Evaluating Complex MAC Protocols for Sensor Networks with APMC

Michaël Cadilhac\textsuperscript{1}  Thomas Hérault\textsuperscript{2}  Richard Lassaigne\textsuperscript{3}  Sylvain Peyronnet\textsuperscript{1,3}  Sébastien Tixeuil\textsuperscript{2}

Abstract

In this paper we present an analysis of a MAC (Medium Access Control) protocol for wireless sensor networks. The purpose of this protocol is to manage wireless media access by constructing a Time Division Media Access (TDMA) schedule. APMC (Approximate Probabilistic Model Checker) is a tool that uses approximation-based verification techniques in order to analyse the behavior of complex probabilistic systems. Using APMC, we approximately computed the probabilities of several properties of the MAC protocol being studied, thus giving some insights about its performance.

Keywords: Wireless sensor networks, approximate verification

1 Introduction

Wireless sensor networks are networks using a large number of machines that embed processors, sensors, actuators and radio communication capabilities. These networks are usually used for environment monitoring and asset tracking. Individual nodes are usually deployed in an \textit{ad hoc} manner, so they must organize themselves to form a multi-hop wireless communication network \cite{967000}.

A major issue in such self-organized wireless networks is the access to the communication medium, as simultaneous wireless transmissions between neighboring nodes result in collisions that garble exchanged messages. Collision management and avoidance are fundamental issues in wireless network protocols, and medium access control (MAC) protocols are distributed solutions to this problem.

In \cite{410972}, Busch, Magdon-Ismail, Sivrikaya and Yener propose a MAC protocol for sensor networks. This protocol provides many interesting properties: it gives...
guarantees about the bandwidth that is allocated to each node, it is fully distributed,
and does not require the existence of a global clock.

To analyze the correctness and performance of such distributed protocols, a lot
of methods can be used. One of them is probabilistic model checking, a class of
algorithmic methods for the verification of probabilistic systems with respect to
quantitative properties. Most of these methods are based on the construction of
a mathematical model of the system and on the expression of the specification in
some temporal language. This model represents all the possible configurations of
the system, and the probabilities of the transitions that can occur between these
states.

The problem of evaluating the satisfaction probability of a temporal property to
be checked, is reduced to the resolution of a system of linear equations over the
state space. However, due to the state space explosion phenomenon during the
modeling step, the representation of the transition matrix can be so large that
the verification becomes intractable. To overcome this phenomenon, symbolic and
numerical methods have been introduced in tools such as PRISM [14]. In the
last years a completely different model checking technique emerged: Approximate
Probabilistic Model Checking. Using this technique we can approximately compute
the probability that a model satisfies a specification [10]. With this method, the
computation time is not necessarily lowered, but the memory consumption becomes
very low (or constant in some cases). Indeed the space complexity of the method is
independent of the size of the model.

The results we present in this paper are twofold: we model a complex contention-
free MAC protocol for wireless sensor networks [4], and we perform various experi-
ments with this model. We thus show the interest of using approximate probabilistic
model checking for the verification and analysis of protocols for sensor networks.

The structure of the paper is as follows. In Section 3 we review the framework of
approximate probabilistic verification and present the tool APMC. Then we briefly
describe the MAC protocol of [4] (Section 4) and its modeling (Section 5). Finally,
Section 6 gives results of several experiments performed on our model using APMC.

2 Related Work

Networks now being imagined for sensors [26] and small devices [5] require en-
ergy conservation, scalability, tolerance to transient faults, and adaptivity to topol-
ogy change. There essentially exists two kinds of MAC algorithms for sensor net-
works [24]:

(i) contention-based protocols: nodes compete for a shared channel, resulting in
probabilistic coordination, e.g. ALOHA [2] and carrier sense multiple access
(CSMA [13]). The CSMA protocol has been widely studied and extended, and
is at the core of widely used standards such as IEEE 802.11 [1].
(ii) scheduled protocols: nodes collaborate to plan communications so that no col-
lisions occur. This plan can be based on frequency (FDMA, frequency division
multiple access), code (CDMA) or time (TDMA), that are widely used in mod-
ern cellular communication systems [20]. Essentially, collisions are avoided by scheduling communications in virtual sub-channels that are obtained by dividing the real channels either by time, frequency, or orthogonal codes. Since virtual sub-channels do not interfere with each other, MAC protocols in this group are essentially collision-free.

Time Division Media Access (TDMA) is a reasonable technique for managing wireless media access, however the priorities of scalability and fault tolerance are not emphasized by most previous research. Recent analysis [9] of radio transmission characteristics typical of sensor networks shows that TDMA may not always substantially improve bandwidth when compared to randomized collision avoidance protocols, however fairness and energy conservation considerations remain important motivations. In applications with predictable communication patterns, a sensor may even power down the radio receiver during TDMA slots where no messages are expected; such timed approaches to power management are typical of the sensor regime. Among TDMA protocols especially designed for sensor networks, [11,12] were studied both theoretically and practically (to get quantitative measures), while the rather complex protocol presented in [4] was only studied by hand, in particular a quantitative analysis was not provided.

The research in the field of methods for approximating probabilistic model checking is quite young and there is only a few approaches other than ours. In [25], a procedure is described for verifying properties of discrete event systems based on Monte-Carlo simulation and statistical hypothesis testing. In [22], a statistical method is proposed for model checking of black-box probabilistic systems. These approaches differ strongly from ours by using statistical hypothesis testing instead of randomized approximation schemes. More recently, in [8], a randomized algorithm for probabilistic model checking of safety properties expressed as $LTL$ formulas was given. In another approach [19] both random testing and abstract interpretation are used for the verification of C programs. Only a few attempts have been made at verifying communications protocols using approximation-based model checking. For instance, APMC was already used for this purpose in [7,6]. A lot of work has been done on the formal verification of communication protocols using probabilistic model checking. One can mention the work of the PRISM team (see, for instance, [15,16]).

3 Approximate Probabilistic Verification and APMC

In this section, we recall the framework of approximate probabilistic verification and present quickly the tool APMC (Approximate Probabilistic Model Checker). APMC is based on a randomised algorithm to approximate the satisfaction probability of a temporal specification, by using sampling of execution paths of the system.
3.1 Models and specifications

The approximation method of APMC can handle any probabilistic system that supports the generation of execution paths. The input language is the same as PRISM and allows to describe in a modular way either discrete-time or continuous-time Markov chains. The specification language can express satisfaction probabilities of temporal logic formulas.

Let $M$ be a discrete-time Markov chain, $s$ be an initial state, and $\psi$ be a linear temporal logic formula. We denote by $Path(s)$ the set of execution paths whose first state is $s$. The probability measure $Prob$ over the set $Path(s)$ is defined classically and we denote by $Prob[\psi]$ the measure of the set of paths satisfying the formula $\psi$. Let $Path_k(s)$ be the set of all paths of length $k > 0$ starting at $s$ and $Prob_k[\psi]$ be the measure of paths satisfying $\psi$ in $Path_k(s)$.

Using the approximation method of APMC, we can approximate, with any degree of accuracy, the satisfaction probability of a temporal formula. In the next subsection, for the sake of clarity, we consider only bounded temporal properties (that is properties of finite paths) and we describe a randomised algorithm to approximate $Prob_k[\psi]$.

3.2 Approximate verification

In order to estimate the probability $p$ of a bounded property $\psi$ with a randomised algorithm, we generate random paths of $Path_k(s)$ and compute a random variable $X$ which estimates $p = Prob_k[\psi]$. We say that an estimation $X$ is $\varepsilon$-good means if the output value of the algorithm is in $[p - \varepsilon, p + \varepsilon]$.

**Definition 3.1** A randomised approximation scheme (RAS) for $p$ is a randomised algorithm $A$ that takes as input a representation of the system, a property $\psi$, two real numbers $\varepsilon, \delta > 0$ and produces a value $X$ such that: $Pr(X \in [p - \varepsilon, p + \varepsilon]) \geq 1 - \delta$. If the running time of $A$ is polynomial in $k, \frac{1}{\varepsilon}$ and $\log(\frac{1}{\delta})$, $A$ is said to be fully polynomial.

The probability $Pr$ is taken over the random choices of the algorithm. The approximation is $\varepsilon$-good with confidence $(1 - \delta)$ after a number of samples polynomial in $\frac{1}{\varepsilon}$ and $\log(\frac{1}{\delta})$. The main advantage is that, in order to design a path generator, we only need to simulate the behaviour of the system using a succinct representation of it (called the diagram). The generic approximation algorithm of APMC is the following.

<table>
<thead>
<tr>
<th>Generic approximation algorithm $GA$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Input:</strong></td>
</tr>
<tr>
<td><strong>Output:</strong></td>
</tr>
<tr>
<td>$N := \ln(\frac{2}{\delta})/2\varepsilon^2$ ; $A := 0$</td>
</tr>
<tr>
<td>For $i = 1$ to $N$ do $A := A + \text{Random Path}(\text{diagram}, k, \psi)$</td>
</tr>
<tr>
<td>Return $Y = A/N$</td>
</tr>
</tbody>
</table>

where the function Random Path is:
### Random Path

**Input:** diagram, k, ψ

**Output:** samples a path π of length k and check formula ψ on π

(i) Generate a random path π of length k (with the diagram)

(ii) If ψ is true on π then return 1 else 0

In [17], it is proven that $G\AA$ is a fully polynomial RAS for computing $p = \text{Prob}_k[\psi]$ whenever $\psi$ is a bounded temporal property and $0 < p < 1$. The randomised algorithm of APMC allows to approximate in polynomial time, and with very high confidence, satisfaction probabilities of temporal properties. The main advantage is to eliminate space complexity by using path generation and efficiently bounding sample size. Moreover this approach is highly parallelizable and APMC uses a distributed computation model to distribute path generation and formula verification on a cluster of workstations.

### 3.3 Architecture and Implementation

APMC includes two independent components: the compiler and the deployer. The APMC compiler takes the model description written with the PRISM language (a variant of Reactive Modules), and a list of temporal properties to check on this model. It produces an ad-hoc verifier for this set of properties over the given model. The output of the compiler is in fact a set of functions in ANSI C suitable for verifying the properties on the model. This file lacks a main function and an engine to produce the verification.

Providing the engine and the missing functions for the ad-hoc verifier is the goal of the deployer. It produces a stand-alone binary which takes only three parameters: the approximation parameter, the confidence and the path length. It then runs the simulation and outputs the approximated probabilities for each of the temporal formulas. Thus, the deployer provides the working program suitable for a distributed verification inside a LAN. This distributed deployment strategy runs in parallel these components on all the participating nodes and provides the same result with a linear acceleration.

### 4 Sketch of the protocol

In [4], a probabilistic distributed algorithm for constructing a TDMA schedule is presented. For each sensor, the time is divided into frames (that need not be of same length at each node), that in turn are divided into slots (whose size is for simplicity considered as the same for each sensor). Essentially, the algorithm has two layers:

(i) the first layer, **LooseMAC** constructs a TDMA schedule where every sensor in a neighborhood is able to communicate with no conflicts; however, some bandwidth may be wasted in this process.

(ii) the second layer, **TightMAC** allocates the remaining slots so that the wasted
bandwidth is actually used.

In order to cope with sensor removal or arrival, those two layers are alternately used.

The **LooseMAC** layer can informally be described as follows. Each node considers that frames are divided into slots of equal size. Then, until a node has not found a free slot, it randomly chooses a slot and emits a message in this slot. If no collision is detected (because a neighbor emitted a message at the same time) or reported (because a neighbor received two messages at the same time, and is experiencing the so-called hidden terminal effect), the time slot is selected by the node. The protocol is illustrated in Figure 1. During Frame A, nodes randomly choose slots. The second and third nodes notice that a collision occurs, while the first node found a collision-free slot. In Frame C, the second and third nodes randomly choose a new slot. The second node knows that the first node chose the third slot (because it is a direct neighbor of this node), so it is able to choose a non conflicting slot. However, the third node chooses a slot that conflicts with the first one. The second node is then able to report a conflict. In Frame E, the third node is eventually able to pick a free slot, and the **LooseMAC** algorithm terminates for all nodes.

The **TightMAC** layer can informally be described as follows. Each node repeatedly uses the **LooseMAC** mechanism until all nodes in the neighborhood at distance 3 have successfully executed **LooseMAC**. Then, nodes compute how many nodes exists in their neighborhood at distance two, that is, the number $\phi$ of nodes they will compete with for obtaining TDMA slots. They then randomly choose slots in the remaining free slots, but no more than $1/\phi$ overall. If no collisions are detected or reported, the additional TDMA slots are gained.

Additional details about the two layers are provided in [4].

## 5 Modeling

In this section, we describe the modeling of the **LooseMAC** and **TightMAC** protocols using the Reactive Module Language.

### 5.1 LooseMAC

Each sensor is modelled by an independent module. A sensor can be in one of three modes:

- **NEWSLOT**, if it has changed its randomly chosen time slot in the current frame,
- **WATCH**, after a whole frame passed since the node in the **NEWSLOT** mode chose
its time slot,

- READY, if no conflict has occurred in a whole frame and the node was in the WATCH state.

When every node has reached the READY state, the system is stable, i.e. the slots chosen are conflict-free. A proof of this protocol is given in [4].

The LooseMAC protocol has been modelled with particular attention paid to keep the original algorithm apparent. In this perspective, its three main subparts as described in the appendixes of [4] are clearly distinguished. For a node $i$, we have the following:

- **Send()**: broadcasts the state of the node $i$ in its current time slot. The state sent consists of the identifier $i$ of the node, its freshness flag, and if it has detected any conflict in the previous phase. The node actually broadcasts if it is in the NEWSLOT mode or if it has detected a conflict.

- **Receive()**: at every time slot, the node checks for conflicts or for the arrival of new nodes. A conflict is detected if the node received noise, i.e. more than one message, if some message is received in $i$’s time slot or if $i$ already received a message in this time slot before. The node stores any correspondence seen between node identifiers and time slots.

- **UpdateMode()**: updates the mode of the node $i$ in its current time slot, according to the mode switches presented before.

The behavior of the main function of this TDMA protocol is as follows:

<table>
<thead>
<tr>
<th>LooseMAC</th>
</tr>
</thead>
<tbody>
<tr>
<td>(i) Initialize some internal values</td>
</tr>
<tr>
<td>(ii) (a) Call sequentially Send(), Receive() and UpdateMode(),</td>
</tr>
<tr>
<td>(b) Increment a local time reference,</td>
</tr>
<tr>
<td>(c) Loop to (a).</td>
</tr>
</tbody>
</table>

The protocol assumes that every node increments at *regular intervals* a local time reference, interval in which we are sure every possible action are made. To model this, we define a global time reference. All nodes use this time reference to wait for the completion of a time frame and increment their local time reference. This is not a limitation since the difference between absolute and relative slot position would only have resulted in a translation of the index of the slots.

Reactive Modules is a reactive language, so the sequentiality of the operations has to be simulated. We use the usual transformation to implement the sequentiality of events through a *state* which describes the current progress of each module. This state represents the execution of each algorithm presented before: a node is in a specific state when it is executing one of them. Also, each algorithm has its own sub states. The equivalence between the execution of LooseMAC and the internal state of a node $i$ is as follows:
The last state, namely \textbf{Ended}, indicates that the node has finished the loop iteration. In the original algorithm, it is the moment when the local time reference is incremented. In our modeling, it corresponds to a \textit{synchronization} between all the processes, so that the global time reference is correctly updated.

As an example of a sub state in an algorithm, \texttt{UpdateMode()} uses a \texttt{DoUpdate-mode} state in which the newly computed mode, which has been stored in \texttt{NewMode}, is assigned to the actual mode of the node. This avoids clashes when the mode is tested within \texttt{UpdateMode()}.  

Since it is assumed that we know a maximum time interval for a communication to complete, they are made quasi-synchronously. When a node $i$ sends a message to another one, $i$ stores directly its data in the variables of its recipient and increments a counter of received message in it. Moreover, a reading is made if and only if all sending are made: this helps knowing if a node actually received \textit{noise}, that is to say, more than one message, or a single message.

In theory, all nodes should execute the same algorithms: the code made so far is expected to be seen only once in the source files. However, the broadcasting part of \texttt{Send()} needs a node-dependent information, that is known at compile-time: its neighbors.

As a consequence, each node module is unique and has to be generated according to a graph file. XRM\textsuperscript{4}, a Reactive Modules preprocessor, has been used to simplify their writing.

\begin{figure}[h]
\centering
\begin{tabular}{|c|c|}
\hline
\textbf{Internal State} & \textbf{LooseMAC Main Algorithm} \\
\hline
\begin{enumerate}
\item \textbf{NotIn} \hspace{0.5cm} // Node is not in the network \\
\item \textbf{Init} \\
\hspace{1cm} \texttt{Mode} $\leftarrow$ \texttt{NEWSLOT} \\
\hspace{1cm} \texttt{$\sigma$} $\leftarrow$ random slot \\
\hspace{1cm} \texttt{while true do} \\
\hspace{2cm} \texttt{Send()} \\
\hspace{2cm} \texttt{Receive()} \\
\hspace{2cm} \texttt{UpdateMode()} \\
\hspace{1cm} \texttt{end while} \\
\end{enumerate}
\end{tabular}
\caption{Node's internal state and LooseMAC algorithm}
\end{figure}

\textsuperscript{4} XRM: eXtended Reactive Modules, \url{http://www.lrde.epita.fr}.
5.2 **TightMac**

TightMac and LooseMAC execute themselves in parallel. As a result, the code for TightMac is an addendum of the previous one. TightMac simply uses the slot found by LooseMAC to compute another conflict-free slot on a smaller frame using the knowledge of the local neighbourhood.

As a simplification of the process, the necessary calculations are made at compile time. This is made possible thanks to the global knowledge of the graph at this moment.

6 **Experiments**

We performed an analysis of the model using APMC. All experiments were run using a set of heterogenous workstations ranging from a simple ATHLON 2 Ghz 512 MB to a bi-Xeon 3.2Ghz 4 GB. We set the approximation parameter $\varepsilon = 10^{-2}$ and the confidence parameter $\delta = 10^{-5}$. We study two topologies: peer-to-peer communication over a generated dense graph, and peer-to-peer communication over a sparse graph. We give results for both topologies. The computation time needed for all experiments is around 76 hours.

6.1 **Checked properties**

In this subsection, we describe the properties we verified. The first three properties were checked on LooseMAC, and the last one on TightMac. For each property, we set the length of the path to 32000, which corresponds to 1000 time unit.

**Experiment 1. Contention-free from the initial state.**

The first experiment was to verify the correctness of the protocol. That is to check that eventually every nodes will become ready, meaning that each node is given a unique slot for communicating, thus avoiding conflicts. More precisely, our goal with this experiment is to check (experimentally) the validity of the lemma 1 of the extended version of [4].

**Experiment 2. A fresh node breaks the stability momentary.**

The algorithm LooseMAC is supposed to handle the case of a node joining or leaving dynamically the network. When a node leave the network, the algorithm is trivially robust. The problem is when a node join the network: potentially there could be a conflict. LooseMAC is designed to be self-stabilizing by forcing some nodes to become non-ready. We verified this assumption by computing the probability that, when a new node join the network, some nodes become non-ready (this is the lemma 2 in the extended version of [4]).

**Experiment 3. Contention-freeness after a node joining the network.**

The last experiment we did on LooseMAC was to check wether all nodes become ready after a new arrival, thus ensuring that the algorithm is self-stabilizing.

**Experiment 4. Stability of the network**

We computed the probability for a node to successfully choose a conflict-free time slot when running TightMac. This ensure that the network state is stable...
6.2 Experimental results

Since the properties we checked on the protocol have been proven by hand in [4], the experimental results are not surprising.

Results for LooseMAC

Figures 3 and 4 present the experimental results for experiments 1 to 3 over sparse and dense networks for two frame sizes (32 and 64 respectively). The network contains 10 nodes.

Experiment 2 shows that the probability that the system detects the apparition of a node is one. It means that, when a new sensor enters the network, at least one of its neighbours acknowledge this by selecting the state NON-READY. This result holds for all topologies, and whatever the frame size. This is a direct consequence of the protocol, since the new process will send a beacon message when entering the network. All the neighbour processes will be aware of it either by noting that one of the slots that was not used before is now used, or that there is a collision on the slot randomly chosen by the new process.

Experiment 1 measures the probability that the system converge (all nodes are in the READY state) according to the time. This curve exhibits a logarithmic convergence time, quickly reaching the probability one after a warmup time. Comparing figures 3(a) and 3(b), or figure 4(a) and 4(b), one can see that the average convergence time is quicker for sparse graphs than dense graphs. This is due to the fact that the probability of conflict is lower with a small neighborhood.

Comparing figures 3(a) and 4(a), or figures 3(b) and 4(b), one can see that with a larger frame size, the convergence time is slower. This is not obvious, since a larger frame size induces a larger space to place the slots, thus a higher probability that a node chooses an empty slot. However, since the protocol works in phases and each node waits for the end of the frame before testing the collision status, a larger frame size implies a slower convergence time. This illustrates a tradeoff on this frame size. It should be short enough to provide a good response time, but large enough to ensure a good probability of finding an empty slot.

Experiment 3 measures the probability that the system converges after a new arrival. One can see that the convergence time exhibits the same behavior as from an initial configuration where none of the nodes are in the READY state, but with a longer warmup phase. The warmup phase is longer since the arrival of the new node can imply a modification of the slots of all the nodes of the network. Moreover, the failure detection first has to propagate to all the nodes.

Results for TightMac

Figure 5 presents the result of experiment 4 over a sparse network with a LooseMAC framesize of 32, as in figures 3(b). One can see that the convergence time of the TightMac algorithm follows the same behavior as the LooseMAC algorithm. This is natural since the TightMac algorithm can converge only when the LooseMAC algorithm did.
In this paper, we presented an analysis, using approximate probabilistic model checking, of the contention-free MAC protocol of [4]. We showed that this method allows to efficiently verify/analyze the correctness and performance of complex distributed algorithms over sensor networks. This method does not suffer from the state space explosion phenomenon arising with classical model checking methods. However the numerical results we give are accurate only with respect to an approximation parameter ($\varepsilon$ here). Moreover, in order to model efficiently the protocol we were required to add some constraints (such as a global timer) to make the model tractable for APMC.
Fig. 4. Experimental results - LooseMAC

Acknowledgement

We thank Benoit Sigoure for providing us pre-processing tools for Reactives Modules.

References


Fig. 5. Experimental results - TightMac


