

Impact of Network Dynamics on Tardiness of Data in Sensor Networks

Tarun Banka, and Anura P. Jayasumana

{tarunb, anura}@enr.colostate.edu

Department of Electrical and Computer Engineering, Colorado State University,
Fort Collins CO 80523 USA

Abstract - Impact of random delays and losses in sensor networks manifests in the form of tardiness of data used for processing at the sink nodes. The age of data used by the end application can impact the accuracy of the end results, and may produce detrimental consequences for many real-time sensing applications. This paper uses a tardiness measure for quantitatively capturing the lateness of the data due to network dynamics, and presents an analytical model relating the network delay, network packet loss rate, packet reordering, and sampling rate to the tardiness. We extend this model to provide aggregate weighted tardiness of data at fusion nodes. The tardiness model is validated using simulation results. We also investigate the tradeoffs between energy consumption and tardiness when desired tardiness is achieved by adjusting sampling rate and transmission power of the sensors nodes. Other potential applications of proposed tardiness measure includes determination of the active/sleep period of MAC layer to meet application goals, and comparison of routing protocols based on their impact on tardiness, and hence the application in real-time sensor networks.

Keywords: Freshness, Tardiness, Sensor Networks, QoS

1. Introduction

Sensor networks are an emerging class of networks with a wide range of practical applications in the real-world. The well-known applications include structure and environmental monitoring, and target tracking [1,15]. In many applications, the sensor network collects information about a physical phenomenon, processes the data and then takes appropriate actions based on the processed data [2]. In many such applications, action should be taken in bounded time for the proper operation of the system. Input data may be useless for such applications if it arrives and is processed after a critical deadline. Therefore, it is important to be aware of the age of the data that is used for processing and computing results. Data freshness has been studied in the context of information system such as data integration system (DIS), and Data Warehouse [4]. In that context, data freshness is considered as a critical component that determines the success of many information systems. A significant amount of research has been done for studying efficient refresh policies for web crawlers to keep the local copies of the remote source data fresh [6]. The key factor that impact the age of the data in such traditional information systems is the rate of change of data at the remote source node. However, in sensor networks besides the rate of change of data some of the other factors that may impact the age of the data are the high network delays, random loss of packets, and packet re-ordering. It is significantly more challenging to quantify and understand the impact of each of

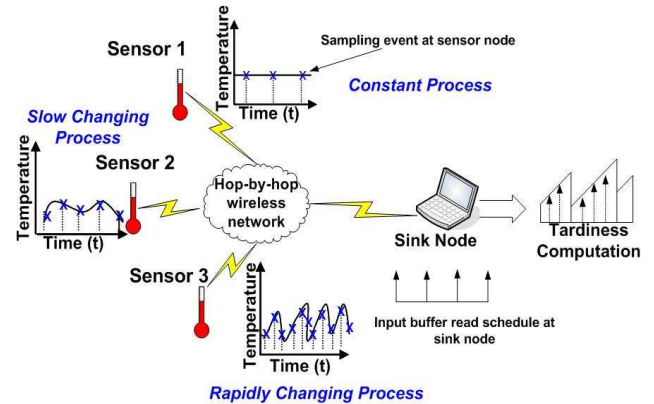


Figure 1. Process monitoring sensor network

these parameters on the age of the data in sensor networks. Our focus on understanding impact of network dynamics on the age of the data makes it different from the existing work that is focused on investigating freshness of data in the context of information systems.

Sensor networks are typically resource constrained in terms of computation capability and available energy. Moreover, most sensor networks use error prone wireless links for communication. Different energy conserving transport, routing, and MAC protocols are used for transferring data from source nodes to the sink nodes [11,12,19,21]. Alternatively, there is an emerging class of real-time sensor networks that uses Internet to distribute sensor data to different computing nodes and end users. Depending on the applications, sensor nodes may be configured to monitor the environment for rare and ephemeral events [7,17]. Whenever such events occur, they are detected by one or more sensing nodes, and the event information is transmitted to a data fusion node or a remote sink node for further processing. Alternatively, in many sensor networks, sensor nodes continually sample environmental processes such as temperature or vibrations as shown in Fig. 1 [8]. Sink nodes such as fusion nodes may access receive buffer for the data after receiving an interrupt due to arrival of a new packet. Alternatively, sink nodes may poll their receive buffers recurrently for the arrival of the data. In both scenarios, data available in the buffer at the time of 'read' is used for the computation. However, it is possible that data read from the input buffer at the sink node may be different from what is available at a particular instant of time at the sensor node. This scenario is possible because either the data gets late within the sensor network due to network delays [5,22] or gets dropped because of reasons such as network congestion or wireless link

errors in the sensor network [11]. When data is dropped in the network, the prior sample of data available is used until new data arrives. In many closed-loop applications, old copy of the data that does not reflect the current state of the environment may be used to generate actuating signal to control the remote environment. This has the potential to compromise the integrity of such systems. It is therefore desired to get a quantitative estimate of age of the data used in the computation at every ‘read’ event at the sink node.

This paper develops a framework to capture network introduced tardiness of data used for end computations in sensor networks. A tardiness measure is defined to capture the age of the data used for computation under different network conditions. An analytical model is proposed that relates network delays, wireless loss rate, degree of packet re-ordering, and sampling rate with the observed tardiness of the data. Moreover, we study the tradeoffs between energy consumption and tardiness of the data delivered to the end user. Such an analysis will abstract the impact of the sensor network characteristics, such as losses and delay due to routing scheme employed, in terms of statistics of tardiness. The statistical characteristics of tardiness may then be used to evaluate the accuracy and reliability of the application, without delving into detailed network characteristics such as the routing protocol. We envision wider applications of the tardiness measure. Tardiness may be used to evaluate and compare the performance of routing protocols in terms of age of the data delivered to the sink node, adaptive sampling techniques [9,16], effect of network topologies in meeting real-time requirements of the applications [10]. This work may also help in configuration of sampling rates, transmission energy, sleep/active schedules at MAC layer, and input buffer read frequency at sink nodes for minimizing the error in the end results due to tardiness of the data. Section 2 describes the tardiness measure and analytical model. Section 3 presents results for the model verification. Section 4 discusses tradeoffs between tardiness and energy consumption in sensor networks. Conclusions are presented in Section 5.

2. Tardiness Measure

The difference between an ideal monitoring system, in which the processing/decision node has instantaneous access to the values/events/parameters being monitored, and a distributed sensor network lies in the age or tardiness of data available at the processing node. With latter, the available data from different sensor nodes have different ages, which depend on network characteristics and protocols.

For example, consider the scenario where the processing/decision nodes receives data from multiple sensors. Let at time t , $X_1(t), \dots, X_N(t)$ be the data values available at the sink node from N different source nodes about the phenomena under observation. Let $F(t)$ be the processing/decision function that combines the most recent data from N source nodes at time t . For an ideal monitoring system, as described above,

$$F(t) = h(X_1(t), X_2(t) \dots X_N(t))$$

i.e., instantaneous data from all sensors is available at the processing node. However, in distributed sensor network based system,

$$F(t) = h(X_1(t-\Delta_1), X_2(t-\Delta_2) \dots X_N(t-\Delta_N))$$

Where Δ_i corresponds to the tardiness of data from source i to the processing/decision sink node. The Δ_i values form a random process that is affected by the network protocols, losses, sampling frequency, sleep schedule, etc. Characterizing this tardiness process allows the impact of network characteristics to be summarized in a way that its influence on different applications (based on different decision functions) can be evaluated more conveniently.

Tardiness measure captures the age of the data used for computation at the receiver node. Age of the data is defined as the time lag from the time data is generated at the sensor node to the time data is used at the sink node by the application.

2.1. Tardiness under dynamic network conditions

We consider a process monitoring sensor network shown in Fig. 1. In this scenario, sensor node periodically samples the physical environment for the data every ‘S’ time units as shown by the vertical dotted lines in Fig. 2 and Fig. 3. At every sampling instant, generated sample is packetized, and is immediately transmitted over single or multi-hop wireless network towards the sink node. In this example, the sink node periodically accesses the input buffer for the received data every ‘R’ interval as shown by red color dots. For the purpose of tardiness evaluation it is assumed that each transmitted sample is time-stamped at the sender node, and all nodes in the sensor network are time synchronized [18]. In Fig. 2 and Fig. 3, D_i is the network delay suffered by the sample i , and $TD[i,k]$ is the tardiness of the data read at k^{th} successive read attempt when sample i is present in the buffer. Dotted vertical lines with arrows shows the age of the data read from the buffer at a particular input buffer read attempt. We now illustrate how random delay and random losses impact the age of the data at the sink node. Two cases are considered; in the first case packets only suffer random delay and no packet losses. In the second case, packet suffers both random delay and random packet losses.

2.1.1. Tardiness Measure under Random Delay and No Network Packet Loss

Fig. 2 shows the case when samples suffer random delay in the network and does not suffer any network packet loss. Network delay depends on factors such as active/sleep schedule in MAC layer, paths selected by the routing protocol, and node distribution in the sensor network. Under these conditions samples arrive at the sink node after suffering random delays. Depending on the arrival time and the periodic read interval time ‘R’, same sample can be read multiple times by the end application. As shown in the Fig. 2, the age of the data increases linearly with time until the next sample arrives.

**Tardiness of Data in Process Monitoring Sensor Networks
(Case: Random Network Delay)**

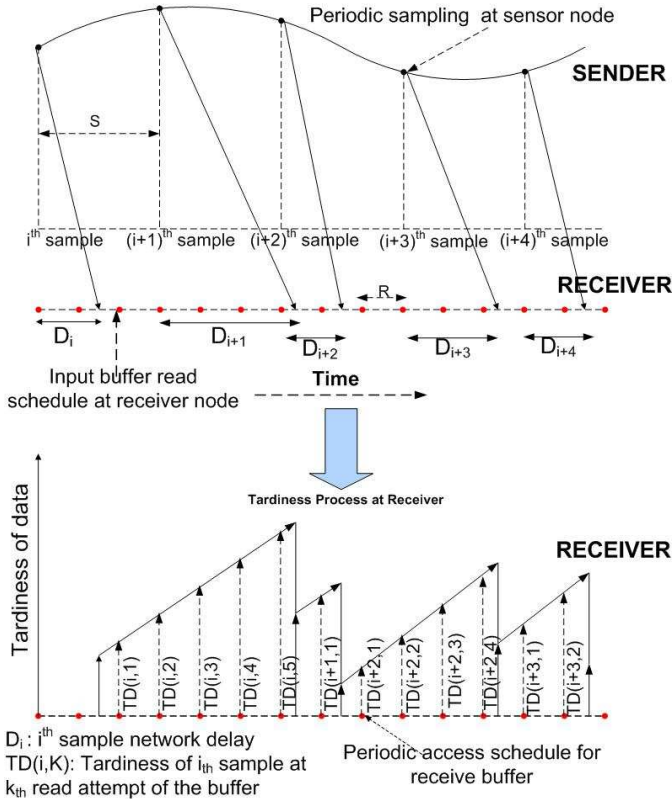


Figure 2. Tardiness measure due to random network delay in process monitoring sensor networks

**Tardiness Measure in Process Monitoring Sensor Networks
(Case: Random Data Loss and Random Delay)**

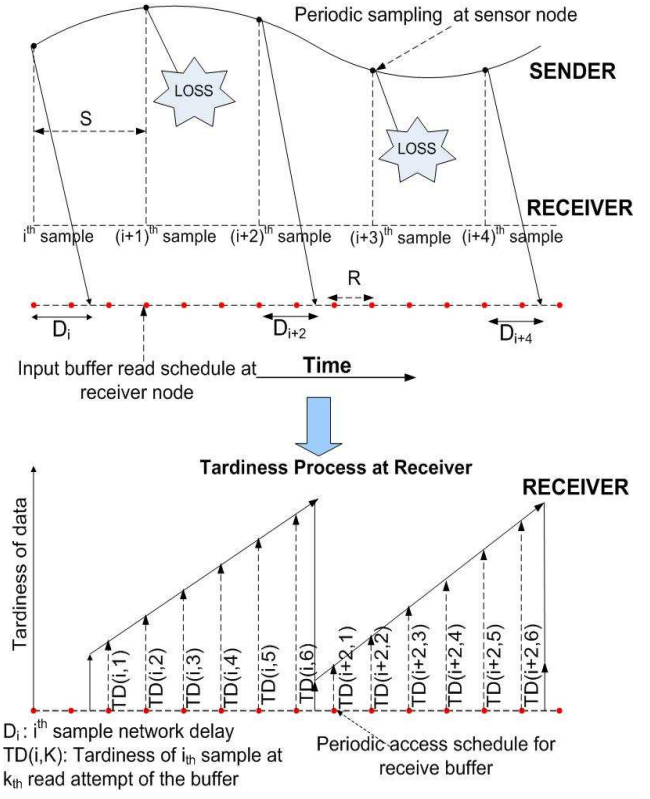


Figure 3. Tardiness measure under random data loss in process monitoring sensor networks

Analytical Model for Tardiness

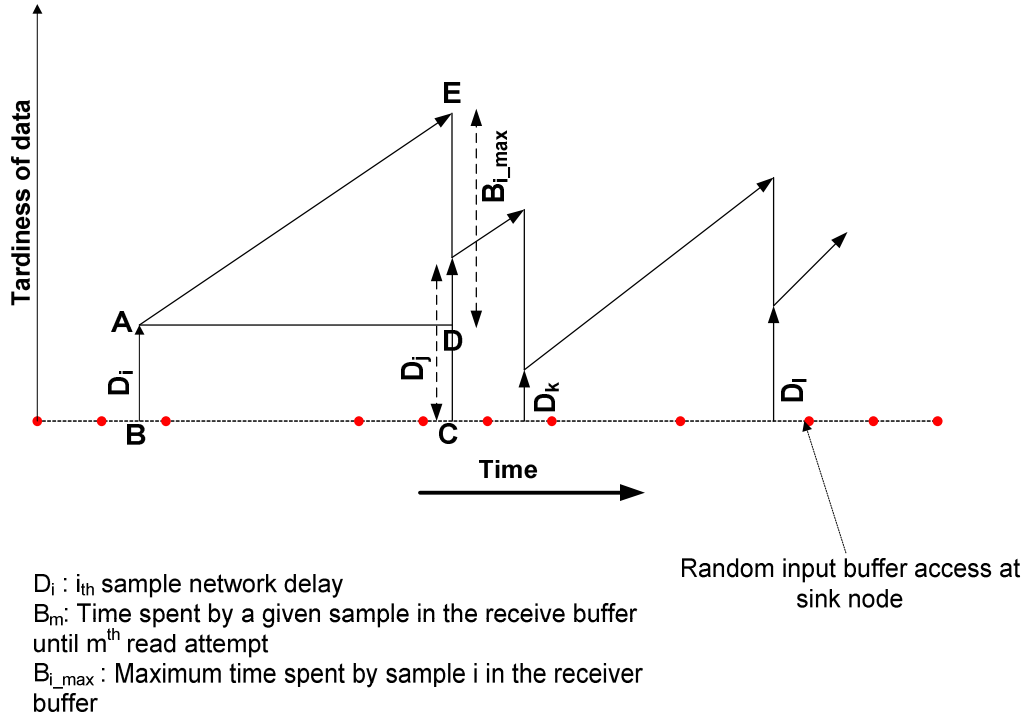


Figure 4. Tardiness Computation at Sink Node, where sample i, j, k and l are sampled at absolute time S_i, S_j, S_k , and S_l at source node (not shown in the figure)

Parameter	Description
D_i	Network delay of sample i
A_i	Arrival time of sample i at sink node
B_m	Time spent by a given sample in the receive buffer until m^{th} read attempt
$B_{i_{\text{max}}}$	Maximum time sample i spends in the receive buffer
P_L	Total packet loss perceived by the application at a sink node
P_N	Network packet loss probability
L	Random variable that models loss characteristics of the network
D	Random variable that model network delays
S	Sampling Interval used at the source node to periodically sample environment
$P_{I/A}$	Total probability that packet is in-order on condition that it arrives at the sink node
$P_{I/A}^{jS}$	Probability that packet is in-order on the condition that it arrives at the sink node and it suffers network delay between $(j-1)S$ and jS
$M(t)$	Total packet arrivals in interval $[0,t]$ at a sink node
W_k	Weight of the data generated by sensor k in the network
$f_d(c)$	Delay distribution of the sensor network
$F_D(c)$	CDF of delay in sensor network
$P_{RR}[d]$	Packet reception rate at distance d from the source node
$\gamma(d)$	SNR at distance d from the source node
f	Frame Length
l	Preamble Length
τ	Absolute time at which input buffer is read and tardiness is computed at sink node

Table 1. List of Parameters for the Analytical Model

Note that age of the data is computed every time the input buffer is read at the receiver node. In Fig. 2, sample $i+1$ suffers higher delay than sample $i+2$ sample. In this case sample i is read multiple times at the scheduled read times until the sample $i+1$ arrives at the sink. Alternatively, sample $i+1$ is read only once as the sample $i+2$ arrives immediately after the first read of the sample $i+1$. Thus in this case tardiness of data read from the buffer is more when sample i is present in the buffer compared to when sample $i+1$ is present in the buffer. Since sampling times at the sensor and read from buffer occur independently, the area under the curve divided by the time interval gives the average tardiness of data from the source.

2.1.2. Tardiness Measure under Random Delay and Network Packet Loss

Fig. 3 illustrates the tardiness process when packets suffer from both random delays and network packet losses. In wireless sensor networks packet losses depend on the network congestion, collisions, and wireless link errors. When a data sample is lost then the prior sample of data is used until the new data arrives at the receiver node. This is not different from the way tardiness is computed in lossless but random delay case in Section 2.1.1. In Fig. 3, sample $i+1$ and sample $i+3$ are randomly dropped in the network. It results in increase in tardiness at successive read as sample i is read multiple time until sample $i+2$ arrives. In Fig. 3, relative larger area of trapezoid indicates greater tardiness compared to Fig. 2.

From Fig. 2 and Fig. 3 it can be inferred that tardiness measure will be impacted by the random delays and the random losses suffered by the data in the sensor network. Intuitively, it is a function of network delay and packet loss probability, read buffer frequency, and sampling frequency. In Section 2.2 we derive an analytical model for the tardiness of the data in process monitoring sensor networks that relates different network parameters to the age of the data.

2.2. Analytical Model for Tardiness of Data in Process Monitoring Sensor Network

In this section we first derive an analytical model for the tardiness of data from a single source node to a single sink node and then extend it for a multi-source to single sink node scenario. Analytical model for tardiness is a function that relates random network delay, random packet loss probability, and sampling interval to the mean age of the data at the sink node. Table 1 shows list of different parameters used for the analytical model. Following assumptions are considered for deriving the tardiness analytical model:

- (i) Sensor node samples the environment periodically every ‘S’ interval and transmit data to a single sink node.
- (ii) Sink node may randomly access the input buffer for data.
- (iii) We consider a case where source node transmits one sample per packet, so sample and packets are used interchangeably in this paper.
- (iv) All nodes in the sensor network are assumed to be time synchronized for measurement purposes. This does not mean that all the nodes carry out sampling at the same instant.
- (v) It is assumed that packets arrives in-order at a sink node.
- (vi) Data may get dropped in the network because of wireless link errors, collisions, or network congestion.

(1) Tardiness of Data from a Single Source:

Fig. 4 illustrates a tardiness computation at the sink node. It considers a scenario when sink node receives samples i , j , k , and l from a single source node without re-ordering after suffering random delays and random network packet losses.

Red solid dots are the random time at which input buffer is accessed for the data at the sink node. Sample 'i' which arrives at sink node at time indicated by 'B' is randomly read from the input buffer until sample 'j' arrives at time 'C' at a sink node. It is assumed that most recently received sample remains in the input receive buffer until next in-order sample arrives at the sink node. At every read schedule, tardiness of the data read from the buffer is evaluated. Average height of a trapezoid in Fig. 4 indicates the average tardiness during each consecutive read attempt until next in-order packet arrives.

In Fig. 4, let D_i be the random delay suffered by the sample i in the network. Let B_m be the total time spent by the latest sample in the input receive buffer at sink node at the time of m^{th} periodic read. Consider a case when the latest sample read is i . Let $TD(i,m)$ be the instantaneous tardiness of data read during m^{th} read attempt when sample i is present in the input receive buffer. Let τ be the absolute time at which m^{th} read of input buffer is performed. Then the tardiness of data during read operation performed at time τ is $T(\tau)$, i.e.,

$$T(\tau) = TD(i, m) = D_i + B_m \quad (1)$$

For each arriving sample i at the sink node $0 \leq B_m \leq B_{i_max}$, where B_{i_max} is the maximum time spent by sample i in the buffer until a sample $j > i$ arrives. Thus the tardiness experienced at the sink changes with time as shown in Fig. 4.

By the application of the law of large numbers, we can assume that long-term time averages can be replaced by the ensemble averages [3]. We can compute mean tardiness using graphical argument. Using Fig. 4, time average of the tardiness $T(\tau)$ in the interval $[0, t]$

$$\frac{1}{t} \int_0^t T(\tau) d\tau = \left(\frac{M(t)}{t} \right) \left(\frac{\sum_{i=0}^{M(t)} B_{i_max} D_i + \frac{B_{i_max}^2}{2}}{M(t)} \right) \quad (2)$$

where $M(t)$ is the total sample arrivals in the interval within $[0, t]$ for which tardiness is computed. Maximum time B_{i_max} that sample i stays in the buffer is determined by the difference in the arrival time of the sample i and sample j where $j > i$.

$$B_{i_max} = (S_j - S_i) + (D_j - D_i) \quad (3)$$

Where S_i and S_j are the sampling time at the source node and D_i and D_j are the network delays of sample i and sample j respectively. Under certain conditions such as no network packet losses and when delays are within certain bounds, $j = i + 1$, i.e., adjacent samples are received at the sink node without re-ordering or loss. Replacing B_{i_max} in Eq. 2 with Eq. 3 and taking $\lim_{t \rightarrow \infty}$, we get

$$\lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t T(\tau) d\tau = E[T] = \mu \left(\frac{E[L^2]}{2} + E[L]E[D] \right) \quad (4)$$

where $L = (j-i) \cdot S$, such that S is the periodic sampling interval and i and j are the adjacent samples received at the sink node where $j > i$. L is a r.v. that models loss characteristics of the network. D is a r.v. that models random delays of different samples that arrive at sink node and μ is the arrival rate of the samples at sink node. Let P_L be the packet loss probability as perceived by the sink node. Then we can determine expectation of mean time interval between two losses and its second moment at the sink node as follows:

$$\begin{aligned} E[L] &= \sum_{i=1}^{\infty} i(P_L^{i-1})(1-P_L)S \\ &= \frac{S}{1-P_L} \end{aligned} \quad (5)$$

$$\begin{aligned} E[L^2] &= \sum_{i=1}^{\infty} i^2 P_L^{i-1} (1-P_L) S^2 \\ &= (1+P_L) \left(\frac{S}{1-P_L} \right)^2 \quad \text{and} \end{aligned} \quad (6)$$

Reciprocal of $E[L]$ given by Eq. 5 is the mean arrival rate μ , i.e.,

$$\mu = \frac{1-P_L}{S} \quad (7)$$

Combining Eq. (4)–(7), we can determine expectation of the tardiness T , i.e., mean tardiness of the data is:

$$E[T] = E[D] + \left(\frac{1}{2} \right) \left(\frac{1+P_L}{1-P_L} \right) S \quad (8)$$

From Eq. 8, we conclude that tardiness of the data at each read attempt depends on the sampling rate, network delay characteristics and network packet loss probability. Increase in network delay, network packet losses and sample time period results in an increase in the tardiness of the data in sensor networks.

(2) Aggregate Tardiness of Data from Multiple Sources to a Single Sink:

The model for tardiness given by Eq. 8 is valid for tardiness for the source-sink pair. In sensor networks, many-to-one data transfer is the common data flow scenario. We consider a case when sink node such as fusion node in the network receives data from multiple sources. As tardiness associated with different sources may not be similar, it is necessary to evaluate the aggregate tardiness of the result computed using inputs from multiple sources. Let there be N sensors from where data is aggregated at the sink node. Let W_k be the weight of the data from the sensor k . The weight can be assigned to each sensor according to the criticality of the data generated by each sensor node. When the data gathered from different sensors are equally important, equal weights may be used. Then aggregate weighted tardiness is defined as:

$$E[T_{aggregate}] = \frac{\sum_{k=1}^N E[T_k] \cdot W_k}{\sum_{k=1}^N W_k} \quad (9)$$

2.3. Consideration for Re-ordered Packets at a Sink Node

In a sensor network packets may arrive out-of-order because of random delays suffered by them in the network. Depending on the routing algorithm, different packets may follow different paths thus may suffer variable delays. For certain real-time applications, out-of-order packet arrival may not be acceptable and are treated as lost packets. For example, in real-time target tracking applications, it is important to have most recent estimate of the position of the target for real-time tracking and prediction. Therefore, all late arrivals of packets with old information may not be of any use to the application and can be treated as lost. In this section, we consider impact of re-ordered packets on the tardiness of data and adapt the tardiness model given by Eq. 8 to provide accurate estimate of tardiness in presence of packet reordering.

There are two steps involved in developing model that consider re-ordering of packets (i) Estimation of packet loss probability, i.e., P_L perceived by the application at the sink node, and (ii) Estimation of mean delay of the packets that arrive in-order at a sink node, i.e., $E[D|I]$ which is conditional expectation of the delay given that packet arrives in-order.

(i) **Estimation of P_L :** When packet-reordering is not considered then P_L is equal to the network packet loss probability P_N . However, when packets are treated as lost because of out-of-order arrival then packet loss probability is:

$$P_L = 1 - p_{I/A} (1 - p_N) \quad (10)$$

where $p_{I/A}$ is the probability that packet is in-order given that it has arrived at the sink node. P_N is the probability that packets are lost in the network because of wireless link errors. It is assumed that network is not congested and there are no collisions. We estimate conditional probability of in-order arrival, i.e., $p_{I/A}$ as follows:

Consider packet i , $i+1$, and $i+2$ in the order of generation at a source node. Let S_i , S_{i+1} , and S_{i+2} be their generation time such that

$$S_{i+1} - S_i = S_{i+2} - S_{i+1} = S \quad (11)$$

Let D_i be the random network delay suffered by the i^{th} packet then its arrival time A_i is:

$$A_i = S_i + D_i \quad (12)$$

By definition, packet i arrives in order if $A_i < A_{i+j}$ where $1 \leq j \leq N$ where N is the total number of arrivals after packet i . A delay distribution of a network provides an estimate of delay suffered by different packets transmitted from the sensor node. Given the delay suffered by a packet generated by the sensor, then that packet arrive in-order if all future packets delays are such that they arrive after the current packet. Note that we

consider a scenario when a sensor node periodically generates a packet every S time interval.

Let $f_d(c)$ be the delay distribution of a sensor networks and let P_N be the network packet loss probability because of wireless link errors. Consider different possible ranges of delay of size S that a given packet i may suffer in the network, and for each possible delay range, probability of the packet i arriving in-order is computed as follows:

Case $0 \leq D_i < S$: In this case packet i is not re-ordered, because packet i arrives at the sink node before packet $i+1$ and subsequent packets are generated at the source node. Then probability of in-order arrival of packet i , when packet delay is between 0 and S is:

$$p_{I/A}^S = \int_0^S f_d(c) dc \quad (13)$$

Case $S \leq D_i < 2S$: In this case, packet i suffers delay between S and $2S$. In this case packet i arrives in-order if either packet $i+1$ is lost in the network or when packet $i+1$ arrives at sink node such that arrival time $A_i < A_{i+1}$. Note that all other future packets other than $i+1$ will always arrive after packet i as they are generated after the worst case arrival time of packet i , i.e., when delay suffered by packet i approaches $2S$. Therefore, probability of packet i arriving in-order when its delay is between S and $2S$ is:

$$\begin{aligned} p_{I/A}^{2S} &= p_N \int_S^{2S} f_d(c) dc + (1 - p_N) \int_S^{2S} p(D_{i+1} > c - S) f_d(c) dc \\ &= p_N \int_S^{2S} f_d(c) dc + (1 - p_N) \int_S^{2S} (1 - F_D(c - S)) f_d(c) dc \end{aligned} \quad (14)$$

where $F_D(a)$ is the CDF of the r.v. D that models delay in the network.

Case $2S \leq D_i < 3S$: Similarly i^{th} sample arrives in-order when both $i+1$ and $i+2$ samples are dropped in the network or if they arrive then their arrival time is greater than A_i . Therefore probability of in-order arrival after considering all possible combinations of arrival and loss of $i+1$, $i+2$ samples is given by:

$$\begin{aligned} p_{I/A}^{3S} &= p_N^2 \int_{2S}^{3S} f_d(c) dc + p_N (1 - p_N) \int_{2S}^{3S} (1 - F_D(c - S)) f_d(c) dc + \\ & p_N (1 - p_N) \int_{2S}^{3S} (1 - F_D(c - 2S)) f_d(c) dc + \\ & (1 - p_N)^2 \int_{2S}^{3S} (1 - F_D(c - S)) (1 - F_D(c - 2S)) f_d(c) dc \end{aligned} \quad (15)$$

Case $k \leq D_i < (k+1)S$: For a general case, packet i arrives in order if for all possible combinations of arrival and loss of future k packets, the arrival time $A_i < A_{i+j}$ where $1 \leq j \leq k$. For lost packets, arrival time is treated as infinite. Therefore

$$\begin{aligned}
P_{I/A}^{(k+1)S} &= P_N^k \int_{ks}^{(k+1)s} f_d(c) dc + P_N^{k-1} (1-P_N) \int_{ks}^{(k+1)s} \alpha(1-F_D(c-S)) f_d(c) dc + \\
&P_N^{k-1} (1-P_N) \int_{ks}^{(k+1)s} \alpha(1-F_D(c-2S)) f_d(c) dc + \dots \quad (16) \\
&+ (1-P_N)^k \int_{ks}^{(k+1)s} \alpha(1-F_D(c-S))(1-F_D(c-2S)) \dots (1-F_D(c-kS)) f_d(c) dc
\end{aligned}$$

We can thus conclude that conditional probability that packet is in-order given that it has arrived is given by summation of all terms given by Eq. 13-16:

$$P_{I/A} = \sum_{j=1}^{\infty} P_{I/A}^{jS} \quad (17)$$

Next step is to estimate network packet loss probability P_N using realistic wireless link loss model. We use the model presented in [23], which provides packet reception rate as a function of distance from the transmitter when Manchester encoding and NCFSK modulation schemes are used, i.e.,

$$P_{RR}[d] = \left(1 - \frac{1}{2} \exp\left(-\frac{\gamma(d)}{2} \frac{1}{0.64}\right)\right)^{16f-8l} \quad (18)$$

where $\gamma(d)$ is the SNR at a distance d between source node and the sink node, f is the frame size, and l is the preamble length. SNR at a distance d is a function of transmission power. Using Eq. 18, network packet loss probability between source node and sink node separated by distance d is determined as:

$$P_N = 1 - P_{RR}[d] \quad (19)$$

Substituting $P_{I/A}$ and P_N from Eq. 17, and Eq. 19 in Eq. 10m we get total packet loss probability P_L due to packet reordering and network losses, i.e.,

$$P_L = 1 - \left(\sum_{j=1}^{\infty} P_{I/A}^{jS} \right) (P_{RR}[d]) \quad (20)$$

(ii) Estimation of mean delay of packets given that they arrive in-order $E[D|I]$: In tardiness model given by Eq. 8, $E[D]$ is the expectation of the delay of all the packets that arrive at the sink node irrespective of their order of arrival. However, for applications that cannot tolerate reordered packets, the delay of reordered packets should not be included in the estimation of the expected delay at the sink node. However when we estimate expected delay based on the delay distribution of a network, it also includes delay for packets that arrive out-of-order at the sink node. Therefore, given the delay distribution $f_d(c)$ of the network, it is necessary to include correction to estimate expected delay for all packets that arrive in-order. We can estimate the conditional expected delay $E[D|I]$ given that packet arrives in-order with the similar approach used for computing $P_{I/A}$ given by Eq. 17. We get

$$E[D | I] = \frac{\sum_{j=1}^{\infty} P_{D/I}^{jS}}{P_{I/A}} \quad (21)$$

where

$$\begin{aligned}
P_{D/I}^{jS} &= P_N^{j-1} \int_{(j-1)s}^{jS} f_d(c) dc + P_N^{j-2} (1-P_N) \int_{(j-1)s}^{jS} \alpha(1-F_D(c-S)) f_d(c) dc + \\
&P_N^{j-2} (1-P_N) \int_{(j-1)s}^{jS} \alpha(1-F_D(c-2S)) f_d(c) dc + \dots \quad (22) \\
&+ (1-P_N)^{j-1} \int_{(j-1)s}^{jS} \alpha(1-F_D(c-S))(1-F_D(c-2S)) \dots (1-F_D(c-(j-1)S)) f_d(c) dc
\end{aligned}$$

Eq. 20 and Eq. 21 gives the new estimates of packet loss probability and the conditional expectation of delay respectively that considers impact of packet reordering. Now mean tardiness as given by Eq. 8 can be written as follows that is valid for estimating mean tardiness for applications that does not tolerate reordered packets. Therefore, new estimate of mean tardiness is:

$$E[T] = E[D | I] + \frac{1}{2} \left(\frac{2 - \left(\sum_{j=1}^{\infty} P_{I/A}^{jS} \right) (P_{RR}[d])}{\left(\sum_{j=1}^{\infty} P_{I/A}^{jS} \right) (P_{RR}[d])} \right) S \quad (23)$$

3. Verification of Analytical model for Tardiness

Analytical model for estimation of mean tardiness given by Eq. 23 for the tardiness measure is validated using simulation results. The model considers impact of random network packet losses, network delay, packet reordering, and periodic sampling interval on the tardiness of the data between a source-sink pair. Source node is configured to generate samples periodically after S interval. We consider uniform network packet loss probability P_N and exponential distribution for the packet delays with mean delay D . Fig. 6 compares the simulation and tardiness model results for single source case under different network packet loss and delay conditions. For these experiments we consider an application that does not tolerate packet re-ordering, i.e., one that uses most recent measurement available in the input receive buffer, and all late out-of-order packets are treated as lost. In Fig. 6, model is verified for three cases 1-3 in order of increase in packet re-ordering. Intuitively, degree of re-ordering depends on the standard deviation of the delays suffered by packets between source-sink pair. Case 1 is an example of no packet re-ordering where sampling interval $S = 5.0$ seconds and standard deviation and mean delay is 0.1 seconds. However for Case 2, packet reordering increases where sampling

interval $S = 5.0$ seconds and mean delay and standard deviation of exponential delay distribution is 20.0 seconds. We consider high network delays based on the observation that in sensor networks age of the data can be in tens of seconds [25]. For Case 3, $S=5.0$ seconds, and mean delay and standard deviation of delay is 50.0 seconds. Case 3 corresponds to highest degree of packet re-ordering and Case 1 corresponds to no packet reordering. As seen in Fig. 6, simulation results and model results for tardiness are in close agreement with each other for all three cases of varied degree of reordering. For case 1 of no packet reordering, simulation and model results overlap each other. As seen in the figure, for a given network packet loss probability P_N packet reordering may lead to significant increase in the tardiness of the data. Fig. 7 corroborates the impact of standard deviation of the network delay on the fraction of the packets that arrives out-of-order. For a given network packet loss probability, fraction of packets that arrive out-of-order, increases with increase in standard deviation. Note that exponential delay distribution is considered in Fig. 7. In the second set of experiment for model verification, sampling interval S is varied for a given network packet loss probability and mean delay, and its impact on the tardiness is studied for a source-sink pair. Fig. 8 shows the simulation and model results of tardiness under varying sampling interval S . In this experiment $P_N = 0.0$, and mean exponential delay $D=20.0$ seconds. As seen in the figure, tardiness increases with the increase in sampling interval and model results are in close agreement with the simulation results.

4. Tradeoffs between Energy Consumption and Tardiness of Data

Most of the sensor networks are energy constrained. In this section, we investigate the tradeoffs between energy consumption and the tardiness of the data. As seen in Section 3, and Section 4 packet loss probability, network delays, and

Tardiness Model Verification under different Network Loss and Network Delay
 $S=5.0$ sec. and Exponential Delay Distribution
 Case 1: Mean Delay $D=0.1$ sec., Case 2: Mean Delay=20.0 sec.,
 Case 3: Mean Delay=50.0 sec.

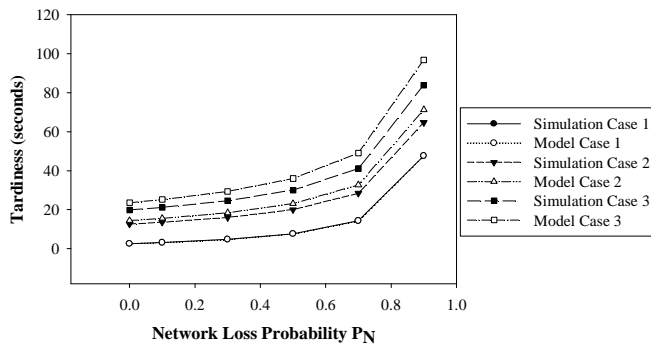


Figure 6. Verification of Tardiness model under for different Network Loss Rates and Network Delays, Case 1 corresponds to random losses and no packet reordering, Case 2 and Case 3 corresponds to random network losses and high to very high degree of reordering

Impact of Mean Network Delay and Standard Deviation on Packet re-ordering

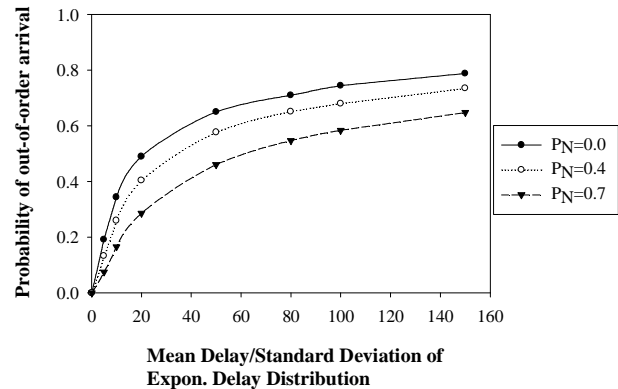


Figure 7. Impact of mean delay and standard deviation on the packet re-ordering under different network packet loss conditions

Tardiness Model Verification under Varying Sampling Interval S
 $P_N = 0$, Mean Delay=20.0 sec

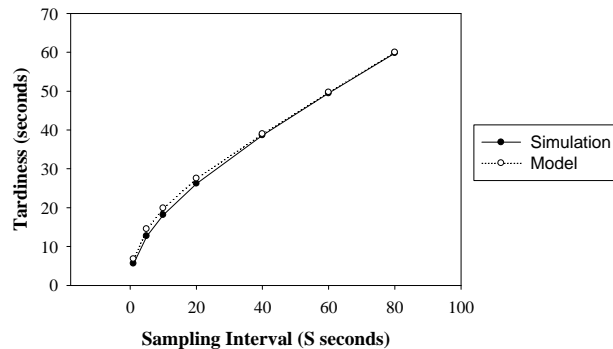


Figure 8. Verification of tardiness model with varying sampling interval

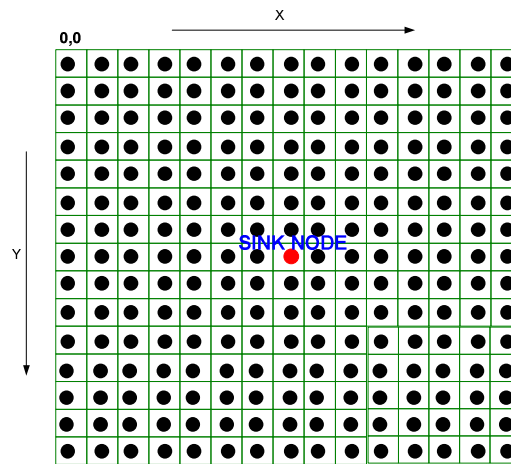
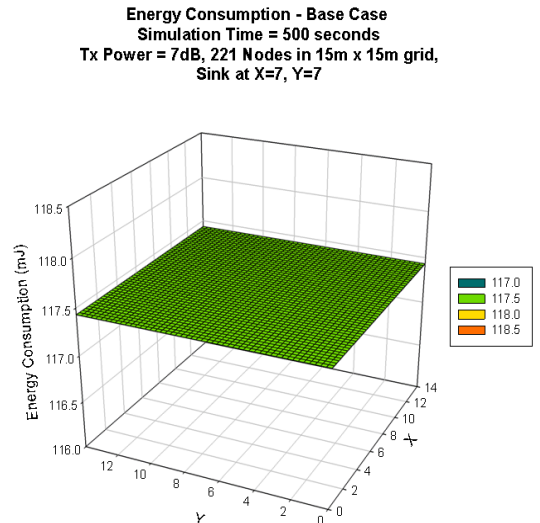
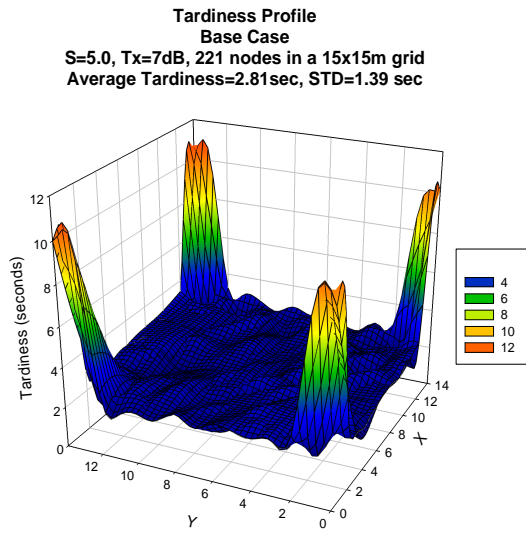


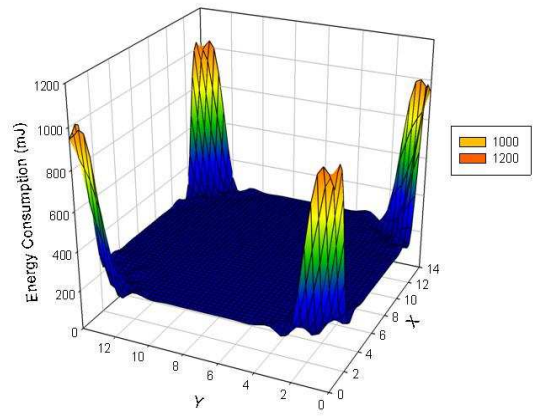
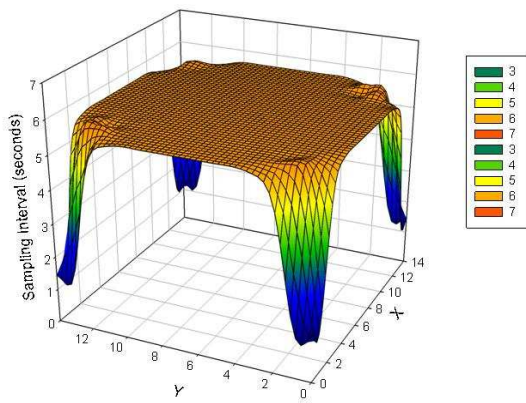
Figure 9. Network with 221 nodes in a 15m x 15m grid, sink at $X=7$, $Y=7$

sampling interval are the key factors that impact the tardiness of the data. Application can achieve desired tardiness bound between source and a sink node by configuring sampling interval, by controlling network packet losses (by adjusting transmit power for example), or network delays (via routing



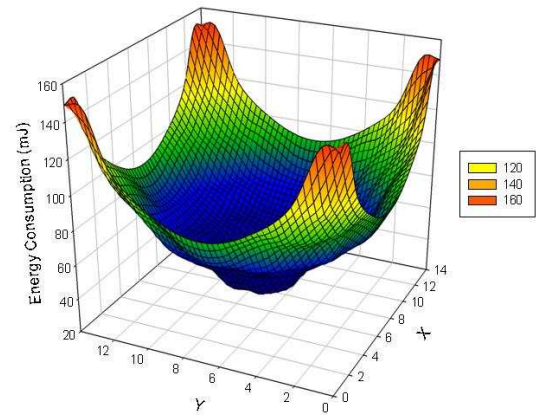
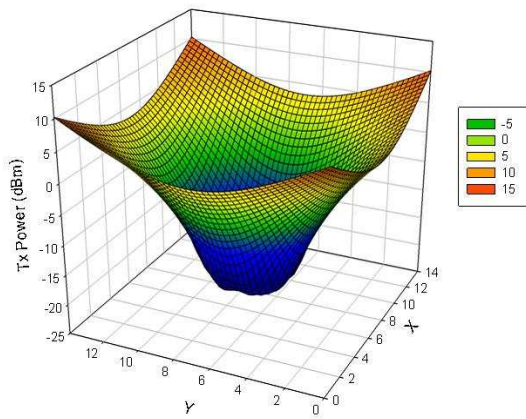
(a) **Adaptive Sampling**
 Tardiness = 2.81seconds
 221 Nodes in a 15m x 15m Grid with Sink at
 X=7, Y=7

(b) **Energy Consumption Profile - Adaptive Sampling**
 Average Energy= 271mJ, STD=131.27,
 Tardiness = 2.81sec, Tx=7dB



(c) **Adaptive Transmission Power**
 221 Nodes in 15m x 15m Grid and Sink at
 X=7, Y=7, Tardiness = 2.81seconds, S=5 seconds

(d) **ENERGY PROFILE: ADAPTIVE TRANSMISSION**
 S=5.0, Tardiness=2.81 sec, Simulation time=500 sec
 221 nodes in 15m x 15m grid
 Average = 77.42 mJ, STD=26.55



(e)

(f)

Figure 10. Application of Tardiness measure in Adapting Sampling Rate and Transmission Energy of source nodes to achieve the desired Tardiness 2.81 seconds with standard deviation =0.

Parameter Name	Parameter Value
Preamble	18 bytes
Frame Length	36 bytes
Encoding	Manchester (2:1)
Modulation	NCFSK
Packet Time	23.3ms
Noise Floor	-105.0dBm
Output Power	-20dBm to 10dBm (Chipcon CC1000 radio 433/315 MHz)
Path Loss Exponent	4.7 (outdoor)
Shadowing Standard Deviation	3.2
Close in reference distance	1m
Close in reference power	55dBm

Table 2. Parameters to determine packet reception rate for MICA2 platform [23,24]

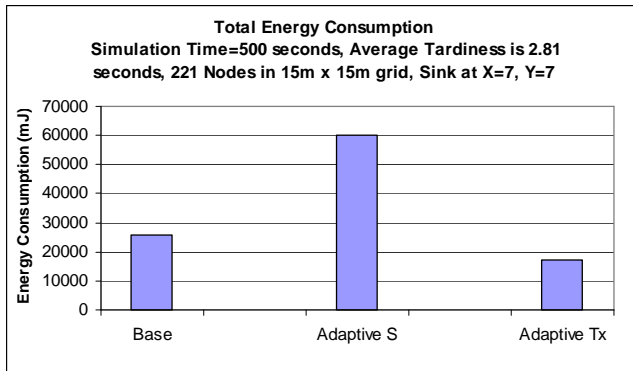


Figure 11. Comparison of total energy consumption in a 500 second interval for three different sensor network configurations with similar average tardiness characteristics

algorithms, adjusting sleep schedule etc.). Configuring each of these parameters to achieve desired tardiness may impact the total energy consumption in the sensor network. Fig. 9 shows the simulation network used for performance analysis. It consists of 221 nodes in a grid of 15m x 15m sensing field. All data to the sink node at the center of the grid, indicated by red dot at X=7, Y=7. In this section, we consider single hop transmission. We consider a scenario when sink node periodically reads input buffer for the most recent available data with read interval 'R'. Each packet transmitted by a source node to the sink node may suffer random losses in the network. In this paper we consider network traffic such that it does not lead to packet losses because of network congestion. All packet losses in the network are considered to be due to wireless link errors. We use a wireless link loss model given by Eq. 18 and Eq. 19 for simulating wireless link losses in the simulator. Table 2 shows different operating parameters used for determining packet loss rate as a function of distance between source and a sink node for a MICA2 hardware platform. As seen the table, we consider outdoor wireless

environment and packet length is 54 bytes. In Fig. 9, maximum distance between any source-sink pair is less than 10m and all nodes can communicate with the sink node in a single hop. Therefore, for this case we ignore delay suffered by packets between source and sink node. Moreover, all packets arrive in-order in a single hop communication. Three sets of experiments are performed for understanding tradeoffs between energy consumption and tardiness.

In the first experiment, all source nodes in Fig. 9 sample their environment periodically with sampling interval $S=5.0$ seconds. Moreover, transmission power of each source node is kept constant, i.e., $T_x=7\text{dBm}$. Average loss rate is computed between all source-sink pairs for a network topology shown in Fig. 9 using Eq. 18, and Eq. 19 over 500 iterations. Fig. 10 (a) and Fig. 10(b) shows results for the experiment 1. Fig. 10(a) shows tardiness of the data from each source to the sink node at X=7, and Y=7 in Fig. 9. In this case average tardiness for all source-sink pairs is 2.81 seconds with standard deviation of 1.39 seconds. In Fig. 10(a), results show that tardiness of different source-sink pairs varies widely between 2.0 and 10.17 seconds in a grid. Fig. 10(b) shows the corresponding energy consumption at each node in the network after 500 seconds of simulation time. As seen in Fig. 10(b), when sampling rate and transmission power of all nodes is same, then each node consumes constant amount of energy, i.e., 117.432mJ in 500 seconds.

Consider a data fusion application where it is necessary to have data with similar tardiness at the sink node. Using Eq. 23, desired tardiness can be achieved by adapting network delay, total packet losses or sampling interval. In the second experiment, sampling interval S of each source node is adapted to achieve the constant mean tardiness of 2.81 seconds between all source-sink pairs. In this case transmission power of each node is kept constant, i.e., $T_x=7\text{dBm}$. Fig. 10(c) shows the sampling interval of all source nodes in a grid to achieve the desired tardiness. As seen in the Fig. 10(c), sensor nodes that are closer to the sink may sample environment at a slower rate. However, as the distance between source node and sink node increases, sampling interval decreases, resulting in higher sampling frequency. The intuition behind increasing the sampling frequency with distance is that when there is high network packet loss probability then sending larger number of packets has the potential to deliver more information to the sink node hence results in decreasing tardiness. It is important to note that this study does not consider occurrence of hot-spots, i.e., network congestion in the network due to increase in sampling rate at certain regions of the network. In all experiments, link bandwidth is not exceeded and there are no packet collisions. All losses are due to wireless link errors. Fig. 10(d) shows the energy consumption at all nodes in a grid when sampling intervals are configured to achieve the desired tardiness. As seen in the figure, energy consumption shows significant amount of variation depending on sampling frequency of the node. Nodes that are closer to the sink node consume significantly less amount of energy compared to nodes at corner of the network grid. Average energy consumption is 271 mJ with high

standard deviation of 131.27 mJ in 500 seconds of simulation time.

In the third experiment, transmission power of each source node is varied while keeping sampling interval S constant, i.e., $S=5.0$ seconds to achieve the constant tardiness of 2.81 seconds for each source-sink pair. By changing the transmission power of the source node, packet loss rate between source-sink pair can be adjusted to achieve the desired tardiness. Aforementioned, packet delivery rate depends on different environment factors such as distance between source and sink node, path loss exponent, and the transmission power. As a first step, for each source node, acceptable network packet loss probability P_N is estimated using tardiness Eq. 23 to achieve the desired tardiness of 2.81 seconds. In the next step transmission power is estimated to achieve the acceptable P_N (calculated in the first step) using wireless loss model proposed in [23] given by Eq. 18 and Eq. 19. Fig. 10(e) shows the transmission power of each source node such that mean tardiness between source-sink pair is 2.81 seconds. MICA2 Chipcon CC1000 radio can transmit at power between -20dBm and 10dBm at 433/315 MHz [24]. In Fig. 10(e), all nodes that require transmission power below -20dBm are configured to operate at -20dBm. In this experiment none of the source node needs transmission power greater than 8.75 dBm. For energy consumption computations, we have used the power consumption at each node which is not the same as the transmitted power. The relationship between actual power required and transmitted power for MICA2 specified in [24] was used assuming that MICA2 mote operates at 3V. Note that nodes that are closer to the sink node require low transmission power compared to nodes at farther distance from the sink node in order to maintain same tardiness between source-sink pair. Note that we have assumed that increase in transmission power does not lead to interference with the neighboring node. This is possible when a TDM based MAC protocol is used for packet transmission. Fig. 10(f) shows the energy consumed by the different nodes in the grid during 500 seconds of simulation time. Note that average transmission energy is 77.42 mJ with standard deviation 26.55 mJ. This is significantly better than the case when sampling interval was adjusted to meet the desired tardiness. Fig. 11 shows the total energy consumed in 500 seconds of simulation time for all three experiments when similar average tardiness was achieved. As seen in the figure, adaptive transmission power performs best in terms of energy efficiency while maintaining same tardiness for all source link pairs. Adaptive sampling does not help in conserving energy but in this case all source sink pairs have the same tardiness of 2.81 seconds. This study illustrates one of the applications of the tardiness model to optimize network configuration parameters to achieve tardiness goals of an application.

5. Conclusions

In sensor networks, network dynamics, networking protocols have significant impact on the tardiness of the data delivered to the processing nodes. The application requirements and the

characteristics of the process being monitored will impose a limit on the tardiness that can be tolerated. This paper presented an analytical model for the tardiness that relates age of the data used for computation with network characteristics such as the network delays and loss rates, and transmission power, as well as the sampling frequency. Analytical model was validated using simulation results. We then applied the model to evaluate tardiness of data received at a sink node. The use of model to evaluate alternate strategies to achieve tardiness targets required by an application was also illustrated. Simulation results for the parameters considered show that adaptive transmission power scheme is more energy efficient compared to adaptive sampling for meeting the desired tardiness requirements of the end users. There are many potential applications of the tardiness measure such as configuring sleep/active schedules of the MAC layer, sampling rate and transmission energy to meet real-time requirements. The tardiness process characteristics effectively capture the impact of network characteristics on the data used at sink nodes for decision making. Application tardiness bound requirements can be used to tune the network parameters to achieve sensor network application goals.

References

- [1] Akyildiz, I., Su, W., Sankarasubramaniam, Y., and Cayirci, E., "A Survey on Sensor Networks," IEEE Communications Magazine, August 2002
- [2] Akyildiz, I., and Kasimoglu, I., "Wireless Sensor and Actor Networks: Research Challenges," Ad-Hoc Networks Journ. (Elsevier), Vol 2, No. 4, pp. 351-367, Oct. 2004
- [3] Bertsekas, D., and Gallager, R., "Data Networks," 2nd Edition, Prentice Hall
- [4] Bouzeghoub, M., and Peralta, V., "A Framework for Analysis of Data Freshness," Proc. of Intl. Workshop on Information Quality in Information Systems, IQIS 2004, pp 59-67, Paris, France, 2004
- [5] Chiasserini, C.-F., and Garetto, M., "Modeling the Performance of Wireless Sensor Networks," Proc. of IEEE Infocom 2004, Vol. 1, March 2004
- [6] Cho, J., and Garcia-Molina, H., "Effective Page Refresh Policies for Web Crawlers," ACM Trans. On Database Systems, Vol. 28, Issue 4, pp. 390-426, December 2003
- [7] Dutta, P., Grimmer, M., Arora A., Bibyk S., and Culler D., "Design of a Wireless Sensor Network Platform for Detecting Rare, Random, and Ephemeral Events," Special Track on Platform Tools and Design Methods for Network Embedded Sensors (SPOTS), Los Angeles, April 2005
- [8] Estrin, D., Culler, D., Pister, K., and Sukhatme, G., "Connecting the Physical World with Pervasive Networks," IEEE Pervasive Computing, Vol. 1, No. 1, pp. 59-69, 2002
- [9] Ganesan, D., Ratnasamy, S., Wang, H., and Estrin, D., "Coping with Irregular Spatio-temporal Sampling in Sensor Networks," 2nd Workshop on Hot Topics in Networks (HotNets-II) 2003, Cambridge (MA) USA, November 2003
- [10] He, T., Stankovic J.A., Lu, C., and Abdelzaher, T., "Speed: A stateless Protocol for Real-time Communication in Sensor Networks," in Proc. of Intl. Conf. on Distributed Computing Systems (ICDCS 2003), Providence, RI, May 2003
- [11] Hull, B., Jamieson, K., and Balakrishnan, H., "Mitigating Congestion in Wireless Sensor Networks," ACM SenSys 2004, Baltimore, MD, Nov., 2004
- [12] Intanagonwiwat, C., Govindan, R., Estrin, D., Heidemann, J., and Silva, F., "Directed Diffusion for Wireless Sensor Networking," ACM/IEEE Transactions on Networking, Vol. 11, No. 1, pp. 2-16, Feb. 2003

- [13] Jain, A. and Chang, E. Y., "Adaptive Sampling for Sensor Networks," Proc. of the 1st Intl. workshop on Data Management for Sensor Networks (DMSN '04), 2004
- [14] Krishnamachari, B., Estrin, D., and Wicker, S.B., "The Impact of Data Aggregation in Wireless Sensor Networks," Proc. of the 22nd Intl. Conf. on Distributed Computing Systems, pp. 575-578, Vienna, Austria, July 2002
- [15] Li, D., Wong, K., Hu, Y.H., and Sayeed, A., "Detection, Classification and Tracking of Targets in Distributed Sensor Networks," IEEE Signal Processing Magazine, Volume: 19 Issue: 2, Mar 2002
- [16] Marbini, A. D., and Sacks, L. E., "Adaptive Sampling Mechanisms in Sensor Networks," London Communications Symposium, London, UK, 2003
- [17] Simon G., Maroti, M., Ledeczi, A., Balogh, G., Kusy, B., Nadas, A., Pap, G., Sallai, J., Frampton, K., "Sensor Network based Counter Sniper System," Proc. of 2nd ACM Conf. on Embedded Networked Sensor Systems (SenSys 2004), Baltimore, MD, Nov. 2004
- [18] Sivrikaya, F., and Yener, B., "Time Synchronization in Sensor Networks: A Survey," Proc. of IEEE Networks, Vol. 18, Issue 4, pp. 45-50, July-Aug. 2004
- [19] Wan, C., Eisenman, S. B., and Campbell, A. T., "CODA: Congestion Detection and Avoidance in Sensor Networks," in Proc. of ACM SenSys 2003, Los Angeles, Nov. 2003
- [20] Xu, N., Rangwala, S., Chintalapudi, K. K., Ganesan, D., Broad, A., Govindan, R., and Estrin, D., "A wireless sensor network for structural monitoring," in Proc. of the 2nd Intl. Conf. on Embedded Networked Sensor systems (SenSys), 2004
- [21] Ye, W., Heidemann, J., and Estrin, D., "Medium Access Control with Co-ordinated Adaptive Sleeping for Wireless Sensor Networks," IEEE/ACM Trans. on Networking, Vol. 12, pp. 493-506, June 2004
- [22] Zhao, J., and Govindan, R. "Understanding Packet Delivery Performance in Dense Wireless Sensor Networks," Proc. of 1st Intl. Conf. on Embedded Networked Sensor Systems (SenSys 2003), Los Angeles, Nov. 2003
- [23] Zuniga, M., and Krishnamachari, B., "Analyzing the Transitional Region in Low Power Wireless Links," 1st IEEE International Conf. on Sensor and Ad hoc Comm. and Networks (SECON), Santa Clara, CA, Oct. 2004
- [24] Chipcon CC1000 Data Sheet, <http://www.chipcon.com>
- [25] Exploratory Project: Heterogeneous Sensor Networks, Intel, <http://www.intel.com/research/exploratory/heterogeneous.htm>