This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2016.2586581, IEEE Internet of Things Journal

Robust Relay Selection for Large-Scale Energy-Harvesting IoT Networks

Hiroki Kawabata, *Student Member, IEEE,* Koji Ishibashi, *Member, IEEE,* Satyanarayana Vuppala, *Member, IEEE,* and Giuseppe Abreu, *Senior Member, IEEE,*

***Abstract*—We consider the relay selection problem in large-scale energy-harvesting (EH) networks. It is known that if channel state information (CSI) is available at EH relays, a diversity order equal to the number of relays can be obtained, however at the penalty of a feedback overhead (necessary to obtain accurate CSI) which is not suitable for energy-limited devices intended $e.g.$ for internet-of-things (IoT) applications. In this paper, we therefore propose a new EH relay selection scheme which is based on the residual energy at each relay's battery, and on information on the distribution of the channels between relays and the destination. The method thus minimizes both the outage probability and the feedback cost. Where previous work relay selection based on channel distribution information (CDI) consider only small-scale fading distribution, we employ a stochastic geometry approach to consider jointly the geometrical distribution ($i.e.$, large-scale fading) and small-scale fading yielding a simple relay selection criterion that furthermore utilizes only rough information on the relay's location, $i.e.$, an ordinal number from the destination. The outage probability of the proposed relay selection scheme is analytically derived, and the achievable diversity order of the proposed approach is investigated. Computer simulations confirm our theoretical analyses and show that our approach is robust against errors in the estimation of the distances between nodes.**

***Index Terms*—opportunistic relaying, channel distribution information (CDI), stochastic geometry, energy harvesting.**

I. Introduction

In 2011, Cisco Internet Business Solution Group (IBSG) predicted that *there will be 50 billion devices connected to the internet by 2020* [1]. This vast network of devices, would enable us to gain information on virtually anything, creating enormous opportunities for new services and applications, for things such as logistics, transportation, health care, agriculture and so on. For such an internet-of-things (IoT) applications, the quality of wireless communications and low energy consumption are fundamental, since the former defines applicable areas, while the latter the readiness of energy-limited devices.

Manuscript submitted on Dec. 31, 2015; revised on June 3, 2016; accepted on xxx. This work was supported in part by Semiconductor Technology Association Research Center (STARC).

H. Kawabata and K. Ishibashi are with the Advanced Wireless & Communication Research Center (AWCC), The University of Electro-Communications, Tokyo 182-8585, Japan. E-mails: hiroki@awcc.uec.ac.jp and koji@ieee.org, respectively.

S. Vuppala is with the School of Engineering, Institute for Digital Communications (IDCOM), The University of Edinburgh, EH9 3JU, UK. E-mail: S.Vuppala@ed.ac.uk

G. Abreu is with the School of Engineering and Sciences, Jacobs University Bremen, 28759 Bremen, Germany. G. Abreu is simultaneously with Department of Electrical Engineering and Electronic Engineering, Ritsumeikan University, Shiga 525-8577 Japan. E-mail: g.abreu@jacobs-university.de and g-abreu@fc.ritsumei.ac.jp, respectively.

Typical wireless channels, however, suffer from multipath fading and shadowing, which significantly reduce communication capacity for a given average transmission power and hinder reliable transmission. Although an effective option is to use multiple antennas to obtain spatial diversity gain [2], in practice, it is difficult equip small IoT devices with multiple antennas due to their size, complexity, and cost. Hence, another concept has been proposed: when the source cannot reliably communicate directly with its destination, other nodes temporarily serve as relays in order to support the communication.

This *cooperative diversity* approach allows devices to enjoy spatial diversity gain without the need to equip them with additional antennas [3], and it has been shown that if sufficiently many relays are available, *opportunistic relying* can attain a diversity order equal to the number of relays itself [4]. In opportunistic relaying, the *best* relay among those available is chosen based on the perfect knowledge of the instantaneous channel state information (CSI), where best is defined in terms of the corresponding instantaneous signal-to-noise ratio (SNR) at the destination. However, in a cooperative diversity system, relays consume their own battery order to support other nodes' communication, so that if multiple nodes drain their batteries at the same time, the network life-time or its topology may quickly deteriorate.

A remedy for this crucial issue is the use of energy harvesting (EH) [5] in combination with opportunistic relaying [6], [7]. Energy harvesting makes it possible to use solar, kinetic, wind, electromagnetic, or other types of energy sources to recharge the relay nodes' batteries. The result of using EH in opportunistic relaying is that non-selected relays efficiently use their inactive time to recharge their batteries while the selected relay forwards the source's information in order to obtain diversity. In [6], the symbol error rate (SER) of a cooperative network with EH relays was derived theoretically, and the advantages of opportunistic relaying was shown for the case in which the selection is based on the current available energy and the CSI. If channel fluctuation is slow, however, there is a possibility that the same relay is selected repeatedly, so that its activity surpasses its ability to harvest energy to recharge, such that considering both the relays suitability and its energy state in the is necessary selection process.

In light of the above, a new relay selection scheme based on the relative throughput gain of each relay and its energy state information (ESI) was proposed in [7]. This method of relay selection was shown to improve short-term performance, since the harvested energy can be efficiently utilized.

A remaining limitation of the aforementioned method is however, the assumption of perfect CSI at relays. Indeed, even if perfect channel reciprocity holds, this assumption implies the need that every relay accurately estimates its channel to every other relay (possible destination), and forward the information to its peers. In other words, the assumption perfect CSI is virtually unfeasible in the context of IoT networks, not to mention that perfect-CSI schemes are known to suffer from significant performance degradation in face of channel estimation erros.

To overcome this challenge, opportunistic relaying methods avoiding the requirement for perfect CSI have been proposed [8]–[10]. In [8], a scheme relying on the average channel gains only was presented. To some extent, average CSI does indeed indirectly capture the contribution of network topology (specifically the distance between relays), but leaves out the contribution of fading.

A complementary approach was proposed in [9], [10], where CSI knowledge was replaced by *channel distribution information (CDI)*. This method was shown that this results in a better performance than when the selection is based on the mean. However, the approach (like its predecessor) requires that relays know the actual distance between nodes, which in fact defeats the very purpose of opportunistic relaying, since actual location information also implies a heavy overhead.

In light of all the above, the contributions of this paper are as follows:

- We introduce a stochastic geometry to model large-scale networks with EH relays. This model captures both the topological and fading contributions to channel fluctuation that affect relay selection.
- We propose a new EH relay selection scheme based on residual batteries of relays and the CDI of both small-scale and large-scale fading. This requires neither extra communication nor does it result in increased computational complexity, and thus it is suitable for IoT devices.
- The closed form of the end-to-end outage probability is derived. Based on this result, we devise a simple relay selection rule that only requires approximate information about the relay's location; specifically, it only requires the ordinal number of the relays.
- The achievable diversity order of the proposed approach is also theoretically investigated.
- We show that our selection scheme is robust against errors in estimating the distances between nodes.

This paper is organized as follows. Section II describes the system model that we use throughout this paper. In Section III, the new EH selection rule is described, and then we derive the outage probability for a given relay. Based on this result, we obtain a simple rule for the proposed EH selection. Moreover, the overall outage probability of our proposed system is derived theoretically by analyzing the EH process and considering the achievable diversity order of the proposed approach. In Section IV, we present numerical results that confirm the theoretical analyses, and we clarify the advantages of the proposed system. Finally, we summarize our conclusions in Section V.

II. System Model

A. Network Model

We consider a network that is composed of a single source (S), a single destination (D), and multiple relays. In particular, we consider a toroidal network in an $M \times M$ $[\mathrm{m}^2]$ Euclidean space of dimension $d = 2$, modeled by a stationary Poisson point process (PPP) of intensity λ in $\mathbb{R}^d$ [12]. In this network, the source is located at the coordinate $(M/2, M/2)$, the destination is located at the origin, and the relays are independently located following a uniform distribution, as shown in Fig. 1. We define the distance between the k-th relay and the destination as $r_{k\mathrm{D}}$, such that $0 < r_{1\mathrm{D}} < r_{2\mathrm{D}} < \cdots$.

We will assume frequency nonselective block Rayleigh fading in all of the links between the transmitter A$\in \{\mathrm{S}, k\}$ and the receiver B$\in \{k, \mathrm{D}\}$, and thus the coefficients are constant during each transmission block. The fading coefficient h_{AB} follows a complex Gaussian distribution with zero mean and unit variance and is available at the receiver. We will assume that the path loss degrades the received power by $1/(1+r_{\mathrm{AB}}^{\beta})$ according to the distance between A and B, where β denotes the path-loss exponent, which, in practice, ranges from 2 to 4. Without loss of generality, β will be assumed to be equal to 2 for the remainder of this paper. Note that the path loss coefficients regularly include the effects of the transmit and receive antenna gains. However, our interest here is to evaluate how the geometrical relationships between the nodes affects the overall performance rather than how it affects the antenna gains. Thus, the path loss is modeled simply as $1/(1 + r_{\mathrm{AB}}^2)$. Furthermore, due to the half-duplex constraint, every transmission is performed in two phases. In the first phase, the source broadcasts its own information. In the second phase, one selected relay forwards this information to the destination, using variable gain amplify-and-forward (VG-AF), which scales and amplifies the received signals [13]. Since the relay has CSI for the channel between the source and itself, the variable amplification coefficient $\alpha_{\mathrm{S}k}$ is given as

$$\alpha_{\mathrm{S}k} = \left(\sqrt{\frac{P_{\mathrm{S}}|h_{\mathrm{S}k}|^2}{1+r_{\mathrm{S}k}^2} + N_0}\right)^{-1}, \tag{1}$$

where P_{A} is the power of the transmitter, and N_0 is the density of additive white Gaussian noise (AWGN).

For the remainder of the paper, we will assume that $N_0 = 1$. Also, we will assume that a direct link between S and D is not available. Thus the instantaneous SNR from the source to the destination through the relay can be expressed as [3]

$$\Gamma_{\mathrm{AF}}^{k} = \frac{\Gamma_{\mathrm{S},k}\Gamma_{k,\mathrm{D}}}{1+\Gamma_{\mathrm{S},k}+\Gamma_{k,\mathrm{D}}}, \tag{2}$$

where $\Gamma_{\mathrm{A,B}}$ is the instantaneous SNR from node A to B given by

$$\Gamma_{\mathrm{A,B}} = \frac{P_{\mathrm{A}}|h_{\mathrm{AB}}|^2}{1+r_{\mathrm{AB}}^2} \triangleq P_{\mathrm{A}}G_{\mathrm{AB}}. \tag{3}$$

The channel gain G_{AB} between node A and node B is defined as $G_{\mathrm{AB}} \triangleq |h_{\mathrm{AB}}|^2/(1+r_{\mathrm{AB}}^2)$. Furthermore, each relay has knowledge of only the probability distributions of $h_{k\mathrm{D}}$ and $r_{k\mathrm{D}}$.

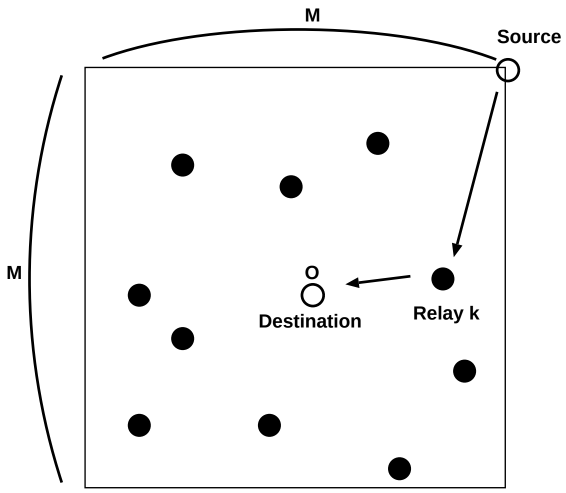

Fig. 1. Network model of proposed system where $d = 2$.

B. Energy Arrival Model

Generally, energy arrivals follow a stochastic process in time, and the quantity of harvested energy differs according to the location of the relays and the time [14]. In this paper, we will assume that the arrival energy is defined by averaging and integrating the arrival energy over time. Thus, the energy harvested by each relay is defined to be a constant value E. In each time slot, each relay recharges its own battery via energy harvesting. Using the harvesting efficiency $\rho > 0$ and transmitting power of the source, we define the energy harvested by the relay in a duration corresponding to two phases as

$$E \triangleq \rho P_{\mathrm{S}} \times T_{\mathrm{S}}, \tag{4}$$

where T_{S} is the length of the time slot, and without loss of generality, we can assume $T_{\mathrm{S}} = 1$.

Due to hardware limitations, all of the relays cannot be simultaneously harvesting energy and forwarding information [15]. The selected relay assists the transmission from the source while consuming all of the energy stored in its own battery; meanwhile, the nonselected relays are harvesting energy. Thus, the relay's transmission power P_k becomes

$$P_k \triangleq N_k \times \frac{E}{T_{\mathrm{S}}} = N_k \times \rho P_{\mathrm{S}}, \tag{5}$$

where N_k is the number of times the battery has been recharged with EH.

In order to make the analysis tractable, the capacity of the battery is assumed to be infinite [16], and the residual battery of each relay is set to zero at the beginning of the transmission, since each relay only utilizes the energy harvested from the environment in order to forward the information.

III. EH Relay Selection Using CDI

A. Method of EH Relay Selection

In this section, we describe the use of CDI for selecting the EH relay. As with the original opportunistic relaying proposed in [4], our approach intends to minimize the outage probability. Unlike the original one, however, in our approach, only the corresponding CDI of the channel gain and the distance between the destination and itself are available to each relay. Therefore, the outage probability of the proposed system must be analyzed in order to determine the relay selection criterion.

An outage event occurs when the channel capacity defined by the instantaneous SNR is less than a given target rate R bits per channel use (bpcu). From (2), once relay k has been selected, the outage probability can be written as

$$P^k_{\mathrm{out}} = \Pr\left[\frac{\log_2(1+\Gamma^k_{\mathrm{AF}})}{2} < R\right] = \Pr\left[\Gamma^k_{\mathrm{AF}} < 2^{2R}-1\right]. \tag{6}$$

Since each relay has only knowledge of probability distributions of $h_{k\mathrm{D}}$ and $r_{k\mathrm{D}}$, (6) becomes

$$\begin{aligned} P^k_{\mathrm{out}} &= \begin{cases} \Pr\left[\Gamma_{k,\mathrm{D}} < \dfrac{a(1+\Gamma_{\mathrm{S},k})}{\Gamma_{\mathrm{S},k}-a}\right], & (\Gamma_{\mathrm{S},k} > a), \\ \Pr\left[\Gamma_{k,\mathrm{D}} < \dfrac{a(1+\Gamma_{\mathrm{S},k})}{a-\Gamma_{\mathrm{S},k}}\right], & (\Gamma_{\mathrm{S},k} < a), \end{cases} \\ &= \begin{cases} \Pr\left[G_{k\mathrm{D}} < \dfrac{a(1+\Gamma_{\mathrm{S},k})}{\Gamma_{\mathrm{S},k}-a}\dfrac{1}{P_k}\right], & (\Gamma_{\mathrm{S},k} > a), \\ \Pr\left[G_{k\mathrm{D}} < \dfrac{a(1+\Gamma_{\mathrm{S},k})}{a-\Gamma_{\mathrm{S},k}}\dfrac{1}{P_k}\right], & (\Gamma_{\mathrm{S},k} < a), \end{cases} \end{aligned} \tag{7}$$

where $a = 2^{2R} - 1$.

Hereinafter, we will assume that $\Gamma_{\mathrm{S},k} > a$. Also, when $\Gamma_{\mathrm{S},k} < a$, the sign is reversed in the final equation.

We note that $|h_{k\mathrm{D}}|^2$ and $r^2_{k\mathrm{D}}$ follow an exponential probability distribution and a generalized gamma distribution, respectively, as follows [12]:

$$|h_{k\mathrm{D}}|^2 \sim f(x) \triangleq e^{-x}, \tag{8}$$

$$1 + r^2_{k\mathrm{D}} \sim g(x;k,\lambda) \triangleq \tfrac{(\pi\lambda)^k (x-1)^{k-1}}{\Gamma(k)} e^{-\pi\lambda(x-1)}, \tag{9}$$

where x and y denote random variables, and $\Gamma(\cdot)$ denotes the gamma function.

Since (8) and (9) are independently distributed random variables, the distribution of $G_{k\mathrm{D}}$ can be written as [17]

$$\begin{aligned} p(x;k,\lambda) &= \int_0^\infty |y| f(xy) g(y;k,\lambda) dy \qquad (10) \\ &= \frac{(\pi\lambda)^k e^{\pi\lambda}}{\Gamma(k)} \int_0^\infty y(y-1)^{k-1} e^{-y(x+\pi\lambda)} dy = \frac{e^{-x}(bx+bk+1)}{(bx+1)^{k+1}}. \end{aligned}$$

where $b = \frac{1}{\pi\lambda}$.

The values of $\Gamma_{\mathrm{S},k}$ and P_k are known at each relay. The outage probability needed for the relay selection can be obtained by integrating (10) over the region defined by right-hand side of the event in (7). However, this integration does not admit a closed-form. Hence, we shall further simplify the distribution to obtain the closed-form expression of the outage probability that will take place in the sequel.

Then, (10) can be rewritten as

$$p(x;k,\lambda) = \frac{e^{-x}}{(bx+1)^k} + \frac{bke^{-x}}{(bx+1)^{k+1}}. \tag{11}$$

If k is large, the second term of (11) obviously tends to zero. If k is small, it is reasonable to assume $x \gg k$ so that $(bx + bk + 1) \approx (bx + 1)$. Consequently, we have

$$p(x; k, \lambda) \approx \frac{e^{-x}}{(bx+1)^k}. \tag{12}$$

The first two derivatives of (12) are respectively given by

$$p^{(1)}(x; k, \lambda) = \frac{-e^{-x}(bx+bk+1)}{(bx+1)^{k+1}} \approx \frac{-e^{-x}}{(bx+1)^k}, \tag{13}$$

$$p^{(2)}(x; k, \lambda) \approx \frac{e^{-x}}{(bx+1)^k}. \tag{14}$$

Thus, the Taylor expansion of (10) becomes

$$p(x; k, \lambda) \approx 1 - x + \frac{x^2}{2!} + \cdots = \sum_{n=0}^{\infty} \frac{(-x)^n}{n!} = e^{-x}, \tag{15}$$

which reveals that the distribution given by (10) can be approximated by an exponential distribution.

Hence, (7) can be approximately calculated by [18]

$$P_{\text{out}}^k \approx 1 - \exp\left[-\frac{1}{\mathbb{E}\left[G_{k\text{D}}\right]} \frac{a(1+\Gamma_{\text{S},k})}{\Gamma_{\text{S},k} - a} \frac{1}{P_k}\right], \tag{16}$$

where $\mathbb{E}[z]$ denotes the mean of a random variable z.

Since $G_{k\text{D}}$ is the ratio of $|h_{k\text{D}}|^2$ and $(1+r_{k\text{D}}^2)$, $\mathbb{E}\left[G_{k\text{D}}\right]$ can be derived with the Taylor expansion [19]. Let us temporarily consider the function $f(N, D) \triangleq N/D$. Then, the Taylor expansion of $f(N, D)$ around the point (μ_N, μ_D) is given by

$$\begin{aligned} f(N, D) = &\frac{\mu_N}{\mu_D} + (N - \mu_N) \left.\frac{\partial f(N, D)}{\partial N}\right|_{(\mu_N, \mu_D)} \\ &+ (D - \mu_D) \left.\frac{\partial f(N, D)}{\partial D}\right|_{(\mu_N, \mu_D)} \\ &+ \frac{1}{2}(N - \mu_N)^2 \left.\frac{\partial^2 f(N, D)}{\partial N^2}\right|_{(\mu_N, \mu_D)} \\ &+ \frac{1}{2}(D - \mu_D)^2 \left.\frac{\partial^2 f(N, D)}{\partial D^2}\right|_{(\mu_N, \mu_D)} \\ &+ (N - \mu_N)(D - \mu_D) \left.\frac{\partial^2 f(N, D)}{\partial N \partial D}\right|_{(\mu_N, \mu_D)} + \cdots \end{aligned} \tag{17}$$

An approximation of the mean of $f(N, D)$ can be obtained using the second-degree truncation of the above Taylor series:

$$\mathbb{E}[f(N, D)] \approx \frac{\mu_N}{\mu_D}\left(1 + \frac{\sigma_D^2}{\mu_D^2}\right) - \frac{\text{Cov}(N, D)}{\mu_D^2}, \tag{18}$$

where σ_D^2 denotes the variance of D, and $\text{Cov}(N, D)$ is the covariance of N and D.

The variables in (18) have the following values:

$$\begin{cases} \mu_N = 1, \\ \mu_D = 1 + \dfrac{k}{\pi\lambda}, \\ \sigma_D^2 = \dfrac{k}{(\pi\lambda)^2}, \\ \text{Cov}(N, D) = 0. \end{cases} \tag{19}$$

The derivation of each of these variables is shown in Appendix A.

From (19), $\mathbb{E}\left[G_{k\text{D}}\right]$ is given by

$$\mathbb{E}[G_{k\text{D}}] \approx \frac{[(\pi\lambda + k)^2 + k]\pi\lambda}{(\pi\lambda + k)^3}. \tag{20}$$

Hence, the outage probability when relay k is selected is

$$P_{\text{out}}^k \approx 1 - \exp\left[-\frac{a(1+\Gamma_{\text{S},k})}{P_k(\Gamma_{\text{S},k} - a)} \frac{(\pi\lambda + k)^3}{[(\pi\lambda + k)^2 + k]\pi\lambda}\right]. \tag{21}$$

Using (21), each relay calculates its own outage probability and sets its timer. When the timer becomes zero, the relay starts to forward the information via VG-AF. The other relays can recognize this transmission and return to the EH mode. For this example, the timer function is defined as

$$T(P_{\text{out}}^k) = \eta \times P_{\text{out}}^k, \tag{22}$$

where $\eta > 0$ is an arbitrary scaling factor.

B. Derivation of the Overall Outage Probability

In this section, the overall outage probability of this system will be theoretically derived. Using S_k, the probability that the k-th relay is selected, the desired outage probability can be written as follows:

$$P_{\text{out}} = \sum_{k=1}^{K} S_k \mathbb{E}[P_{\text{out}}^k], \tag{23}$$

where K denotes the number of relays, the mean of K is given by $\mathbb{E}[K] = \lambda M^2$, and $\mathbb{E}[P_{\text{out}}^k]$ is the mean of P_{out}^k, given by

$$\begin{aligned} \mathbb{E}[P_{\text{out}}^k] &= 1 - \exp\left[-\frac{a(1+\mathbb{E}[\Gamma_{\text{S},k}])}{\mathbb{E}[P_k](\mathbb{E}[\Gamma_{\text{S},k}] - a)} \frac{1}{\mathbb{E}[G_{k\text{D}}]}\right] \\ &= 1 - \exp\left[-\frac{a(1+P_{\text{S}}\mathbb{E}[G_{\text{S}k}])}{\mathbb{E}[P_k](P_{\text{S}}\mathbb{E}[G_{\text{S}k}] - a)} \frac{1}{\mathbb{E}[G_{k\text{D}}]}\right]. \end{aligned} \tag{24}$$

To calculate the above equation, we must analytically derive the probability that the relay k is selected. Let us consider the case where P_{out}^j is less than P_{out}^i for an arbitrary set (i, j) and $i \neq j$. The probability of this event is

$$\begin{aligned} &\Pr[P_{\text{out}}^i > P_{\text{out}}^j] = \\ &= \Pr\left[\frac{1}{\mathbb{E}[G_{i\text{D}}]} \frac{a(1+\Gamma_{\text{S},i})}{\Gamma_{\text{S},i} - a} \frac{1}{P_i} > \frac{1}{\mathbb{E}[G_{j\text{D}}]} \frac{a(1+\Gamma_{\text{S},j})}{\Gamma_{\text{S},j} - a} \frac{1}{P_j}\right] \\ &= \Pr\left[P_j > \frac{\mathbb{E}[G_{i\text{D}}]}{\mathbb{E}[G_{j\text{D}}]} \frac{\Gamma_{\text{S},i} - a}{1+\Gamma_{\text{S},i}} \frac{1+\Gamma_{\text{S},j}}{\Gamma_{\text{S},j} - a} P_i\right]. \end{aligned} \tag{25}$$

Assuming $\Gamma_{\text{S},k} \to \infty$ and that (5) holds, the above equation can be approximated by

$$\begin{aligned} \Pr[P_{\text{out}}^i > P_{\text{out}}^j] &\approx \Pr\left[P_j > \frac{\mathbb{E}[G_{i\text{D}}]}{\mathbb{E}[G_{j\text{D}}]} P_i\right] \\ &= \Pr\left[N_j > \frac{\mathbb{E}[G_{i\text{D}}]}{\mathbb{E}[G_{j\text{D}}]} N_i\right]. \end{aligned} \tag{26}$$

Equation (26) implies that the number of times EH is required for relay j is greater than the number of times it is required for relay i, and thus the outage probability at relay j is smaller than that at relay i. Table I shows the number of times that EH is required at relay j, $j = 1, 2, \cdots 10$, when $N_i = N_1 = 1$.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2016.2586581, IEEE Internet of Things Journal

TABLE I
FREQUENCY OF EH OPERATIONS

j	N_j	j	N_j
1	1.0000	6	2.1794
2	1.2214	7	2.4263
3	1.4537	8	2.6746
4	1.6921	9	2.9240
5	1.9344	10	3.1742

From these results, the probability that a given relay k is selected can be approximated as follows:

$$S_k \approx \Pr[\text{Relay } k \text{ is selected}] \approx \frac{1}{N_k}\left(\sum_{l=1}^{K}\frac{1}{N_l}\right)^{-1}. \quad (27)$$

Since only one relay is selected at a time, and that relay consumes all of the residual battery at each time slot while the other relays are recharging their own battery, $\mathbb{E}[P_k]$ can be derived by using the relay selection probability, as follows:

$$\mathbb{E}[P_k] = \frac{(\text{Energy harvested in two steps})}{(\text{Relay selection probability})} = \frac{\rho P_{\mathrm{S}}}{S_k}. \quad (28)$$

From (23), (24), (27), and (28), we can finally obtain the overall outage probability as

$$P_{\mathrm{out}} \approx \sum_{k=1}^{K} S_k \left(1 - \exp\left[-\frac{S_k}{\rho P_{\mathrm{S}}}\frac{(\pi\lambda+k)^3}{\{(\pi\lambda+k)^2+k\}\pi\lambda} \times \frac{a(r_{\mathrm{SD}}^2 - 2r_{\mathrm{SD}}\frac{\Gamma(k+\frac{1}{2})}{\sqrt{\pi\lambda}\Gamma(k)} + \frac{k}{\pi\lambda} + P_{\mathrm{S}})}{P_{\mathrm{S}} - a(r_{\mathrm{SD}}^2 - 2r_{\mathrm{SD}}\frac{\Gamma(k+\frac{1}{2})}{\sqrt{\pi\lambda}\Gamma(k)} + \frac{k}{\pi\lambda})}\right]\right), \quad (29)$$

where the deviation of $\mathbb{E}[G_{\mathrm{S}k}]$ is shown in Appendix B.

C. Achievable Diversity Order and Coding Gain

We will now further investigate the diversity order and the coding gain of the proposed opportunistic relaying method to determine the performance of the system. The outage probability of the system given by (29) can be approximated by (30) at the top of the next page, which is obtained by a Taylor expansion. Therefore, (30) becomes the equation for P_{S}^{-1} when the SNR is high. Thus, this system has an achievable diversity order of one. We also note that it is obvious that as ρ increases, P_{out}^k decreases.

IV. NUMERICAL RESULTS

In this section, we present numerical results of the proposed relay selection using CDI, which confirm the derivation of the outage probability of the proposed relay selection method. We also evaluate the advantages of the proposed system via computer simulations. The parameters used in simulations are listed in Table II. We assume that $M = 10\,\mathrm{m}$, $\lambda = 1.0$, and $R = 1.0\,\mathrm{bpcu}$ as typical values. We also assume the harvesting efficiency $\rho = 1.0$ and 0.1 as examples. Note that, when $\rho = 1.0$, only one EH time is required for relays to achieve the same transmission power as the source, i.e., high-efficiency EH. Meanwhile, the case with $\rho = 0.1$ corresponds to low-efficiency EH since more EH times are needed to achieve the same transmission power as the source.

For comparison, we also present results of direct transmission, EH relay selection using the instantaneous channel gain [4], one using the mean of the channel gain [8], and one using both the CDI of the small-scale fading and the mean of the channel gain [9]. Although those conventional works did not consider EH, we here simply assume that selections are performed considering the relay's transmission power constraint given by (5) for a comparison purpose.

A. Outage Performance Comparisons

Figure 2 shows the outage probability of the EH relay selection methods when ρ is assumed to be 1.0. Note that the horizontal axis denotes the SNR between the source and the destination P_{S}/N_0.

As can be seen in Fig. 2, the outage probability of the proposed system, derived as (29), shows good agreement with that obtained by Monte Carlo simulations. Moreover, as compared with the direct transmission, the proposed relay selection remarkably improves the outage probability performance. For example, the gain is about 20.0 dB at 10^{-2}. Similarly, comparing our proposed method with the EH relay selection using the channel mean, the additional gain is about 10.0 dB at 10^{-3}. Compared with the conventional relay selection using CDI, the outage probability performance of the proposed one is identical while, in our relay selection, the relays only use the rough location information, i.e., ordinal number k, instead of the mean (or the actual distance). The EH relay selection using the instantaneous CSI achieves the best performance. This is because the best relay in terms of the end-to-end SNR is always selected with the aid of the CSI, while some other relay might be selected when only statistical information such as CDI is used. For a similar reason, the diversity order of the proposed system is one, while that of the system with the instantaneous CSI is equal to the number of relays. However, in the proposed relay selection, each relay has to know only the ordinal number k from the destination, and this significantly reduces the overhead for transmissions and the load for relays, which is suitable for IoT devices.

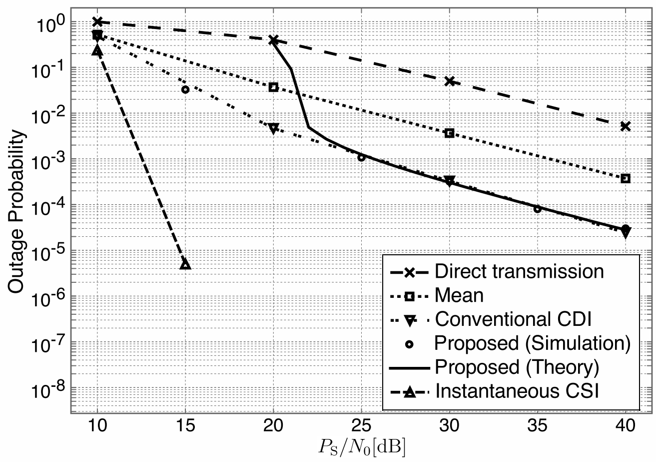

Fig. 2. Outage probability performance of direct transmission, EH relay selection using the instantaneous channel gain, EH relay selection using the channel mean, EH relay selection using the conventional CDI, and the proposed EH relay selection; $\rho = 1.0$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2016.2586581, IEEE Internet of Things Journal

$$
\begin{aligned}
P_{\mathrm{out}} \approx \sum_{k=1}^{K} S_k & \left[\left(\frac{S_k}{\rho P_{\mathrm{S}}} \frac{(\pi\lambda+k)^3}{\{(\pi\lambda+k)^2+k\}\pi\lambda} \frac{a(r_{\mathrm{SD}}^2 - 2r_{\mathrm{SD}}\frac{\Gamma(k+\frac{1}{2})}{\sqrt{\pi\lambda}\Gamma(k)} + \frac{k}{\pi\lambda} + P_{\mathrm{S}})}{P_{\mathrm{S}} - a(r_{\mathrm{SD}}^2 - 2r_{\mathrm{SD}}\frac{\Gamma(k+\frac{1}{2})}{\sqrt{\pi\lambda}\Gamma(k)} + \frac{k}{\pi\lambda})} \right) \right. \\
& \left. - \frac{1}{2} \left(\frac{S_k}{\rho P_{\mathrm{S}}} \frac{(\pi\lambda+k)^3}{\{(\pi\lambda+k)^2+k\}\pi\lambda} \frac{a(r_{\mathrm{SD}}^2 - 2r_{\mathrm{SD}}\frac{\Gamma(k+\frac{1}{2})}{\sqrt{\pi\lambda}\Gamma(k)} + \frac{k}{\pi\lambda} + P_{\mathrm{S}})}{P_{\mathrm{S}} - a(r_{\mathrm{SD}}^2 - 2r_{\mathrm{SD}}\frac{\Gamma(k+\frac{1}{2})}{\sqrt{\pi\lambda}\Gamma(k)} + \frac{k}{\pi\lambda})} \right)^2 + \cdots \right] \\
\approx \sum_{k=1}^{K} S_k & \left[\frac{S_k}{\rho P_{\mathrm{S}}} \frac{a(\pi\lambda+k)^3}{\{(\pi\lambda+k)^2+k\}\pi\lambda} \right] \quad \text{s.t.} \quad \left(\frac{P_{\mathrm{S}}}{N_0} \to \infty \right). \qquad (30)
\end{aligned}
$$

Figure 3 depicts the outage probabilities with $\rho = 0.1$, and it can be seen that all performances except that of direct transmission are degraded due to less energy being harvested. Since, once ρ becomes small, the outage probability of the proposed selection P_{out} becomes large as discussed in Section III-C. Furthermore, the performance improvement compared with the selection with the mean becomes smaller than the case with $\rho = 1.0$.

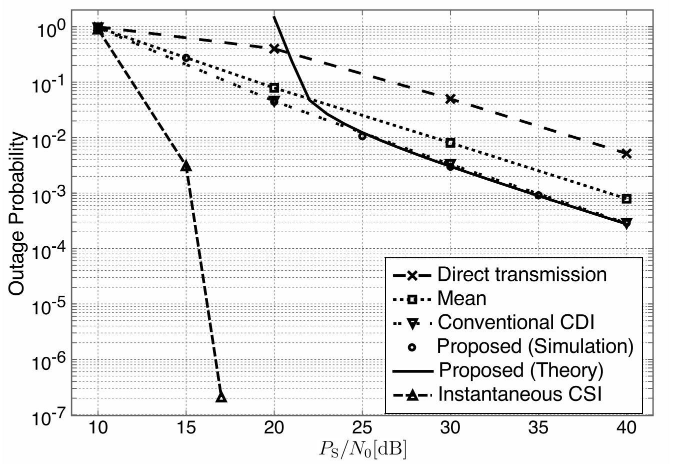

Fig. 3. Outage probability performance of direct transmission, EH relay selection using the instantaneous channel gain, EH relay selection using the channel mean, EH relay selection using the conventional CDI, and the proposed EH relay selection; $\rho = 0.1$

TABLE II
SIMULATION PARAMETERS

M [m]	10.0
λ	1.0
ρ	1.0, 0.1
R [bpcu]	1.0
N_0	1.0

B. Relay Selection Probability

The relay selection probability, that is, the probability that the relay k is selected, is shown in Fig. 4. We assumed that $P_{\mathrm{S}}/N_0 = 40\,\mathrm{dB}$ and $\rho = 1.0$; the vertical axis indicates the relay selection probability, and the horizontal axis indicates the relay index k.

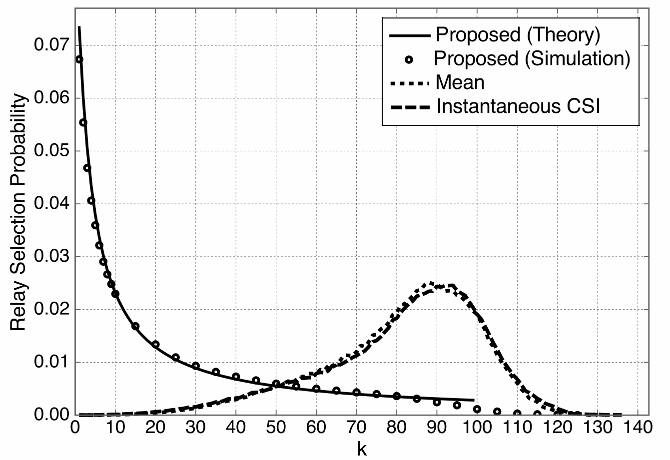

Fig. 4. Relay selection probabilities for the direct EH relay selection using the instantaneous channel gain, EH relay selection using the channel mean, EH relay selection using the conventional CDI, and the proposed EH relay selection; $P_{\mathrm{S}}/N_0 = 40$[dB]$, \rho = 1.0$

From the figure, we can see that when k is small, the relay selection probability of the proposed system is in good agreement with that obtained by Monte Carlo simulations; however, when k is large, the results diverge. The probability mass function (pmf) of the conventional EH relay selection using the mean is almost same as that of one using the instantaneous CSI. However, the pmf of the proposed approach has a completely different shape from those. Meanwhile, our proposed selection exhibits the lower outage probability than one using the mean. This gain comes from the use of statistical information of channels, and it reduces the erroneous selection of relays, compared with the mean-based selection.

C. Effect of Errors in Estimating the Distance

As discussed in the previous section, our proposed system shows the same outage probability as the conventional EH relay selection using the CDI. In this section, we will show that, compared with the conventional relay selection using the CDI proposed in [9], our selection scheme is more robust against errors in estimating the distance between the nodes. Consider a case in which each relay is equipped with a global positioning system (GPS) and is thus able to transmit its own coordinates to a common destination.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2016.2586581, IEEE Internet of Things Journal

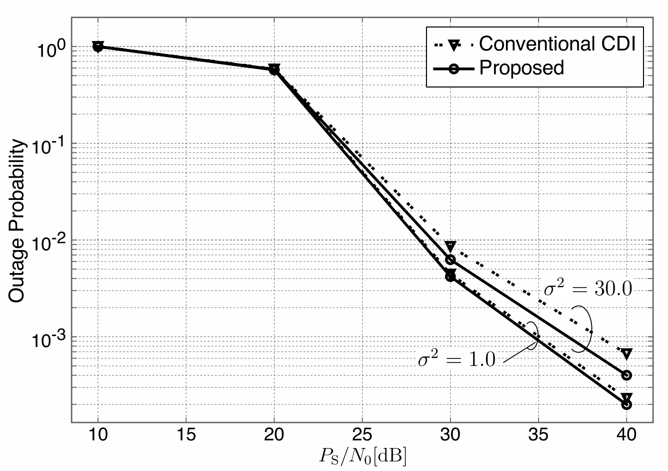

Fig. 5. Outage probability performance of EH relay selection using the conventional CDI and the proposed method; $M = 20.0, \lambda = 0.1, \sigma^2 = 1.0, 30.0$

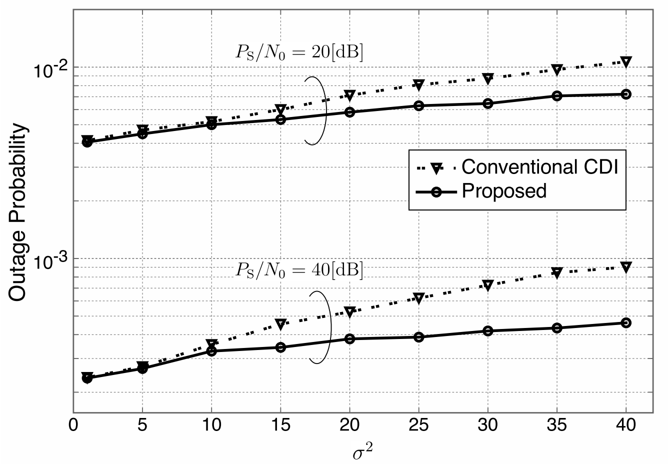

Fig. 6. Outage probability performance of EH relay selection using the conventional CDI and the proposed method; $M = 20.0, \lambda = 0.1, P_\mathrm{S}/N_0 = 20\mathrm{dB}, 40\mathrm{dB}$

From the coordinates of each relay, the destination can calculate the distance between that relay and itself; this information is then returned to the relay. If we further assume (w.l.g.) that the destination is located at the origin, and that it knows its own coordinates, the estimated distance between the k-th relay and the destination, $\hat{r}_{k\mathrm{D}}$, becomes

$$\hat{r}_{k\mathrm{D}} = \sqrt{(x_k + n_{x_k})^2 + (y_k + n_{y_k})^2}, \tag{31}$$

where (x_k, y_k) are the true coordinates of the kth relay, and n_{x_k} and n_{y_k} denote their estimation errors, assumed to be i.i.d zero-mean Gaussian with variance $\sigma^2/2$.

The selection method using the CDI uses $\hat{r}_{k\mathrm{D}}$ directly for relay selection. In the proposed scheme, the relay selection uses instead the order of the relay from the destination, k, which is derived from $\hat{r}_{k\mathrm{D}}$.

Figures 5 and 6 show the outage probabilities for the proposed scheme and for the conventional relay selection using the CDI; here, $\sigma^2 = 1.0, 30.0$, $M = 20.0\,\mathrm{m}$, $\lambda = 0.1$, and the other parameters are the same as in the previous section. Note that, when $\sigma^2 = 30.0$, a probability that the error in the estimated distance is less than about 15.5 m becomes 95.4%, while the localization accuracy of a typical off-the-shelf GPS (Garmin 18x) is 15 m RMS 95% [20]

From Fig. 5, we can see that there is no significant difference between the two outage probability performances when $\sigma^2 = 1.0$. However, when $\sigma^2 = 30.0$, the performance gap between the two outage probabilities is about 2 dB in the high P_S/N_0 region. It can also be seen in Fig. 6 that as both P_S/N_0 and σ^2 increase, the gap becomes large. Thus, our proposed scheme is robust against distance estimation error compared with the conventional selection using CDI, since our proposed scheme uses k, which is derived from $\hat{r}_{k\mathrm{D}}$, and the conventional method that uses the CDI uses $\hat{r}_{k\mathrm{D}}$ directly.

V. Conclusion

In this paper, we have proposed a new EH relay selection scheme that is based on a residual battery of relays and the CDI of both small-scale and large-scale fading. Furthermore, we derived a simple selection rule and a closed solution for the end-to-end outage probability. As discussed in Section III-C, this system can achieve a diversity order of one, even though each relay can exploit knowledge of the entire distribution between the destination and itself. These results show that if we wish to enhance the spatial diversity, it is important to obtain the instantaneous CSI. On the other hand, considering the practical limitations of IoT devices, our approach is still an attractive option that realizes low-power consumption and highly reliable communications.

Appendix A

Below, we derive the values of each variable used in (18). Let us consider

$$\frac{N}{D} = G_{k\mathrm{D}} = \frac{|h_{k\mathrm{D}}|^2}{1 + r_{k\mathrm{D}}^2}. \tag{32}$$

As shown in (8), $N = |h_{k\mathrm{D}}|^2$ follows an exponential distribution with a single mean:

$$\mu_N = \mathbb{E}\left[|h_{k\mathrm{D}}|^2\right] = 1, \tag{33}$$

and, as shown in (9), $D = 1 + r_{k\mathrm{D}}^2$ follows a generalized gamma distribution:

$$\begin{aligned}\mu_D &= 1 + \mathbb{E}\left[r_{k\mathrm{D}}^2\right] \\ &= 1 + \int_0^\infty x \frac{(\pi\lambda)^k x^{k-1}}{\Gamma(k)} e^{-\pi\lambda x} \\ &= 1 + \frac{(\pi\lambda)^k}{\Gamma(k)} \int_0^\infty x^k e^{-\pi\lambda x} dx \\ &= 1 + \frac{k}{\pi\lambda}.\end{aligned} \tag{34}$$

Then, the variance of D is

$$\begin{aligned}\sigma_D^2 &= \mathbb{E}[D^2] - \mathbb{E}[D]^2 \\ &= 1 + 2\mathbb{E}[r_{k\mathrm{D}}^2] + \mathbb{E}[r_{k\mathrm{D}}^4] - \mu_D^2,\end{aligned} \tag{35}$$

where $r_{k\mathrm{D}}^4$ is derived by substituting $y = x^2$ for $r_{k\mathrm{D}}^2$:

$$r_{k\mathrm{D}}^4 \sim \frac{(\pi\lambda)^k y^{k-1}}{\Gamma(k)} \frac{\sqrt{y}^k}{2y} e^{-\pi\lambda\sqrt{y}}, \tag{36}$$

$$\begin{aligned}\mathbb{E}[r_{k\mathrm{D}}^4] &= \frac{(\pi\lambda)^k}{2\Gamma(k)} \int_0^\infty \sqrt{y}^k \exp(-\pi\lambda\sqrt{y}) dy \\ &= \frac{k(k+1)}{(\pi\lambda)^2}.\end{aligned} \tag{37}$$

Therefore,

$$\sigma_D^2 = \frac{k}{(\pi\lambda)^2}. \tag{38}$$

Finally, $\mathrm{Cov}(N, D)$ is derived as follows:

$$\begin{aligned}\mathrm{Cov}(N, D) &= \mathbb{E}[N \times D] - \mu_N \mu_D \\ &= \mathbb{E}[|h_{k\mathrm{D}}|^2] + \mathbb{E}[|h_{k\mathrm{D}}|^2]\mathbb{E}[r_{k\mathrm{D}}^2] - \mu_D = 0.\end{aligned} \tag{39}$$

APPENDIX B

Here we describe the derivation of $\mathbb{E}[G_{\mathrm{S}k}]$, given by

$$\mathbb{E}[G_{\mathrm{S}k}] = \mathbb{E}\left[\frac{|h_{\mathrm{S}k}|^2}{1 + r_{\mathrm{S}k}^2}\right]. \tag{40}$$

To begin with, the distribution of $1 + r_{\mathrm{S}k}^2$ should be investigated. It is remarkably difficult to derive the distribution, analytically. Our interest here is to obtain the closed form solution of $\mathbb{E}[G_{\mathrm{S}k}]$ to see the achievable diversity order of the proposed approach at the sufficiently high SNR region. Thus, we focus only on a rough approximation of $\mathbb{E}[G_{\mathrm{S}k}]$.

If the ordinal number from the source denoted by k' is given, $(1 + r_{\mathrm{S}k'}^2)$ also follows a generalized gamma distribution by moving the origin to the source [12]. Conditioned on k, the probability that the k'th closest node to the source becomes the kth closest node to the destination is calculated by taking the average of the generalized gamma distribution with k' over the conditional probability $\Pr(k'|k)$. Considering that the shape of generalized gamma distributions with different but close k' is almost same, and the set of possible k' should be composed of close numbers due to the shape of the distribution of $r_{k\mathrm{D}}$, the resulting distribution can be approximated by a generalized gamma distribution. Therefore, from (17), an approximation of the mean of $G_{\mathrm{S}k}$ can be roughly obtained using the first-degree truncation of the Taylor series:

$$\mathbb{E}[G_{\mathrm{S}k}] \approx \frac{\mathbb{E}[|h_{\mathrm{S}k}|^2]}{1 + \mathbb{E}[r_{\mathrm{S}k}^2]}. \tag{41}$$

Like $|h_{k\mathrm{D}}|^2$, $|h_{\mathrm{S}k}|^2$ follows an exponential distribution with a single mean:

$$\mathbb{E}[|h_{\mathrm{S}k}|^2] = 1. \tag{42}$$

On the other hand, from the law of cosines,

$$\begin{aligned} r_{\mathrm{SD}}^2 &= r_{\mathrm{S}k}^2 + r_{k\mathrm{D}}^2 - 2r_{\mathrm{S}k}r_{k\mathrm{D}}\cos\theta, \\ r_{\mathrm{SD}}^2 &\le (r_{\mathrm{S}k}^2 + r_{k\mathrm{D}}^2)^2, \\ r_{\mathrm{S}k} &\le r_{\mathrm{SD}} - r_{k\mathrm{D}}, \end{aligned}$$

where θ denotes the angle between $r_{\mathrm{S}k}$ and $r_{k\mathrm{D}}$. $\mathbb{E}[r_{\mathrm{S}k}^2]$ becomes

$$\begin{aligned}\mathbb{E}[r_{\mathrm{S}k}^2] &\approx \mathbb{E}[r_{\mathrm{SD}}^2 - 2r_{\mathrm{SD}}r_{k\mathrm{D}} + r_{k\mathrm{D}}^2] \\ &= \mathbb{E}[r_{\mathrm{SD}}^2] - 2\mathbb{E}[r_{\mathrm{SD}}]\mathbb{E}[r_{k\mathrm{D}}] + \mathbb{E}[r_{k\mathrm{D}}^2],\end{aligned} \tag{43}$$

where $r_{k\mathrm{D}}$ is derived by substituting $y = \sqrt{x}$ for $r_{k\mathrm{D}}$:

$$r_{k\mathrm{D}} \sim \frac{2(\pi\lambda y^2)^k}{y\Gamma(k)} e^{-\pi\lambda y^2}, \tag{44}$$

$$\begin{aligned}\mathbb{E}[r_{k\mathrm{D}}] &= \frac{2(\pi\lambda)^k}{\Gamma(k)} \int_0^\infty \sqrt{y}^{2k} \exp(-\pi\lambda y^2) dy \\ &= \frac{\Gamma(k + \frac{1}{2})}{\sqrt{\pi\lambda}\Gamma(k)}.\end{aligned} \tag{45}$$

Thus, $\mathbb{E}[G_{\mathrm{S}k}]$ can be written as follows:

$$\mathbb{E}[G_{\mathrm{S}k}] \approx \frac{1}{1 + r_{\mathrm{SD}}^2 - 2r_{\mathrm{SD}}\frac{\Gamma(k+\frac{1}{2})}{\sqrt{\pi\lambda}\Gamma(k)} + \frac{k}{\pi\lambda}}. \tag{46}$$

REFERENCES

[1] Dave Evans, "The Internet of Things: How the Next Evolution of the Internet Is Changing Everything," April 2011. Available: http://www.cisco.com/web/about/ac79/docs/innov/IoT_IBSG_0411FINAL.pdf

[2] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communication: Performance criterion and code construction," *IEEE Trans. Inform. Theory*, vol. 44, no. 2, pp. 744–765, Mar. 1998.

[3] J. N. Laneman, D. N. C. Tse, and G. W. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," *IEEE Trans. Inform. Theory*, vol. 50, no. 12, pp. 3062–3080, Dec. 2004.

[4] A. Bletsas, H. Shin and M. Z. Win "Cooperative communications with outage-optimal opportunistic relaying," *IEEE Trans. Wireless Commun.*, vol. 6, no. 9, pp.3450–3460, Sep. 2007.

[5] J. A. Paradiso and T. Starner, "Energy scavenging for mobile and wireless electronics," *IEEE Perv. Comput.*, vol. 4, pp. 18–27, Jan-Mar. 2005.

[6] B. Medepally and N. B. Mehta, "Voluntary energy harvesting relays and selection in cooperative wireless networks," *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3543–3553, Sep. 2010.

[7] Y. Luo, J. Zhang, and K. Letaief, "Relay selection for energy harvesting cooperative communication systems," in *IEEE Global Commun. Conf. (GLOBECOM)*, pp. 2514–2519, Dec. 2013.

[8] A. Tukmanov, S. Boussakta, Z. Ding and A. Jamalipour, "On the impact of relay-side channel state information on opportunistic relaying," Proc. *IEEE ICC WCS*, pp.4071–4075, June 2013.

[9] Z. Fang, X. Zhou, X. Bao and Zongxin, "Outage minimized relay selection with partial channel information," Proc. *IEEE ICASSP*, pp.2617–2620, Apr. 2009.

[10] H. Jiang, S. Zhang and W. Zhou, "Outage analysis of distributed scheme on opportunistic relaying with limited CSI", *IEEE Commun. Lett.*, vol. 15, no. 9, pp.935–937, July 2011.

[11] H. Jiang, S. Zhang and W. Zhou "Outage analysis of distributed scheme on opportunistic relaying with limited CSI," *IEEE Commun. Lett.*, vol. 15, no. 9, pp.935–937, Sep. 2011.

[12] M. Haenggi, Stochastic Geometry for Wireless Networks, Cambridge University Press, 1st ed., 2012.

[13] K. Ishibashi and H. Ochiai "Performance analysis of amplify and forward cooperation over peak-power limited channels," *IEEE Trans. Wireless Commun.*, vol. 11, no. 1, pp. 258–265, Jan. 2012.

[14] O. Ozel, K. Tutuncuoglu, J. Yang, S. Ulukus and A. Yener, "Transmission with energy harvesting nodes in fading wireless channels: optimal policies", *IEEE J. Sel. Areas Commun.*, vol. 29, no. 8, pp.1732–1743, Sep. 2011.

[15] R. Zhang and C. K. Ho, "MIMO broadcasting for simultaneous wireless information and power transfer," in Proc. *IEEE Global Commun. Conf. (GLOBECOM)*, Houston, USA, pp. 1–5, Dec. 2011.

[16] A. A. Nasir, X. Zhou, S. Durrani, and R. A. Kennedy, "Throughput and ergodic capacity of wireless energy harvesting based DF relaying network," in Proc. of *IEEE ICC'14*, (Sydney, Australia), pp. 4066–4071, June 2014.

[17] S. Vuppala and G. Abreu, "Secrecy outage in random wireless networks subjected to fading," in Proc. *IEEE Personal Indoor Mobile Radio Communication (PIMRC)*, pp. 441–445, Sep. 2013.

[18] I. Krikidis, S. Timotheou, and S. Sasaki, "RF energy transfer for cooperative networks: Data relaying or energy harvesting?," *IEEE Commun. Lett.*, vol. 16, no. 11, pp. 1772–1775, Nov. 2012.

[19] G. M. P. Kempen and L. J. Vliet, "Mean and variance of ratio estimators used in fluorescence ratio imaging," *Cytometry*, vol. 39, pp. 300–305, Apr. 2000.

[20] GPS 18x Technical Specifications
Available: http://static.garmin.com/pumac/GPS_18x_Tech_Specs.pdf