Performance of Random Routing on Grid-Based Sensor Networks

Dulanjalie C. Dhanapala
Anura P. Jayasumana
Department of Electrical and Computer Engineering,
Colorado State University, Fort Collins, CO 80523, USA

Qi Han
Department of Mathematical and Computer Sciences,
Colorado School of Mines, 1500 Illinois Street, Golden,
CO 80401, USA

***Abstract* - Random routing protocols in sensor networks forward packets to randomly selected neighbors. These packets are agents carrying information about events, or queries seeking such information. We derive the probability of a packet visiting a given node in a given step as well as the rendezvous probability of agents and queries within a specific number of hops at a given node(s) in a 2-D grid-based sensor network. The utility of the model is demonstrated by determining the protocol parameters to optimize performance of rumor routing protocol under different constraints, e.g., to evaluate the number of queries and agents to maximize the probability of rendezvous for a given amount of energy. Monte Carlo simulations are used to validate the model. The closed form exact solution presented, unlike existing models relying on asymptotic behavior, is applicable to small and medium-scale networks as well. An upper bound is provided for the case where the packet is not sent back to its immediate forwarding node. Simulation results indicate that the model is a good approximation even for sparse arrays with 75 % of the nodes. The model can be used to set parameters and optimize performance of several classes of random routing protocols.**

Keywords-sensor networks, Wireless, Routing,Random routing

I. INTRODUCTION

Wireless Sensor Networks (WSNs) are expected to connect a large number of smart sensor and actuator devices, self organize, sense events, exchange data over wireless medium, make collaborative decisions to monitor phenomena, and even interact with the environment. Routing protocols for WSNs can be broadly classified into two categories. First category relies on some type of structural or location information. A structure may be imposed via an initial self-organization phase, e.g., by forming a cluster tree [4] or assigning logical coordinates [7]. Geographical location information may be obtained during deployment, or via a localization scheme. The second category of routing protocols does not rely on geographical or a logical coordinate system for routing, rather they use random routing strategies such as random walks or gossiping [2],[5],[6],[8],[9],[10],[16].

As random routing does not guarantee the delivery or discovery of information of interest, one has to rely on models to tune the protocol parameters and to obtain required level of reliability and power efficiency. Reliability here refers to the probability of an event agent and a query meeting. Rumor routing algorithm [6], for example, uses multiple agents and multiple queries in order to achieve higher reliability. A formal model is necessary to select parameters such as the number of queries, number of agents, and Time-To-Live (TTL), i.e., the maximum number of hops a packet can traverse, for different scenarios to meet the performance goals. With rumor routing, the length of agent packet increases with each hop and directly affects the energy/cost. Therefore a more accurate and flexible model for energy evaluation is required when tuning these parameters. We develop a model that can be used to determine such parameters under different conditions on grid based sensor networks.

There has been many research-works on modeling random routing protocols. The Brownian motion based model in [14] provides asymptotic results for source-only search, source-receiver "sticky" search, and spatially-periodic caching in a rectangular grid, whereas our model provides an **exact** solution for these cases using a much simpler model. A random walk based routing algorithm is proposed in [18], which also evaluates the probability of a packet moving between two nodes in a rectangular grid. However this model over-estimates the probability by neglecting certain path combinations and joint probability. A three-way handshake random walk based protocol is proposed in [1] for un-structured systems for one shot queries. Though no mathematical model is presented, they have studied the performance in a realistic system which is useful in assessing how much theoretical results deviate from the actual results due to real environmental issues that are difficult to model mathematically. Similarly, [11] and [15] propose random walk based routing algorithms for WSNs. The algorithm in [15] achieves load balancing in WSNs but no method is presented to tune the parameters for different scenarios. The data gathering problem of the networks with static nodes and a mobile patrol node is addressed in [11], providing analytical **bounds** for coverage in an unconstrained random walk. Moreover, [3] considers dynamic environments with failure recovery mechanisms and gives an upper bound for the time required to cover only a constant fraction of the network (partial cover time). The analytical model proposed in [12] evaluates the performance in terms of the mean system data gathering delay and the induced spatial distribution of energy consumption. An analysis of impact of the requisite knowledge on the routing efficiency in periodic lattice network of finite square cells is provided in [13]. Even though these and other publications address performance analysis for particular network scenarios, none of them proposes a general-purpose and flexible method of tuning the parameters to

This research is supported in part by NSF Grant CNS-0720889 and CNS-0720875

optimize performance for different topologies in order to improve the WSN lifetime.

In this paper we propose a method of evaluating **exact** probability of a packet visiting a node within given number of hops in random routing using a much simpler model. Using this result we find the probability of a packet visiting a node within a given number of hops. We then derive the probability of an agent meeting a query within a given number of hops.

In Section II, we introduce the analytical model, and then in Section III we present simulation results. Finally Section IV provides conclusions.

II. ANALYTICAL MODEL

An exact model for the propagation of queries and agents is presented in this section. In the rectangular topology as shown in Fig. 1, every node has four equidistance neighbors in its communication range. A packet is routed in a random manner with a set maximum hop count. The network boundary is not considered below, yet can be accommodated easily. Two cases are analyzed. In Case (1), each node forwards the packet to one of its four neighbors with an equal probability. In Case (2) a node sends to its neighbors except the sender with equal probability. If not mentioned specifically, Case (1) is the default case.

A. Probability of a packet visiting (I,J) in the H^{th} hop

Without loss of generality, consider a packet originating from *(0,0)*. Let $P_H(I,J)$ be the probability that it reaches *(I,J)* in the H^{th} hop (Fig. 1),

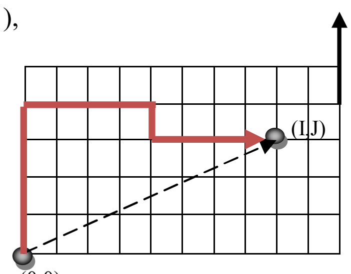

Figure 1: Rectangular grid.

Case (1): As minimum number of hops required is $I+J$ $(=K)$, H-K additional hops are left assuming that $H>K$. For the packet to end up at *(I,J)*, H-K must be even. Let the number of movements in east (E), west (W), north (N) and south (S) directions be e, w, n and s respectively, with $e+w+n+s=H$. To reach *(I,J)*, the packet needs to move I net units to east and J net units to north, thus e-w $=$ I, and n-s=J. So the total numbers of movements in E, W, N and S directions are $I+i$, i, $J+j$, and j respectively. i and j are the additional hops taken in E, W, N and S directions in order to get to node *(I, J)* in H hops.

Using multinomial distribution [19], $P_H(I,J)$ is given by

$$P_H(I,J) \quad = \quad \sum_{i=0}^{(H-K)/2} \frac{H!}{(I+i)!\,i\,!(J+j)!\,j!}\left(\frac{1}{4}\right)^H \quad ; i+j=(H-K)/2 \qquad (1)$$

Using Vandermonde's Convolution [20]

$$P_H(I,J) \quad = \quad \frac{H!\,{}^H C_M}{(M-I)!(M-J)!}\left(\frac{1}{4}\right)^H \quad ; M=\frac{H+I+J}{2} \qquad (2)$$

Case (2): We derive an upper bound for $P_H(I,J)$.

In this case, except for the initial step, the probability of selecting a particular next hop is 1/3. A path consists of an H-tuple, with each element in $\{E,W,N,S\}$. In this pattern, there are H-1 adjacent pairs. Since a packet is not forwarded to the node it came from, none of those pairs can be movements in *(N,S)* , *(S,N)* , *(E,W)* or *(W,E)*. Let the number of movements in *N,S,E* and *W* be *n,s,e* and *w;* then the probability of occurrence of *N,S,E,W* are n/H, s/H, e/H and w/H respectively. Therefore the probability of getting *(N,S)* or *(S,N)* illegal patterns is $2(ns/H^2)$, while the probability of getting *(E,W)* or *(W,E)* is $2(we/H^2)$. Here the probability of having illegal pairs in one pattern is over-estimated due to assumption that all the pairs are considered independent and equally probable. Since the number of movements $n=J+j$, $s=j$, $e=I+i$ and $w=I$, the probability of a path valid for Case 1, having illegal pairs for Case 2 is bounded by,

$$2\left(\frac{(I+i)i}{H^2}+\frac{(J+j)j}{H^2}\right)$$

Therefore upper bound of the probability of not containing an illegal pair in any pattern is

$$\left(1-2\left(\frac{(I+i)i}{H^2}+\frac{(J+j)j}{H^2}\right)\right)^{H-1}$$

Now we can obtain the following bound for this case.

$$P_H(I,J) \le \sum_{i=0}^{(H-K)/2} \frac{H!}{(I+i)!i!(J+j)!j!}\frac{1}{4}\left(\frac{1}{3}\right)^{H-1}\left(1-\frac{2}{H^2}((I+i)i+(J+j)j)\right)^{H-1} \qquad (3)$$

The following analysis uses $P_H(I,J)$, and for Case 1 it is given by (2) and for Case 2 we use the bound of (3).

B. Probability of a packet visiting (I, J) within H hops

$Q_H(I,J)$: Probability of a packet visiting *(I,J)* within H hops

$p_h(I,J)$: Probability of a packet visiting *(I,J)* for the first time in the h^{th} hop. Then,

$$P_H(I,J) \;=\; p_H(I,J)\cdot P_0(0,0)+p_{H-2}(I,J)\cdot P_2(0,0) + p_{H-4}(I,J)\cdot P_4(0,0)+\ldots+p_0(I,J)\cdot P_H(0,0)$$

$$P_H(I,J) \;=\; \sum_{i=0}^{H} p_{H-i}(I,J)\cdot P_i(0,0)$$

where, $P_h(0,0)$ is the probability of a packet revisiting itself within h hops, which can be obtained from

$$P_H(0,0) \;=\; \left(\frac{H!}{(H/2)!^2}\right)^2\left(\frac{1}{4}\right)^H$$

The probability of a packet reaching *(I,J)* within H hops, is given by

$$Q_H(I,J) \;=\; \sum_{h=K}^{H} p_h(I,J) \qquad (4)$$

In order to find $Q_H(I,J)$ we need to find $p_h(I,J)$ $\forall h$

$$p_{K+2m}(I,J) = P_{K+2m}(I,J) + \sum_{i=0}^{m-1} P_{2j}(0,0)\, p_{K+2i}(I,J) \quad ; i+j=m \qquad (5)$$

As can be seen, Equation (5) is a recursive formula, with the initial condition,

$$p_K(I,J) = P_0(0,0)\, p_K(I,J) = P_K(I,J)$$

Therefore Equation (4) can be evaluated.

C. Probability of an agent meeting a query

When a node detects an event it will send one or more agents with a set TTL informing other nodes about the event. If any other node is interested in the event and if it is not aware of the event initially, it will send a query. Schemes such as rumor routing send more than one agent or query, thereby intending to improve performance. The probability of an agent meeting query is zero when the sum of maximum hop counts of the two is smaller than the Manhattan distance between the event and query nodes.

Fig. 3 shows the coordinate relationship with respect to event and query, with the rendezvous point having coordinates *(I,J)* and *(I'J')* with respect to origin of query and event respectively.

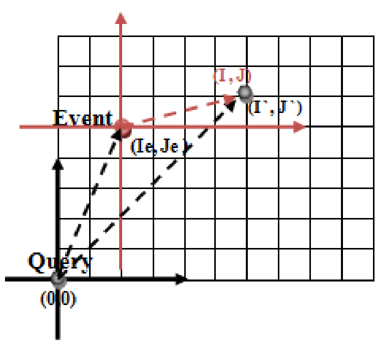

Figure 3: Coordinate relationship for event, query and their meeting point.

Subscript '*e*' denotes event agent and '*q*' denotes query. Then,

I_e, J_e (I_q, J_q) : Event (Query) location
H_e (H_q) : Time to live of agent (query)
h_e (h_q) : Instantaneous hop count 0 *to* $H_e(H_q)$

Without loss of generality let the query origin be *(0,0)*.

We define

$M_{h_e,h_q}(I,J)$ as the probability of the agent NOT meeting query within h_e and h_q at node *(I, J)*

Since in the rumor routing protocol [6], an agent generated travels until its TTL expires, we are interested in $h_e = H_e$

The query continues till it meets the event or TTL expires, i.e., $h_q \le H_q$. Since spreading of agent and query are independent, P [Agent NOT meeting query within h_q hops at *I,J*]

$$= 1 - Q_{H_e}(I,J)\, Q_{h_q}(I',J')$$

A query not meeting an agent at node (I_1, J_1) is independent of the query not meeting agent at (I_2, J_2) within h_q hops.

Therefore, P [Agent NOT meeting query anywhere within h_q hops]

$$M_{H_e,h_q}(I,J) = \prod_{\forall (I+J)\le h_q} \left(1 - Q_{H_e}(I,J)\, Q_{h_q}(I',J')\right) \qquad (6)$$

Therefore, P [Agent meeting query anywhere for the first time within h_q hops]

$$R_{H_e,h_q}(I,J) = 1 - M_{H_e,h_q}(I,J)$$

Also P [Agent meeting query anywhere for the first time at h_q th hop] $= R_{H_e,h_q}(I,J) - R_{H_e,h_q-1}(I,J)$ (7)

Let the probability of at least one of N packets originating at *(0,0)* visiting *(I,J)* in h hops be $Q_h^{(N)}(I,J)$. Since each packet is independent and identical,

$$Q_h^{(N)}(I,J) = 1 - (1 - Q_h(I,J))^N \qquad (8)$$

Advantage of (7) is, we need not to assume the agent path is already setup because we can vary the agent TTL as well after finding $Q_h^{(N)}(I,J)$ using (8) for given number of queries and agents. Thus the rendezvous probability of any of the *Ne* agents and any of the *Nq* queries is given by

$$M_{H_e,h_q}\Big|_{N_e,N_q} = \prod_{\forall (I+J)\le h_q} \left(1 - Q_{H_e}^{(N_e)}(I,J)\, Q_{h_q}^{(N_q)}(I',J')\right) \qquad (9)$$

III. SIMULATION RESULTS

We now verify the model using Monte Carlo simulation, performed using MATLAB 7.5 for a rectangular grid of 2500 nodes. For illustration purposes we use agent and query TTL values *Hq* and *He* as 30 and 170 respectively.

A. Model Verification

The query generating node is assumed to be at the origin. Due to symmetry, we present results only for *(+I,+J)*

Case 1: Exact $Q_H(I,J)$

Fig. 4 shows the probability of a packet originating from *(0,0)* to visiting node *(I,J)* with respect to different *(I,J)*. The analytical results are exaclty the same as the results from the Mote Carlo model of Random Routing in rectangular grid.

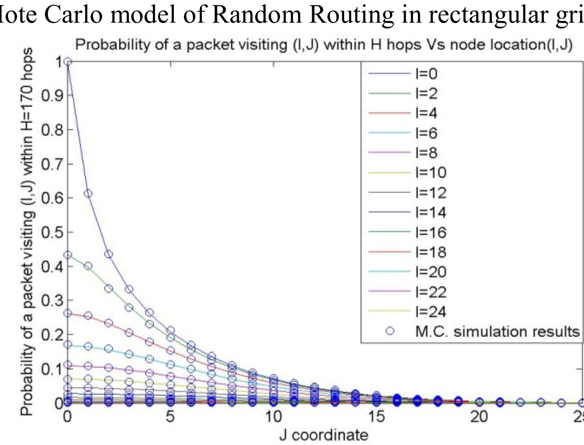

Figure 4: Variation of $Q_H(I,J)$ with respect to node location.

Case 2: Upper bound for $Q_H(I,J)$

Equations (3) and (4) were evaluated under the same network specification. Exact $P_H(I,J)$ and $Q_H(I,J)$ for the case where packet is not passed to the immediate node that it came from were evaluated using a Monte Carlo simulation, to evaluate the tightness of upperbound of (4).

$Q_H(I,J)$ vs J is presented in Fig. 5 for *I=0*, similar tightness has been observed in the bound for other combinations of *I,J* as well.

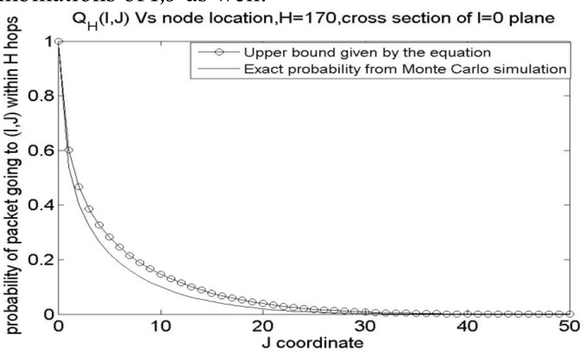

Figure 5: Variation of $Q_H(I,J)$ evaluated using the equation and Monte Carlo simulation in I=0 plane.

Next we consider sparse networks with probability of a node being present at a grid point being *90%, 80%* and *50%.* Probability of a packet visiting *(I,J)* within H hops is shown in Fig. 6. As can be seen, probabilities for 90% and 80% denisities are almost the same as that for 100% nodes availability case. Additional simulation results indicate that probability $Q_H(I,J)$ is more or less the same as 100% node available grid, if the availability of nodes is more than 75%, thus making it possible to apply the model for such arrays as well. In sparse arrays some nodes may have all neighbors while some nodes have less. In fact, the model may be used to evaluate a broader class of networks. For example, a network in which nodes go to sleep mode periodically is similar to a sparse-array.

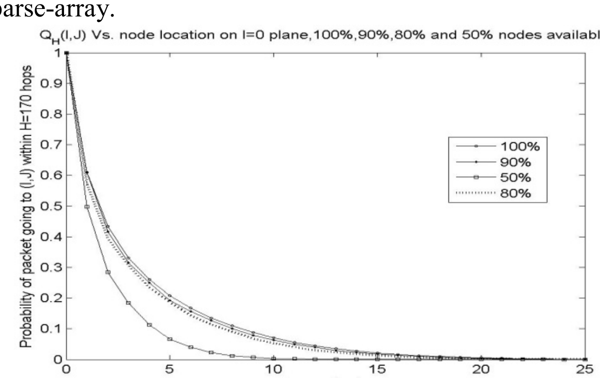

Figure 6: Variation of $Q_H(I,J)$ on I=0 plane when 100%, 90%, 80% and 50% of nodes are present.

B. Applications of the model:

Main importance of the proposed model is its utility to set parameters to achieve optimum performance under required conditions, e.g., for a given reliability, for fixed total energy, or for a given delay bound. In the following, we illustrate how to obtain the optimal number of agents and queries with fixed energy budget.

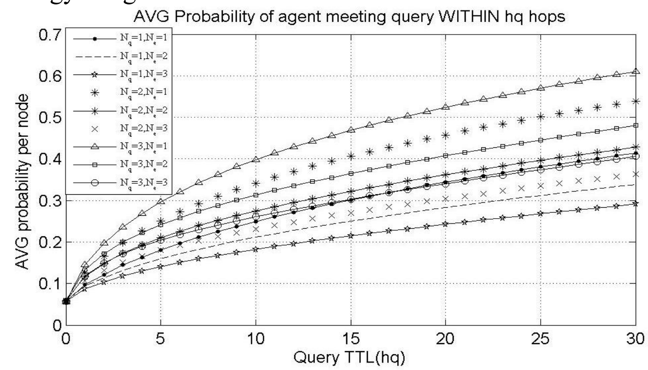

Figure 7:Variation of Reliability with Query TTL when number of queries and agents changes.

The total energy consumption is the sum of energy used for agents and energy used for queries. In this analysis we keep the total energy for all agents fixed.

$$\text{Energy of the agents} = \mathrm{E}_\alpha \times \left(\frac{TTL}{N_e}\right) N_e = \mathrm{E}_\alpha \times H_e$$

where E_α is the energy consumption per packet.

Fig. 7 illustrates the variation of probability of an agent meeting a query within a given number of query hops. When total agent cost remains fixed, the performance degrades as the number of agents increases. This is due to the fact that for one rumor, the variance of the probability of it visiting nodes in the area of interest is lower than that when N_e agents are sent, each with TTL$1/N_e$ thus utilizing the same power. Thus we can conclude that under fixed agent energy, the best performance is given by a **single agent.**

We also investigated the impact of query packet length on the optimal number of agents and queries. In many routing algorithms such as rumor routing [6], the length of the packet increases as the query/agent moves on. In such applications the energy used increases exponentially with TTL. Whereas in many other algorithms, a node keeps information only of its neighbors, and therefore the packet length and energy is fixed. Thus we evaluated these two cases.

With the first case, as the agent propagates the packet length of agent/query remains the same. Since energy spent on agents is fixed, to hold the total energy per event constant, the query energy also should be fixed. Given that the query packet length is a constant, for example, if $N_q=1$ uses 30 hops, using the same energy for $N_q=2$ case a query can go only15 hops, and similarly for $N_q=3$ case for 10 hops. As shown above $N_e=1$ gives the highest reliability, so let $Ne=1$. From Fig. 7, the reliability for $(N_q=1, N_e=1)|_{hq=30}$, $(N_q=2, N_e=1)|_{hq=15}$ and $(N_q=3, N_e=1)|_{hq=10}$ are the same, i.e. 0.4. This holds for all the h_q values as well. So it can be concluded that for a given energy allocation for queries, the reliability that can be achieved is independent of N_q.

The second case is where the agent packet length keeps growing as it propagates from node to node. For example, let us consider initial packet length of 5 words, and at each hop one word will be added to the packet. And let the energy per word be E_w. So the total energy is the sum of an arithmatic seris with difference H_q . Therefore,

Energy for queries $= E_w \times \left(\frac{H_q}{2} (10 + H_q) \right) N_e$

Fig. 8 illustrates the variation of the reliability with number of queries under different total energy values (for query and agent). For this example, the performance is optimum when the number of queries is 4, and the number of agents is 1, given that agent TTL is 170, and expecting the agent to meet a query within 30 hops.

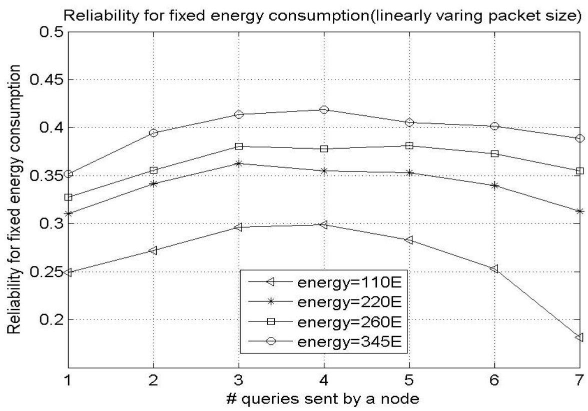

Figure 8: Variation of reliability as number of queries changes under different fixed energy value.

IV. CONCLUSIONS

We derived the exact probabilities of a packet, using random routing on a rectangular grid, visiting a node of interest within given number of hops, and the agents meeting queries. The equations derived facilitate the evaluation of many other useful probabilities for the network such as the probability of a query failing to find an informed node, the probability with which a node should generate the query to achieve a given service requirement, and the probability of a node locating an event without a-priori event information. Since all these probabilities are given as a function of network parameters, model can be used to select parameters for optimum performance, and to observe the performance as these parameters vary. For example we have shown how to achieve the optimal point for fixed energy. Similar model can be used to evaluate the performance in terms of delay, reliability, etc. Finally, we have shown that the model results hold even for sparse networks of more than 75% node available. The model thus could be expected to hold for networks where the sleep/wakeup schedules of neighbors overlap significantly as well.

Main advantages of the model are its simplicity and adaptability. Further research is needed to handle weighted random routing schemes where the probability of next hop depends on other information. We have extended the model for triangular, hexagonal and 3-D grids. Extending the analysis to evaluate the probability of networks using geographical coordinates or logical coordinates is also of interest.

REFERENCES

[1] J. Ahn, S. Kapadia, S. Pattem, A. Sridharan, M. Zuniga, and Jung-Hyun Jun, "Empirical evaluation of querying mechanisms for unstructured WSNs," ACM SIGCOMM Computer Communication Review, Vol. 3, No 3, Jul. 2008

[2] I. F. Akyildiz, S. Weilian, Y. Sankarasubramaniam, and E. E. Cayirci, "A survey on sensor networks," IEEE Communications Magazine, vol. 40, no. 8, pp. 102–114, 2002.

[3] C. Avin and C. Brito, **"**Efficient and robust query processing in dynamic environments using random walk techniques,**"** 3rd Int. Symp. Information Processing of Sensor Network, pp. 277-286, Apr. 2004.

[4] H.M.N.D. Bandara and A. P. Jayasumana, "An enhanced top-down cluster and cluster tree formation algorithm for wireless sensor networks," In Proc. 2nd Int. Conf. on Industrial & Information Systems (ICIIS 2007), Peradeniya, Sri Lanka, pp. 37-42, Aug., 2007.

[5] T. Banka, G. Tandon, and A. P. Jayasumana, "Zonal rumor routing for wireless sensor networks," In Proc. Int. Conf. Information Technology: Coding and Computing (ITCC'05), Vol II**,** pp: 562 – 567, 2005.

[6] D. Braginsky and D. Estrin, "Roumor routing algorithm for sensor networks," In Proc. 1st Workshop on Sensor Networks and Applications (WSNA), 2002.

[7]Q. Cao and T. Abdelzaher, "A scalable logical coordinates framework for routing in wireless sensor networks," 25th IEEE International Real-Time Systems Symposium, pp. 349-358, Dec. 2004.

[8] C. Intanagonwiwat, R. Govindan, and D. Estrin, "Directed diffusion: a scalable and robust communication paradigm for sensor networks," In Proc. 6th Int. Conf. on Mobile Computing and Networking (MobiCom'00), NY, USA, pp. 56–67, 2000.

[9] J. Kulik, W. Heinzelman, and H. Balakrishnan, "Negotiation-based protocols for disseminating information in wireless sensor networks," Wirel. Netw., vol. 8, no. 2/3, pp. 169–185, 2002.

[10] S. Lindsey and C. S. Raghavendra, "Pegasis: Power-efficient gathering in sensor information system," In Proc. e IEEE Aerospace Conf., Mar. 2002.

[11] L. Lima and Jo˜ao Barros **"**Random walks on sensor networks," 5th Int. Symposium on Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks and Workshops (WiOpt 2007), pp 1-5, 2007.

[12] I. Mabrouki , X. Lagrange, and G. Froc, "Random walk based routing protocol for wireless sensor networks," In Proc. 2nd Int. Conf. on Performance Evaluation Methodologies and Tools, Nantes, France, 2007 , Article No. 71.

[13] I. Mabrouki, G. Froc,and X. Lagrange, "Biased random walk model to estimate routing performance in sensor networks," In Proc. 2nd Int. Conf. Performance evaluation methodologies and tools, Nantes, France, Vol. 321, Article No. 71, 2007.

[14] S. Shakkottai, "Asymptotics of query strategies over a sensor Network," In Proc. INFOCOM 2004, vol. 1, pp. 548-557, Mar. 2004

[15] S.D. Servetto, G. barrenechea, "Constrained random walks on random raphs: routing algorithm for large scale WSNs," In Proc.1st ACM Int. Workshop on WSNs and Applications, Atlanta, GA, Sep. 2002.

[16] H. O. Tan and I. Korpeoglu, "Power efficient data gathering and aggregation in wireless sensor networks," SIGMOD Rec., vol. 32, no. 4, pp. 66–71, 2003.

[17] H. Tian and H. Shen, "An optimal coverage scheme for wireless sensor networks," In Proc. Int. Conf. Networking, pp. 722–730, 2005.

[18] H. Tian, H. Shen, and T. Matsuzawa, "Random walk routing for wireless sensor networks," In Proc. 6th Int. Conf, Parallel and Distributed Computing, Applications and Technologies (PDCAT'05), pp. 196-200, Dec. 2005.

[19]Wolfram Math World, "Multinomial Distribution,"Available: http://mathworld.wolfram.com/MultinomialDistribution.html

[20] Ken Ward's Mathematics Pages, "Series Sums of Binomial Coefficients," 2007, Available: http://www.trans4mind.com/personal_development/mathematics/series/summingBinomialCoefficients.htm.