Introduction to Model-based Reliability
Evaluation of Wireless Sensor Networks

C. Jäggle * J. Neidig * T. Grosch * F. Dressler **

* Siemens AG, Industry Sector, Nuremberg, Germany (e-mail: {claudia.jaeggle, joerg.neidig, thomas.grosch}@siemens.com).
** Computer Networks and Communication Systems, Dept. of Computer Science, University of Erlangen, Germany (e-mail: dressler@informatik.uni-erlangen.de).

Abstract: A high level of reliability is a significant requirement for using wireless sensor networks in industrial environments. Model-based evaluation is usually applied in conventional systems to estimate the reliability. In contrast, for analyzing sensor networks, these methods are hardly tested and proven due to the unique properties of that kind of network. This paper presents a first model-based approach to quantitatively assess the reliability of sensor networks. Based on an abstract model of a sensor network, the degree of detail with respect to the characteristic aspects of sensor networks is increased in a stepwise manner. If network topologies are taken into account, analytical methods fail and the assessment of reliability measures has to be done numerically. Finally, a Monte Carlo Simulation is conducted to also cover dynamics in the topology. In conclusion, this paper shows that, by knowledge of link probabilities and lifetime distributions of single sensor nodes, model-based analysis may be used to estimate reliability measures of sensor networks.

Keywords: Network reliability, Sensor networks, Stochastic modeling, Monte Carlo Simulation

1. INTRODUCTION

A visionary idea within the scope of logistics and stock keeping envisions to equip products, or generally speaking objects, with intelligence, communications capabilities and sensors, i.e. with sensor nodes, to automate industrial processes. This idea shifts the automation intelligence, information collection, storage, and retrieval concepts from a centralized approach to a highly decentralized and object-oriented approach. By this means e.g. stored blood bottles could measure the environment temperature and alert in case of improper values. Packets in a logistics center could communicate with the infrastructure to determine their transportation ways automatically.

The idea is to use sensor nodes as intelligent labels and to attach them directly to products. Thus, product information like manufacturer, date of expiry or destination as well as automation instructions for further product processing steps may be saved at the products themselves. The on-board processor and sensors enable the node to interact with its environment. To benefit from that technique, it has to be assured that the acquired and stored data can be read and processed reliably. However, harsh industrial environments often constrain communication, especially the direct access to reading devices. For that reason, the sensor nodes of several objects have to build up a multi-hop network to ensure continuous connectivity. Another problem is that cheap sensor nodes tend to be not very reliable: communication links may break, nodes may fail or get lost during transportation. To be able to persistently access the data acquired and stored by each object even in case that some nodes fail, the information has to be stored redundantly on multiple nodes. The system is functional as long as sufficient sensor nodes are alive to deliver the entire product information of the network to an external reader. Our main objective is to assess the MTTF of such a system.

In order to be applicable in industrial environments, such systems must be sufficiently reliable. In conventional systems, model-based evaluation is usually applied to estimate the reliability. In contrast, for analyzing wireless sensor networks, these methods are hardly tested and proven. Characteristic aspects of sensor networks such as limited accessibility and dynamic topology changes complicate the usage of traditional reliability evaluation methods. Default techniques, for example Boolean algebra (Whitesitt, 1995) or fault tree analysis (International Electrotechnical Commission, 1990), usually presuppose that network nodes may fail while network links are fully reliable. Recently developed approaches concentrate on the reliability assessment for networks with unreliable links but assume fully reliable network nodes. For instance, Chiu et al. (2001) present a heuristic algorithm to obtain the K-terminal reliability for a distributed system assuming that each node is perfectly reliable whereas operational probabilities are assigned to the links. The same assumptions are made by Hardy et al. (2007), where the K-terminal reliability is evaluated using Binary Decision Diagrams. Elmallah (1992) and Shpungin (2008) consider unreliable nodes as well as unreliable links. The first approach models distributed networks as probabilistic graphs that are represented by polygon diagrams, and provides polynomial time algorithms to solve K-terminal reliability problems with...
node failures. The second one uses Monte Carlo Simulation and a combinatorial approach to efficiently estimate the system reliability. Random lifetimes are assigned to both, network edges and network nodes. By contrast, the approach presented in this paper assumes discrete link probabilities such that links may fail and recover. Additionally, since limited accessibility is not a topic in conventional systems, this special aspect of sensor networks is included in none of the mentioned analysis but will be addressed in this paper.

To overcome the shortcomings of traditional reliability evaluation methods, novel modeling techniques with focus on wireless sensor networks have been proposed. Sha and Shi (2005) present a lifetime model based on the energy consumption of individual nodes, the importance of different sensors based on their positions, the link quality, the connectivity, and the coverage of the sensor network. Rai and Mahapatra (2005) consider a sensor node as alive until running out of energy. The lifetime of a whole sensor network is determined analytically depending on the fraction of alive sensor nodes. A mathematical analysis of network reliability based on the number of transmitted messages from source nodes to sink is given by Durmaz and Baydere (2004). The lifetime of event-driven wireless sensor networks has been analyzed by Noori and Ardakani (2008). Depending on the node deployment, the initial energy of the sensors, the packet generation model, and number of sensors, an analytical expression has been derived for the complementary cumulative density function of the network lifetime. Based on such quite diverse definitions of the sensor network lifetime, Dietrich and Dressler (2009) proposed a more comprehensive lifetime metric including energy issues and application-dependent quality of service demands.

This paper presents a first model-based approach to assess the reliability of redundant sensor networks in automation industry in a stepwise manner. Based on an abstract model of a sensor network, the degree of detail with respect to the characteristic aspects of sensor networks is successively increased. Mainly two characteristics of sensor networks are taken into account: the limited accessibility and dynamic topology. The former signifies that not every node of a sensor network may be addressed directly by a reading device due to shading effects or limited transmission ranges. The latter expresses that links undergo occasional failures, followed by repair, such that links may be either available or unavailable. We show that for a simplified system the reliability assessment can be done analytically. If specific network topologies are considered, analytical methods fail and reliability measures have to be assessed numerically. To cover both, limited accessibility and dynamic topologies, a Monte Carlo Simulation is conducted. For all the evaluations, we assume that link probabilities and lifetime distributions of single sensor nodes are known. Link probabilities specify the existence of communication connections between sensor nodes as well as communication connections between a node and a reading device. In this paper, we demonstrate that by knowledge of these parameters model-based analysis may be used to estimate reliability measures for sensor networks.

The remainder of this paper is organized as follows. First, the system to be analyzed is presented in Section 2. In Section 3, needed definitions and notations are fixed. The reliability assessment will be done in Section 4. Finally, conclusions are drawn in Section 5.

2. SYSTEM DESCRIPTION

The system to be analyzed throughout this paper is a sensor network similar to that given in Fig. 1. The network consists of a set of sensor nodes attached to objects on a pallet. Each node is able to store and to relay data. There are two different types of communication links: connections between two arbitrary sensor nodes and connections between sensor nodes and a retrieval system. Since sensor nodes and network links are not very reliable and may fail, data is stored redundantly on several nodes. The system is considered as functional as long as sufficient sensor nodes are alive, connected with each other, and connected with an external reader to be able to deliver the entire system information whenever it is demanded by the user. The lifetime distribution and the mean time to failure has to be calculated to make a statement about system reliability.

Fig. 1. Sensor network example

3. NOTATIONS AND DEFINITIONS

Throughout this paper, the system to be analyzed is defined by the triple $\Sigma = \{S, S_0, E\}$. It consists of a network $S$ of individual sensor nodes $\{S_1, S_2, \ldots, S_N\}$, where $N$ denotes the number of sensor nodes. Furthermore, the reading device is said to be a special network node, the so-called reader node $S_0$. The third system element is the set $E$ of direct communication links $E_{i,j}$ between any two different sensor nodes $S_i$ and $S_j$ as well as between any sensor node and the reader node. The undirected graph $G$ depicts the topology of the network. It holds that $E_{i,j} = E_{j,i}$. A list of symbols is given in Table 1.

The lifetime of an individual sensor node is characterized by a failure distribution $F(t)$ or by a reliability function $R(t)$. The failure distribution function is the probability of a sensor node failing in the time interval $[0, t]$. The reliability function or survivor function is the probability of a sensor node not failing in that time interval. The relationship between the failure and the reliability function is the following:

$$R(t) = 1 - F(t).$$

The failure distribution incorporates every random event that disrupts the functionality of a sensor node, namely...
In contrast to sensor nodes, the reader node is considered as fully reliable. Thus, the mean time to failure (MTTF) of a sensor node is:

$$MTTF = \int_0^\infty R(t) \, dt.$$  \hfill (2)

In contrast to sensor nodes, the reader node is considered as fully reliable.

Analogously, the failure behavior of the system as a whole may be specified by $F_S(t)$, $R_S(t)$, and $MTTF_S$. To be regarded as fully functional, the system has to meet two conditions: a subset of at least $K$ sensor nodes has to be connected to be able to communicate with each other and to extract the total amount of system information out of the redundantly on multiple nodes stored data (condition 1). Furthermore, the connected subnetwork must be accessible by the retrieval system. Accordingly, at least one sensor node within the subset must have a direct communication link to the reader node (condition 2).

Since the wireless communication range of sensor nodes is limited and transmission paths may be seriously disturbed, it is likely that not every possible direct communication link $E_{i,j}$ exists. The link probability $p_{\text{link}}^{i,j}$ with $i, j \in [1, \ldots, N]$ and $i \neq j$ denotes the probability that two sensor nodes $S_i$ and $S_j$ may directly communicate with each other. Correspondingly, $p_{\text{critical}}^{i}$ depicts the probability that an individual sensor node $S_i$ with $i \in [1, \ldots, N]$ is directly connected to the reader node $S_0$. Due to the fact that contact points to the reader significantly influence system reliability, these sensor nodes are referred to as critical nodes and $p_{\text{critical}}^{i}$ is called criticality probability.

## 4. RELIABILITY ASSESSMENT

The main objective of this paper is to assess the reliability of the sensor network given in Fig. 1 in terms of lifetime distribution and mean time to failure. The modeling and calculation process is done in a stepwise manner. First, a simplified system is analyzed. Then, limited accessibility is taken into account. As another intermediate step networks with a fixed topology are examined. Finally, dynamic topology is considered.

### 4.1 Abstract Model

A possibility to keep the model simple is to assume unlimited accessibility and full meshing. Setting

$$p_{\text{link}}^{i,j} = 1 \quad \forall i, j \in [1, \ldots, N], \quad i \neq j \hfill (3)$$

$$p_{\text{critical}}^{i} = 1 \quad \forall i \in [1, \ldots, N] \hfill (4)$$

results in a conventional K-out-of-N-redundancy: First, every pair of sensor nodes is directly connected; secondly, every sensor node is able to establish contact with the reader node. Each sensor node is a critical node, i.e. $C = N$. Reliability calculation for that kind of redundant system can be done analytically by using binomial coefficients or default methods from reliability theory like fault tree analysis (International Electrotechnical Commission, 1990). However, for the considered system, unlimited accessibility and full meshing are not realistic such that the corresponding model is far too abstract.

### 4.2 Limited Accessibility

In this subsection, full meshing is assumed for sensor node to sensor node communication, but the constraint on the accessibility is softened, i.e. $C < N$, so that only a subset of sensor nodes is directly connected to the reader. The assumption is made that $C$ is known and constant during operation time. $p_{\text{critical}}^{i}$ will only take two different values:

$$p_{\text{critical}}^{i} = \begin{cases} 1 & \text{for } i \in [1, \ldots, C] \\ 0 & \text{for } i \in [C + 1, \ldots, N] \end{cases} \hfill (5)$$

To fulfill condition 2, at least 1-out-of-$C$ critical nodes has to be active. The probability to meet that condition is expressed by:

$$R_C(t) = \sum_{i=1}^{C} \binom{C}{i} R(t)^i F(t)^{C-i} \hfill (6)$$

Depending on the number of critical sensor nodes $i$, there have to be at least $(K - i)$-out-of-$(N - C)$ noncritical nodes, which are functional to additionally fulfill condition 1. By virtue of the special case that $i > K$, i.e. the
number of active critical nodes is high enough to meet condition 1 and 2, a case differentiation is needed. For that special case, the noncritical nodes may take arbitrary states irrespective of the number of active critical nodes: $i$:

$$R_{K-1}(t) = \begin{cases} 
   \sum_{j = K-1}^{N-C} \binom{N-C}{j} R(t)^j F(t)^{N-C-j} & \text{if } i \leq K \\
   \sum_{j = 0}^{N-C} \binom{N-C}{j} R(t)^j F(t)^{N-C-j} & \text{if } i > K 
\end{cases}$$

(7)

The reliability of a $K$-out-of-$N$-redundancy with $C$ critical nodes may be calculated by combining (6) and (7):

$$R_S(t) = \sum_{i=1}^{C} \left[ \binom{C}{i} R(t)^i F(t)^{C-i} R_{K-1}(t) \right]$$

(8)

That approach results in a closed form solution. Thus, system reliability can be expressed analytically and determined exactly under the assumption of full meshing and limited accessibility. The main disadvantage is that one reliability function must be fixed for all sensor nodes.

Example. The procedure will be illustrated on an example. Given the network of Fig. 2 consisting of four sensor nodes, where $S_1$ and $S_2$, printed in bold, are critical nodes, the reliability for a 2-out-of-4-redundancy is evaluated as follows:

$$R_S(t) = \sum_{i=1}^{2} \left[ \binom{2}{i} R(t)^i F(t)^{2-i} R_{2-i}(t) \right]$$

$$= 5R(t)^2 - 6R(t)^3 + 2R(t)^4$$

(9)

Fig. 2. Network example with full meshing and two critical nodes

Assuming an exponential distribution with a mean of two years for the reliability function of individual sensor nodes results in

$$R_S(t) = 5 \left( e^{-0.5t} \right)^2 - 6 \left( e^{-0.5t} \right)^3 + 2 \left( e^{-0.5t} \right)^4.$$  

(10)

The corresponding curve is depicted in Fig. 3. At a time of about seven years the system reliability reaches zero. The mean time to system failure is

$$MTTF_S = \int_0^\infty R_S(t) \, dt$$

$$= \int_0^\infty 5 \left( e^{-0.5t} \right)^2 - 6 \left( e^{-0.5t} \right)^3 + 2 \left( e^{-0.5t} \right)^4 \, dt$$

$$= 2 \text{ years.}$$

(11)

4.3 Fixed Network Topology

We now include topology information in the reliability assessment of a sensor network. Instead of assuming a full mesh, a fixed topology is considered where $p_{ij}$ is set to one for some pairs of $S_i$ and $S_j$ and set to zero for the remaining links. Thus, the network graph $G$ is given as input to the evaluation. Since the connectivity of individual nodes is different, the binomial distribution cannot be used to summarize several system states. In fact, every possible system state has to be evaluated, i.e. fulfillment of condition 1 and 2 has to be checked. For that reason, the reliability evaluation has to be done numerically rather than analytically. The procedure is described in Algorithm 1.

**Algorithm 1: System reliability for fixed network**

1. input : $N$, $K$, $C$, $F(t)$, $R(t)$, $G$
2. output: $R_S(t)$
3. begin
4. $R_S(t) \leftarrow 0$
5. foreach system state do
6. if system is functional then
7. $v \leftarrow$ number of active sensor nodes
8. $w \leftarrow$ number of inactive sensor nodes
9. $R_S(t) \leftarrow R_S(t) + R(t)^v \cdot F(t)^w$
10. end
11. end
12. end

Line 6 in Algorithm 1 is very expensive in terms of computational time because of checking condition 1. The probability that a set of $K$ nodes is connected in a distributed system is referred to as $K$-terminal reliability (Chiu et al., 2001). In general, the computational complexity of evaluating $K$-terminal network reliability measures is $\mathcal{NP}$-hard (Ball, 1986). For that reason, any exact algorithm requires exponential computing time and the reliability assessment will be feasible for small problems only. In contrast to the analytical solution of subsection 4.2, this approach is not limited to a single reliability function for all sensor nodes.

Example. To illustrate how to evaluate numerically the reliability of sensor networks with fixed topologies, the network of Fig. 4 with four nodes is analyzed. There are two critical nodes: $S_1$ and $S_2$. The system reliability.
Fig. 4. Network example with fixed topology and two critical nodes

for a 2-out-of-4-redundancy is calculated by summing up the probability for all states that represent a functional system. If e.g. all four nodes are alive, the system is functional. Therefore $R(t)^4$ has to be included in the calculation whereas the term $F(t)^4$ is ignored because four dead nodes do not constitute a functional system. $R_S(t) = R(t)^4 + R(t)^3 F(t) + R(t)^3 F(t) + R(t)^3 F(t) + R(t)^2 F(t)^2$ leads to $R_S(t) = (e^{-0.5t})^4$ and $F(t) = 1 - e^{-0.5t}$ and $R_S(t) = 3R(t^2) F(t)^2$ (12)

Setting $R(t) = e^{-0.5t}$ and $F(t) = 1 - e^{-0.5t}$ leads to $R(t) = (e^{-2t}) - 3(e^{-1.5t}) + 3(e^{-t})$ (13) and a mean time to system failure of $MTTF_S = 1.5$ years. (14)

4.4 Dynamic Topology

The computational complexity of evaluating the reliability for the given scenario increases if dynamic network topologies are considered. As every sensor node and every link may take two states, the number of possible system states is extremely high. It is not feasible to carry out an operational check for every existing system state. For that reason Monte Carlo Simulation is used to estimate the system reliability. The simulation procedure is described in Algorithm 2. A given basis topology $G$ indicates which links are physically possible and which sensor nodes may be critical nodes. Depending on the set $E$ of direct communication links in $G$, $p_{critical}$, and $p_{link}$, a random network topology is generated for every simulation replication. Additionally, a random lifetime $l_i$ is sampled from $F(t)$ for every sensor node. The sampled lifetimes are sorted in ascending order. Whenever the end of a node lifetime is reached, the fulfillment of the two functionality conditions is rechecked. In this way the point in time of the transition from correctness to malfunction is updated. Thus each replication determines a single simulated time to system failure $L_m$. By conducting a large number of $M$ replications, the estimators $MTTF_S$ and $R_S(t)$ for the mean time to system failure and the system reliability may be assessed. In order to get significant results, the sample size has to be chosen properly. According to Law and Kelton (2000), the minimum number of required samples is determined by the following formula:

$$M = \min \left\{ i \geq S^2(n) \left( \frac{z_{1-\alpha/2}}{\epsilon} \right)^2 \right\}$$ (15)

$S^2(n)$ represents the sample variance observed on a test simulation with $n$ replications, whereas $\epsilon$ indicates the maximum absolute error affecting the estimated mean time to system failure. The level of significance is denoted by $\alpha$ and the $(1-\alpha/2)$-quantile of the standard normal distribution by $z_{1-\alpha/2}$. Variable $i$ signifies any possible integer that is greater than or equal to the given expression.

The proposed Monte Carlo Simulation allows to estimate the reliability of a network of sensor nodes attached to objects of one pallet where limited accessibility and dynamic topologies play a significant role. The reliability function may be chosen independently for each sensor node. Likewise, link probabilities may be fixed arbitrarily and independently. The given analysis approach may be extended by considering further characteristic aspects of sensor networks. For instance the knowledge of sensor node positions would allow to calculate the distance between nodes and incorporate fading aspects. Concerning the computational simplicity, the Monte Carlo Simulation was feasible for networks with up to one hundred sensor nodes. To increase the efficiency, variance reduction methods may be used.

Example. At this point the general procedure of Algorithm 2 is applied to the specific basis topology of Fig. 5. Altogether, the network consists of 30 sensor nodes, out of which 6 are critical and at least 15 are required to deliver the entire information of the pallet as a whole to the reading device. The remaining parameters are set according to Table 2.
improvement. Interestingly, according to the simulation results, the net-

The estimated system reliability is depicted in Fig. 6. The accomplished 110000 replications took 56 min and 50 sec on a 1.6 GHz machine with 1 GB RAM. The estimated mean time to system failure was 0.949 years. The estimated system reliability is depicted in Fig. 6. Interestingly, according to the simulation results, the network reliability is very poor in comparison with the mean time to failure of single sensor nodes. A suggestion for future research may be to investigate schemes of reliability improvement.

5. CONCLUSION

This paper is a first contributing effort to assess the reliability of redundant sensor networks in automation industry using a model-based approach. It was shown that analytical methods are appropriate for abstract models only. If network topologies have to be taken into account, analytical methods fail and the assessment of reliability measures has to be done numerically. Furthermore, it became apparent that simulative methods are favorable when considering dynamic topologies. First simulation results are surprising insofar as system reliability is very poor in comparison with the reliability of single sensor nodes. For that reason one may reasonably assume that the reliability of single sensor nodes is not the dominating factor of sensor network reliability.

REFERENCES


