On the Design of Automatic Link Establishment in High Frequency Networks

Bruno Baynat, Hicham Khalife, Vania Conan, Catherine Lamy-Bergot, Romain Pouvez

International Journal of Networking and Computing – www.ijnc.org
ISSN 2185-2839 (print) ISSN 2185-2847 (online)
Volume 7, Number 2, pages 419–446, July 2017

On the Design of Automatic Link Establishment in High Frequency Networks

Bruno Baynat
Université Pierre et Marie Curie
Paris, France

Hicham Khalife, Vania Conan, Catherine Lamy-Bergot and Romain Pouvez
Thales Communications Security
Paris, France

Received: December 16, 2016
Revised: April 9, 2017
Revised: June 8, 2017
Accepted: June 21, 2017
Communicated by Jacir L. Bordim

Abstract

Most High Frequency (HF) communications systems deployed on the field today implement Automatic Link Establishment (ALE) techniques in order to help the HF stations automatically set up a link with good properties. Two generations (so called 2G and 3G ALE) have been standardized since the 90's, and are today being revisited due to the emergence of wideband HF waveforms. In this paper, we develop Markovian models of the 2G ALE procedure, which is nowadays the most widely used as it can operate while being completely asynchronous. Our models are "channel oriented", i.e., they observe the system from channel occupation perspective regardless of node status. We show by comparison with high-level OMNET++ simulations that our models provide fast and accurate estimation of all performance parameters of interest, and capture the main characteristics of the ALE process and the interactions between their numerous parameters. We believe that our work constitutes a useful tool to help operator plan and dimension HF networks. We also exploit the model to give some insight on the limitations of current 2G ALE, helping the design of future ALE strategies.

1 Introduction

HF radio communications have long been the only solution for wireless communications beyond line-of-sight (BLOS) with none or minimal infrastructure. Supporting communications over several thousands of kilometers, HF propagation channel is however highly variable and error-prone. It tends to make HF communications unreliable and difficult to establish, especially when used with none or poor knowledge of the propagation conditions. Their BLOS capability explains their importance in military communications.

ALE solutions have been historically developed to provide automation and ease of use to end-users having less time and less skills to operate HF communication systems. In practice, ALE-able radios are given a pre-defined and shared set of frequencies that are scanned for incoming calls by all stations in the same network when they are idle. Any station wanting to establish a link determines

the first frequency it should use, based on a Link Quality Analysis (LQA) mechanism. The radio then begins its call and tries to establish a link with another radio. Obviously, this automatic (without human intervention) selection of a frequency between a caller and a called radio should be done as quickly as possible to allow the end-user to place his call and provide information (whether voice or data) to its correspondent.

Two generations of ALE standards, MIL-STD-188-141A [21] denoted as the ALE 2G and more recently the STANAG 4538 [15] known as ALE 3G co-exist on the field. Several comparisons between these two standards exist today in the literature [23, 12]. Their main findings is that 3G outperforms the 2G ALE in dense collision prone networks, however the two protocols provide close results in large unidirectional networks, for instance BRASS-type naval scenarios. In addition, one main advantage of 2G ALE over 3G ALE is that it can operate while being completely asynchronous. For all these reasons, it is still nowadays the most widely used and *de-facto* interoperability standard [21].

Enhancing the performance of ALE mechanisms remains a tough challenge to overcome. Inspired by the recent progress in wireless networking and communication domains, several intiatives to improve the efficiency of existing standards and sometimes propose completely new solutions started to arise. These proposals rely on two different paradigms: i) exploring cognition instigated by the recent progress in the cognitive radio domain in order to learn and optimize selecting/exploiting the existing channels for communications [17], or ii) investigating wideband transmissions for higher throughputs hence new applications [6]. Nevertheless, to the best of our knowledge, except few simulation studies [23], no existing work have tried to mathematically model the ALE standards.

In this paper, we develop Markovian models of the ALE 2G procedure, based on our prior work on HF modeling [16]. The proposed models are "channel oriented", i.e., observe the system from channel occupation perspective regardless of node status. We first propose a Continuous Time Markov Chain (CTMC) model that, at the cost of some simplifying assumptions that are clearly enumerated and discussed, provides all the performance parameters of interest, e.g., the ALE duration, and the success and failure probabilities of the procedure. Then, in order to reduce the complexity of this first detailed CTMC, we developed an aggregated CTMC that enables to drastically reduce the number of states of the chain and thus the complexity, without introducing a significant error in most cases. In order to validate our models, we have conducted high-level OMNET++ simulations in which many of the assumptions made for the models are reproduced, e.g., Markovian assumptions and probabilistic success or failure for transmissions. Nevertheless, these simulations show that by reproducing a per node behavior of the ALE procedure, we can quite closely reproduce models that looks at the system from a channel perspective regardless of the state of nodes. Our work allows to investigate the network performance for different traffic loads, number of channels, and communication durations. More generally, our models enable the analysis of the complex interplay between the different ALE parameters and their impact on the system behavior, and provide a way to help operators plan and dimension HF 2G deployments.

The remainder of the paper is structured as follows. Section 2 discusses the similarities of the HL ALE with other communications systems and presents the related state of the art. In Section 3, we precisely describe the ALE 2G system and the assumption we make in order to develop the so called "detailed" Markovian model presented in Section 4. This model is validated in Section 5. Section 6 presents an "aggregated" model that aims at reducing the complexity of the detailed model, as well as some other extensions. In Section 7, we explore how the model can be exploited in order to dimension and configure HF networks. Finally Section 8 concludes this paper.

2 State of the Art

Several research papers have focused on modeling multi-channel (multi-frequency) wireless networks such as cognitive radio networks. Even though the HF ALE was not specifically studied, some resemblance to our work can be noted.

2.1 Modeling cognitive radio networks

In [19] authors propose a dynamic channel-selection solution for autonomous wireless users transmitting delay-sensitive multimedia applications in cognitive radio networks. To do so, the radio interface evaluates the expected delays experienced by various priority traffics using priority queuing analysis that considers the wireless environment, traffic characteristics, and the competing users behaviors in the same frequency channel. However, this work assumes information exchange between wireless nodes. Such assumption cannot be made in HF networks. A Preemptive Resume priority M/G/1 queue to model the total service time for various target channels sequences based on the activity pattern and interactions between the primary and secondary users is presented in [26]. Based on this model the optimal channel sequence can be derived. In fact, this work is very interesting for selecting the best channels sequence for a communication in order to reduce the service time. Nevertheless, two main issue may render such model not applicable on ALE procedures. First, when it comes to voice communications, potential interruptions when switching from one frequency to another cannot be tolerated. Second, designing a signaling protocol to dynamically exchange the channel sequence between the sender and the receiver constitutes a great challenge. Also in the area of cognitive radios, [11] investigates the performance of the secondary network in the case when channels are opportunistically available for secondary users using different channel bonding or aggregation strategies, and no spectral handover is implemented. Continuous Time Markov Chain (CTMC) models are used to model the spectrum occupation in this study. More precisely, three channel strategies are explored where the simplest one corresponds to no assembling, i.e., traditional multi-channel network. However, as in all cognitive radio networks, two categories of users with different priorities are considered. This assumption deeply changes the models compared to the flat same priority users.

2.2 MAC layer protocols for cognitive radios

New MAC layer protocols for cognitive radio networks were proposed in [22]. These MAC protocols enable the secondary users to identify and utilize the leftover frequency spectrum in a way that constrains the level of interference to the primary users. More interestingly, this work proposes a Markov chain model and a $M/G^Y/1$-based queuing model to characterize the performance of the proposed multi-channel MAC protocols under the two types of channel-sensing policies, for the saturation network and the non-saturation network scenarios, respectively. Indeed, these two sensing strategies (a random and a collaborative policies) are based on the information obtained from a dedicated control channel. Although the random policy is completely distributed and presents some similarities with HF systems, the existence of the control channel makes transmission over data channel contention free. Such assumption cannot be made in today's HF networks. A novel three dimensional discrete-time Markov chain to characterize the process of spectrum handoffs and analyze the performance of unlicensed users was proposed in [20]. Since in real cognitive radio networks, a dedicated common control channel is not practical, the model implements a network coordination scheme where no dedicated common control channel is needed. Nevertheless, this work assumes that all nodes are synchronized and follow the same frequency hopping scheme. This is not the case in ALE 2G standard.

2.3 Rendezvous schemes for cognitive radios

In cognitive radio networks, rendezvous schemes present some resemblance to the HF ALE mechanism. Rendezvous is a distributed algorithm implemented on individual cognitive radios helping them to dynamically select the same channel for communication. Most popular techniques are based on channel hopping schemes somehow synchronized [5] [24] [7] or without assuming time synchronization between radios [14] [1]. Indeed, finding a rendezvous is obviously very similar to the ALE procedure that strives to find a common channel for communication between a sender and a receiver. However, the lack of synchronization in the ALE 2G as well as the absence of primary users that have higher priorities in channel access constitute important differences. In fact, in cognitive radio transmitting long enough in order to allow the receiver to round robin across all existing channels

cannot be tolerated. Clearly the presence of primary users whose appearance can force cognitive (secondary) radio to vacante renders such solution not applicable. For this reason the rendezvous techniques proposed for cognitive radio try to reduce the handshake period which is not the case of the ALE 2G standard. This observation constitutes a major difference between the ALE and any other multichannel solution and should be properly taken into account in the proposed Markovian models. Nevertheless, to the best of our knowledge, no modeling iniatives of rendezvous took place in the literature.

2.4 Other modeling initiatives

Besides, in recent cellular networks, bandwidth is partitioned into tens to hundreds of parallel channels, each of which can be allocated to a possibly different user in each time slot. It is the case for 4G systems such as LTE where each base-station employs an OFDM (Orthogonal Frequency Division Multiplexing) based slotted-time air-interface at the base-station. In this context, [3] compares various scheduling algorithms for these small buffer multi-channel systems. Using queuing theory and Markov chains, this works highlights that a class of iterative algorithms (iLQF ? iterated Longest Queues First) are rate-function optimal in the many-channels regime. Nevertheless, in these cellular networks, perfect synchronization is logically considered.

Prior to these initiatives, researchers have tried to model multi-channel Slotted Aloha and IEEE 802.11-like protocols [18]. Their assumptions do not hold with the ALE procedure in a HF environment. Tzamaloukas *et al.* proposed RICH-DP, a receiver based MAC protocol in [25]. Their solution, based on a receiver-initiated collision-avoidance handshake, does not require carrier sensing or the assignment of unique codes to nodes in order to ensure collision-free reception. A common frequency hopping to all receivers allows a perfect synchronization between nodes. They have solved analytically their proposal (with the help of Markov chains) and through simulations. However, the perfect synchronization assumptions makes the considered model differ significantly from the ALE 2G model studied in this paper. The multi-channel MAC protocol (MCMAC) [13] extends the IEEE 802.11 MAC to use multiple physical-layer channels. The protocol uses a single control channel, and multiple data channels whereby each data transfer has a control phase and a data exchange phase. The performance of this protocol was analytically derived in [8]. In fact, authors present an analytical framework for evaluating multi-channel MAC protocols using M/G/1 queue. To model the dynamics of the protocol and to obtain the performance measures, they apply Stochastic Reward Net (SRN) modeling technique that is an extension of stochastic Petri nets (SPN). Clearly, using a control channel renders such model unsuitable in our context. More generally, authors in [27] propose a book that draws the performance analysis of many multi-channel access protocols. The HF ALE does not fall into the scope of their study. Rendezvous based MAC protocol were also proposed in the context of classical multi-channel ad hoc networks [4]. In the CQM protocol proposed here, similarly to most rendezvous schemes proposed for cognitive radios, a strict time synchronization is assumed between nodes. This assumption cannot be made with the ALE 2G standard.

Finally, many initiatives have modeled the IEEE 802.11 standard and more precisely its Distributed Coordinated Function (DCF) mode in single cell and multihop networks [2, 9, 10]. However, this protocol exploits a single channel for communication thus differs in essence from the considered multi-channel case.

3 System Description

In the whole paper we focus on ALE 2G, which is nowadays, as aforementioned, the most widely used standard. The ALE 2G standard defines both the physical layer as well as the access technique of the MAC layer for the participating stations also referred to as ALE technique. At the physical layer, the ALE 2G employs an 8-ary FSK (Frequency Shift Keying) modulation whereby each symbol is coded over 3 bits. Every symbol has an 8 ms duration what yields 125 symbols per seconds. The hence obtained data rate equals $3 \times 125 = 375$ bits/s. We focus below on the access technique description that constitutes the MAC layer mechanism of the standard.

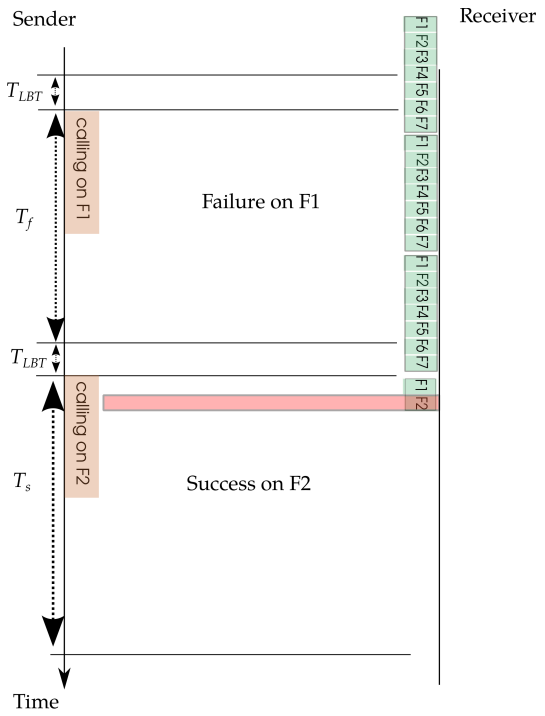


Figure 1: ALE 2G general concept time diagram

The main advantage of the ALE 2G standards stems from its ability to operate while being completely asynchronous. In other words, at time t, a node can be listening or transmitting on any existing channel without any information on the status of the other nodes. More precisely, the source node sends a call request on a channel for a time duration long enough to enable the receiver to scan all available channel during the emitter transmission. Therefore the size of call request frame depends on the number of available channels for communication in the system. If the receiver is able to detect the call (failure can be due to channels conditions at the receiver side), a handshake is undergone that leads to the establishment of the call. If no answer is received for a call request, the sender moves to the next available channel and initiates a call request on this channel (if any). Figure 1 exemplifies the 2G ALE procedure.

3.1 ALE 2G access mechanism

We consider a HF network composed of M nodes. These nodes can exploit a set of N channels for communication and reception. In the ALE 2G, a node selects a single channel i corresponding to a mid-band frequency f_i, $i \in \{0, 1, ..., N\}$, for transmitting or receiving. In fact, as illustrated in Figure 2, a node can be in one of the four following states:

- **Listening state.** A node that has nothing to transmit and that is not receiving, listens continuously on the N available channels. Listening is done sequentially by sensing a channel for a short period of time before moving to the next band. The sensing period is set in such a way as to detect transmissions over this particular band. A station leaves the listening state in two cases. First, if while scanning a particular channel, it detects a transmission corresponding to a call request with its own address as destination, in which case the station moves to the called state. Second, if it receives internally (from higher layers) a call request towards another participating node, in which case the station moves to the calling state.

- **Calling state.** When a node needs to initiate a call it enters in the calling state and follows a procedure to identify a channel on which to communicate with the receiver. For that purpose

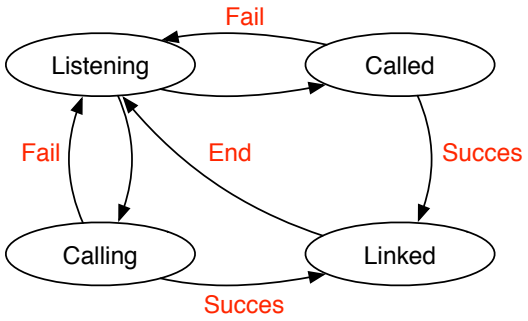


Figure 2: Node state diagram

the caller tests all N channels in sequence: first it checks if the channel f_1 is available: it performs a carrier sensing process for a time denoted as T_{LBT} ("LBT" stands for "Listen Before Talk"). If successful this means that channel f_1 is free (on the caller's side) and then it sends a call initiation request message on that channel containing the receiver's address. The main part of the request message is repeated for a significantly long time so as to allow any free receiver node to scan all of the N channels during this transmission. If the receiver receives the request with its own address, it accepts the request and sends back an acknowledgement. In this case the successful handshake lasts for a time denoted as T_s. If the receiver is busy or if propagation conditions on that channel are bad, the request message is not answered by the caller. The handshake is considered as a failure after a timeout denoted as T_f. The sender then repeats the procedure sequentially on all N channels.

- **Called state.** A node leaves the listening state to the called state after detecting (through sensing) that he is the intended receiver of a sent call request frame. A called station then replies to the caller and awaits for the confirmation from the latter as in classical 3-way handshake procedures. In case the handshake is not completed successfully, the call request is aborted forcing the caller to find another available channel for making its call and pushing the receiver back to the listening state. Note here that many factors can provoke the failure of this handshake. The impact of ionospheric propagation, the long distance attenuation over the HF bands in addition to the receiver unfavorable state/conditions are the most common causes.

- **Linked state.** Following a successful 3-way handshake exchange (i.e., the end of the ALE), both sender and receiver enter the linked state. Nodes remain in this state for the communication duration.

In summary, an idle node that wants to establish a new communication with a destination node, first listens to channel 1 during a time T_{LBT}. If it senses a communication (of any kind) on the corresponding frequency band, it moves to channel 2, and so on, until it finds a free channel (if the N channels are busy, the call request is dropped). The source node then starts a handshake procedure by sending an establishment request frame on the found free channel. If it receives a positive answer from the destination node, the communication between the two nodes starts on the chosen channel. The whole handshake on this channel lasts for a time T_s. If the source node receives no comprehensible answer from the destination node, it considers the handshake as a failure after a timeout T_f. It then tries to establish the communication on another frequency, by sensing the remaining channels one by one. The procedure as well as relative durations are shown in Figure 1

3.2 Frame structure and impact on call requests

As described earlier, after an LBT, the call request is repeated for a duration long enough in order to allow the receiver to scan all available channels and hook on the call during this transmission.

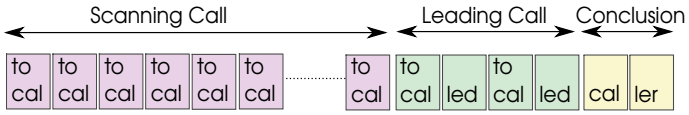


Figure 3: ALE 2G call request frame structure

One may wonder how a receiver node that did not start receiving the call request from its beginning will be able to decode correctly the received information. The answer to this question is in fact in the frame structure (and its length) defined for the HF ALE standard and depicted in Figure 3. In practice, the call request has a specific frame structure that allows the destination, by repeating multiple times the same information, to correctly start receiving the sent frame. As shown in Figure 3, this frame is constituted by three separate parts. The first part called *Scanning Call* enables the receivers to synchronize by starting to receive at the beginning of the request. Indeed, this part incorporates a repetition of the first 3 characters of the destination address. Note here that stations in the ALE standard are identified by a string of characters of variable length such as "GEORGE" for instance. By repeating the first 3 characters of the destination (here GEO) multiple times for a period long enough to allow scanning stations to loop across all available channels, all stations which address start by GEO will stop scanning and hook on this frequency waiting for the following parts of the frame. Note that the duration of scanning call is computed in order to allow all scanning stations to run through all available frequencies at least one time. After this step, the *Leading Call* contains the remaining parts of the destination address repeated twice for safety reasons. This allows the stations that matches the first 3 characters to decide whether they are the destination of this call request. After this step only the station which address is "GEORGE" is still locked on this frequency; all other resume listening iteratively on the system available channels. The *Conclusion* part of the frame structure defines the caller ID thus allowing the receiver to answer back the sender and finalize the call initiation handshake.

It is worth saying here that the *Leading Call* and *Conclusion* parts are of a variable size that depends on the number of characters of the destination and the source addresses. Therefore call requests have variable size lengths (and durations) what allows to tie break after a collision between requests initiated in the same time without the need for a backoff. In fact, after a collision one transmitter will move to the next frequency before the other node implicated in the collision. Consequently, the latter when moving the following frequency and undergoing an LBT will find this band occupied and will naturally move to the next available one.

More precisely, each *Scanning Call* block (i.e. first 3 characters of the destination address) has a duration of 784 ms. A scanning call has a total duration of $N \times 784$ where N is the total number of available channels for the ALE procedure. At the receiver side, a scanning rate for a station in listening mode is limited to 500 ms per frequency. In order to loop over all the available frequency bands, a receiver requires in the worst case $500 \times N$ ms. This duration is shorter than $784 \times N$ required to transmit accross all the available channels what ensures a transmission long enough for the receiver to catch the beginning of a scanning call. As for the *Leading call*, it is constituted of the remaining characters of the called address as well as the caller address what forms a variable length block size. Similarly, the *Conclusion* part has a variable lenth that depends on the source address number of characters. In a rough calculation, if the caller and the called addresses are 6 characters length and for N of 10 (10 available channels), the call frame duration is about 15 seconds. However, due to the multiple repetition of the same information, if the called address length is 8 characters, the frame duration reaches 18 seconds. Note that this computation does not include the length of the acknowledgment that also includes addresses and has variable duration. In brief, the variable frame length/duration coupled to the low data rate reduces dramatically the collisions between transmitting nodes without the need for a backoff mechanism.

3.3 Assumptions on the system

Equipped with a single transceiver capable to operate on a single channel, a node can manage only one communication at a time. More precisely, a station in the called state, calling state or linked state, is unable to detect another transmissions addressed to it. These call requests are therefore rejected.

Assumption 1 *A node already in communication cannot respond to a call request coming from another node.*

Besides, a calling station may receive locally other call requests that can either be buffered or dropped (depending on the underlying application). Here we assume that these calls are also discarded. This assumption can be typically made for voice traffic, however extensions to buffered data call requests will be considered as a future work.

Assumption 2 *A new call request arriving locally on a node that is already communicating is lost.*

As a consequence, in our studied system, any call request arriving at a node either immediately triggers an ALE procedure or is dropped.

Let us finally remind the system assumption related to the sensing order of frequencies:

Assumption 3 *All nodes sense the channels sequentially in the same order, from* 1 *up to* N*.*

4 Detailed Model

4.1 State description

The model we propose is "channel oriented". This means that it describes the evolution of the state of the N channels without structurally including the state of the M nodes ("listening", "calling", "called" and "linked"). In other words, the model will explicitly describe neither the identity or the state of a node that has initiated a communication, nor the identity or the state of the node to which the communication is addressed. As a result, the model will be useful to derive the performance parameters of the channels (e.g., frequencies occupation) and of the ALE procedure (e.g., probability of success or average ALE duration), but will not directly address the nodes performance (e.g., call attempts rejected because a node is already communicating). However, we will see that some of these nodes performance parameters can be derived from the channels performance parameters.

The considered state of the system is thus a vector $\vec{n}$ of N components, each one corresponding to a given channel i, $i \in \{0, 1, ..., N\}$, and in which each component can take three values:

- idle: simply denoted as f_i and meaning that there is currently no communication or call attempt on channel i;
- used for a call attempt: denoted as $\hat{f}_i$ and meaning that there is a node currently trying to establish a communication with another node on channel i (i.e., there is an ongoing 3-way handshake on channel i that has not yet lead to a success or to a failure);
- used for a communication: denoted as $\bar{f}_i$ and meaning that there is an ongoing communication between two nodes on channel i.

From this state description we derive the state diagram illustrated in Figure 4. Note that the number of states of this diagram is 3^N. On the figure we represent the transitions out of a particular state $(\hat{f}_1, \bar{f}_2, f_3, \bar{f}_4, \hat{f}_5)$ of a system made of $N = 5$ channels. In this state, channel 3 is idle, channels 2 and 4 are occupied by a communication (between two nodes that are not specified), and channels 1 and 5 are used by nodes that are currently making a 3-way handshake on these two frequencies (again the two calling nodes and the two called nodes are not specified in the state description). From this state, different events may occur. First, one of the two ongoing communications may terminate leading to one of the two upper states, $(\hat{f}_1, f_2, f_3, \bar{f}_4, \hat{f}_5)$ (if the communication on f_2 ends

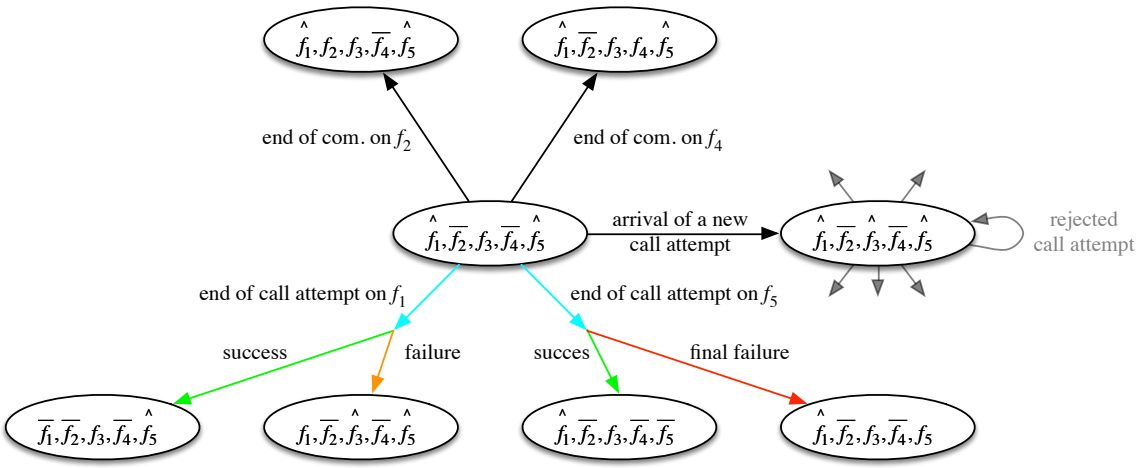


Figure 4: Channel state diagram

first) or $(\hat{f}_1, \bar{f}_2, f_3, f_4, \hat{f}_5)$ (if the communication on f_4 ends first). Then, a new call attempt may arrive on one idle node, leading to right state $(\hat{f}_1, \bar{f}_2, \hat{f}_3, \bar{f}_4, \hat{f}_5)$ where the new calling node tries to establish a communication on the only idle frequency he has found, f_3. Third, the call attempt on frequency f_1 may end either because the corresponding 3-way handshake has lead to a success, in which case a new communication begins on frequency f_1, leading to state $(\bar{f}_1, \bar{f}_2, f_3, \bar{f}_4, \hat{f}_5)$, or because the handshake has failed, in which case the call attempt is placed on the next idle frequency, f_3, leading to state $(f_1, \bar{f}_2, \hat{f}_3, \bar{f}_4, \hat{f}_5)$. Finally, the call attempt on frequency f_5 may end either because the corresponding handshake has been successful, leading to state $(\hat{f}_1, \bar{f}_2, f_3, \bar{f}_4, \bar{f}_5)$, or because the handshake has failed, in which case the call attempt is definitely rejected, leading to state $(f_1, \bar{f}_2, \hat{f}_3, \bar{f}_4, f_5)$.

It is very important to emphasize that this state diagram implicitly relies on the assumption that the "Listen Before Talk" time, T_{LBT}, is negligible with regard to other times involved in the modeling. Indeed, if we do not consider this time as negligible, we cannot consider anymore that the arrival of a new call attempt directly makes the system move from state $(\hat{f}_1, \bar{f}_2, f_3, \bar{f}_4, \hat{f}_5)$ to state $(\hat{f}_1, \bar{f}_2, \hat{f}_3, \bar{f}_4, \hat{f}_5)$, and must introduce intermediate states where the source node listen on channel 1, then listen on channel 2, and finally listen on channel 3 and finds it idle.

Assumption 4 *The "Listen Before Talk" time T_{LBT} is neglected.*

An extension of the model taking into account a non negligible LBT time is presented in Section 6.3.

4.2 Markovian model

As previously mentioned, any new call request arriving on a node that is not idle is rejected (Assumption 2). In the model we will assume that arrival of new requests on idle nodes can be modeled as a Poissonian process.

Assumption 5 *The global arrival process of new call requests in the whole system (on any of idle nodes) is assumed to be a Poisson process with rate λ.*

The poissonian hypothesis is a reasonable assumption as soon as the arrival process of new requests result of the superposition of many independent renewal processes on each node.

Next, we assume that the communication time between two nodes can be modeled by an exponential distribution.

Assumption 6 *The communication time is assumed to be exponential with rate μ.*

This is a very classical assumption, that we have no reason not to make without any further specifications on the system behavior.

One important modeling assumption concerns the probability that a handshake between a given source node and its destination node on a given frequency results in a success or in a failure. Indeed two factors are necessary in order for this handshake procedure to succeed. First, as said before, the success is very likely related to the propagation conditions and to the distance between nodes. Second, a success is also conditioned by the fact that the destination node is idle. These two events can reasonably be considered as independent, and the probability of both occurring is thus the product of the probabilities of each of them taken individually. If the first one can be characterized by a fixed probability that is independent of the state of the system (e.g., estimated from simulations), the second one should depend on the load of the system. However, when the number M of communicating nodes is high with regards to the number N of frequencies, which is the most likely scenario, we can assume that the success of the handshake procedure can be characterized by a constant (and state-independent) probability.

Assumption 7 *A 3-way handshake between a source node and a destination node (on any free channel) has a probability p_s to succeed and to result in a communication between the two nodes, and a probability $p_f = 1 - p_s$ to fail and to force the source node to find another free channel to establish the communication.*

With all these assumptions, the state diagram depicted in Figure 4 can directly be transformed into a Continuous-Time Markov Chain (CTMC) illustrated in Figure 5. The rates of the transitions from state $(\hat{f}_1, \bar{f}_2, f_3, \bar{f}_4, \hat{f}_5)$ to one of the four lower states include the inverse of the average time until a handshake ends (by either a success or a failure), $\frac{1}{p_s T_s + p_f T_f}$, multiplied by the corresponding probabilities p_s or p_f. The upper transitions correspond to the end of a communication (either on f_2 or on f_4) and have thus an associated rate of μ, and the right transition corresponds to a call request arrival (on an idle node) and has thus an associated rate of λ.

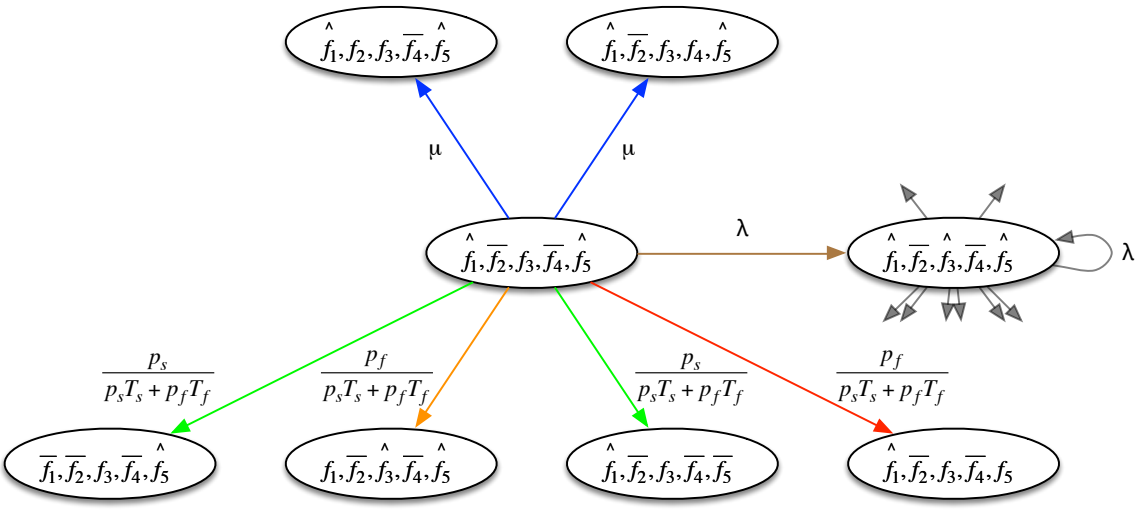


Figure 5: Markovian model

The CTMC has 3^N states and can be solved using any appropriate numerical technique (such as the Gauss-Seidel technique). Note however that the number of state increases very rapidly with the number N of channels. This is the reason for the derivation of the so-called "aggregated model" presented in Section 6.1.

4.3 Performance parameters

The steady-state solution of the CTMC provides the stationary probabilities $p(\vec{n})$ of all states $\vec{n}$ of the chain. We can derive from these probabilities all the performance parameters of interest as follows.

First, we define $ni(\vec{n})$ as the number of idle channels in a given state $\vec{n}$, $nc(\vec{n})$ as the number of channels used for a communication, and $nh(\vec{n})$ as the number of channels used for a handshake. Obviously, for any state $\vec{n}$ of the chain, $ni(\vec{n}) + nc(\vec{n}) + nh(\vec{n}) = N$ at any time. As an example, for the vector $\vec{n} = (\hat{f}_1, \bar{f}_2, f_3, \bar{f}_4, \hat{f}_5)$, we have $ni(\vec{n}) = 1$, $nc(\vec{n}) = 2$ and $nh(\vec{n}) = 2$.

An arriving call request can eventually result in three events:

1. The call request can be rejected if it arrives when there is currently no idle channel. This is illustrated on Figure 5 for the right state $(\hat{f}_1, \bar{f}_2, \hat{f}_3, \bar{f}_4, \hat{f}_5)$.

2. The call request can eventually result in a success if the source node manage to place a successful handshake on a free channel. This event corresponds to the crossing of a "green transition" in the CTMC illustrated in Figure 5.

3. The call request can eventually result in a failure if the source node does not manage to receive a comprehensible answer from its destination on all tested channels. This event corresponds to the crossing of a "red transition" in the CTMC.

We then define X_r, the average number all call requests rejected by unit of time, X_s the average number of call requests leading to a success (meaning to a communication) by unit of time, and X_f the average number of call requests leading to a failure by unit of time. These throughputs can be estimated as follows. First, X_r is just the number of "loops" crossed by unit of time in the CTMC (as the one illustrated on state $(\hat{f}_1, \bar{f}_2, \hat{f}_3, \bar{f}_4, \hat{f}_5)$ in Figure 5):

$$X_r = \sum_{\vec{n} \,|\, ni(\vec{n})=0} p(\vec{n})\lambda. \tag{1}$$

Then, X_s is the number of crossing of "green transitions" by unit of time:

$$X_s = \sum_{\vec{n}} p(\vec{n}) nh(\vec{n}) \frac{p_s}{p_s T_s + p_f T_f}. \tag{2}$$

In order to derive X_f, we first need to define the function $nr(\vec{n})$ as the number of channels used for a handshake in state $\vec{n}$ that are not followed by an idle channel. As an illustration, in state $\vec{n} = (\hat{f}_1, \bar{f}_2, f_3, \bar{f}_4, \hat{f}_5)$, $\hat{f}_1$ is followed by the idle channel f_3, but $\hat{f}_5$ is not followed by any idle channel. As a result, $nr(\hat{f}_1, \bar{f}_2, f_3, \bar{f}_4, \hat{f}_5) = 1$. In fact, $nr(\vec{n})$ corresponds to the number of "red transitions" out of state $\vec{n}$, a red transition corresponding to a "final failure", as illustrated in Figure 4. The throughput X_f can now be computed as the number of crossing of "red transitions" by unit of time:

$$X_f = \sum_{\vec{n}} p(\vec{n}) nr(\vec{n}) \frac{p_f}{p_s T_s + p_f T_f}. \tag{3}$$

Obviously, the conservation of flows implies that $X_r + X_s + X_f = \lambda$.

From these throughputs, we can now evaluate P_r, the rejection probability of a call request, P_s, the probability that a call request result in a success, and P_f, the probability that a call request result in a failure:

$$P_r = \frac{X_r}{\lambda},\ P_s = \frac{X_s}{\lambda},\ P_f = \frac{X_f}{\lambda}. \tag{4}$$

In order to calculate another performance parameter of interest, namely the average ALE time, we first calculate Q_h, the mean number of channels used for a handshake:

$$Q_h = \sum_{\vec{n}} p(\vec{n}) nh(\vec{n}). \tag{5}$$

We then derive from Little's law the average duration of an ALE procedure:

$$R_{ALE} = \frac{Q_h}{\lambda}. \tag{6}$$

We now turn our attention to determining the average duration of an ALE procedure conditioned by the fact that this establishment procedure is a success, denoted by $R_{ALE \mid s}$, conditioned by the fact that it is a failure, denoted by $R_{ALE \mid f}$, and conditioned by the fact that the call request has been rejected, denoted by $R_{ALE \mid r}$. These three quantities are related to the average ALE duration R_{ALE} thanks to the law of total probability:

$$R_{ALE} = R_{ALE \mid s}\ P_s + R_{ALE \mid f}\ P_f + R_{ALE \mid r}\ P_r. \tag{7}$$

As we assume that the LBT time is negligible ($T_{LBT} = 0$), $R_{ALE \mid r} = 0$. We thus have:

$$R_{ALE} = R_{ALE \mid s}\ P_s + R_{ALE \mid f}\ P_f. \tag{8}$$

Conditioned by the fact that an arriving call request finds n idle channels ($n > 0$), the probability that the ALE procedure is a failure is ${p_f}^n = (1 - p_s)^n$ (as it must fail on all n idle channels), and the corresponding lost time of this ALE procedure (knowing that it results in a failure) is nT_f. Now, because arrivals of requests follow a Poisson process, the PASTA theorem tell us that the probability that an arriving request finds n idle channel is just the stationary probability that there are n idle channel at any time, denoted as $pi(n)$, that can be directly obtained from the stationary probabilities of the Markov chain by summation:

$$pi(n) = \sum_{\vec{n} \mid ni(\vec{n}) = n} p(\vec{n}). \tag{9}$$

We can then estimate the average duration of an ALE procedure conditioned by the fact that the ALE procedure is a failure as:

$$R_{ALE \mid f} = \frac{\sum_{n=1}^{N} pi(n) {p_f}^n n T_f}{\sum_{n=1}^{N} pi(n) {p_f}^n}. \tag{10}$$

It is then straightforward to calculate the average duration of an ALE procedure conditioned by the fact that the ALE procedure is a success, from relation 8:

$$R_{ALE \mid s} = \frac{R_{ALE} - R_{ALE \mid f}\ P_f}{P_s}. \tag{11}$$

Finally we can calculate the average number of free channels, Q_i, as well as the average number of channels used for a communication, Q_c:

$$Q_i = \sum_{\vec{n}} p(\vec{n}) ni(\vec{n}). \tag{12}$$

$$Q_c = \sum_{\vec{n}} p(\vec{n}) nc(\vec{n}). \tag{13}$$

As a channel used for a communication involves exactly two nodes, we can deduce from Q_c the average number of nodes in communication, Q_n:

$$Q_n = 2Q_c. \tag{14}$$

4.4 Asymptotic behavior at low load and high load

We develop in this subsection the asymptotic expressions of the performance parameters of interest in the two extreme cases of a very low load and a very high load.

In the case of a very low load, i.e., when λ tends to zero, the rejection probability P_r obviously tends toward 0 and the probability P_f that a call request results in a failure tends toward ${p_f}^N$. Indeed, when a new call request arrives it has a very high chance to find all channels idle and the only way for the call request to result in a final failure is that it fails on all tested channels. As a consequence, the probability P_s that a call request results in a success tends toward $1 - {p_f}^N$:

$$P_r^{\lambda \to 0} = 0,\ P_f^{\lambda \to 0} = {p_f}^N,\ P_s^{\lambda \to 0} = 1 - {p_f}^N. \tag{15}$$

For similar reasons, the average duration $R_{ALE \mid f}$ of an ALE procedure conditioned by the fact that it is a failure is NT_f:

$$R_{ALE \mid f}^{\lambda \to 0} = NT_f. \tag{16}$$

In order to give the expression of the (unconditioned) average ALE duration R_{ALE}, we again use the fact that a call request has a very high chance to find all N channels idle upon arrival. If the call succeeds on channel 1 (with a probability p_s) the average ALE duration is T_s; if it fails on channel 1 and then succeed on channel 2 (with a probability $p_f p_s$) the average ALE duration is $T_f + T_s$; and so on until the penultimate case where the request fails successively on the first $N-1$ channels and succeed on the last one (with a probability ${p_f}^{N-1} p_s$), that corresponds to an average ALE duration of $(N-1)T_f + T_s$. Finally, the last case corresponds to a final failure of the request (with a probability ${p_f}^N$) and corresponds to an average ALE duration of NT_f. The average ALE duration can thus be expressed as:

$$R_{ALE}^{\lambda \to 0} = \left(\sum_{n=0}^{N-1} {p_f}^n p_s (nT_f + T_s) \right) + {p_f}^N NT_f. \tag{17}$$

We can derive similarly the expression of the average duration $R_{ALE \mid s}$ of an ALE procedure conditioned by the fact that the ALE procedure is a success, by just removing the last listed case and renormalizing the probabilities:

$$R_{ALE \mid s}^{\lambda \to 0} = \sum_{n=0}^{N-1} \frac{{p_f}^n p_s}{1 - {p_f}^N} (nT_f + T_s). \tag{18}$$

In the case of a very high load, i.e., when λ tends to infinity, the rejection probability P_r obviously tends toward 1, and, as a result, both probabilities P_f and P_s tend toward 0:

$$P_r^{\lambda \to \infty} = 1,\ P_f^{\lambda \to \infty} = 0,\ P_s^{\lambda \to \infty} = 0. \tag{19}$$

Now when considering a very high load, conditioned by the fact that an arriving request is not rejected, it has a very high chance to find only one idle channel. As a result the average ALE durations conditioned by the fact that the ALE procedure is either a success or a failure have the following very simple expressions:

$$R_{ALE \mid s}^{\lambda \to \infty} = T_s,\ R_{ALE \mid f}^{\lambda \to \infty} = T_f. \tag{20}$$

And using the law of total probability (8) we get:

$$R_{ALE}^{\lambda \to \infty} = 0. \tag{21}$$

5 Validation of the detailed model

In order to validate our Markovian model, we solve numerically the stationary equations associated with the chain via MALTAB (using the Gauss-Seidel technique), and compare the performance

metrics to OMNeT++ simulations. We have used OMNeT version 4.6 with the objective of exploiting the engine of the discrete event simulator and developed our own physical and MAC layers without using any of the frameworks available for this purpose. We apply to simulations the same assumptions as those used for deriving the Markovian model: Assumption 1 to 7 as labelled in Sections 4.1 and 4.2. Importantly, contrarily to the Markovian model which is channel oriented, simulation describes the evolution of the state of each of the M nodes (as detailed in Section 3.1), with realistic transmission over the N available channels. In particular, in the considered discrete event simulations, the ALE algorithm is played separately by every node based on the observed state of every frequency and its variation.

The physical layer is simulated through the transmission success (respectively failure) probability p_s (respectively p_f) without producing a transmission between the sender and the receiver, i.e., no real channels are simulated. As for the MAC layer, we have implemented the ALE 2G standard whereby available nodes scan continuously a set of available channels and transmitters transmit in a round robin fashion over available frequency identified through their listening mechanism (LBT).

We consider in our validation a HF system made of $M = 40$ nodes communicating through $N = 5$ channels. The mean communication time according to is set to $1/\mu = 13.3$ s, this value was obtained from measured operational HF data where the average HF call duration is around 13 seconds. Moreover, based on the 2G standard [21], we set $T_s = 24$ s and $T_f = 21$ s. To fulfil Assumption 4, we neglect in our simulations the LBT duration ($T_{LBT} = 0$ s). Note that an extension to a non negligible LBT duration is presented in Section 6.3. Following Assumption 7, we also reproduced in simulation the fact that a handshake procedure succeeds on a given free channel with a fixed probability p_s. We have conducted several runs while changing the value of p_s (0.1, 0.5 and 0.9). Performance parameters are computed with a varying load λ in the interval $]0; 1]$ call demands per second. Simulations have a variable length, according to λ, in order to get sufficient data to compute performance parameters. Additionally, in all simulation plots, each point is the average of 10 simulation runs. Since our simulations are on node basis, in most simulations in order to obtain an aggregated results, similar to the model proposed, obtained values (such as duration R_{ALE} for instance) are averaged over all the nodes participating in the simulation then averaged again for the 10 simulation runs.

5.1 Channels occupancy and acceptance rate

Figure 6 shows the occupancy of channels as a function of the load when the success probability $p_s = 0.5$. Recall here that a channel can be in one of the three possible states: idle, in handshake or in communication. The figure compares the average number of channels in each state, Q_i, Q_h and Q_c, derived for the Markovian model (relations 12, 5 and 13), to those obtained from simulation. From this figure one can first observe that our Markovian model matches very accurately the simulations. More precisely, the average relative error between model and simulation is less than 1%, with a maximum error around 3%.

From figure 6 one can easily notice that the number of idle channels drops quickly with the load. Most importantly, most of the busy channels are occupied by handshake procedures while few of them are used for communications. This can be seen as a suboptimal use of available channels. Figure 6 also shows quite intuitively that the higher the value of λ the more difficult the success of an ALE on a free channel.

Figure 7 compares the probabilities of rejection, success and failure, respectively P_r, P_s and P_f, obtained by the model (equation 4) and by simulations, still for $p_s = 0.5$. When the load increases, we can see that the failure rate first increases up to a maximum, then decreases toward zero. In the first phase, the increase of λ implies a raise of the number of calls, so more failures occur. Nevertheless, the more λ grows the more channels are occupied, that translates into a raise of the rejection rate and consequently a decrease of the failure rate. Besides, the success (resp. rejection) rates decrease (resp. increase) with the overall system load. Note that these results are corroborated by the asymptotic behavior developed in Section 4.4 for the rejection probability: $P_r^{\lambda\to 0} = 0$ and $P_r^{\lambda\to\infty} = 1$; for the failure probability: $P_f^{\lambda\to 0} = 0.5^5 = 0.03125$ and $P_f^{\lambda\to\infty} = 0$; and for the success probability: $P_s^{\lambda\to 0} = 1 - 0.5^5 = 0.96875$ and $P_s^{\lambda\to\infty} = 0$. As previously mentioned, model and

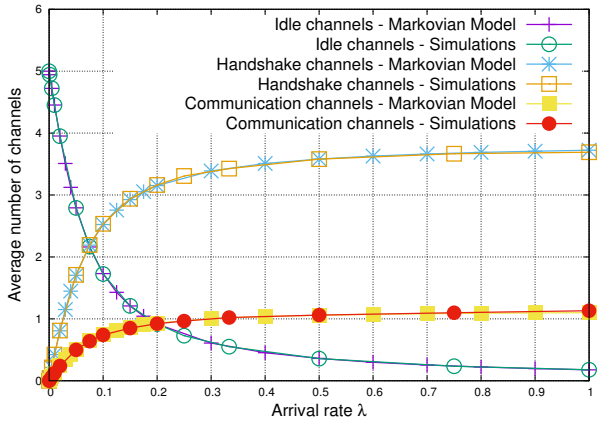


Figure 6: State of channels function of the arrival rate for $p_s = 0.5$

simulations match perfectly, the average relative error between both being less than 2%, with a maximum error around 5% for P_f.

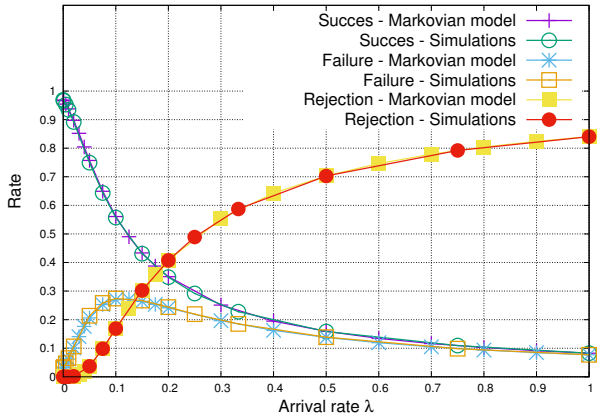
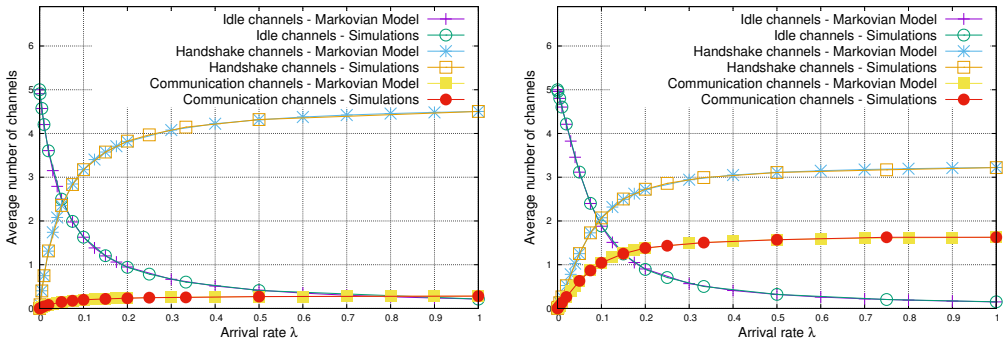


Figure 7: ALE acceptance function of the arrival rate for $p_s = 0.5$

We have also investigated in Figure 8 the channels occupancy for different success probabilities p_s. Unsurprisingly, with a lower success probability ($p_s = 0.1$), it becomes more difficult to establish a communication. Consequently, the high number of failing attempts increases the number of channels occupied by handshake and reduces the number of those exploited for effective communications. In contrast, when $p_s = 0.9$, more communications will be possible on more channels with less handshakes taking places.

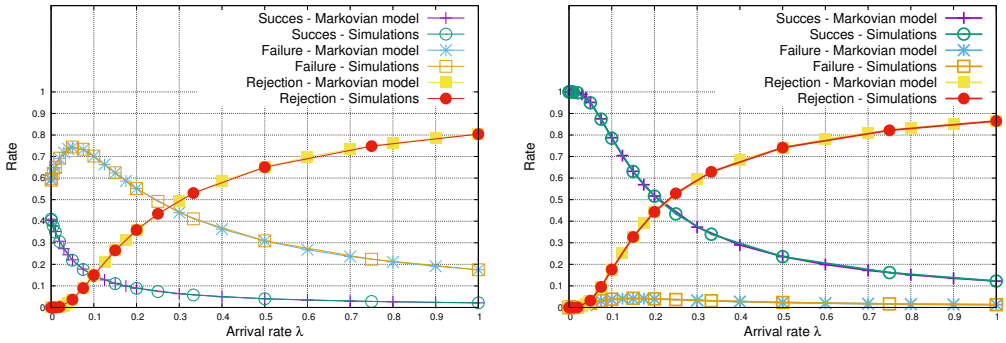
In practice, these results highlight the impact of the handshake on the ALE 2G procedure. Indeed, even when the success probability of a call is very high ($p_s = 0.9$) more handshakes than effective communications are taking place on the available channels. More precisely, no less than 3 over the 5 available channels are used for handshakes with a relatively low arrival rate (starting from $\lambda = 0.2$) . Same remarks can be made on the call success probability given on Figure 9 for different p_s. Quite logically here, higher values of p_s yield lower failures and higher success ratios. More surprising is the rejection rate that grows faster with λ when calls have higher probability to succeed ($p_s = 0.9$). Indeed, such behavior can be explained by the higher number of succeeding calls that occupy faster the available channels pushing new arrivals to be rejected by finding all the frequencies already in use. Let us finally highlight that the perfect match between the results of our model and simulations remains true for different values of p_s.



(a) State of channels function of the arrival rate ($p_s = 0.1$)

(b) State of channels function of the arrival rate ($p_s = 0.9$)

Figure 8: Channels states for different success probabilities



(a) ALE success, failure and rejection rates ($p_s = 0.1$)

(b) ALE success, failure and rejection rates ($p_s = 0.9$)

Figure 9: ALE sucess, failure and rejection for different success probabilites p_s

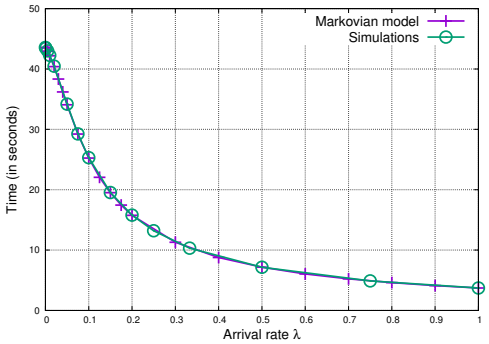
5.2 ALE duration

We investigate here the average ALE duration, R_{ALE}, regardless of the outcome of the procedure (unconditioned), as well as the ALE duration in the case it leads to a success or to a failure, denoted by $R_{ALE \mid s}$ and $R_{ALE \mid f}$ respectively. These quantities are derived from the Markovian model (relations 6, 11 and 10) and compared to simulations. R_{ALE} is depicted in Figure 10(a) that shows that, in average, the duration of an ALE procedure decreases with λ. In fact, the more λ grows, the less channels are available, therefore less channels are tested; at high load, no more channels are available and, as the LBT time is neglected, the call is immediately rejected, thus R_{ALE} decreases towards 0s. Two main observations can be made here. First, these curves confirm again the accuracy of our Markovian model compared to simulations, with less than 10% maximum relative errors. Second, also encouraging, the coherence of the limiting values in Figure 10(a) with the asymptotic behavior computed in Section 4.4, where $R_{ALE}^{\lambda \to 0} = 43.594$s and $R_{ALE}^{\lambda \to \infty} = 0$s.

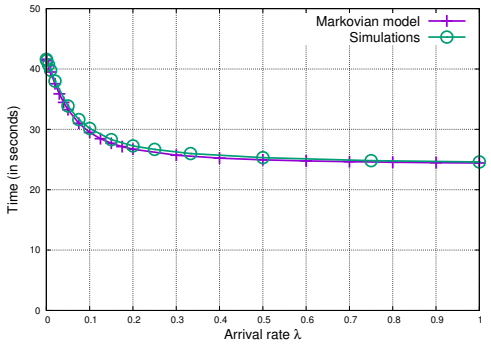
Similarly in Figure 10(b), the average duration $R_{ALE \mid s}$ of an ALE procedure conditioned by the fact that the ALE procedure is a success decreases with λ, however at high load, i.e., when the number of free channels is low, the ALE time converges to the ALE duration over a single available channel since the outcome here is definitely a success. Note that this value corresponds to the asymptotic behavior $R_{ALE \mid s}^{\lambda \to \infty} = 24$ s.

In the case the final result of the ALE procedure is a failure, Figure 10(c) converges to testing a single channel that eventually fails in perfect coherence with $R_{ALE \mid f}^{\lambda \to \infty} = 21$ s.

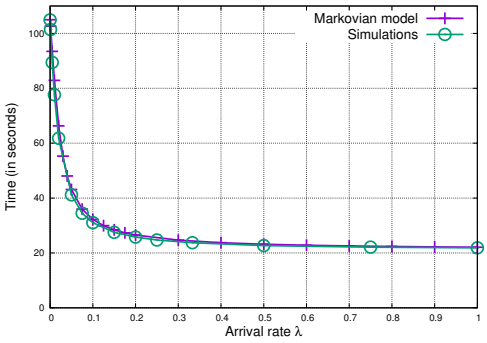
When modifying the success probability p_s the same general tendency is observed. Looking at



(a) R_{ALE} function of the arrival rate



(b) $R_{ALE \mid s}$ function of the arrival rate



(c) $R_{ALE \mid f}$ function of the arrival rate

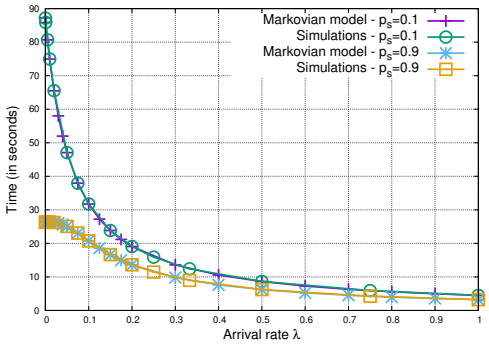
Figure 10: ALE duration for $p_s = 0.5$

Figures 11(a), 11(b) and 11(c), the duration of the ALE process (either unconditioned or conditioned by a success or a failure) is longer if p_s is close to 0. This is due to the fact that the more channels are during the process, the longer is the ALE procedure. Particularly interesting is the case of $R_{ALE \mid s}$ when $p_s = 0.9$ (Figure 11(b)) that highlights that in such situation often only 1 channel is tested regardless of the system load. Indeed, the ALE duration stable around 24 s (the value of T_s) corresponds to testing successfully a single frequency for every arriving call that succeeds independently of the system state.

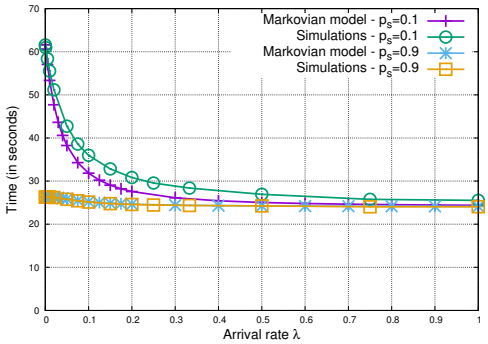
Besides, we can observe in Figure 11(c) that the estimation by the model of $R_{ALE \mid f}$ is less accurate for $p_s = 0.9$ (than the one calculated when $p_s = 0.1$ or 0.5). The relative error between analytical results and simulation reach a maximum of 35% around a load $\lambda = 0.05$. Similarly the estimation of $R_{ALE \mid s}$ is less accurate for $p_s = 0.1$ (see Figure 11(b)), with a relative error close to 13%. This highlights the fact that relation 10 introduces an additional approximation with regard to all other performance parameters, inherent to the fact that this relation implicitly assumes that the state of the channels remains unchanged during the whole ALE procedure of an arriving request, which is actually not the case. Note however, that this approximation has no effect on low load and high load situations (both asymptotes of the model match simulation), and affects only the radius of curvature of the curves.

6 Aggregated model and extensions

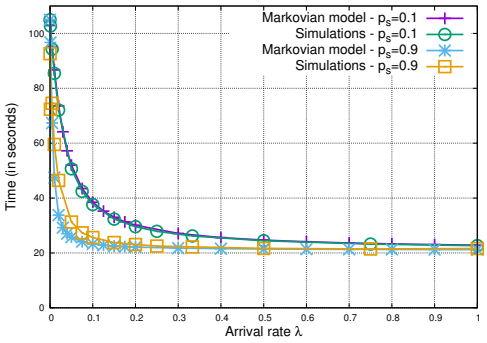
As mentioned earlier, the number of states of the Markov chain associated with the model developed in Section 4.2 increases very rapidly with the number N of channels. In order to overcome this problem we present in this section an approximate aggregated Markov chain model.



(a) R_{ALE} function of the arrival rate



(b) $R_{ALE \mid s}$ function of the arrival rate



(c) $R_{ALE \mid f}$ function of the arrival rate

Figure 11: ALE duration for different success probabilities p_s

6.1 Description of the aggregated model

A state $\vec{m}$ of this aggregated model is obtained from the detailed state description of the previous model by only considering the number of channels in one of the three possible states, idle, used for a communication, used for a handshake. $\vec{m}$ is thus a vector of three components (n_i, n_c, n_h), where n_i is the number of idle channels, n_c is the number of channels used for a communication, and n_h is the number of channels used for a handshake. Of course any possible vector $\vec{m}$ is such that $n_i + n_c + n_h = N$. As a result, the number of states is thus reduced from 3^N in the previous detailed model to $C_{N+2}^N = \frac{(N+2)(N+1)}{2}$ (the number of ways to place N objects in 3 boxes) in this aggregated model. As an illustration, for $N = 10$, the number of states is drastically reduced from 59.049 to 66.

The Markov chain associated with the aggregated model is illustrated in Figure 12 that represents the output transition of a generic state $\vec{m} = (n_i, n_c, n_h)$ such that $n_i \geq 1$. The colors of the transitions of this aggregated Markov chain match the colors of the transitions of the original detailed Markov chain. Brown corresponds to the arrival of a new call request, blue to the end of a communication, green to the success of an ALE procedure, orange to the failure of a handshake that lead the source node to try on another free channel, and red to the failure of a handshake that leads to a failure of an ALE procedure. The rate associate with the brown transition is obviously λ, and the rate associated with the blue transition is $n_c\mu$ as there are n_c ongoing communications in state $\vec{m}$. The rate associated with the green transition is just the rate of any green transition of the original Markov chain, multiplied by the number n_h of green transitions out of any state $\vec{n}$ of this chain such that $nh(\vec{n}) = n_h$, i.e., $\frac{n_h p_s}{p_s T_s + p_f T_f}$. The only one difficulty of the aggregated model is to estimate the rate of the red transition. Indeed with the aggregated state description $\vec{m}$, the orange transition does not make the system change state, and is thus a loop. And it is well known that loops in CTMCs can be removed without changing in any way the behavior of the chain.

In order to estimate the rate associated with the red transition, let us first define $\alpha(\vec{m})$, the

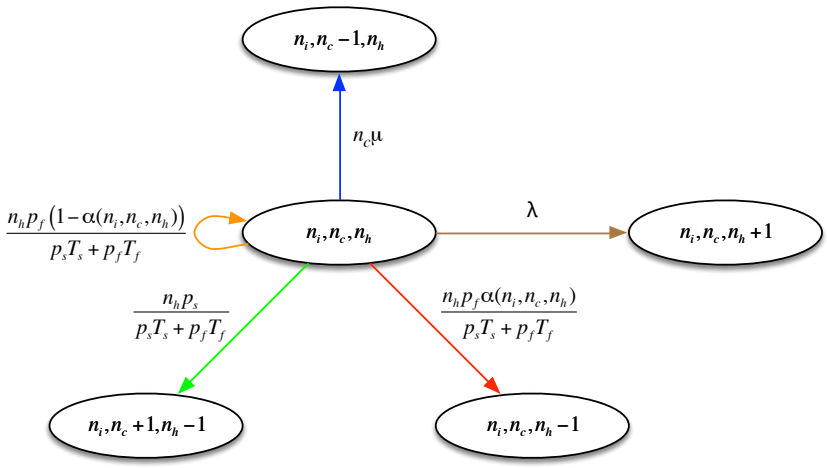


Figure 12: Aggregated Markov chain model

proportion of failures of a handshake that leads to a final failure, when the system is in state $\vec{m}$. In the original detailed Markov chain, this proportion corresponds to the proportion of channels used for a handshake which failure will necessarily lead to the failure of the corresponding ALE procedure. In other words, this can be seen as the proportion of red transitions among the orange and red transitions, which in turn maps to the proportion of channels used for a handshake that are not followed by an idle channel. As an illustration, for the state $\vec{n} = (\hat{f}_1, \bar{f}_2, f_3, \bar{f}_4, \hat{f}_5)$ depicted in Figure 5, $\alpha(\vec{n}) = \frac{1}{2}$.

In order to derive the exact value of $\alpha(n_i, n_c, n_h)$, one would need to calculate the sum of the $\alpha(\vec{n})$ for all vectors $\vec{n}$ such that $ni(\vec{n}) = n_i$, $nc(\vec{n}) = n_c$ and $nh(\vec{n}) = n_h$, weighted by the corresponding stationary probabilities:

$$\alpha(n_i, n_c, n_h) = \sum_{\vec{n} \,|\, ni(\vec{n})=n_i,\ nc(\vec{n})=n_c,\ nh(\vec{n})=n_h} p(\vec{n})\alpha(\vec{n}). \tag{22}$$

This is obviously not a possible solution as it would require to solve the original detailed Markov chain to obtain the probabilities $p(\vec{n})$, and this is exactly what we want to avoid. In order to work around the problem, we are going to assume that all states $\vec{n}$ corresponding to a given state $\vec{m}$ (i.e., all states such that $ni(\vec{n}) = n_i$, $nc(\vec{n}) = n_c$ and $nh(\vec{n}) = n_h$) are equiprobable. This is actually not true, and we will evaluate the impact of this assumption on numerical results.

What is important in order to compute the probability $\alpha(\vec{n})$ associated with a given detailed state $\vec{n}$ is the proportion of channels used by a handshake that are followed by at least one idle channel. As a result, the positions of the channels used for a communication in $\vec{n}$ are not important in the derivation of $\alpha(\vec{n})$. We are thus going to consider a reduced state description $\vec{n}'$ having $n_i + n_h$ components (corresponding to channels that are not used for a communication), each one being a "I" if a channel is idle or a "H" is it is used for a handshake. As an example, the reduced state corresponding to state $\vec{n} = (\hat{f}_1, \bar{f}_2, f_3, \bar{f}_4, \hat{f}_5)$ is $\vec{n}' = (H, I, H)$. With this reduced state description is it obvious that one "H" is followed by a "I" and the other is not, and thus $\alpha(\vec{n}') = \frac{1}{2}$.

We list in Table 1 all reduced states $\vec{n}'$ associated with a given state $\vec{m} = (n_i, n_c, n_h)$ of the aggregated model (with $n_i \geq 1$ and $n_h \geq 1$). Such a state $\vec{n}'$ has exactly n_i "I" components and n_h "H" components. For all of these states ending by a "I" component, the corresponding $\alpha(\vec{n}') = 0$, as any "H" component of $\vec{n}'$ is at least followed by the last "I" component. And, as illustrated in the table, for all states ending by a "I" component followed by exactly k "H" components ($k \leq n_h$), the corresponding $\alpha(\vec{n}') = \frac{k}{n_h}$, as any of the first $n_h - k$ "H" is at least followed by the last "I", and the k remaining "H" are not followed by any "I". Table 1 also gives the number of states of each kind. For example, the numbers of states $\vec{n}'$ ending by a "I" components is $C^{n_h}_{n_h+n_i-1}$, as it corresponds to the number of way to choose the n_h "H" components among the $n_h + n_i - 1$ first components of vector $\vec{n}'$. Note finally that the number of all states $\vec{n}'$ is $C^{n_h}_{n_h+n_i}$.

Table 1: Reduced states $\vec{n}'$ corresponding to a given state $\vec{m} = (n_i, n_c, n_h)$ (with $n_i \geq 1$ and $n_h \geq 1$)

states $\vec{n}'$ ending by	number of such states	$\alpha(\vec{n}')$
$(..., I)$	$C^{n_h}_{n_h+n_i-1}$	0
$(..., I, H)$	$C^{n_h-1}_{n_h-1+n_i-1}$	$\frac{1}{n_h}$
$(..., I, H, H)$	$C^{n_h-2}_{n_h-2+n_i-1}$	$\frac{2}{n_h}$
$(..., I, H, H, H)$	$C^{n_h-3}_{n_h-3+n_i-1}$	$\frac{3}{n_h}$
...		
$(I, ..., I, H, ..., H)$	1	1

The proportion $\alpha(n_i, n_c, n_h)$ can thus be derived from the following simple arithmetic average:

$$\alpha(n_i, n_c, n_h) = \frac{\sum_{k=1}^{n_h} C^{n_h-k}_{n_h-k+n_i-1} \frac{k}{n_h}}{C^{n_h}_{n_h+n_i}}. \tag{23}$$

Finally, the rate associated with the red transition out of state $\vec{m} = (n_i, n_c, n_h)$ of Figure 12 can be estimated by the product of the rate of any red transition out of a corresponding state $\vec{n}$ of the original model ($\frac{p_f}{p_s T_s + p_f T_f}$) by the average number of red transitions out of all such states $\vec{n}$ ($n_h \alpha(n_i, n_c, n_h)$), i.e., $\frac{n_h p_f \alpha(n_i, n_c, n_h)}{p_s T_s + p_f T_f}$.

As for the detailed Markov chain, this aggregated Markov chain can be solved using any appropriate numerical technique, but much faster than the original one. We also need to adapt the expressions of the performance parameters (described in Section 4.3) to the resulting stationary probabilities $p(\vec{m})$.

The average number all call requests rejected by unit of time, X_r, the average number of call requests leading to a success (meaning to a communication) by unit of time, X_s, and the average number of call requests leading to a failure by unit of time, X_f, can be expressed as follows:

$$X_r = \sum_{\vec{m} \mid n_i = 0} p(\vec{m}) \lambda, \tag{24}$$

$$X_s = \sum_{\vec{m}} p(\vec{m}) \frac{n_h p_s}{p_s T_s + p_f T_f}, \tag{25}$$

$$X_f = \sum_{\vec{m}} p(\vec{m}) \frac{n_h p_f \alpha(n_i, n_c, n_h)}{p_s T_s + p_f T_f}. \tag{26}$$

And the mean number of channels used for a handshake, Q_h, or for a communication, Q_c, are:

$$Q_h = \sum_{\vec{m}} p(\vec{m}) n_h, \tag{27}$$

$$Q_c = \sum_{\vec{m}} p(\vec{m}) n_c. \tag{28}$$

It is worthwhile noting that, as all vectors $\vec{m} = (n_i, n_c, n_h)$ appearing in the multiple summations of relations 24 to 28 are such that $n_i + n_c + n_h = N$, these expressions only involve a double summation (e.g., for $n_i = 0$ to N and for $n_c = 0$ to $N - n_i$).

All the remaining performance parameters keep the same expression as those developed in Section 4.3.

6.2 Numerical results of the aggregated model

We follow here the same procedure (used for the detailed model in Section 5) to validate our aggregated model. The steady-state equations associated with the chain are again solved numerically via MATLAB, but much faster than those associated with the detailed model. Actually, all the analytical results presented in this subsection were obtained in about 5 seconds, instead of 5 minutes with the detailed model with the same number of channels $N = 5$ (and the difference would be much bigger with a higher value of N). We draw the same performance parameters as those presented in Section 5, for a load $\lambda \in]0;1]$. Here we only give the representative curves corresponding to $p_s = 0.5$. The results corresponding to $p_s = 0.1$ or to $p_s = 0.9$ are indeed very similar and lead to comparable errors when compared to simulation.

Figure 13 depicts the occupancy of channels as a function of the load. The relative differences between the performance parameters derived from the aggregated model and those obtained from simulation (for each of the three parameters), are not significantly bigger than those obtained with the detailed model (compare to Figure 6). The average relative error between model and simulation is now less than 5%, with a maximum error of around 10%. This highlights the accuracy of the aggregated model for estimating the occupancy of channels.

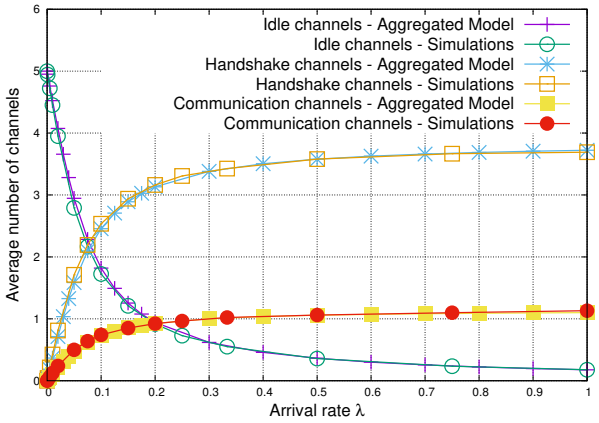


Figure 13: State of channels function of the arrival rate for $p_s = 0.5$

Figure 14 compares the probabilities of rejection, success and failure, obtained from the aggregated model to simulations. Here we observe a notable difference at low load, especially for the two last parameters (success and failure probabilities), between the model and simulation. This difference is actually the result of the main approximation of the aggregated model, that assumes the equiprobability of all states $\vec{n}$ of the detailed model corresponding to a given state $\vec{m}$ of the aggregated model (see Section 6.1). This equiprobability is particularly inaccurate when the load is low. Indeed for a very low load, channels with a small number (close to 1) are more likely to be used than channels with a high number (close to N). On the other hand, the curves of the aggregated model merge rapidly with those corresponding to simulation when load increases. This highlights the fact that the aggregated model can produce poor performance in the case of a low load, but remains very accurate when the load ranges between medium to high. However, we must recall that at very low load, the performance of the system are very precisely estimated by the close-form asymptotic values derived in Section 4.4. This greatly diminishes the importance of the low accuracy of the aggregated model in low load situations.

Figure 15(a), 15(b), 15(c) show the average duration of an ALE procedure, unconditioned (R_{ALE}), conditioned by success ($R_{ALE \mid s}$), and conditioned by a failure ($R_{ALE \mid f}$), respectively. Here again, we can see that the aggregate model deviates significantly from simulation at low load (especially for the first two parameters). A relative error of about 45% is observed on $R_{ALE \mid s}$ when load tends to zero. But again, in such limiting cases, the three performance parameters are very well estimated by the asymptotic expressions given in Section 4.4.

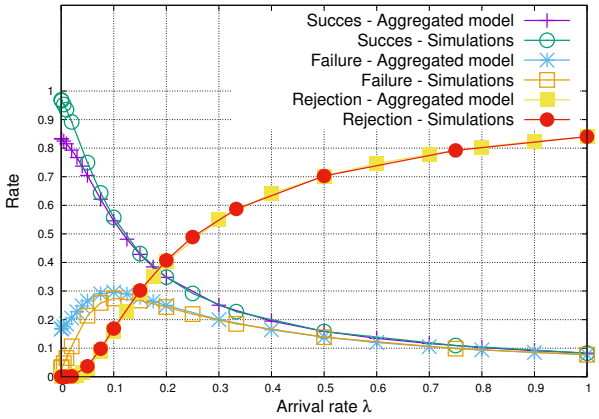


Figure 14: ALE acceptance function of the arrival rate for $p_s = 0.5$

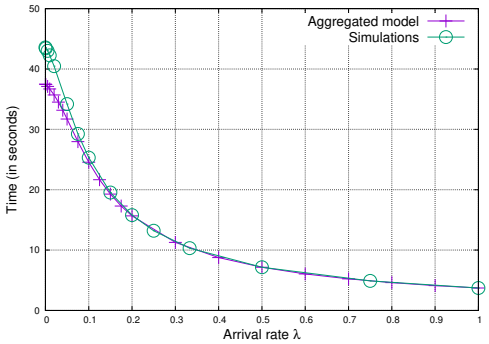
6.3 Modeling non negligible LBT time

We detail here how to easily relax Assumption 4 assuming that the "Listen Before Talk" time, T_{LBT}, is negligible. It is worthwhile noting that this extension is presented in the context of the aggregated model, but can be applied in a very similar way to the detailed model of Section 4. In order to take into account a non negligible LBT time into models, we first need to point out that a node that is listening on a given channel does not directly impact the behavior of other nodes but only introduces an additional delay in its own behavior. As an illustration, if a new call request arrives on a given node, and if this request finds the system in state $(\hat{f}_1, \bar{f}_2, f_3, \bar{f}_4, \hat{f}_5)$ illustrated in Figure 4 for instance, the node will first have to listen on channel 1, then on channel 2 and finally on channel 3, in order to realize that channel 3 is the first idle channel at that time. It is thus only after a delay of $3T_{LBT}$ that the node can try to establish the communication on channel 3.

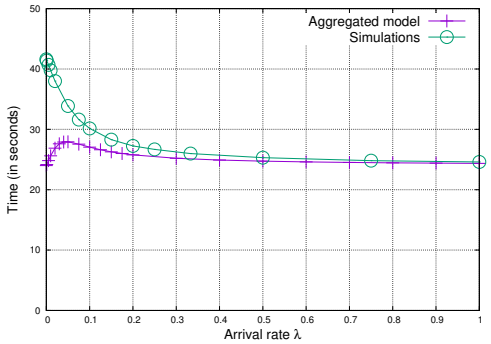
We need to take into account this additional delay in our models. The first idea would be to include in the state description of the models (either the detailed model or the aggregated model) information about channels that are currently sensed by nodes that are willing to initiate a new communication. This however would increase drastically the number of states of the underlying Markov chains hence preventing their resolution in a reasonable time. Instead, we exploit the fact that the dynamic of the system is not really impacted by nodes that are in listening state. We can thus use previous models (with a negligible T_{LBT}) to obtain all performance parameters of interest (e.g., P_r, P_s and P_f), and finally readjust the ALE durations (R_{ALE}, $R_{ALE \mid r}$, $R_{ALE \mid s}$ and $R_{ALE \mid f}$) by just adding to them the average delay involved by the Listen Before Talk procedure of nodes.

The readjustment of $R_{ALE \mid r}$ and $R_{ALE \mid f}$ is trivial. Indeed, provided that a call request has been rejected or has resulted in a final failure, all N channels have been necessarily sensed by the node that has initiated the request. As a result, we must add a value of NT_{LBT} to both $R_{ALE \mid r}$ (that was previously null) and $R_{ALE \mid f}$ (obtained from relation 10).

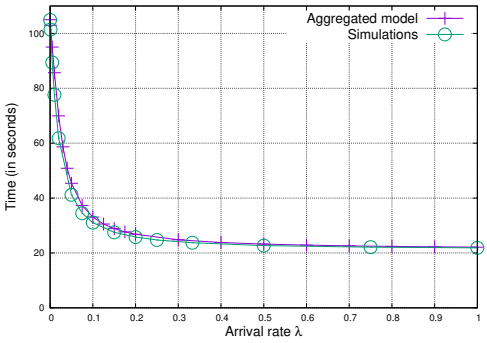
In contrast, the readjustment of R_{ALE} or $R_{ALE \mid s}$ needs more attention. Note however that once one of these values is calculated, it will be straightforward to obtain the other using the law of total probability (relation 7). Both can be derived in a very similar way. Here we detail the readjustment of $R_{ALE \mid s}$. Conditioned by the fact that an arriving call request finds the system in state $\vec{m} = (n_i, n_c, n_h)$, meaning, thanks to PASTA property, with a probability $p(\vec{m})$, the calling node will try to place successively the call on the n_i free frequencies. In the case where the ALE procedure finally succeeds on the n^{th} free frequency (for any $n = 1, ..., n_i$), i.e., with a probability $\frac{{p_f}^{n-1} p_s}{1-{p_f}^{n_i}}$ (remember the ALE is conditioned by a success), the node had to sense the n first free channels, as well as all non-free channel that precede the n^{th} chosen free channel. Knowing that there are $N - n_i$ non-free channels, and $n_i + 1$ possible locations of any of these non-free channels among the n_i free channels, the average number of non-free channels before the n^{th} chosen free



(a) R_{ALE} function of the arrival rate



(b) $R_{ALE \mid s}$ function of the arrival rate



(c) $R_{ALE \mid f}$ function of the arrival rate

Figure 15: ALE duration for $p_s = 0.5$

channel can be estimated as $n \times \frac{N-n_i}{n_i+1}$. As a result, the readjustment of $R_{ALE \mid s}$ consists in adding to the expression obtained with a model assuming a negligible LBT, an additional time given by relation 29:

$$\sum_{\vec{m} \mid n_i \neq 0} p(\vec{m}) \left[\sum_{n=1}^{n_i} \frac{{p_f}^{n-1} p_s}{1 - {p_f}^{n_i}} \left(1 + \frac{N - n_i}{n_i + 1} \right) n T_{LBT} \right]. \tag{29}$$

Note that this last equation can be equivalently expressed as:

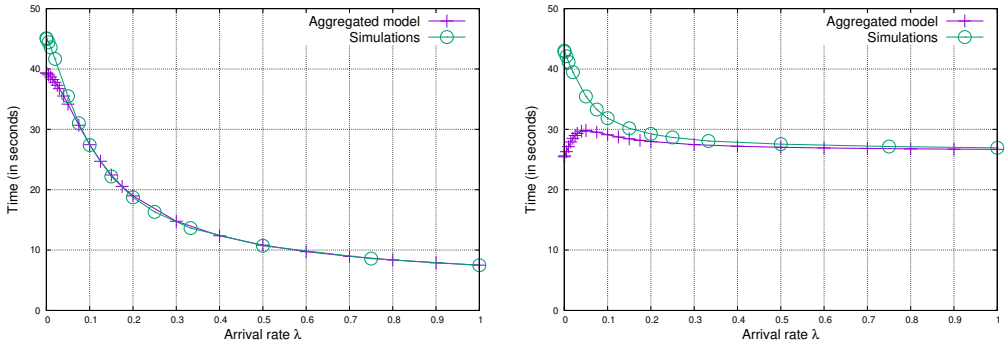
$$\sum_{\vec{m} \mid n_i \neq 0} p(\vec{m}) \frac{\left(1 - (n_i + 1){p_f}^{n_i} + n_i {p_f}^{n_i+1}\right)(N+1)}{(1 - p_f)(1 - {p_f}^{n_i})(n_i + 1)} T_{LBT}. \tag{30}$$

The average ALE duration R_{ALE} is obtained thanks to the relation 8 using the law of total probability.

Figures 16(a), 16(b), 16(c) and 16(d) show the average duration of an ALE procedure, unconditioned (R_{ALE}), conditioned by success ($R_{ALE \mid s}$), conditioned by a failure ($R_{ALE \mid f}$) and conditioned by a rejection ($R_{ALE \mid r}$), respectively, all considering $T_{LBT} = 0.784s$. These curves show that the proposed readjustment is very accurate, enabling the aggregated model to provide performance parameters with comparable average errors (when compared to simulation). We obtain very similar results (not shown here) in the case where $p_s = 0.1$ and 0.9, as well as for all other performance parameters.

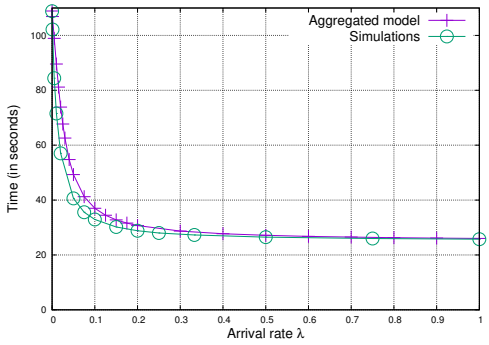
7 Models exploitation

Now that we have validated our models, the focus is on highlighting how these models can be used to help dimension and configure HF networks. In the following, we first present an example of

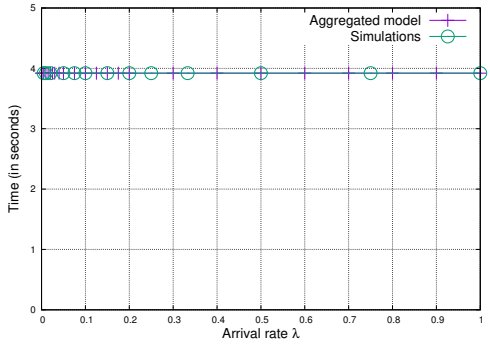


(a) R_{ALE} function of the arrival rate

(b) $R_{ALE \mid s}$ function of the arrival rate



(c) $R_{ALE \mid f}$ function of the arrival rate



(d) $R_{ALE \mid r}$ function of the arrival rate

Figure 16: ALE duration for $p_s = 0.5$ and $T_{LBT} = 0.784s$

exploitation of the detailed model for comparing channel selection strategies, then we highlight a potential use of the aggregated model for the dimensioning of the system according to a given QoS criterion.

7.1 Comparing channel selection strategies

In order to investigate the influence of channel selection strategies, we have first extended our detailed model (presented in Section 4 and validated in Section 5) to account for different success probabilities p_s on the different channels. The extension is straightforward and consists only on indexing p_s and p_f with channel numbers and using appropriate values in the Markov chain. Note that we could not have integrated this feature to the aggregated model. We consider a system made of $N = 5$ channels, with a success probability vector $p_s = (0.1, 0.3, 0.5, 0.7, 0.9)$ corresponding to the success probability on the 5 channels. We then compare three possible techniques in selecting channels in the ALE process: "increasing order" refers to the case where the channels are chosen in increasing order of the success probability, i.e., worst first, "decreasing order" refers to the case where the best channel is selected first and the worst last, and finally the "random order" selects channels randomly without considering their success probability.

In Figure 17, we compare the ALE duration for these three selection strategies. From these curves, two main conclusions can be made. First, at low loads (large sparse networks), selecting first the best available channels for transmissions can significantly reduce the ALE handshake duration. This observation is particularly true when comparing to the increasing order (the worst first), the gap being lower when compared to a random channel selection strategy. Indeed, when the load is low, the probability that an arriving call request finds all channels free is high, and as a results testing first "good" channels (corresponding to the decreasing order strategy) will increase the chance of quickly establishing a communication on one of the N free channels. Second, at high loads all selection strategies perform similarly in terms of link establishment duration. Indeed, when

the load is high, few channels are available and all strategies have approximately the same chance of finding a free channel for transmission. As a result all strategies lead to a comparable ALE duration. Interestingly, the performance gap between the three strategies decreases quickly. In fact, this observation is coherent with previous results showing that the number of available channels for communication drops sharply when increasing the load.

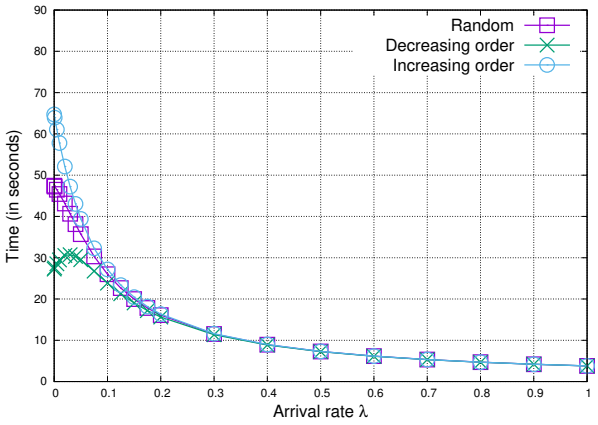


Figure 17: ALE duration function of the load, for different frequency selection strategies

Figure 18 also investigates the ALE success probabilities for these three selection strategies. As shown by the figure, the frequency selection impact is negligible when considering the success rate of the ALE. In other words, clever channel choice can sometimes make the process faster but its impact on the ALE outcome remains very limited.

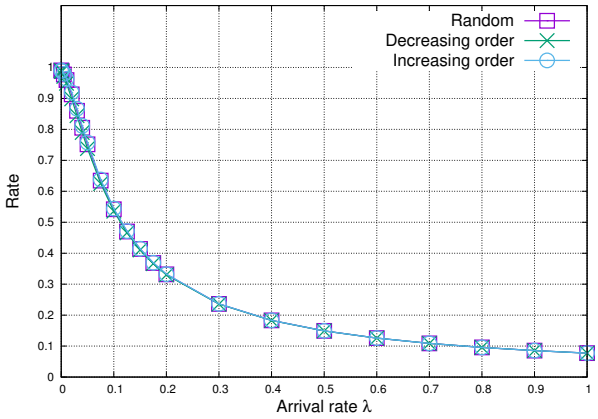


Figure 18: P_s function of the load, for different frequency selection strategies

7.2 Mission planning

Next we use our aggregated model to derive Erlang-like curves in order to dimension the parameters of the HF system so as to ensure a given level of QoS. Note that this kind of dimensioning diagrams requires to run a high number of configurations, forcing the use of a very efficient (fast) performance evaluation tool and definitely prohibiting the use of simulation. We thus naturally chose to use our aggregated model.

As an example, the curves represented in Figure 19 enable to derive the minimal number of channels ensuring that at least 80% of call requests result in a success, i.e., $P_s > 80\%$. The curves are drawn for a success probability $p_s = 0.5$ that is supposed to be the same for all channels.

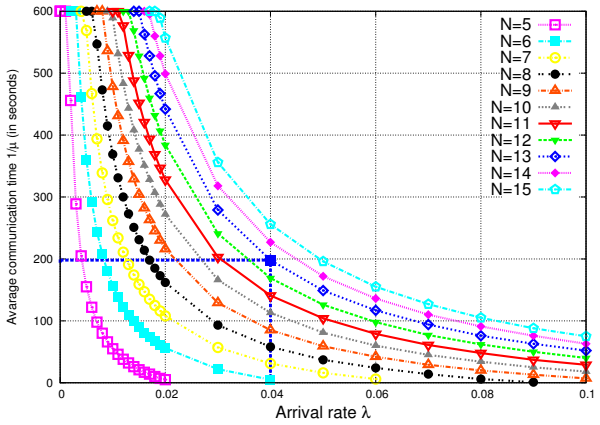


Figure 19: Dimensioning curves to ensure $P_s > 80\%$, for $p_s = 0.5$

They can be used as follows. If the arrival rate of calls is $0.04s^{-1}$ (corresponding to one call every $25s$ in average) and a communication lasts for $200s$ in average, the intersection point is between the two curves corresponding to $N = 12$ and $N = 13$. As a result, a minimum of 13 channels is necessary to ensure 80% of success in the communication establishment. This type of charts serves as a benchmark for worst case scenario. It actually enables, based on predicted link quality (p_s) as well as operational data such as calls arrival rate and their mean durations, the estimation of the number of frequencies required for an operational mission.

8 Conclusion and Future Work

Since late 90's, two standards for HF communications have been proposed. However, in order to prepare the new generation of HF standards capable to take advantage of recent advances in wireless communication and networking, thorough understanding of existing standards and their limitations deems necessary. n this paper, we have developed Markovian models for analyzing the most widely used Automatic Link Establishment (ALE) procedure in HF networks, namely the 2G ALE. We have compared the performance derived from our models to OMNet++ simulations and shown their accuracy. Our models allows us to set guidelines for designing the next generation of ALE standards. In particular, the handshake duration should be reduced in order to make more channels available for communication. Conversely, there is no need for a clever channel choice strategy, as even if it can sometimes make the process faster, its impact on the ALE outcome remains very limited. Besides, our model enables the analysis of the complex interplay between different ALE parameters and their influence on the system capabilities. As such they provide powerful tools to plan and dimension ALE 2G networks. In the future, we plan to enrich our models by relaxing some assumptions, e.g., relating the success of a 3-way handshake to the load of the system or considering buffered calls. Adapting our model to 3G ALE is also an ongoing work.

References

[1] K. Bian and J. Park. Maximizing rendezvous diversity in rendezvous protocols for decentralized cognitive radio networks. *IEEE transactions on Mobile Computing, Vol.12, No.7*, 2013.

[2] G. Bianchi. Performance analysis of the ieee 802.11 distributed coordination function. *IEEE Journal on Selected Areas in Communications (JSAC), Vol.18, No.3*, 2000.

[3] S. Bodas, S. Shakkottai, L. Ying, and R. Srikant. Scheduling in multi-channel wireless networks: Rate function optimality in the small-buffer regime. In *Proceedings of the ACm Sigmetrics/Performance*, 2009.

[4] C. Chao and H. Tsai. A channel-hopping multichannel mac protocol for mobile ad hoc networks. *IEEE transactions on Vehicular Technology, Vol.63, No.9*, 2014.

[5] L. DaSilva and I. Guerreiro. Sequence-based rendezvous for dynamic spectrum access. In *Proceedings of the IEEE DySpan*, 2008.

[6] W.N. Furman, E. Koski, and J.W. Nieto. Design concepts for a wideband hf ale capability. In *Proceedings of the 12th IET International conference on ionospheric radio systems and techniques (IRST)*, 2012.

[7] Z. GU and H. Pu adn Q-S. Hua. Improved rendezvous algorithms for heterogeneous cognitive radio networks. In *Proceedings of the IEEE Infocom conference*, 2015.

[8] V. Gupta, M. Gong, S. Dharmaraja, and C. Williamson. Analytical modeling of bidirectional multi-channel ieee 802.11 mac protocols. *International Journal on Communication Systems (IJCS), Vol.24, No.5*, 2011.

[9] J. He and H. Punng. Performance modeling and evaluation of ieee 802.11 distributed coordination function in multihop wireless networks. *Elsevier Computer Communications Journal, Vol.29, No.9*, 2006.

[10] R. Jayaparvathy, S. Anand, S. Dharmaraja, and S. Srikanth. Performance analysis of ieee 802.11 dcf with stochastic reward nets. *International Journal on Communication Systems (IJCS), Vol.20, No.3*, 2007.

[11] L. Jiao, V. Pla, and F.Y. Li. Analysis on channel bonding/aggregation for multi-channel cognitive radio networks. In *Proceedings of the European Wireless conference*, 2010.

[12] E. Koski, S. Chen, S. Pudlewiski, and T. Melodia. Network simulation for advanced hf communications engineering. In *Proceedings of the 12th IET International conference on ionospheric radio systems and techniques (IRST)*, 2012.

[13] T. Kuang and C. Williamson. A bidirectional multi-channel mac protocol for improving tcp performance on multihop wireless ad hoc networks. In *Proceedings of the ACM Ininternational symposium on Modeling, analysis and simulation of wireless and mobile systems (MSWIM)*, 2004.

[14] Z. Lin, H. Liu, and X. Chu. Enhanced jump-stay rendezvous algorithm for cognitive radio networks. *IEEE Communications Letters, Vol.17, No.9*, 2013.

[15] NATO. Stanag4538 technical standards for an automatic radiocontrol system (arcs) for hf communication links. *Ed 1*, 2000.

[16] R. Prouvez, B. Baynat, H. Khalife, V. Conan, and C. Lamy-Bergot. Modeling automatic link establishment in hf networks. In *Proceedings of the IEEE Military Communications Conference (MILCOM)*, 2015.

[17] A. Shahid, S. Ahmad, A. Akram, and S. A. Khan. Cognitive ale for hf radios. In *Proceedings of the second international conference on computer engineering and applications*, 2010.

[18] D. Shen and V. Li. Performance analysis for a stabilized multi-channel slotted aloha algorithm. In *Proceedings of the IEEE Pimrc conference*, 2003.

[19] H. Shiang and M. Van der Schaar. Queuing-based dynamic channel selection for heterogeneous multimedia applications over cognitive radio networks. *IEEE Transactions on multimedia, Vol.10, No.5*, 2008.

[20] Y. Song and J. Xie. Performance analysis of spectrum handoff for cognitive radio ad hoc networks without common control channel under homogeneous primary traffic. In *Proceedings of the IEEE Infocom conference*, 2011.

[21] Military Standard. Mil-std-188-141 a, interoperability and performance standards for medium and high frequency radio systems. *Department of Defense*, 1999.

[22] H. Su and X. Zhang. Cross-layer based opportunistic mac protocols for qos provisionings over cognitive radio wireless networks. *IEEE Journal on Selected Areas in Communications (JSAC), Vol.26, No.1*, 2008.

[23] D. Tait, A.F.R. Gillespie, and S.E. Tinder. Modelling 2g and 3g ale: A quantitative comparison. In *Proceedings of the 12th IET International conference on ionospheric radio systems and techniques (IRST)*, 2012.

[24] N. Theis, R. Thomas, and L. DaSilva. Rendezvous for cognitive radios. *IEEE tranactions on Mobile Computing, Vol.10, No.2*, 2010.

[25] A. Tzamaloukas and J.J. Garcia-Luna-Aceves. A receiver-initiated collision-avoidance protocol for multi-channel networks. In *Proceedings of the IEEE Infocom conference*, 2001.

[26] C.-W. Wang, L.-C. Wang, and F. Adachi. Modeling and analysis for proactive-decision spectrum handoff in cognitive radio networks. In *Proceedings of the IEEE globecom conference*, 2010.

[27] W. Yue and Y. Matsumoto. Performance analysis of multi-channel and multi-traffic on wireless communication networks. *Book Kluwer Academic Publishers*, 2002.