

Research Article

A Fast and Reliable Hybrid Data Delivery Protocol for Large-Scale Heterogeneous Sensor Networks

Sanggil Kang,¹ Yujin Lim,² Wilmarc Lopez,¹ and Hoyoung Hwang³

¹Department of Computer Science and Computer Engineering, Inha University, Incheon 420-751, Republic of Korea

²Department of Information Media, University of Suwon, Hwaseong-si 445-743, Republic of Korea

³Department of Multimedia Engineering, Hansung University, Seoul 136-792, Republic of Korea

Correspondence should be addressed to Yujin Lim, yujin@suwon.ac.kr

Received 14 September 2011; Accepted 16 October 2011

Academic Editor: Junyoung Heo

Copyright © 2012 Sanggil Kang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

We propose a hybrid data delivery method for the large-scale heterogeneous sensor networks, which is a fast and reliable delivery protocol for the aggregated data from the sinks to the GW. We develop a new multicriteria-ranking algorithm which determines multiple forwarders for each hop by ranking neighbor nodes. To rank the nodes, we compute the fitness value using features for each node such as the received signal strength, nodal delay, and hop distance. We determine the time of sending among forwarders using the waiting time assignment algorithm. In the experimental section, we show that our method outperforms conventional data delivery protocols in terms of data delivery ratio and end-to-end delay.

1. Introduction

Wireless sensor networks (WSNs) are multihop wireless networks which main purpose is to deliver sensed data collected from multiple sensors to one or more data-collecting devices (or sinks). Its inherent power limitation [1], wireless connectivity and vast number of applications have been attracting much attention from a good number of researchers in the recent years [2, 3]. WSNs are used in acquiring information from inaccessible or perilous areas. Sensed data collected at a sensor are delivered to a remote monitoring center through an intermediate system.

Large-scale heterogeneous sensor network, one type of WSN, has a number of heterogeneous wireless nodes deployed over a wide geographical region. The WSN is commonly made up of two subnetworks: a sensor network and a distribution network. Sensors in the sensor network collect or sense data and deliver them to the nearest sink. In general, sensors are static and densely deployed so they can be easily connected to a sink. The distribution network on the other hand is responsible for delivering data gathered from sinks to the remote monitoring center through a gateway (GW). The GW in turn is connected to an external network like the Internet where the remote monitoring center gets the

data. Issues on the deployment of nodes and data collection in the sensor network have already been well addressed in a number of literatures [4–7]. In general, data delivery from sensors to the nearest sink is assumed to be loss tolerant due to the sheer amount of correlated data [8]. However, the aggregated data in the distribution network are not. Furthermore, the significant distance between the monitoring center and the GW is also a challenge. But since they are connected through a high-speed wired network, we focus on the data delivery between the sinks and the GW. In addition, it is desirable for the protocol to be relatively fast because a remote monitoring center may need to take appropriate actions based on the information provided by the distribution network.

Hence in this paper, we develop a fast and reliable delivery protocol for the aggregated data from the sinks to the GW. We obtain good data delivery ratio and end-to-end delay by solving the drawbacks of conventional protocols. We also present a new multicriteria ranking paradigm and present a waiting time assignment algorithm. In the experimental section, we show that our protocol outperforms conventional flooding and relaying techniques.

This paper is organized as follows. Section 2 presents the drawbacks of related works. In Section 3, we describe

in detail our proposed protocol followed by the analysis of experimental results in Section 4. We conclude in Section 5.

2. Related Works

In general, wireless data delivery protocols can be divided into two major categories, (1) relaying and (2) flooding approaches [9]. Relay methods send messages once to a single forwarder. Usually, these methods make use of a multi-criteria fitness function to determine the link cost of its neighbors [10, 11]. Such criteria are usually extracted through a network discovery process commonly done by sending control messages. The more the number of control messages sent, the better the decision process of choosing the next forwarder reflects the state of the network. However, too many of these messages might cause data delivery ratio to fall due to increase in the probability of collision. And since packets are sent only once to a single forwarder, the probability of packet lost is also high. Moreover, the delivery time may suffer from the overhead of the network discovery process.

Flooding techniques on the other hand, in their purest sense, broadcast messages epidemically by forwarding the data to all of its neighbors [12]. In flooding techniques, data is delivered relatively fast because there is minimal or no overhead [13]. However, flooding protocols are notorious for causing *broadcast storms* which lowers deliver ratio due to high chance of collision between messages. Usual attempts to solve the problem like RBVT [13, 14], MHVB [12, 15] and RSB [16] assign neighboring nodes a *waiting time* before sending. It has a value commonly dependent on the forward progress given by $d_{SD} - d_{N_iD}$, where d_{SD} is the distance between the sink S and the destination D and d_{N_iD} is the distance between the receiving neighbor N_i and D [12–16]. The best neighbor, through computation of its own waiting time, should in principle have the least value. By this, the best node broadcasts first. The other neighbors cancel their timers and suppress rebroadcasting or reset their timer to a random value [16], after receiving the duplicate broadcast from the winning node. Such technique can be generally referred to as the election approach [14]. The approach has also an advantage of implicitly having multiple backup forwarders.

The following section explains how to solve the drawbacks of both relay and flooding approaches for the WSN's distribution network.

3. Multicriteria Sender-Side Election Protocol

To explain our proposed protocol named multi-criteria sender-side election (MCSS) protocol, we discuss how we model the local network using neighbor information then present how it is used to rank the neighboring nodes. We present how we assign justifiable waiting time to the best k ranking sinks. Lastly, we discuss how the idea fits into the whole delivery protocol.

3.1. Neighbor Table. Instead of building a path from the sink to the GW, our protocol selects its next forwarder

on each hop in order to improve data delivery ratio and end-to-end delay. For each hop, the forwarding node ranks its one-hop neighbors with respect to how much they are able to receive the message. We call it the fitness value of a neighboring node. The fitness value reflects the communication environment of the node. To do this, the Neighbor Table (NT) contains three fields: received signal strength (RSS), nodal delay of messages received from the neighbor, and hop distance from the GW.

The RSS of a message measures the signal strength between the sink and the neighbor. The higher the RSS, the better the chance that broadcasts are successful. The nodal delay describes the amount of time it takes for a message to be delivered from the sink to the neighbor. It includes the processing delay, the queuing delay, the transmission delay, and the propagation delay. The less is the nodal delay of messages received, the less is the end-to-end delay. Since packet loss probability in a wireless multi-hop communication environment increases with the number of hops [17], we choose the hop distance between the neighbor and the GW as one of the criteria. To infer the hop distance of each node, we periodically flood a PROBE message in the distribution network.

The sinks update their own NTs for every message received. The three fields of the received messages are averaged using the Exponential Moving Average (EMA). We think of the NT as an ordered set of the i th neighbor vector $\vec{v}_i = \langle c_{i,j} \rangle$, $i = 1, 2, \dots, n$ and $j = 1, 2, 3$ with $c_{i,1}$ as the RSS, $c_{i,2}$ as the nodal delay, and $c_{i,3}$ as the hop distance of neighbor i from the GW. Each value is updated as follows:

$$\bar{c}_{i,j}(t) = \alpha \bar{c}_{i,j}(t-1) + (1-\alpha)c_{i,j}(t), \quad (1)$$

where $\bar{c}_{i,j}(t)$ and $\bar{c}_{i,j}(t-1)$ are the new and old average values at time t , respectively. $c_{i,j}(t)$ is measured value at time t and the constant α is a smoothing factor.

3.2. Selection of Forwarders. The conventional schemes selecting a single forwarder can increase the chance of data delivery failure because the forwarder can miss the message in dynamic communication environment. Thus, we choose more than one forwarder by ranking neighbor nodes. To rank the nodes, we compute the fitness value using the three fields in NT for each neighbor. Note that the number of forwarders depends on the degrees of connection. In our experiments, we set it the half of the number of sink's neighbors.

In general, the larger is the value of RSS and the smaller are the nodal delay and hop distance, the better is the performance. For the consideration, we inverse the nodal delay and hop distance using

$$c_{i,j} = \begin{cases} \frac{1}{(c_{i,j} + \lambda)}, & j = 2, 3 \\ c_{i,j}, & j = 1, \end{cases} \quad (2)$$

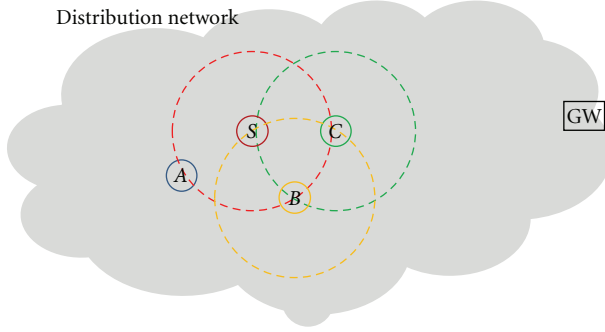


FIGURE 1: Sink S sends a message to the GW through one of its neighbors A , B or C .

TABLE 1: Neighbor table of S in Figure 1.

Neighbor	RSS (dBmV/m)	Nodal delay (microseconds)	Hop distance (hops)
A	3	4.5	2
B	2.5	4	1
C	2	4.3	1

TABLE 2: Inverse values for Table 1 with $\lambda = 1 \times 10^{-9}$.

Neighbor	RSS*	Nodal delay	Hop distance
A	3	0.22	0.5
B	2.5	0.25	1
C	2	0.23	1

where λ is a very small positive number used to avoid division by zero. The fitness value F_i of node i is given as

$$F_i = \sum_{j=1}^3 \Gamma_j(c_{i,j}) \cdot w_{i,j}, \quad (3)$$

where $\Gamma_j(c_{i,j})$ is the min-max normalization function of $c_{i,j}$. The weight $w_{i,j}$ can be thought of as the importance of field j to the overall evaluation and it is calculated as

$$w_{i,j} = \frac{\max_{i \in n} c_{i,j}}{\sum_{i \in n} c_{i,j}}. \quad (4)$$

The top k ranked neighbors are chosen as forwarders. To illustrate our method, we use the node setup in Figure 1 and NT in Tables 1–3. Here, the sink S sends a data to the GW and ranks its neighbors A and B . The final ranking makes B the overall best neighbor with a fitness value equal to 0.95. If $k = 1$, then only B is considered as a forwarder.

The responsibility of computing the waiting time is moved from the receiving neighbors to the sink. After choosing the forwarders, the sink assigns each forwarder with a waiting time [18]. We assign the best ranked forwarder with zero waiting time. For the waiting time gaps between adjacently ranked forwarders, we use the nodal delays between each of them.

TABLE 3: Scores and fitness values of the neighbors of S based on Table 2 with $w_1 = 0.4$, $w_2 = 0.35$, and $w_3 = 0.4$.

Neighbor	RSS Score	Nodal delay Score	Hop distance Score	Fitness value
A	1	0	0	0.4
B	0.5	1	1	0.95
C	0	0.37	1	0.52

TABLE 4: Waiting time table with $n = 2$ ($\beta = 0$).

Neighbor	Waiting Time
B	0
C	4.15

The waiting time assigned to a forwarder with rank $r = 1, 2, \dots, k$ is given as

$$(r - 1) \frac{\sum_{i=1}^{k-1} c_{i,2}}{k - 1} + \beta, \quad (5)$$

where β is a small random number. The waiting time table as shown in Table 4 is appended to and sent together with the data.

3.3. The Delivery Protocol. This section discusses the overall data delivery protocol from a sink to the GW. When a sink collects and aggregates sensed data from the sensors in its service area, it sends the aggregated data to the GW. We refer to the aggregated data as the message. We use the hybrid sender-side election approach discussed above. Additionally, each sink node keeps a record of the messages and sends and forwards in a message table (MT). An entry in the MT contains a message ID and the last known time-to-live (L.TTL). An ACK field is also included to serve as a crossing-out mechanism for successfully acknowledged one-hop broadcasts and is initially set to NO_ACK.

The sink node computes its waiting time and attaches it to the message. The message is broadcasted to its neighbors. The sink enters waiting mode for the message by starting a timer initially set to the mean of its forwarders' nodal delay. And the sink creates an entry in the MT with L.TTL equals to the maximum TTL. If a duplicate message is received during the time and the TTL is less than L.TTL, the sink stops its timer and exits waiting mode. The message is acknowledged by setting ACK to YES_ACK and ignores future duplicates. In other words, the received message is already a rebroadcast from one of its forwarders and has already progressed by one hop. We refer to the condition as the suppression condition. Otherwise, if the timer expires, the message is rebroadcasted. Figure 2 illustrates the data delivery process.

4. Experimental Results and Analysis

In this section, we present our experiments done using the NS-2 simulator [19]. We used the shadowing propagation model [20] with path loss exponent $\beta = 3.0$ and shadowing deviation $\sigma_{dB} = 5.0$ to model a noisy outdoor environment.

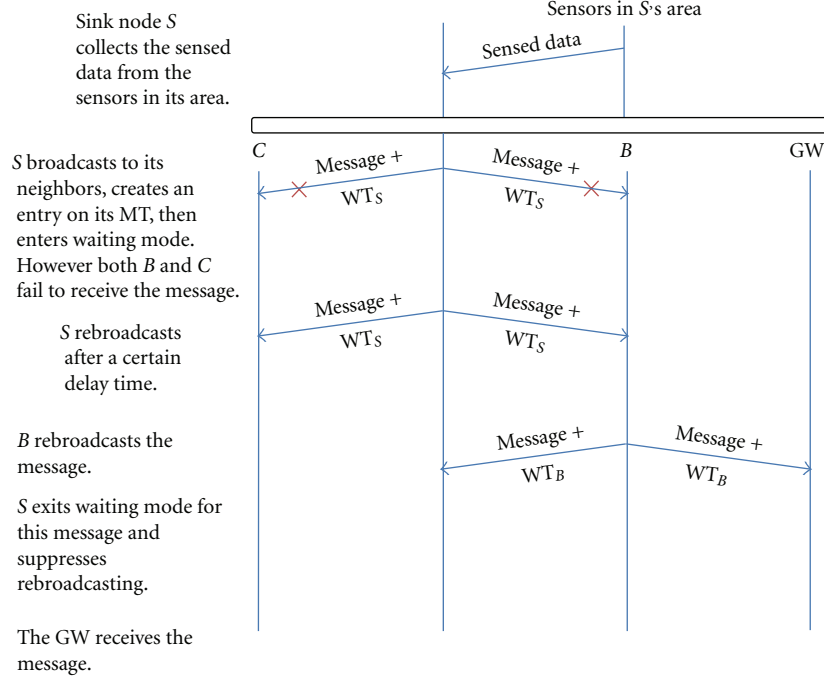


FIGURE 2: Sequence diagram for an example data delivery scenario for Figure 1.

Sinks are placed in a 2000 m by 2000 m area and a transmission range of each node is 250 m. A sink sends a message to the GW every second at constant bit rate (CBR) for 100 seconds. Additionally, we use IEEE 802.11 [21] as the MAC layer.

We first experimented on both conventional flooding and relay approaches to serve as basis for comparison with our protocol. For the relay approach, we investigated four combinations of criteria setups. In all setups, nodes send a control message every 100 milliseconds. For the flooding approach, we experimented on two weight combinations of the RBVT protocol [13]. This protocol uses forward progress, RSS and optimal transmission area given by

$$f_{\text{trans}}(x) = \begin{cases} x + d_{\text{trans}}, & x \leq d_{\text{opt}}, \\ -x + d_{\text{max}}, & x > d_{\text{opt}}, \end{cases} \quad (6)$$

where $d_{\text{opt}} = 100$ m and $d_{\text{max}} = 250$ m are the optimal and the estimated maximum transmission ranges, respectively, and $d_{\text{trans}} = 150$ m is the translation distance.

We evaluated the performance of the protocols through the following performance factors.

- (1) Data delivery ratio—the value computed by dividing the number of data packet successfully received at the GW by the total number of data packets sent by sinks.
- (2) End-to-end delay—the time needed to forward data from a sink to the GW.

In Figure 3, we can see that the relay approach using only RSS performs best in terms of delivery ratio compared to other conventional setups. Its performance falls with the

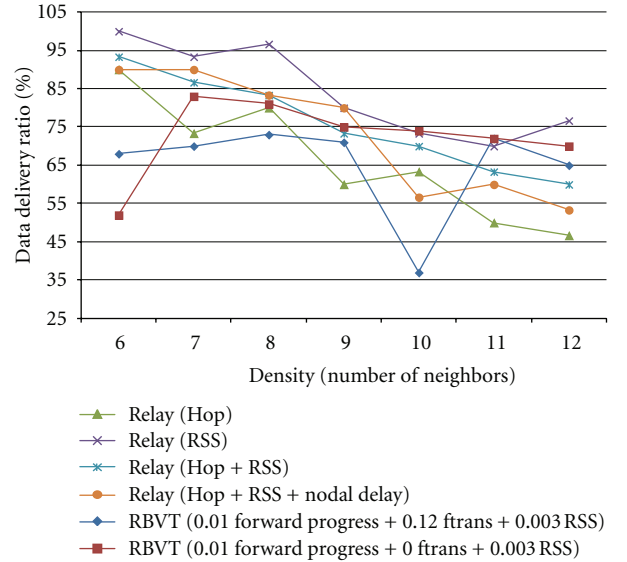


FIGURE 3: Data delivery ratio of conventional protocols.

increase of neighbor density. This is due to collisions from the intense number of control messages. On the other hand, though RBVT does not send control messages, its delivery ratio might have suffered from huge waiting time values being directly equal to the sum of the three criteria.

In Figure 4, relay approach using only RSS is still the best performing protocol in terms of end-to-end delay. Though RBVT does not have a discovery process, the distance-dependent waiting time makes it perform worst. Additionally, it is also important to note that considering RSS

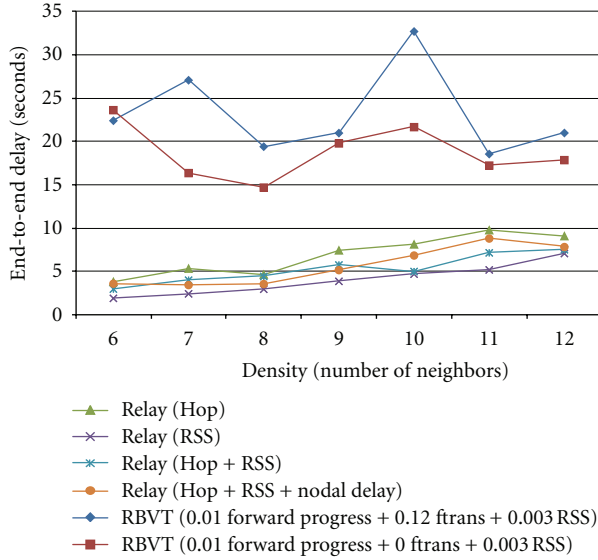


FIGURE 4: End-to-end delay of conventional protocols.

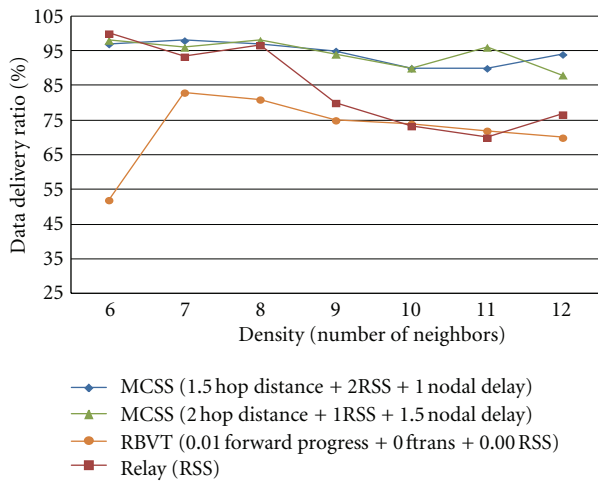


FIGURE 5: Data delivery ratio of MCSS and conventional protocols.

more makes the overall performance of the relay approach better and somewhat less spiky. This gives us an idea on how important the assignment of weights is.

Using the best setups for both conventional approaches, we compared the performance of our MCSS protocol using two weight combinations. In Figure 5, we can see that our protocol achieves an increase of about 20% in data delivery ratio. It is caused by the reduction of control messages. Additionally, the assignment of zero waiting time value for the best forwarders avoids unnecessary pending for message. Moreover, the improved performance can also be attributed to having the $k - 1$ ranking forwarders as backup resenders.

In Figure 6, we can see that MCSS decreases the end-to-end delay of relay approach and RBVT protocol by about 100% and 500%, respectively. It is due to the waiting time values being well gapped, that is, nodal delay dependent and sender-side assigned. Additionally, note that varying

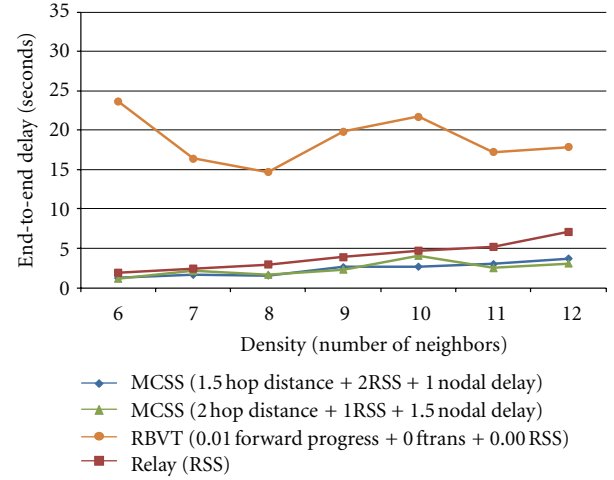


FIGURE 6: End-to-end delay of MCSS and conventional protocols.

the weights in our protocol has minor effect on its overall performance. Hence MCSS is resilient to poor weight assignments with respect to the protocols experimented upon.

We therefore conclude from the above experimental results that our MCSS protocol improves the quality of data delivery between the sinks and the GW in terms of end-to-end delivery ratio and delay.

5. Conclusion

We developed a fast and reliable data delivery protocol for large-scale heterogeneous sensor networks. The summary and contributions of this paper are as follows.

- (1) We identified the data delivery problem in the distribution network of large-scale heterogeneous WSNs.
- (2) We analyzed conventional wireless data delivery protocols and enumerated their weaknesses.
- (3) We developed a selection algorithm of forwarders.
- (4) We developed a sender-side waiting time assignment algorithm for forwarders.
- (5) We developed a general hybrid multi-criteria sender-side election protocol which utilizes the two developed ideas.
- (6) We applied our approach to the data delivery process between the sinks and the GW in the distribution network.
- (7) We showed that our protocol outperforms conventional data delivery protocols in terms of data delivery ratio and end-to-end delay.

In our approach, we used three criteria to model WSN. As a future work, we will do research on more possible criteria and on ways to assign weights dynamically to correspond to a better performance.

Acknowledgment

This work was supported by the GRRRC program of Gyeonggi province ([GRRRC SUWON2011-B4], Research on Precision Location Tracking System for Real-Time Situation).

References

- [1] H. Chen, C. K. Tse, and J. Feng, "Minimizing effective energy consumption in multi-cluster sensor networks for source extraction," *IEEE Transactions on Wireless Communications*, vol. 8, no. 3, Article ID 4801500, pp. 1480–1489, 2009.
- [2] D. Culler, D. Estrin, and M. Srivastava, "Overview of sensor networks," *IEEE Computer*, vol. 37, no. 8, pp. 41–49, 2004.
- [3] K. Romer and F. Mattern, "The design space of wireless sensor networks," *IEEE Wireless Communications*, vol. 11, no. 6, pp. 54–61, 2004.
- [4] D. Niculescu, "Communication paradigms for sensor networks," *IEEE Communications Magazine*, vol. 43, no. 3, pp. 116–122, 2005.
- [5] J. N. Al-Karaki and A. E. Kamal, "Routing techniques in wireless sensor networks: a survey," *IEEE Wireless Communications*, vol. 11, no. 6, pp. 6–28, 2004.
- [6] F. Ye, H. Luo, J. Cheng, S. Lu, and L. Zhang, "A two-tier data dissemination model for large-scale wireless sensor networks," in *Proceedings of the 8th Annual International Conference on Mobile Computing and Networking (MobiCom '02)*, pp. 148–159, Atlanta, Ga, USA, September 2002.
- [7] D. De Couto, D. Aguayo, J. Bicket, and R. Morris, "A high-throughput path metric for multi-hop wireless routing," *Wireless Networks*, vol. 11, no. 4, pp. 419–434, 2005.
- [8] Y. Sankarasubramaniam, O.B. Akan, and I.F. Akyildiz, "ESRT: event-to-sink reliable transport in wireless sensor networks," in *Proceedings of the Annual ACM International Conference on Mobile Computing and Networking (MobiCom '03)*, pp. 177–188, Annapolis, Md, USA, June 2003.
- [9] P. Cataldi, A. Tomatis, G. Grilli, and M. Gerla, "A novel data dissemination method for vehicular networks with rateless codes," in *Proceedings of the IEEE Wireless Communications and Networking Conference (WCNC '09)*, pp. 1–6, April 2009.
- [10] B. K. Szymanski and G. G. Chen, "Computing with time: from neural networks to sensor networks," *Computer Journal*, vol. 51, no. 4, pp. 511–522, 2008.
- [11] W. Zhao, D. Liu, and Y. Jiang, "Distributed neural network routing algorithm based on global information of wireless sensor network," in *Proceedings of the WRI International Conference on Communications and Mobile Computing (CMC '09)*, pp. 552–555, Yunnan, China, January 2009.
- [12] T. Osafune, L. Lin, and M. Lenardi, "Multi-hop vehicular broadcast (MHVB)," in *Proceedings of the 6th International Conference on ITS Telecommunications (ITST '06)*, pp. 757–760, Chengdu, China, June 2006.
- [13] J. Nzouonta, N. Rajgure, G. Wang, and C. Borcea, "VANET routing on city roads using real-time vehicular traffic information," *IEEE Transactions on Vehicular Technology*, vol. 58, no. 7, pp. 3609–3626, 2009.
- [14] K. Egoh and De Swades, "A multi-criteria receiver-side relay election approach in wireless ad hoc networks," in *Proceedings of the ITST 2006 Military Communications Conference (MILCOM '06)*, Washington, DC, USA, October 2006.
- [15] M. O. Cherif, S. M. Secouci, and B. Ducourthial, "How to disseminate vehicular data efficiently in both highway and urban environments?" in *Proceedings of the 6th Annual IEEE International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob '10)*, pp. 165–171, Niagara Falls, Canada, October 2010.
- [16] J. Lee and W. Chen, "Reliably suppressed broadcasting for Vehicle-to-Vehicle communications," in *Proceedings of the 71st IEEE Vehicular Technology Conference, VTC 2010-Spring*, pp. 1–7, Taipei, Taiwan, May 2010.
- [17] S. Toumpis and A. J. Goldsmith, "Capacity regions for wireless ad hoc networks," *IEEE Transactions on Wireless Communications*, vol. 2, no. 4, pp. 736–748, 2003.
- [18] W. Wang, F. Xie, and M. Chatterjee, "Small-scale and large-scale routing in vehicular ad hoc networks," *IEEE Transactions on Vehicular Technology*, vol. 58, no. 9, Article ID 5089413, pp. 5200–5213, 2009.
- [19] "The network simulator, ns-2," <http://www.isi.edu/nsnam/ns/>.
- [20] T. S. Rappaport, *Wireless Communications, Principles and Practice*, Prentice Hall, New York, NY, USA, 1996.
- [21] IEEE Std. 802.11, "Wireless LAN medium access control (MAC) and physical layer (PHY) specification: higher speed physical layer (PHY) extension in the 2.4 GHz Band," 1999.

