Received Date: July 28, 2011, Revised Date: April 17, 2012.

Survivability Analysis Using Probabilistic Model Checking: A Study on Wireless Sensor Networks

Sophia Petridou, *Member, IEEE,* Stylianos Basagiannis, *Member, IEEE,* and Manos Roumeliotis *Member, IEEE*

***Abstract*—Survivability of a wireless sensor network (WSN) reflects the ability of the network to fulfill its mission despite the presence of abnormal events such as failures. Given that sensor networks are receiving increasing attention due to the wide range of their applications, which include the critical areas of health, military and security, survivability constitutes a key property for their study. This paper proposes a quantitative analysis for survivability evaluation of wireless sensors networks using probabilistic model checking. We define network survivability in line with four measures, namely, the frequency of failures, the data loss, the delay and the compromised data due to a variety of failures. In particular, three types of failure events are considered, namely node, link and attack failures, which are due to power faults, communication faults and black hole attacks, correspondingly. Then, we represent network's behavior with Continuous-Time Markov Chains (CTMCs) and we randomly inject the aforementioned faults and attacks in the network, in order to derive results which quantify the impact of them. Although the proposed study considers and provides results for a wireless sensor network architecture, it has the potential of being exploited in different networks with their own specifications.**

***Index Terms*—Availability, probabilistic model checking, survivability, wireless sensor networks, PRISM**

I. Introduction

WIRELESS sensor networks (WSNs) are continuously gaining popularity, since modern society depends on large-scale information systems, which embed WSNs, to conduct business, government and defence [1], [2]. In recent years, a variety of critical applications, including health monitoring, industrial control, battlefield surveillance, transportation and environmental monitoring [3], [4], [5], are conducted by WSNs. Sensor networks are differentiated by other wireless environments in that, their nodes are left to operate without human intervention for weeks or months at a time. Therefore, wireless sensor nodes are vulnerable to a number of dangers, including attacks and failures. These events entail that a number of sensors will stop functioning correctly, affecting the way that network delivers its services.

All the above have motivated the research community to intensively work over WSNs' survivability, since survivability analysis addresses the issue of a network coping with failures, recovering from them and minimizing their impact [6], [7]. In [1], [2], survivability is defined as a key property of a network, which reflects its ability to fulfill its mission, in a timely manner, despite the presence of abnormal events such as failures. A more mathematically restricted definition of survivability is introduced by [7] providing the framework to quantify the concept of network's survivability.

In this paper, we proceed to the survivability analysis of a WSN, according to the definition [7], using probabilistic model checking [8], [9]. For this purpose, we consider a network environment consisting of a number of sensor nodes, distinguished as critical and simple sensors, and a central base station connected to a database. Links between sensors are wireless, while critical sensors are wired connected to the base station. Both critical and simple nodes sense data, which in turn are aggregated to the base station through critical nodes. We model such an environment as a Continuous-Time Markov Chain (CTMC) and, then, we inject faults and attacks in the developed CTMC in order to study their impact on the WSN. We use the term fault to denote incidents, e.g., power or communication outages, and attack to indicate intentional abnormal events, e.g., black hole attacks. Both faults and attacks lead to network failures.

More specifically, inspired by [10], WSN survivability is defined in line with the frequency of failures and the impact of them. In our analysis, we consider three types of failure events, namely node, link and attack failures, due to power faults, communication faults and black hole attacks, respectively. Consequently, data loss, delay and compromised data reflect the impact of failures on the WSN. Therefore, frequency of failures, data loss, delay and compromised data are the four measures that our analysis is based on. The proposed survivability analysis is automatically performed within the PRISM model checker [8] and produces the full state space of the developed CTMC. Solving the aforementioned CTMC, each state is assigned a unique probability enabling the analyst to derive probabilistic results. We verify the produced state space using logical formulas defined in the Continuous Stochastic Logic (CSL) [11] and, in this way, we calculate the impact of failures on WSN. Although our study considers a WSN, it provides a general framework that can be exploited by network designers to analyze critical infrastructures under their specifications.

The remainder of the paper is organized as follows. A related work overview along with our contribution are discussed in Section II. The survivability specification definition, our analysis is based on, is presented in Section III. Section IV elaborates on details concerning the WSN environment, the implementation of our CTMC model and the faults and attacks injected in it. Probabilistic verification results derived by the CTMC model of a WSN are presented and discussed in

S. Petridou is with the Department of Informatics, Aristotle University of Thessaloniki, Greece, e-mail: spetrido@csd.auth.gr.

S. Basagiannis is with the United Technologies Research Centre (UTRC), Cork, Ireland, e-mail: basagis@utrc.utc.com.

M. Roumeliotis is with the Department of Technology Management, University of Macedonia, Thessaloniki, Greece, e-mail: manos@uom.gr.

Section V. The paper concludes in Section VI with the impact of the proposed approach and our future work insights.

II. Related Work

Survivability analysis of WSNs is strongly associated with the fact that, there is a high rate of failures, affecting the ability of the network to continue operating in order to fulfill its mission, e.g., data aggregation. Some well-known research papers in the past and recent bibliography [1], [2], [6], [12] address issues and challenges concerning survivability analysis of networked systems. They provide a good description of the concept of survivability, along with a theoretical framework, which motivates us to focus on a specific analysis coping with WSNs and their technical specifications. Indicative papers of the latest bibliography [10], [13], [14], which concern not only WSNs but networks in general, also deal with the survivability issue with a more application-oriented way.

In particular, authors in [13] present a survivability analysis approach, which is illustrated on an abstract model of a networked system, in their case the United States Payment System. Their model is based upon Constraint Markov Decision Processes (CMDPs), which constitute an arduous handcrafted approach and, as such, they present limited scalability. Furthermore, CMDPs have the drawback that they do not consider time, but according to [1], [2] timeliness is a basic survivability requirement. In [10], authors study the survivability performance of wireless ad-hoc networks using an analytical CTMC model. Although their approach is interesting, it becomes a highly intensive manual task for complex networks, since it requires the solution of complicated Markov chains for deriving steady-state probabilities. For example, the authors consider only faults as abnormal events and they do not inject attacks in their model. On the other hand, authors in [14] consider only attack events, while their focus is on WSNs rejuvenation after them. While their point of view, that is to derive the expected cost due to sensors' downtime, is insightful, it does not capture all survivability concept aspects. The current work proposes a methodology, which studies the issue of WSN survivability considering both faults and attacks.

A. Motivation and Contribution

Our motivation is to address the scalability issues concerning the survivability analysis of a WSN. In our model, scalability refers to the network components (i.e., sensor nodes and their links), to the failure events (i.e., faults and attacks), as well as to the analysis results (i.e., the size of the WSN model's state space and the time to resolve this model).

This paper provides the following contributions:

1) The adoption of formal methods and especially model checking [15], [16] for survivability analysis of a WSN, through the full state space exploration of its corresponding model. Probabilistic model checking has also been applied by the authors [17], [18], [19] to quantitatively analyze security protocols, in order to study the effects of security parameters in mobile environment, such as key lengths and cipher suites, as well as the cost of security threats.
2) The integration of probabilistic model checking in a mathematical framework [7], defining survivability specification, in order to build a finite-state CTMC model. In this way, we proceeded to the implementation of a fault and attack injection mechanism which embedded in the CTMC model. To the best of our knowledge this is the first time that, a CTMC model along with a fault and attack injection mechanism is automatically verified using the PRISM model checker.
3) A case study and an evaluation which includes quantitative results over the full state space of the CTMC model. The evaluation proves the feasibility of our approach on a 1000-nodes testbed.
4) The development of a flexible framework that can be exploited by networks and protocols' designers towards evaluating their products prior to the implementation phase. Insights for this are derived by our results, which relate a network's tolerance to its size and complexity. Thus, the current work is protocol independent opposed to [17], [18], [19] which analyze specific protocols.

III. Survivability Characterization Using CTMC

There are a number of definitions of network survivability [2], [6], [13]. The one introduced by Software Engineering Institute emphasizes timeliness and survivability under attacks and failures [2], considering in this way the survivability as a means of assessing the extent to which a network supports its services under abnormal events. According to [10], although the above definitions provide a good description of the concept of network survivability, they do not supply a mathematical framework which allows the quantitative study of survivability. For this purpose, Knight and Sullivan [7] introduced a definition of survivability with the mathematical precision that enables its quantitative analysis. Thus, in this paper, we employ the definition of [7] and we proceed to the survivability characterization of a WSN, by building the full state space of a stochastic model and solving it for the probability that the network resides in states of our interest.

Definition 1: A survivability specification $\mathcal{S}$ for a wireless sensor network is defined as a four-tuple, i.e., $\mathcal{S} = (\mathcal{E}, \mathcal{R}, \mathcal{P}, \mathcal{M}_{CTMC})$ where:

- $\mathcal{E}$ is a set of specifications and assumptions that adhere to the environment of the WSN to be analyzed. This may include details of network architecture, components, topology and infrastructure for data transmission and connectivity among components. Furthermore, it may refer to any anticipated changes that might occur in the network affecting its operation, as well as to various hazards to which the network might be exposed.
- $\mathcal{R}$ is a set of specifications, each of which is a complete statement that represents a service, that the network environment $\mathcal{E}$ continues to provide despite the presence of failures. But, to describe network's tolerance in the environment $\mathcal{E}$, the fault model $\mathcal{F}$ of the network should also be defined. $\mathcal{F}$ refers to the set of failures F_i that might occur in the analyzed network, where $F_i \in \mathcal{F}$ and $\forall F_i$ there is an impact function $Im(F_i)$, which quantifies

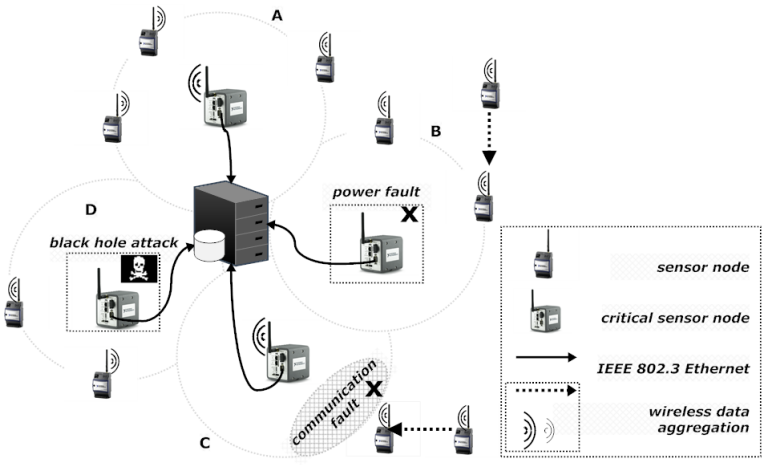

Fig. 1. Wireless sensor network architecture

the impact of a failure to the network. The above entails that $\mathcal{F} \subset \mathcal{R}$.

- $\mathcal{P}$ is the probability mass function across the set of specifications $\mathcal{R}$. Since a probability is assigned to each element of the set $\mathcal{R}$, the sum of probabilities over the whole set $\mathcal{R}$ is 1. The probability associated with each $F_i \in \mathcal{F}$ defines the fraction of operating time during which failures occur in the network.
- $\mathcal{M}_{CTMC}$ is a finite-state CTMC model composing a wireless sensor network's operation in line with the aforementioned specifications, namely $\mathcal{E}, \mathcal{R}$ and $\mathcal{P}$. A $\mathcal{M}_{CTMC}$ is defined as a four-tuple $(S, \bar{s_i}, Rt, L)$, where:
  - S is a finite set of states with a unique label assigned to each of them and which compose the set of specifications $\mathcal{R}$,
  - $\bar{s_i} \in S$ is the initial state of the CTMC model,
  - $Rt : S \times S \rightarrow \Re_{\geq 0}$ is the transition rate matrix and
  - $L : S \rightarrow 2^{AP}$ is the labeling function of atomic propositions AP that are true in S.

A WSN is survivable if it complies with the above survivability specification. By this definition, to reside in each state of the $\mathcal{M}_{CTMC}$ has a unique probability, and the network's survivability is determined by the probability on the network residing in preferred states. The probability mass function $\mathcal{P}$ is computed by automatically solving the $\mathcal{M}_{CTMC}$ model for the set of states in S, as it is described in Section IV-C. Once the $\mathcal{M}_{CTMC}$ model is built, the properties of our interest are verified using CSL formulas [11]. Thus, quantitative results over the full state space of the $\mathcal{M}_{CTMC}$ model are produced.

IV. Survivability Analysis of Wireless Sensor Networks According to $\mathcal{S}$

Both the survivability definition of [2] and the survivability specification $\mathcal{S}$ dictate that failures' occurrence is of utmost importance in network survivable analysis, since failures reflect the performance degradation of the network under abnormal events. In fact, according to [12], network survivability evaluation should emphasize on the assessment of the frequency of failures as well as on the assessment of their impact. Definition 1, which is employed throughout this paper, is connected with both frequency of failures and their impact in that, the first one is associated with the probability of network residing in undesirable states, whilst the latter is related to the probability that the network will accomplish its mission, e.g., data aggregation, in a finite amount of time. Measures that reflect the impact of failures are data loss, delay and compromised data. Thus, our network survivability analysis is based on four measures, namely the frequency of failures, the data loss, the delay and the compromised data due to failures, while three types of failures are considered in the fault model $\mathcal{F}$, i.e., node, link and attack failures, as it is described in Section IV-B.

TABLE I
Basic symbols' notation

Parameter	Notion
cs	number of critical sensor nodes
s	number of sensor nodes
λ_1	rate of node failure F_1
λ_2	rate of link failure F_2
λ_3	rate of attack failure F_3
μ	repair rate of failure F_2

A. Environment

Fig. 1 illustrates the environment $\mathcal{E}$ that it is considered in survivability specification $\mathcal{S}$ of our analysis. In particular, we assume a WSN consisting of a number of nodes, distinguished as critical sensor nodes cs and simple sensor nodes s, according to the notation of Table I, and a central base station connected to a database. Typically, the mission of such a network of tens or hundreds of nodes, that operate without human interaction for weeks or months at a time, is monitoring the environment through data aggregation [20], [21], [22]. In our scenario, sensor nodes s, which are randomly located in and out of domains A, B, C and D, collect data, such as light, temperature and humidity and sent them to critical sensor nodes cs through a routing tree. Data is also gathered by the cs nodes. Once the critical sensor nodes cs aggregate data from sensor nodes s, they submit them, along with their own data,

to the central base station though an IEEE 802.3 connection, as shown in failure-free domain A of Fig. 1.

A WSN architecture, such as the one of Fig. 1, should fulfill its mission with little or no human intervention, since placing new nodes or recharging batteries of nodes in bird nets, earthquake test sites or heating ducts is time consuming and expensive [20]. Thus, it is essential that failures be identified and their impact be studied when we proceed to survivability analysis of such networks.

B. Fault Model

The fault model $\mathcal{F}$ refers to the set of failures F_i that might occur in the network environment $\mathcal{E}$, described in Section IV-A. As depicted in Fig. 1, the fault model $\mathcal{F}$ injects faults and attacks in $\mathcal{M}_{CTMC}$ which lead to the following failures:

- F_1: *Node failure* is due to power fault. F_1 indicates that a sensor node cs may be out of order because of its battery discharging, as shown in domain B of Fig. 1. We assume that recharging of batteries cannot be tolerated in the network of our interest and this constitutes cs nodes critical, since, their operation to sent the collected data to the base station affects network's mission. According to Table I, λ_1 is the rate of F_1 occurrence.
- F_2: *Link failure* is due to communication fault. F_2 denotes that a sensor node s may lose its wireless link connection with a critical node cs, as shown in domain C of Fig. 1. We assume that a link connection can be re-established in a given time boundary. Link failure F_2 occurs with a rate λ_2 and is repaired with a rate μ.
- F_3: *Attack failure* is due to black hole attacks. F_3 indicates that a critical sensor node cs may be compromised by an adversary, as shown in domain D of Fig. 1. In this case, the malicious node refuses to forward every packet received, creating a black hole in WSN. We assume that a cs node is compromised with a rate λ_3 and it cannot be recovered to its proper function [23], [24].

Given that failures F_1, F_2 and F_3 influence network's mission, we define their impact $Im(F_1), Im(F_2), Im(F_3)$ respectively, as follows:

- $Im(F_1)$: failure F_1 lead to *data loss*. Once a critical node cs becomes powered-off, it loses all collected data, which, in turn, will not be submitted to the base station.
- $Im(F_2)$: failure F_2 leads to *data delay*. Once a sensor node s loses its link connection, it will be delayed in submitting the data that gathers, since it should wait for link repair.
- $Im(F_3)$: failure F_3 lead to *compromised data*. Once a critical node cs becomes compromised, it refuses to forward the data that aggregate. $Im(F_3)$ is differentiated from $Im(F_1)$, since the submitted data is not lost but kept by a malicious node.

Thereby, $Im(F_1), Im(F_2)$ and $Im(F_3)$ along with frequency of them are the four measures that our analysis is based on.

C. Implementation of $\mathcal{M}_{CTMC}$ model

As stated in Section III, we modeled the WSN depicted in Fig. 1 using a Continuous-Time Markov Chain (CTMC) [25]. A CTMC model is considered to be a stochastic process [26], that is expected to verify properties of interest, e.g., the steady-state probability of a WSN to correctly collect all data from its nodes in a finite amount of time.

Our $\mathcal{M}_{CTMC}$ is build according to a number of assumptions. Node failures F_1 concern only the critical sensor nodes cs failures, since these nodes affect decisively network's operation towards its mission. We consider two operational modes for critical nodes cs, i.e. {active, failed}, the local variables Act_{cs} and Fld_{cs} which represent the number of cs in the corresponding mode and the constant max_cs which denotes the maximum number of cs. $\mathcal{M}_{CTMC}$ model begins with $Act_{cs} = max\_cs$, while the range of values of Act_{cs} and Fld_{cs} is $[0, max\_cs]$. At any given time it stands that $Act_{cs} + Fld_{cs} = max_{cs}$. Then, critical sensor nodes cs may make the transition to failed mode at any time. Thus, in our finite state model, we allow a nondeterministic transition to a state where mode of a cs is equal to failed, i.e., $Fld_{cs} = Fld_{cs} + 1$ and $Act_{cs} = Act_{cs} - 1$. When a critical node cs makes the transition to failed mode, it does not contribute to the network's mission and it stays in that mode forever, i.e., for time instances t_1, t_2 with $t_2 > t_1$ it stands that $Fld_{cs}(t_2) \geq Fld_{cs}(t_1)$.

Link failures F_2 concern the connectivity among a number of sensor nodes s and the critical node cs to whom they report their collected data. We consider two operational modes for active critical nodes Act_{cs}, i.e. {linked, unlinked}, in line with F_2. Local variables Lnk_{cs} and $Ulnk_{cs}$ denote the number of Act_{cs} in the corresponding mode. The $\mathcal{M}_{CTMC}$ model begins with no link failures, i.e., $Lnk_{cs} = Act_{cs}$, the range of values of Lnk_{cs} and $Ulnk_{cs}$ is $[0, max_{cs}]$ and at any given time it stands that $Lnk_{cs} + Ulnk_{cs} = Act_{cs}$. Then, a link failure may occur at any time and then an active critical sensor node Act_{cs} makes the transition to unlinked mode. Thus, in our finite state model, we allow a nondeterministic transition to the state where the mode of Act_{cs} is equal to unlinked, i.e., $Ulnk_{cs} = Ulnk_{cs} + 1$ and $Lnk_{cs} = Lnk_{cs} - 1$. When an active critical node Act_{cs} makes the transition to unlinked mode, it does not gather data from nodes s. However, a link failure may be randomly restored with a rate μ, as shown on Table I. During a link failure, nodes s keep their data and submit them after link re-establishment.

Attack failures F_3 concern only cs nodes and especially Lnk_{cs}, since an adversary can influence the network's mission by compromising linked critical nodes. In line with F_3, Lnk_{cs} are distinguished in {compromised, uncompromised} and the local variables Cmp_{cs}, $Ucmp_{cs}$ denote the number of Lnk_{cs} in the corresponding mode. Model begins with no attacks, i.e., $Ucmp_{cs} = Lnk_{cs}$, the range of values of Cmp_{cs} and $Ucmp_{cs}$ is $[0, max\_cs]$ and, at any given time, it stands that $Cmp_{cs} + Ucmp_{cs} = Lnk_{cs}$. In our finite state model, a black hole attack is represented by an Lnk_{cs} node making a transition to compromised mode, i.e., $Cmp_{cs} = Cmp_{cs} + 1$ and $Ucmp_{cs} = Ucmp_{cs} - 1$. When a cs node is compromised,

TABLE II
MODEL CHECKING RESULTS OF M_{CTMC}

Sensor nodes s	Total states of S	Transitions	Iterations	Time (sec)
10	2073	5768	16	0.009
50	9846	28168	24	0.021
100	19566	56228	34	0.135
1000	194526	561306	214	4.624
10000	1944126	5612108	2014	548.588

it drops the packets it receives and thus it does not contribute to the network's mission. We assume that it stays in that mode forever, i.e., for time instances t_1, t_2 with $t_2 > t_1$ it stands that $Cmp_{cs}(t_2) \geq Cmp_{cs}(t_1)$.

The developed model contains one (1) module, namely $M = \{M_{wsn}\}$, which represents the behavior of the WSN of Fig. 1. The local variables s, cs affect the state space S of module M_{wsn}, while the variables $Act_{cs}, Fld_{cs}, Lnk_{cs}, Ulnk_{cs}$ regulate the paths between discrete states of the produced state space. In general, the waiting time t of a transition from a state i to a state j, where $i, j \in S$, is mapped by a negative exponential distribution $e^{-Rt(i,j)\cdot t}$, the parameter of which is the transition rate, i.e., $Rt(i,j)$. In our case, this exponential distribution is defined in line with the rates $\lambda_1, \lambda_2, \lambda_3, \mu$, which express the rate of node failure F_1, link failure F_2, attack failure F_3 and repair of F_2, respectively.

Once the finite-state model $\mathcal{M}_{CTMC}$ is build, each state is assigned with a unique probability according to the probability mass function $\mathcal{P}$. By solving the produced state space, we can derive probabilistic results, such as the probability of network residing in undesirable states and the probability that the network will accomplish its mission, e.g., data aggregation. According to Definition 1, the first probability is associated with the frequency of failures, while the latter is related to their impact, i.e., data loss $Im(F_1)$, delay $Im(F_2)$ and compromised data $Im(F_3)$. These four measures used for producing the survivability evaluation results presented in Section V.

V. SURVIVABILITY EVALUATION

This section presents the survivability evaluation results derived by the quantitative analysis of Definition 1. The behavior of the WSN, presented in Section IV-A, is modeled in the environment of the PRISM model checker using CTMCs, as described in Section IV-C. Our $\mathcal{M}_{CTMC}$ allows parameterization, since the local variables s and cs, which denote the scale of the network under analysis and affect the produced state space S, can be varied. The same stands for the values of rates λ_1, λ_2, λ_3 and μ. In results that follow, λ_1, λ_2 and λ_3 are varied in order to pinpoint the $Im(F_1)$, $Im(F_2)$ and $Im(F_3)$, while the repair rate of F_2 is fixed at $\mu = 0.05$, which corresponds to an average repair time equal to 20 sec. We also define the transmission rate to be $push = 10$ nodes per second. Values of μ and $push$ are indicative and derived by [10], but they can obviously be adjusted according to the network requirements.

Results derived by a dual-core 2.5GHz machine under Ubuntu 10.04 distribution with 4GB of RAM. Table II presents numerically the way that PRISM model checker build the state space of $\mathcal{M}_{CTMC}$ for different values of s. For these results, we assume that $cs = 5$, $\lambda_1 = 0.2$ nodes/sec and $\lambda_2 = 0.5$ links/sec. Time results indicate the scalability of our $\mathcal{M}_{CTMC}$ model, which allows the survivability analysis of complex networks of the scale of 10^4 sensor nodes. The remaining values of Table II concern the total number of the produced states, along with the transitions between them, and the iterations needed to solve the CTMC model. The magnitude of these numbers denotes the depth of our analysis.

We proceed to the quantitative analysis of the developed M_{CTMC} model defining the properties of the WSN as queries in stochastic logic formulas. These properties represent partial specifications of the steady-state and the transient behavior of M_{CTMC}, as defined in survivability specification $\mathcal{S}$. Then, the probabilistic model checker PRISM resolves the actual probability of network residing in states of interest based upon these formulas. Queries are formed using Continuous Stochastic Logic (CSL) [11].

According to our implementation assumptions of Section IV-C, the $\mathcal{M}_{CTMC}$ model begins with all critical sensors being active, i.e., $Act_{cs} = max\_cs$, and no link and attack failures, i.e., $Lnk_{cs} = Ucmp_{cs} = Act_{cs}$. Then, failures F_1, F_2 and F_3 occur with rates λ_1, λ_2 and λ_3, and only F_2 are repaired with rate μ. Since, failures F_1 and F_3 are not fixed, the steady state availability of the network depends on the availability of critical sensors cs. If the network's availability and non-availability are denoted by $P_{Availability}$ and $P_{NonAvailability}$, respectively, then it holds that:

$$P_{Availability} + P_{NonAvailability} = 1 \tag{1}$$

$P_{Availability}$ of Eq. 1 is defined as the probability of the network residing in states of S, where at least one (1) critical sensor is active and uncompromised, i.e., $Act_{cs} \geq 1$ and $Ucmp_{cs} \geq 1$. Thus, it can be calculated as follows:

$$P_{Availability} = \sum_{i,j} P_{(i,j)},\ \forall i,j \in S : Rt(i,j) > 0 \\ and\ Act_{cs} \geq 1\ and\ Ucmp_{cs} \geq 1 \tag{2}$$

Eq. 2 entails that $P_{NonAvailability}$ is associated with the probability of none cs nodes being active due to F_1 or all active cs nodes being compromised due to F_3. Thus:

$$P_{NonAvailability} = \sum_{i,j} P_{(i,j)},\ \forall i,j \in S : Rt(i,j) > 0\ and \\ ((Fld_{cs} = max_{cs})\ or\ (Act_{cs} \geq 1\ and\ Cmp_{cs} = Act_{cs})) \tag{3}$$

Based on Eq. 3, we can determine all the paths of the produced state space of the $\mathcal{M}_{CTMC}$ which lead to the

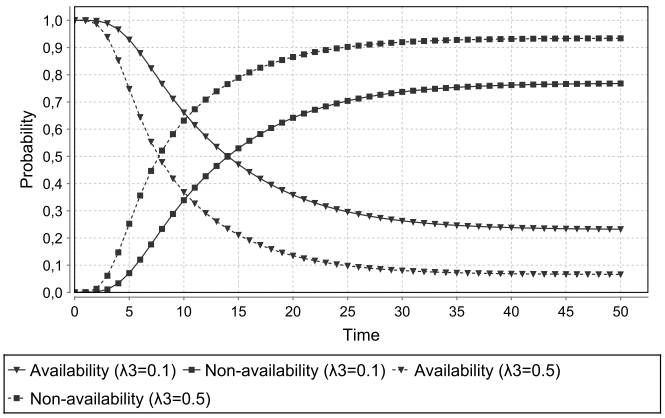

Fig. 2. Network's availability and non-availability expressed in line with λ_3

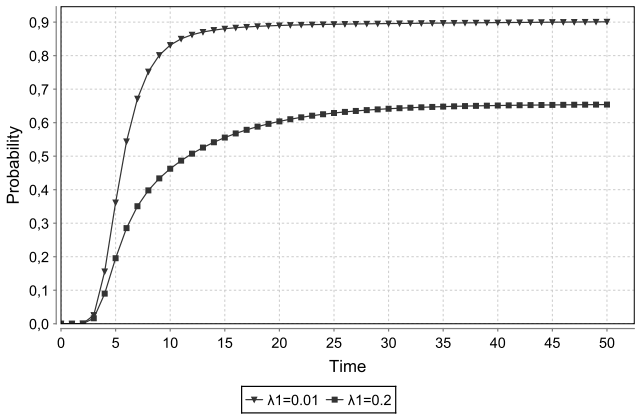

Fig. 3. Data loss, i.e., $Im(F_1)$, is expressed in line with the probability that the network will accomplish data gathering at any given time step, when the probability is calculated as a function of critical sensors' failure rate λ_1

network residing in undesirable states (Eq. 3). Such a property is expressed by $Query\_1$ as follows:

$Query\_1 : P_{NonAvailability} =?\ [true\ U <= C_0\ ((Fld_{cs} = max\_cs)\ or\ (Act_{cs} \geq 1\ and\ Cmp_{cs} = Act_{cs}))]$,
$s = 500, cs = 5, \lambda_1 = 0.2, \lambda_2 = 0.5$

$Query\_1$ is applied to a network of $s = 100$ sensor nodes, $cs = 5$ critical nodes which has $\lambda_1 = 0.2$ nodes/sec and $\lambda_2 = 0.5$ links/sec. It searches out the probability of a path appeared in the state space S, where network will stop operating towards its mission (Eq. 3), for different values of λ_3. A similar query is launched for computing $P_{Availability}$ and the results are shown in Fig. 2. A first observation is that, at any given time step t the curves of Fig. 2, which depict $P_{Availability}$ and $P_{NonAvailability}$, obey Eq. 1, indicating the proper operation of $\mathcal{M}_{CTMC}$. It is natural, that as time passes, the network's availability declines, since F_1 and F_3 are not repaired, and for the same reason, non-availability is increased. At the same time, Fig. 2 depicts the influence of black hole attacks (F_3) in WSN. Network's availability drops dramatically as λ_3 increases.

Once we quantify network's availability (Fig. 2), we can proceed to the first measure evaluation, i.e., frequency of failures. In our $\mathcal{M}_{CTMC}$ the failures' frequency is defined as $\lambda_1 \cdot P_{Availability} + \lambda_3 \cdot P_{Availability}$ and refers to the frequency of F_1 and F_3. This is due to the fact that cs nodes affect decisively the network's operation towards its mission and their failures, either because of power fault (F_1) or because of black hole attack (F_3), are not repaired. In the above definition $\lambda_1, \lambda_3 \in [0..1]$. Thus, at the one end, when $\lambda_1 \approx 0$ we evaluate the frequency of the attack failures for different values of λ_3, whilst, at the other end, when $\lambda_3 \approx 0$ we evaluate the frequency of the node failures for different values of λ_1. For any other values of λ_1, λ_3 failures frequency definition considers both F_1 and F_3. In general, as rates λ_1, λ_3 increase the network's availability decreases. However, although failures will be more frequent at the beginning of $\mathcal{M}_{CTMC}$, this trend is reversed with time. The explanation is that, as network operates and its availability decreases, less Act_{cs} will be available either to be failed (F_1) or to be compromised (F_3).

Apart from the frequency of failures, data loss, delay and compromised data are the other three measures which were evaluated in our survivability analysis. As described in Section IV-B, data loss is due to $Im(F_1)$, delay expresses the $Im(F_2)$ and compromised data is owed to $Im(F_3)$. Thus, in order to study the impact of failures, we launched PRISM executions for different values of λ_1 for $Im(F_1)$, λ_2 for $Im(F_2)$ and λ_3 for $Im(F_3)$. For data loss and delay, the property that is quantitatively verified is the one that explores the produced state space of the $\mathcal{M}_{CTMC}$ for the probability of all sensor nodes s successfully submitting their data to the critical sensors cs in a predefined time boundary. This property is expressed by $Query\_2$ as follows:

$Query\_2 : P_{Success} =?\ [true\ U <= C_0\ Success = true]$,
$s = 100, cs = 5$

The interpretation of $Query\_2$ is: "which is the actual probability of the all network's sensor nodes s to submit their data to the critical nodes cs in a finite amount of time?". In this query, *Success* is a logical formulae defined in our $\mathcal{M}_{CTMC}$ model, in order to control that nodes s will successfully submit their data.

When this probability is calculated in line with critical sensors' failure rate λ_1, it expresses data loss, since the impact of F_1 is data loss, as defined in Section IV-B. Fig. 4 depicts the results derived for $\lambda_1 = 0.01$ and 0.2, nodes/sec while λ_2, λ_3 are fixed at 0.1. Since $\lambda_1 = 0.01$ is a relatively low value of critical sensors' failure rate, it is expected that the corresponding curve in Fig. 3 reaches probability 1. This means that all sensors s will succeed in submitting their data after 50 time steps, according to the graph, since a low value of λ_1 entails that many cs will remain active to aggregate the data. However, due to λ_3, some of the Act_{cs} will be compromised and they will not deliver the packets that receive causing $P_{Success} < 1$. On the other hand, a high value of nodes' failure rate dramatically affects the network's behavior towards accomplishing its mission. Thus, the curve for $\lambda_1 = 0.2$ falls significantly, reaching the probability of

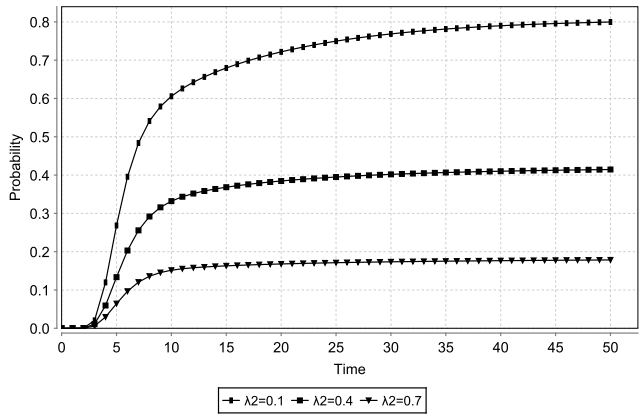

Fig. 4. Delay, i.e., $Im(F_2)$, is expressed in line with the probability that the network will accomplish data gathering at any given time step, when the probability is calculated as a function of link failure rate λ_2

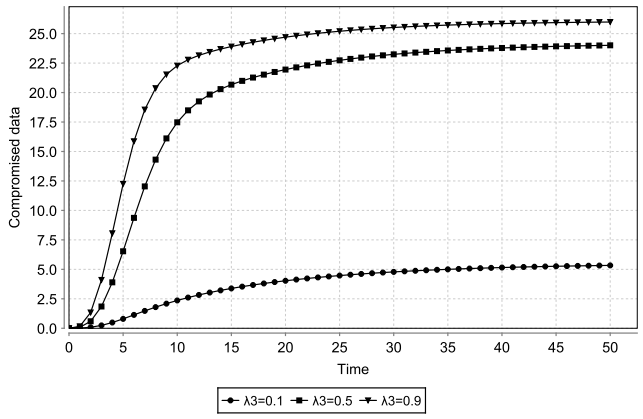

Fig. 5. Compromised data, i.e., $Im(F_3)$, are calculated in line with the number of s nodes whose data are aggregated by compromised sensors cs nodes for different values of attack failure rate λ_3

0.65 after 50 times steps.

In order to study network's delay, we calculate the probability of $Query\_2$ for different values of links' failure rate λ_2. Failures F_2 in contrast to F_1 and F_3 are repaired and, consequently, they affect the timing of network's mission accomplishment. Thus, we re-launched $Query\_2$ for $\lambda_2 = 0.1, 0.4$ and 0.7 links/sec, while λ_1, λ_3 are fixed at 0.1. The fact that no curve of Fig. 4 reaches probability 1 indicates the impact of F_1 and F_3. The impact of F_2 is evident from the curves' arrangement, which denotes that the successful data aggregation will be delayed according to λ_2. In particular, higher values of links' failure rate entail delay, as indicated by lower probability $P_{Success}$, since unlinked cs nodes ($Ulnk_{cs}$) do not contribute to the network's mission.

Besides the above quantitative probabilistic results, one of the benefits of probabilistic model checking is the capability of the analysis to produce quantitative results derived by user-defined reward structures [27]. For the last measure, i.e., compromised data, we define an instantaneous reward structure. This structure attaches to the produced state space S weights equal to the number of s nodes whose data has been aggregated by compromised sensors cs due to F_3. We calculate the compromised data, by defining the instantaneous reward query $Query\_3$, as follows:

$Query\_3 : R =?\ [I = C_0]$,
$C_0 = 50, s = 100, cs = 5, \lambda_1 = 0.1, \lambda_2 = 0.1$

$Query\_3$ is applied to a network of $s = 100$ sensor nodes and $cs = 5$ critical nodes, where $\lambda_3 = 0.1, 0.5$ and 0.9 attacks/sec and λ_1, λ_2 are 0.1. The instantaneous reward property of the type $R =?\ [\mathrm{I}^{=\mathrm{t}}]$, used in this query, is associated with each path of the model for up to a given time boundary t and it is calculated for different values of attacks' failure rate λ_3. The results of network's behavior towards the last measure is depicted in Fig. 5. The ascending trend of all three curves denotes the effect of black hole attack, i.e., as time passes more s are "deceived" and submit their data to compromised cs, increasing the last measure. The impact of λ_3 is evident from the curves' arrangement, since higher attacks' failure rates lead to more compromised data.

Finally, an interesting study for the analyzed network is to probabilistically quantify its success in aggregating data, according to the ratio of sensors s and critical sensors cs in a given time boundary. For this purpose, we re-run query $Query\_2$ defining $\lambda_1 = 0.2$ nodes/sec, $\lambda_2 = 0.5$ links/sec and $\lambda_3 = 0.1$ attacks/sec. Sensor nodes s are varied from 50 to 250 with a step of 50, while critical sensors cs, deployed at the x-axis of Fig. 6, range from 1 to 10.

In Fig. 6, we indicatively observe that the probability of $s = 50$ sensors successfully submitting their data in a network with $cs = 5$ critical sensors is 0.7, while for $s = 100$ the probability falls at 0.25. This is natural, since for the same number of cs, the more the sensors s, the lower the probability of the network accomplishing data aggregation. On the other hand, the ascending trend of each curve indicates that for the same number of s, the more the critical sensors cs, the higher the aforementioned probability. Although, adding critical nodes cs may increase network's survivability according to the measures defined in this paper, such an action will affect both the cost of the network's implementation and complexity. The latter query may be of great importance for the network design analysts, since, prior to the implementation phase, they will have the capability of studying their network's design in line with the tradeoff between maximizing network's operation and minimizing its implementation cost.

The aforementioned capability along with the scalabity of the proposed $\mathcal{M}_{CTMC}$ differentiate the current work from the related research studies [10], which are also based on the use of Markov stochastic chains. The analytical approach of [10] implies restrictions considering both the scale and the complexity of the analyzed network. The proposed methodology creates flexible CTMC models that can incorporate additional features, which mutually affect each other. Thus, we can enlarge the complexity of WSN without sacrificing the capability of analyzing it, since the produced CTMCs can be solved automatically and timely (Table II) using probabilistic model checking. Thereby, instead of proposing a solution to an instance of WSN, we provide a methodology which can be exploited by network designers in order to fit their

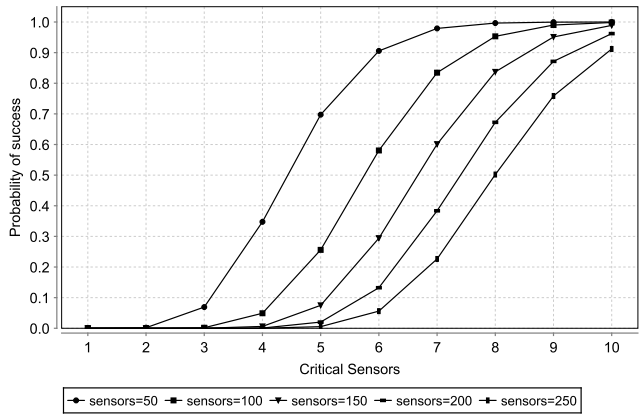

Fig. 6. Tuning the number of critical sensors cs and sensor nodes s determines the probability of the network accomplishing data gathering at a given time step

requirements to the proposed model.

VI. Conclusions and Future Work

In this paper, we introduce an automated quantitative analysis using probabilistic model checking as a means of evaluating survivability of WSNs. Exploiting a well-defined mathematical framework for survivability characterization, we proceed to the development of a CTMC model, that evaluates WSN survivability in line with four measures, namely the frequency of failures, data loss, data delay and compromised data due to a variety of failures. Our approach provides a series of probabilistic results that pinpoint the analysis accuracy when a number of network parameters are varied. Our method is flexible enough to be utilized under different type of networks with their own specifications. As a future perspective, we aim at enhancing the proposed quantitative analysis with energy consumption parameters, such as battery life of sensor nodes, as well as with issues regarding the cost of faults recovery.

VII. Acknowledgment

The authors would like to thank Fotis Loukos for his fruitful discussions with respect to the study on WSNs.

References

[1] R. Ellison, R. Linger, T. Longstaff, and N. Mead, "Survivable network system analysis: A case study," *IEEE Softw.*, vol. 16, no. 4, pp. 70–77, 1999.

[2] R. Ellison, D. Fisher, R. Linger, H. Lipson, T. Longstaff, and N. Mead, "Survivable network systems: An emerging discipline," Software Engineering Institute, Carnegie Mellon University, Tech. Rep. CMU/SEI-97-TR-013, 1997.

[3] M. S. Obaidat, "Future and challenges of the security of e-systems and computer networks," in *Proc. of the Int. Conference on Security and Cryptography*, Spain, Jul. 2007, pp. 15–16.

[4] A. Meier, M. Woehrle, M. Weise, J. Beutel, and L. Thiele, "Nose: Efficient maintenance and initialization of wireless sensor networks," in *In Proc. of the 6th Annual IEEE Communications Society Conf. on Sensor, Mesh, and Ad Hoc Communications and Networks*, Italy, Jun. 2009, pp. 1–9.

[5] J. Alonso-Zarate, E. Stavrou, A. Stamou, P. Angelidis, L. Alonso, and C. Verikoukis, "Energy-efficiency evaluation of a medium access control protocol for cooperative arq," in *to appear in Proc. of IEEE Int. Conference on Communications*, Japan, Jun. 2011.

[6] J. Sterbenz, R. Krishnan, R. Hain, A. Jackson, D. Levin, R. Ramanathan, and J. Zao, "Survivable mobile wireless networks: Issues, challenges, and research directions," in *Proc. of the 3rd ACM Workshop on Wireless Security*, Atlanta, GA, USA, Sep. 2002, pp. 31–40.

[7] J. Knight and K. Sullivan, "On the definition of survivability," Department of Computer Science, University of Virginia, Tech. Rep. CS-TR-33-00, 2000.

[8] M. Kwiatkowska, G. Norman, and D. Parker, "Prism 2.0: A tool for probabilistic model checking," in *Proc. of the 1st Int. Conf. on Quantitative Evaluation of Systems*, Netherlands, sep 2004, pp. 322–323.

[9] M. Kwiatkowska, "Quantitative verification: models techniques and tools," in *Proc. 6th joint meeting of the European Software Engineering Conf. and the ACM SIGSOFT Symposium on the Foundations of Software Engineering*, Croatia, sep 2007, pp. 449–458.

[10] D. Chen, S. Garg, and K. Trivedi, "Network survivability performance evaluation: A quantitative approach with applications in wireless ad-hoc networks," in *Proc. of the 5th ACM Int. Workshop on Modeling Analysis and Simulation of Wireless and Mobile Systems (MSWiM'02)*, Atlanta, GA, USA, Sep. 2002, pp. 61–68.

[11] A. Aziz, K. Sanwal, V. Singhal, and R. Brayton, "Model-checking continous-time markov chains," *ACM Transactions on Computational Logic*, vol. 1, no. 1, pp. 162–170, 2000.

[12] A. Zolfaghari and F. Kaudel, "Framework for network survivability performance," *IEEE J. Sel. Areas Commun.*, vol. 12, no. 1, pp. 46–51, 1994.

[13] S. Jha and J. Wing, "Survivability analysis of networked systems," in *Proc. of the 23rd Int. Conf. on Software Engineering*, Canada, May 2001, pp. 307–317.

[14] S. Parvin, D. Kim, S. Lee, and J. Park, "Achieving availability and survivability in wireless sensor networks by software rejuvenation," in *In Proc. of the 4th Int. workshop on Security, privacy and trust in pervasive and ubiquitous computing*, Italy, Jul. 2008, pp. 13–18.

[15] E. Clarke, O. Grumberg, and D. Peled, *Model Checking*. The MIT Press, 2000.

[16] C. Baier and J. Katoen, *Principles of Model Checking*. The MIT Press, 2008.

[17] S. Basagiannis, S. Petridou, N. Alexiou, G. Papadimitriou, and P. Katsaros, "Quantitative analysis of a certified e-mail protocol in mobile environments: A probabilistic model checking approach," *Computers & Security, Elsevier*, vol. 30, no. 4, pp. 257–2772, 2011.

[18] S. Petridou, S. Basagiannis, N. Alexiou, G. Papadimitriou, and P. Katsaros, "Quantitative model checking of an rsa-based email protocol on mobile devices," in *Proc. of the 16th IEEE Symbosium on Computers and Communications*, Greece, Jun. 2011, pp. 639–845.

[19] S. Basagiannis, P. Katsaros, A. Pombortsis, and N. Alexiou, "Probabilistic model checking for the quantification of dos security threats," *Computers & Security*, vol. 28, no. 6, pp. 450–465, 2009.

[20] S. Madden, M. Franklin, J. Hellerstein, and W. Hong, "Tinydb: An acquisitional query processing system for sensor networks," *ACM Transactions on Database Systems*, vol. 30, no. 1, pp. 122–173, 2005.

[21] A. Cerpa, J. Elson, , M. Hamilton, J. Zhao, D. Estrin, and L. Girod, "Habitat monitoring: Application driver for wireless communications technology," in *ACM SIGCOMM Workshop on Data Communications in Latin America and the Caribbean*, Costa Rica, Apr. 2001, pp. 20–41.

[22] P. Papadimitratos, J. Luo, and J. Hubaux, "A randomized countermeasure against parasitic adversaries in wireless sensor networks," *IEEE J. Sel. Areas Commun.*, vol. 28, no. 7, pp. 1036–1045, 2010.

[23] C. Karlof and D. Wagner, "Secure routing in wireless sensor networks: Attacks and countermeasures," *Ad Hoc Networks, Elsevier*, vol. 1, no. 2–3, pp. 293–315, 2003.

[24] N. Ahmed and S. K. anf S. Jha, "The holes problem in wireless sensor networks: a survey," *Mobile Computing and Communications Review*, vol. 1, no. 2, pp. 4–18, 2005.

[25] W. Stewart, *Introduction to the numerical solution of Markov chains*. Princeton University Press, 1994.

[26] J. Katoen, M. Kwiatkowska, G. Norman, and D. Parker, "Faster and symbolic ctmc model checking," in *Proc. 1st Joint Int. Workshop on Process Algebra and Probabilistic Methods, Performance Modeling and Verification*, Aachen, Germany, Sep. 2001, pp. 23–38.

[27] M. Kwiatkowska, G. Norman, and A. Pacheco, "Model checking expected time and expected reward formulae with random time bounds," *Computers & Mathematics with Applications*, vol. 51, no. 2, pp. 305–316, 2006.

Sophia Petridou ($M'08$) received the Diploma and Ph.D. degrees in Computer Science from the Department of Informatics, Aristotle University of Thessaloniki, Greece in 2000 and 2008 respectively. Her main research interest are in the areas of MAC protocols of optical networks, wireless networks, probabilistic model checking of protocols and clustering algorithms. She has published more than 20 papers in refereed journals and at conference proceedings.

Stylianos Basagiannis ($M'12$) is a senior researcher working at United Technologies Research Centre in Cork, Ireland (UTRC). He holds a Ph.D. (2010) and a Diploma (2004) in Computer Science both awarded from the Department of Informatics, Aristotle University of Thessaloniki, Greece, and an MSc (2005) in Distributed and Multimedia Information Systems from the Department of Computer Science, at Heriot-Watt University of Edinburgh, UK. He has published more than 20 research articles in international journals and conference proceedings in the areas of formal methods, security and safety analysis.

Manos Roumeliotis ($M'81$) received his Diploma in EE from the Aristotle University of Thessaloniki, Greece in 1981, and the MS and Ph.D. degrees in Computer Engineering from Virginia Tech (1983 and 1986). He taught at the Department of Electrical and Computer Engineering of West Virginia University and the Department of Applied Informatics of the University of Macedonia (UoM), Thessaloniki, Greece. Since 2008 he is a Professor at the Department of Technology Management of UoM. His research interests include digital logic simulation, computer architecture, and parallel processing. He has published more than 80 papers in refereed journals and at conference proceedings.