Formal Probabilistic Analysis of Lifetime for a WSN for Border Monitoring

Maissa Elleuch[1,3], Osman Hasan[2], Sofiène Tahar[2], and Mohamed Abid[1]

[1] CES Laboratory, National School of Engineers of Sfax, Sfax University
Soukra Street, 3052 Sfax, Tunisia
maissa.elleuch@ceslab.org
mohamed.abid@enis.rnu.tn
[2] Dept. of Electrical & Computer Engineering, Concordia University
1455 de Maisonneuve W., Montreal, Quebec, H3G 1M8, Canada
{melleuch,o_hasan,tahar}@ece.concordia.ca
[3] Digital Research Center of Sfax
Technopark of Sfax, Tunisia

Abstract. Scheduling sensor nodes in Wireless Sensor Networks (WSN) for lifetime management purposes is a simple and intuitive approach. However, it is also crucial to not compromise on the main performance requirements of the considered application. For mission-critical WSN applications, different Quality of Service (QoS) requirements on network performance have to be satisfied. Nevertheless, traditional techniques usually focus on the average performance values without considering the targeted QoS requirements. In this paper, we provide rigorous formalizations in higher-order logic of the network lifetime maximization problem, under QoS constraints, for randomly-scheduled wireless sensor networks. We also use natural deduction based reasoning to verify the desired properties using theorem proving. In particular, we build upon our earlier developments on coverage and detection analysis to formally analyze the lifetime maximization problem for a border monitoring application.

Keywords: Wireless sensor networks, Performance analysis, Theorem proving, Nodes Scheduling, Network lifetime, Border monitoring

1 Introduction

Wireless Sensor Networks (WSNs) have emerged as a key enabler technology for various surveillance applications [41] including environmental monitoring and object tracking. Since sensors are basically battery-powered, energy saving arises as the most critical requirements. In a WSN for forest fire detection, where sensors are randomly and densely deployed, the network should be able to ensure the monitoring of the area while being functional for a long period. As a wild fire occurs occasionally, some sensors can be intuitively deactivated by partitions to save the whole network energy, and thus extend the network lifetime [35]. In this context, the k-set randomized scheduling [21] is an efficient scheduling approach, which mainly consists in randomly organizing the set of nodes into k subsets.

Scheduling sensors for lifetime management is surely a simple approach, however, it is also crucial to not compromise on the performance of the application. For mission-critical WSN applications, different Quality of Service (QoS) requirements have to be usually satisfied [5, 37]. More generally, QoS is regarded as "the capability of providing assurance that the service requirements of applications can be satisfied" [5]. For example, in a forest fire application, where alarm packets are vital, the WSN should not only cover the whole area, but, ensure also that the fire outbreak is detected within the shortest time with a high probability. Hence, besides the network lifetime, the coverage and the detection performances are critical requirements. Nevertheless, for the k-set randomized scheduling, these performance metrics are completely probabilistic [21, 39]. Hence, some fire outbreaks may not be effectively covered if the surrounding nodes are inactive, due to random scheduling.While the probabilistic aspect poses real challenges on the analysis of WSNs, missing fire intrusion, can have devastating consequences.

The performance of the randomized scheduling has been generally analyzed using paper-and-pencil based probabilistic technique followed by some simulations [33, 18, 20]. However, both paper-and-pencil proof and simulation methods cannot be regarded as completely accurate mainly due to the error proneness of the former and the in-exhaustive nature of the later. Compared with traditional simulation, formal methods are less frequently used for the validation of WSNs. Based on mathematical techniques, formal methods [14] rigorously analyze the theoretical model of the given system. Recently, formal methods have gained a growing interest in the context of WSNs to analyze their functional or quantitative correctness [29, 3, 42], but most of the existing work is focused on the validation of their functional aspects only. Nevertheless, reliable performance evaluation of WSNs constitutes also an extremely challenging aspect.

In this paper, we provide an accurate formal analysis of the network lifetime for randomly-scheduled WSNs. In particular, we are interested in the higher-order-logic formalizations of the lifetime maximization problem, given in [39], under QoS constraints. The main performance requirements here are associated to the network coverage, the detection probability and the detection delay. In earlier work [6, 9], we have presented a formalization of the k-set randomized scheduling algorithm and its main performance properties based on the recent probability theory formalizations [27] in the HOL theorem prover. The practical interest of these developments has been illustrated through the formal analysis of various WSN applications [7–9]. We build upon these theoretical developments to formally show that the optimal solution for the lifetime maximization problem exists, and give the conditions under which the optimal solutions exist. This formal analysis is illustrated through a border security monitoring application.

The rest of this paper is organized as follows. Section 2 reviews some related work. We summarize, in Section 3, the main requirements of this work. Section 4 describes the lifetime maximization problem under QoS requirements. In Section 5, the higher-order-logic formalizations of this problem are provided for a WSN application for border monitoring. Section 6 is devoted to discussions, before concluding the paper in Section 7.

2 Related Work

Theoretical analysis, also known as paper-and-pencil based probabilistic technique, has been widely used to validate randomized scheduling algorithms for WSN. Such analysis consists in constructing a pure theoretical model where the required random variables are determined together with the associated performance metrics. Afterwards, an accurate probabilistic based study is achieved. For validation purposes, simulation, using the Monte Carlo method [24], is finally carried out. The analysis of the randomized scheduling has been usually done using the paper-and-pencil based probabilistic technique [36, 18, 21, 40, 23], followed by simulations on some network scenarios for the main performance metrics. For example, Mamun [25] evaluated the coverage using a pure mathematical model while simulations have been run with specific network sizes and sensing ranges.

Model checking technique [2] has been successfully explored for the validation of various aspects in the WSN context. In [29], the formal analysis of the Optimal Geographical Density Control (OGDC) algorithm, which is a kind of randomized scheduling algorithm, has been performed within the RT-Maude rewriting tool [30]. Several other prominent works reported on the use of model checking for the analysis of WSN protocols include [34, 11, 22], or for the development of formal frameworks [15, 43]. While the main strength of all these works is their formal models and automatic verification, they suffer from the common model checking related problem of state space explosion [2]. Hence, the analysis of the OGDC algorithm [29] has been limited for WSNs with only 6 nodes within a monitored region of $15m \times 15m$. On the other hand, none of the previous works provided a sound modelling of the randomness aspect in WSNs, which constitutes a real limitation since most of the WSN algorithms are probabilistic. In [29], a random function, assumed to be 'good', has been used to model the probabilistic behavior of interest. For Uniform distributions, a sampling value generated by the same random function on a given interval is selected.

To cope with these major problems, probabilistic model checking [31] has also been used for the probabilistic functional analysis of wireless systems [11, 12, 42]. Probabilistic model checking captures the probability modelling for both the system and the property of interest. Nevertheless, the reasoning support for statistical quantities in most of model checkers suffers from many shortcomings. Indeed, expected performance values are usually obtained through several runs on the built model [3, 42]. The obtained results can hardly be termed as exhaustive and thus formally verified.

On the other hand, very few works based on theorem proving [13] exist in the open literature. A synchronization protocol for WSNs, has been analyzed using the Isabelle/HOL theorem prover [16]. The work in [4] built a theorem proving based framework for WSN algorithms based on the PVS system. Nevertheless, the randomness aspect in this work has been characterized by a pseudo-random generator, while the nodes mobility specified through a simple recursive function. Furthermore, the uniform probability, considered for link quality changes, has been just instantiated by a given value throughout the analysis. The analysis

results using the PVS framework can not be hence considered as reliable versus the probability modelling.

Unlike previous works, we provide rigorous formalizations of the network lifetime maximization problem [39], under QoS constraints, for randomly-scheduled WSNs, and use natural deduction based reasoning to verify the desired properties. Traditionally, the simulation-based analysis is usually made for different performance metrics to validate their average values without considering their potential relationship and the desired QoS requirements. In the open literature, few works deal with the formal analysis of QoS properties in WSN. In [34], the authors analyzed Biomedical Sensor Networks (BSN) in terms of QoS requirements on packet delivery ratio, network connectivity and end-to-end delay. Using the model checker UPPAAL, they validate worst-case scenarios of these metrics, and compare the soundness of their results to a well-known WSN simulator. The work in [10] verified the same QoS properties, while focusing on decreasing the power consumption. Although the scalability of the built model is acceptable for BSN, the probabilistic aspect is not considered at all. Due to the sound formalization of probability and its reasoning support available in the HOL theorem prover, the formalizations, given in this paper, are rigourous. In addition, the presented formalizations are generic and completely valid for all values.

3 Preliminaries

In this section, we introduce the probabilistic analysis in the HOL theorem prover. Then, we briefly describe the k-set randomized scheduling algorithm.

3.1 Probabilistic Analysis in HOL

In this work, we utilize the recently developed and most generic probability theory developed by Mhamdi [26], within the HOL theorem prover. By including a Borel space, Mhamdi generalized the previous HOL formalization of measure theory. After specifying the extended real numbers in HOL, he formalized measure, Lebesgue, probability and information theories. The formalization of probability theory in HOL is hence based on the Kolmogorov axiomatic definition of probability. Such formalization thus provides a unified framework for discrete and continuous probability measures.

A probability measure P is a measure function on the sample space Ω and an event is a measurable set within the set F of events which are subsets of Ω. Thus, (Ω, F, P) is a probability space iff it is a measure space whose measure is 1, i.e., $P(\Omega) = 1$. A random variable is a measurable function, satisfying the condition that the inverse image of a measurable set is also measurable (Definition 1).

Definition 1.

```
⊢ ∀X p. real_random_variable X p =
              prob_space p ∧
              (∀x ∈ p_space p ⇒ X x ≠ NegInf ∧ X x ≠ PosInf) ∧
              X ∈ measurable (p_space p,events p) Borel.
```

where X designates the random variable, p is a given probability space, $NegInf$ and $PosInf$ are the higher-order-logic formalizations of negative or positive infinity. $Borel$ is the HOL definition of the Borel sigma algebra which is the smallest sigma algebra generated by the open sets.

The probability distribution of a random variable is the function that accepts a random variable X and a set s and gives the probability of the event $\{X \in s\}$.

Definition 2.

```
⊢ ∀X p.
  distribution p X = (λs. prob p (PREIMAGE X s ∩ p_space p)).
```

The expectation of a random variable X is defined in HOL [26] as its Lebesgue integral with respect to the probability measure p.

$$E[X] = \int_{\Omega} X dp. \tag{1}$$

which has been formalized in HOL, in the discrete case, as follows.

Theorem 1.

```
⊢ ∀X p. (real_random_variable X p) ∧ FINITE (IMAGE X (p_space p))
  ⇒ (expectation p X =
      ∑_(IMAGE X (p_space p)) (λr. r × Normal (distribution p X {r}))).
```

where `(IMAGE X (p_space p))` designates the values of the random variable X over the sample space `(p_space p)`. In the discrete case, this list has to be finite, i.e., `(FINITE (IMAGE X (p_space p)))`. The HOL function `Normal` allows the conversion of the real-valued `distribution` to its corresponding extended real.

3.2 The *k*-set Randomized Scheduling Algorithm

Consider a WSN that is formed by randomly deploying a set Sn of n sensor nodes over a field of interest of size a. Every sensor can only sense the surrounding environment and detect events within its circular sensing area of size r. We suppose that the nodes are uniformly and independently deployed. During the setup stage, the k-set randomized scheduling is run in parallel on every node as follows [19]. Each node starts by randomly picking a number, denoted by i, ranging from 0 to $(k-1)$, where k is the number of subsets or partitions. A node s_j is thus assigned to the i^{th} sub-network, designated by S_i, and will activate itself only during the scheduling round of that subset. At the end, k disjoint sub-networks are created to work alternatively.

Fig. 1 shows a small WSN of eight sensor nodes, which is randomly portioned into two sub-networks; S_0 and S_1. Each node randomly chooses a number 0 or 1 in order to be assigned to one of these two sub-networks. Suppose that nodes 0; 2; 5, randomly choose the number 0 and thus join the subset S_0, whereas nodes 1; 3; 4; 6; 7, select the number 1 and will be in the subset S_1. These two sub-networks will work by rounds, i.e., once the nodes 1; 3; 4; 6; 7, illustrated by the dashed circles, will be active, the remaining nodes 0; 2; 5, will be at the sleep state, and vice-versa.

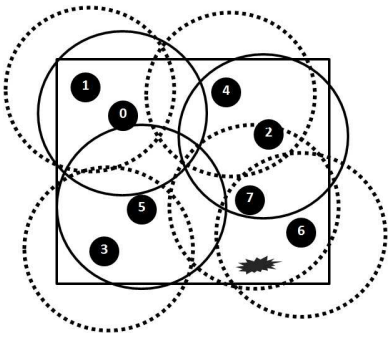

Fig. 1. The k-set randomized scheduling for ($n = 8$) nodes and ($k = 2$) subsets.

4 The Optimal Lifetime Problem under QoS Constraints

In the context of a WSN using the randomized scheduling, the network lifetime is "the elapsed time during which the network functions well" [38, 39]. The network lifetime, denoted by T_{Nlife}, has been mathematically defined as follows [38, 39].

$$T_{Nlife} = k \times T_{Slife} \tag{2}$$

where k is the number of subsets and T_{Slife} is the average lifetime of a sensor.

In [6, 9], we developed the higher-order-logic formalizations of the k-set randomized scheduling and three of its performance aspects within the sound core of the HOL theorem prover. The relevant metrics of interest are the network coverage, the detection probability and the detection delay, denoted as C_n, P_d, and D, respectively. In particular, we formally analyzed the minimum number of nodes to deploy in order to ensure a network coverage intensity C_n of at least t, denoted here as C_{nreq}, for a given number of sub-networks k [6].

$$n \geq \left\lceil \frac{\ln(1 - C_{nreq})}{\ln\left(1 - \frac{q}{k}\right)} \right\rceil. \tag{3}$$

where n is the total number of nodes, k; the number of subsets and q designates the probability that a given event is covered by at least one sensor.
While a coverage of C_{nreq} is achieved, the other detection metrics, are not guaranteed. Hence, deploying this lower bound n_{min} nodes may lead to worst values for the detection metrics, which is not desired.

Since the main goal of the k-set randomized scheduling is extending the network lifetime [19, 21], most related performance metrics should have appropriate values. These appropriate values, designated as Quality of Service (QoS) constraints, mainly depend on the application requirements, and are set according to some pre-defined values.

The lifetime problem [38, 39] initially consists in maximizing the network lifetime T_{Nlife} while minimizing the delay D, maximizing the detection probability P_d and the network coverage intensity C_n.

$$\begin{cases} 1.\ D \leq QoS_{DD} \\ 2.\ P_d \geq QoS_{DP} \\ 3.\ C_n \geq QoS_{C_n} \\ 4.\ n = c. \end{cases} \tag{4}$$

where QoS_{DD}, QoS_{DP}, and QoS_{C_n} are predefined QoS constraints associated to the detection delay D, the detection probability P_d, and the network coverage C_n, respectively, and c is a constant value.

According to Equation (2), maximizing the network lifetime T_{Nlife} is to maximize the number of subsets k. Nevertheless, the detection delay D will intuitively increases when k is growing, which is not suitable for WSN applications. There is thus an upper bound on the k-values so that a good coverage C_n can be ensured with acceptable delay D and detection probability P_d. Consequently, the main issue rather consists in optimizing the network lifetime to find the set of k-values that satisfy the main QoS constraints.

5 Application: Border Security Monitoring

Continuous surveillance along country borders is usually a high-priority concern, especially given the critical terrorism world context. Deployed along the borders, smart sensors can thus stop intruding objects including illegal immigrants, terrorists, and forces or vehicles in a military context [17]. Due to the safety-critical feature of the target application, sensors should have a smart behavior regarding the power availability while satisfying the main QoS requirements. Deployed WSNs for border monitoring usually suffer from limited lifetime [1], e.g, a REMBASS sensor can be operational for 30 days only [17]. Thus, the k-set randomized scheduling algorithm has been proposed for use to save energy for a border monitoring application [40].

In [9], we presented our higher-order-logic formalizations of the detection performances for randomly-scheduled WSNs. The practical effectiveness of these developments, have been then illustrated, through analyzing a WSN for border surveillance [40, 32]. In this paper, we focus on formally analyzing the optimal lifetime problem, presented in Section 4, for the same WSN-based application for border security monitoring. Hence, the nodes have a sensing range of $30m$, and are deployed into an area of size $a = 10000m^2$, whereas, the success probability q of a sensor covering a point, is $q = 0.28$. In the context of this application, the detection probability should be very high ($P_d > QoS_{DP} = 0.95$), whereas the detection latency as the shortest possible ($D < QoS_{DD} = 15s$) [1]. The QoS value for the network coverage intensity C_n, is not given in the reference paper, and is thus kept as generic for the considered application.

According to the definition of the network lifetime, given in Equation (2), optimizing T_{Nlife} basically depends on optimizing the corresponding k-values. An optimal solution exists, if there exist values of k satisfying the three first conditions of the problem, presented in Equation (4), for a given number of nodes ($n = c$) [38, 39].

In Theorem 2, we formally verify the main condition so that the lifetime problem, has an optimal solution [38, 39].

Theorem 2.

$$\begin{aligned} S_a = &\{k \mid D \leq QoS_{DD} < \frac{(Q-1+s)(Q^2-1+s)}{2Q(Q+1)}\left[1-(1-q)^n\right], \\ &1-(1-q)^n \geq P_d \geq QoS_{DP} > 0, \\ &1 \leq k \leq \frac{q}{(1-(1-QoS_{C_n})^{\frac{1}{n}})}, 0 < QoS_{C_n} < 1, n = c\} \end{aligned} \tag{5}$$

is bounded and non-empty.

where L is the duration of an occurring event, T is the length of a scheduling cycle, $Q = \lceil \frac{L}{T} \rceil$, and s is the remainder of the intrusion period L in terms of the number of slots T. The parameter $s = \frac{L}{T} + 1 - \lceil \frac{L}{T} \rceil$.

Proof. Each condition of the problem (Equation (4)) produces a set of k-values, which has to be proved as bounded and non-empty. The term bounded, used here, basically means "bounded above". Unfortunately, the reference textbooks [38, 39] provide a very abstract proof deducing that the big set S_a is bounded and non-empty. Larger investigations from the mathematical view as well as the WSN one, has been necessary to be able to understand the whole reasoning and switch it into the HOL theorem prover.

It is worth mentioning that, for space constraints, we will only involve the main mathematical assumptions related to the used variables. The interest reader can refer to [6, 9] for further details.

5.1 The Detection Delay

The optimization problem (Equation (4)) generates the following set of k-values for the detection delay.

$$S_D = \{k \mid D \leq QoS_{DD} < \frac{(Q-1+s)(Q^2-1+s)}{2Q(Q+1)}\left[1-(1-q)^n\right], n = c\} \tag{6}$$

To prove that the set S_D is bounded on k, the first intuitive way is to look for these concrete bounds. However, given the complexity of the delay expression [9], such bounds are seemingly very hard to obtain. Through a deeper analysis, we find out that the main proof depends on two main results. Indeed, if we can find the limit of the set sequence (Here $D(k)$) versus the parameter k, then we can get that this set is finite (Theorem 3). The second result states that if the set is finite then it is obviously bounded (Theorem 4).

Theorem 3 (*Finite set upon a limit*). *If a given sequence $U_n \to a$, then $\forall \varepsilon > 0$, there are only finitely many n for which $\mid U_n - a \mid \geq \varepsilon$.*

```
⊢ ∀U (ε:real) (a:real). (0 ≤ ε) ∧ (U → a)
   ⇒ FINITE {(n : num) : ε ≤ | U(n) − a |}.
```

Proof. Consider $\varepsilon > 0$, and the set $A_\varepsilon = \{n \in \mathbb{N} : \mid U_n - a \mid \geq \varepsilon\}$. Using the definition of the limit for the real sequence U_n, we have: $\forall \varepsilon > 0$, there exists N such that $\forall n.\, n \geq N$, we have $\mid U_n - a \mid < \varepsilon$. The set of n for which $\mid U_n - a \mid \geq \varepsilon$ will be contained in the set $\{1, 2, ..., N\}$, and hence finite.

Theorem 4 (*Upper bound of a finite integer set*). *Every finite set of integer s is bounded.*

$\vdash$ ∀(s:num->bool). FINITE s ⇒ BOUNDED s.

where the HOL function BOUNDED specifies a bounded set of integers.

Lemma 1 (*The set S_D is bounded*).

$\vdash$ ∀n k q s L Ts QoSDD. (0 < s < 1) ∧ (0 < L) ∧ (0 < Ts)
⇒ (BOUNDED {k | DD p D n k q ≤ QoSDD }).

Proof. We require the limiting value of the detection delay D versus k (Lemma 2), as well as the asymptotic behavior of the delay D on k (Lemma 3). Then, considering Theorem 3 for the sequence $D(k)$, with the right value of ε, we can get that the set S_D is bounded. Indeed, since $D(k)$ is increasing (Lemma 3), the maximum possible values is $\lim_{k\to\infty} D$, which is given in Lemma 2. We thus get $QoS_{DD} < lim_{k\to\infty} D$. Plugging in Theorem 3 with $\varepsilon = (\lim_{k\to\infty} D) - QoS_{DD} = \frac{(Q-1+s)(Q^2-1+s)}{2Q(Q+1)}\left[1-(1-q)^n\right] - QoS_{DD}$, we can obtain that the set S_D is finite. Finally, based on Theorem 4, we deduce that S_D is bounded.

Lemma 2 (*Limit of the detection delay when k is very large*).

$\vdash$ ∀n q s L Ts. (1 ≤ n) ∧ (0 < s < 1) ∧ (0 < L) ∧ (0 < Ts) ∧
(0 < q < 1)
⇒ ($\lim_{k\to\infty}$ DD = $\frac{(Q-1+s)(Q^2-1+s)}{2Q(Q+1)}\left[1-(1-q)^n\right]$).

where $Q = \left\lceil \frac{L}{T} \right\rceil$.
Proof. We verified Lemma 2 using an alternate proof since the original proof, based on the Mean Value Theorem (MVT), was not possible in HOL. Indeed, while the MVT theorem in HOL is available for constant real bounds, these bounds are considered as variables in the paper-and-pencil proof [39].

Lemma 3 (*The detection delay is increasing as k increases*).

$\vdash$ ∀n q s L Ts. (1 ≤ n) ∧ (0 < s < 1) ∧ (0 < L) ∧ (0 < Ts) ∧
(0 < q < 1)
⇒ (mono_incr (λk. real (DD p D n k q))).

where the HOL function mono_incr denotes an increasing natural sequence.
Proof. The proof of the above lemma is based on the derivative of the corresponding real functions. The reasoning thus involved a large amount of real analysis with very complicated mathematical expressions including summations and using various properties of sequences and series of real numbers. It is important to note that the original proof of the above lemma in [39] was missing a whole fraction term, which is fortunately positive and thus does not finally affect the validity of the function monotonicity.

We conclude that S_D is non-empty, using the monotonicity of $D(k)$ on k (Lemma 2), along with some reasoning on the quality of service constraints. Indeed, $D(k)$, increasing versus k, means that the minimum delay value, is induced

for $(k = 1)$, i.e, $D(1)$. The values of $D(k)$; including QoS_{DD}, cannot go below $D(1)$. Hence, we always have $D(k) > D(1)$, which gives $QoS_{DD} > D(1)$. This ensures that $(k = 1) \in S_D$, and hence S_D is non-empty.

5.2 The Detection Probability

Based on the lifetime problem (Equation 4), we have:

$$S_{Pd} = \{k \mid P_{d|k=1} = (1 - (1 - q)^c) \geq P_d \geq QoS_{DP} > 0, n = c\} \tag{7}$$

which is required to be verified as bounded and non-empty.

Lemma 4 (*The set S_{Pd} is bounded*).

$\vdash \forall$ q n s L Ts QoSDP. (1 $\leq$ n) $\wedge$ (0 $<$ s $<$ 1) $\wedge$ (0 $<$ L) $\wedge$
(0 $<$ Ts) $\wedge$ (0 $<$ q $<$ 1) $\wedge$ ($\forall$k. L $<$ k$\times$Ts) $\wedge$ (0 $<$ QoSDP $<$ 1)
$\Rightarrow$ BOUNDED $\{$k $\mid$ QoSDP $\leq$ Pd p n k s L Ts q$\}$.

Proof. We first achieve the proof that S_{Pd} is finite using Theorem 3 such that $A = 0$ and $\varepsilon = QoS_{DP}$ which is > 0. For that, the behavior of the detection probability P_d regarding the parameter k is required (Lemmas 5 and 6). We finally establish that the set S_{Pd} is bounded using Theorem 4 together with the latter result.

Lemma 5 (*Limit of the detection probability as k is infinite*).

$\vdash \forall$ q n s L Ts. (1 $\leq$ n) $\wedge$ (0 $<$ s $<$ 1) $\wedge$ (0 $<$ L) $\wedge$ (0 $<$ Ts) $\wedge$
(0 $<$ q $<$ 1) $\wedge$ ($\forall$k. L $<$ k$\times$Ts)
$\Rightarrow \lim_{\text{k}\to+\infty}$ (λk. Pd p n k s L Ts q) = 0.

Lemma 6 (*The detection probability is decreasing versus k*).

$\vdash \forall$ q n s L Ts. (1 $\leq$ n) $\wedge$ (0 $<$ s $<$ 1) $\wedge$ (0 $<$ L) $\wedge$ (0 $<$ Ts) $\wedge$
(0 $<$ q $<$ 1) $\wedge$ ($\forall$k. L $<$ k$\times$Ts)
$\Rightarrow$ (mono_decr (λk. Pd p n k s L Ts q)).

Since the detection probability is decreasing with k (Lemma 6), the best detection probability value is ensured for $(k = 1)$. So, we have $P_d(1) > P_d(k)$. The QoS_{DP} values cannot go above $P_d(1)$, i.e, $Pd(1) > QoS_{DP}$. Hence, $(k = 1) \in S_{Pd}$, which guarantees that the set S_{Pd} is non-empty.

5.3 The Network Coverage

Unlike the detection metrics, the upper bound of the k-values for the coverage set; S_{Cn}, can be obtained through some mathematical operations.

$$S_{Cn} = \{k \mid 1 \leq k \leq \frac{q}{(1 - (1 - QoS_{C_n})^{\frac{1}{n}})}, n = c\} \tag{8}$$

Theorem 5 (*The set S_{Cn} is bounded*).

```
⊢ ∀p q n s QoSCn. (1 ≤ n) ∧ (0 < q < 1) ∧ (0 < QoSCn < 1)
⇒ BOUNDED {k | QoSCn ≤ Cn p X k s C n q}.
```

Proof. The proof is mainly based on Theorem 4, together with some real analysis about the floor function and subsets.

The set S_{Cn} can be simply deduced as non-empty. Similarly, as the network coverage is decreasing versus the parameter k [7], the best coverage is then achieved for $(k = 1)$. We hence target a good QoS value for coverage, but which can not exceed $C_n(1)$.

Finally, we can deduce that the big set with the generic QoS values;

$$S_a = S_D \cap S_{Pd} \cap S_{Cn}$$

is bounded and non-empty, using the above reasoning on the three sets S_D, S_{Pd} and S_{Cn}, i.e, Theorems 1, 4, and 5, respectively, together with the fact that $(k = 1)$ is shown to be in each of the three sets, and hence in their intersection.

Based on that, we can easily establish that, for our border security monitoring application, we have:

$$\begin{aligned} S_{app} =& \{k \mid D \leq (QoS_{DD} = 15) < \frac{(Q-1+s)(Q^2-1+s)}{2Q(Q+1)}\left[1-(0.72)^n\right], \\ & 1-(0.72)^n \geq P_d \geq (QoS_{DP} = 0.95) > 0, \\ & 1 \leq k \leq \frac{0.28}{(1-(1-QoS_{C_n})^{\frac{1}{n}})}, 0 < QoS_{C_n} < 1, n = c\} \end{aligned} \tag{9}$$

is bounded and non-empty.

In this section, we formally illustrate the analysis of the optimal lifetime problem, given in Equation (2), for a border security monitoring WSN application [1] such that $(QoS_{DP} = 0.95)$ and $(QoS_{DD} = 15s)$. It is worth to mention that the formal developments of lifetime can be quite valuable to analyze any randomly-scheduled application like a general surveillance framework for WSN.

6 Discussion

In this paper, we have been able to formally analyze, within the HOL theorem prover, the optimal lifetime problem (Equation 4) under Quality of Service (QoS) constraints, for wireless sensor networks using the k-set randomized scheduling. These QoS constraints are associated with the key performance metrics, i.e., the network coverage, the detection probability and the detection delay. More particularly, there are two main conditions on the k-values, under which the optimal lifetime solution exists for such problems. These conditions require that the big set S_a of k-values, shown in Equation (5), is non-empty and bounded. For that, we built upon our higher-order-logic foundations, developed in [6, 8, 9], to verify this minimal set of conditions, and illustrate this analysis through a border security monitoring application with concrete QoS values for the detection probability and the detection delay.

The current lifetime analysis, presented in this paper, primarily illustrates the great value of the existing higher-order-logic developments for the other performance metrics. Indeed, the lifetime verification has been possible thanks to the sound and complete formalizations of the network coverage, done in [6–8], together with the detection probability and delay, presented in [9]. The successful verification of the lifetime optimization problem thus clearly highlights the main advantages of our theoretical developments of the coverage and detection attributes in terms of precision and coherence. Hence, it would not have been possible to effectively achieve the main lifetime proof if, for example, there was a missing assumption on one of the design parameters in the detection part.

While the main goal of the previous formalizations on coverage and detection [6, 9] was to formally verify the expressions associated with the probabilistic attributes of interest, the lifetime problem is considered in a completely different way. Indeed, the lifetime definition of a randomly-partitioned wireless network, as specified in the paper-and-pencil probabilistic models [38, 39], is very simple (Definition 2) and does not require any investigation from the formalization side. However , it was found to be quite interesting to tackle the formal analysis of the lifetime optimization problem (Equation 5) under quality of service constraints. Clearly, the higher-order-logic formalization process for the network lifetime is quite different from the three other performance metrics, where the main idea was to formally analyze the conditions under which the optimal network lifetime exists, rather than verify the lifetime in itself.

Comparably to the other performance aspects, many difficulties have been implied in the lifetime verification. Although the lifetime proof seems simple, there were many hidden steps making the understanding of the main proof quite challenging. Hence, except for the coverage set where the concrete bounds on k were simple to get, the other sets on the delay D and the detection probability P_d have been directly deduced to be non-empty and bounded. These deductions, based on some missing steps in the corresponding paper-based proof [38, 39], involved significant mathematical investigations. No indication was given about which mathematical result is applied. Nevertheless, it is very common that some details which seem obvious for mathematicians turn out to be very hard to follow from the reader's side.

Secondly, the high degree of interactivity required within a theorem prover in general and in HOL, in particular, was also a huge obstacle for a quick formalization. Hence, tedious mathematical efforts may be needed to prove a basic result or just to correctly handle complicated summations. For instance, the proof of Lemma 3, which occupied about half a page in the original textbook [39], took about 12 pages of HOL code. For the same lemma, we discovered that a whole fraction term was missing in the original mathematical analysis [39]. This discrepancy would have had a crucial impact on the final result if the term was of opposite sign. On the other hand, it is clear that it would not have been possible to catch this error based on a manual inspection unless the proof is redone step by step. Such interesting finding clearly highlights the main strength of formal methods guaranteeing accurate and complete results.

7 Conclusions

In this paper, we presented a reliable approach for the formal analysis of the the network lifetime for randomly-scheduled WSNs. Hence, based on our earlier work [6, 9], we provided the higher-order-logic formalizations of the lifetime maximization problem [39], under Quality of Service (QoS) constraints related to the network coverage and the detection performances. These formalizations enable us to formally verify the network lifetime related characteristics of a border security monitoring application using the k-set randomized scheduling.

Compared with the existing approaches such as traditional paper-and-pencil probabilistic modelling, simulation and probabilistic model checking, our theorem-proving based approach allows a generic formal verification of randomly-scheduled WSNs regardless of the values of the design parameters. Besides, due to the sound support of probability theory available in the HOL theorem prover, our approach enables much more reliable validation of the probabilistic performance attributes of interest including statistical quantities. Finally, unlike most of the previous work focusing on the validation of the functional aspects of WSNs, our work is distinguishable by addressing the performance aspects.

As future work, the formalization of the optimal detection probability [28], can be also investigated in the same way of the network lifetime, achieved in this paper. The whole proposed approach, described in [8], can be also generalized to tackle the formal analysis of a variant of the k-set randomized scheduling [18].

References

1. Arora, A.: A Line in the Sand: a Wireless Sensor Network for Target Detection, Classification, and Tracking. Computer Networks 46(5), 605–634 (2004)
2. Baier, C., Katoen, J.P.: Principles of Model Checking. The MIT Press (2008)
3. Ballarini, P., Miller, A.: Model Checking Medium Access Control for Sensor Networks. In: Proceedings of the Symposium on Leveraging Applications of Formal Methods, Verification and Validation. pp. 255–262. IEEE Computer Society (2006)
4. Bernardeschi, C., Masci, P., Pfeifer, H.: Analysis of Wireless Sensor Network Protocols in Dynamic Scenarios. In: Stabilization, Safety, and Security of Distributed Systems, Lecture Notes in Computer Science, vol. 5873, pp. 105–119. Springer (2009)
5. Chen, D., Varshney, P.K.: QoS Support in Wireless Sensor Networks: A Survey. In: Proceedings of the International Conference on Wireless Networks. pp. 227–233. CSREA Press (2004)
6. Elleuch, M., Hasan, O., Tahar, S., Abid, M.: Formal Analysis of a Scheduling Algorithm for Wireless Sensor Networks. In: Formal Methods and Software Engineering, Lecture Notes in Computer Science, vol. 6991, pp. 388–403. Springer (2011)
7. Elleuch, M., Hasan, O., Tahar, S., Abid, M.: Formal Probabilistic Analysis of a Wireless Sensor Network for Forest Fire Detection. In: Symbolic Computation in Software Science, Electronic Proceedings in Theoretical Computer Science, vol. 122, pp. 1–9. Open Publishing Association (2013)
8. Elleuch, M., Hasan, O., Tahar, S., Abid, M.: Towards the Formal Performance Analysis of Wireless Sensor Networks. In: Proceedings of the Workshop on Enabling

Technologies: Infrastructure for Collaborative Enterprises. IEEE Computer Society (2013)
9. Elleuch, M., Hasan, O., Tahar, S., Abid, M.: Formal Probabilistic Analysis of Detection Properties in Wireless Sensor Networks. Formal Aspects of Computing 27(1), 79–102 (2015)
10. Fanourgakis, E.: Modelling and Verification of QoS properties of a Biomedical Wireless Sensor Network. Project Work, University of Hamburg-Harbug (2012)
11. Fehnker, A., Hoesel, L.V., Mader, A.: Modelling and Verification of the LMAC Protocol for Wireless Sensor Networks. In: Integrated Formal Methods, Lecture Notes in Computer Science, vol. 4591, pp. 253–272. Springer (2007)
12. Fruth, M.: Probabilistic Model Checking of Contention Resolution in the IEEE 802.15.4 Low-rate Wireless Personal Area Network Protocol. In: Proceedings of the 2nd symposium on Leveraging Applications of Formal Methods, Verification and Validation. pp. 290–297. IEEE Computer Society (2006)
13. Gordon, M., Melham, T.: Introduction to HOL: A Theorem Proving Environment for Higher-order Logic. Cambridge Univ. Press (1993)
14. Gupta, A.: Formal Hardware Verification Methods: a Survey. Formal Methods in System Design 1(2-3), 151–238 (1992)
15. Hanna, Y., Rajan, H., Zhang, W.: Slede: a Domain-specific Verification Framework for Sensor Network Security Protocol Implementations. In: Proceedings of the Conference on Wireless Network Security. pp. 109–118. ACM (2008)
16. Heidarian, F., Schmaltz, J., Vaandrager, F.: Analysis of a Clock Synchronization Protocol for Wireless Sensor Networks. Theoretical Computer Sciences 413(1), 87–105 (2012)
17. Hewish, M.: Reformatting Fighter Tactics. Jane's International Defense Review (2001)
18. Hsin, C., Liu, M.: Network coverage using low duty-cycled sensors: Random & coordinated sleep algorithms. In: Proceedings of the Symposium on Information Processing in Sensor Networks. pp. 433–442 (2004)
19. Liu, C.: Randomized Scheduling Algorithm for Wireless Sensor Networks. In Project Report of Randomized Algorithm, University of Victoria, B.C., Canada (2004)
20. Liu, C., Wu, K., King, V.: Randomized Coverage-preserving Scheduling Schemes for Wireless Sensor Networks. In: Networking Technologies, Services, and Protocols; Performance of Computer and Communication Networks; Mobile and Wireless Communication Systems, Lecture Notes in Computer Science, vol. 3462, pp. 956–967. Springer (2005)
21. Liu, C., Wu, K., Xiao, Y., Sun, B.: Random Coverage with Guaranteed Connectivity: Joint Scheduling for Wireless Sensor Networks. IEEE Transactions on Parallel and Distributed Systems 17(6), 562–575 (2006)
22. Liu, S., Ölveczky, P., Meseguer, J.: Formal Analysis of Leader Election in MANETs using Real-Time Maude. In: Software, Services, and Systems, Lecture Notes in Computer Science, vol. 8950, pp. 231–252. Springer (2015)
23. Liu, Y., Gu, Y., Chen, G., Ji, Y., Li, J.: A Novel Accurate Forest Fire Detection System Using Wireless Sensor Networks. In: Proceedings of the Conference on Mobile Ad-hoc and Sensor Networks. pp. 52–59. IEEE Computer Society (2011)
24. MacKay, D.: Introduction to Monte Carlo Methods. In: Proceedings of NATO Advanced Study Institute on Learning in Graphical Models. pp. 175–204. Kluwer Academic Publishers (1998)
25. Mamun, Q.: A Coverage-Based Scheduling Algorithm for WSNs. International Journal of Wireless Information Networks 21(1), 48–57 (2014)

26. Mhamdi, T.: Information-Theoretic Analysis using Theorem Proving. Ph.D. thesis, Concordia Univ., Montreal, QC, Canada (December 2012)
27. Mhamdi, T., Hasan, O., Tahar, S.: Formalization of Entropy Measures in HOL. In: Interactive Theorem Proving, Lecture Notes in Computer Science, vol. 6898, pp. 233–248. Springer (2011)
28. Olteanu, A., Xiao, Y., Wu, K., Du, X.: Weaving a Proper net to Catch Large Objects in Wireless Sensor Networks. IEEE Transactions on Wireless Communications 9(4), 1360–1369 (2010)
29. Ölveczky, P., Thorvaldsen, S.: Formal Modeling and Analysis of the OGDC Wireless Sensor Network Algorithm in Real-time Maude. In: Formal Methods for Open Object-based Distributed Systems, Lecture Notes in Computer Science, vol. 4468, pp. 122–140. Springer (2007)
30. The Real-Time tool (2013), http://heim.ifi.uio.no/peterol/RealTimeMaude/
31. Rutten, J., Kwaiatkowska, M., Normal, G., Parker, D.: Mathematical Techniques for Analyzing Concurrent and Probabilisitc Systems. CRM Monograph Series, American Mathematical Society (2004)
32. Sun, Z., Wang, P., Vuran, M., Al-Rodhaan, A., Al-Dhelaan, A., Akyildiz, I.: BorderSense: Border Patrol through Advanced Wireless Sensor Networks. Ad Hoc Networks 9(3), 468–477 (2011)
33. Tian, D., Georganas, N.: A Coverage-preserving Node Scheduling Scheme for Large Wireless Sensor Networks. In: Proceedings of the International Workshop on Wireless Sensor Networks and Applications. pp. 32–41. ACM (2002)
34. Tschirner, S., Xuedong, L., Yi, W.: Model-based Validation of QoS Properties of Biomedical Sensor Networks. In: Proceedings of the International Conference on Embedded Software. pp. 69–78. ACM (2008)
35. Wang, L., Xiao, Y.: A Survey of Energy-efficient Scheduling Mechanisms in Sensor Networks. Mobile Networks and Applications 11(5), 723–740 (2006)
36. Wu, K., Gao, Y., Li, F., Xiao, Y.: Lightweight Deployment-Aware Scheduling for Wireless Sensor Networks. Mobile Networks and Applications 10(6), 837–852 (2005)
37. Xia, F.: QoS Challenges and Opportunities in Wireless Sensor/Actuator Networks. Sensors 8(2), 1099–1110 (2008)
38. Xiao, Y., Chen, H., Wu, K., Liu, C., Sun, B.: Maximizing Network Lifetime under QoS Constraints in Wireless Sensor Networks. In: Proceeding of the Global Telecommunications Conference. pp. 1–5. IEEE Computer Society (2006)
39. Xiao, Y., Chen, H., Wu, K., Sun, B., Zhang, Y., Sun, X., Liu, C.: Coverage and Detection of a Randomized Scheduling Algorithm in Wireless Sensor Networks. IEEE Transactions on Computers 59(4), 507–521 (2010)
40. Xiao, Y., Zhang, Y., Peng, M., Chen, H., Du, X., Sun, B., Wu, K.: Two and Three-dimensional Intrusion Object Detection under Randomized Scheduling Algorithms in Sensor Networks. Computer Networks 53(14), 2458–2475 (2009)
41. Yick, J., Mukherjee, B., Ghosal, D.: Wireless Sensor Network Survey. Computer Networks 52(12), 2292–2330 (2008)
42. Zayani, H., Barkaoui, K., Ayed, R.B.: Probabilistic Verification and Evaluation of Backoff Procedure of the WSN ECo-MAC Protocol. International Journal of Wireless & Mobile Networks 12(1), 156–170 (2010)
43. Zheng, M., Sun, J., Liu, Y., Dong, J., Gu, Y.: Towards a Model Checker for NesC and Wireless Sensor Networks. In: Formal Methods and Software Engineering, Lecture Notes in Computer Science, vol. 6991, pp. 372–387. Springer (2011)