frac{k-1}{4}.$$

Define $\widetilde{Z}^t = E[Z | X_i^m, m \leq t]$ and the following equation holds:

$$E[\widetilde{Z}^t | \widetilde{Z}^{t-1}] = \widetilde{Z}^{t-1},$$

which indicates that the sequence forms a martingale. Then we apply *Azuma's inequality* [20] to $\widetilde{Z}^t$ and we have

$$Pr[|S^W| < \frac{n}{2}] \leq Pr[Z < n(3\log_L n - \frac{1}{2})]$$
$$= Pr[\widetilde{Z}^W < n(3\log_L n - \frac{1}{2})].$$

Since $E[\widetilde{Z}^W] = E[Z] \geq \frac{(k-1)W}{4} = \frac{\eta n \log\log n}{4}$, we get

$$Pr[\widetilde{Z}^W < n(3\log_L n - \frac{1}{2})] \leq Pr[\widetilde{Z}^W < E[\widetilde{Z}^W] - \frac{nM}{4}]$$
$$\leq \exp(-\frac{n^2 M^2 (k-1)}{32\eta k^2 n \log\log n}),$$

where $M = \eta \log\log n - 3\log_L n + \frac{1}{2}$.

Taking the condition given in the theorem we can get

$$Pr[|S^W| < \frac{n}{2}] < e^{-\frac{n}{k \log\log n}}.$$

■

We note that previous mentioned strategy to handle the case when n is unknown in Subsection III-C can also be used in the protocol with a feedback mechanism. Similarly we can observe a factor of two slowdown in this case.

V. Performance Evaluation

In this section, we validate our theoretical results by simulations. In our simulations we assume that nodes are all in a clique, and the size of the clique n is known beforehand. The cases when n is unknown and nodes are in multi-hop networks are well discussed and verified in [7, 9, 12, 13], thus we omit them. Each data plot in the figures stands for an average result over 20 runs for accuracy.

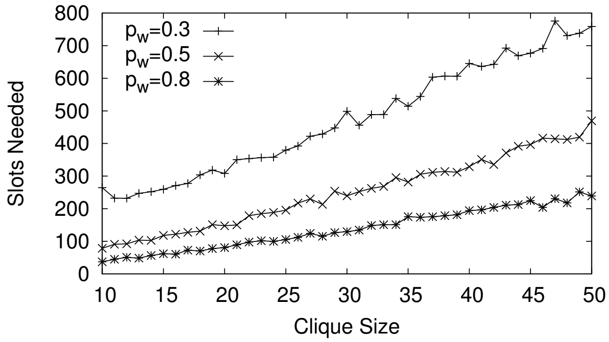

Figure 4. Comparison of Different Duty Cycles (ALOHA-like)

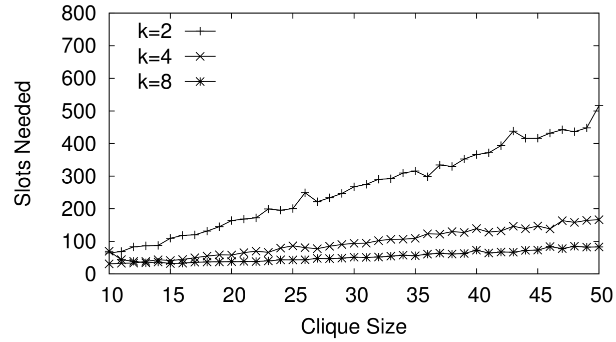

Figure 5. Comparison of Different k-MPR (ALOHA-like)

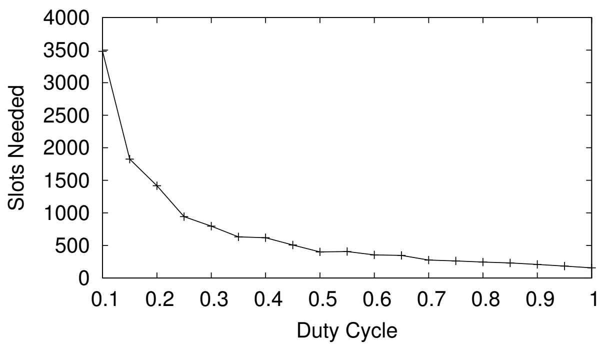

Figure 6. Comparison of Different Duty Cycles (ALOHA-like, $n = 50$)

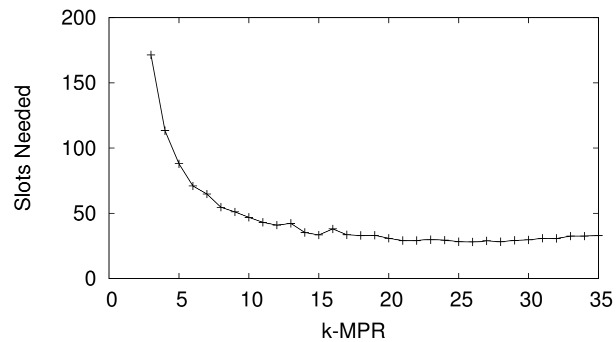

Figure 7. Comparison of Different k-MPR (ALOHA-like, $n = 50$)

A. Validation of Theoretical Results

Figure 2 and 3 are the simulation results for the ALOHA-like protocol and the adaptive ALOHA-like protocol, respectively. In both figures the parameter $p_w = 0.5$ and $k = 3$. Figure 2 shows the trend of the number of time slots needed to discover all nodes with increasing size of the clique. We can see that the simulation results well fit with the corresponding theoretical values. The deviation is due to the reason that the closed form of the generalized *K Coupon Collector's Problem*'s expected time is non-trivial to be derived, thus we can only give the asymptotic results, but it is still able to prove the correctness of our derivation.

Figure 3 provides a link-based view of the ND process. In the link-based view we regard the connection between any two nodes in the clique as a link, and it is easy to see there are $n(n-1)$ links in a clique of n nodes. We present the ratio of discovered links in the given time with different sizes of cliques, where the given time is determined by the adaptive ALOHA-like protocol (We set $\eta = 5$ in this figure.). We can see from the figure that the discovery ratio is very close to 1 when the clique size is relatively large. Nevertheless the ratio is not that acceptable when n is small. It is reasonable because our results are all asymptotic results and the results match our derivation well when n is large.

B. Different Settings for ALOHA-like Protocol

Now we analyze the performance of ALOHA-like protocol when different duty cycles and k-MPR are deployed in the clique. In this simulation, when comparing different duty cycles we set $k = 3$. When comparing different k we set the duty cycle $p_w = 0.8$.

Figure 4 shows the comparison among three different settings of duty cycles when $k = 3$. Figure 6 shows the trend of the time slots needed with increasing duty cycles when $n = 50$. When the duty cycle increases, the total time needed to discover all nodes decreases. It is predictable because low duty cycle means many nodes may be dormant at a single slot and the transmitting nodes' discovery messages cannot be received by most nodes, indicating that transmitters need more transmissions to ensure that the discovery messages have been received by all nodes at least once.

The comparison among different settings of k-MPR is shown in Figure 5 (Note that now $p_w = 0.8$). Figure 7 shows the trend of the time slots needed with increasing k when $n = 50$. When k increases, the total time needed decreases. Furthermore, when k doubles, the speed of ND is about twice faster. This coincides with our theoretical result and thus proves the correctness of our theorems.

On the other hand we must point out that the time

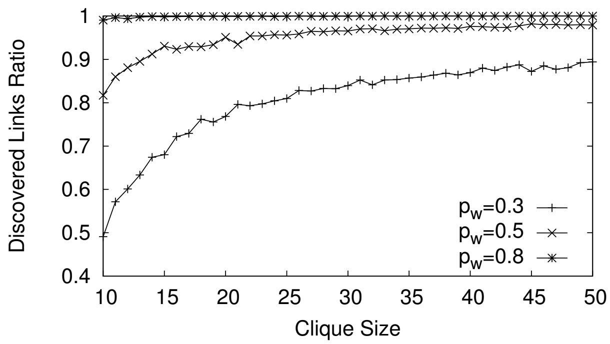

Figure 8. Comparison of Different Duty Cycles (Adaptive ALOHA-like)

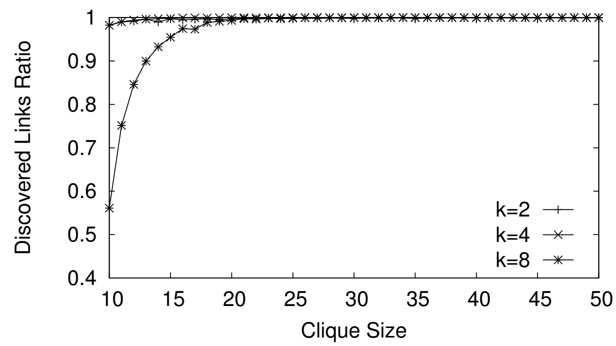

Figure 9. Comparison of Different k-MPR (Adaptive ALOHA-like)

needed is not always decreasing as k increases. Note that the beginning part of the case $k = 8$ is higher than the case $k = 4$. This is mainly because the transmitting probability is based on k and thus this probability is relatively high if k is large. When number of nodes is small, at a slot most nodes are transmitting and few nodes are receiving. Consequently, transmitting nodes need to spend more time letting all other nodes receive their discovery messages. We can observe this from Figure 7 obviously. When $k > 15$, the total time needed fluctuates instead of keeping decreasing.

C. Different Settings for Adaptive ALOHA-like Protocol

In this subsection we begin to analyze the performance of the adaptive ALOHA-like protocol with different settings of duty cycles and k-MPR. Similarly, when comparing duty cycles we set $k = 3$. When comparing different k we set $p_w = 0.8$.

Figure 8 shows the discovery ratio of three scenarios with different duty cycles. It is clear that the discovery ratio increases as the duty cycle p_w increases. Again we observe that when the size of the clique is small the ratio turns out to be relatively low because of our asymptotic analysis. When the duty cycle approches 0.8, the adaptive ALOHA-like scheme ensures that almost all nodes can be discovered in the given time slots. In addition, we can see from the figure that as the size of clique rises, the discovery ratio also rises. This coincides with our aysmptotic analysis and shows that the validity of the adaptive ALOHA-like protocol.

Figure 9 shows the discovery ratio of three scenarios with different settings of k. In this figure, when the clique size is 10 with $k = 8$, the discovery ratio is only less than 0.6, whereas the ratio is almost 1 when $k = 4$ and $k = 2$. This result may not seem to cater our intuition, but it indeed can justify our assertion again: when the clique size is small, it is not a good idea to use large k. Too many transmitting nodes and too few receiving nodes will bring the side effect which prolongs the process of ND, because a node need to transmit a lot of times to make itself heard by all other nodes.

VI. RELATED WORK

Many works have focused on the problem of ND and various protocols have been proposed and analyzed to adapt to different situations and assumptions. Basically, protocols of ND can be classified into three classes: deterministic protocols [1], multi-user detection-based protocols [2–4], and randomized protocols [5–7, 9–13]. Deterministic protocols usually need a leader, which is aware of the whole topology of the network and schedule the transmitting and receiving beforehand to total avoid collisions. This kind of scheduling costs lot of time and it is hard to implement it in a large scale distributed system. The multi-user detection-based protocols need complicated signal processing techniques and require that each node keeps all other nodes' signal signatures, which is unrealistic in many scenarios. Compared with the previous two kinds of protocols, randomized protocols are widely deployed due to their effectiveness and low cost.

The milestone of the randomized protocols of ND is the Birthday Protocol proposed in [5] by McGlynn *et al.*, who consider the randomized strategy in a synchronous system to avoid collisions in a clique. In birthday protocol, each node transmits its discovery message by probability p and receives other nodes' messages by probability $1 - p$ in a slot. Furthermore, the authors proved that the optimal transmission probability $p = 1/n$, where n is the size of the clique.

Based on the birthday protocol, Vasudevan *et al.* [6] proposed a similar randomized strategy when directional antennas are used instead of omnidirectional antennas. However the authors did not analyze the expected time formally in this paper. Later in [7], the authors first theoretically analyzed the time upper bound of the birthday protocol by reducing the ND problem to the classical *Coupon Collector's Problem*. When there are n nodes in the clique in a synchronous system, the expected time needed to discover all nodes is neH_n where H_n is the n-th Harmonic number. In [7], the authors also proposed methods to handle more realistic situations where n is unknown beforehand, the system is asynchronous [11] and a feedback mechanism is introduced into the system [10]. Basically, not knowing n beforehand and the asynchronous system leads to no more than a factor of two slowdown respectively, and there will be a $\ln n$ improvement if a feedback mechanism is brought in. In addition, the author also proposed a method to determine when to terminate the ND process when n is unknown. Sun *et al.* proposed a refined ALOHA-like protocol to compete a long-existing problem in the traditional ALOHA-like protocols.

Zeng *et al.* first extended the results of [5, 7] to the k-MPR situation. In contrast to previous works that are all based on the assumption that there is a collision if two or more nodes transmit simultaneously in a clique, k-MPR allows at most k ($k \geq 2$) nodes in a clique transmit simultaneously. The authors proved that the expected time needed to discover all nodes is $\Theta(n \ln n/k)$. Ideally, if $k \geq n$, the expected time is shortened to $\Theta(\ln n)$. Similarly, the lack of knowledge of n, the asynchronous system and the import of feedback mechanisms result in the same factors of slowdown or speedup as they are in [7].

You *et al.* [13] considered a different MPR model in comparison with [9]. In [13], there are k channels. At each slot each node can transmit on one of the k channels or receive on all channels simultaneously. As a result a node can receive at most k packets successfully if k nodes choose mutually exclusive channels to transmit their messages (Note that there is only one channel in [9].). The authors got the same time complexity $\Theta(n \ln n/k)$.

Recently many works have focused on the proper ND protocols for WSNs. Many sensor nodes work in a duty cycle because of the shortage of power supply. Hence, the protocols which assume that a node has been discovered by all other nodes if it has transmitted successfully only once no longer work in the low-duty-cycle WSNs, since some nodes may be dormant and cannot receive anything at some time instants.

You *et al.* [12] extended the discussion of [7] to the low-duty-cycle case. By reducing the problem to the *K Coupon Collector's Problem* [19], the authors proved that when the duty cycle is $1/2$, the upper bound is $ne(\log_2 n + (3\log_2 n - 1)\log_2 \log_2 n + c)$ with a constant c.

Besides these works which are aimed at accelerating the process of ND, there are also many other researches that discuss other problems about ND such as the security [15], energy consumption [16], etc. Due to the space limitation we will not introduce them in detail.

VII. Conclusion and Future Work

In this paper, we analyzed the neighbor discovery problem in low-duty-cycle WSNs, and derived the time complexity for two protocols respectively. For the ALOHA-like protocol, the expected time to finish ND is $O(\frac{n \log n \log\log n}{k})$ with k-MPR. Furthermore, if a feedback mechanism is introduced into the system, the expected time is $O(\frac{n \log\log n}{k})$. In addition, the lack of knowledge of n results in a factor of two slowdown in comparison with the n-known case. All our theoretical results are verified by extensive simulations.

In the future, we would like to evaluate these protocols by doing test-bed experiments. Also we would like to extend the protocols to some more realistic situations, e.g. nodes with different clocks, nodes with different duty cycles and more realistic radio models.

References

[1] A. Keshavarzian and E. Uysal-Biyikoglu. "Energy-Efficient Link Assessment in Wireless Sensor Networks". In *Proc. of IEEE INFOCOM*, 2004.

[2] D. Angelosante, E. Biglieri, and M. Lops. "Neighbor Discovery in Wireless Networks: A Multiuser-Detection Approach". In *Information Theory and Applications Workshop*, 46-53, 2007.

[3] J. Luo and D. Guo. "Neighbor Discovery in Wireless Ad Hoc Networks Based on Group Testing". In *Annual Allerton Conference*, 2008.

[4] L. Zhang, J. Luo, and D. Guo. "Compressed Neighbor Discovery for Wireless Networks". Preprint, http://arxiv.org/abs/1012.1007.

[5] M. J. McGlynn and S. A. Borbash. "Birthday Protocols for Low Energy Deployment and Flexible Neighbor Discovery in Ad Hoc Wireless Networks". In *Proc. of ACM MobiHoc*, 137-145, 2001.

[6] S. Vasudevan, J. F. Kurose, and D. F. Towsley. "On Neighbor Discovery in Wireless Networks with Directional Antennas". In *Proc. of IEEE INFOCOM*, 2005.

[7] S. Vasudevan, D. Towsley, D. Goeckel, and R. Khalili. "Neighbor Discovery in Wireless Networks and the Coupon Collector's Problem". In *Proc. of ACM MobiCom*, 181-192, 2009.

[8] G. Sun, F. Wu, X. Gao, and G. Chen. "PHED: Pre-Handshaking Neighbor Discovery Protocols in Full Duplex Wireless Ad Hoc Networks'. In *Proc. of IEEE GLOBECOM*, 2012.

[9] W. Zeng, X. Chen, A. Russell, S. Vasudevan, B. Wang, and W. Wei. "Neighbor Discovery in Wireless Networks with Multipacket Reception". In *Proc. of ACM MobiHoc*, 3:1-10, 2011.

[10] R. Khalili, D. Goeckel, D. Towsley, and A. Swami. "Neighbor Discovery with Reception Status Feedback to Transmitters". In *Proc. of IEEE INFOCOM*, 2010.

[11] S. A. Borbash, A. Ephremides, and M. J. McGlynn. "An Asynchronous Neighbor Discovery Algorithm for Wireless Sensor Networks". Elsevier Ad Hoc Networks, 5(7): 998-1016, 2007.

[12] L. You, Z. Yuan, P. Yang, and G. Chen. "ALOHA-Like Neighbor Discovery in Low-Duty-Cycle Wireless Sensor Networks". In *Proc. of IEEE WCNC*, 749-754, 2011.

[13] L. You, X. Zhu, and G. Chen. "Neighbor Discovery in Peer-to-Peer Wireless Networks with Multi-Channel MPR Capability". In *Proc. of IEEE ICC*, 2012.

[14] J. Jeon and A. Ephremides. "Neighbor Discovery in a Wireless Sensor Network: Multipacket Reception Capability and Physical-Layer Signal Processing". In *48th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, 2010.

[15] M. Poturalskim, P. Papadimitratos, and J. Hubaux. "Secure Neighbor Discovery in Wireless Networks: Formal Investigation of Possibility". In *Proc. of ASIACCS*, 2008.

[16] M. Kohvakka, J. Suhonen, M. Kuorilehto, V. Kaseva, M. Hännikäinen, and Timo D. Hämäläinen. "Energy-efficient neighbor discovery protocol for mobile wireless sensor networks" Ad Hoc Networks, 7(1): 24-41, 2009.

[17] W. Xu and A. K. Tang. "A Generalized Coupon Collector Problem". In *Journal of Applied Probability*, 48(4): 1081-1094, 2011.

[18] K. Hamza. "The Smallest Uniform Upper Bound on the Distance between Binomial and Poisson Distributions". In *Statistics & Probability Letters*, 23(1): 21-25, 1995.

[19] P. Erdös and A. Rényi. "On a Classical Problem of Probability Theory". Magyar Tud. Akad. Mat. Kutato Int. Kozl, 1961.

[20] R. Motwani and P. Raghavan. "Randomized Algorithms". Cambridge University Press, 1995.