Alano: May You Discover Your Neighbors

Abstract—Neighbor discovery is a crucial step in constructing wireless networks. Various methods have been proposed to promote the discovery rate and minimize the discovery latency. However, none of them has considered the connectivity and the distribution of the nodes in a large-scale network. In a large-scale network, each node has a capacity to sense the nodes within its transmission range which are called neighbors and other nodes are connected by multi-hop communication.

In this paper, we propose Alano, the first nearly optimal probability based algorithm to discover neighbors in large-scale networks. To begin with, we consider the distribution of the nodes in the networks and compute the expected number of neighbors n using the local information. Then, we prove the discovery latency of Alnano is bounded by a low latency $O(nlnn)$. Finally, we propose a Relaxed Difference Set based Alano algorithm (RDS-Alano) to achieve low-latency neighbor discovery process in the symmetric energy-efficient networks and a Traversing Pointer based Alano algorithm (TP-Alano) in the asymmetric energy-efficient networks. Our evaluation shows Alano achieves 31.35% to 32.32 times lower latency than existing methods and holds higher performance in quality, scalability and robustness.

I. INTRODUCTION

The growing interest in the Internet of Things (IoT) has resulted in a number of wide-area deployments of wireless networks [1], such as wireless sensor networks [2], mobile campus networks [3], mobile gaming community [4], etc. All these realistic networks possess multi-hop and large-scale characteristics obviously.

Neighbor discovery is a fundamental step of constructing a wireless network, based on which the network can implement further applications such as routing and broadcasting. The core target is for each node in the network to discover the nodes in its radio sensing range with one-hop communication, which are called neighbors. A number of existing methods [5]–[15] have been proposed to deal with this issue, some of which are based on deterministic techniques, those turn on the radio with deterministic sequences, while others are probabilistic approaches, those turn on the radio with different probabilities.

Unfortunately, despite great efforts, neighbor discovery in a realistic network remains an open problem. The key issue lies in the collision condition in the large-scale networks. In a large-scale network, each node is only able to sense the nodes if their Euclidean distance is no more than a given threshold based on received signal strength [16]–[18]. We call this multi-hop wireless network connectivity as *partially-connected*.

On the one hand, the deterministic approaches only deal with the neighbor discovery problem for two nodes. When transferred to multiple nodes scenario, they can not solve the collision issue when more than one neighbors are transmitting simultaneously.

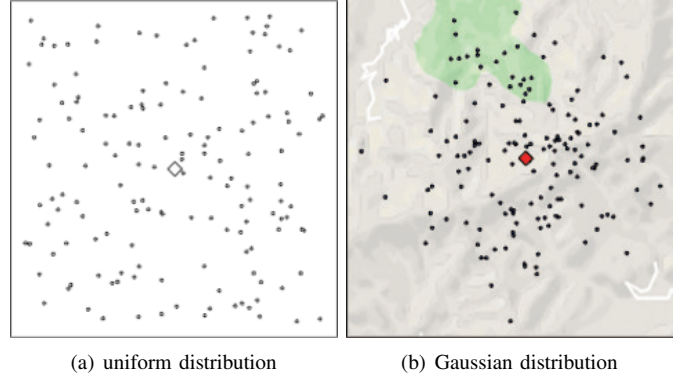(a) uniform distribution   (b) Gaussian distribution

Fig. 1. WSN deployments following uniform and Gaussian distribution.

On the other hand, probabilistic approaches can deal with multiple nodes discovery well but only consider the network is fully-connected, the topology of which is a complete graph. Therefore in the large-scale network where the nodes are partially-connected, the probability adopted in these approaches can not be desirable and thus can not reduce the collision efficiently. To our best knowledge, no relevant researches have focused on the neighbor discovery problem in the large-scale networks where nodes are partially-connected.

Our initiative insight is to take the distribution of the nodes in the network into fully consideration, since in a wireless network the distribution of nodes' deployment directly plays a vital role in determining the intrusion detection capability of a communication device.

In this paper, we propose Alano[1], a nearly optimal probability based algorithm to discover neighbors in large-scale networks. As fully studied in [17] , the nodes in wireless sensor networks are likely to follow a uniform or a Gaussian distribution as showed in Fig. 1, depending on the specific network applications, which is taken into consideration of the proposed Alano algorithm.

In addition, among the large-scale networks, there is a special type named energy-efficient network [19]. Wireless sensor network is a typical energy-efficient network that all the sensor nodes have to maintain strict power budgets to attain years of lifetime [20]. Duty cycle mechanism, the fraction of time the radio is turned on, is utilized to raise the power-awareness of the nodes in the network. Correspondingly, the neighbor discovery process needs adjustments to deal with the dilemma between a balance of energy-efficiency and low-latency.

---

[1]Alano is the god of luck in Greek mythology.

For energy-efficient networks, we design deterministic methods to align the wake-up time slots between the neighbors to achieve lower latency bound. Specifically, We propose RDS-Alano algorithm for the symmetric energy-efficient networks, where all the nodes share a identical duty cycle θ and TP-Alano for the asymmetric energy-efficient networks where each node possesses a respective duty cycle θ_i.

Our simulation shows the proposed Alano algorithms hold significant strengths compared to the state-of-the-art methods, based on the evaluation of speed, quality, scalability and robustness. Alano achieves 31.35% to 32.32 times lower latency and has higher discovery rate during the whole process of neighbor discovery, no matter in symmetric or asymmetric scenario and in uniform or Gaussian distribution. When the number of nodes increases and the network becomes denser, Alano still keeps its high performance.

The main contributions of this paper are summarized as follows:

1) We consider the distribution of nodes and propose Alano, a low-latency algorithm to achieve neighbor discovery process in a large-scale network.

2) We propose a Relaxed Difference Set based Alano algorithm (RDS-Alano) to achieve low-latency neighbor discovery process in the symmetric energy-efficient networks.

3) We propose a Traversing Pointer based Alano algorithm (TP-Alano) to achieve low-latency neighbor discovery process in the asymmetric energy-efficient networks.

The remainder of the paper is organized as follows. The next section highlights some related work and puts forward some serious problems. Some notion definitions and the system model are given in Section III. We analyse the node's expected number of neighbors and propose Alano algorithm in IV as a foundation. Section V describes the RDS-Alano algorithm for symmetric scenario and TP-Alano algorithm for asymmetric scenario respectively in energy-efficient networks. We have conducted extensive simulations, and the results are shown in Section VI. Finally, we conclude the paper in Section VII.

II. RELATED WORK

Neighbor discovery problem has raised a great deal of attention of scholars [21]. A number of neighbor discovery methods have been proposed in the past decade. Technically, these approaches can be classified into two categories, probabilistic and deterministic.

In the deterministic methods [5]–[11], some mathematical techniques, such as co-primality, quorum system, etc., are utilized to promote the discovery performance. The deterministic methods holds an obvious advantage that they can achieve neighbor discovery process for two nodes within a bounded time latency.

Nevertheless, there exists some crucial weak points in the deterministic algorithm. Firstly, Disco [5] proposes a discovery protocol that each node has a capability to send a beacon (one or a few bits) at both beginning and end of an active slot, which is widely adopted by the later algorithms such as SearchLight [7], BlindDate [10] and Hello [8], Nihao [11]. It is quite an efficient way for two nodes to discover each other within an ideal time latency. However, they do not solve the collision issues when receiving packages from multiple neighbors. Furthermore, when they are extended for multiple nodes, only sending a beacon to discover the neighbors is totally insufficient. A node needs to send a complete package (some papers still call beacon) containing all its information, otherwise the neighbor can not identify which neighbor the beacon belongs to [22]. Thus a complete time slot is necessary for a node to transmit a package or listen to the channel to receive a package.

Another category is probabilistic methods [12]–[15]. These approaches utilize probability techniques to promote the randomness to discover the neighbors. Different from the deterministic algorithms, this kind of method shows an significant strength in the multiple nodes scenario. However, probabilistic methods only present an expectation discovery latency and can not guarantee a latency bound in the worst case. In addition, almost all the existing methods consider the network is fully-connected, the topology of which is a complete graph. Deploying a fully-connected network in a large-scale area is technically impractical due to the limited sensing range of devices communication. How far the other node can be detected as a neighbor for a mobile equipment depends on criterion such as the received signal strength [23]. From our analysis and simulations, they show a poor performance in the large-scale networks, where the nodes are partially-connected.

To the best of our knowledge, neither deterministic nor probabilistic methods are designed for the networks which are partially-connected. In this paper, we present a low-latency, energy-efficient neighbor discovery algorithm for the large-scale networks.

The proposed RDS-Alano and TP-Alano in this paper are a combination of both two categories, and thus we compare our proposed algorithms with both deterministic and probabilistic methods. Particularly as mentioned above, the deterministic approaches need some adjustments when transferred to large-scale networks since the protocol can not work. Details will be introduced in Section VI.

III. PRELIMINARIES

In this section, we first describe the network and node model. Then we formulate the Neighbor Discovery problem formally. The notations are listed in Table I.

A. Network and Node Model

Communication model. We study the neighbor discovery problem under the unit-disk model [16], [18], since it is a popular model that enabled the development of efficient algorithms for crucial networking problems, and other models such as *SINR* which lacks of good algorithmic features can be transformed to it by particular means [24]. In the unit-disk model, two network nodes are directly connected if and only if their Euclidean distance is no more than a given threshold

Notation	Description
N	The number of nodes in the network
u_i	node u_i with ID i
u_{ij}	the jth neighbor of node u_i
δ_{ij}	The transmission time drift between u_i and u_j
$n*$	The average number of neighbors is n, $n = p_n N$
t_0	The length of a time slot
$L(i,j)$	The discovery latency that node u_i discovers node u_j
$L(i)$	The discovery latency that node u_i discovers all neighbors
θ	The pre-defined global duty cycle
θ_i	Node u_i's local duty cycle
W	The time slot spent by a node discovering all neighbors
M	Neighboring matrix, $M_{ij} = 1$ means u_i and u_j are neighbors

based on the received signal strength, which is called the transmission range.

Network connectivity. In a large-scale network under the unit-disk model, each node has a capacity to sense the nodes within its transmission range. The nodes which can be connected by one-hop communication are called neighbors, and other nodes can be connected by multi-hop communication. This wireless network connectivity is defined as *Partially-Connected*, contrary to the fully-connected scenario where all the nodes can communicate with each other by one-hop communication.

Multi-hop wireless networks are often modeled as graphs. A symmetric matrix $M_{N \times N}$ is used to record the neighboring relations as:

$$M_{i,j} = \begin{cases} 1 & connected \\ 0 & disconnected \end{cases}$$

Energy-Efficient Networks. In the large-scale networks, there is a particular type called *Energy-Efficient Networks*. A typical one of the energy-efficient networks is the wireless sensor networks. The wireless sensor network consists of a number of sensors distributed separately in a target area. The deployed sensor nodes keep their most time in sleep pattern to avoid quick energy consumption and wake up timely to work on duty.

Network distribution. In a network, the location of the nodes are likely to obey uniform distribution[][], Gaussian distribution[][] or other combinatorial distributions. For a certain, the expected number of neighbors can be calculated the local network density.

Time model. Time is divided into slots of equal length t_0, which is sufficient to finish one communication process (transmit or receive a piece of package). In each time slot, a node transform its pattern according to a pre-defined duty schedule.

Node transmission model. In this paper, we denote the node set in the network as $U = \{u_1, u_2, ..., u_N\}$. When a node wake up in a time slot, it can turn to either the transmitting state or listening state.

- **Transmitting state**. A node turn to transmitting state will broadcast a package containing its own identify information to all neighbors.
- **Listening state**. A node turn to listening state will monitor the frequency channel to collect its neighbors' packages.

Transiting between the states only costs little time, compared to one complete time slot.

Transmission drift. Each node may start their neighbor discovery process at different time slots and we denote δ_{ij} as the asynchronous time drift between a pair of neighbor nodes (u_i, u_j). $\delta_{ij} = 0$ represents a synchronous case.

Transmission collision. Under the unit-disk model, collisions will occur to a node in its listening state when two or more neighbor nodes transmit concurrently and thus no valid information will be gathered.

Definition 1: **Duty Schedule** is a pre-defined sequence $S = \{s^t\}_{0 \le t < T}$ of period T and

$$s^t = \begin{cases} S & Sleep \\ T & Transmit \\ L & Listen \end{cases}$$

Note that the sleep state is only for the energy-efficient networks. Each node construct its own duty schedule according to a specific strategy and repeats it until finding all the neighbors.

Since the waking-up duration has a significant affect on the battery's lifetime, duty cycle is utilized to restrict the energy consumption in the energy-efficient networks.

Definition 2: **Duty Cycle** represents the fraction of one period T where a node turns its radio on. It can be formulated as:

$$\theta = \frac{|\{t : 0 \le t < T, s^t \in \{T, L\}\}|}{T}.$$

B. Problem Definition

We consider a large-scale and multi-hop network, where two nodes are neighbors if they locate within the radio range of each other.

Each node follows its duty schedule to achieve neighbor discovery. Note that the neighbor discovery process is not bidirectional, which means any pair of neighbors need to find each other separately. The time slots within which a node u_i find one of its neighbors u_j can be formulated as $L(i,j)$. Then we define the discovery latency that node u_i discovers all neighbors as:
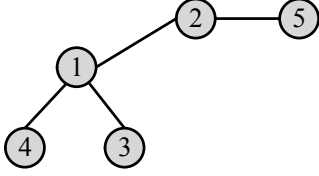
Definition 3: **Discovery Latency** of node u_i is the time to discover all neighbors:

$$L(i) = \max_{j : M_{i,j} = 1} L(i, j).$$

Thus the neighbor discovery problem can be formulated as:

Problem 1: For a node u_i with its neighbor set $S = \{u_{i1}, u_{i2}, ...u_{ij}, ...\}$, design a strategy to construct a duty schedule, which satisfies $\forall$ neighbor nodes u_{ij} :

$$\exists t \ s.t. : \quad S_i(t) = L, S_{ij}(t) = T, \forall k \ne j : S_{ik}(t) \in \{L, S\}.$$

(a) The topology of a simple wireless network

Time	...	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	...
Node 1			T	S	S	S	S	S	L	S	S	L	S	S	S	T	...
Node 2		S	S	L	S	S	S	S	T	S	L	S	S	S	S	L	...
Node 3	...	S	S	S	T	S	S	S	T	S	L	S	S	S	L	S	...
Node 4			S	S	S	S	S	S	S	S	S	S	T	T	T	T	...
Node 5							S	S	L	S	S	S	T	S	S	S	...

(b) Neighbor discovery process

Fig. 2. An example of neighbor discovering process. S, T and L represents Sleep pattern, Transmitting state and Listening state in wake-up pattern respectively.

An example of neighbor discovery process is given in Fig.2. Fig.2(a) shows the topology of a simple wireless network, which consists of 5 nodes. Fig.2(b) describes the neighbor discovery process in the asynchronous scenario, as we can see the nodes start their process at different time slot. Duty cycle is set as 0.25 which implies a node will wake up once during a period of four time slots. The duty schedule of node u_1, for example, is $S_1 = \{T, S, S, S, S, S, S, T, S, S, T, S, S, S, T, ...\}$. The time drift between node u_1 and $u_4 : \delta_{14} = 1$ while u_1 and u_2 start their process synchronously. At time slot 12, node u_5 discovers its neighbor node u_2 while node u_1 could not discover node u_2 due to a collision from its another neighbor node u_3.

IV. LARGE-SCALE NETWORKS

In a large-scale network and the unit-disk model, all the nodes are partially-connected with each other. A node can discover the nodes within its sensing range.

In the network for $\forall$ node u_i, it is possible for the nodes to be aware of its position coordinate (x_i, y_i) and thus get the local density by a pre-defined density function, based on the deployment of the network application.

The expected number of neighbors can be computed as follows [18], [25].

Denote the density function as :

$$f(x,y) = \begin{cases} \varphi(x,y) & (x,y) \in D \\ 0 & (x,y) \notin D \end{cases}$$

where D is the network covering area.

$\forall$ node $u_i(x_i, y_i)$, its sensing range area R_i can be formulated as :

$$(x - x_i)^2 + (y - y_i)^2 \leq r^2$$

where r is the detection radius.

Thus, we can obtain the expected number of neighbors of node u_i as :

$$NB(u_i) = N \iint_{R_i} f(x,y)\, dx\, dy - 1.$$

In the large-scale networks, we ignore the boundary area of the network and assume the nodes in the network is of an enormous quantity, so the expected number of neighbors can be formulated as:

$$NB(u_i) = N \iint_{R_i} \varphi(x,y)\, dx\, dy.$$

Note that, when the network area is far more larger than the sensing area of the nodes, we can approximately get:

$$NB(u_i) = N\pi r^2 \varphi(x,y).$$

Then we propose **Alano**, a randomized neighbor discovery algorithm. We describe the algorithm for $\forall$ node u_i in Alg. 1. Alano algorithm indicates that what probability for a node choose to turn to transmitting state or listening state is determined by the expected number of neighbors varying from node to node.

Algorithm 1 Alano Algorithm

1: $\hat{n}_i = N \iint_{R_i} \varphi(x,y)\, dx\, dy$;
2: $p_t^i = \frac{1}{\hat{n}_i}$;
3: **while** $True$ **do**
4:      A random float $\epsilon \in (0,1)$;
5:      **if** $\epsilon < p_t$ **then**
6:          Transmit a message containing node information of u_i;
7:      **else**
8:          Listen on the channel and decode the node information if receive a message successfully;
9:      **end if**
10: **end while**

In the following section IV-A, we first consider a general situation that the nodes in the network are uniform distributed. We derive a proof that the probability chosen in Alano is the optimal one and show the bounded latency will not be much larger than its expectation. Then in the section IV-B we describe a more common situation that the nodes in the network obey Gaussian Distribution, we present a approximation analysis that the discovery latency will not be much larger than uniform distribution.

A. Uniform Distribution

Uniform distribution is a basic one for the deployment of wireless networks. For instance, consider there is a wireless sensor network carrying out a task of measuring temperature and humidity in a target area, thus the sensors are supposed to be evenly deployed and the density function can be formulated as:

$$f(x) = \begin{cases} \frac{1}{A} & (x,y) \in D \\ 0 & (x,y) \notin D \end{cases}$$

where A is the area of D.

Every node in the network has the same expected number of neighbors and transmit with the same probability as:

$$\hat{n} = \frac{N\pi r^2}{A}, \quad p_t = \frac{1}{\hat{n}} = \frac{A}{N\pi r^2}.$$

According to Alano, the probability that node u_i discover a specific neighbor successfully in a time slot can be formulated as:

$$p_s = p_t(1 - p_t)^{\hat{n}-1}.$$

Let:

$$p_s' = (1 - p_t)^{\hat{n}-1} - (\hat{n} - 1)p_t(1 - p_t)^{\hat{n}-2} = 0.$$

It is easy to confirm that when

$$p_t = \frac{1}{\hat{n}}.$$

p_s gets the maximum value:

$$p_s = \frac{1}{\hat{n}}(1 - \frac{1}{\hat{n}})^{\hat{n}-1} \approx \frac{1}{\hat{n}e}.$$

Thus we can conclude that the probability chosen in Alano to transmit is the optimal one.

Next we analyse the expected latency for a node to discover all its neighbors. We denote W_j to be a random variable representing the number of the time slots needed to discover a new neighbor after $(j - 1)$ neighbors have been discovered, which follows Geometric distribution with parameter $p(j)$: $p(j) = (\hat{n} - j + 1)p_s$. Then the expectation of W_j is computed as:

$$E[W_j] = \frac{1}{p(j)} = \frac{1}{(\hat{n} - j + 1)p_s}$$

The expectation time latency of discovering all the neighbors can be formulated as:

$$E[W_j] = \sum_{j=1}^{\hat{n}} \frac{1}{p_s} H_n \approx ne(lnn + \Theta(1)) = \Theta(nlnn).$$

where H_n is the n-th Harmonic number, i.e., $H_n = lnn + \Theta(1)$.

We get the expectation discovery latency is within O(nlogn) and then we show the bounded latency will not be much larger than its expectation.

If W_i is given, the value of W_j will not be affected for $i < j$. That is, for $i \neq j$, W_i and W_j are independent and they satisfy $P(W_j = w_j | W_i = w_i) = P(W_j = w_j)$. Since W_j follows Geometric distribution, and $Var[W_j] = \frac{1-p_j}{p_j^2}$, the variance of W is

$$Var[W] = \sum_{j=1}^{n} Var[W_j] \leq \frac{\pi^2}{6p_{suc}^2} - \frac{H_n}{p_{suc}}.$$

With *Chebyshev's inequality*, the probability that the discovery time is 2 times larger than the expectation is

$$P[W \geq 2E[W]] \leq \frac{Var[W]}{E[W]^2} \leq \frac{\pi^2}{6H_n^2} - \frac{p_{suc}}{H_n}.$$

For large n, $P[W \geq 2E[W]]$ is close to 0. That is, the time for a node to find all neighbors is very likely to be smaller than 2 times of expected latency. Therefore,

$$W = O(nlnn).$$

B. Gaussian Distribution

A common scenario is that the nodes in a network obey 2D Gaussian distribution. For example, an intrusion detection application may need improved detection probability around important entities [17].

In this section, we present a theoretical proof that the latency performance will not be much larger than uniform distribution. The analysis is not only applicable for Gaussian distribution but also flexible for all the other distributions in a large-scale network.

We first denote the approximate neighbors of node u_i as set $S(u_i) = \{u_{i1}, u_{i2}, ..., u_{i\hat{n}_i}\}$. When the nodes obey Gaussian distribution, according to Alano, the probability that node u_i discovers a certain neighbor node u_{ij} successfully in a time slot can be formulated as:

$$p_{suc} = (1 - p_t^i)p_t^{ij} \prod_{k=1}^{\hat{n}_i, k \neq j} (1 - p_t^{ik})$$

Denote:

$$p_t^{imax} = \max_{1 \leq j \leq \hat{n}_i} \{p_t^{ij}\}, \quad p_t^{imin} = \min_{1 \leq j \leq \hat{n}_i} \{p_t^{ij}\}.$$

Thus for $\forall j, 1 \leq j \leq n_i$:

$$(1 - p_t^i)p_t^{imin}(1 - p_t^{imax})^{\hat{n}_i - 1}$$

$$\leq (1 - p_t^i)p_t^{ij} \prod_{k=1}^{\hat{n}_i, k \neq j} (1 - p_t^{ik})$$

$$\leq (1 - p_t^i)p_t^{imax}(1 - p_t^{imin})^{\hat{n}_i - 1}$$

Denote:

$$P = (1 - p_t^i)p_t^{imin}(1 - p_t^{imax})^{\hat{n}_i - 1}$$

$$Q = (1 - p_t^i)p_t^{imax}(1 - p_t^{imin})^{\hat{n}_i - 1}$$

Thus we get:

$$\frac{1}{\hat{n}_i Q} \leq E[W_1] \leq \frac{1}{\hat{n}_i P}$$

$$\frac{1}{(\hat{n}_i - 1)Q} \leq E[W_2] \leq \frac{1}{(\hat{n}_i - 1)P}$$

$$..........$$

$$\frac{1}{Q} \leq E[W_{\hat{n}_i}] \leq \frac{1}{P}$$

Sum all the equations above, and we will get:

$$\frac{1}{Q}H_n \leq E[\sum_{j=1}^{\hat{n}_i} W_j] \leq \frac{1}{P}H_n$$

Since the sensible neighbors are within a close distance of the node compared to the total network area, which implies the density function values are within the same order of

magnitude. Thus we can conclude that the expectation of the time latency in the normal distributed networks are still $E[W] = O(nlnn)$. Similarly the bounded latency can be proved to be $W = O(nlnn)$ in the same way as uniform distribution.

V. ENERGY-EFFICIENT NETWORKS

In an energy-efficient network (i.e., wireless sensor networks), the battery consumption is a crucial factor to be taken into account. Duty cycle is a key technique to deal with the dilemma between a balance of energy-efficiency and low-latency.

We first consider the symmetric energy-arrangement situation that all the nodes share a global duty cycle θ, and propose a RDS based Alano algorithm. Then we propose a traversing pointer based Alano algorithm for a more general scenario, where nodes have asymmetric battery-scheduling capability with local duty cycle θ_i.

Our initiative idea is to align the wake-up slots of the neighbor nodes within a bounded time, and then invoke the Alano algorithm to achieve neighbor discovery process w.h.p. More specifically, we utilize the property of RDS and traversing pointer to guarantee a wake-up slot rendezvous in each period T.

A. Symmetric scenario: RDS-Alano Algorithm

When a global duty cycle θ is shared by all the nodes in the network, we utilize relaxed difference set (RDS) to align the wake-up time slots.

Relaxed difference set (RDS) is an efficient tool to construct cyclic quorum systems [26], [27]. The definition can be described as:

Definition 4: A set $R = \{a_1, a_2, ..., a_k\} \subseteq Z_n$ (the set of all nonnegative integers less than n) is called a Relaxed Difference Set (RDS) if for every $d \neq 0$ (mod n), there exists at least one ordered pair (a_i, a_j) such that $a_i - a_j \equiv d$ (mod n), where $a_i, a_j \in D$.

It has been proved that any RDS must have cardinality $|R| \geq \sqrt{N}$ [27]. We present a simple linear algorithm for RDS construction under Z_N with $\lceil \frac{3\sqrt{N}}{2} \rceil$ cardinality in Alg. 2.

Algorithm 2 RDS construction under Z_N

1: $R := \emptyset$;
2: $\lambda := \lceil \sqrt{N} \rceil, \mu := \lceil \frac{\lceil \sqrt{N} \rceil}{2} \rceil$;
3: **for** $i = 1 : \lambda$ **do**
4: $\quad R := R \cup i$;
5: **end for**
6: **for** $j = 1 : \mu$ **do**
7: $\quad R := R \cup (1 + j * \lambda)$;
8: **end for**

The initiative idea of Alg. 2 can be described as Fig. 3. The framed elements are selected as in Alg. 2 Line. 4 and Line. 7.

We give a formal correctness proof of the construction as following:

1	2	3	...	λ
$1+\lambda$	...		...	...
$1+2\lambda$		...	...	
...	...		...	...
$1+(\mu-1)\lambda$		N/2	...	
$1+\mu\lambda$				
...	...		...	...
$1+(\lambda-1)\lambda$	...		N	λ²

Fig. 3. An Sketch of RDS construction in Alg. 2

Theorem 1: The set $R = \{r_0, r_1, ..., r_{\lambda+\mu-1}\}$ constructed in Alg. 2 is a RDS, where $|R| = \lambda + \mu = \lceil \sqrt{N} \rceil + \lceil \frac{\lceil \sqrt{N} \rceil}{2} \rceil \approx \lceil \frac{3\sqrt{N}}{2} \rceil$.

Proof: We first reach a consensus that if there exists one ordered pair (a_i, a_j) satisfying $a_i - a_j \equiv d$ (mod N), then we can get an opposite pair (a_j, a_i) such that $a_j - a_i \equiv (N - d)$ (mod N). Thus we only need to find at least one ordered pair (a_i, a_j) for each d from 1 to $\lfloor N/2 \rfloor$.

The λ in Line 2 is the smallest integer satisfying $\lambda^2 \geq N$. Then every d from 1 to $\lfloor N/2 \rfloor$ can be represented as: $d = 1 + j \times \lambda - i$, where $1 \leq j \leq \mu, 1 \leq i \leq \lambda$. Thus there exists $a_j = 1 + j \times \lambda$ added in Line. 4 and $a_i = i$ added in Line. 7 satisfying $a_j - a_i \equiv d$. ∎

Next, we present a RDS based Alano algorithm (RDS-Alano) in Alg. 3, to achieve neighbor discovery process in a partially-connected and energy-efficient network with global duty cycle θ.

In Alg. 3, RDS is used to construct a deterministic schedule for the node to wake up in every period T, and Alano is utilized as a probabilistic strategy to determine the transmission state (transmit or listen) in each wake-up slot.

Algorithm 3 RDS Based Alano Algorithm

1: $T := \lceil \frac{9}{4\theta^2} \rceil$;
2: Invoke Alg. 2 to construct the RDS $R = r_0, r_1, ..., r_{\lceil \frac{3\sqrt{T}}{2} \rceil}$ under Z_T;
3: $t := 0$;
4: **while** $True$ **do**
5: $\quad$ **if** $(t+1) \in R$ **then**
6: $\quad\quad$ Invoke Alg. 1 to determine transmission state;
7: $\quad$ **else**
8: $\quad\quad$ Sleep;
9: $\quad$ **end if**
10: $\quad$ t := (t + 1) % T;
11: **end while**

We show a proof of time latency bound for Alg. 3 as following:

Theorem 2: Alg. 3 guarantees the discovery latency to be bounded within $O(\frac{nlog}{\theta^2})$ w.h.p.

Proof: It is easy to confirm that the duty cycle $\widetilde{\theta}$ in Alg.

3 corresponds to θ as:

$$\widetilde{\theta} = \frac{|RDS|}{|T|} = \frac{\lceil \frac{3\sqrt{T}}{2} \rceil}{T} = \theta.$$

For any pair of neighbor nodes (u_i, u_j), we can find an ordered pair (r_i, r_j) from their respective RDS such that $r_i - r_j \equiv \delta_{ij}$ (mode T), which indicates any neighbor pair can wake up in the same time slot at least once in every period T. Regarding the whole period T as a time slot in Alg. 1, we obtain the latency bound as $O(\frac{nlogn}{\theta^2})$ w.h.p. ∎

There is a detail to be noticed when regarding the whole period T as a time slot in Alg. 1. Due to the periodicity, a node will rendezvous the same neighbors in the identical wake-up slot in each period. Since for a certain node there are more than one wake-up slots in each period, the rendezvoused neighbors in each wake-up time slot may be equal or less than the total number of its neighbors, resulting less collisions and lower latency compared to when all the neighbors wake up in the same time slot. Thus the latency bound can not be larger.

B. Asymmetric scenario: TP-Alano Algorithm

For a more practical scenario, the nodes in a wireless sensor networks for instance, are assigned to diverse tasks such as temperature measurement, sunshine collection, etc., and thus ought to have asymmetric capability of battery-management with local duty cycle θ_i.

We propose a traversing pointer based Alano algorithm(TP-Alano) in Alg. 4. In each period T, every node wakes up in two different time slots, one of which is the first slot of each period and another is a traversing slot different from period to period, as Alg. 4 Line. 5 indicates.

Algorithm 4 Traversing Pointer Based Alano Algorithm

1: $T :=$ Find the smallest prime $\geq \frac{2}{\theta_i}$;
2: $t := 0$;
3: **while** $True$ **do**
4:     $t_1 := t\%T$;
5:     $t_2 := \lfloor t/T \rfloor \%(T-1) + 1$;
6:     **if** $t_1 = 0 || t_1 = t_2$ **then**
7:         Invoke Alg. 1 to determine transmission state;
8:     **else**
9:         Sleep;
10:     **end if**
11:     $t := t + 1$;
12: **end while**

We call the first time slot in each period T as a *fixed pointer* and the traversing slots as a *traversing point*. These pointers are used to guarantee a wake-up time rendezvous in every period $T_i T_j$. A sketch of the pointers is described in Fig. 4.

Remark 1: Since the period of T is selected as: *find the smallest prime $\geq \frac{2}{\theta_i}$*, which is likely to result in the consequence that the duty cycle $\widetilde{\theta}_i$ in Alg. 4 is smaller than the expected θ. This can be easily solved by selecting some

random wake-up time slots in each period T to conform to duty cycle θ.
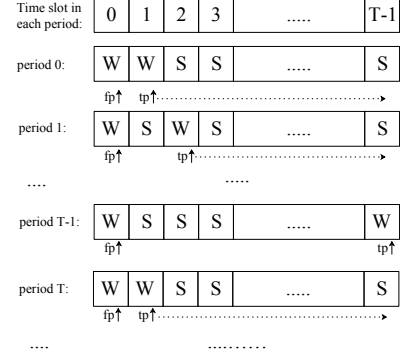

Fig. 4. An Sketch of TP construction in Alg. 4

We give a correctness proof of the time bound to achieve neighbor discovery process as following:

Theorem 3: Alg. 4 guarantees the discovery latency to be bounded within $O(\frac{nlogn}{\theta_i \theta_j})$ w.h.p., where θ_i and θ_j are the duty cycles of a pair of neighbor nodes (u_i, u_j) respectively.

Proof: We first prove that any pair of nodes (u_i, u_j) wake up at the same time slot every period $T_i T_j$.

Case 1: $T_i \neq T_j$. Since T_i and T_j are different primes, according to Chinese remainder theorem, there exists a time slot $t_\tau \in [0, T_i T_j)$ satisfying:

$$0 = t_\tau \quad mod \quad T_i. \tag{1}$$

$$\delta_{ij} = t_\tau \quad mod \quad T_j. \tag{2}$$

Equation1 and 2 implies there exists a fixed pointer of node u_i and a fixed pointer of node u_j rendezvous in every $T_i T_j$.

Case 2: $T_i = T_j$. Since $T_i = T_j = T$, if the asynchronous drift between node u_i and u_j : $\delta_{ij} = 0$, the fixed pointers of node u_i and node u_j will rendezvous with each other in every period T. Otherwise since the traversing point will traverse all the time slots once during period $(T - 1)T$, there exists a traversing point of node u_i rendezvous with a fixed pointer of node u_j every period $(T - 1)T$, and a traversing point of node u_j will consequentially rendezvous with a fixed pointer of node i once every period $(T - 1)T$ as well.

Thus for any pair of neighbor nodes (u_i, u_j), they can wake up at the same time slot at least once in every period $T_i T_j$. Regarding the whole period $T_i T_j$ as a time slot in Alg. 1, we obtain the latency bound as $O(\frac{nlogn}{\theta_i \theta_j})$ w.h.p. ∎

VI. EVALUATION

We implemented Alano in C++ and evaluated the algorithms in a cluster of 9 servers, each equipped with an Intel Xeon 2.6GHz CPU with 24 hyper-threading cores, 64GB memory and 1T SSD.

We simulated the network that follows uniform distribution and normal distribution respectively. We consider 500 nodes with radio range 10 in 100×100 area following uniform distribution, and more generally, 1000 nodes with radio range 5 following normal distribution $N(50, 15^2)$. The duty cycle is

0.1, and each time slot represents $20ms$. These settings make the network more complicated and realistic than [6]–[15].

We evaluated discovery latency of Alano, Aloha-like [14], Hello [8], Hedis [9], and Searchlight [7] in partially-connected network. As the deterministic algorithms, Hello, Hedis and Searchlight, only have two states, $\{ON, OFF\}$. To compare fairly, we generally assume that when nodes are in $\{ON\}$ state, they transmit and listen with equal probability. We show that Alano has lower latency, higher discovery rate, better scalability, and robustness.
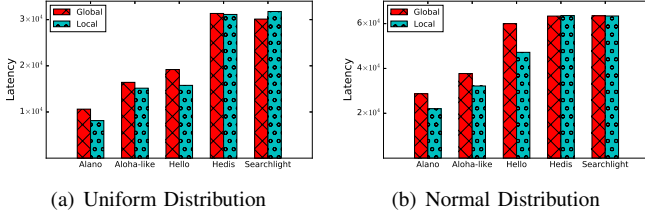
A. Speed: Discovery Latency

(a) Uniform Distribution (b) Normal Distribution

Fig. 5. Alano achieves lower latency.

In Fig. 5, when nodes follow uniform distribution, Alano has 54.64% to 1.95 times lower latency with gloabal duty cycle, and 85.25% to 2.91 times lower latency with local duty cycle. When nodes follow normal distribution, Alano has 31.35% to 1.21 times lower latency with gloabal duty cycle, and 45.94% to 1.88 times lower latency with local duty cycle. The deterministic algorithms Hello, Hedis and Searchlight have high latency, because their design just considered the bounded latency within two nodes. When the network becomes denser and some nodes have more than one neighbors, they cannot discover rapidly, because collisions happen so frequently.

B. Quality: Discovery Rate

(a) Uniform Distribution with Global Duty Cycle

(b) Normal Distribution with Global Duty Cycle

(c) Uniform Distribution with Local Duty Cycle
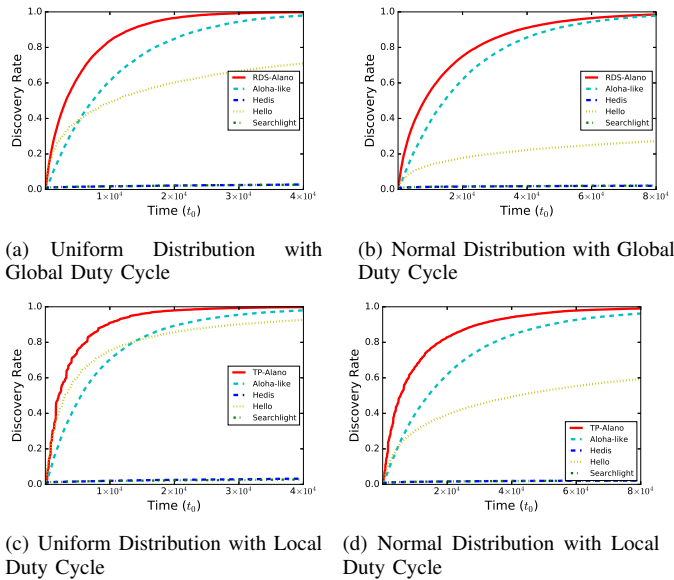
(d) Normal Distribution with Local Duty Cycle

Fig. 6. Alano achieves higher discovery rate in larger networks.

When we increase the number of nodes in the uniform distributed network from 500 to 1000, and the number of nodes in the normal distributed network from 1000 to 2000, Fig. 6 shows that Alano with either global or local duty cycle, has higher discovery rate during the whole course of neighbor discovery in both uniform and normal distribution. The deterministic algorithms Hello, Hedis and Searchlight cannot discover all channels, because of the occurence of collisions.

C. Scalability: Duty Cycle and Network Density

1. Duty Cycle

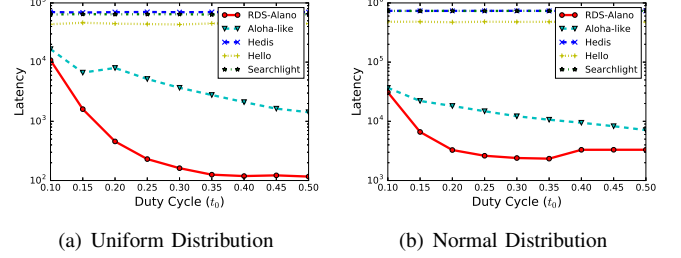(a) Uniform Distribution (b) Normal Distribution

Fig. 7. Alano achieves lower latency in different duty cycle.

With different duty cycles, Fig. 7 shows that Alano has lower latency. Compared with Aloha, Alano has from 53.66% to 11.23 times lower latency. The latency of Alano and Aloha generally decreases as the duty cycle increases, while Hello, Hedis and Searchlight have high latency due to the collision. In normal distribution, Alano has a small twist with duty cycle 0.35, because when the duty cycle increases, nodes are more likely to transmit and therefore collide.

2. Network Density

When the number of nodes increases and the network becomes denser, Alano still shows 4.68 times to 6.51 times lower latency than Aloha-like in uniform distribution, and 25.23% to 1.03 times lowever latency in normal distribution in Fig. 8. Here we compare Alano with Aloha-like, because Hello, Hedis and Searchlight can hardly discover neighbors in denser networks.
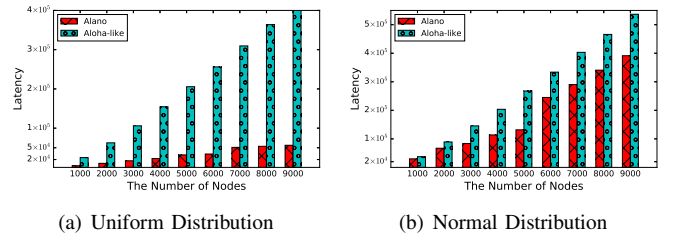

(a) Uniform Distribution (b) Normal Distribution

Fig. 8. Alano achieves lower latency with different number of nodes.

D. Robustness

In reality, nodes exhaust the energy now and then, and it is important for neighbor discovery protocols to remain good performance when nodes go off duty. Fig. 9 shows that when 5% to 30% nodes die, Alano still reach low latency robustly.
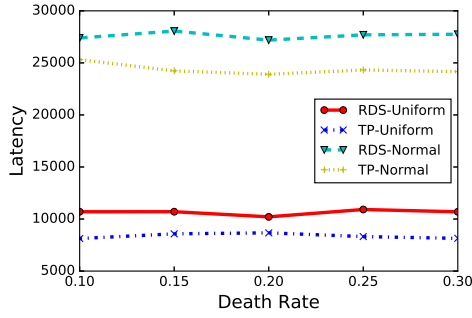
Fig. 9. Alano still reaches low latency when nodes die.

VII. CONCLUSION

The conclusion goes here.
..
..
..
..
..
..
..
..
..

REFERENCES

[1] Z. Qin, G. Denker, C. Giannelli, P. Bellavista, and N. Venkatasubramanian, "A software defined networking architecture for the internet-of-things," in *Network Operations and Management Symposium (NOMS), 2014 IEEE*. IEEE, 2014, pp. 1–9.

[2] J. Yick, B. Mukherjee, and D. Ghosal, "Wireless sensor network survey," *Computer networks*, vol. 52, no. 12, pp. 2292–2330, 2008.

[3] F. Hernandez-Campos and M. Papadopouli, "A comparative measurement study the workload of wireless access points in campus networks," in *Personal, Indoor and Mobile Radio Communications, 2005. PIMRC 2005. IEEE 16th International Symposium on*, vol. 3. IEEE, 2005, pp. 1776–1780.

[4] R. Cunningham, "Optimizing location-based mobile gaming applications," May 31 2002, uS Patent App. 10/160,365.

[5] P. Dutta and D. Culler, "Practical asynchronous neighbor discovery and rendezvous for mobile sensing applications," in *Proceedings of the 6th ACM conference on Embedded network sensor systems*. ACM, 2008, pp. 71–84.

[6] A. Kandhalu, K. Lakshmanan, and R. R. Rajkumar, "U-connect: a low-latency energy-efficient asynchronous neighbor discovery protocol," in *Proceedings of the 9th ACM/IEEE international conference on information processing in sensor networks*. ACM, 2010, pp. 350–361.

[7] M. Bakht, M. Trower, and R. H. Kravets, "Searchlight: Won't you be my neighbor?" in *Proceedings of the 18th annual international conference on Mobile computing and networking*. ACM, 2012, pp. 185–196.

[8] W. Sun, Z. Yang, K. Wang, and Y. Liu, "Hello: A generic flexible protocol for neighbor discovery," in *INFOCOM, 2014 Proceedings IEEE*. IEEE, 2014, pp. 540–548.

[9] L. Chen, R. Fan, K. Bian, M. Gerla, T. Wang, and X. Li, "On heterogeneous neighbor discovery in wireless sensor networks," in *Computer Communications (INFOCOM), 2015 IEEE Conference on*. IEEE, 2015, pp. 693–701.

[10] K. Wang, X. Mao, and Y. Liu, "Blinddate: A neighbor discovery protocol," *IEEE Transactions on Parallel and Distributed Systems*, vol. 26, no. 4, pp. 949–959, 2015.

[11] Y. Qiu, S. Li, X. Xu, and Z. Li, "Talk more listen less: Energy-efficient neighbor discovery in wireless sensor networks," in *Computer Communications, IEEE INFOCOM 2016-The 35th Annual IEEE International Conference on*. IEEE, 2016, pp. 1–9.

[12] M. J. McGlynn and S. A. Borbash, "Birthday protocols for low energy deployment and flexible neighbor discovery in ad hoc wireless networks," in *Proceedings of the 2nd ACM international symposium on Mobile ad hoc networking & computing*. ACM, 2001, pp. 137–145.

[13] S. Vasudevan, D. Towsley, D. Goeckel, and R. Khalili, "Neighbor discovery in wireless networks and the coupon collector's problem," in *Proceedings of the 15th annual international conference on Mobile computing and networking*. ACM, 2009, pp. 181–192.

[14] L. You, Z. Yuan, P. Yang, and G. Chen, "Aloha-like neighbor discovery in low-duty-cycle wireless sensor networks," in *Wireless Communications and Networking Conference (WCNC), 2011 IEEE*. IEEE, 2011, pp. 749–754.

[15] T. Song, H. Park, and S. Pack, "A probabilistic neighbor discovery algorithm in wireless ad hoc networks," in *Vehicular Technology Conference (VTC Spring), 2014 IEEE 79th*. IEEE, 2014, pp. 1–5.

[16] T. Moscibroda and R. Wattenhofer, "The complexity of connectivity in wireless networks." in *Infocom*, 2006, pp. 1–13.

[17] Y. Wang, W. Fu, and D. P. Agrawal, "Gaussian versus uniform distribution for intrusion detection in wireless sensor networks," *IEEE Transactions on Parallel and Distributed systems*, vol. 24, no. 2, pp. 342–355, 2013.

[18] L. Wang, P.-J. Wan, and W. Washington, "Connectivity of multihop wireless networks with log-normal shadowing," *Wireless Networks*, vol. 21, no. 7, pp. 2279–2292, 2015.

[19] C. E. Jones, K. M. Sivalingam, P. Agrawal, and J. C. Chen, "A survey of energy efficient network protocols for wireless networks," *wireless networks*, vol. 7, no. 4, pp. 343–358, 2001.

[20] A. Dunkels, "The contikimac radio duty cycling protocol," 2011.

[21] W. Sun, Z. Yang, X. Zhang, and Y. Liu, "Energy-efficient neighbor discovery in mobile ad hoc and wireless sensor networks: A survey," *IEEE Communications Surveys & Tutorials*, vol. 16, no. 3, pp. 1448–1459, 2014.

[22] G. Zhou, T. He, S. Krishnamurthy, and J. A. Stankovic, "Impact of radio irregularity on wireless sensor networks," in *Proceedings of the 2nd international conference on Mobile systems, applications, and services*. ACM, 2004, pp. 125–138.

[23] V. Daiya, J. Ebenezer, S. S. Murty, and B. Raj, "Experimental analysis of rssi for distance and position estimation," in *Recent Trends in Information Technology (ICRTIT), 2011 International Conference on*. IEEE, 2011, pp. 1093–1098.

[24] E. Lebhar and Z. Lotker, "Unit disk graph and physical interference model: Putting pieces together," in *Parallel & Distributed Processing, 2009. IPDPS 2009. IEEE International Symposium on*. IEEE, 2009, pp. 1–8.

[25] R. Meester and R. Roy, *Continuum percolation*. Cambridge University Press, 1996, vol. 119.

[26] J.-R. Jiang, Y.-C. Tseng, C.-S. Hsu, and T.-H. Lai, "Quorum-based asynchronous power-saving protocols for ieee 802.11 ad hoc networks," *Mobile Networks and Applications*, vol. 10, no. 1-2, pp. 169–181, 2005.

[27] W.-S. Luk and T.-T. Wong, "Two new quorum based algorithms for distributed mutual exclusion," in *Distributed Computing Systems, 1997., Proceedings of the 17th International Conference on*. IEEE, 1997, pp. 100–106.