Neighbor Discovery in Wireless Networks Using Compressed Sensing with Reed-Muller Codes

Lei Zhang and Dongning Guo
Department of Electrical Engineering and Computer Science
Northwestern University
Evanston, IL 60208, USA

***Abstract*—A novel scheme for full-duplex neighbor discovery in wireless networks is proposed. The scheme allows all nodes to simultaneously discover one-hop neighbors and identify their network interface addresses (NIAs) within a single frame of transmission, which typically consists of no more than a few thousand symbols. The key technique is to assign each node a unique *on-off signature* derived from a second-order Reed-Muller code and let all nodes simultaneously transmit their signatures. Despite that the radio is half-duplex, each node observes a superposition of its neighbors' signatures (partially) through its own off-slots. To identify its (small number of) neighbors out of a large network address space, each node solves a *compressed sensing* (or *sparse recovery*) problem using a chirp reconstruction algorithm. A network of over one million Poisson distributed nodes (with 20-bit NIAs) is studied numerically, where each node has 30 neighbors on average, and the channel between each pair of nodes is subject to path loss and Rayleigh fading. Within a single frame of 4,096 symbols, nodes can discover their respective neighbors with on average 99.9% accuracy at 16 dB signal-to-noise ratio (SNR). The algorithm is scalable to networks of virtually any size of practical interest due to its sub-linear complexity. The new scheme requires much fewer transmissions than conventional random-access discovery schemes to achieve the same performance.**

I. Introduction

In many wireless networks, such as a mobile ad hoc network (MANET), a node has direct radio link to only a small number of other nodes due to path loss and other forms of attenuation. Those nodes are called its *neighbors (or peers)*. Before efficient routing or other network-level activities are possible, nodes have to discover and identify the network interface addresses (NIAs) of their neighbors. This is called *neighbor discovery (or peer discovery)*. The problem is crucial in self-organizing networks without pre-existing infrastructure. It is becoming more important in infrastructure networks due to development of heterogeneous cellular networks with unsupervised picocells and femtocells.

State-of-the-art neighbor discovery protocols, such as that of the IETF MANET working group [1] and the ad hoc mode of IEEE 802.11 standards, can basically be described as follows: A *query node* broadcasts a probe request. Each neighbor responds by repeatedly transmitting its NIA in probe response frames interleaved with random delays. If a response frame does not collide with any other frame, the corresponding NIA is successfully received. After long enough time, all neighbors can be discovered with high probability. We refer to such a scheme as *random-access neighbor discovery*. Because of intermittent transmissions, all nodes can be query nodes and also be discovered by its neighbors at the same time during the neighbor discovery period. Several such algorithms which operate in or on top of medium access control (MAC) layer have been proposed in the literature [2]–[5].

We make several important observations: 1) Each query node makes measurements through a multiaccess channel, where the received signal is a superposition of transmissions from its neighbors, corrupted by noise and interference from non-neighbors; 2) For each query node, to acquire neighbors' NIAs is equivalent to identifying, out of the list of all NIAs, which ones are used by its neighbors; 3) The number of neighbors a node has is typically orders of magnitude smaller than the size of the NIA space, so neighbor discovery is by nature a *compressed sensing (or sparse recovery)* problem [6], [7]. This is the reason why the number of measurements (i.e., the number of symbol intervals in a discovery period) needed can be dramatically smaller than the NIA space.

The key to efficient neighbor discovery is thus to design node responses, which we shall refer to as node signatures, as well as the compressed sensing algorithm. Random-access discovery can be regarded as requiring node response to assume a specific format, namely, a node's signature over the discovery period consists of repetitions of its NIA interleaved with periods of silence. Although such signature format allows the NIA to be directly read out from a successfully received frame, it is rather wasteful because of the structural restrictions. In fact, most of the retransmissions are redundant to most neighbors.

In this work, we propose a novel neighbor discovery scheme which assigns the nodes a set of deterministic on-off signatures based on a second-order Reed-Muller (RM) code. RM codes date back to 1950s and have been shown to be excellent for sparse recovery [8]. A second-order RM code consists of quadrature phase-shift keying (QPSK) symbols. The encoding complexity is polynomial in the code length. Importantly, the reconstruction complexity is sub-linear in the NIA space, thus the scheme is scalable to extremely large networks.

A crucial constraint addressed in this work is the half-duplex nature of wireless nodes, which prevents a radio from receiving any useful signal at the same time and over the same frequency band within which it is transmitting. The physical reason is that during transmission, a radio's own signal picked up by its receive antenna is many orders of magnitude stronger than

978-1-61284-824-2/11/$26.00 ©2011 IEEE

the signals from other nodes, such that the desired signals are lost due to the limited dynamic range of the radio frequency (RF) circuits. This is likely to remain a key constraint in the foreseeable future. If the original RM code consisting of QPSK symbols is used, a transmitting node cannot hear other nodes. In order to achieve full-duplex discovery, we introduce off-slots by replacing a half of the QPSK symbols by zeros, so that a node can receive useful signals through them. The chirp reconstruction algorithm of [8] is modified to perform well (with a longer code length) despite that half of the symbols are erased.

The proposed scheme is computationally feasible for extremely large networks. Numerical results in Section V verify the feasibility and effectiveness of the scheme for a network with over one million nodes. It is, in principle, straightforward to extend the scheme to support networks with 2^{48} addresses (i.e., all IEEE 802.11 MAC addresses) or more. As is shown in a comparison in Section VI, the proposed scheme needs much fewer transmissions than conventional random-access schemes for achieving the same error performance. In addition, the new scheme entails much less transmission overhead (such as preambles and parity checks), because it takes a single frame of transmission as opposed to many frame transmissions in random-access discovery.

II. Related Work and Contributions of This Paper

In today's wireless networks, the de facto standard signaling of half-duplex radios is to alternate between transmitting frames and receiving frames. In this case, in order for neighboring nodes to hear and discovery each other without other coordination, random access with repeated and randomly delayed frame transmissions is an evident solution. Several random-access discovery schemes have been analyzed in the literature (see, e.g., [9]). The scenario that nodes have a reliable collision detection mechanism is also considered in [10], [11]. Such schemes typically operate in the medium access control (MAC) or higher layers under the assumption of a collision channel model, where overlapping frames at a receiver obliterate each other. Most recently, Qualcomm has developed the FlashLinQ technology [12] based on orthogonal frequency-division multiplexing (OFDM). It carries out neighbor discovery over a large number of orthogonal time-frequency slots [13], which is essentially random access, where multiple subcarriers allow efficient use of channel resources.

A precursor of the current work is [14], [15], in which we proposed a scheme which assigns each node a pseudo-random *on-off signature*. The number of on-slots is a small fraction of the total number of slots, so that the signatures are sparse. The superposition of the neighbors' signatures is a denser sequence of pulses, in which a pulse is seen at a slot if at least one of the neighbors sent a pulse during that slot. A simple decoding procedure via eliminating non-neighbors was developed based on algorithms originally introduced for *group testing*. The problem with random signatures is that decoding requires visiting all signatures so that the complexity is at least linear in the address space, which renders the scheme infeasible for more than 20-bit NIAs, or a million nodes. Indeed, although random codes generally perform as good as the best code according to Shannon's random coding argument, it is evident that, some structure has to be introduced in the codes for efficient decoding. The on-off signatures based on Reed-Muller codes proposed in this work carry the needed structure to resolve the scalability issue. The chirp reconstruction algorithm entails a sub-linear complexity, which is scalable to extremely large NIA space. Moreover, the high SNR required by the modified chirp algorithm is much lower than that required by the group testing algorithm.

Some of the unique contributions of this work are highlighted as follows:

1) Despite the half-duplex nature of wireless nodes, full-duplex neighbor discovery is achieved using on-off signaling;
2) Reed-Muller codes are used to solve the scalability problem of neighbor discovery;
3) This work considers a more realistic model than [14], [15] by including path loss and the near-far effect;
4) In [14], [15], reliable discovery is achieved only for small neighborhoods (fewer than 10 neighbors) and at high SNR, whereas this work allows discovery of 30 neighbors at low SNR.

We also note that the problem of inferring about the inputs to a noisy linear system from the outputs have been studied in many contexts. For example, the group testing algorithm has been used to solve an RFID problem [16]. One important area relevant to the model (1) is multiuser detection. References [17], [18] considers a related user activity detection problem in cellular networks, and suggests the use of coherent multiuser detection techniques, which is not applicable for neighbor discovery because of unknown channel coefficients. Reference [19] takes channel estimation into account, but the algorithm does not scale well with the network size.

III. The Channel and Network Models

A. The Linear Channel

Consider a wireless network where each node is assigned a unique network interface address from the address space $\{0, 1, \dots, N\}$. If the space consists of all IEEE 802.11 MAC addresses then $N = 2^{48} - 1$. The actual number of nodes can be much smaller than N, but since a node may take any address in the space, it is equivalent to assume that there are exactly $N + 1$ nodes. We shall later discuss the problem of having all nodes simultaneously discover their respective neighborhoods, but for now let us assume that node 0 is the only query node. Over a neighbor discovery period of M symbol intervals, node n sends a signal, referred to as its *signature*, which can be expressed as $\boldsymbol{S}_n = [S_{1n}, \dots, S_{Mn}]^\top$. For the time being let us ignore the propagation delay and assume symbol-synchronous transmissions from all nodes. The received signal of the query node can thus be expressed as a column vector of length M:

$$\boldsymbol{Y} = \sqrt{\gamma} \sum_{n \in \boldsymbol{N}_0} U_n \boldsymbol{S}_n + \boldsymbol{W} \tag{1}$$

where γ denotes the average channel gain in SNR, $\boldsymbol{N}_0$ denotes the set of NIAs in the neighborhood of node 0, U_n denotes the complex-valued coefficient of the wireless link from node n to node 0, and $\boldsymbol{W}$ consists of M independent unit circularly symmetric complex Gaussian random variables, with each entry $W_m \sim \mathcal{CN}(0, 1)$. For simplicity, transmissions from non-neighbors, if any, are accounted for as part of the additive Gaussian noise. Such interference is modeled explicitly in a more comprehensive study [20].

For convenience, we introduce binary variables B_n, which is set to 1 if node n is a neighbor of node 0, and set to 0 otherwise. Let $\boldsymbol{X} = [B_1U_1, \ldots, B_NU_N]^\top$ and $\underline{\boldsymbol{S}} = [\boldsymbol{S}_1, \ldots, \boldsymbol{S}_N]$. Then model (1) can be rewritten as

$$\boldsymbol{Y} = \sqrt{\gamma}\, \underline{\boldsymbol{S}} \boldsymbol{X} + \boldsymbol{W} \; . \tag{2}$$

B. Propagation Model and Near-Far Problem

Previous work [15] simply assumes the channel gain U_n to be a Rayleigh fading random variable. In this work, we incorporate the effect of network topology and propagation loss into the channel model (1). Suppose all nodes transmit at the same power. Moreover, large-scale attenuation follows power law with path-loss parameter α, and small-scale attenuation is modeled as independent fading.

For simplicity, we consider a network consisting of N nodes on an area of size A, where each node's location is uniformly distributed in the area and independent of the other nodes. In other words, the N nodes form a homogeneous Poisson point process. Suppose also that N and A are large enough for the boundary effect to be ignored. Consider a uniformly and randomly selected pair of nodes and let r stand for the distance between them. The channel power gain between them is $Gr^{-\alpha}$, where G denotes the small-scale attenuation. The nodes are called neighbor of each other if the channel gain between them exceeds a certain threshold, i.e., $Gr^{-\alpha} > \eta$ for some fixed threshold η. The average number of neighbors is related to the threshold η by $c = N\,\mathsf{P}(Gr^{-\alpha} > \eta)$.

Since nodes are independent, the indicators $B_1, \ldots, B_N$ are identically distributed (i.i.d.) Bernoulli random variables with $\mathsf{P}\{B_1 = 1\} = c/N$. It is fair to assume that the coefficients $\{U_n\}_{n \in N_0}$ of all nodes are i.i.d. complex-valued circularly symmetric random variables. For simplicity, we further assume that the distribution of the amplitude $|U_n|$ is identical to the distribution of $\sqrt{Gr^{-\alpha}}$ conditioned on that it exceeds $\sqrt{\eta}$. It is straightforward to show that $\mathsf{P}(Gr^{-\alpha} > u^2 \big| Gr^{-\alpha} > \eta) = u^{-\frac{4}{\alpha}} \big/ \eta^{-\frac{2}{\alpha}}$ for every $u \geq \sqrt{\eta}$. Thus, the probability density function (pdf) of $|U_n|$ is

$$p(u) = \begin{cases} \frac{4}{\alpha} \frac{u^{-4/\alpha - 1}}{\eta^{-2/\alpha}}, & u \geq \sqrt{\eta}; \\ 0, & \text{otherwise.} \end{cases} \tag{3}$$

Interestingly, the distribution does not depend on the fading statistics.

The near-far situation, namely that some neighbors can be much stronger than others is inherently modeled in (1)–(3). As we shall see, RM codes with erasure and the corresponding chirp reconstruction algorithm are highly resilient to the near-far problem. One of the reasons is that the gain of strong neighbors can be estimated quite accurately so that their interference to weaker neighbors can be removed.

C. The Neighbor Discovery Problem

The model (1) represents a familiar noisy linear measurement system. We shall refer to $\boldsymbol{Y} = [Y_1, \ldots, Y_M]^\top$ as the measurements. Given $\boldsymbol{Y}$, the SNR γ, and knowledge of the measurement matrix $\underline{\boldsymbol{S}} = [\boldsymbol{S}_1, \ldots, \boldsymbol{S}_N]$, the goal of neighbor discovery is to determine which entries of $\boldsymbol{X}$ are nonzero, i.e., to recover the support of $\boldsymbol{X}$. It is fair to assume that the signatures $\{\boldsymbol{S}_n\}$ are known input to the query node. This is true even in the case of random-access neighbor discovery, where $\boldsymbol{S}_n$ consists of repetitions of an error-control-coded NIA interleaved with delays, and synchronization flags and/or training symbols are embedded so that the delays can be measured accurately in absence of collision.

Two types of errors are possible: If an actual neighbor is eliminated by the algorithm, it is called a *miss*. On the other hand, if a non-neighbor survives the algorithm and is thus declared a neighbor, it is called a *false alarm*. The *rate of miss* (resp. *rate of false alarm*) is defined as the average number of misses (resp. false alarms) in one node's neighborhood divided by the average number of neighbors the node has.

D. Propagation Delay and Synchronicity

The assumption that all neighbors' signatures are received synchronously at a query node is reasonable. First, the duration of an on/off-slot should be chosen to be much larger than the maximum propagation delay between neighbors. For example, in a network where the communication range is 300 meters so that the propagation delay is within 1 microsecond, the slot interval can be 10 microseconds by design. More pronounced propagation delays can also be explicitly addressed in the physical model, but this is out of the scope of this paper. Secondly, nodes shall be synchronized and transmit their signatures simultaneously. This can be achieved using a common clock provided by the global positioning system (GPS) or by using a distributed algorithm for reaching average consensus (see e.g., [21]). At any rate, there is no fundamental disadvantage by using on-off signaling, because any alternate scheme has to deal with the same timing uncertainty.

E. Full-Duplex Network-wide Discovery

Unlike in [14], [15], this work also considers the problem that many or all nodes need to discover their respective neighborhoods at the same time. It is important to note that the physical *half-duplex* constraint prevents a wireless node receiving any useful signal at the same time and over the same frequency band on which it is transmitting. A random-access scheme naturally supports network-wide discovery. This is because each node transmits its NIA intermittently, so that it can listen to the channel to collect neighbors' NIAs during its own epochs of non-transmission. Collision is inevitable, but if each node repeats its NIA for sufficient number of times with enough (random) spacing, then with high probability it can be received by every neighbor once without collision.

Random access discovery tends to waste a substantial amount of channel resources due to retransmissions. Using on-off signatures enables network-wide discovery more efficiently because a node can receive a small amount of useful information during every one of its own off-slots. The impact of the half-duplex constraint is in effect a reduction of the number of useful measurements. From the viewpoint of any query node, once the erasures are purged, model (1) still applies. To trigger network-wide discovery, nodes can be programmed to simultaneously transmit their signatures at regular, pre-determined time instants. In this way neighbor discovery does not interfere with data transmission. Although mobility is not explicitly treated in the paper, sufficiently frequent execution of the discovery scheme can track the neighborhood of a node accurately.

IV. Reed-Muller Signatures and Chirp Reconstruction

Previous works [14], [15] proposed random on-off signatures, where the measurement matrix $\underline{\boldsymbol{S}}$ consists of i.i.d. Bernoulli entries. In absence of noise and mutual cancellation, the measurement $Y_m = \sqrt{\gamma}\sum_{n=1}^{N} S_{mn}X_n$ is nonzero if any node from the group $\{n : n = 1, \dots, N, \text{ and } S_{mn} \neq 0\}$ is a neighbor. The group testing algorithm in [15] visits every measurement Y_m with its power below a threshold T to eliminate all nodes who would have transmitted energy at time slot m from the neighbor list. Those nodes which survive the elimination process are regarded as neighbors. The group testing algorithm requires only noncoherent energy detection and is remarkably simple. However, the algorithm visits each entry of the measurement matrix once, so that their computational complexity is $\mathcal{O}(MN)$. A general purpose processor may handle up to 16-bit NIAs ($N = 65{,}536$) in real time, where M is a few thousand, but a much larger NIA space becomes infeasible. Moreover, the group testing algorithm achieves reliable discovery only if the SNR is very high. In fact the error performance is unacceptable for many practical scenarios [15].

In this work, we propose to use on-off signatures based on (deterministic) Reed-Muller codes. The advantage over the random signatures of [15] is that the corresponding chirp reconstruction algorithm has sub-linear complexity, which allows practical, efficient discovery even if the NIA space is 2^{48} or larger. The key to this scheme is to first construct a deterministic measurement matrix based on a second order RM code. These signatures, however, consist of all QPSK (nonzero) entries, which prevents a transmitting node from simultaneously discovering its neighborhood. In order to enable network-wide neighbor discovery, we propose to introduce zero entries by erasing about 50% of the symbols in each signature.

We first discuss the original RM code without erasure. Such a code is sufficient for a single silent query node to acquire its neighborhood. The construction of the signatures is described in detail in [22]. We only provide a sketch of the construction as follows. Description of the signature generation and chirp reconstruction schemes shall be given shortly as Algorithms 1 and 2, respectively, which allows someone without familiarity with RM code to implement the schemes.

Given a positive integer m, construct the Kerdock set $\mathcal{K}(m)$ consisting of 2^m binary symmetric $m \times m$ matrices. Roughly speaking, they are binary Hankel matrices where the top row consists of arbitrary entries and each of the remaining reverse diagonals is computed from a fixed linear combination of the entries in the top row. It is known that $\mathcal{K}(m)$ is an m-dimensional vector space with basis $\{\boldsymbol{P}(e_m^i)\}_{i=1}^m$, where $\boldsymbol{P}(e_m^i)$ corresponds to the $m \times m$ matrix in which all but the i-th entry in the top row are 0 and the i-th entry is set to 1. For any $l \leq m$, matrix $\boldsymbol{P} \in \mathcal{K}(l)$ can be padded to an $m \times m$ symmetric matrix represented by $\mathcal{P}(\boldsymbol{P})$, where the lower left $l \times l$ submatrix is $\boldsymbol{P}$ and the other entries are all zeros if any. Putting together all padded bases $\{\mathcal{P}(\boldsymbol{P}(e_m^i))\}_{i=1}^m$, $\{\mathcal{P}(\boldsymbol{P}(e_{m-1}^i))\}_{i=1}^{m-1}$, …, $\{\mathcal{P}(\boldsymbol{P}(e_1^1))\}$ forms a new set $\boldsymbol{B}$, which can be regarded as basis of the $m(m+1)/2$-dimensional vector space formed by all $m \times m$ symmetric matrices. Given $n \leq m(m+1)/2$, any $\boldsymbol{c} = (c_1, \dots, c_n)^{\mathsf{T}} \in \mathbb{Z}_2^n$ corresponds to an $m \times m$ symmetric matrix represented by $\boldsymbol{P}(\boldsymbol{c}) = \sum_{i=1}^n c_i \boldsymbol{B}(i) \mod 2$, where $\boldsymbol{B}(i)$ is the i-th element in the set $\boldsymbol{B}$.

Each node can then map its NIA into a signature as follows. Let the NIA consist of $n_1 + n_2$ bits, which is divided into two binary vectors: $\boldsymbol{b}' \in \mathbb{Z}_2^{n_1}$ and $\boldsymbol{c} \in \mathbb{Z}_2^{n_2}$, with $n_1 \leq m$ and $n_2 \leq m(m+1)/2$. Let $\boldsymbol{b} \in \mathbb{Z}_2^m$ be formed by appending $m - n_1$ zeros after $\boldsymbol{b}'$. The corresponding signature is of 2^m symbols, whose entry indexed by each $\boldsymbol{a} \in \mathbb{Z}_2^m$ is given by

$$\phi_{\boldsymbol{P}(\boldsymbol{c}),\boldsymbol{b}}(\boldsymbol{a}) = \exp\left[j\pi\left(\frac{1}{2}\boldsymbol{a}^{\mathsf{T}}\boldsymbol{P}(\boldsymbol{c})\boldsymbol{a} + \boldsymbol{b}^{\mathsf{T}}\boldsymbol{a}\right)\right] . \tag{4}$$

The system can thus accommodate up to $2^{m(m+3)/2}$ nodes with distinct signatures, each of length 2^m. For example, if $m = 10$, we have up to 2^{65} signatures of length $2^{10} = 1{,}024$.

The chirp reconstruction algorithm developed in [8] can be applied for the query node to recover its neighborhood based on the received signal. The general idea of the iterative algorithm is as follows: 1) initialize the residual signal to $\boldsymbol{Y}$, 2) take the Hadamard transform of the auto-correlation function of the residual signal to expose the coefficient of the digital chirps, 3) sequentially extract the Kerdock matrix $\boldsymbol{P}(\boldsymbol{c}_k)$ and $\boldsymbol{b}_k$ corresponding to the largest energy component in the residual signal and determine the corresponding X_k, and 4) update the residual signal and repeat steps 2 to 4 until the residual falls below a threshold. The query node has thus acquired the signatures of all peers with high probability, and can then recover the NIAs if needed.

The drawback of using the original RM code is that the signatures defined by (4) consist of QPSK symbols, so that a node cannot simultaneously receive useful signal while transmitting its own signature. In order to achieve network-wide neighbor discovery, we apply 50% erasures to each of the original RM codewords to obtain an on-off signature, so that nodes can listen during their own off-slots. The signature of each node consists of as many off-slots as on-slots, thus two nodes can receive pulses from each other over about 25% of the slots.

Algorithm 1 Signature Generation Algorithm

Input: n-bit NIA [As a trivial example just for illustration, let $n = 2^4 = 16$.]

Choose m such that $n = n_1 + n_2$ with $n_1 \leq m$ and $n_2 \leq \frac{m_0}{2}(2m - m_0 + 1)$ where $m_0 \leq m/2$. [In the example, $n_1 = 2, n_2 = 2, m = 2, m_0 = 1$.]

Divide n-bit NIA into two vectors $\boldsymbol{b}' \in \mathbb{Z}_2^{n_1}$ and $\boldsymbol{c} \in \mathbb{Z}_2^{n_2}$. Let $\boldsymbol{b} \in \mathbb{Z}_2^m$ be formed by appending $m - n_1$ zeros after $\boldsymbol{b}'$.

Generate the original RM code $\phi_{\boldsymbol{P}(\boldsymbol{c}),\boldsymbol{b}}$ of length 2^m according to (4). [In the example, there are 16 codewords listed in the following table.]

$\boldsymbol{b}$	$\boldsymbol{c}$	$\phi_{\boldsymbol{P}(\boldsymbol{c}),\boldsymbol{b}}$	$\boldsymbol{b}$	$\boldsymbol{c}$	$\phi_{\boldsymbol{P}(\boldsymbol{c}),\boldsymbol{b}}$
00	00	1 1 1 1	01	00	1 -1 1 -1
00	01	1 j 1 -j	01	01	1 -j 1 j
00	10	1 j j -1	01	10	1 -j j 1
00	11	1 1 1 -1	01	11	1 -1 j j
10	00	1 1 -1 -1	11	00	1 -1 -1 1
10	01	1 j -1 j	11	01	1 -j -1 -j
10	10	1 j -j 1	11	10	1 -j -j -1
10	11	1 1 -1 1	11	11	1 -1 -j -j

Generate the erasure pattern $\boldsymbol{r}$ of length 2^m as follows: Let the first segment of 2^{m-m_0} bits be i.i.d. Bernoulli random variables with parameter $1/2$ and repeat the segment 2^{m_0} times to form the 2^m bits of $\boldsymbol{r}$.

Output: The on-off signature of length 2^m is the pointwise product of $\phi_{\boldsymbol{P}(\boldsymbol{c}),\boldsymbol{b}}$ and $\boldsymbol{r}$.

In the chirp reconstruction algorithm described in [8], in order to recover the i-th row of the Kerdock matrix corresponding to the largest energy component in the residual signal, the auto-correlation is computed between the residual signal of length 2^m and its shift by 2^{m-i}. It is advisable to guarantee that the positions of erasures in the received signal and its shift are perfectly aligned. Therefore, we apply random erasures to the signatures in the following simple manner: Suppose n_2 is chosen such that the Kerdock matrix generated by each node is determined by its first $m_0 \leq m/2$ rows. For node k, the erasure pattern $\boldsymbol{r}_k$ of length 2^m is constructed as follows: Divide $\boldsymbol{r}_k$ into 2^{m_0} segments with equal length 2^{m-m_0}, let the first segment be i.i.d. Bernoulli random variables with parameter $1/2$ and all remaining segments be identical copies of the first segment. It is easy to see that after introducing erasures into the signatures, the network can still accommodate $2^{\frac{1}{8}m(3m+10)}$ nodes. For example, if $m = 12$, we have up to 2^{69} signatures of length $2^{12} = 4,096$. The signature generation for each node in the network is summarized in Algorithm 1. Along with the general algorithm, we provide a specific example in braces to illustrate the construction.

Algorithm 2 Chirp Reconstruction Algorithm

Input: received signal $\boldsymbol{Y}$ in (2), signature matrix of all other nodes $\underline{\boldsymbol{S}}$ and its own erasure pattern $\boldsymbol{r}$.

Choose three parameters: residual power indicator ϵ, maximum number of iterations $T_{\max}$ and the neighbor power indicator η_0.

Initialize the number of iterations t to 0 and the residual signal $\boldsymbol{Y}_r$ to $\tilde{\boldsymbol{Y}}$, which is the pointwise product of $\boldsymbol{Y}$ and $1 - \boldsymbol{r}$.

Initialize the neighbor set $\boldsymbol{N} = \emptyset$ and the coefficient vector $\boldsymbol{C} = \emptyset$.

Main iterations:

while $\|\boldsymbol{Y}_r\|_2 \geq \epsilon$ and $t \leq T_{\max}$ **do**

  for $i = 1, 2, \ldots, m_0$ **do**

    Compute the pointwise multiplication of $\boldsymbol{Y}_r$ with its shift in the amount of 2^{m-i}.

    Compute the fast Walsh-Hadamard transform of the computed auto-correlation.

    Find the position of the highest peak in the frequency domain. Based on the peak location, decode the i-th row of an $m \times m$ matrix $\boldsymbol{P}(\boldsymbol{c}_k)$, which corresponds to a certain node k.

  end for

  Use the first m_0 rows of $\boldsymbol{P}(\boldsymbol{c}_k)$ from the preceding iterations to determine the remaining rows of $\boldsymbol{P}(\boldsymbol{c}_k)$.

  Compute $\boldsymbol{S}_k^0(\boldsymbol{a}) = \exp\left[j\pi\left(\frac{1}{2}\boldsymbol{a}^\top \boldsymbol{P}(\boldsymbol{c}_k)\boldsymbol{a}\right)\right]$ for all $\boldsymbol{a} \in \mathbb{Z}_2^m$ and apply Hadamard transform to the pointwise product of $\boldsymbol{Y}_r$ and $\overline{\boldsymbol{S}_k^0}$;

  Recover $\boldsymbol{b}_k$ by finding the highest peak in the frequency domain.

  Compute $\phi_{\boldsymbol{P}(\boldsymbol{c}_k),\boldsymbol{b}_k}$ according to (4) and recover the on-off signature $\boldsymbol{S}_k$ by pointwise product of $\phi_{\boldsymbol{P}(\boldsymbol{c}_k),\boldsymbol{b}_k}$ and $\boldsymbol{r}_k$.

  Add node k to the neighbor set $\boldsymbol{N}$ and add a corresponding 0 to the coefficient vector $\boldsymbol{C}$.

  Put together all signatures of nodes in $\boldsymbol{N}$ to form a matrix $\boldsymbol{S}$. Construct $\tilde{\boldsymbol{S}}$ by pointwise multiplying each column in $\boldsymbol{S}$ with $1 - \boldsymbol{r}$.

  Determine the value of vector $\boldsymbol{X}$ which minimizes $\|\boldsymbol{Y}_r - \tilde{\boldsymbol{S}}\boldsymbol{X}\|_2$. Update the coefficient vector $\boldsymbol{C}$ by $\boldsymbol{C} + \boldsymbol{X}$.

  Update the residual signal $\boldsymbol{Y}_r$ by $\boldsymbol{Y}_r - \tilde{\boldsymbol{S}}\boldsymbol{C}$.

end while

Output: Find all elements in $\boldsymbol{N}$ such that the corresponding coefficients in $\boldsymbol{C}$ are no less than η_0.

The chirp reconstruction algorithm with some modifications can then be applied here for any node (say, node 0) to recover its neighborhood based on the observations (denoted as $\tilde{\boldsymbol{Y}}$) through its own off-slots. The details are provided in Algorithm 2. In light of the original chirp algorithm, we emphasize the key steps as follows:

1) Initialize the residual signal $\boldsymbol{Y}_r$ to $\tilde{\boldsymbol{Y}}$;
2) Take the Hadamard transform of the auto-correlation function of $\boldsymbol{Y}_r$ and its shift by $2^{m-i}, i = 1, 2, \ldots, m_0$, so that the first m_0 rows of $\boldsymbol{P}(\boldsymbol{c}_k)$ can be recovered according to the chirps exposed in the frequency domain, then the whole matrix $\boldsymbol{P}(\boldsymbol{c}_k)$ can be determined;
3) Compute $\boldsymbol{S}_k^0(\boldsymbol{a}) = \exp\left[j\pi\left(\frac{1}{2}\boldsymbol{a}^\top \boldsymbol{P}(\boldsymbol{c}_k)\boldsymbol{a}\right)\right]$ for all $\boldsymbol{a} \in \mathbb{Z}_2^m$ and apply Hadamard transform to the pointwise product of $\boldsymbol{Y}_r$ and $\overline{\boldsymbol{S}_k^0}$ to recover $\boldsymbol{b}_k$;
4) Recover the erased signature $\boldsymbol{S}_k$ by pointwise product of $\phi_{\boldsymbol{P}(\boldsymbol{c}_k),\boldsymbol{b}_k}$ and $\boldsymbol{r}_k$, then put together all signatures already recovered to form a matrix $\boldsymbol{S}$. Determine the value of $\boldsymbol{X}$ which minimizes $\|\boldsymbol{Y}_r - \tilde{\boldsymbol{S}}\boldsymbol{X}\|_2$, where $\tilde{\boldsymbol{S}}$ is constructed by setting all rows of $\boldsymbol{S}$ but those which correspond to the off-slots of node 0 to zero, and then update all corresponding coefficients;
5) Update the residual signal and repeat steps 1 to 4 until the residual falls below a threshold or the total number of iterations exceeds a threshold.

Algorithm 2 provides enough details about the scheme and is ready for implementation.

V. A Numerical Example

We illustrate the performance using the following example. Suppose there are 2^{20} valid NIAs in the network. Without loss of generality, we consider a typical node located at the origin and let all other nodes be distributed as a Poisson point process on a square centered at the origin. Let the path loss be $\alpha = 3$ and the neighbor cut-off power be $\eta = 0.05$. First let the density of the nodes be such that node 0 has on average $c = 10$ neighbors. Choose $m = n_1 = n_2 = 10$, then the signature length is $2^m = 1,024$. During all 100 trials at different SNRs, there are no false alarms registered. The rate of miss is plotted in Fig. 1 against the SNR. We find that the total error rate can be lower than 0.2% at SNR=11.5 dB. In contrast, if random on-off 1,000-bit signatures of [15] are used instead, at least 23 dB SNR is needed to achieve the same error rate, even if the size of the address space is only 10,000.

In Fig. 1, we also repeat the simulation with the number of average neighbors changed to $c = 30$ and the parameters m, n_1, n_2 changed to $m = n_2 = 12, n_1 = 8$. In this case, the signature length is $2^m = 4,096$. During all 50 trials, there are still no false alarms and the total error rate can be lower than 0.1% at SNR=16 dB, which is quite desirable in practice.

VI. Comparison with Random Access

We compare the performance of the compressed neighbor discovery schemes described in Section IV with that of traditional random-access discovery schemes. It is important to note that, only one frame is needed by compressed neighbor discovery, as opposed to many frames in the case of random access, which offers significant additional reduction of synchronization and error-control overhead embedded in every frame.

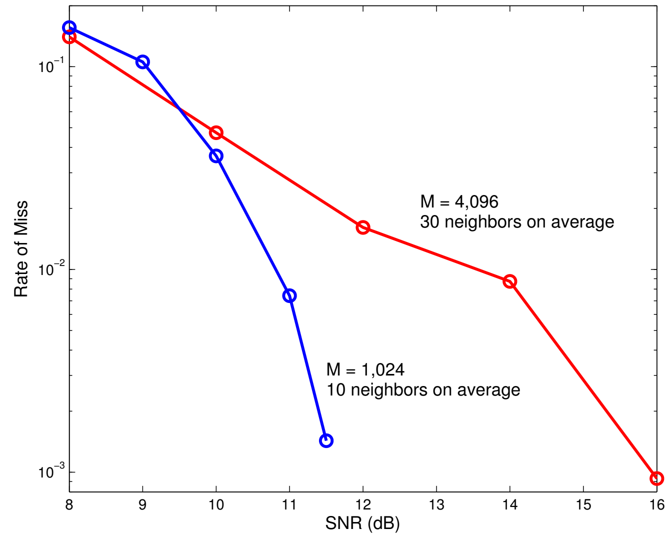

Fig. 1. Rate of miss versus SNR

A. Comparison with Generic Random-Access Discovery

Suppose a random-access discovery scheme is used. Nodes contend to announce their NIAs over a sequence of t contention periods. In each period, each neighbor independently chooses to either transmit (with probability θ) or listen (with probability $1 - \theta$). Let $\rho = c/N$. The error rate can be lower bounded by the probability of one given neighbor being missed, which is given by

$$\sum_{z=1}^{N} \binom{N}{z} \rho^z (1-\rho)^{N-z} \left[1 - \theta(1-\theta)^{z-1}\right]^t . \qquad (5)$$

Consider a network with 2^{20} nodes and on average 10 nodes around node 0. We assume time is slotted and transmission of each bit takes one slot. Suppose 1,024-bit signatures are used and the SNR is 11.5 dB. The error rate achieved by using on-off RM signatures is no greater than 0.002 (see Fig. 1(a)). If random-access discovery is carried out, the smallest number of contention periods that brings the error rate below 0.002 is 194, which is obtained when θ is set to 0.072. In each contention period, the number of bits transmitted is at least 20 just to carry the NIA. With binary signaling, the total overhead is at least $194 \times 20 = 3880$ slots, which is 3.8 times that needed by compressed neighbor discovery with on-off signaling.

In fact, the efficiency of compressed neighbor discovery can be an order of magnitude higher than that of random access if all overhead is accounted for. This is because that transmitting a 20-bit NIA reliably over a fading channel generally requires a few hundred symbols with all overhead included.

B. Comparison with IEEE 802.11g

It is also instructive to compare compressed neighbor discovery with the popular IEEE 802.11g technology. Consider the ad hoc mode of 802.11g with active scan, which is basically a random-access discovery scheme. The signaling rate is 4 μs per orthogonal frequency division multiplexing (OFDM) symbol. One probe response frame takes about 850 μs. (The response frame includes additional bits but is dominated by the NIA.) Thus it takes at least 850 $\mu s \times 194 \approx 165\ ms$ for a query node to discovery 10 neighbors with error rate 0.002 or lower. If compressed neighbor discovery with on-off RM signature is used, 1,000 symbol transmissions suffice to achieve the same error rate. Using 802.11g symbol interval (4 μs), reliable discovery takes merely 4 ms. A highly conservative choice of the symbol interval is 30 μs, which includes carrier (on-off) ramp period (say 10 μs) and the propagation time (less than 1 μs for 802.11 range). Compressed neighbor discovery then takes a total of 30 $\mu s \times 1000 \approx 30\ ms$, less than $1/5$ of that required by 802.11g.

VII. Concluding Remarks

The on-off signaling of the neighbor discovery scheme in this paper was first proposed in [23] and referred to as *rapid on-off-division duplex (RODD)*. Such signaling departs from the collision model and exploits the multiaccess nature of the wireless channel. Moreover, it is recognized that a half-duplex radio need not switch between transmitting and receiving modes at the frame level. Instead, we interleave the transmitted symbols with off symbol slots (in similar fashion as on-off keying (OOK)). The received signal consists of useful signals acquired during one's own off-slots as well as self-interference during the on-slots. To address the concern that switching the transmitting or receiving circuits on and off is difficult and may consume much energy, we note that no switching is needed; rather, both circuits operate continuously and the decoding algorithm simply discards the received signals over one's own on-slots, or equivalently, regards them as erasures.

In theory, on-off signaling achieves higher throughput than random access schemes [23]. More recently, the idea of using RODD signaling and sparse recovery technique is extended to carry out full-duplex data exchanges between neighboring wireless nodes [24].

Acknowledgement

We thank Robert Calderbank and Sina Jafarpour for sharing their MATLAB code of the chirp reconstruction algorithm. We also thank Dr. Jun Luo for assistance in carrying out the comparison in Section VI.

References

[1] *RFC 3684: Topology Dissemination Based on Reverse-Path Forwarding (TBRPF)*, MANET Working Group Std., mANET Working Group, The Internet Engineering Task Force (IETF).

[2] 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*. Long Beach, CA, USA, Oct. 2001, pp. 137–145.

[3] S. Vasudevan, J. Kurose, and D. Towsley, "On neighbor discovery in wireless networks with directional antennas," in *Proc. IEEE INFOCOM*, vol. 4, 2005, pp. 2502–2512.

[4] E. Felemban, R. Murawski, E. Ekici, S. Park, K. Lee, J. Park, and Z. Hameed, "SAND: Sectored-antenna neighbor discovery protocol for wireless networks," in *Proc. IEEE Conf. Sensor Mesh and Ad Hoc Communications and Networks*, 2010, pp. 1–9.

[5] X. An, R. V. Prasad, and I. Niemegeers, "Neighbor discovery in 60 GHz wireless personal area networks," in *Proc. IEEE Int'l Symp. World of Wireless Mobile and Multimedia Networks*, 2010, pp. 1–8.

[6] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inform. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.

[7] E. J. Candes and T. Tao, "Near-optimal signal recovery from random projections: Universal encoding strategies?" *IEEE Trans. Inform. Theory*, vol. 52, no. 12, pp. 5406–5425, Dec. 2006.

[8] S. D. Howard, A. R. Calderbank, and S. J. Searle, "A fast reconstruction algorithm for deterministic compressive sensing using second order Reed-Muller codes," in *Proc. Conf. Inform. Sciences & Systems*, Mar. 2008, pp. 11–15.

[9] S. A. Borbash, A. Ephremides, and M. J. McGlynn, "An asynchronous neighbor discovery algorithm for wireless sensor networks," *Ad Hoc Networks*, vol. 5, pp. 998–1016, Sep. 2007.

[10] S. Vasudevan, D. Towsley, D. Goeckel, and R. Khalili, "Neighbor discovery in wireless networks and the coupon collector's problem," in *Proc. 15th Annual Int'l Conf. Mobile Comput, Network.*. Beijing, China, 2009, pp. 181–192.

[11] R. Khalili, D. Goeckel, D. Towsley, and A. Swami, "Neighbor discovery with reception status feedback to transmitters," in *Proc. IEEE INFOCOM*, Mar. 2010, pp. 1–9.

[12] M. Corson, R. Laroia, J. Li, V. Park, T. Richardson, and G. Tsirtsis, "Toward proximity-aware internetworking," *IEEE Wireless Commun.*, vol. 17, no. 6, pp. 26–33, Dec. 2010.

[13] J. Ni, R. Srikant, and X. Wu, "Coloring spatial point processes with applications to peer discovery in large wireless networks," *IEEE/ACM Trans. Networking*, vol. PP, no. 99, p. 1, 2010.

[14] J. Luo and D. Guo, "Neighbor discovery in wireless ad hoc networks based on group testing," in *Proc. Allerton Conf. Commun., Control, and Computing*. Monticello, IL, USA, 2008.

[15] ——, "Compressed neighbor discovery for wireless ad hoc networks: the Rayleigh fading case," in *Proc. Allerton Conf. Commun., Control, and Computing*. Monticello, IL, USA, Oct. 2009.

[16] M. Kodialam, W. C. Lau, and T. Nandagopal, "Identifying RFID tag categories in linear time," in *WiOpt*, 2009.

[17] D. D. Lin and T. J. Lim, "Subspace-based active user identification for a collision-free slotted ad hoc network," *IEEE Trans. Commun.*, vol. 52, pp. 612–621, 2004.

[18] D. Angelosante, E. Biglieri, and M. Lops, "A simple algorithm for neighbor discovery in wireless networks," in *Proc. IEEE Int'l Conf. Acoustics, Speech and Signal Processing*, vol. 3, Apr. 2007, pp. 169–172.

[19] ——, "Neighbor discovery in wireless networks: A multiuser-detection approach," *Physical Communication*, vol. 3, no. 1, pp. 28–36, 2010.

[20] L. Zhang, J. Luo, and D. Guo, "Compressed neighbor discovery for wireless networks," *preprint*.

[21] O. Simeone, U. Spagnolini, Y. Bar-Ness, and S. Strogatz, "Distributed synchronization in wireless networks," *IEEE Signal Processing Mag.*, vol. 25, no. 5, pp. 81–97, Sep 2008.

[22] A. R. Calderbank, A. C. Gilbert, and M. J. Strauss, "List decoding of noisy reed-muller-like codes," *CoRR*, vol. abs/cs/0607098, 2006.

[23] D. Guo and L. Zhang, "Rapid on-off-division duplex for mobile ad hoc neworks," in *Proc. Allerton Conf. Commun., Control, and Computing*, Monticello, IL, USA, Sep. 2010.

[24] L. Zhang and D. Guo, "Wireless peer-to-peer mutual broadcast via sparse recovery," *http://arxiv.org/abs/1101.0294*.