This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE Secon 2009 proceedings.

Featherlight Information Network with Delay-Endurable RFID Support (FINDERS)

Zhipeng Yang and Hongyi Wu
The Center for Advanced Computer Studies
University of Louisiana at Lafayette
Email:{zxy1767,wu}@cacs.louisiana.edu

***Abstract*— This research centers on the Featherlight Information Network with Delay-Endurable RFID Support (FINDERS), composed of passive RFID tags which are ultra light, durable, and flexible, without power supply for long-lasting applications under strict weight constraints and harsh environments. It expands the use of RFID gear for wireless network construction, aiming to find events of interest and gather aggregate information. FINDERS faces unprecedented challenges in communication and networking, due to its sporadic wireless links, unique asymmetric communication paradigm, intermittent computation capability, and extremely small memory of tags. Analytic and simulation data are collected to show the feasibility and efficiency of FINDERS.**

I. Introduction and Motivation

Radio frequency identification (RFID) has received growing attention lately, as a result of reduced tag costs and sizes as well as expanded reader communication ranges. This research deals with a *Featherlight Information Network with Delay-Endurable RFID Support (FINDERS)*, composed of passive RFID tags which are ultra light, durable, and flexible, without power supply for long-lasting applications. It investigates into the use of RFID gear for wireless network construction, aiming to find events of interest and gather aggregate information. Like modern smart sensors [1], RFID tags and readers together may constitute a distributed sensor system for pervasive information gathering, subject to different data acquisition, delivery, and storage approaches. Unlike typical active sensors that rely on battery power supply and indispensably require sturdy casing for their protection, however, a passive RFID tag is particularly suitable for environments with strict weight constraints, e.g., in the studies of small wildlife animals, where the weight of the sensor must be under 5% of the weight of the animal to avoid hindering the wildlife movement or the welfare of the animals themselves. This requires very small sensors to be employed, which is in sharp contrast to other projects that target at large animals including zebras [2], whales [3], and deers [4], and human beings [5]. Earlier investigation has revealed that most of these small animals cannot carry any active devices (such as GPS receivers or the smallest Crossbow sensors) with battery and casing, since overweight sensors often lead to high mortality rates of the animals being studied. Similar weight constraint also applies in many other fields of scientific studies and industrial applications. While efforts have been made to develop miniature sensors in laboratories, the lowest weight of any active sensor that aims to achieve a given communication range and lifetime is bounded inevitably by its battery and casing, where the latter must be heavy-duty for the protection of power source and powered electronic circuits under harsh environments. This becomes a key hurdle that limits the applicability of active sensors in various applications with strict weight constraints, not to mention extra hassle involved in ensuring adequate battery power.

This work is supported in part by the National Science Foundation under grant CNS-0831823.

FINDERS consists of two types of nodes, readers and tags, as illustrated in Fig. 1. The readers are deployed at strategic locations according to specific applications. For example, in wildlife and biological studies, readers can be set up at natural "choke points" where the animals have to move past or through, because of significant movement barriers otherwise. Since FINDERS is often deployed under harsh and complex field environments, it is not practically viable for most readers to establish reliable connections to communicate with each other or to access the backbone networks. As a result, they become isolated readers, or IR's (see IR's 1-4 in Fig. 1). Power supply is challenging for the IR's too. Most of them rely on solar panels and high capacity batteries, which can be recharged/replaced but will post a limit on the daily duty circle of the readers. Only a few readers at convenient locations are equipped with reliable power source and network connections, and they are dubbed gateway readers (GR's; see GR's 1 and 2 in Fig. 1). GR's serve as the gateways to deliver data from FINDERS to the destination (e.g., a data server). The readers have relatively large storage space compared with tags, and each of them maintains a data queue. While all readers are fixed, the tags are attached to moving targets and thus become mobile (see Tags 1-8 in Fig. 1). Many off-the-shelf passive tags that are light and durable and have sufficient reading/writing ranges can be employed in FINDERS. For example, we have adopted the Alien passive RFID system for this research, which consists of ALR-9900 reader and ALN-9540 Squiggle^{TM} (passive) tags. The ALR-9900 reader is a powerful device that possesses 64 MB RAM and 64 MB flash memory and supports up to 4 antennae and 50 communication channels. Being very thin and light, an ALN-9540 tag measures $8.15 \times 94.8 \times 0.05$mm and weights less than one gram. Its memory can hold 14 bytes of user data. With such small storage space, the tag maintains a simple buffer that can be read and written in one operation, without queuing.

978-1-4244-2908-0/09/$25.00 ©2009 IEEE

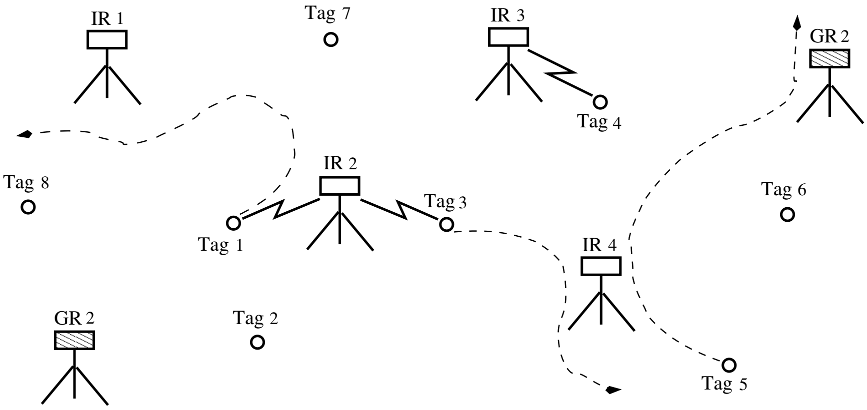


Fig. 1. An overview of FINDERS.

The tag is well engineered and packaged, exhibiting excellent survivability under such harsh environments as underwater, underground tunnels, and extreme temperatures ranging from -25^{o}C to +50^{o}C. Our lab and field experiments have revealed that the Alien system can achieve a reading/writing distance of some 20 ft. Since the readers are static and most of them are isolated from each other, the communications in FINDERS rely on the mobility of the tags to establish time-varying opportunistic links with nearby readers, forming a delay-tolerant network (DTN [6]) for data delivery. For example, in the scenario shown in Fig. 1, Tags 1 and 3 are within the reading/writing range of IR 2. Thus IR 2 may read the data from Tag 1 and write them into Tag 3. When Tag 3 passes by IR 4, it unloads its data to the latter, which in turn writes the data into Tag 5 when it comes into the writing range. Tag 5's trajectory passes through GR 2. The data can thus be delivered to the destination via this gateway.

In general, FINDERS may serve to find events of interest and to gather aggregate system information through sporadic wireless communications between its constituent passive tags and readers. FINDERS is a very unique wireless information network that distinguishes itself from conventional communication networks, sensor networks, and delay-tolerant networks by the following characteristics:

(1) Nodal heterogeneity: FINDERS consists of two very different types of nodes. A reader is a static and powerful device, with large storage, high computing power, and long-lasting (but still limited) battery power. On the other hand, the passive tag can be mobile and has extremely limited resource.

(2) Asymmetric communication: The communication in FINDERS can be established between a tag and a reader only, but not tags to tags or readers to readers. The communication must be initiated by a reader, in contrast to symmetric transmissions in most conventional networks.

(3) Intermittent connectivity: The connectivity of FINDERS is very low and intermittent, forming a sparse network where a tag is connected to a reader only occasionally. In fact, FINDERS is a special DTN with unique communication and storage constraints.

(4) Intermittent computation: Besides connectivity, the computation at the tag is also intermittent. It is available only when the tag is powered up by a nearby reader. Thus, such continuous functions necessary to many protocols as counters and timers cannot be implemented here.

(5) Critical network resource: The fundamental philosophy in DTN is to trade storage for sporadic connectivity. The DTN node is assumed to have sufficient storage space, which can hold a large volume of data to alleviate the needs of immediate transmission. This, however, is no longer valid in FINDERS, because the buffer space of the tags (the main vehicle for data transportation) is so limited that it may become the critical network resource and communication bottleneck.

(6) Delay tolerability: Data delivery delay in FINDERS is potentially high, due to loose connectivity and extremely limited network resources. However, such delay, though not desirable, is usually tolerable by the applications which aim at pervasive information gathering from a statistical perspective.

(7) Fault tolerability: Redundancy may exist in FINDERS during data acquisition and delivery. Thus, a data packet may be lost without degrading information gathering performance. At the same time, there are critical needs in data filtering and aggregation at IR's and GR's, due to the likely existence of multiple copies of the same data.

The above characteristics make the development of FINDERS a very unique, interesting, and challenging problem, calling for effective solutions to overhaul the data acquisition and delivery schemes in such a featherlight information network with extremely limited resources. While various sensor networks [1]–[5] have been investigated and deployed, data collection, storage and communication solutions devised therein are not directly applicable to FINDERS. Little earlier studies have ever been conducted on information networks composed of RFID gears with sporadic communication between moving tags and nearby readers. In particular, FINDERS faces the following unprecedented challenges. First, the sporadically available wireless links render it impossible to form a well connected mesh network for end-to-end communications, which are the basis of mainstream sensor network technologies [1]. Moreover, the unique asymmetric communication paradigm, the intermittent computation capability and the extremely small buffer size of a passive tag overthrow the fundamental principle of DTN (where ample buffering is employed under the intermittent connectivity to alleviate the needs of immediate transmission), resulting in inefficiency if the existing DTN protocols are employed in FINDERS. To this end, a series of interesting problems related to data delivery are investigated in this research:

(1) Routing metrics: Routing metrics are usually adopted by routing protocols as an indicator of the available resource of a given path or link. For example, hop count and delay are popularly used in conventional networks. These metrics, however, do not reflect the unique network resource of FINDERS, and thus may lead to poor network performance or even failure if used for routing. In this research, we investigate effective routing metrics based on meeting probabilities between tags and readers, and devise mechanism that consumes insignificant storage space for maintaining and updating such metrics.

(2) Duplication control: In FINDERS and many other DTN networks, replication is necessary during data delivery for achieving a given success ratio. However, replication increases overhead, and worse yet, excessive replication may even degrade the delivery ratio due to frequent buffer overflow. Duplication control thus becomes a key issue to be tackled in this research. It arrives at a system-wide optimization problem. While the data delivery schemes in FINDERS are (and have to be) much simpler compared with the protocols for other wireless or wired networks, such optimization is non-trivial and its performance largely determines the effectiveness and feasibility of FINDERS.

(3) Queue management: Each reader maintains a data queue. Given the limited communication resource (due to sporadic connectivity and extremely limited storage of tags) and the relatively high loading factor (due to newly acquired and duplicated data), it is an overriding design issue in FINDERS to differentiate the packets in the queue by a simple and efficient parameter which signifies their importance and determines which packet to transmit if a communication opportunity becomes available or which packet to drop if the queue is full. In this research, we propose the prioritization parameter based on the probability that at least one copy of given information can be delivered to the GR's. Note that the goal of such queue management is to facilitate efficient communication in the network with redundancy, in contrast to the QoS-aware algorithms which aim to prioritize raw information.

(4) Protocol design: Routing metrics, data duplication, and queue management are not independent. For instance, data duplication will clearly affect queue management, and at the same time, itself is also influenced by routing. Built on these three basic components, we carefully devise the data delivery protocol by synergizing the interaction among them, in order to achieve efficient network resource utilization. Furthermore, several key constraints and enhancements must be investigated and explored to optimize network performance. In particular, we study issues related to power efficiency, synchronization, and various reading/writing strategies.

The rest of this paper is organized as follows. Sec. II discusses the state-of-the-art technologies related to FINDERS. Sec. III presents the analysis-based general feasibility study. Sec. IV introduces the proposed communication protocol. Sec. V presents our simulations and results. Finally, Sec. VI concludes the paper.

II. Related Work

This research is related to two emerging technologies: RFID and DTN. RFID has gained increasingly wide-spread adoption recently, with several interesting research problems been explored during the past few years. Tag arbitration has been extensively studied, yielding solutions fall into two categories, slotted ALOHA [7]–[10] and tree splitting algorithms [10]–[13]. The problem of counting RFID tags by using statistic approaches was studied in [14], [15]. Meanwhile, [16] and [17] intend to discover missing tags and tag popularity, respectively. A load balancing scheme is introduced in [18] to distribute the communication load evenly among readers. In addition, several algorithms [19]–[22] have been developed to locate the tags, given that object tracking and localization is one of the most important applications of the RFID system. Despite those earlier studies on RFID, however, as far as we know, this is the first work that investigates the RFID-based delay-tolerant mobile sensor networks.

DTN is a sporadically connected network with frequent partitions [6], [23]. Originally developed for deep space communication in high-delay environments, DTN technology has been recently introduced into wireless sensor networks [2], [3], [5], [24]–[32] and mobile ad hoc networks [33]–[42]. Almost all of these studies assume sufficient storage space at the DTN nodes, which can hold a large volume of data to alleviate the needs of immediate transmission. But this fundamental philosophy of paying storage for sporadic connectivity is no longer valid in FINDERS, because the available memory could be extremely limited at the tags. Several DTN-based network architectures most relevant to FINDERS are briefed below. Infostation is proposed as a complement to the cellular network, to provide high speed wireless access in isolated coverage areas [33]. It aims to support information distribution, instead of data acquisition and delivery. Data Mule [29] is proposed to collect data in sparse sensor networks, where a mobile entity (called a data mule) travels in the network, receives data from its nearby sensors, and delivers them to sinks. The data mule has sufficient storage space and battery power, and its mobility is controlled in a way to reach efficient data transportation. Similarly, Message Ferrying [41] intends to exploit non-random nodal mobility to help deliver data in partitioned mobile ad hoc networks. Those earlier DTN systems are in sharp contrast to FINDERS, where a tag moves at will (dictated by the object to which it is tagged) and usually possesses intermittent computing capability and only some tens of bytes of memory for user data storage.

III. General Feasibility Study

While the featherlight weight of passive RFID tags provides unique opportunities to support various applications with strict weight constraints, the capacity of FINDERS is unsurprisingly low compared with many other data networks, due to its extremely limited network resources. To understand its capacity and accordingly the applications that it can support, we first carry out a feasibility study of the FINDERS system before moving into the detailed protocol design and evaluation. This feasibility study aims to gain insights into the basics of FINDERS and unveil its general capacity, without consideration of specific protocols and network environments.

To this end, we need to investigate why a data delivery effort may fail and where the data packet may be dropped. First, since the GR's connect to the conventional networks, they usually have sufficient communication bandwidth (for our target applications) and can safely deliver the received data to end users. The tags keep data packets in their nonvolatile memory. As long as they are not falsely overwritten, the data will not be lost (if we ignore any physical failures). On the other hand, each IR maintains a data queue, which may overflow when packet arrival is too high. Therefore, we focus on the queueing behavior of the IR's in our feasibility study.

(1) Queueing Model. We consider a general queuing model (i.e., a $G/G/1$ queue) for each IR. While it is difficult to derive solutions for the G/G/1 queue without the detailed knowledge of packet arrival and service (which heavily depend on the data transmission protocol), we intend to obtain an estimation of the network capacity (i.e., the amount of data that can be handled by the system) based on the condition of queue stability. More specifically, the arrival rate must be lower than the service rate in order to keep the queue stable. This study only requires to discover the average data service rate, as to be discussed next.

(2) Average Data Service Rate. A data packet is served (i.e., sent out from the IR's queue) when the IR meets a suitable tag. Thus, it is critical to understand how often the meeting between the IR and a suitable tag can happen. Consider a FINDERS with K IR's, J GR's, and T tags, where each tag has a capacity of m, i.e., the tag can store up to m data packets. For simplicity, we assume that a tag can receive data (i.e., becomes a suitable tag) only if it is empty (while more sophisticate schemes for data transmission will be discussed in Sec. IV). Moreover, we consider that Tag i moves by following a randomly chosen path. For analytic tractability, the time to travel through such a path is assumed to be a random variable under exponential distribution with the rate of λ_i. This assumption, thought not always true under all network environments, greatly reduces the analytic complexity. Note that our feasibility study does not intend to provide an accurate analysis, but instead to have a general observation of the capacity of FINDERS. Besides, we assume Tag i has a long term statistical probability of γ_{ik}^I to meet IR k and the probability of γ_{ij}^G to meet GR j. It is easy to verify that the time intervals to meet IR k and GR j are exponentially distributed with the rate of λ_{ik}^I and λ_{ij}^G, respectively, where $\lambda_{ik}^I = \lambda_i \gamma_{ik}^I$ and $\lambda_{ij}^G = \lambda_i \gamma_{ij}^G$. In addition, let $\lambda_i^I = \sum_{k=1}^K \lambda_{ik}^I$, which is the total rate for Tag i going to any IR's, and similarly $\lambda_i^G = \sum_{j=1}^J \lambda_{ij}^G$.

Based on above assumptions, a Markovian model (see Fig. 2) is established to analyze the meeting events between Tag i and the readers, and accordingly the data service rate of the IR's queue. The model consists of $K+2$ states. State S_k ($1 \le k \le K$) is the state that IR k can transmit m packets to Tag i (i.e., the tag is empty). State S_N is for the state where Tag i meets any IR's, but it is not empty and thus cannot receive data packets. State S_G indicates that Tag i meets a GR and delivers the data packets it carries. Once the packets are delivered to the GR, they are removed from the tag, which thus becomes empty and ready for receiving new data packets.

Let's denote P_k, P_N, and P_G the steady state probabilities of States S_k, S_N, and S_G, respective. Based on the Markovian model, we can derive the following state equations:

$$\begin{cases} P_k(\lambda_i^G + \lambda_i^I) = \lambda_{ik}^I P_G, \forall 1 \le k \le K, \\ P_N \lambda_i^G = \sum_{k=1}^K \lambda_i^I, \\ P_G \lambda_i^I = P_N \lambda_i^G + \sum_{k=1}^K (\lambda_{ik}^I P_k), \\ \sum_{k=1}^K P_k + P_N + P_G = 1. \end{cases} \quad (1)$$

Solve them, we have $P_G = \frac{\lambda_i^G}{\lambda_i^I + \lambda_i^G}$, $P_k = \frac{\lambda_{ik}^I \lambda_i^G}{(\lambda_i^G + \lambda_i^I)^2}$, and $P_N = \frac{{\lambda_i^I}^2}{(\lambda_i^G + \lambda_i^I)^2}$. To facilitate the understanding of FINDERS' general capacity, we further simplify the above results by letting $\gamma_{ik}^I = \gamma_{ij}^G = \gamma$ (i.e., Tag i have equal probability to meet all readers), and accordingly $\lambda_{ik}^I = \lambda_{ij}^G = \lambda$. Then we arrive at $P_k = \frac{J}{(K+J)^2}$, which is the steady state probability that Tag i becomes empty and arrives at an IR. With an overall rate of $(K+J)\lambda$ for Tag i to visit any readers, it has a rate of $(K+J)\lambda P_k$ to meet IR k when it is empty. Assume all tags are independent, the aggregated arrival rate of empty tags at an IR is $T(K+J)\lambda P_k$. Since each tag can carry m packets, we have the service rate of an IR's queue, $\mu = \frac{J}{J+K}\lambda m T$.

(3) Estimated Capacity. To ensure a stable queue at the IR, the service rate must be greater than the arrival rate. Therefore, the entire FINDERS with K IR's can accept no more than $\frac{JK}{J+K}\lambda m T$ packets, or

$$A \le \frac{JK}{J+K}\lambda m T, \quad (2)$$

which can serve as an estimation of the network capacity. It shows that, under typical scenarios, FINDERS can easily achieve a data rate of multiple kbps or higher (see Figs. 3 & 4 where by default a tag can hold 96 bits, K=5, m=1, T=100, λ=0.05, and J=2), which is sufficient to support a wide range of applications, such as wildlife tracking and lightweight monitoring. The capacity increases linearly with m, λ, and T, and nonlinearly with J and K.

In addition to the overall network capacity, we also have several interesting findings on the data generation in individual events. In general, new data are fed into the IR's queue only when it meets a tag. Without loss of generality, let's assume α packets arrive at the IR when such meeting event occurs. Thus, we have $\alpha \lambda T \le \mu$, or $\alpha \le \frac{Jm}{J+K}$. In other words, when a tag meets an IR, no more than $\frac{Jm}{J+K}$ packets can be generated and fed into the IR's queue. This result provides a guideline for traffic control in FINDERS and leads to two interesting observations. First, α is not related to the number of tags. This is because the tags contribute to both data arrival and service of individual IR. Second, increasing the number of IR's, K, results in a lower α. Though this is a little anti-intuitive (as the IR's help data relaying), it can be explained below. The IR's directly contribute to the network traffic load (when they meet tags), but do not directly consume the data (which can be successfully delivered only when they arrive at the GR's but not IR's). Thus, more IR's decrease the maximum allowed α. However, a larger K does help improve the overall network capacity, because they relay and buffer data packets, thus increasing the opportunity to deliver data to GR's. Additionally, a sufficiently large K is usually required by the applications, in order to achieve the needed coverage and granularity in data acquisition.

IV. Proposed Data Delivery Protocol

With extremely limited resources in buffer space, communication links, and battery power, FINDERS faces unprecedented challenges in communication and networking. As we have discussed in Sec. I, the sporadically available wireless links render it impossible to form a well connected mesh network for end-to-end communications. Moreover, the basic principle of DTN is to employ ample buffering for alleviating the needs of immediate transmission. This, however, becomes challenging in FINDERS, due to the extremely small memory and intermittent computing capability of its constituent tags (which are responsible for data delivery). In this research, we have explored novel adaptive approaches for efficient data transmission in FINDERS, by dynamically creating redundancy, managing data queues, and routing data packets. In the rest of this section, we will first introduce two basic parameters that are key to our proposed scheme, followed by the protocol description, implementation, and further discussions.

A. Two Key Parameters

1) Effective Delivery Capability (EDC): FINDERS is an opportunistic network, where the communication links exist with certain probabilities. It is essentially important to define an appropriate routing metrics, to represent the key resource and to indicate the nodal delivery capability in such networks. To this end, several approaches have been investigated, and they are largely based on the nodal meeting probabilities (or more sophisticated parameters derived from nodal meeting probabilities) in contrast to end-to-end delay or path length adopted in conventional networks. For example, in [2], a node estimates its capability of transmitting data to a destination by its direct contact probability with the destination. While simple, it has been shown inefficient [5], especially when the source is far away from the destination. Another approach is to let each node maintain the pair-wise contact probability of any two nodes in the network. Then, the shortest path algorithm is applied to identify the path with the highest probability between a given source and destination pair [42]. Its drawback is the high overhead in storage and communications, which is unaffordable in FINDERS where the tags have extremely limited storage space. To lower overhead, [43] employs a transitive equation to update the routing metrics to each possible destination. In addition, several systems [3], [36] do not maintain any routing metrics. All links and nodes are treated equally in making routing decisions. This is inefficient in many practical applications where nodes have nonequal probabilities to reach a destination.

In this research we have investigated simple and efficient online estimation approaches to obtain the effective delivery capability (EDC) of tags and readers. Such online estimation usually involves tradeoffs among complexity, stability, agility, and required memory space. While many online estimators are available, the exponentially weighted moving average (EWMA) is one of the most effective schemes and has been employed in various applications (including TCP RTT calculation). EWMA is simple, reacts effectively to small shifts, and is memory efficient, requiring only constant space. Note that although EWMA is a well-defined method, it remains challenging to design appropriate online updating mechanism to obtain EDC in FINDERS. In particular, since the tag cannot perform continuous computation, effective schemes must be devised for calculating EDC, based on the necessary state information saved in its nonvolatile memory. Let ξ_i denote the parameter of Node i, which may be a reader or a tag. ξ_i is to be averaged and estimated dynamically. A GR always has its $\xi_i = 1$. For an IR (or a tag), ξ_i is initialized as zero and updated upon an event of meeting a tag (or a reader). Denote t_o the time at which the last event occurred at Node i and thus ξ_i was updated. When a reader and a tag meet, the reader acquires the current EDC and t_o of the tag and calculates the new EDC's for both of them. More specifically, if Node i meets Node j at time t_1, ξ_i is updated as follows,

$$\xi_i = (1-\eta)^{\lfloor \frac{t_1 - t_o}{\Delta} \rfloor}[\xi_i] + \eta\hat{\xi}, \tag{3}$$

where $0 \leq \eta \leq 1$ is a constant weight to keep partial memory of historic status, Δ is a constant that regulates the decrease of ξ_i, $[\xi_i]$ is the parameter of Node i before it is updated, and $\hat{\xi} = [\xi_j]$ if $[\xi_j] > [\xi_i]$ or $\hat{\xi} = 0$ otherwise. The updated EDC and t_o are kept at the reader and written into the tag's nonvolatile memory, respectively. Clearly, this scheme requires time synchronization among readers, which can be achieved by various means with different degrees of accuracy.

Our studies show that EWMA effectively captures the special characteristics of FINDERS, with two interesting findings outlined below:

Result 1: the expectation of ξ_i converges to constant under statically distributed mobility. More specifically, let $\tilde{\xi}_i$ denote the limit of the expectation of ξ_i, we have, $\tilde{\xi}_i = \lim_{t\to\infty} E(\xi_i(t))$. Due to the space limit, detailed proof is omitted here. This result shows the stability of EDC, which is essentially important for it to serve as the routing parameter in FINDERS, ensuring efficient and stable routing.

Result 2: EDC is an effective estimation of end-to-end data delivery rate. As revealed by a separate study of ours under the

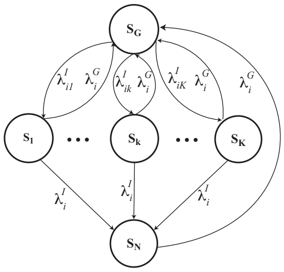


Fig. 2. Markovian model for Tag i.

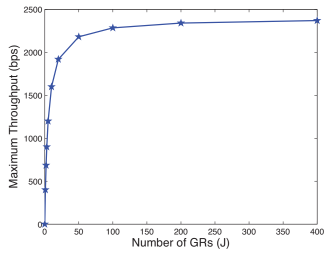


Fig. 3. Throughput vs. J.

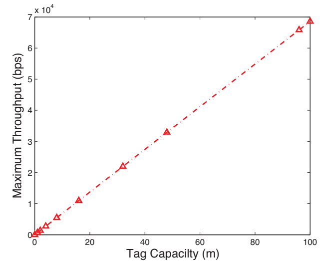


Fig. 4. Throughput vs. m.

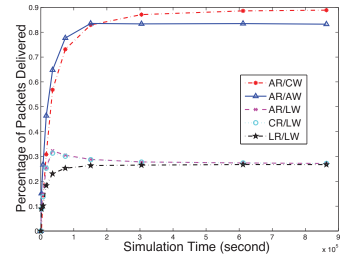


Fig. 5. Reading/writing strategies.

real mobility trace [44], EDC exhibits excellent consistency and closely follows the actually delivery rate. Thus, when EDC serves as the routing metrics, data packets always flow from the nodes with lower delivery rate to the nodes with higher delivery rate, achieving efficient data delivery in FINDERS.

2) Fault Tolerance Degree (FTD): In a typical store-and-forward network, the packets are deleted from the buffer after they are transmitted to the next hop successfully. In FINDERS (like many other DTN networks), however, the reader may still keep a copy of the data after its transmission, in order to maintain necessary redundancy for achieving a given success ratio. Therefore, multiple copies of the data could be created and stored by different readers and transported by different tags in the network. The fault tolerance is introduced to represent the amount of redundancy and to indicate the importance of a given data packet. We assume that each packet carries a field that keeps its fault tolerance degree (FTD). Let $\mathcal{F}_i^M$ denote the FTD of Packet M in the queue of Reader i. It is defined to be the probability that at least one copy of Packet M is delivered to a GR by the tags in the network. When a packet is generated, its FTD is initialized to be zero and updated during transmissions. Let's consider a general scenario where Reader i transmits Packet M to Tag j. This transmission essentially creates two copies of Packet M, each of which shares part of the responsibility to deliver the data to the GR. An appropriate FTD needs to be assigned to them, respectively. More specifically, the packet transmitted to Tag j is associated with an FTD of $\mathcal{F}_j^M$,

$$\mathcal{F}_j^M = 1 - (1 - [\mathcal{F}_i^M])(1 - \xi_i), \qquad (4)$$

and the FTD of the packet at Reader i is updated as

$$\mathcal{F}_i^M = 1 - (1 - [\mathcal{F}_i^M])(1 - \xi_j), \qquad (5)$$

where $[\mathcal{F}_i^M]$ is the FTD of Packet M at Reader i before the transmission. Eqs. (4) & (5) are derived based on the definition of FTD introduced above to reflect the probability that the corresponding data is delivered. The update of FTD repeats each time upon a data transmission. In general, the more times the data packet has been forwarded, the more copies are created, thus increasing its delivery probability. Accordingly, it is then associated with a larger FTD. As to be discussed next, FTD is important in the control of duplication redundancy and the queue management.

B. An Overview of Proposed Data Delivery Protocol

The data transmission protocol needs to be carefully devised based on EDC and FTD, in order to achieve efficient network resource utilization. To this end, we propose a simple and effective scheme, as outlined below, with six steps:

- Step 1: Data Retrieval. Each reader periodically broadcasts query signals. When nearby tags are detected, the reader reads in their data and EDC parameters.
- Step 2: Update of EDC. EDC's are calculated for both the reader and the tags according to Eq. (3).
- Step 3: Data Queuing. The reader inserts the data into its queue, which is sorted according to FTD.
- Step 4: Receiver Identification. One or multiple nearby tags will be identified for data transportation.
- Step 5: Data Duplication. The data packets at the top of the queue are selected for transmission. For each such packet, two or more copies are created and processed (including the update of their FTD's according to Eqs. (4) and (5)).
- Step 6: Data transmission. For each chosen tag, a package that includes the updated EDC of the tag and the selected data packets with updated FTD's is created and transmitted via a writing operation.

Next we elaborate each of the six steps, and several problems, tradeoffs and alternatives involved therein. To facilitate our discussion, we start with Step 3 (Data Queuing), followed by Steps 4 & 5, and finally Steps 1 & 6. Step 2 is done straightforwardly according to Eq. (3).

1) Queue Management: Data queuing and queue management are essentially important to the efficiency of FINDERS. As indicated in Step 3, once the reader acquires data from nearby tags, it inserts the data into its queue. The queue management is meant to appropriately sort the data packets in the queue, determine which data packet to be sent, and decide which data packet to drop if the queue is full. Our proposed queue management scheme is based on FTD, which signifies how important a data packet is. The packet with a smaller FTD is more important and should be transmitted with a higher priority. This is done by sorting the packets in the queue in an increasing order of their FTD's. A packet with the smallest FTD is always at the top of the queue and transmitted first. A packet is dropped on the following two occasions. First, when the queue is full, the packet with the highest FTD is dropped. Second, if the FTD of a packet is larger than a threshold, it is dropped, even if the queue is not full. This is to reduce

TABLE I

COMBINATIONS OF READING AND WRITING STRATEGIES.

	Limited-Read (LR)	Conservative-Read (CR)	Aggressive-Read (AR)
Limited-Write (LW)	LR/LW	CR/LW	AR/LW
Conservative-Write (CW)	LR/CW	CR/CW	AR/CW
Aggressive-Write (AW)	LR/AW	CR/AW	AR/AW

transmission overhead, given that the data contained in the packet will be delivered to the GR's with a high probability by other nodes that carry duplicate copies of the data.

2) Receiver Identification: In Steps 4 and 5, the reader needs to identify one or multiple tags for data transportation and to determine the amount of duplication, by addressing the tradeoff between overhead and delivery probability. Without loss of generality, let us consider Reader i, which has a set of tags in its reading/writing range. Reader i first learns their EDC's, and sorts them in a descending order. Let Ξ designate the set of sorted tags whose EDC's are greater than ξ_i. Next, Reader i simply picks up Packet M at the top of its queue and transmits it to the first tag in Ξ. As we have discussed earlier, this transmission essentially creates two copies of the data. Each of them is associated with an FTD calculated according to Eqs. (4) and (5). Then, the copy at Reader i is put back into its queue at the appropriate position based on its newly updated FTD (i.e., $\mathcal{F}_i^M$). As indicated in the queue management, this copy can be dropped if $\mathcal{F}_i^M$ is larger than a threshold, in order to reduce unnecessary transmission overhead. This process repeats until all tags in Ξ have been occupied or the queue of Reader i becomes empty. Note that if the original FTD of a data packet is very low, it may be put back to the top of the queue again after a transmission and thus be transmitted multiple times, which is necessary to achieve a given data delivery probability.

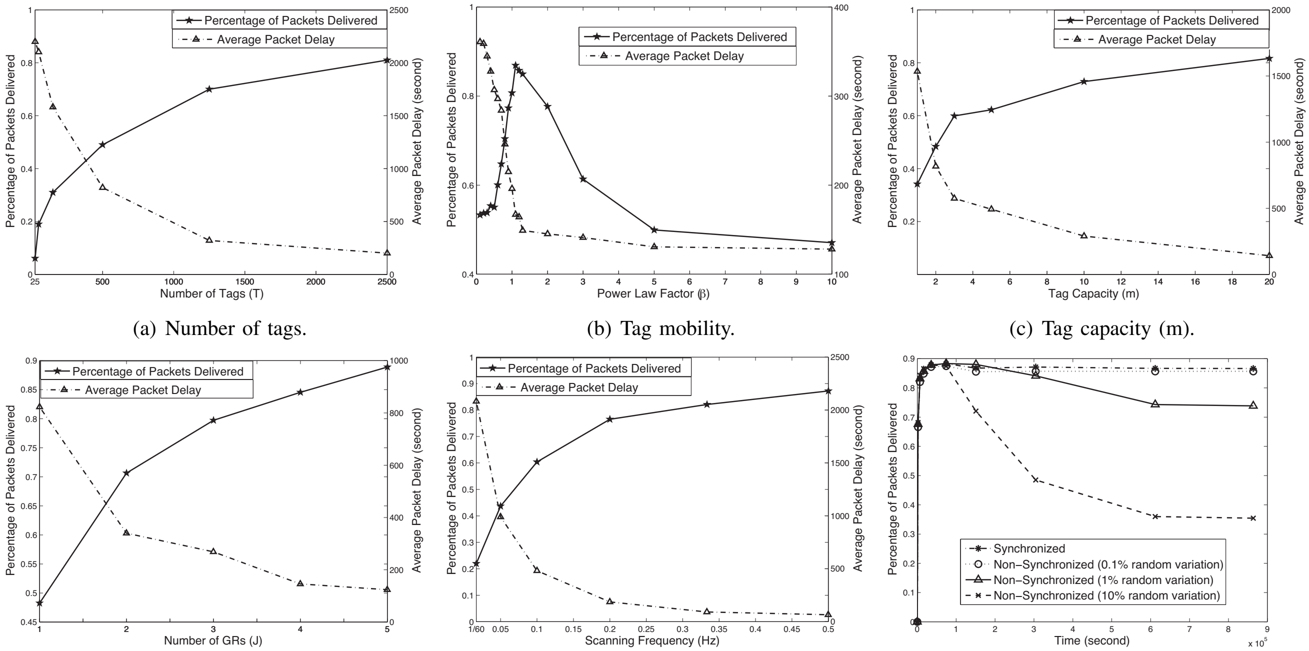
3) Reading and Writing Strategies: Data retrieval (i.e., Step 1) and data transmission (i.e., Step 6) are realized via reading and writing operations, respectively. Our studies have shown that the reading and writing strategies may significantly affect the communication efficiency of FINDERS. In general, three strategies are available for reading and writing, respectively, as outlined below: (a) Limited-Read (LR): only GR's (but not IR's) read data from tags; (b) Conservative-Read (CR): any reader with a higher EDC than that of the tag can read data from the tag; (c) Aggressive-Read (AR): the reader reads data from any neighboring tags; (d) Limited-Write (LW): the reader writes data into a tag only if the tag is empty; (e) Conservative-Write (CW): the reader only writes its data to the tag which has higher EDC; and (f) Aggressive-Write (AW): the reader writes data to any neighboring tags.

These reading and writing strategies yield totally nine combinations as shown in Table I. Among them, "LR/LW", "CR/LW", and "AR/LW" are obviously inefficient, because the reader is allowed to write data into a tag only when the tag is empty. Under this strategy, only the tags that commute directly between the GR's and the IR's are used for data delivery. This seriously limits the utilization of the tags, which are the critical transportation vehicles in FINDERS. We further observe that "LR/CW", "LR/AW", "CR/CW", and "CR/AW" are inefficient either, as the reader may overwrite the tag without reading out its data in some or all scenarios, resulting in high data lost rate. Therefore, we finally arrive at two feasible approaches: "AR/CW" and "AR/AW". The difference between them lies in their writing strategies. Intuitively, "CW" is a better approach, as it generates less redundancy than that of "AW" and avoids unnecessary data duplications, by controlling the data packets to flow from the nodes with lower EDC to the nodes with higher EDC. In order to gain a quantitive observation of the impact of these reading/writing strategies, we carry out simulations with 1 GR, 4 IR's, and 500 tags. The detailed description of our simulation setup is deferred to Sec. V. The results are shown in Fig. 5. As can be seen, "AR/CW" achieves the highest data delivery rate. Therefore, it is adopted for FINDERS by default in our later discussions.

V. SIMULATION RESULTS

Extensive simulations are indispensable for a comprehensive evaluation of FINDERS with a large number of readers and tags, which are not affordable to build in labs. Our simulated network consists of 25 readers, uniformly deployed in a bounded area. Each reader can serve as either an IR or a GR. A number of tags are distributed in the network, and move according to power-law distribution, which is deemed as one of the most realistic mobility models for delay-tolerant mobile networks [45]. More pecifically, the network area is partitioned into cells. Each tag is randomly assigned a home cell. Tag i makes its decision to move into one of the neighboring cells or stay at the current cell in every time slot. For instance, if it is currently in Cell 0, it may move into one of four adjacent cells (i.e., Cells 1-4) or stay in the current cell (Cell 0) in the next time slot. Its probability to be in Cell x in the next time slot is $P_i(x|0) = P_i(x)/\sum_{z=0}^{4} P_i(z)$, where $x = 0, 1, 2, 3, 4$ and $P_i(x) = k_i(\frac{1}{d_i(x)})^{\beta}$. k_i is a constant and β is the exponent of the power-law distribution, respectively. $d_i(x)$ denotes the distance from Cell x to the home cell of Tag i. The default parameters used in our simulations are summarized below. One GR and 500 tags are simulated. β=1.2. The tag capacity is m=2 packets. The data arrival rate at a reader is 1 packet per 10 time slots. We are mainly interested in the network throughput (defined as the percentage of data packets that can be delivered) and the average packet delay.

We first study the performance of FINDERS by varying several parameters related to the tags. As shown in Fig. 6(a), the network throughput increases and the average packet delay decreases with the number of tags increasing from 25 to 2500, as more tags result in more opportunities for data communication between readers and tags (i.e., higher service rate). However, the gain becomes smaller when the tag density is already high. This is because when the tag density is high enough, additional tags will carry unnecessarily duplicated data, and thus do not help further improve network



(a) Number of tags.

(b) Tag mobility.

(c) Tag capacity (m).

(d) Number of GR's.

(e) Scanning frequency.

(f) Synchronization errors.

Fig. 6. Impact of several key parameters on the performance of FINDERS.

performance. The power-law factor β determines the mobility pattern of the tags. Under very small β, all sensors tend to have similar (uniform) mobility and accordingly very close EDC. Thus packet relay among nodes is ineffective, leading to low delivery rate (see Fig. 6(b) where J=4). With very large β, the tag has a high chance to stay at its home location, which means less mobility. Therefore, the throughput decreases too. Meanwhile, although it is anti-intuitive to observe the delay decreases with the increase of β, it is in fact reasonable because the average delay is calculated based on the delivered data packets only. When β is large, most packets are delivered by the tags whose home locations are close to the GR's and thus experience short delay. In addition, as shown in Fig. 6(c), increasing the tag capacity can improve the performance of FINDERS by allowing more packets be read from and written into the tag. Fig. 6(d) illustrates the results with the number of GR's increasing from 1 to 5. Clearly, more GR's improve the chance for the tags to meet GR's, and thus higher data deliver rate and lower packet delay. We have also studied the impact of readers' scanning frequency (see Fig. 6(e)). As the scanning frequency increases from 1/60 to 1/2 Hz, the readers can grasp more opportunities to communicate with tags and deliver data packets, leading to higher throughput and lower delivery delay.

The impact of synchronization errors is also studied by our simulations, as shown in Fig 6(f) (where J=4). As we have discussed in Sec. IV-A.1, time synchronization among the readers is required to update EDC according to Eq. (3). Any inaccuracy in time synchronization will lead to errors in the calculation of EDC. For example, consider an IR with a clock faster than the clocks of other readers in the network. Its EDC will be less than its true value, because t_1 in Eq. (3) becomes bigger due to the fast clock. Similarly, a slow clock leads to a higher EDC. A small error of EDC may not hurt the network performance, as long as the order of the EDC's remain the same (i.e., the node with higher true EDC still maintains a higher EDC although with errors). In practice, the readers are synchronized to the standard global time before they are deployed. Once a reader (especially the IR) is deployed, however, it is difficult to keep the synchronization, because it is isolated from the time servers. Hereafter, the clock skew will lead to bigger and bigger errors in synchronization. But fortunately, our simulation results show that if the clock skew is no more than 0.1%, the system can healthily run for at least one month with no significant performance degradation. Considering the typical precision of a CPU clock, the skew may range from 10^{-6} to 10^{-20} seconds per second. Therefore the system can maintain its highly efficient data delivery for years without the need of resynchronization, which is sufficient in most applications.

VI. Conclusion

We have proposed a Featherlight Information Network with Delay-Endurable RFID Support (FINDERS), composed of passive RFID tags which are ultra light, durable, and flexible, without power supply for long-lasting applications under strict weight constraints and harsh environments. It expands the use of RFID gear for wireless network construction, aiming to find events of interest and gather aggregate information. FINDERS faces unprecedented challenges in communication and networking, due to its sporadic wireless links, unique asymmetric communication paradigm, intermittent computation capability, and extremely small memory of tags. We have conducted analysis and simulations to show the feasibility and efficiency of FINDERS.

REFERENCES

[1] I. Akyildiz, W. Su, and Y. Sankarasubramaniam, "A Survey on Sensor Networks," *IEEE Comm. Magazine*, vol. 40, no. 8, pp. 102–114, 2002.
[2] http://www.princeton.edu/ mrm/zebranet.html.
[3] T. Small and Z. J. Haas, "The Shared Wireless Infostation Model – A New Ad Hoc Networking Paradigm (or Where there is a Whale, there is a Way)," in *Proc. of ACM International Symposium on Mobile Ad Hoc Networking and Computing (MOBIHOC)*, pp. 233–244, 2003.
[4] http://www.wu.ece.ufl.edu/projects/DeerNet/DeerNet.html.
[5] Y. Wang and H. Wu, "DFT-MSN: The Delay Fault Tolerant Mobile Sensor Network for Pervasive Information Gathering," in *Proc. of IEEE Conference on Computer Communications (INFOCOM)*, pp. 1–12, 2006.
[6] V. Cerf, S. Burleigh, A. Hooke, L. Torgerson, R. Durst, K. Scott, K. Fall, and H. Weiss, "Delay Tolerant Network Architecture." draft-irtf-dtnrg-arch-02.txt, 2004.
[7] K. Ali, H. Hassanein, and A.-E. M. Taha, "Rfid anti-collision protocol for dense passive tag environments," in *Proc. of IEEE Conference on Local Computer Networks (LCN)*, pp. 819–824, 2007.
[8] M. Jacomet, A. Ehrsam, and U. Gehrig, "Contactless identification device with anti-collision algorithm," in *Proc. of International Technical Conference on Circuits, Systems, Computers and Communications (CCNC)*, 1999.
[9] H. Vogt, "Efficient object identification with passive rfid tags," in *Proc. of International Conference on Pervasive Computing*, 2002.
[10] J. Myung and W. Lee, "Adaptive splitting protocols for rfid tag collision arbitration," in *Proc. of ACM International Symposium on Mobile Ad hoc Networking and Computing (MOBIHOC)*, pp. 202–213, 2006.
[11] D. H. amd C. Wood, "Analysis of tree algorithms for rfid arbitration," in *Proc. of IEEE International Symposium on Information Theory*, p. 107, 1998.
[12] C. Law, K. Lee, and K.-Y. Siu, "Efficient memoryless protocol for tag identification," in *Proc. of 4th international workshop on Discrete algorithms and methods for mobile computing and communications*, pp. 75–84, 2000.
[13] V. Namboodiri and L. Gao, "Energy-aware tag anti-collision protocols for rfid systems," in *Proc. of IEEE International Conference on Pervasive Computing and Communications (PERCOM)*, pp. 23–36, 2007.
[14] M. Kodialam and T. Nandagopal, "Fast and reliable estimation schemes in rfid systems," in *Proc. of ACM International Symposium on Mobile Ad hoc Networking and Computing (MOBIHOC)*, pp. 322–333, 2006.
[15] C. Qian, H.-L. Ngan, and Y. Liu, "Cardinality estimation for large-scale rfid systems," in *Proc. of IEEE International Conference on Pervasive Computing and Communications (PERCOM)*, 2008.
[16] C. C. Tan, B. Sheng, and Q. Li, "How to monitor for missing RFID tags," in *Proc. of International Conference on Distributed Computing Systems (ICDCS)*, 2008.
[17] B. Sheng, C. C. Tan, Q. Li, and W. Mao, "Finding popular categoried for RFID tags," in *Proc. of ACM International Symposium on Mobile Ad Hoc Networking and Computing (Mobihoc)*, 2008.
[18] Q. Dong, A. Shukla, V. Shrivastava, D. Agrawal, and S. Banerjee, "Load balancing in largescale rfid systems," in *Proc. of IEEE Conference on Computer Communications (INFOCOM)*, pp. 2281–2285, 2007.
[19] D. Hahnel, W. Burgard, D. Fox, K. Fishkin, and M. Philipose, "Mapping and localization with RFID technology," in *Proc. of IEEE International Conference on Robotics and Automation*, pp. 1015–1020, 2004.
[20] K. Yamano, K. Tanaka, M. Hirayama, E. Kondo, Y. Kimuro, and M. Matsumoto, "Self-localization of mobile robots with RFID system by using support vector machine," in *Proc. of IEEE/RSJ Int'l. Conference on Intelligent Robots and Systems*, pp. 3756–3761, 2004.
[21] Y. C. L. Linel M. Ni, Yunhao Liu and A. P. Patil, "Landmarc: Indoor location sensing using active RFID," in *Proc. of IEEE Conference on Pervasive Computing and Communications*, pp. 407–415, 2003.
[22] C. Wang, H. Wu, and N.-F. Tzeng, "RFID-Based 3-D Positioning Schemes," in *Proc. of IEEE Conference on Computer Communications (INFOCOM)*, pp. 1235–1243, 2007.
[23] K. Fall, "A delay-tolerant Network Architecture for Challenged Internets," in *Proc. of ACM SIGCOMM Conference*, pp. 27–34, 2003.
[24] http://www.cens.ucla.edu/.
[25] A. Mainwaring, J. Polastre, R. Szewczyk, D. Culler, and J. Anderson, "Wireless Sensor Networks for Habitat Monitoring," in *Proc. of ACM International Workshop on Wireless Sensor Networks and Applications (WSNA)*, pp. 88–97, 2002.
[26] http://down.dsg.cs.tcd.ie/sendt/.
[27] http://www.sics.se/cna/dtnsn/index.html.
[28] M. Ho and K. Fall, "Poster: Delay Tolerant Networking for Sensor Networks," in *Proc. of IEEE Conference on Sensor and Ad Hoc Communications and Networks*, 2004.
[29] R. C. Shah, S. Roy, S. Jain, and W. Brunette, "Data MULEs: modeling a three-tier architecture for sparse sensor networks," in *Proc. of the First International Workshop on Sensor Network Protocols and Applications*, pp. 30–41, 2003.
[30] H. Wu, Y. Wang, H. Dang, and F. Lin, "Analytic, Simulation, and Empirical Evaluation of Delay/ Fault-Tolerant Mobile Sensor Networks," *IEEE Transactions on Wireless Communications*, vol. 6, no. 9, pp. 3287–3296, 2007.
[31] Y. Wang and H. Wu, "Delay/Fault-Tolerant Mobile Sensor Network (DFT-MSN): A New Paradigm for Pervasive Information Gathering," *IEEE Transactions on Mobile Computing*, vol. 6, no. 9, pp. 1021–1034, 2007.
[32] Y. Wang, H. Wu, F. Lin, and N.-F. Tzeng, "Cross-layer Protocol Design and Optimization for Delay/Fault-Tolerant Mobile Sensor Networks," *IEEE Journal on Selected Areas in Communications (JSAC), Special Issue on Delay and Disruption Tolerant Wireless Communication*, June 2008. A preliminary version was presented in IEEE ICDCS'07.
[33] D. J. Goodman, J. Borras, N. Mandayam, and R. Yates, "Infostations: A new system for data and messaging services," in *Proc. of IEEE Vehicular Technology Conference*, pp. 969–973, 1997.
[34] Y. Wang, S. Jain, M. Martonosi, and K. Fall, "Erasure-Coding Based Routing for Opportunistic Networks," in *Proc. of ACM SIGCOMM workshop on Delay Tolerant Networking and Related Topics*, pp. 229 – 236, 2005.
[35] B. Burns, O. B. Brian, and N. Levine, "MV Routing and Capacity Building in Disruption Tolerant Networks," in *Proc. of IEEE Conference on Computer Communications (INFOCOM)*, pp. 398– 408, 2005.
[36] T. Spyropoulos, K. Psounis, and C. S. Raghavendra, "Spray and Wait: An Efficient Routing Scheme for Intermittently Connected Mobile Networks," in *Proc. of ACM SIGCOMM workshop on Delay Tolerant Networking and Related Topics (WDTN)*, pp. 252 – 259, 2005.
[37] P. Hui, A. Chaintreau, J. Scott, R. Gass, J. Crowcroft, and C. Diot, "Pocket Switched Networks and Human Mobility in Conference Environments," in *Proc. of ACM SIGCOMM workshop on Delay Tolerant Networking and Related Topics (WDTN)*, pp. 244–251, 2005.
[38] M. Musolesi, S. Hailes, and C. Mascolo, "Adaptive Routing for Intermittently Connected Mobile Ad Hoc Networks," in *Proc. of IEEE 6th International Symposium on a World of Wireless, Mobile and Multimedia Networks (WOWMOM)*, pp. 1–7, 2005.
[39] J. LeBrun, C.-N. Chuah, and D. Ghosal, "Knowledge Based Opportunistic Forwarding in Vehicular Wireless Ad Hoc Networks," in *Proc. of IEEE Vehicular Technology Conference*, pp. 1–5, Spring, 2005.
[40] A. Lindgren, A. Doria, and O. Schelen, "Probabilistic Routing in Intermittently Connected Networks," in *Proc. of The First International Workshop on Service Assurance with Partial and Intermittent Resources (SAPIR 2004)*, 2004.
[41] W. Zhao, M. Ammar, and E. Zegura, "A Message Ferrying Approach for Data Delivery in Sparse Mobile Ad Hoc Networks," in *Proc. of the 5th ACM International Symposium on Mobile Ad Hoc Networking and Computing (MOBIHOC)*, pp. 187–198, 2004.
[42] J. Ghosh, J. Philip, and C. Qiao, "Sociological Orbit Aware Location Approximation and Routing (SOLAR) in MANET," *Ad Hoc Networks*, vol. 5, no. 2, pp. 189–209, 2007.
[43] A. Lindgren, A. Doria, and O. Schelen, "Probabilistic Routing in Intermittently Connected Networks," in *Proc. of the First Int'l Workshop on Service Assurance with Partial and Intermittent Resources*, 2004.
[44] http://www.cs.dartmouth.edu/ campus/data.html.
[45] J. Leguay, T. Friedman, and V. Conan, "DTN Routing in a Mobility Pattern Space," in *Proc. of ACM SIGCOMM05 Workshop on Delay Tolerant Networking and Related Topics*, pp. 276 – 283, 2005.