Anycast-Based Routing Scheme with Restricted Flooding for Wireless Sensor Networks

Beran Necat, Alexander Kostin, Yasemin Fanaeian
Management Information Systems Department
Girne American University
Kyrenia, Cyprus

Abstract—An improved method of dynamic route establishment between sensors and sinks is proposed for wireless sensor networks. The method is based on the use of an anycast-based restricted flooding technique and does not require the accumulation and maintenance of state information in network nodes. The proposed scheme is implemented as a simulation model in terms of extended Petri nets. The developed model was investigated in extensive simulation experiments. The behaviour of the model was evaluated with the use of a few practically important performance metrics.

Keywords-component: Anycast routing, Modeling, Petri nets, Simulation, Wireless sensor networks

I. INTRODUCTION

Wireless sensor networks are becoming increasingly popular in different application areas. These areas include (but are not limited to) military, environment, health, home, and commercial applications such as education, security, and business [1], [2]. Typically, a wireless sensor network (WSN) consists of a sufficiently large number of sensor nodes and a relatively small number of sink nodes, distributed in some area. It is not uncommon that nodes of the WSN are not fixed, but move in the area with some predefined pattern or randomly.

The main work of each sensor node in the WSN is to collect some data (such as temperature, humidity, pressure, noise levels, vehicle movement, the presence or absence of objects, soil makeup, etc.) and to transmit the data to one of the sink nodes for the subsequent processing and, possibly, for the forwarding of the processed data to a main information center.

Since sensor nodes typically have low power batteries, it is highly important to consume, by each sensor node, as little energy as possible, because replacement or recharging of batteries in sensors can be practically complex or impossible.

In addition to the transmission of its own data, each sensor usually works also as a router for forwarding data from other sensor nodes since, with a small transmission distance, a sensor is often not capable of directly reaching a sink node.

To transmit collected data, the sensor node needs to preliminary establish a route to some sink node or to determine such a route every time, when it is ready to transmit its collected data. These two methods of localization of a sink node by sensor nodes are known as proactive and reactive techniques, respectively [3], [4].

Each of the two methods of sink localization, proactive and reactive, has positive and negative features. In the proactive method, each sensor establishes a route to a sink and then uses the established route to transmit all available data to the sink or receive some control information from the sink. State information about the established route is stored in all nodes on the route from this sensor to the sink. Unfortunately, if transmission conditions in the network are not stable (for example, as a result of changing weather or expiration of batteries in nodes), then a new route (possibly to another sink) must be established. Such a change of routes can happen quite frequently, which will considerably overload the network and decrease the effective data transmission rate. In addition, since state information must be stored in nodes for each separate sensor, such a method requires considerable memory in nodes and is not scalable with the increase in the number of sensors in the network.

In the reactive method, a sensor establishes a route to a sink only when it is ready to transmit data. After the transmission, the sensor can become idle (sleeping) during a long time interval and does not need to keep the previously established route since it can be completely obsolete when the sensor is ready to transmit new data. This means that, with this method, nodes need to store route state information only during a relatively short time interval. However, in general, to establish a route every time on demand, many sensors can be involved in non-restricted transmissions, that will result in high network traffic and, consequently, require high energy consumption in involved nodes.

In this paper, we propose an improved reactive method of localization of sinks by sensors, with the hop-restricted flooding. The method is based on the idea of an expanding coverage ring. This is done by variation of the time-to-live (TTL) value in each search request, starting with TTL = 1. If, with the given TTL value, no sink replies, then the search request is repeated by the sensor with TTL = 2. Thus, sequentially incrementing TTL value to some maximum in the search requests, the sensor can eventually receive a reply from at least one sink and thus localize it.
As will be shown, this method has a good performance considerably reducing the network traffic.

The rest of the paper is organized as follows. In Section 2, the network structure, its operation, mobility scheme and internode communication are considered. Section 3 describes an energy consumption model. Section 4 is devoted to the organization of the model of the wireless sensor network. Section 5 presents used performance metrics. Section 6 covers simulation setup, results of simulation and their discussion for some combination of parameters. Section 7 concludes the paper and outlines possible expansion of the developed model.

II. STRUCTURE AND OPERATION OF THE INVESTIGATED NETWORK

It is assumed that the wireless sensor network under study covers some area and is populated by sensor nodes and sink nodes. These nodes are initially distributed uniformly in the given area.

Without the loss of generality, the network area is assumed to be rectangular, with coordinates \( x_{\text{min}} \) and \( x_{\text{max}} \) for \( X \) dimension, and \( y_{\text{min}} \) and \( y_{\text{max}} \) for \( Y \) dimension. The nodes communicate with each other through bidirectional high frequency radio communication channels. As is known, at high radio frequencies the distance of transmission by each node is limited. This is especially true for wireless sensor networks, in which sensor nodes typically have low power batteries.

For generality and in contrast with the majority of published works on wireless sensor networks, it is assumed also that all network nodes in the area are mobile, so that the relative positions of nodes vary in time according to a chosen mobility model.

There exist many different mobility models for wireless networks. For the wireless sensor network under study, the random direction mobility model is used [5].

According to this model, each node, starting from a random initial position in the network area, randomly chooses a direction of its movement and moves in this direction with some speed until it reaches the area border. Once the area border is reached, the node randomly chooses its new direction of movement in the network area. The maximal speed of node movement is limited, but actual speed of each node is random within the specified ranges \( V_{\text{max}}(x) \) and \( V_{\text{max}}(y) \) for two dimensions.

In general, radio transmission on high frequencies not only fades quickly with a distance, but this distance also depends on the direction of transmission due to different obstacles, fading effects and interferences in different directions from the transmitter. This means that the reliability of internode communication in a wireless network depends on the direction from the sending node to each receiving node. That is, the success of a node to receive a message from a transmitting node depends not only on the distance to the transmitting node, but also on its orientation with respect to the transmitter. To represent internode communication in a more realistic way, we associate with each receiving node eight directed sectors, as shown in Fig. 1. Clearly, each sector of a receiving node covers 45 degrees. In terms of communication, each sector represents a communication link of the receiving node with transmitting nodes in the network area.

It is assumed further, that these communication links behave differently from a reliability point of view. Specifically, at any moment of time, each link can be in one of two states: ON (active) and OFF (inactive). In the developed model, it is assumed that the duration of states ON and OFF are independent random exponentially distributed variables with mean values \( 1/\mu \) and \( 1/\lambda \) respectively, where \( 1/\mu \) and \( 1/\lambda \) are mean durations of ON and OFF states of each link.

Under the stated conditions, the link behaviour represents a continuous-time Markov process with two states. Now, the link can be formally characterized by the expression

\[
I = \frac{\lambda}{\lambda + \mu},
\]

where \( I \) is the ratio of time the link is operational and thus represents the link availability.

In the model, each sent message contains coordinates of the sending node. Clearly, each node knows its own coordinates. Now, each time the receiving node has a message at its input, it initially determines a sector from which the message has arrived. Knowing the sector, the receiving node checks the current state of the oriented link corresponding to this sector. If its current state is ON, then the message is accepted for the further processing by the receiver. Otherwise, the message is discarded. Such a scheme with multiple oriented sectors implicitly reflects...
many reliability aspects of internode communication and takes into account the appearance of different obstacles between receiving and sending nodes in different directions during the motion of nodes in the network area.

We take into account also a special case, when communicating nodes are very close to each other. In such a case, the communication between nodes is usually quite reliable. The distance of reliable communication is random in general and can be up to a few meters. It is assumed that the distance of reliable communication is different in different links and is distributed uniformly from zero to some small value which is a configuration parameter of the model. Fig. 2 shows an area of diameter $D_{\text{max}}^a$, with one receiving node $r$ and three sending nodes $s_1$, $s_2$, and $s_3$, located at a distance not more than $D_{\text{max}}^a$ from node $r$. Even with a small value $D_{\text{max}}^a$, only nodes $s_1$ and $s_2$ in the dashed-curve area can reliably transmit messages to node $r$. Messages from node $s_3$ cannot be received by node $r$ (for example, due to some obstacle between $s_3$ and $r$).

We assume that the investigated network behaves in the following way. Every time, when a sensor wishes to establish a route to some sink (with the purpose of the subsequent transmission of a ready data packet), it broadcasts a search request message with $\text{TTL} = 1$, starts a time-out, and waits for a reply from any sink that can be at a distance of one hop from the sensor. In general, at this distance there can be zero, one or more sinks. If at least one sink replies, then the search request is recorded as successful. On the other hand, if no reply is received during the specified time-out, then the next search request with the incremented $\text{TTL}$ value is transmitted by the searching sensor. With the incremented $\text{TTL}$ value, more sensors will be involved in the retransmission of the search message. Such a procedure is repeated up to some maximal value of $\text{TTL}$ with two possible outcomes. In the first outcome, with some $\text{TTL}$ value, at least one sink will reply with a message, containing a complete path from the searching sensor to the replying sink. If more than one sink replies, then only the earliest reply is accepted. If no reply is received even with the maximal $\text{TTL}$ value, then this search request is considered as unsuccessful.

Thus, in general, in the search for a sink, the sensor can transmit a few search messages with the increasing $\text{TTL}$ value in subsequent search messages. The search for a path to a sink is done by each sensor node in the network. In addition to the generation of its own search messages, each sensor works also as a router of search messages, generated by other sensor, and of replies from sink nodes to other sensors.

Fig. 3 shows a network area with two sink nodes $a$ and $b$ and a number of sensor nodes $i, j, k, \ldots$.

Sensor $i$ establishes a path to a sink $a$ via sensors $n, o,$ and $p$. Sensor $l$ creates its path to sink $b$ via sensors $q$ and $r$. On the other hand, sensor $k$ has two paths to sink $a$, one short (double dashed) path via sensors $m$ and $s$ and a longer (dashed) path via sensors $t, u,$ and $v$. As a result, sink $a$ will reply to sensor $k$ two times, but the second reply, corresponding to the longer path, will be discarded by sensor $k$. Note that sensor $j$ will probably establish its path to sink $a$ via sensors $m$ and $s$.

III. ENERGY CONSUMPTION MODEL

Energy consumption is one of the most important characteristics of any WSN. Since, in general, each sensor node has a low power battery, it is highly desirable that the sensor consumes as low energy as possible during its lifetime and especially for transmission of messages. Power in sensors is necessary for such purposes, as transmission and receiving of messages, for processing of data, for sensing of the network. There are different energy consumption schemes proposed or described in literature [1], [6], [7], [8], [9], [10], [11], [12], [13], [14]. It is beyond the scope of this paper to conduct a comparative analysis of these schemes. Instead, a very simple energy consumption model will be used in this paper and, based on this model, dependence of energy consumption on different configuration parameters of the WSN will be studied in simulation experiments.
It is assumed, that each sensor in the network model generates and transmits its routing search requests with period $P$ that is a configuration parameter of the model. Period $P$ consists of two intervals ON and OFF with durations $T(\text{ON})$ and $T(\text{OFF})$ respectively, so that

$$P = T(\text{ON}) + T(\text{OFF}).$$

(2)

During ON interval, the sensor is alive and can transmit and receive messages. On the other side, during OFF interval the sensor is basically idle, it does not receive anything, but may complete transmissions, started during ON interval.

The ratio

$$\varepsilon = \frac{T(\text{ON})}{T} = \frac{T(\text{ON})}{T(\text{ON}) + T(\text{OFF})}$$

(3)

can be considered as a sensor availability. Clearly, the smaller the value of $\varepsilon$, the less energy is consumed by the sensor.

It is assumed in the network model that at the beginning of each period $P$ (that is, at the beginning of ON interval) the sensor always generates and transmits its routing request message. Then, during ON interval, it can receive replies from sink nodes to its request, messages from other sensors and work as a router, retransmitting these messages to other network nodes. One operational period $P = t_2 - t_1$ for some sensor node is shown in Fig. 4.

Denote by

$$E_i(t) = aT(\text{ON}) + \beta t_a n_{ij}$$

(4)

the energy consumed by sensor $i$ during $j$th operational period. In expression (4), $a$ and $\beta$ are coefficients, $t_a$ is the average duration of one transmission and $n_{ij}$ is the number of all transmissions by this sensor during time $T(\text{ON})$ of the current period. The first component in the sum (4) represents the energy consumed in one period for receiving. For simplicity, it is assumed that this energy is proportional to the duration of ON interval. The second component in the sum (4) is the energy consumed for all transmissions during one operational period. Clearly, $n_{ij} \geq 1$, since the sensor transmits at least its own request (but can retransmit also messages received from other nodes).

It is reasonable to choose coefficients $a$ and $\beta$ in (4) to satisfy the relation

$$aT(\text{ON}) \ll \beta t_a n_{ij}.$$ 

(5)

This means, that the major part of energy during one operational period is consumed for transmission, which is true for real world wireless sensor networks [15].

Since, as was stated before, $n_{ij} \geq 1$, the expression (5) can be reduced to

$$aT(\text{ON}) \ll \beta t_a.$$ 

(6)

With $N$ operational periods, sensor $i$ will consume energy

$$E_i = \sum_{j=1}^{N} [aT(\text{ON}) + \beta t_a n_{ij}] = aNT(\text{ON}) + \beta t_a \sum_{j=1}^{N} n_{ij}.$$ 

(7)

With $M$ sensors in the network, each of which operates during $N$ periods, the total energy consumed by the sensors of the networks is

$$E = \sum_{i=1}^{M} E_i = aMN T(\text{ON}) + \beta t_a \sum_{j=1}^{N} n_{ij}.$$ 

(8)

In this expression, the double sum is the total number of all transmissions by sensors in the network during $N$ operational periods. Here transmissions of sinks are not taken into account, since typically they have much more powerful energy sources.

Now, since $n_{ij} \geq 1$, the minimal energy consumed by all sensors in the network cannot be less than

$$E_{\text{min}} = aMN T(\text{ON}) + \beta t_a MN = MN (aT(\text{ON}) + \beta t_a).$$

(9)

Assume for simplicity that

$$0 < aT(\text{ON}) \leq \beta t_a.$$ 

(10)

This means that all receivings during $T(\text{ON})$ need energy not more than the energy for one transmission. Let

$$aT(\text{ON}) = \gamma \beta t_a,$$

(11)

where $\gamma$ is a scaling coefficient, with $0 \leq \gamma \leq 1$. Here $\gamma = 1$ means that energy consumed by a sensor during one operational period for receiving is equal to the energy for one transmission, while $\gamma = 0$ means that the energy for receiving is negligibly small and can be ignored.

Then expression (8) can be rewritten as

$$E = \beta t_a (MN + \sum_{j=1}^{N} \gamma n_{ij}).$$

(12)
and $E_{\text{min}}$ in (9) becomes

$$E_{\text{min}} = \beta_{\text{tn}} \left( N \Delta t + M \Delta t \right). \quad (13)$$

Now, as an energy consumption metric it is reasonable to use the relative energy consumption per sensor:

$$\epsilon = \frac{E}{E_{\text{min}}} = \frac{N \Delta t + \sum_{M} \sum_{p} t^{\text{p}}}{M \Delta t (1 + \epsilon)}. \quad (14)$$

This performance metric can be easily calculated in the network model and was used in simulation experiments.

IV. ORGANIZATION OF THE MODEL OF A WIRELESS SENSOR NETWORK

The model of a wireless sensor network has been developed in terms of a class of extended Petri nets. The detailed description of these Petri nets, with numerous examples of their use for modelling and simulation of different types of information systems, including networking systems, is given in [16].

The general structure of the developed model is shown in Fig. 5. The model consists of two types of modules: node module and switching module. The node module implements the functionality of each network node. Recall that there are two types of network nodes in the WSN: sensor node and sink node. The node module of the model exists for each network node, but its functionality is somewhat different for sensors and sinks. The total number of node modules in Fig. 5 is $W$. With $M$ sensor nodes, the model has $W \cdot M$ sink nodes.

The only switching module performs all switching operations to provide communication between network nodes. In addition, the switching module controls movement of nodes in a given network area, traces coordinates of nodes and, for each transmitted message, determines which nodes are potentially reachable from the transmitting node and passes the transmitted message to the reachable nodes for subsequent processing. Note that even if a receiving node is reachable from the transmitting node, the received message can be accepted for the receiving and transmitting nodes and on the current state of the receiving node (ON or OFF).

Each of the two types of modules, shown in Fig. 5, is represented in the model as a segment in terms of extended Petri nets. It is beyond the scope of this paper to describe the both types of modules in more detail. The book [16] contains the detailed description of the switching module that was used in another type of wireless network. This description includes also the complete Petri-net scheme of the switching module and its source text.

V. PERFORMANCE METRICS

To investigate the proposed model of a WSN, four important performance metrics were used. The first performance metric is the relative response ratio. For each sensor, its response ratio represents the number of successful route searching messages, i.e. the number of sink replies to route search requests. If the sensor receives more than one reply to its search request from the same or different sinks, then only the first reply is counted.

Denote by $f_i$ the number of search requests generated by sensor $i$, and by $r_i$ the number of first replies received by this sensor. For convenience, the relative response ratio will be used. Then, with $M$ sensors in the network, the relative response ratio is represented by the expression

$$R = \frac{\sum_{i=1}^{M} r_i}{\sum_{i=1}^{M} f_i}. \quad (15)$$

Clearly, possible values of $R$ are in the range $(0, 1)$, where maximal value corresponds to a situation when each sensor receives a reply for each transmitted search request. Note that, for example, with a small value of TTL in the search message of the sensor, there can be no reply from any sink to this sensor at all. This will result in the decreased value of $R$.

Depending on the current sensor availabilities and link availabilities, there can be no reply also with a sufficiently large TTL value.

The second performance metric, chosen in this paper, is the response time. For each sensor, the response time related to a route search request, is the time interval since the moment of transmission of the request message up to the moment of receiving of the first reply, if any, from a sink. Denote by $d_i$ the average response time for sensor $i$. This response time is measured in the model for all successful search requests generated by sensor $i$. With $M$ sensors in the network, the overall response time is represented by expression

$$D = \left( \sum_{i=1}^{M} d_i r_i \right) / \sum_{i=1}^{M} r_i. \quad (16)$$

where $r_i$ has the same meaning as in expression (15).

The next performance metric is the relative network traffic. This metric represents the number of transmissions in the network per each route request message generated by
a sensor. Clearly, this number should be kept as low as possible, since the increased number of transmissions will require a corresponding increase in energy consumption. In the model, this metric is represented by expression

\[ G = f_2 \cdot \sum_{m \in M} f_3 \]  

(17)

where \( f_2 \) is the total number of transmissions in the network and \( f_1 \) has the same meaning as in expression (15).

The last performance metric considered in the behaviour of the network model in this paper is relative energy consumption \( e \). The derivation and explanation of this performance metric are given in Section 3.

VI. SIMULATION SETUP AND RESULTS OF SIMULATION

As was stated in Section 4, the model of the wireless sensor network under study has been developed with the use of a class of extended Petri nets. The designed model was implemented and investigated in the simulation system Winsim [16] according to the following setup. The network area is assumed to be a square of 300 m \( \times \) 300 m, with the total number of mobile nodes \( W = 50 \). In the simulation, the number of sensor nodes is set to \( M = 45 \). Remaining \( W - M = 5 \) are sink nodes.

With \( W = 50 \) nodes, the given area can be approximated as a Poisson point field [17]. It is known that, for a two-dimensional Poisson point field, probability density function of the distance from any point to its nearest point is

\[ f(r) = \begin{cases} \frac{2\pi \lambda e^{-2\pi \lambda r^2}}{r}, & r \geq 0, \\ 0, & r < 0, \end{cases} \]  

(18)

where \( \lambda \) is the density of the Poisson field.

For the network area 300 m \( \times \) 300 m, with \( M = 50 \) nodes, \( \lambda = \frac{1}{11000} \) nodes per square meter. Using this value in (18), one can find that the mean value of the distance between pairs of nearest stations is about 32 m. Based on this, we assume in the simulation, that each station in the WSN has a maximal transmission distance of 50 m. With this distance, each transmitting node can reach, with sufficiently high probability, one or more neighbour nodes.

All nodes in the network area move according to the pattern described in Section 2. The speed of movement of each node is random, but it is limited by 5 km/h. In its movement, each sensor node periodically transmits routing search requests to localize a sink and, at the same time, retransmits messages on behalf of other nodes. The period of generation of search requests is 5000 ms. To accumulate sufficient statistical data, each sensor node generates at least 200 request messages. With \( M = 45 \) sensor nodes, this will result in 45 \( \times \) 200 = 9000 request messages in total. Thus, with period \( P = 5000 \) ms, the simulation interval for each simulation run should be not less that 200 \( \times \) 5000 ms.

In the simulation experiments the simulation interval was set to be 200 \( \times \) 5000 ms + 10000 ms, where 10000 ms is a small margin. Thus, actually each sensor node will generate 202 search requests. Distance of reliable communication between close nodes is random, with the maximal value in the range between 5 and 10 meters.

In all conducted simulation experiments the link availability (see expression (1)) is fixed at \( l = 0.8 \), with mean duration of link in ON state equal to 10000 ms and, correspondingly, with mean duration of link in OFF state equal to 2500 ms. As was stated in Section 2, durations of ON and OFF intervals are random exponentially distributed variables.

For the sensor availability \( s \) (see expression (3)), three different values 0.5, 0.7, and 0.9 were used in different simulation runs. In particular, with fixed value \( P = 5000 \) ms, \( \lambda \) (ON) = 2500 ms for \( s = 0.5 \). Further, in all simulation experiments the interval of checking the states of oriented communication links of each node was set at 2000 ms. At the end of this interval, the link state can be changed to OFF (instead of ON) or ON (instead of OFF). Finally, we assume for simplicity that transmission time of any type of message by each network node is random and distributed according to the uniform probability distribution in the range (5, 30) ms.

In all simulation experiments the varied parameter was TTL. The values of TTL were 1, 2, 3, 4, and 5.

The simulation results are shown in the form of graphs in Figs. 6, 7, 8, and 9.

On the base of the simulation results, we can arrive at the following comments and observations.

1. As one could expect, all used performance metrics depend on TTL values. But the character of this dependence is different for different metrics.

2. As Fig. 6 demonstrates, the relative traffic in the network increases approximately four times for sensor availability \( s = 0.5 \), when TTL value increases from 1 to 5. It is slightly more than one for TTL = 1, but it almost does not change when TTL value increases from 4 to 5. However, for sensor availability \( s = 0.9 \) the relative traffic increases almost linearly for all TTL values, with the maximal value more than 8.

4. According to Fig. 7, the average response ratio behaves almost in the same way for sensor availability values 0.5 and 0.7, but it is higher for \( s = 0.9 \) and it remains the same when TTL increases from 4 to 5. Actually, the average response ratio can be considered as an evaluation of probability of establishment of a route to a sink by a sensor. As is known from the probability theory [18], if some event happens with probability \( p \), then the expected number of attempts up to the first appearance of this event is \( 1/p \). Thus, for example, for TTL = 4 and \( s = 0.5 \), the
average response time is about 0.46, so that $1/0.46 \approx 2.2$. This means, that a sensor needs to make 2.2 attempts on average to establish a route to a sink.

6. Overall response time (Fig. 8) also increases with the increase of TTL value, but this increase is less than two times, when TTL value changes from 1 to 5. The increase of response time can be explained by the fact that, with larger values of TTL, more distance network nodes will respond and contribute to the response time.

8. Relative energy consumption (Fig. 9) has its minimal value 1 for TTL = 1 and then increases with the increase of TTL, since more and more nodes will be involved in the retransmission of messages. The minimal value 1 of this metric is due to the fact that, with TTL = 1, each sensor transmits only its own messages and does not retransmit messages of other sensors.

Fig. 6. Relative traffic for different values of sensor availability $x$.

Fig. 7. Average response ratio of the network.

Fig. 8. Overall average response time of the network.

Fig. 9. Relative energy consumption of the network for $p=0.5$ in (11).

9. As graphs in Figs. 6 - 9 demonstrate, all investigated performance metrics do not change essentially when TTL value increases from 4 to 5. This means that, for the given network, there is no necessity to try values of TTL larger than 4.

VII. CONCLUSION

An anycast-based routing scheme with a restricted flooding for wireless sensor networks is proposed and investigated with simulation. In the network, a realistic internode communication scheme is implemented that takes into account reliability of receiving messages by each node depending on its orientation relative to transmitting nodes and, thus, relatively easy reflects the effect of different obstacles between receiving and transmitting nodes.

The simulation model of the network is developed in terms of a class of extended Petri nets that give a possibility to clearly represent parallelism of events and processes in the model.

During extensive simulation, practically important performance metrics were investigated and their dependence of the value of TTL was shown and explained.

The developed model can be used as a starting point for the implementation of more complex networks, their behaviour and more advanced protocols of routing in wireless sensor networks.

REFERENCES


