Formal Probabilistic Performance Verification of Randomly-scheduled Wireless Sensor Networks

Maissa Elleuch*

CES Laboratory, National School of Engineers of Sfax, Sfax University
Soukra Street, 3052 Sfax, Tunisia
and
Digital Research Center of Sfax, Technopark of Sfax, Sfax, Tunisia
E-mail: maissa.elleuch@crns.rnrt.tn
*Corresponding author

Osman Hasan and Sofiène Tahar

Dept. of Electrical & Computer Engineering, Concordia University
1455 de Maisonneuve W., Montreal, Quebec, H3G 1M8, Canada
E-mail: o_hasan@ece.concordia.ca
E-mail: tahar@ece.concordia.ca

Mohamed Abid

CES Laboratory, National School of Engineers of Sfax, Sfax University
Soukra Street, 3052 Sfax, Tunisia
E-mail: mohamed.abid@enis.rnu.tn

Abstract:

Energy efficiency in Wireless Sensor Networks (WSN) is one of the most critical issue regardless of the target application. While scheduling sensors by partitions to preserve energy is a simple and intuitive approach in this context, it is also important 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 met. Besides, various assumptions, may effectively impact the sensing performance capabilities of the network. Nevertheless, most analysis techniques focus on the average performance values, and do not consider neither the targeted QoS requirements, nor the probabilistic feature of the algorithm. Based on the theorem proving approach, we first provide, in this paper, an accurate formal analysis of the network lifetime maximization problem, under QoS constraints, for randomly-scheduled wireless sensor networks. After that, we tackle the higher-order-logic formalization of the intrusion coverage intensity, for a modified version of the randomized scheduling, with more realistic assumptions for the intrusion object, in a two or three dimensional plane.

Keywords: Theorem proving ; Wireless sensor networks ; Node scheduling ; Performance analysis ; Network lifetime ; Intrusion coverage.

Reference to this paper should be made as follows: Elleuch, M., Hasan, O., Tahar, S. and Abid, M. (2018) 'Formal Probabilistic Performance Verification

of Randomly-scheduled Wireless Sensor Networks', *International Journal of Critical Computer-Based Systems*, Vol. x, No. x, pp.xxx–xxx.

Biographical notes: Maissa Elleuch is currently an assistant professor in Computer Sciences at the Digital Research Center of Sfax (CRNS). She received her Ph.D. degree in Computer Systems Engineering from the National School of Engineers of Sfax, in 2015. Before that, She received her Engineer degree in Computer Sciences in 2006 from the National Engineering School in Computer Sciences (ENSI). She is also a member of the Computer and Embedded Systems (CES) Laboratory at the National School of Engineers of Sfax (ENIS). Her research interests include formal methods, system verification and wireless sensor networks.

Osman Hasan received the B.Eng. (Hons.) degree from the N-W.F.P University of Engineering and Technology, Pakistan, in 1997, and the M.Eng. and Ph.D. degrees from Concordia University, Canada, in 2001 and 2008, respectively. He served as an ASIC design engineer from 2001 to 2003 in the industry prior to joining Concordia University in 2004 for his Ph.D. degree. Currently, he is an assistant professor at the National University of Science and Technology, Pakistan and affiliate assistant professor at Concordia University, Canada. His current research interests include formal methods, higher-order-logic theorem proving and probabilistic analysis.

Sofiène Tahar received the Diploma degree in computer engineering from the University of Darmstadt, Germany, in 1990, and the Ph.D. degree with distinction in computer science from the University of Karlsruhe, Germany, in 1994. Currently, he is a professor and research chair in formal verification of system-on-chip at the Department of Electrical and Computer Engineering, Concordia University, Canada. His research interests are in the areas of formal hardware verification, system-on-chip verification, analog and mixed signal circuits verification, and probabilistic, statistical and reliability analysis of systems.

Mohamed Abid Head of "Computer Embedded System" laboratory CES-ENIS, Tunisia. Mohamed ABID is working now as a professor at the National Engineering School of Sfax (ENIS), University of Sfax, Tunisia. He received the Ph.D. degree from the National Institute of Applied Sciences, Toulouse (France) in 1989. His current research interests include: hardware-software co-design, System on Chip, Reconfigurable System, and Embedded System, etc. Dr. Abid has served also as Guest professor at several international universities and as a consultant to research and development in Telnet Incorporation.

This paper is a revised and expanded version of a paper entitled 'Formal Probabilistic Analysis of Lifetime for a WSN-based Monitoring Application' presented at 10th International Workshop on Verification and Evaluation of Computer and Communication Systems, Tunis, Tunisia, October 6-7, 2016.

1 Introduction

Monitoring an environment continuously is a very challenging task but Wireless Sensor Networks (WSN) (Yick et al., 2008) have emerged as a key enabler technology for various applications. The sensor devices can accurately sense their environment, and then forward the collected measures to a base station. Some common applications include disaster

detection, habitat and traffic monitoring, agriculture and marine monitoring (Xu et al., 2014), and weather forecasting.

Regardless of the target application, energy efficiency arises as an invariable issue, especially given the major difficulties for replacing or recharging the sensor batteries in harsh environmental conditions. For example, in a WSN deployed for forest fire detection, sensors should not only ensure the appropriate monitoring of the whole area, but also be functional for a long period. In this context, a very intuitive way to preserve energy is to activate sensors by partitions, and thus extend the network lifetime (Wang and Xiao, 2006). This is a very attractive idea, based on the observation that most of the intrusion events, like a fire outbreak, occur occasionally. Therefore, considering a smaller number of sensors active at any given time, the lifetime of the overall system increases, at the cost of lower performance. The k-set randomized scheduling (Liu et al., 2006) is a node scheduling approach, which mainly consists in randomly organizing the set of sensors into k subsets. The formed subsets of nodes work alternatively within their allocated time slot, so that the overall network energy is preserved.

Randomly scheduling sensors for lifetime management purposes is a very efficient approach, but it is also important to carefully consider the performance of the application. Hence, safety-critical WSN applications have to usually meet various Quality of Service (QoS) requirements (Chen and Varshney, 2004; Xia, 2008). These QoS requirements can be regarded as "the capability of providing assurance that the service requirements of applications can be satisfied" (Chen and Varshney, 2004). For the same example of the forest application, sensing the fire outbreak on time and delivering the alarm within the shortest delay and with a high probability, is of utmost importance. Other assumptions, related to the intrusion object itself, can definitely affect the sensing capabilities of randomly-scheduled WSN, and thus the network performance. Indeed, as the intrusion object is larger, the more likely it will be detected. Hence, while the intrusion object should have a certain form, most previous studies (Liu et al., 2006; Xiao et al., 2010) just consider the intrusion as a point. In real WSN applications, an intrusion object can not be assimilated to a point and is much larger than that. For instance, in a WSN deployed for border security monitoring, intruders crossing country borders are very different and may have various sizes and shapes.

Besides the network lifetime, the coverage and the detection performances arise also as critical requirements. Nevertheless, for the k-set randomized scheduling, these performance metrics are completely probabilistic (Liu et al., 2006; Xiao et al., 2010). It may happen that some fire outbreak is not effectively covered, if no nodes are deployed around the fire because of the random node deployment, or if the surrounding nodes are inactive, due to randomized scheduling. While the probabilistic aspect poses real challenges on the analysis of WSN, missing some critical intrusions, can have devastating consequences.

Generally, the performance of the randomized scheduling is analyzed using probabilistic analytical models (Tian and Georganas, 2002; Hsin and Liu, 2004; Liu et al., 2005). Nevertheless, the complete correctness of analytical models is apparently hard to establish. The reliability of the theoretical built models is consolidated through simulation using the Monte Carlo method (MacKay, 1998). This simulation-based method usually consists on approximately answering a query on a probability distribution by repetitively analyzing a large number of samples. Due to the inherent incompleteness of this technique coupled with the rounding errors of computer arithmetics, the simulation approach, which is also used to validate the analytical results, may give inaccurate results.

Compared with traditional simulation, formal methods are less frequently used for the validation of WSN. Based on mathematical techniques, formal methods (Gupta, 1992)

rigorously analyze the theoretical model of the given system. These methods have been proposed as an efficient solution to validate a wide range of hardware and software systems increasing thus the analysis reliability. Recently, formal methods have gained a growing interest in the context of WSN to analyze their functional or quantitative correctness (Ölveczky and Thorvaldsen, 2007; Ballarini and Miller, 2006; Zayani et al., 2010), but most of the existing work is focused on the validation of functional aspects only. Nevertheless, reliable performance evaluation of WSN constitutes also an extremely challenging aspect.

In this paper, we first provide an accurate formal analysis of the network lifetime for randomly-scheduled WSN (Liu et al., 2006; Xiao et al., 2010). In particular, we are interested in the higher-order-logic formalizations of the lifetime maximization problem, given in (Xiao et al., 2010), under QoS constraints. The earlier work of (Elleuch et al., 2011, 2015, 2016a,b) described a formalization of the k-set randomized scheduling algorithm and its main performance properties based on the recent probability theory formalizations (Mhamdi et al., 2011) in the HOL theorem prover. We will build upon these theoretical developments to formally show that the optimal solution for the lifetime maximization problem exists. As proposed in (Liu et al., 2006; Xiao et al., 2010), the algorithm is abstracting intrusion objects by points, which is an over simplification that is not reflected in all cases of real WSN applications. In the second part of this paper, we tackle the formalization of the intrusion coverage intensity, for a modified version of the algorithm (Xiao et al., 2008a,b, 2009), considering object size and shape in a two or three dimensional plane.

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 formal analysis of the lifetime maximization problem under QoS requirements. In Section 5, the higher-order-logic formalization of the intrusion coverage under different object forms is provided. Section 6 is devoted to discussions.

2 Related Work

The analysis of the randomized scheduling has been usually done using the paper-and-pencil based probabilistic technique (Wu et al., 2005; Hsin and Liu, 2004; Liu et al., 2006; Xiao et al., 2009). Such analysis consists in constructing a theoretical model where the required random variables are specified and analyzed together with the associated performance metrics. For validation purposes, simulation, using the Monte Carlo method (MacKay, 1998), is finally carried out, to validate the proposed model. In (Wu et al., 2005), a randomized scheduling, is studied via analytical modeling, where many statistical quantities like the expectation of non-covered area, have been proved. The resilience of the same algorithm regarding clock asynchrony has been mathematically studied in (Liu et al., 2006). Mamun (Mamun, 2014) evaluated the coverage using a mathematical model while simulations have been run with specific network sizes and sensing ranges. Besides, previous studies on analyzing randomly-scheduled WSN usually ignore the potential form of the intrusion object, and abstract it using a point (Liu et al., 2006; Xiao et al., 2010; Hsin and Liu, 2004). Few studies address the size and the shape of the intrusion object. For example, the work of (Olteanu et al., 2010) analytically considers the impact of the size and the shape of the intrusion object to analyze the optimal detection probability. More recently, (Wang et al., 2017) has proposed a mathematical framework, called E-HIPA, for an accurate target localization in the spatial domain.

Model checking technique (Baier and Katoen, 2008) has been successfully explored for the validation of various aspects in the WSN context. In (Ölveczky and Thorvaldsen, 2007), 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 (RTMaude). Several other prominent works reported on the use of model checking for the analysis of WSN protocols include (Tschirner et al., 2008; Fehnker et al., 2007; Liu et al., 2015), or for the development of formal frameworks (Hanna et al., 2008; Zheng et al., 2011). 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 (Baier and Katoen, 2008). Hence, the analysis of the OGDC algorithm (Ölveczky and Thorvaldsen, 2007) has been limited for WSN with only 6 nodes within a monitored region of $15m \times 15m$. In addition, the work in (Zheng et al., 2011) has reported over 1 million generated states for the verification of a simple property. Some additional temporal abstractions have been applied to correctly achieve the analysis. On the other hand, none of the previous studies provided a sound modelling of the randomness aspect in WSN, which constitutes a real limitation since most of the WSN algorithms are probabilistic. In (Ölveczky and Thorvaldsen, 2007), a random function, assumed to be 'good', has been used to model the probabilistic behavior of interest. Moreover, the PMaude tool has been proposed to enhance the accuracy of the probabilistic analysis in (Ölveczky and Thorvaldsen, 2007; Liu et al., 2014).

To overcome these major problems, probabilistic model checking (Rutten et al., 2004) has also been used for the probabilistic functional analysis of wireless systems (Fehnker et al., 2007; Fruth, 2006; Zayani et al., 2010). Similar to the traditional technique, probabilistic model checking is based on an exhaustive exploration of the system model to check the validity of a given probabilistic property. The model checker PRISM (PRISM) has been applied quite frequently for the validation of Medium Access Control (MAC) protocols for WSN (Fehnker et al., 2007; Fruth, 2006; Zayani et al., 2010). 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 (Ballarini and Miller, 2006; Zayani et al., 2010). The obtained results can hardly be termed as exhaustive and thus formally verified.

On the other hand, very few works based on theorem proving (Gordon and Melham, 1993) exist in the open literature. As example, a synchronization protocol for WSN, has been analyzed using the Isabelle/HOL theorem prover (Heidarian et al., 2012). The work in (Bernardeschi et al., 2009) 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, is just instantiated by a given value throughout the analysis. The analysis results obtained can not be hence considered as reliable versus the probability modelling.

Unlike previous works, we provide rigorous formalizations of the network lifetime maximization problem (Xiao et al., 2010), under QoS constraints, for randomly-scheduled WSN, 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 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 (Tschirner et al., 2008), 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. The work in (Fanourgakis, 2012) 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, we first provide, in this paper, a formal analysis of the lifetime for randomly-scheduled WSN. Then, we develop the intrusion coverage formalizations of a novel version of the randomized algorithm (Xiao et al., 2009), under new assumptions related to the intrusion object.

3 Preliminaries

In this section, we first introduce the HOL theorem prover (HOL4), followed by its probabilistic foundations. Then, we briefly describe the k-set randomized scheduling algorithm.

3.1 HOL Theorem Prover

The HOL theorem prover (HOL4) is a proof assistant of higher-order logic. The verification approach of HOL consists in three main steps: describing the system to be verified in higher-order logic, formalizing the properties of interest as proof goals of higher-order-logic and finally verifying these goals as theorems within HOL. Furthermore, the HOL theorem prover includes a very rich library of theories. A theory can be defined as a set of pre-verified theorems for a given domain, function or operation. When needed, a HOL theory can be loaded and used, which greatly aids the verification process. Additionally, users may be assisted by automatic proof procedures (Gordon and Melham, 1993), which are a collection of steps in a single command. Despite the existence of all these theories and automatic procedures, most of the time, proofs in HOL are interactive and require the intervention of user. Various proof techniques, such as rewriting, simplification, specialization, generalization and mathematical induction, are available in HOL to aid the verification process. We summarize, in Table 1, some of the HOL symbols used in this paper and their corresponding mathematical signification (Gordon and Melham, 1993).

3.2 Probabilistic Analysis in HOL

In this work, we utilize the recently developed and most generic probability theory developed by Mhamdi (Mhamdi, 2012; Hasan and Tahar, 2015), 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 3.1).

HOL Symbol	**Standard**	**Meaning**
$\wedge$	and	Logical and
SUC n	$n+1$	Successor of a num
count n	$\{m \mid m < n\}$	Set of all m strictly less than n
PREIMAGE $f\ s$	$\{x \mid f\ x \in s\}$	The inverse image of the subset s
$\{x \mid P(x)\}$	$\{\lambda x.P(x)\}$	Set of all x that satisfy the cond. P
$x\ pow\ n$	x^n	$real\ x$ raised to num power n
$exp\ x$	e^x	Exponential log. on x
SIGMA $f\ s$	$\sum_s f$	Sum of the sequence $f(x); x \in s$
$lim(\lambda n.f\ n)$	$\lim_{n\to\infty} f(n)$	Limit of the $real$ sequence f

Table 1 HOL Symbols.

Definition 3.1:

```
⊢ ∀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 3.2:

```
⊢ ∀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 (Mhamdi, 2012) 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.

3.3 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 (Liu, 2004). 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 $(0 \leq j < n)$ 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. These subsets will be working independently and alternatively in a round-robin fashion. During a given working round T_i, only the nodes belonging to S_i are turned on to detect a potential event. Otherwise, the nodes of the subset S_i will fall asleep. Whereas, during all the other scheduled rounds, the nodes of the subset S_i will fall asleep.

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.

4 The Optimal Lifetime Problem under QoS Constraints

4.1 Problem Description

In the context of a WSN using the randomized scheduling, the network lifetime is "the elapsed time during which the network functions well" (Xiao et al., 2006, 2010). The network lifetime, denoted by T_{Nlife}, has been mathematically defined as follows (Xiao et al., 2006, 2010).

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

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

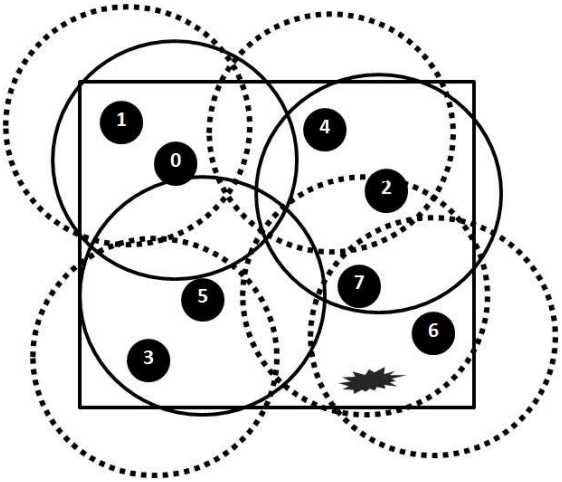

Figure 1 The k-set randomized scheduling for ($n = 8$) nodes and ($k = 2$) subsets.

The work in (Elleuch et al., 2011, 2015) 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, this work 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 (Elleuch et al., 2011).

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

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 (Liu, 2004; Liu et al., 2006), 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 (Xiao et al., 2006, 2010) 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.

Based on 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.

4.2 Formalization in HOL

The definition of the network lifetime, given in Equation (2), implies that 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 ($n = c$) (Xiao et al., 2006, 2010).

In Theorem 2, we formally verify the main condition so that the lifetime problem has an optimal solution (Xiao et al., 2006, 2010).

Theorem 2:

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

is bounded and non-empty.

$Qep = \frac{(Q-1+s)(Q^2-1+s)}{2Q(Q+1)}$, where $Q = \lceil \frac{L}{Ts} \rceil$. We recall that L is the duration of an occurring event and Ts is the length of a scheduling cycle. The parameter s is the remainder of the intrusion period L in terms of the number of slots Ts, such that $s = \frac{L}{Ts} + 1 - \lceil \frac{L}{Ts} \rceil$.

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 (Xiao et al., 2006, 2010) provide a very abstract proof deducing directly 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.

4.2.1 The Detection Delay

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

$$S_D = \{k \mid D \leq QoS_{DD} < Qep\left[1-(1-q)^n\right], n = c\} \tag{5}$$

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 (Elleuch et al., 2015), 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 is bounded.

```
⊢ ∀(s:num->bool). FINITE s ⇒ BOUNDED s.
```

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

Lemma 5: ***The set S_D is bounded***

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

Proof. The proof is based on Theorem 4, which further requires to satisfy the assumptions of Theorem 3. In particular, we need to adjust the right value of ε, in order to prove that the set S_D is finite. We recall our results on the limiting value of the detection delay DD, as well as the asymptotic behavior of the delay D versus k (Elleuch et al., 2015). Since we have formally verified that $\lim_{k\to\infty}$ DD = $\frac{(Q-1+s)(Q^2-1+s)}{2Q(Q+1)}\,[1-(1-q)^n]$, we choose $\varepsilon = (\lim_{k\to\infty} DD) - QoS_{DD}$. Moreover, since $DD(k)$ is increasing on k, the maximum possible values is $\lim_{k\to\infty} D$. We thus get the quality of service value $QoS_{DD} < lim_{k\to\infty} D$. We can deduce that the set S_D is finite, and thus bounded.

We conclude that S_D is non-empty, using the monotonicity of the detection delay $DD(k)$ on k (Elleuch et al., 2015), along with some reasoning on the quality of service constraints. in fact, when the detection delay $DD(k)$ is increasing versus k, the minimum delay value, is induced for $(k = 1)$, i.e, $D(1)$. The values of $D(k)$; including QoS_{DD}, cannot hence go below $D(1)$. 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.

4.2.2 The Detection Probability

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

$$S_{Pd} = \{k \mid P_{d|k=1} \geq P_d \geq QoS_{DP} > 0, n = c\} \tag{6}$$

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

Lemma 6: ***The set*** S_{Pd} ***is bounded***

```
⊢ ∀q n s L Ts QoSDP. (1 ≤ n) ∧ (0 < s < 1) ∧ (0 < L) ∧
 (0 < Ts) ∧ (0 < q < 1) ∧ (∀k. L < k×Ts) ∧
 (0 < QoSDP < 1)
 ⇒ BOUNDED {k | QoSDP ≤ Pd p n k s L Ts q}.
```

Proof. Similar to the proof of Theorem 5, we establish that the set S_{Pd} is bounded using Theorem 4. For that, we first achieve the proof that S_{Pd} is finite. Hence, we consider Theorem 3 such that $(a = 0)$ and $(\varepsilon = QoS_{DP})$. The limiting behavior of the detection probability P_d regarding the parameter k is required. In (Elleuch et al., 2015), we have already shown that $\lim_{k\to\infty} P_d = 0$.

A very simplified proof sketch of Lemma 6 is provided in Figure 2, through a backward chaining. Using the HOL tactic `MATCH_MP_TAC`, together with rewriting `RW_TAC`, the main goal is matched, through Theorem 4, into `FINITE SPd'`. During this step, the detection probability P_d is simultaneously substituted with its main mathematical expression, using our development from the detection probability theory (Elleuch et al., 2015). Theorem 3 is then instantiated using appropriate values of $(\varepsilon = QoS_{DP})$ (`MP_TAC`). The corresponding assumption regarding the limiting behavior of the detection probability should be now satisfied. The latter result is available in the detection probability theory (Elleuch, 2013). Finally, applying a solving tactic called `METIS_TAC`, along with some real analysis, the main goal can be achieved.

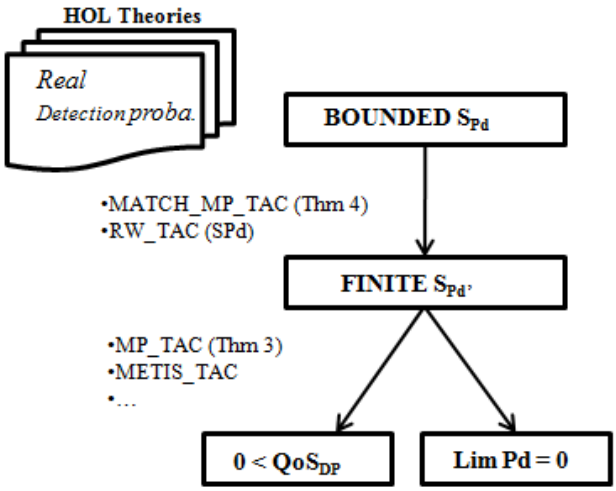

Figure 2 A Simplified Proof Sketch for Lemma 6.

Since the detection probability is decreasing with k (Elleuch et al., 2015), the best detection probability value is ensured for $(k = 1)$. So, we get $P_d(1) > P_d(k)$. On the other

hand, it is clear that the quality of service value; QoS_{DP}, 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.

4.2.3 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.

$$S_{Cn} = \{k \mid 1 \leq k \leq \frac{q}{(1-(1-QoS_{C_n})^{\frac{1}{n}})}, n = c\} \tag{7}$$

Lemma 7: ***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 (Elleuch et al., 2013a), 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 5, 6, and 7, respectively, together with the fact that $(k = 1)$ is shown to be in each of the three sets, and hence in their intersection.

The formal analysis of the optimal lifetime problem required many intermediate results in relation with the behavior of the three performance attributes, which are the detection delay, the detection probability and the network coverage. A brief summary of the required lemmas is provided in Table 2 (Elleuch et al., 2013a, 2015), and the interested reader can access them from (Elleuch, 2013).

Lemma	**Formulation**
DD is an increasing function of k	`mono_incr (DD)`
Limit of DD when k is very large	$\frac{(Q-1+s)(Q^2-1+s)}{2Q(Q+1)}\left[1-(1-q)^n\right]$
Pd is a decreasing function of k	`mono_decr (Pd)`
Pd definitely decreases when k is very large	$\lim_{k \to +\infty} Pd = 0$
C_n is a decreasing function of k	`mono_decr (Cn)`

Table 2 Required Lemmas for the Different Sets.

4.3 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 (Hewish, 2001). 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 WSN for border monitoring usually suffer from limited lifetime (Arora et al., 2004), e.g, a REMBASS sensor can be operational for 30 days only (Hewish, 2001). Thus, the k-set randomized scheduling algorithm has been proposed for use to save energy for a border monitoring application (Xiao et al., 2009).

(Elleuch et al., 2015) presented the higher-order-logic formalizations of the detection performances for randomly-scheduled WSN. The practical effectiveness of these developments, have been then illustrated, through analyzing a WSN for border surveillance (Xiao et al., 2009; Sun et al., 2011). In this paper, we focus on formally analyzing the optimal lifetime problem, presented in Section 4.1, 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 ($D < QoS_{DD} = 15s$) (Arora et al., 2004). 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.

Based on the formalizations developed in the previous section, we can easily establish that, for our border security monitoring application, we have:

$$\begin{aligned} S_{app} = \{k \mid & D \leq (QoS_{DD} = 15) < Qep\left[1 - (0.72)^n\right], \\ & 1 - (0.72)^n \geq P_d \geq (QoS_{DP} = 0.95) > 0, \\ & 1 \leq k \leq \frac{0.28}{(1 - (1 - QoS_{C_n})^{\frac{1}{n}})}, \\ & 0 < QoS_{C_n} < 1, n = c\} \end{aligned} \tag{8}$$

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 (Arora et al., 2004) such that ($QoS_{DP} = 0.95$) and ($QoS_{DD} = 15s$). It is worth to mention that the formal analysis of the network lifetime, developed so far, can be quite valuable to analyze any randomly-scheduled application like a general surveillance framework for WSN.

5 The Intrusion Coverage under Object size

5.1 Problem Description

The performance of the randomized nodes scheduling has been usually studied assuming that the intrusion object is simply a point in a plane monitored area of size a (Liu et al., 2006; Xiao et al., 2010). Nevertheless, in practical applications, such assumptions are not realistic at all, and have important impacts on the performance of the whole application. Obviously, detected vehicles in a traffic monitoring application are completely different in

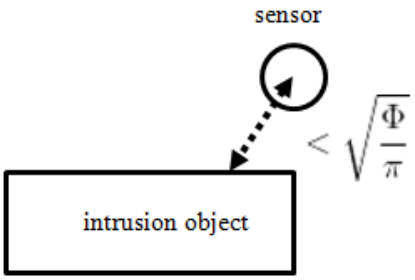

Figure 3 A Rectangular Intrusion Object.

size and shapes from the enemy tanks crossing a country border. Modelling such objects by points is an over-approximation, that does not reflect the real world.

In particular, the coverage property of the network heavily depends on some of the intrusion object attributes, as well as, the monitored area. In general, a given point is said to be covered, if any occurring event at this point, is detected by at least one active node with a given probability (Liu et al., 2006; Xiao et al., 2010). Previously, the probability that a sensor covers a given point was specified as $\frac{r}{a}$, where r is the size of the sensing area of each sensor, a is the size of the monitored area. Under the new assumptions, the probability that a sensor covers a given point in the field, is no longer significant (Liu et al., 2006; Xiao et al., 2010), and has been rather redefined into the probability that a sensor node can detect an intrusion object with size o (Xiao et al., 2008a, 2009). Considering different shapes of the plane with different object forms, the new probability that a sensor node can detect an intrusion object with a certain size, is described in the sequel.

Assume a two-dimensional plane, the object to detect has been abstracted to 2 forms: a circle or a rectangle (Xiao et al., 2008a, 2009). These can be direct approximations of various potential objects, such as a land mine or a military vehicle. The intrusion object of size o, is modelled either as a circle having as a diameter $2\sqrt{\frac{o}{\pi}}$, or a rectangle with length b and width $\frac{o}{b}$. The sensing area of a sensor is circular and is denoted here as Φ. In order to detect the intrusion, the center of the sensor node should intuitively overlap the area, or at least the borderline, of the intrusion object. In other words, the distance from the sensor center to the object border should be less than $\sqrt{\frac{\Phi}{\pi}}$ (Figure 4).

The sensor detection probability in case the intrusion object is a circle is defined as:

$$\begin{aligned} q2c &= \frac{\pi}{a}\left(\sqrt{\frac{\Phi}{\pi}}+\sqrt{\frac{o}{\pi}}\right)^2 \\ &= \frac{1}{a}\left(\sqrt{\Phi}+\sqrt{o}\right)^2, circle \end{aligned} \tag{9}$$

In case the intrusion object is a rectangle, the detection probability of a sensor is:

$$q2r = \frac{1}{a}\left(o+2(b+o/b)\sqrt{\Phi/\pi}+\Phi\right), rectangle \tag{10}$$

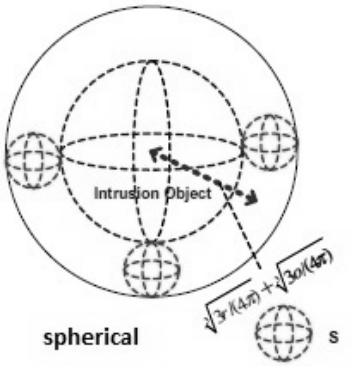

Figure 4 A Spherical Intrusion Object (Xiao et al., 2009).

Since the approximation of the area into a 2D plane does not hold for all WSN applications, better abstractions have been investigated in (Xiao et al., 2008b, 2009) to meet the requirements of different applications like underwater monitoring. Hence, in a 3-dimensional case, the sensing area of an individual sensor is of a spherical shape, and an object, can nearly be of two shapes: spherical or cuboid.

In case the object is abstracted to a sphere, which occupies an area of size o, its radius is mathematically $\sqrt[3]{3o/(4\pi)}$. Otherwise, if the object has a cuboid form, the length, width, and height are denoted as b, h, and $o/(bh)$, respectively. Similarly, a sensor node, with a sensing area of Φ, cannot overlap an intrusion object if it is far from the boundary of the intrusion object, by a distance more than its radius, i.e., $\sqrt[3]{3\Phi/(4\pi)}$. (Figure 5).

When the intrusion object is of spherical form, the detection probability of a sensor is specified in the following equation.

$$\begin{aligned} q3c &= \frac{4\pi}{3a}(\sqrt[3]{3\Phi/(4\pi)} + \sqrt[3]{3o/(4\pi)})^3 \\ &= \frac{1}{a}(\sqrt[3]{3\Phi} + \sqrt[3]{o})^3, spherical \end{aligned} \tag{11}$$

The detection probability of a sensor, in case the intrusion object is a cuboid, is expressed by:

$$\begin{aligned} q3s &= \frac{1}{a}[o + 2(bh + o/b + o/h)\sqrt[3]{3\Phi/(4\pi)} \\ &\quad +\Phi + (b + h + o/(bh))\pi(\sqrt[3]{3\Phi/(4\pi)})^2] \\ &= \frac{1}{a}[o + 2(bh + o/b + o/h)\sqrt[3]{3\Phi/(4\pi)} \\ &\quad +\Phi + (b + h + o/(bh))\sqrt[3]{9\Phi^2/16\pi^2}], cuboid \end{aligned} \tag{12}$$

Next, we will make use of the new expressions of intrusion detection probability, specified above, to develop the higher-order-logic formalizations of the intrusion coverage under two and three dimensional spaces. We will also formally analyze the influence of the new assumptions on the behavior of the coverage property.

5.2 Formalization of the Intrusion Coverage

We are first interested in formally verifying the probability that the k-set randomized node scheduling produces an empty partition or sub-network. Then, we will describe the higher-order-logic formalizations of the intrusion coverage under new assumptions associated to the intrusion object (Xiao et al., 2009).

5.2.1 The probability of an empty scheduled partition

In the k-set randomized scheduling algorithm, each sensor node randomly selects a unique number i out of the k available options. The k generated subsets of nodes $\{S_i, 0 \leq i \leq (k-1)\}$ are thus disjoint, i.e., a given node belongs to one subset at once. Afterwards, these node subsets are scheduled to work alternatively within their scheduling time slots $\{T_i, 0 \leq i \leq (k-1)\}$.

To emphasize on the impact of the random feature inherent to the k-set randomized scheduling algorithm, we consider the example of a randomly-scheduled WSN where the set of n nodes is partitioned into $(k = 3)$ subsets: S_0, S_1 and S_2 (see Fig. 3). Let t_0 be any reference time, while an intrusion event e, lasting L, starts at time t_z. Due to the probabilistic feature of the scheduling algorithm, the sub-network S_2 does not contain any node. Since the subsets are working by rounds, a complete time slot is allocated to the subset S_2 at every turn, but there are no active nodes to detect the event e during the whole time slot. In an other scenario, all n nodes may randomly be assigned into the same partition giving only one subset, which is non-empty. In this case, there will be a single round during which an event is likely to be covered. The empty sub-networks of nodes, generated by the randomized scheduling, hence have a significant effect on the overall network performance.

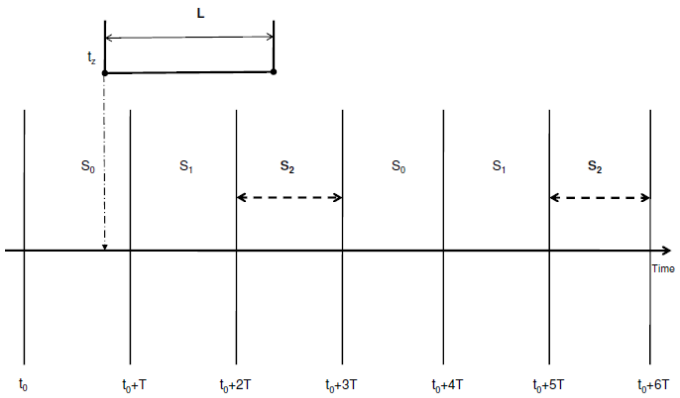

Figure 5 An example of the k-set randomized scheduling for n nodes and $k = 3$.

The assignment of the nodes to the subsets is uniformly done, so that the random organization of the nodes into several sub-networks is fair over the whole network. Each

node intuitively joins a single subset with the same probability $\left(\frac{1}{k}\right)$. The appropriate random variable, required in this formalization, should uniformly distribute the nodes over the k sub-networks, i.e., a Uniform random variable (Definition 5.1).

Definition 5.1: ***The Uniform random variable***

```
⊢ ∀X p k.
 uniform_distr_rv X p k = (real_random_variable X p) ∧
 (IMAGE X (p_space p) = IMAGE (λx.&x) (count (SUC k))) ∧
 (∀m. m ∈ IMAGE X (p_space p) ⇒
     (distribution p X {m} = 1/&k)).
```

where X is a real-valued random variable; `real_random_variable`, which takes values on the integer interval $[0..(k-1)]$, i.e., `(IMAGE (λx.&x) (count (SUC k)))` with the probability distribution; `distribution`, equals $\left(\frac{1}{\&k}\right)$. The operator `&`, used in the above definition, allows the conversion of the natural number k into its extended real number counterpart.

The set of n nodes is uniformly partitioned into k sub-networks, a node hence joins a given subset S_j with the uniform probability $\left(\frac{1}{k}\right)$. The same node will miss the same subset with the complement probability $\left(1-\frac{1}{k}\right)$. Consequently, a given subset S_j is empty if and only if the n sensors do not join, i.e., miss this subset. More formally, consider the event $T_{i,j}$: "The sensor i does not join the subset S_j", we have then:

$$\begin{aligned} Pr(S_j\ is\ empty) &= Pr(n\ sensors\ do\ not\ join\ S_j) \\ &= Pr(T_{0,j} \cap ... \cap T_{(n-1),j}) \end{aligned} \tag{13}$$

where

$$Pr(T_{i,j}) = \left(1 - \frac{1}{k}\right) \tag{14}$$

Since the n sensor nodes miss the subset S_j independently, the events $T_{0,j}$,...,$T_{(n-1),j}$ will be mutually independent, which means that any given event is completely independent of the intersection of any other events (Feller, 1968). Accordingly, we successfully verify, in Theorem 8, the probability that a given subset S_j is empty in a randomly-scheduled WSN.

Theorem 8: ***The basic probability of an empty subset***

```
 ⊢ ∀X p k n j.
 (prob_space p) ∧ (1 < k) ∧ (uniform_distr_rv X p k) ∧
 (j ∈ IMAGE X (p_space p)) ∧
 (∀s m. indep p ({x | X x ≠ m} ∩ (p_space p))
 ({x | subset_empty m (rd_subsets s (X x))} ∩ (p_space p)))
 ⇒ (prob p ({x | subset_empty j (rd_subsets n (X x))} ∩
p_space p) = (1 − 1/&k)^n).
```

where

- The assumption `(1 < k)` ensures that the number of sub-networks is greater than 1 since the randomized scheduling would be meaningless for $(k = 1)$.
- `(uniform_distr_rv X p k)` is the Uniform random variable, given in Definition 5.1.
- The event `({x | subset_empty j (rd_subsets n (X x))} ∩ p_space p)` formally models the event of the probability given in Equation (14), i.e., the event "The subset S_j is empty". The function `rd_subsets` hence generates the output values of the Uniform random variable X ordered as a list of length n, in which the predicate `subset_empty` looks for the index j.
- The last assumption ensures the mutual independence over the set of the $T_{i,j}$ events (Equation (13)) using the HOL function `indep`.

Proof. The proof of the above theorem is based on induction and the multiplication rule, which switches the probability of a set of independent events to the product of their respective probabilities, i.e., $Pr(\bigcap_{i=0}^{(n-1)} T_{i,j}) = \prod_{i=0}^{(n-1)} Pr(T_{i,j})$. To complete the proof, the verification of the probability distribution of the Uniform random variable, $Pr(T_{i,j})$, and its complement, along with set theoretic analysis was required.

Since a sub-network is either empty or not, we can model such behavior by simply a Bernoulli random variable `Y`, with the success probability `(prob p ({x | subset_empty j (rd_subsets n ((X k) x))} ∩ p_space p))`, verified as $\left(1 - \frac{1}{\&k}\right)^{n}$. A non-empty sub-network is thus described by a Bernoulli random variable (Elleuch et al., 2015) with the complement probability of $\left(1 - \frac{1}{k}\right)^{n}$, where n is the number of covering sensors for a given point.

Definition 5.2:

```
⊢ ∀X p k n. sbst_non_empty_rv X p k n =
  bernoulli_distr_rv X p (1 − (1 − 1/&k)^n).
```

While it would have been much simpler to directly model an empty sub-network by a Bernoulli random variable, the above analysis has been useful to concretely show the logical reasoning and justify the origin of the associated probability and parameters. It is important to note that, contrarily to the formalizations presented in (Elleuch et al., 2011) within the probability theory (Hasan, 2008), the above resulting formalizations is completely different by using the most recent probability theory within the HOL theorem prover (Mhamdi, 2012).

5.2.2 The Intrusion Coverage

Generally, the coverage property, is a spatial performance attribute which measures how well the area of interest is monitored or tracked by the sensor nodes (Wang and Xiao, 2006). In Figure 6, the point B is covered by the sensor at the bottom, whereas, point A is left uncoverd since it does not belong to any of the sensing range of the two sensors. If there are uncovered points, then the coverage is said to be partial.

Every point of the monitored area is thus characterized by a coverage value, called the coverage intensity of a specific point C_p, whereas the coverage over the whole network;

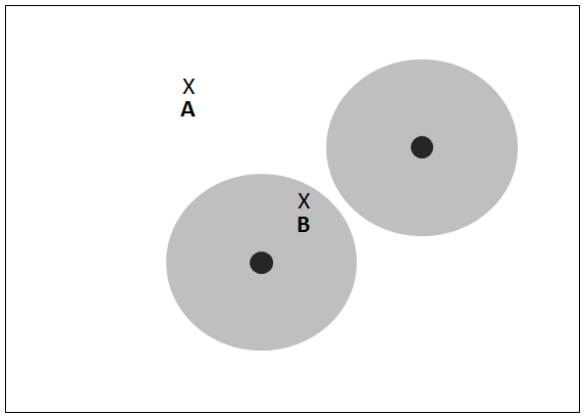

Figure 6 Coverage Example.

C_n, is the average of C_p. Under new assumptions of size and shape, the network coverage intensity, as proposed in (Liu et al., 2006; Xiao et al., 2010), is no longer appropriate for large intrusion objects, and the coverage performance is rather measured through the intrusion coverage/detection, denoted as V_n. Thus, the main metric C_n has been reconsidered, into V_n for intrusion coverage (Xiao et al., 2009) (Equation (15)).

$$V_n = E[V_o] \tag{15}$$

where V_o is the intrusion coverage intensity for a specific object.

Given an intrusion object, of any form, the average time this object is covered in a whole scheduling cycle of length $k \times T$ (Liu et al., 2006; Xiao et al., 2009), defines the intrusion coverage intensity of a specific object, which we denote as V_o. In a WSN deployed with randomized scheduling, the intrusion object is covered if the actual active partition contains at least one covering node, i.e., is not empty of covering nodes. The intrusion coverage intensity of a specific object; V_o, is mathematically specified similarly as the coverage intensity of a specific point; C_p (Liu et al., 2006; Xiao et al., 2010).

$$V_o = \frac{E[\sum_{j=0}^{k-1} Y_j] \times T}{k \times T}. \tag{16}$$

where Y_j is the random variable describing a non-empty subset, and E denotes an expectation.

The following mathematical expression for the intrusion coverage intensity of a given object; V_o, has been formally verified in Theorem 9, similarly as in (Elleuch et al., 2016a).

Theorem 9:

```
⊢ ∀Y p k s c.
(prob_space p) ∧ (FINITE s) ∧ (1 < k) ∧
(CARD s = k) ∧ (∀i. i ∈ s ⇒ sbst_non_empty_rv (Y i) p k c)
⇒ (intrcov_intsty_pt p Y k s c = Normal (1 − (1 − 1/k)^c)).
```

where

- `(sbst_non_empty_rv (Y i) p k c)` describes a non-empty partition of covering nodes, which is basically a Bernoulli random variable with the complement probability of $\left(1-\frac{1}{k}\right)^{c}$, where c is the number of covering sensors for a given point.
- The assumption `(∀i. i ∈ s ⇒ sbst_non_empty_rv (Y i) p k c)` ensures that every element of the set s is a random variable `sbst_non_empty_rv`.
- The predicate `intrcov_intsty_pt` defines the coverage behavior a specific intrusion object, as in Equation (16).
- The HOL function `Normal` is used to convert a real value to its corresponding value in an extended real.

The average or the expectation value of the intrusion coverage intensity ; V_o, over all points of the monitored area, describes the intrusion coverage ; V_n (Equation (15)). Based on the final expression of V_o, shown in Theorem 9, we can rewrite

$$V_n = E[1-\left(1-\frac{1}{k}\right)^{c}]\,. \tag{17}$$

The expression of V_n mainly depends on c which is the number of nodes covering a given intrusion object in the field. The event of covering an intrusion or not is usually characterized using a Bernoulli trial with probability $q=\frac{\Phi}{a}$, where Φ is the size of the sensing range of each sensor, and a the size of the field. When taking into account the outcomes of the parameter c within the n nodes of the network, we find that it describes a Binomial random variable, denoted as C, with arguments n; the number of nodes, and q; the probability that a sensor covers a given intrusion. The probability q is hence exactly the same as the probability that an object is within the sensing range of a sensor. Accordingly, the different probabilities have been already discussed and expressed for different object shapes (Equations (9), (10), (11), and (12)).

In the previous work when abstracting an object to a point (Elleuch et al., 2011, 2015, 2016a), it has been possible to consider the probability of covering a point as a simple constant value q. Giving the assumptions of size and shape of the intrusion, additional parameters should be involved. We now redefine the Binomial random variable with n trials and success probability $P1$, such that $P1$ is a function in order to correctly reflect the variation of the different parameters related to the size and shape of the intrusion object. The function $P1$ thus depends on the input parameters Φ: the sensing range of an individual sensor, o: the size of the intrusion object, a: the size of the area, b: the potential length of the object, and h its potential height. It can hence take one of the probability values that have been already specified for an object of the form of a circle (Equation (9)) and a rectangle (Equation (10)), in a two-dimensional plane, or a spherical (Equation (11)) and a cuboid (Equation (12)), in a three-dimensional plane.

$$(P1\ \Phi\ o\ a\ b\ h) = q2c \vee q2r \vee q3s \vee q3c \tag{18}$$

Definition 5.3:

```
⊢ ∀X p q n Φ o a b h.
 bino_intr_rv X p n (P1 Φ o a b h) =
 (real_random_variable X p) ∧
 (IMAGE X (p_space p) =
 IMAGE (λx.&x) (count (SUC n))) ∧
 (∀m. &m ∈ (IMAGE X (p_space p)) ⇒
 (distribution p X {&m} = &(binomial n m) ×
 (P1 Φ o a b h)^m × (1 − (P1 Φ o a b h))^(n−m))).
```

where p is the probability space over which the real random variable X is defined, `IMAGE (λx.&x) (count (SUC n))` gives the possible values that can take the Binomial. The input argument $(P1\ \Phi\ o\ a\ b\ h)$ is the probability function, which denotes that the intrusion object is within a sensor range. The function `binomial`, used in the above definition, is the higher-order-logic formalization of the binomial coefficient for reals.

Since the variable C is a random variable, the intrusion coverage intensity over the network V_n, shown in Equation (17), is an expectation of a function of the Binomial variable C, and not a simple expectation. The intrusion coverage intensity of the whole WSN with n nodes has been formally formalized through the following function `intrcov_netw`, shown in Definition 5.4.

Definition 5.4:

```
⊢ ∀Y p k s C n Φoabh.
intrcov_netw p Y k s C n (P1 Φ o a b h) =
expectation p (λx. intrcov_intsty_pt p Y k s (num (C x))).
```

where p: the probability space, Y: a random variable that returns an extended real number, s: the summation set whose cardinality is k, C: the Binomial random variable for the number of covering nodes, n: the total number of nodes, and the probability function $P1$ that an intrusion object can be detected. The function `num`, has to convert an extended real; `(C x)`; to its corresponding natural value, using the real function `floor`.

Using the higher-order-logic formalizations developed above, we formally verify in the following theorem, the final network intrusion coverage intensity.

Theorem 10:

```
⊢ ∀p Y k s C n Φoabh.
 (prob_space p) ∧ (0 < (P1 Φ o a b h) < 1) ∧
 (events p = POW (p_space p)) ∧ (1 ≤ n) ∧ (1 < k) ∧
 FINITE s ∧ (CARD s = k) ∧ (sn_covers_p C p (P1 Φ o a b c) n) ∧
 (expectation p C ≠ PosInf) ∧ (expectation p C ≠ NegInf) ∧
 (∀i x. (i ∈ s) ∧ (x ∈ p_space p) ⇒
 sbst_non_empty_rv (Y i) p k (num(C x)))
 ⇒ (intrcov_netw p Y k s C n (P1 Φ o a b h) = Normal
(1 − (1 − ((P1 Φ o a b h)/(&k))^n))).
```

- The assumption `(events p = POW (p_space p))` specifies the events set as the power set of the sample space Ω.

- The assumption `(0 < (P1 Φ o a b h) < 1)` ensures that the intrusion coverage probability lies in $]0..1[$.
- `sn_covers_p` is the Binomial random variable (Definition 5.3) such that its expectation is finite, i.e., `(expectation p C ≠ PosInf) ∧ (expectation p C ≠ NegInf)`. The variables `(PosInf)` and `(NegInf)` are the higher-order-logic formalizations of positive infinity and negative infinity, respectively.

Proof. Besides the linearity of the expectation property, the proof of the above theorem is mainly based on verifying the expectation of a function of a random variable. A considerable amount of real analysis associated to the use of the new probability function $P1$, in particular for the rational power, was also needed.

5.3 Formal Behavioral Analysis of the Intrusion Coverage

Based on the verified expression of V_n, in Theorem 10, the intrusion coverage behavior can now be investigated according to the different design parameters, particularly o: the size of the intrusion object, b: the length of the object, and h its height. Indeed, there is a typical scenario when the intrusion object has a size o, which is as large as the monitored field a. In this case, a unique sensor is intuitively required for detection. Hence, more efficient deployments of the sensors can be studied according to the sizes and shapes of the intrusion.

In Lemma 11, we formally verify the relationship between the parameters o and b, for which the minimum of the intrusion coverage degree V_n is achieved, when the object occupies a rectangular area.

Lemma 11:

Under a two dimensional plane, the lower bound of the intrusion coverage V_n is achieved with $b = \sqrt{o}$ for a rectangular intrusion.

Various asymptotic properties, already verified, for the network coverage can also be directly reverified for the intrusion coverage. Given a value t for the intrusion coverage intensity and an intrusion object size, with attributes o, b, and h, we successfully verify, in Lemma 12, the minimum number of sensors; n_{min}, to deploy to ensure an intrusion coverage V_n as t.

Lemma 12:

```
⊢ ∀p Y s k C n t Φoabh.
(1 ≤ n) ∧ (1 < k) ∧ (0 < (P1 Φ o a b h) < 1) ∧ (0 < t < 1) ∧
 (Normal t ≤ intrcov_netw p Y k s C n (P1 Φ o a b h))
⇒ ⌈ ln(1−t) / ln(1−(P1 Φ o a b h)/k) ⌉ ≤ &n.
```

To achieve an intrusion coverage of at least t, we formally verify, in Lemma 13, the maximum on the number of disjoint subsets k for a given n.

Lemma 13:

$$\vdash \forall p\ X\ s\ k\ C\ n\ \Phi\mathtt{oabh}.\ (1 \leq n) \wedge (1 < k) \wedge (0 < (\mathtt{P1}\,\Phi\mathtt{oabh}) < 1) \wedge (0 < t < 1) \wedge (Normal\ t \leq (intrcov\_netw\ p\ Y\ k\ s\ C\ n\ (\mathtt{P1}\,\Phi\mathtt{oabh}))) \Rightarrow \mathtt{k} \leq \frac{(\mathtt{P1}\,\Phi\mathtt{oabh})}{1-\mathtt{e}^{\frac{\ln(1-\mathtt{t})}{(\delta\mathtt{n})}}}.$$

In this section, we described our higher-order-logic formalizations of the intrusion coverage intensity, for a more realistic randomized algorithm (Xiao et al., 2008a,b, 2009), under assumptions of object size and shape. These formalizations have been then useful to formally reason about the intrusion coverage performance behavior.

6 Discussion

In this paper, we first developed, within the HOL theorem prover, the formal analysis of the optimal lifetime problem (Equation 4) under Quality of Service (QoS) constraints, for wireless sensor networks using the k-set randomized scheduling (Liu et al., 2006; Xiao et al., 2010). The main QoS constraints are related to the most energy-related performance metrics, i.e., the network coverage, the detection probability and the detection delay. For that, we built upon the higher-order-logic formalizations of these performance metrics (Elleuch et al., 2011, 2013b, 2015), to verify the minimal set of conditions on the k-values, and formally verify the network lifetime related characteristics of a border security monitoring application, with real QoS values for the detection probability and the detection delay. The randomized scheduling, formalized so far, usually assumes that the intrusion is simply a point, which does not reflect the practical WSN applications. Intuitively, the sensing abilities of the deployed WSN are depending on the size of the intruding object, which can vary from a vehicle to a tiny particle. For example, for detecting large objects fewer sensors are required in the field. Subsequently, the second part of this paper had been devoted to formalize a modified version of the randomized scheduling (Xiao et al., 2008a,b, 2009), under assumptions of object size and shape in a two or three dimensional plane.

The analysis, presented in this paper, primarily illustrates the effectiveness 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 (Elleuch et al., 2011, 2013a,b), along with the detection probability and delay, presented in (Elleuch et al., 2015). 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. On the other hand, the formalization of the intrusion coverage definitely shows the usefulness of our previous theoretical developments on coverage. Although, the assumptions on the intrusion object are modified, there were some similarities in the specification, which allowed us to leverage upon the previous coverage theory. The whole successful verification, achieved in this paper, thus clearly highlights the main advantages of our previous theoretical developments in terms of precision and coherence.

On the other hand, the theorem proving technique allows us to verify universally quantified generic expressions. Obviously, such generality cannot be achieved by traditional approaches, such as simulation and model checking. Moreover, these traditional approaches usually involve pseudo-random variables, which add another factor of inaccuracy in the analysis. Similarly, manual paper-and-pencil based analysis techniques are prone to human

error and thus the analysis may not include some essentials assumptions. Whereas, due to the formalization support of the probability theory (Mhamdi, 2012; Hasan and Tahar, 2015), available in the HOL theorem prover, we have been able to provide an accurate formalization through an appropriate modeling of randomness.

Comparable to the other performance aspects, many challenges were encountered in the verification, described in this paper. While the lifetime proof seems simple, there were many implicit steps that make the understanding of the main proof quite challenging. Indeed, previous formalizations on coverage and detection (Elleuch et al., 2011, 2015) were focused on formally verifying the expressions associated with the probabilistic attributes of interest. However, 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. In addition, 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 (Xiao et al., 2006, 2010), 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 intrusion coverage part required a considerable effort to correctly adapt the previous HOL code of coverage, and match it to the new context of object size and shape. The higher-order-logic definition of the intrusion coverage intensity over the network should consider the new values of detection probabilities, while taking into account the major modifications in the design parameters. While we previously defined the probability of covering a point as a constant, the same probability has to be reconsidered as a variable function for the intrusion coverage behavior depending on the main arguments; Φ: the sensing range of an individual sensor, o: the size of the intrusion object, a: the size of the area, b: the potential length of the object, and h its potential height. This way we have been able to formally study the intrusion behavior by varying these parameters, which provides us with useful insights for practical WSN applications.

The formalization, presented in this paper, consumed approximatively 100 man hours and 800 lines of HOL code. Indeed, the high degree of interactivity required while reasoning in a theorem prover hinders the formalization process. Enormous efforts are thus sometimes required to prove a basic result or just to correctly handle complicated summations. Some of the lemmas, especially regarding monotonicity, were very tricky to prove and required very lengthy proofs to achieve the real analysis. For instance, the proof of one lemma, which occupied about half a page in the original textbook (Xiao et al., 2010), may took over pages of HOL code. The proof of the presented theorems required many intermediate mathematical results, which formalization was very tedious in HOL. Moreover, in some instances, we had to develop alternative mathematical proofs, completely different from the reference proofs in (Xiao et al., 2008a,b, 2009), because the formal verification of the given proof was not possible in HOL4 due to the unavailabilty of some mathematical results.

7 Conclusions

In this paper, we described a reliable approach for the formal analysis of wireless sensor networks using the k-set randomized scheduling to preserve energy. Based on the earlier

work of (Elleuch et al., 2011, 2015), we provided the higher-order-logic formalizations of the lifetime maximization problem (Xiao et al., 2010), under Quality of Service (QoS) constraints related to the network coverage and the detection performances. After that, we reconsider the previous network coverage formalization into the intrusion coverage, using a modified version of the randomized algorithm (Xiao et al., 2008a,b, 2009), under assumptions of object size and shape in a two or three dimensional plane.

On the other hand, our theorem-proving based approach allows a generic formal verification of randomly-scheduled WSN regardless of the values of the design parameters. Evidently, such results cannot be achieved with existing approaches, such as traditional paper-and-pencil probabilistic modelling, simulation and probabilistic model checking. In addition, 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 WSN, our work is distinguishable by addressing the performance aspects.

Interesting perspectives to the lifetime analysis, presented in this paper, include the formal quantification of the energy consumption and the formalization of the reliability models, for randomly-scheduled WSN. Probabilistic model checkers may be explored for such perspectives. Also, as future work, finding the optimal values of k, can be also investigated. On the other hand, the developed results can be refined by modeling a realistic sensor deployment following a two-dimensional Gaussian distribution. In this regard, the higher-order-logic of the Gaussian random variable, developed in (Qasim et al., 2016), can be very helpful. Also, the current approach for the network lifetime would be very useful to investigate the formalization of the optimal detection probability in (Olteanu et al., 2010).

References

A. Arora, P. Dutta, S. Bapat, V. Kulathumani, H. Zhang, V. Naik, V. Mittal, H. Cao, M. Demirbas, M. Gouda, Y. Choi, T. Herman, S. Kulkarni, U. Arumugam, M. Nesterenko, A. Vora, and M. Miyashita. A Line in the Sand: a Wireless Sensor Network for Target Detection, Classification, and Tracking. *Computer Networks*, 46(5):605–634, 2004.

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

P. Ballarini and A. Miller. Model Checking Medium Access Control for Sensor Networks. In *Proceedings of the Symposium on Leveraging Applications of Formal Methods, Verification and Validation*, pages 255–262. IEEE Computer Society, 2006.

C. Bernardeschi, P. Masci, and H. Pfeifer. Analysis of Wireless Sensor Network Protocols in Dynamic Scenarios. In *Stabilization, Safety, and Security of Distributed Systems*, volume 5873 of *Lecture Notes in Computer Science*, pages 105–119. Springer, 2009.

D. Chen and P. K. Varshney. QoS Support in Wireless Sensor Networks: A Survey. In *Proceedings of the International Conference on Wireless Networks*, pages 227–233. CSREA Press, 2004.

M. Elleuch. *Formalization of the Detection Properties of WSNs in HOL*, 2013. HOL Code: `http://hvg.ece.concordia.ca/projects/prob-it/wsn.php`.

M. Elleuch, O. Hasan, S. Tahar, and M. Abid. Formal Analysis of a Scheduling Algorithm for Wireless Sensor Networks. In *Formal Methods and Software Engineering*, volume 6991 of *Lecture Notes in Computer Science*, pages 388–403. Springer, 2011.

M. Elleuch, O. Hasan, S. Tahar, and M. Abid. Formal Probabilistic Analysis of a Wireless Sensor Network for Forest Fire Detection. In *Symbolic Computation in Software Science*, volume 122 of *Electronic Proceedings in Theoretical Computer Science*, pages 1–9. Open Publishing Association, 2013a.

M. Elleuch, O. Hasan, S. Tahar, and M. Abid. Towards the Formal Performance Analysis of Wireless Sensor Networks. In *Proceedings of the Workshop on Enabling Technologies: Infrastructure for Collaborative Enterprises*, pages 365–370. IEEE Computer Society, 2013b.

M. Elleuch, O. Hasan, S. Tahar, and M. Abid. Formal Probabilistic Analysis of Detection Properties in Wireless Sensor Networks. *Formal Aspects of Computing*, 27(1):79–102, 2015.

M. Elleuch, O. Hasan, S. Tahar, and M. Abid. Formal Probabilistic Analysis of a WSN-based Monitoring Framework for IoT Applications. In *Formal Techniques for Safety-Critical Systems*, volume 694 of *Communications in Computer and Information Science*. Springer, 2016a.

M. Elleuch, O. Hasan, S. Tahar, and M. Abid. Formal Probabilistic Analysis of Lifetime for a WSN-based Monitoring Application. In *Proceedings of the International Workshop on Verification and Evaluation of Computer and Communication System*, volume 1689 of *CEUR Workshop Proceedings*, pages 43–58. CEUR-WS.org, 2016b.

E. Fanourgakis. Modelling and Verification of QoS properties of a Biomedical Wireless Sensor Network. Project Work, University of Hamburg-Harbug, 2012.

A. Fehnker, L. V. Hoesel, and A. Mader. Modelling and Verification of the LMAC Protocol for Wireless Sensor Networks. In *Integrated Formal Methods*, volume 4591 of *Lecture Notes in Computer Science*, pages 253–272. Springer, 2007.

W. Feller. *An Introduction to Probability Theory and its Applications*, volume 1. John Wiley & Sons, 1968.

M. Fruth. 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*, pages 290–297. IEEE Computer Society, 2006.

M. Gordon and T. Melham. *Introduction to HOL: A Theorem Proving Environment for Higher-order Logic*. Cambridge Univ. Press, 1993.

A. Gupta. Formal Hardware Verification Methods: a Survey. *Formal Methods in System Design*, 1(2-3):151–238, 1992.

Y. Hanna, H. Rajan, and W. Zhang. Slede: a Domain-specific Verification Framework for Sensor Network Security Protocol Implementations. In *Proceedings of the Conference on Wireless Network Security*, pages 109–118. ACM, 2008.

O. Hasan. *Formal Probabilistic Analysis using Theorem Proving*. PhD thesis, Concordia Univ., Montreal, QC, Canada, 2008.

O. Hasan and S. Tahar. *Formalized Probability Theory and Applications using Theorem Proving*. IGI Global, 2015.

F. Heidarian, J. Schmaltz, and F. Vaandrager. Analysis of a Clock Synchronization Protocol for Wireless Sensor Networks. *Theoretical Computer Sciences*, 413(1):87–105, 2012.

M. Hewish. Reformatting Fighter Tactics. Jane's International Defense Review, 2001.

HOL4. *The HOL theorem prover*, 2013. http://hol.sourceforge.net/.

C. Hsin and M. Liu. Network coverage using low duty-cycled sensors: Random & coordinated sleep algorithms. In *Proceedings of the Symposium on Information Processing in Sensor Networks*, pages 433–442. ACM, 2004.

C. Liu. Randomized Scheduling Algorithm for Wireless Sensor Networks. In Project Report of Randomized Algorithm, University of Victoria, B.C., Canada, 2004.

C. Liu, K. Wu, and V. King. 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*, volume 3462 of *Lecture Notes in Computer Science*, pages 956–967. Springer, 2005.

C. Liu, K. Wu, Y. Xiao, and B. Sun. Random Coverage with Guaranteed Connectivity: Joint Scheduling for Wireless Sensor Networks. *IEEE Transactions on Parallel and Distributed Systems*, 17(6):562–575, 2006.

S. Liu, P. Ölveczky, and J. Meseguer. A Framework for Mobile Ad hoc Networks in Real-Time Maude. In *Rewriting Logic and Its Applications*, volume 8663 of *Lecture Notes in Computer Science*, pages 162–177. Springer, 2014.

S. Liu, P. Ölveczky, and J. Meseguer. Formal Analysis of Leader Election in MANETs using Real-Time Maude. In *Software, Services, and Systems*, volume 8950 of *Lecture Notes in Computer Science*, pages 231–252. Springer, 2015.

D. MacKay. Introduction to Monte Carlo Methods. In *Proceedings of NATO Advanced Study Institute on Learning in Graphical Models*, pages 175–204. Kluwer Academic Publishers, 1998.

Q. Mamun. A Coverage-Based Scheduling Algorithm for WSNs. *International Journal of Wireless Information Networks*, 21(1):48–57, 2014.

T. Mhamdi. *Information-Theoretic Analysis using Theorem Proving*. PhD thesis, Concordia Univ., Montreal, QC, Canada, December 2012.

T. Mhamdi, O. Hasan, and S. Tahar. Formalization of Entropy Measures in HOL. In *Interactive Theorem Proving*, volume 6898 of *Lecture Notes in Computer Science*, pages 233–248. Springer, 2011.

A. Olteanu, Y. Xiao, K. Wu, and X. Du. Weaving a Proper net to Catch Large Objects in Wireless Sensor Networks. *IEEE Transactions on Wireless Communications*, 9(4): 1360–1369, 2010.

P. Ölveczky and S. Thorvaldsen. Formal Modeling and Analysis of the OGDC Wireless Sensor Network Algorithm in Real-time Maude. In *Formal Methods for Open Object-based Distributed Systems*, volume 4468 of *Lecture Notes in Computer Science*, pages 122–140. Springer, 2007.

PRISM. *The PRISM Model checker*, 2013. http://www.prismmodelchecker.org/.

M. Qasim, O. Hasan, M. Elleuch, and S. Tahar. Formalization of Normal Random Variables in HOL. In *Intelligent Computer Mathematics*, volume 9791 of *Lecture Notes in Computer Science*, pages 44–59. Springer, 2016.

RTMaude. *The Real-Time tool*, 2013. http://heim.ifi.uio.no/peterol/RealTimeMaude/.

J. Rutten, M. Kwaiatkowska, G. Normal, and D. Parker. *Mathematical Techniques for Analyzing Concurrent and Probabilisitc Systems*. CRM Monograph Series. American Mathematical Society, 2004.

Z. Sun, P. Wang, M. Vuran, A. Al-Rodhaan, A. Al-Dhelaan, and I. Akyildiz. BorderSense: Border Patrol through Advanced Wireless Sensor Networks. *Ad Hoc Networks*, 9(3): 468–477, 2011.

D. Tian and N. Georganas. A Coverage-preserving Node Scheduling Scheme for Large Wireless Sensor Networks. In *Proceedings of the International Workshop on Wireless Sensor Networks and Applications*, pages 32–41. ACM, 2002.

S. Tschirner, L. Xuedong, and W. Yi. Model-based Validation of QoS Properties of Biomedical Sensor Networks. In *Proceedings of the International Conference on Embedded Software*, pages 69–78. ACM, 2008.

J. Wang, D. Fang, Z. Yang, H. Jiang, X. Chen, T. Xing, and L. Cai. E-HIPA: An Energy-Efficient Framework for High-Precision Multi-Target-Adaptive Device-Free Localization. *IEEE Trans. Mob. Comput.*, 16(3):716–729, 2017.

L. Wang and Y. Xiao. A Survey of Energy-efficient Scheduling Mechanisms in Sensor Networks. *Mobile Networks and Applications*, 11(5):723–740, 2006.

K. Wu, Y. Gao, F. Li, and Y. Xiao. Lightweight Deployment-Aware Scheduling for Wireless Sensor Networks. *Mobile Networks and Applications*, 10(6):837–852, 2005.

F. Xia. QoS Challenges and Opportunities in Wireless Sensor/Actuator Networks. *Sensors*, 8(2):1099–1110, 2008.

Y. Xiao, H. Chen, K. Wu, C. Liu, and B. Sun. Maximizing Network Lifetime under QoS Constraints in Wireless Sensor Networks. In *Proceeding of the Global Telecommunications Conference*, pages 1–5. IEEE Computer Society, 2006.

Y. Xiao, H. Chen, Y. Zhang, X. Du, B. Sun, and K. Wu. Intrusion Objects with Shapes under Randomized Scheduling Algorithm in Sensor Networks. In *Proceedings of International Conference on Distributed Computing Systems Workshops*, pages 315–320. IEEE, 2008a.

Y. Xiao, H. Chen, Y. Zhang, X. Du, B. Sun, and K. Wu. Three Dimensional Intrusion Objects Detection under Randomized Scheduling Algorithm in Sensor Networks. In *Proceedings of International Conference on Mobile Ad-hoc and Sensor Networks*, pages 16–22. IEEE, 2008b.

Y. Xiao, Y. Zhang, M. Peng, H. Chen, X. Du, B. Sun, and K. Wu. Two and Three-dimensional Intrusion Object Detection under Randomized Scheduling Algorithms in Sensor Networks. *Computer Networks*, 53(14):2458–2475, 2009.

Y. Xiao, H. Chen, K. Wu, B. Sun, Y. Zhang, X. Sun, and C. Liu. Coverage and Detection of a Randomized Scheduling Algorithm in Wireless Sensor Networks. *IEEE Transactions on Computers*, 59(4):507–521, 2010.

G. Xu, W. Shen, and X. Wang. Applications of Wireless Sensor Networks in Marine Environment Monitoring: A Survey. *Sensors*, 14(9):16932–16954, 2014.

J. Yick, B. Mukherjee, and D. Ghosal. Wireless Sensor Network Survey. *Computer Networks*, 52(12):2292–2330, 2008.

H. Zayani, K. Barkaoui, and R. B. Ayed. Probabilistic Verification and Evaluation of Backoff Procedure of the WSN ECo-MAC Protocol. *International Journal of Wireless & Mobile Networks*, 12(1):156–170, 2010.

M. Zheng, J. Sun, Y. Liu, J. Dong, and Y. Gu. Towards a Model Checker for NesC and Wireless Sensor Networks. In *Formal Methods and Software Engineering*, volume 6991 of *Lecture Notes in Computer Science*, pages 372–387. Springer, 2011.

Appendices

Tables

HOL Symbol	**Standard**	**Meaning**
$\wedge$	and	Logical and
SUC n	$n+1$	Successor of a num
count n	$\{m \mid m < n\}$	Set of all m strictly less than n
PREIMAGE $f\ s$	$\{x \mid f\ x \in s\}$	The inverse image of the subset s
$\{x \mid P(x)\}$	$\{\lambda x.P(x)\}$	Set of all x that satisfy the cond. P
$x\ pow\ n$	x^n	$real\ x$ raised to num power n
$exp\ x$	e^x	Exponential log. on x
SIGMA $f\ s$	$\sum_s f$	Sum of the sequence $f(x)$; $x \in s$
$lim(\lambda n.f\ n)$	$\lim_{n\to\infty} f(n)$	Limit of the $real$ sequence f

Table 1 HOL Symbols.

Lemma	**Formulation**
DD is an increasing function of k	`mono_incr (`DD`)`
Limit of DD when k is very large	$\frac{(\mathtt{Q}-1+\mathtt{s})(\mathtt{Q}^2-1+\mathtt{s})}{2\mathtt{Q}(\mathtt{Q}+1)}\left[1-(1-\mathtt{q})^{\mathtt{n}}\right]$
Pd is a decreasing function of k	`mono_decr (`Pd`)`
Pd definitely decreases when k is very large	$\lim_{k\to+\infty} Pd = 0$
C_n is a decreasing function of k	`mono_decr (`Cn`)`

Table 2 Required Lemmas for the Different Sets.

Figures

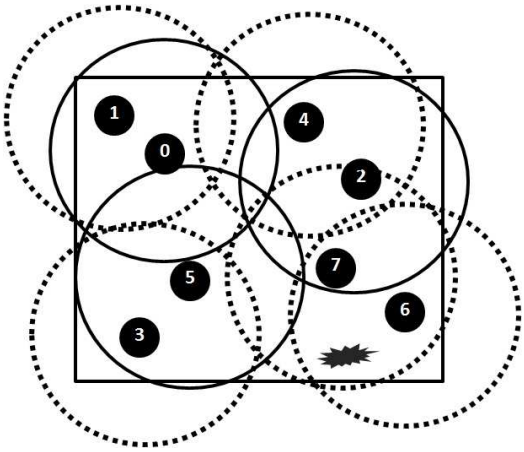

Figure 1 The k-set randomized scheduling for ($n = 8$) nodes and ($k = 2$) subsets.

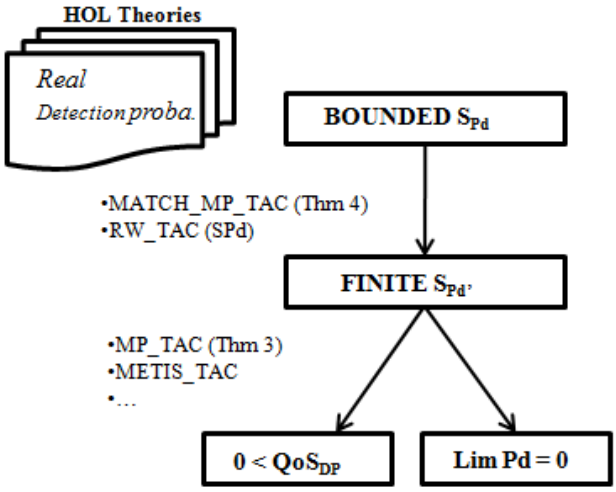

Figure 2 A Simplified Proof Sketch for Lemma 6.

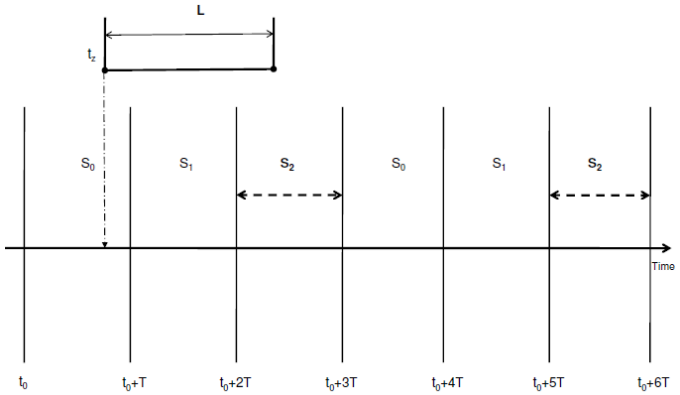

Figure 3 An example of the k-set randomized scheduling for n nodes and $k = 3$.

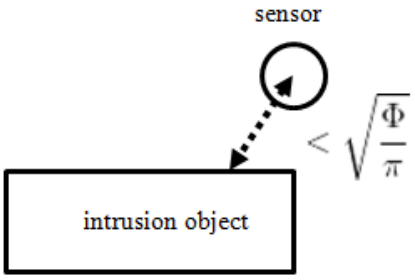

Figure 4 A Rectangular Intrusion Object.

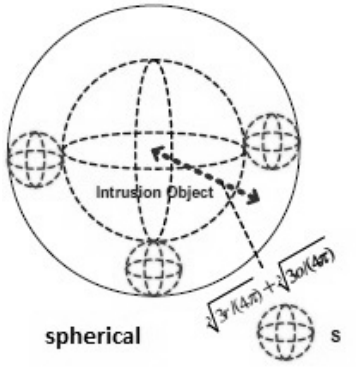

Figure 5 A Spherical Intrusion Object (Xiao et al., 2009).

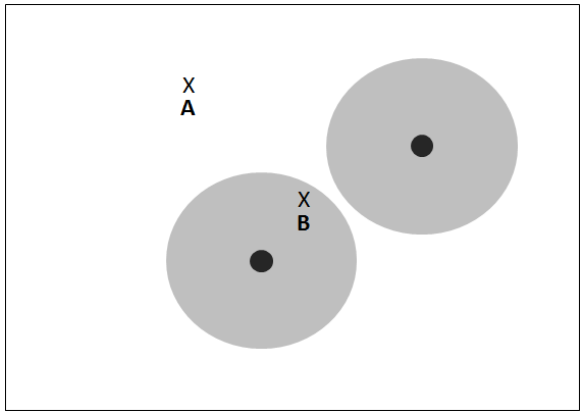

Figure 6 Coverage Example.