FUZZY DIFFUSION FOR DISTRIBUTED SENSOR NETWORKS

Manikanden Balakrishnan and Eric E. Johnson
New Mexico State University
{mabalakr, ejohnson}@nmsu.edu

ABSTRACT

Distributed Sensor Networks (DSNs) are an emerging technology, recently finding extensive application in scientific and military surveillance. DSNs operate under severe energy constraints and are largely characterized by short-range multi-hop radio communications, which drives the need for energy-efficient routing schemes in such networks. Directed diffusion, a data-centric routing approach for application-aware DSNs has been shown to outperform traditional wireless routing schemes and achieves reasonable energy conservation through data-aggregation. However, directed diffusion ignores the energy level of sensor nodes, and we believe that incorporating the knowledge of relative energy reserves into the routing algorithm will improve energy efficiency and significantly prolong network lifetime. In this paper, we introduce fuzzy diffusion, an energy optimization on the directed diffusion scheme, and quantify its energy performance using ns-2 simulations.

1. INTRODUCTION

A Distributed Sensor network (DSN) comprises a multitude of tiny nodes, collaborating in their sensing, processing and communication process to accomplish high-level application tasks. DSNs provide persistent, unattended monitoring of natural and man-made phenomena in applications such as homeland security, law enforcement, military reconnaissance, space exploration, environmental monitoring, and early warning of natural disasters. These applications often demand continuous monitoring of physical phenomena for extended periods of time without the possibility of replenishing the energy supply at each node. In some applications, a renewable energy source is provided (e.g., a solar panel) but the power available is strictly limited. Thus the effectiveness of a DSN depends on its efficiency in using the limited energy supply.

A typical sensor network (for monitoring applications) consists of hundreds of tiny, short-range, energy constrained, wireless sensors deployed densely in the target area to sense and communicate information. As shown in [1] short-range multi-hop sensor communication provides considerable energy savings as compared to long-range communication and this signifies the importance of energy-adaptive routing schemes for DSNs. Directed diffusion, a data-centric routing paradigm for application-aware sensor networks was proposed and evaluated in [2]. The results show (from sensor perspective) that directed diffusion easily outperforms traditional data disseminations schemes (such as flooding and omniscient multicast), and is fairly energy-efficient, achieved through optimal path selection and data aggregation. The initial research on diffusion schemes [2, 3] leads us to further exploration, with a main objective of introducing adaptive behavior based on energy awareness. One approach to this is to explicitly incorporate the knowledge of relative energy reserves in the network into the routing layer to enable energy-adaptive routing decisions (directed diffusion does not consider the energy level of sensor nodes).

In a multi-hop sensor network, nodes forward a considerable volume of packets for other nodes apart from communicating their own application data. Data-forwarding overhead could be fatal for low energy nodes, and the residual energy resources at the weak nodes largely determine the connectivity and lifespan of a DSN. With this motivation, we propose an energy optimization that can be incorporated into existing diffusion schemes: *Fuzzy Diffusion* shifts the energy cost of data forwarding to non-critical nodes (nodes having high residual energy or less data), while still achieving an energy-balance in the network. Critical nodes (low energy nodes with heavy traffic) have reduced data forwarding burden and expend most of their power in sensing and communicating their sensor data, thus seeking to achieve net longevity. Fuzzy diffusion will be ideal for surveillance applications, where sensor nodes are densely deployed for sustained observation of physical events.

1.1 *Related Work*

This work builds upon the directed diffusion paradigm for distributed coordination in sensor networks, developed in [2] and further optimized in [3]. We try to provide explicit energy-adaptations to the existing diffusion schemes for prolonging the network lifetime. We propose to incorporate energy awareness using some of our previous work in fuzzy routing for mobile ad-hoc networks [4, 5], where the routing algorithm employs fuzzy logic to allocate route resources based on message priority and network congestion status. In this work we apply the idea in [4] to sensor networks, with the logic-design parameters and evaluation metrics chosen appropriately for sensor routing schemes.

GEAR [6], an energy-aware routing scheme for sensor networks explicitly uses node-energy and location information to direct the query flood towards target regions at minimal cost. It applies an energy adaptive neighbor selection algorithm to route the packets and reduce unnecessary communication, thereby achieving significant energy conservation. In this work we focus on improving the energy-efficiency of diffusion sensor networks that uses application-specific information, rather than node location or identity information, to reduce communication costs.

2. FUZZY DIFFUSION

In this section we provide a brief overview of the directed diffusion scheme and explain why fuzzy diffusion is important, followed by a detailed introduction to the fuzzy diffusion protocol. For interested readers, detailed explanation and performance analysis of directed diffusion can be found in [2, 3].

2.1 *Why Fuzzy Diffusion?*

Directed diffusion is a data-centric routing scheme designed specifically for wireless sensor applications. In diffusion, data is *named* using attribute-value pairs and disseminated throughout the network. Data is drawn towards *interested* nodes, which is completely unlike the IP mode of communication where nodes are uniquely identified by their addresses. Diffusion gets the name data-centric, in the sense that all communication is for named-data.

Current research in [3] discusses a family of diffusion algorithms, each variant optimized for a specific application. The early work on directed diffusion [2] is now being referred to as the *two-phase-pull* diffusion algorithm. The fuzzy logic introduced here will embed with any diffusion algorithm and we chose the two-phase-pull scheme, since it generally suits periodic monitoring applications (which are of interest here) wherein the task duration is long enough to amortize the initial gradient set-up cost. Hereafter, directed diffusion refers to the two-phase-pull variant of the family of algorithms.

In directed diffusion, querying nodes (sinks) disseminate 'interest' throughout the network, setting up gradients to draw events (named data) towards the sink. A node receiving an interest re-broadcasts it if it cannot accomplish the query task (providing the requested data). This way interests get forwarded towards source nodes that send sensor data back (at slow exploratory rate) to the sink nodes for the entire task duration. The positive reinforcement phase of diffusion subsequently enables flow of data at high rate from the source to sink.

Nodes forwarding interests implicitly agree to forward data, since it sets up gradients that draw data towards the sinks. The gradient set-up costs are huge in diffusion due to the interest flood, and schemes to reduce this initial cost are mandatory in diffusion. In-network processing is generally used to suppress duplicate interests and redundant broadcasts. Location information can also be used to direct the interest flood, avoiding flooding in non-target regions, thereby providing considerable energy savings. However in diffusion, nodes, irrespective of their target region can *individually decide* whether or not to forward the interest based upon their network status, to achieve net longevity.

2.2 *Fuzzy Diffusion Design*

Recognizing that the decision to forward an interest is an offer to expend energy in support of that interest, we seek to inform the interest forwarding decision with knowledge of energy reserves at the node and its neighbors. The goal of fuzzy diffusion is to create energy awareness in routing such that, at any instant in the network, the load of relaying data is assigned to nodes having relatively plentiful energy reserves, while still maintaining an energy-balance among the nodes.

There are several mathematical tools for decision making in the literature. We chose fuzzy logic, as it is simple, exclusively designed for decision making with multiple input parameters and it also makes an efficient trade-off between significance and precision.

Each node in fuzzy diffusion employs fuzzy logic for computing the interest forwarding probability based on its *network energy status* and *pending traffic* in queue. These two parameters represent the input fuzzy variables that are used by the fuzzy controller to estimate the *criticality* of a node, which determines the probability of interest forwarding. In a crisp sense, interests are suppressed with high probability at critical nodes, relieving them of data forwarding burden. This shifts the energy cost of data relaying to non-critical nodes, which should provide useful energy savings in the network.

2.2.1 *Fuzzy Input Variables*

The two linguistic input variables used by the fuzzy controller are:

Relative Energy Level (REL) of a node, defined as the residual energy status of a node with respect to its neighborhood. This factor represents a node's *energy-criticality* and is given by;

$$REL = \frac{E_{node} - E_{\min}}{E_{\max} - E_{\min}}$$

Where,

E_{max} and E_{min} = Maximum and Minimum energy levels in the neighborhood

E_{node}= Node's residual energy level

The *REL* definition is an approximate rank function [7] that indicates the energy ranking of a node among its neighbors. Higher the *REL*, lesser is the energy-criticality of a node. Nodes exchange energy information by piggybacking their current energy levels along with interest and data messages (energy value can also be included as one of the attribute values), and each node maintains a cache for storing neighbor residual energy levels. The implementation complexity is negligible.

Traffic Intensity (TI) of a node, defined as the amount of traffic pending in a node's queue. This includes the application traffic and also the traffic that a node has already committed to forwarding. This factor represents the *traffic burden* on a node and is given by;

$$TI = \frac{Traffic\ in\ node's\ Queue}{Maximum\ Queue\ Size\ of\ the\ node}$$

Lower the *TI*, lesser the load in a node.

Both *REL* and *TI* lie in the range [0-1] and jointly define the criticality of a node. For example, a node with low *REL* and high *TI* is a critical node, since the remaining energy resources might only be sufficient for the large pending traffic and it may therefore be appropriate to reject any further forwarding requests.

Critical nodes are highly conservative in accepting new interests, and conserve their energy to communicate locally produced application data, thereby maintaining network coverage and connectivity for longer periods of time. Fuzzy diffusion does not degrade the information flow even though some critical intermediate nodes decline to participate in data communication, because nodes with relatively higher energy reserves in a neighborhood still communicate data (one being high rate path) and contribute to the application task. The advantage of using *REL* as energy parameter is that it describes a node's energy status with respect to the network.

During the conservation period (period during which a node's status is critical and interests are rejected with high probability), the *REL* of a critical node increases due to the energy drain in other active nodes in the network (critical nodes still receive data from neighbors and hence the current energy information) during that period. This gradually shifts the critical node to normality (properly reflecting the evolving network energy status) and therefore increases the probability of accepting interests. In fuzzy diffusion environment, *REL* is reasonably resistant to outliers [7] due to the energy balance among the network nodes, though incorporating dispersion parameters into the *REL* computation will be an interesting follow up to this introductory work.

However, *REL* alone will not be sufficient to define the node criticality, since if a low-energy node does not have any pending traffic, it might as well forward data and contribute to the application task. Thus *REL* and *TI* are combined to reflect a node's state more richly and form the core decision parameters of the fuzzy controller.

2.2.2 Fuzzy logic for interest forward

The input variables are represented as discrete fuzzy values with a level of resolution defined by the following fuzzy membership functions (The graphical representation of membership functions are shown in figure 1):

REL Membership function associates *REL* with fuzzy energy values; low, medium and high. Degree of Membership (DOM) represents the magnitude of participation of a node's energy level in a fuzzy set. *REL* value is plotted on the x-axis and is projected vertically to the upper set-boundary lines to determine the DOM with each set.

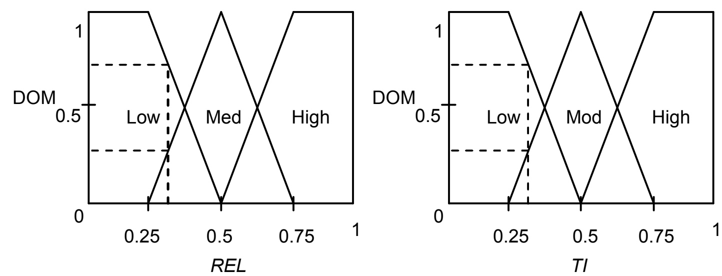

Figure 1. Membership Functions

Similarly the ***TI membership function***, maps the *TI* of a node to three discrete fuzzy values; low, moderate and high. Estimating the DOM of a variable with each fuzzy set identifies the degree of node criticality and a simple straight line fit can be used to compute the DOM values corresponding to *REL* and *TI* values.

The fuzzy decision, whether or not to forward an interest, is computed using a rule base. The decision is expressed as a probability of interest forwarding, p_f. The rules are represented in Mamdani form [8] as:

IF *REL* is a and *TI* is b, THEN P_f is c

where, a, b and c are the linguistic values of the linguistic variables *REL*, *TI* and P_f. The above logic rule can be called a *fuzzy association* and Table 1 shows all the possible associations from the given set of fuzzy values defined by the membership sets.

Table 1. Rule Table: Fuzzy associations for interest forward

TI / *REL*	Low	Moderate	High
Low	*p*	0	0
Medium	1	*p*	0
High	1	1	*p*

The values inside the table are the P_f (fuzzy action) corresponding to the fuzzy variable associations. *p* is a probability representing the line of criticality (its value represents the degree of conservativeness), above which a node moves towards critical stage and has zero probability of forwarding an interest (extremely conservative). p is made tunable in this initial research for identifying the optimal value that can obtain an efficient trade-off between energy conservation and attainable throughput.

2.2.3 *Final Fuzzy Response*

For a given value of *REL* and *TI*, the corresponding DOM values with each fuzzy set are computed. All possible associations between the fuzzy variables are listed and their association weight is computed using the max-min inference method. Then the fuzzy action corresponding to each association is obtained from the rule table. Finally the fuzzy response (P_f) is computed using the centroid method [8]:

$$\text{Probability of interest forward, } P_f = \frac{\sum W_i Z_i}{\sum W_i}$$

W_i = Weight of a rule (association) i = min (DOM$_i$ with *TI* fuzzy sets, DOM$_i$ with *REL* fuzzy sets)
Z_i = Fuzzy action for the rule i (from rule table)

3. IMPLEMENTATION

The entire fuzzy algorithm described in section 2 can be considered to be a fuzzy-controller that can be embedded into any diffusion algorithm. The controller can be invoked when we receive an interest to compute the probability of forwarding the interest. Figure 2 shows a simplified structure where the directed and fuzzy diffusion functionalities differ. U(0,1) is a uniform number, generated between 0 and 1 to implement the probabilistic interest forwarding.

We implemented fuzzy diffusion in *ns-2* [9] simulator (fuzzy logic was embedded into the two phase pull filter application) and obtained performance results for comparison with directed diffusion. The simulations were run with parameter values chosen for a typical monitoring application, and also in accordance with the earlier diffusion research [2, 3] for a consistent comparison. Table 2 lists the chosen configuration parameters.

Sink nodes generate interests (specifying the monitoring rate and duration) and disseminate it throughout the network. Source nodes in the simulation periodically generate events (monitor information) and communicate them to sink nodes throughout the task duration. We modified the messages to include the node's current energy level and all nodes maintain an energy table that lists the collected neighborhood energy information.

The metrics were chosen to analyze the energy efficiency of fuzzy diffusion, and also its potential performance degradation due to the conservative approach. We examine the energy profile (residual energy versus time) of the first node to die, the overall residual energy in the network, and the volume of events reported by directed diffusion and several variations of fuzzy diffusion.

The simulations for fuzzy diffusion were run with three values of p (low p implies highly conservative); 0, 0.5 and 1. The performance results were obtained with 90% confidence intervals.

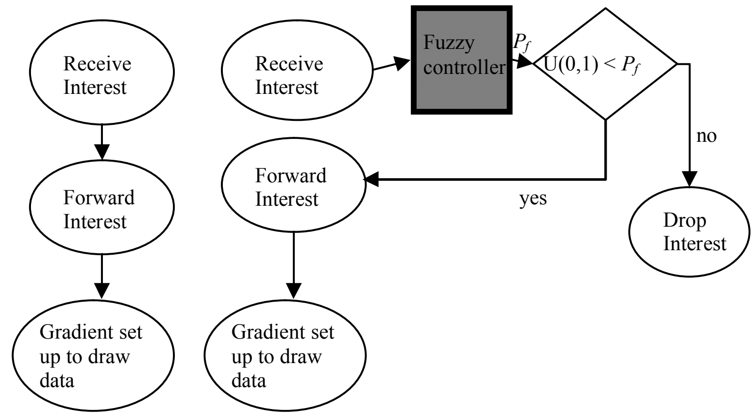

Figure 2. Interest processing structure

Table 2. Simulation Parameters

Number of nodes	100, scattered uniformly in the field
Topography	1400m by 1400m large target area of surveillance
Radio range of nodes	250m
Channel bandwidth	1.6 Mb/s
Simulation run time	600 seconds
Initial energy of nodes	20 Joules
Transmission power consumption	660 mW
Reception power consumption	395 mW
Idle power consumption	39 mW
Number of sources and sinks	10
Event (data message) size	64 bytes
Interest size	36 bytes
Interest generation rate	1 per 30 seconds
Event rates (exploratory and high)	Same as in [1]

3.1 *First-dying node*

Figure 3 shows the energy profile of the first-dying node in the network for all the diffusion protocols. This graph depicts the energy utilization of the diffusion algorithms. The first dying node was chosen, since it is the weakest node in the network at the simulation instant, and how these protocols treat the weakest link is significant for prolonging network lifetime and connectivity.

As expected, directed diffusion has almost linear energy drain even for the weakest node, confirming no explicit energy adaptiveness at the node level. The energy conservation in fuzzy diffusion increases as the value of p

is decreased, since the probability of interest forwarding decreases as p decreases (from fuzzy association table). Nodes become more conservative as we decrease p and hence the node lifetime is prolonged. It is also interesting to see that for $p = 0$ the curve is linear after certain amount of simulation run time. This is because once the weakest node has low *REL* it never accepts interests (P_f is zero for all *TI*) and the power drain is only due to packet reception, idle listening and the data that this node sources (if it's a source or sink).

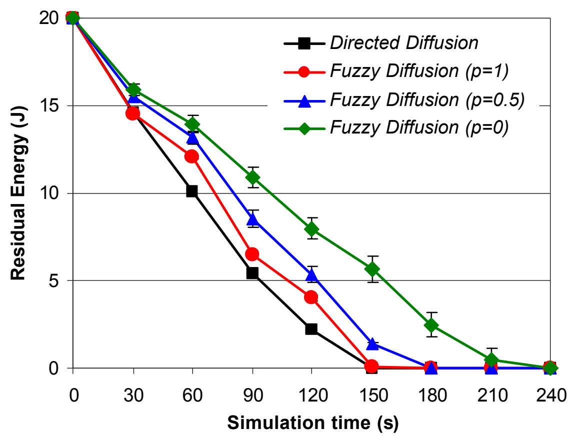

Figure 3. First dying node's energy profile

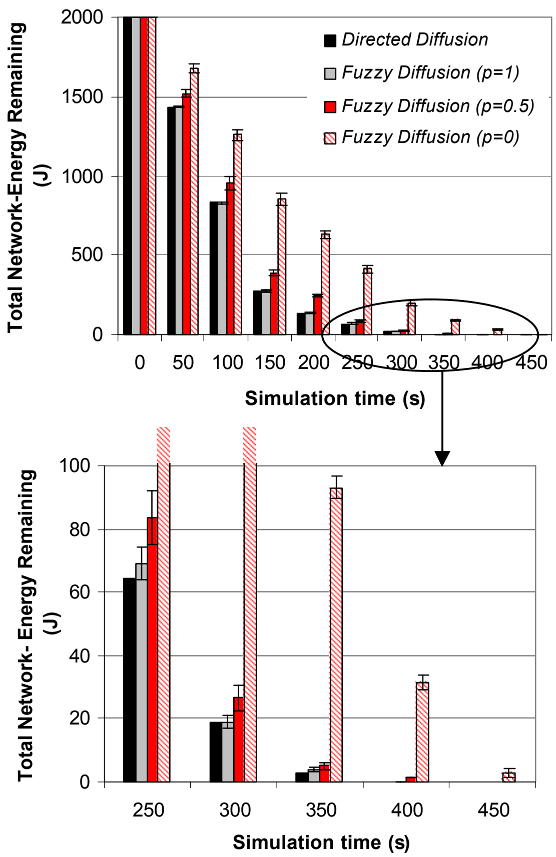

Figure 4. Residual Network Energy

3.2 *Residual energy*

Network energy was extracted from all nodes every 50 sec during the simulation for the purpose of comparing total residual energy in the network versus time for the routing protocols, which are plotted in Figure 4. The results towards the end of simulation are zoomed to clearly depict the extended network life time. Certain results with $p = 0$ in the zoomed version clearly exceed the graph window, but is still included to show the amount of performance difference.

It is clear from the bar-graph that the network life time is prolonged for fuzzy diffusion; in this specific simulation scenario it was extended by 50-100 simulation seconds, which is about 8-17% increase in simulation run time as compared to directed diffusion. The performance improvement might be much higher with optimized fuzzy parameters and configuration parameters, which is subject to further research. The net-energy performance for $p = 0$ is significantly different from other cases, the reason being the same as for figure 4.

Also, as a direct consequence of the prolonged network life time, we can certainly expect the information delivery to improve in fuzzy diffusion as it gets more conservative, since more events are sourced and disseminated throughout the network. This phenomenon is explained in figure 5.

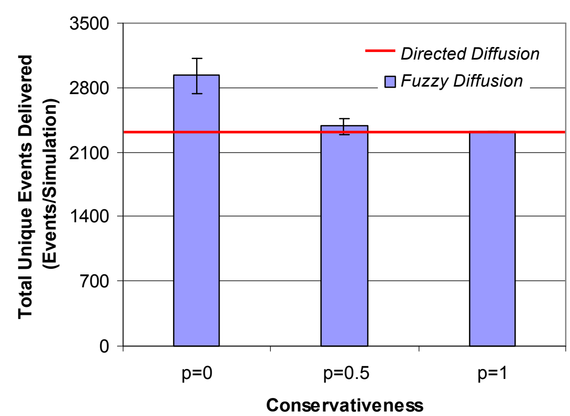

Figure 5. Information Delivery

3.3 *Events reported*

Figure 5 shows the total number of unique events reported at all the sinks during the simulation run and it represents the information delivery efficiency of the diffusion algorithms. The diffusion filter in ns-2 implements a simple in-network processing technique, where intermediate nodes suppress identical events from different upstream neighbors. Each event when generated at a source node is assigned a unique sequence number that enables duplicate event suppression at every hop. A large and dense wireless network comprises several non-disjoint paths from source to sink and this simple technique significantly reduces duplicate transmissions.

Directed diffusion is expected to have the least performance, since it has the lowest network lifetime. It is shown as the benchmark (dark red line) to analyze the performance of fuzzy diffusion. As seen in the graph, for p = 0 the number of events reported is increased significantly due to the considerable amount of net-longevity achieved through extreme conservation. As mentioned before in Sec. 2.2.1, fuzzy diffusion does not degrade the information dissemination even if it gets extremely conservative, since non-critical nodes (higher *REL*) still participate in data forwarding, reinforcing a high data rate path from source to sink.

Critical nodes declining to forward interests, reduces the number of alternate paths (from source to sink) in a large network, thereby reducing several slow exploratory paths. However, the improved network life-time and hence the extended high-data-rate delivery period clearly out shadows the reduction in information delivery due to exploratory paths. Also, the reduced alternate paths indirectly reduce duplicate event delivery at the sinks, further achieving marginal energy savings.

For p = 0.5 and p = 1, the event delivery performance closely approaches that of directed diffusion, while obtaining significant energy savings. From these initial results, p = 0 seems to represent a good trade-off between energy conservation and throughput, although further research is needed. These initial results suggest that in a dense network, diffusion nodes could afford to be highly conservative as there are abundant alternate paths, out of which at least one would be reinforced, contributing to the application task.

The appropriate choice of p depends entirely on network size, density and application requirements. High degree of conservativeness might not be appropriate for a small network with relatively fewer paths, since the reduction in exploratory paths might significantly degrade the information delivery. The above results clearly depict the superior energy performance of fuzzy diffusion over directed diffusion.

4. CONCLUSIONS AND FUTURE WORK

In this paper we have introduced an energy optimization technique, fuzzy diffusion, that can be embedded into any diffusion algorithm and that appears applicable for periodic monitoring applications. The effectiveness of this new technique in prolonging the lifetime of DSNs clearly indicates the importance of explicitly incorporating energy information into the routing layer. Directed diffusion achieves energy savings through in-network data aggregation, which coupled with explicit energy awareness schemes (such as the one proposed here) is a significant and efficient step towards next generation wireless sensor applications.

Further work on fuzzy diffusion will explore various network and cross-layer parameters that influence the criticality of a node. The fuzzy association table is the most important design parameter of fuzzy diffusion and intensive research is needed to obtain its optimal gradient that maximizes energy efficiency and throughput performance. We anticipate that this may vary depending on whether nodes have non-renewable storage or a renewable energy source of finite capacity. Analyzing the performance of fuzzy diffusion under different network sizes and topology, with optimal data-aggregation will quantify any performance degradation due to the conservative approach.

Embedding the fuzzy controller in additional variants of diffusion algorithm [3] and analyzing their performance might provide application-specific optimizations. Fuzzy logic can also be incorporated into the positive reinforcement phase of diffusion such that, at any instant, nodes with high energy reserves are reinforced to send data at high rate. Similar schemes can be designed and implemented using other decision making tools and their performance compared with fuzzy scheme to complement this work.

REFERENCES

[1] Gregory J. Pottie and William J. Kaiser, “Embedding the internet: wireless integrated network sensors”, *Communications of the ACM*, 43(5):51–58, May 2000.

[2] Intanagonwiwat.C, Govindan.R, Estrin.D, “Directed Diffusion: A Scalable and Robust Communication Paradigm for Sensor Networks”, *Proceedings of ACM/IEEE International Conference on Mobile Computing and Networking*, pages56-67, Boston, MA, USA, Aug.2000.

[3] Silva.F, Heidemann.J, Govindan.R, Estrin.D, “Directed Dif fusion”, USC/ISI Technical Report ISI-TR-2004-586.

[4] Alandjani.G and Johnson.E , “Fuzzy Routing in Ad Hoc Networks”, *Proceedings of 2003 IEEE International Performance, Computing, and Communications Conference*, IEEE, Phoenix, AZ, April 2003.

[5] Alandjani, Gasim, *Fuzzy Routing for Ad Hoc Networks*, Thesis Report, New Mexico State University, 2002.

[6] Yan Yu, Ramesh Govindan, and Deborah Estrin. Geographical and energy aware routing: A recursive data dissemination protocol for wireless sensor networks. Technical Report TR-01-0023, University of California, Los Angeles, Computer Science Department, 2001.

[7] R. K. Jain, *The Art of Computer Performance Analysis*, John Wiley and Sons, 1992.

[8] Ross T.J, *Fuzzy Logic with Engineering Applications,* II Edition, John Wiley and Sons, 2004.

[9] Sandeep Bajaj, Lee Breslau, Deborah Estrin, Kevin Fall, Sally Floyd, Padma Haldar, Mark Handley, Ahmed Helmy, John Heidemann, Polly Huang, Satish Kumar, Steven McCanne, Reza Rejaie, Puneet Sharma, Kannan Varadhan, Ya Xu, Haobo Yu, and Daniel Zappala. Improving simulation for network research. Technical Report 99-702b, University of Southern California, March 1999. revised September 1999, to appear in IEEE Computer.