Node Aging Effect on Connectivity of Data Gathering Trees in Sensor Networks

Jae-Joon Lee, Bhaskar Krishnamachari and C.-C. Jay Kuo
Department of Electrical Engineering, University of Southern California
Los Angeles, CA 90089-2564, USA, E-mails: {jaejlee,bkrishna}@usc.edu, cckuo@sipi.usc.edu

***Abstract*— Sensor nodes age over time due to device failure and/or battery energy depletion. Node survival rates affect data communication, sensing coverage, and especially the connectivity of data gathering trees that provide a forwarding path from each source to the sink and enable data aggregation. The node aging effect on the connectivity of a data gathering tree over time is analyzed in this research. First, we discuss the general node aging problem by considering the device failure rate and the energy consumption rate. In the analysis of the energy consumption rate, we examine the effect of the data aggregation degree and the hop distance on the amount of data communication handled by a node of a data gathering tree. Then, we present the survival function and the connectivity probability per hop distance in a data gathering tree. Finally, the resulting non-uniform connectivity over time in a data gathering tree is examined with the comparison between the device failure effect and the energy depletion effect through extensive simulation. It is shown by mathematical analysis as well as simulation that the node aging process has a significant impact on connectivity as the hop distance increases.**

I. Introduction

Most sensor network applications assume that a large number of low-cost sensor devices with limited battery energy are deployed in an unattended manner. Under such an environment, sensor devices are more vulnerable to failures caused by external environment events and/or internal device problems than those in a controlled system. The unexpected failure or battery energy depletion limits the lifetime of sensor nodes. This sensor node aging process has a negative impact on the network operations for applications such as data gathering and sensing coverage.

Recent studies on connectivity have focused on identifying a proper radio range to achieve the connectivity between any two nodes in a uniformly distributed sensor environment at the deployment time [1] [2]. Shakkottai *et al.* [3] considered the node failure rate in bounding the probability for all nodes to be connected. However, many sensor network applications require mainly the connectivity between any source and its sink via a data gathering tree rather than between arbitrary end points [4]. In addition, network dynamics and communication patterns may result in a time-varying survival rate of nodes, which in turn leads to a non-uniform node density and connectivity across the network in the node aging process.

The node aging process is concerned with the survival rate of nodes, which is affected by the device failure probability and the energy consumption rate. To the best of our knowledge, the effect of these factors on the connectivity of a data gathering tree has not yet been examined extensively. Sensor device reliability can be analyzed using conventional reliability models that include various device lifetime functions of a diverse shape and scale [5]. The energy consumption rate in data gathering trees is mainly governed by the energy cost to receive data from descendant nodes and transmit data to the sink. During data gathering, data aggregation can be performed to reduce the energy cost [6]. The data aggregation degree would be an application specific factor that reflects the characteristics of data forwarded to the sink. A different data aggregation scheme affects the energy consumption rate of nodes at various hop distances from the sink, leading to a different node aging process in a data gathering tree.

In this work, issues associated with non-uniform connectivity along the data gathering tree in the sensor network due to the aging of nodes over time are examined as in our earlier work [7] where we presented the preliminary energy consumption rate analysis and the radio range effect on the connectivity through simulations. In this paper, we focus on the effect of different data aggregation and compare the device failure effect and the energy consumption effect on the connectivity. We first discuss device reliability and the energy consumption rate. For device reliability, the Weibull distribution [5] is used to analyze the effect of device failure on connectivity. For the energy consumption rate, we show that nodes at a different hop distance will have different communication requirements to result in a different energy consumption behavior. Based on these two results, we are able to conduct the connectivity analysis of a data gathering tree with respect to the hop distance from the sink, and derive a lower bound. Finally, we perform the computer simulation to confirm analytic results and compare two effects on the connectivity.

The rest of this paper is organized as follows. In Section II, device reliability and the energy consumption rate of a node aging process are discussed. Section III presents the connectivity analysis. Then, simulation results are provided in Section IV. We give concluding remarks as well as some future research topics in Section V.

II. Node Aging Process

We examine the node aging process that is caused either by the device failure or equipped energy depletion in this section.

0-7803-8521-7/04/$20.00 © 2004 IEEE

A. Device Reliability

The low-cost sensor device is vulnerable to failure due to external or internal problems. Sensor devices can be deployed in diverse environments, including a hostile area. The external environment such as temperature, pressure, etc. may result in device malfunctions. It may also have internal hardware or software failure.

Here, we model the reliability of a sensor node using a classical distribution known as the Weibull distribution [5]. The distribution has a diverse shape and scale and is often used to model the lifetime of a device. The probability density function of the Weibull distribution is of the following form:

$$f(t) = \frac{\beta}{\eta}(\frac{t}{\eta})^{\beta-1}e^{-(\frac{t}{\eta})^{\beta}}, \quad (1)$$

where β and η are the shape and the scale parameters. It indicates the likelihood of failure at time t. The reliability function of node i that obeys the Weibull distribution is given by

$$R_i(t) = e^{-(\frac{t}{\eta_i})^{\beta_i}}, \quad (2)$$

which is the complement of the cumulative distribution function of $F(t)$ of the Weibull density function given in (1). It is the probability that a device is functioning at time t.

B. Energy Consumption Rate Analysis

Main energy consumption activities include communications, data processing and sensing. Communications are needed for situations such as data gathering to the sink, in-network data communication within a cluster, and control message exchanges. Since the communication plays a dominant role in the battery energy consumption [8], we focus on the data gathering communication of nodes for the energy consumption analysis.

A data gathering tree is shown in Fig. 1. It is formed such that all deployed sensor nodes are connected to a sink, and can be constructed as follows. First, we find the set of first hop nodes that can directly connect to the sink within its radio range. Then, other nodes select the parent nodes that are within their radio range and have a shorter hop distance from the sink to establish the forwarding path towards the sink [4]. When nodes are densely and uniformly distributed in a complete circle network area with a sink located at the center of the circle as shown in Fig. 1, the area of the h-hop distance from the sink is

$$((h+1)^2 - h^2)\pi R^2 = (2h-1)\pi R^2.$$

The ratio of the number of nodes at each hop distance would be the same as that of the area. Thus, the average number N_h^C of children of a node at hop distance h can be computed as

$$N_h^C = \frac{2h+1}{2h-1}.$$

In a data gathering tree, the number of children in a data gathering tree and the data aggregation degree are the two main factors that determine the time of a node's death due

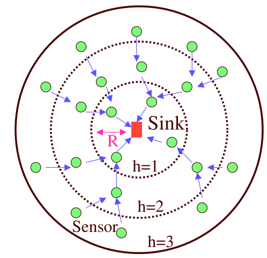


Fig. 1. The illustration of a data gathering tree, where h represents the hop distance from the sink and R indicates the radio range of a sensor device

to energy depletion. We use the data aggregation degree, denoted by k_d with $0 \leq k_d \leq 1$, to characterize how the data from child nodes are aggregated with its own data in a data gathering tree. It is apparent that k_d is an application dependent parameter. When k_d is 1, all data from child nodes can be perfectly aggregated into its own data (*e.g.*, MIN, MAX, SUM, COUNT), and $k_d = 0$ means that no aggregation is performed at all during data gathering.

Let D_i^T and D_i^R be the amount of data to transmit and to receive during one round of data gathering process (*i.e.* time unit in this paper), respectively. D_i^S denotes the amount of data generated by node i per round. Then, the total amount of data to transmit can be written as

$$D_i^T = D_i^S + (1-k_d)D_i^R = D_i^S + (1-k_d)\sum_{j=1}^{N_i^C} D_j^T, \quad (3)$$

where N_i^C is the number of children of node i and j is the jth child node of node i. If all nodes generate one data packet per each data gathering round towards the sink, without data aggregation ($k_d = 0$), the number of packets received by each node per round is the same as that of the descendent of the node. Then, each node forwards received packets in addition to its own data generated at the same time period. In contrast, if the forwarding nodes perfectly aggregate the data from its children ($k_d = 1$), the received packets decrease proportionally with the number of children, and the number of packets to transmit is one per round.

To analyze the energy consumption rate at each hop distance, we can generalize (3) with D_h^T, which is the average amount of data to transmit from a node that is h hop distant from the sink. If all nodes generate D^S per round and nodes at h hop distance away have N_h^C children nodes on the average, then D_h^T can be obtained as

$$\begin{aligned}
D_h^T &= D^S\Big(1 + \sum_{i=h}^{L-1}((1-k_d)^{i+1-h}\prod_{j=h}^{i} N_j^C)\Big) \\
&= D^S\Big(1 + \sum_{i=h}^{L-1}\big((1-k_d)^{i+1-h}\prod_{j=h}^{i}(\frac{2j+1}{2j-1})\big)\Big) \\
&= D^S\Big(1 + \frac{(1-k_d)^{1-h}}{2h-1}\sum_{i=h}^{L-1}((1-k_d)^i(2i+1))\Big) \quad (4)
\end{aligned}$$

where L is the maximum hop distance from the sink. If no data aggregation is performed ($k_d = 0$), then

$$D_h^T = D^S\Big(1 + \frac{L^2 - h^2}{2h-1}\Big). \quad (5)$$

Furthermore, D_h^R can be obtained from (4). That is,

$$\begin{aligned} D_h^R &= N_h^C D_{h+1}^T \\ &= D^S\big(\frac{2h+1}{2h-1}\big) \\ &\Big(1 + \frac{(1-k_d)^{-h}}{2h+1}\sum_{i=h+1}^{L-1}((1-k_d)^i(2i+1))\Big) \end{aligned} \quad (6)$$

The D_h^T and D_h^R given by (4) and (6) represent the average amount of data that are received and transmitted by a node in h hop distance per round in data gathering. They are the key factors to determine the average energy consumption rate of nodes at a certain hop distance. Then, the energy consumption rate for data gathering communication (E_h^G) at h hop distance can be expressed as

$$E_h^G = E^T D_h^T + E^R D_h^R, \quad (7)$$

where E^T and E^R are the energy amounts required for transmitting and receiving unit data.

Fig. 2 shows the average amount of data to be communicated by a node at a different hop distance with several data aggregation degrees. Each node is assumed to generate one data packet per round and the maximum hop distance (L) is set to 6. As shown in this figure, when no data aggregation (i.e. $k_d = 0$) is performed during data gathering, there is a significant amount of data to be communicated and processed by the first hop nodes, leading to much faster energy depletion at the first hop nodes while the majority of other nodes are still alive. When more and more data aggregation is performed, the node aging rate due to energy consumption at a different hop distance becomes closer. This is due to the significant reduction of data communication cost in nodes at a hop distance that is closer to the sink.

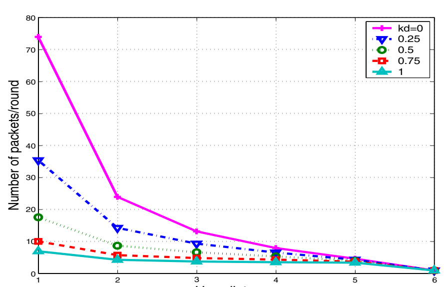


Fig. 2. The amount of data received and transmitted by a node at various hop distances with several data aggregation degrees.

The node aging process at the same hop distance is clearly dependent on the maximum hop distance from the sink (L) and data aggregation degree (k_d). This effect is especially obvious for the first hop nodes. In Table I, we show the effect of the maximum hop distance (L) and the data aggregation degree (k_d) on the first hop node.

TABLE I
THE EFFECT OF THE MAXIMUM HOP DISTANCE (L) AND THE DATA AGGREGATION DEGREE (k_d) ON THE DATA COMMUNICATION RATE (THE NUMBER OF PACKETS PER ROUND) OF THE FIRST HOP NODES.

Data	$k_d = 1$	$k_d = 0.75$	$k_d = 0.5$	$k_d = 0.25$	$k_d = 0$
L=2	4.00	4.75	5.50	6.25	7.00
L=3	4.00	5.38	8.00	11.88	17.00
L=4	4.00	5.59	9.75	17.78	31.00
L=5	4.00	5.66	10.88	23.48	49.00
L=6	4.00	5.69	11.56	28.70	71.00

III. CONNECTIVITY OF DATA GATHERING TREES FOR AGING NETWORKS

Following [4], a data gathering tree is formed by communication paths from each sensor device to a sink by allowing each node to select a parent node in the upper hop level. In a dynamic data gathering tree, if the parent node dies or has no connection to the sink, then the node searches for the new parent node among candidate parent nodes that are still functioning. As the radio range increases, the node would have a larger number of candidate parents.

The sensor node survival function ($S_i(t)$) characterizes the node aging process in a data gathering tree. It is defined to be the probability that node i is functioning at time t (data gathering rounds). As discussed in the previous section, the function is primarily dependent on the energy consumption rate and device reliability of a node. Initially, when $t = 0$, $S_i(0) = 1$. For $t > 0$, $S_i(t)$ is expressed as

$$S_i(t) = \begin{cases} 0, & if\ S_i(t-1) = 0\ or \\ & E_i - \sum_{k=1}^{t}(E_i^G(k) + E_i^{NG}(k)) < 0 \\ R_i(t), & otherwise \end{cases} \quad (8)$$

where E_i is the initial battery energy, E_i^G is the energy required for data gathering communication per round from (7), and E_i^{NG} is the energy required for all other operations for node i, and $R_i(t)$ is the reliability function.

Let $C_i^h(t)$ be the event that node i with h hop distance away from the sink on the data gathering tree is connected to the sink at time t. We use p_i^{h-1} to refer to any candidate ascendent nodes in hop distance $h-1$ of node i in a dynamic data gathering tree. Then, we have

$$Pr(C_i^h(t)) = \begin{cases} S_i^h(t), & h = 1, \\ Pr(\bigcup C_{p_i^{h-1}}^{h-1}(t))S_i^h(t), & h \geq 2. \end{cases} \quad (9)$$

As given above, the probability that a node is connected to the sink is equal to the probability that any of the candidate parents is connected and the node is alive. The lower bound of (9) is the case where there is only one available parent for all nodes in the forwarding path:

$$Pr(C_i^h(t)) \geq (\prod_{k=1}^{h-1} S_{p_i^k}^k(t))S_i^h(t), \quad h > 1 \quad (10)$$

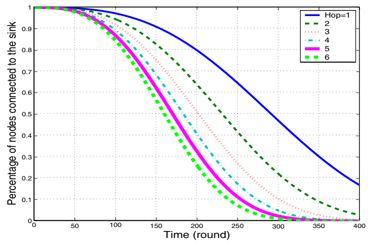


Fig. 3. The analytical lower bound of node connection at each hop distance due to the device failure effect.

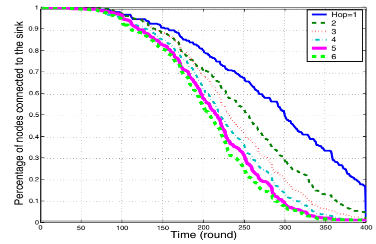


Fig. 4. The percentage of node connection at each hop distance due to the device failure effect.

Fig. 3 shows the lower bound of the connectivity probability in a data gathering tree at each hop distance according to (10). This theoretical lower bound will be compared with experimental results in Fig. 4 in Sec. IV.

IV. Simulation Results

In this section, the node aging effect on the connectivity of a data gathering tree is examined by considering effects of reliability and the energy consumption rate. For the energy consumption rate, we focus on the upper bound of connectivity by applying the perfect data aggregation ($k_d = 1$) during data gathering.

The simulation was set up as follows. The network consists of 400 nodes distributed uniformly at random and the sink is located at the center of the network. All simulation results are obtained as the average of 20 different random node deployments and each simulation runs for 400 rounds. For each random deployment, nodes construct the data gathering tree using the localized parent selection scheme as described in [4]. The radio range is set to provide 12 neighboring nodes within the range on the average. To understand the energy consumption rate effect on connectivity, different data aggregation degrees have been evaluated.

A. Effect of Device Reliability

To study the reliability effect on the connectivity, all nodes have the uniform survival function that follows the Weibull reliability function with $\eta = 330$ and $\beta = 3$ in (2). We observe that the simulation results given in Fig. 4 and the lower bound analytical results given in Fig. 3 match quite well. We see that, as nodes age over time, the gap between the connectivity at a shorter hop distance and that at a longer hop distance increases since the connectivity loss rate of a longer hop distance is higher than that of a shorter hop distance.

B. Effect of Data Aggregation Degree

The impact of the data aggregation degree on connectivity is studied. We show the decreasing behavior of the number of nodes connected to the sink at each hop distance over time in Figs. 5 and 6, with perfect and no data aggregation, respectively.

To find the upper bound of connectivity to the sink at each hop distance affected by energy depletion of nodes, we present the simulation results with perfect data aggregation ($k_d = 1$) case in Fig. 5. As the data aggregation degree decreases (i.e. k_d approaches to zero), significant connectivity loss is observed at much earlier time due to energy depletion as indicated by Fig. 6. This early sharp reduction of connectivity is caused by a much faster energy consumption rate of the first hop nodes as discussed in Section II-B. As compared with Fig. 5, the connectivity loss rate with $k_d = 0$ is over 10 times higher than that with $k_d = 1$. In addition, from these figures, we observe that the connectivity loss rate becomes faster as the hop distance increases.

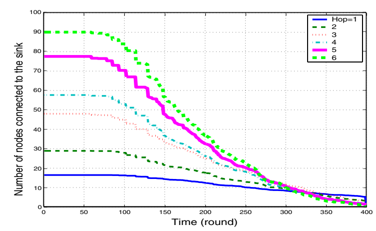


Fig. 5. Connectivity reduction over time at each hop distance with perfect data aggregation ($k_d = 1$) performed.

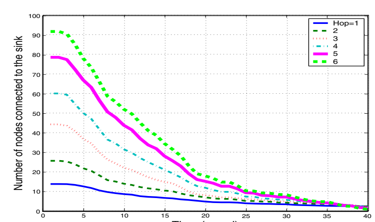
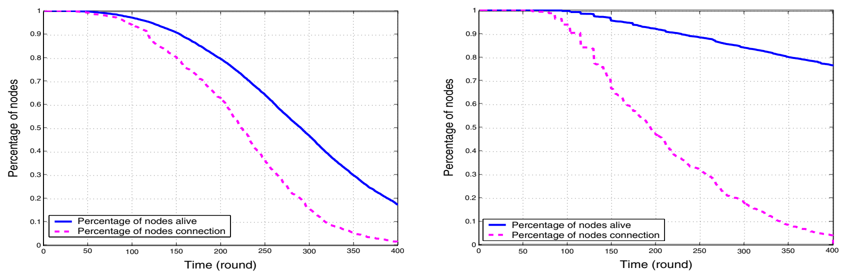


Fig. 6. Connectivity reduction over time at each hop distance, where no data aggregation ($k_d = 0$) is performed and the time scale is 1/10 of that in Fig. 5

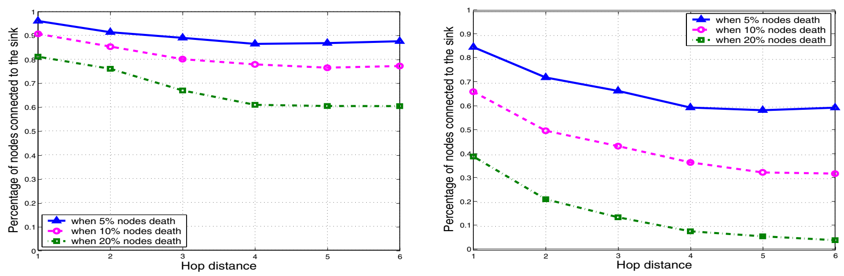
C. Node Aging and Connectivity Aging

Fig. 7 presents the node aging and connectivity aging (*i.e*, how connectivity degrades over time), cased by both device failure and energy depletion. The main difference of these two effects on the connectivity is that the energy depletion case shows a much larger gap between the node aging rate and the connectivity aging rate. This is because the survival rates among each hop distance are different even with perfect data aggregation ($k_d = 1$), In contrast, if reliability is the dominant factor that determines the node aging, the survival rate that follows the device reliability function would be



(a) (b)

Fig. 7. The percentage of nodes that are alive and the percentage of nodes that are connected to the sink due to (a) the device failure effect and (b) the energy depletion effect, where perfect data aggregation is performed.



(a) (b)

Fig. 8. The percentage of node connection at each hop distance when 5%, 10%, and 20% of nodes die due to (a) the device failure effect and (b) the energy depletion effect, where perfect data aggregation is performed.

uniform among all nodes. Faster energy depletion of the first hop nodes leads to the connectivity loss of nodes at a longer hop distance while the majority of nodes are still functioning with remaining energy as shown in Fig. 7 (b).

D. Connectivity Aging at Different Hop Distances

Connectivity aging at each hop distance is shown in Fig. 8. First, higher connectivity is maintained for the device failure effect as compared to that due to the energy depletion effect as nodes age over time. Second, the difference connectivity among each hop distance is more significant in the energy depletion case than in the device reliability case. The reason of lower connectivity and a larger gap in connectivity among different hop distances in the energy depletion case is that most of node death occurs in a closer hop distance from the sink due to faster energy depletion. As the data aggregation degree becomes smaller, the larger portion of node death comes from the first hop nodes, leading to significant connectivity loss of nodes at a longer hop distance as discussed in Section IV-B.

V. Conclusion and Future Work

The node aging process due to device reliability and the energy consumption rate with a different data aggregation degree in a data gathering tree was investigated in this work. The connectivity degrading process over time at each hop distance was examined with both analytical and simulation results. We showed a significant impact of the node aging process on connectivity of a multi-hop data gathering paths.

A good understanding of the node aging process and the impact on the connectivity on a data gathering tree over time would help provide efficient node deployment, topology control and data gathering strategy to achieve energy-saving so as to enhance the network lifetime. We plan to extend the current work to heterogeneous deployment in the near future. In addition, we will examine the network aging problems with data communication patterns and network operations other than the periodic data gather tree application.

Acknowledgment

This work was supported in part by NSF CAREER grant number 0347621.

References

[1] P. Gupta and P. R. Kumar, "Critical power for asymptotic connectivity in wireless network," in *Stochastic Analysis, Control, Optimization and Applications*. Birkhauser, Boston, 1998, pp. 547–566.
[2] C. Bettstetter, "On the minimum node degree and connectivity of a wireless multihop network," in *Proc. ACM MOBIHOC'02*, 2002.
[3] S.Shakkottai, R. Srikant, and N. Shroff, "Unreliable sensor grids: Coverage, connectivity and diameter," in *Proc. IEEE INFOCOM'03*, San Francisco, USA, Apr. 2003.
[4] C. Zhou and B. Krishnamachari, "Localized topology generation mechanisms for self-configuring sensor networks," in *Proc. IEEE GLOBECOM'03*, San Francisco, USA, Dec. 2003.
[5] L. M. Leemis, *Reliability: Probabilistic Models and Statistical Methods*. Englewood Cliffs, NJ: Prentice-Hall, 1995.
[6] B. Krishnamachari, D. Estrin, and S. Wicker, "The impact of data aggregation in wireless sensor networks," in *International Workshop on Distributed Event-Based Systems*, 2002.
[7] J. Lee, B. Krishnamachari, and C. C. Kuo, "Impact of energy depletion and reliability on wireless sensor network connectivity," in *Proc. SPIE Defense & Security*, Apr. 2004.
[8] V. Raghunathan, C. Schurgers, S. Park, and M. B. Srivastava, "Energy-aware wireless microsensor networks," *IEEE Signal Processing Magazine*, vol. 19, pp. 40–50, 2002.