Probabilistic Routing for Small Devices – Formal Model ⋆

Jacek Cichoń and Marek Klonowski

Institute of Mathematics and Computer Science
Wrocław University of Technology
Poland

Abstract. In this paper we study efficiency of fundamental routing protocols in distributed systems assuming that communication between nodes can fail. Described and analyzed models are applicable for wide class of systems built of small devices (like sensors) with energetic constrains. We introduce s formal model and present rigorous analysis of several basic strategies of an oblivious protocol. We show dramatic difference between efficiency of seemingly similar strategies. Since the real-life systems usually consist of moderate number of devices, presented analysis is not limited to asymptotic behavior (in some cases trivial). Instead, where possible exact formulas are given.
We believe that presented results for basic protocols as well as the formal model can be useful building blocks for analysis of other, more sophisticated strategies adjusted for specific goals of a given system. In particular, for network of sensors working in an Ad Hoc radio network in hostile environment which makes the communication channel unreliable.

1 Introduction

In this paper we investigae several basic strategies of transmitting information in a distributed system with unreliable communication. The aim is to transmit a signal through the path of intermediate nodes. However, any transmission may fail with some controlable probability dependent on devoted amount of resources (usually energy). Motivated by the growing importance of extremely constrained devices we restrict our attention to oblivious protocols - in particular, we assume that devices do not send any acknowledgments confirming successful delivery of a message. This assumption is supposed to catch that considered devices are very simple, working in very hard conditions and their behaviour cannot be adjusted after the predeployment phase.

To provide a rigorous analysis of strategies first we introduce a formal model. We assume that devices spread in an environment are logically sliced into disjoint and totally ordered layers. This assumption contradicts the very common paradigm of random allocation of devices which is absolutely unacceptable in most critical applications.

The main goal of the system is to transmit a signal (information) from the first layer to at least one device in the last layer. Each device in a particular layer can transmit a signal only to some specified devices in the next layer.

The transmission is successful (i.e. broadcast message is correctly received in the next layer) with some fixed probability p. This parameter models the risk of failure of a transmission (caused for instance by interference, impact of environment etc.) and is strongly dependent on the power of the transmission signal. That is, the more power is used, the more probable a successful transmission is.

⋆ Partially supported by EU within the 7th Framework Programme under contract 215270 (FRONTS). Partially supported by funds from Polish Ministry of Science and Higher Education – grant No. N N206 2573 35

The problem is that in real world applications devices are powered by an unrechargeable battery and thus algorithms should be as energy–efficient as possible. Therefore we are looking for a practical trade–off between saving energy and probability of transmitting demanded information.

In various scenarios a device located in the layer i may transmit to a subset of nodes in the layer $i+1$. In settings with devices with directional antennae, the number of possible receivers depends on the angle of transmitting.

Note that the wider the angle the smaller the power density per unit area under the same Tx power. Thus designing more links between nodes requires higher energy consumption. Or equivalently, with the wider angle, the range of broadcasting is shorter, which leads to necessity of distributing more sensors.

In our paper we give an analysis of protocols for basic and most natural structures. However, it can be also used for investigating more complex systems in order to improve their efficiency by optimizing parameters. Note that presented results are true already for systems of very moderate size, what makes them very practical.

In many cases we presented a very precise analysis of some parameters of investigated models, in order to make it useful as a building block for other, possibly more complex systems. In such a case very high precision is necessary, since even a very small factor in a single component may have a significant influence on the overall system.

Some of presented results seems to be simple, however to the best of our knowledge they cannot be proved (with such precision) in any simpler way. In particular they **cannot** be obtained using any standard Cheroff-bound-type method.

1.1 Description of the Model

Let $\{\mathcal{X}_i\}_{i\geq 1}$ be a family of finite pairwise disjoint sets representing consecutive layers of devices. Let $V = \bigcup_{i\geq 1} \mathcal{X}_i$. We interpret elements of V as devices. Let $E = \bigcup_{i\geq 1} (\mathcal{X}_i \times \mathcal{X}_{i+1})$. We shall use the graph (V, E) as an underlying geometric structure of the transmission of signals between nodes.

Let $T : E \to [0, 1]$. We interpert the number $T(e)$ as the probability of a successful transmission of the signal from node x to node y, where $e = (x, y)$. We call the triple $TC = (V, E, T)$ a *Transmission Channell*.

Let $(X(e))_{e\in E}$ be a family of independent $0-1$ random variables such that $\Pr[X(e) = 1] = T(e)$. This random variable models if the message is successfully transmitted over the edge e. In this case we call the edge an *active* edge.

We define recursively a random variable $Y : V \to \{0, 1\}$ in the following way: $Y(x) = 1$ for all $x \in \mathcal{X}_1$ and $Y(x) = \max\{Y(y) \cdot X((y, x)) : y \in \mathcal{X}_i\}$ for $x \in \mathcal{X}_{i+1}$. We say that a node x is active if $Y(x) = 1$, what is interpreted as an event that the information was successfully delivered to the node x. In other words, the node x is active if and only if there is a directed path of active edges from any node in the first layer to the node x. Finally, we define the last layer with an active node as:

$$L = \min\{i : (\forall x \in \mathcal{X}_{i+1})(Y(x) = 0)\} \ .$$

Our goal is to investigate the expected value $\mathrm{E}[L]$ of the random variable L for various Transmission Channels.

In the model considered above the transmitted message may be transported through the network in multiple copies. For some reasons (to be explained in the section 4) it is useful to consider also protocols wherein the message is sent through exactly (possibly partially overlapping) paths. In such case we define a *transmission path* in a transmission channel as a $p_l = (x_1^{(l)}, x_2^{(l)}, \ldots, x_n^{(l)})$ such that $x_i^{(l)} \in \mathcal{X}_i$. Usually there are more than one paths in the transmission channel.

1.2 Previous and Related Work

Randomization in the context of routing protocols for sensor networks has already appeared in many papers. To the best of our knowledge, most of them aim at deliberate usage of randomization in order to improve the efficiency (in terms of time and energy usage) of analyzed protocols. In many cases they are particularly useful for ad hoc radio networks with huge number of devices [15,16].

The last protocol, considered in the section 4 with random transmitting paths is similar to idea of onion routing [10]. This paper, however, randomization of routing was a tool for hiding information about communication patterns in order to provide privacy of users. Another anonymity paper with sending a message in several copies is [11]. Of course there is also a long list of distributed algorithms wherein randomization is used to determine the routing path to improve communication. One of notable examples is the Valiant's scheme [12]. There connections have limited capacity and randomization is used to avoid their overloading. In that model connections are reliable - i.e sent message is always delivered.

Described model resembles some of problems often met in reliability theory. On the other hand some of the investigated processes can be seen as a model of a special kind of percolation phenomenon (e.g. [3]). To the best of our knowledge, however, none of the results can be directly applied as a usefull (non-asymptotic and precise) mathematical foundation for systems of constrained devices.

1.3 Notations and Basic Facts

Let $[k] = \{1, \ldots, k\}$. We denote by $[[X]]$ a function equal 1 if the formula X is true and 0 otherwise (the indicator functions). If f and g are two real functions then we say that $f(x) \sim g(x)$ if $x \to x_0$ if $\lim_{x \to x_0} f(x)/g(x) = 1$. By i we denote the imaginary unit and by $\Re(z)$ we denote the real part of the complex number z. We write $[z^n]f(x) = a_n$ if $f(x) = \sum a_i x^i$

The harmonic number H_n is defined as $H_n = \sum_{i=1}^{n} 1/i$. Let us recall that $H_n = \ln n + \gamma + O(1/n)$, where $\gamma = 0.5772\ldots$ is the Euler-Mascheroni constant.

Random variable X has a geometric distribution with parameter p ($X \sim Geo(p)$) if $\Pr[X = a] = (1-p)^{a-1}p$ for $a = 1, 2, \ldots$. If $X \sim Geo(p)$ then $\mathrm{E}[X] = 1/p$ and $\mathrm{var}[X] = (1-p)/p$.

By $\mathrm{B}(x, y)$ we denote the Euler beta function defined as $B(x, y) = \Gamma(x)\Gamma(y)/\Gamma(x+y)$. We will use the fact that the function $B(n+1, z)$ is analytic everywhere except for $z = 0, -1, -2, \ldots$ and that its residue at $z = -k$ is equal to $(-1)^k \binom{n}{k}$.

2 Long Transmission Channel

Let us fix $k \geq 1$ and let $X_i = \{1, \ldots, k\} \times \{i\}$. We call the structure $(\{X_i\}_{i \geq 1}, T)$ a Long Transmission Channell (LTC). In this section we shall discuss properties of two kinds of LTC: the first one consists of independent lines and the second consists of fully dependent blocks.

2.1 Independent Lines

Let $T((x, i), (y, i+1)) = p \cdot [[x = y]]$. We call the structure $SLTC_{k,p} = (\{\mathcal{X}_i\}_{i \geq 1}, T)$ a Simple Long Transmission Channell of width k. In the simplest case this structure may be a model of a network of devices with directional antennas and very small transmission angle. Depending on physical features such devices may have relatively long range of transmission for a fixed energy. Moreover they are usefull in the radio networks with risk of interference between broadcasting

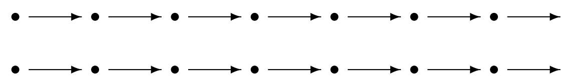

Fig. 1. Simple long transmission channel

stations.

We begin our consideration with some facts about geometric distributions and harmonic numbers. We give a precise formula for an extreme statistic (see e.g. [8]) for geometric distribution and later we prove an asymptotic of this formula for p tending to 1.

Lemma 1. *Let $X_1, \ldots, X_k$ be independent random variables with distribution $Geo(p)$. Let $Y_{k,p} = \max\{X_1, \ldots, X_k\}$. Then*

$$\mathrm{E}[Y_{k,p}] = \sum_{i=1}^{k} \binom{k}{i} \frac{(-1)^{i+1}}{1-q^i} \ ,$$

where $q = 1 - p$.

Proof. Notice that for arbitrary $a \geq 1$ we have $\Pr[Y_k < a] = (\Pr[X_1 < a])^k$, therefore $\Pr[Y_k \geq a] = 1 - (1 - q^{a-1})^k$. Hence $\mathrm{E}[Y_k] = \sum_{a \geq 0}(1 - (1 - q^a)^k)$. Let us fix $A > 1$. Then

$$\sum_{a=0}^{A-1}(1 - (1 - q^a)^k) = A - \sum_{a=0}^{A-1}(1 - q^a)^k =$$

$$A - \sum_{i=0}^{k} \binom{k}{i} \sum_{a=0}^{A-1}(-1)^i q^{ai} = -\sum_{i=1}^{k} \binom{k}{i} (-1)^i \sum_{a=0}^{A-1} q^{ai} = \sum_{i=1}^{k} \binom{k}{i} (-1)^{i+1} \frac{1 - q^{Ai}}{1 - q^i} \ .$$

Therefore

$$\mathrm{E}[Y_k] = \lim_{A} \sum_{i=1}^{k} \binom{k}{i} (-1)^{i+1} \frac{1 - q^{Ai}}{1 - q^i} = \sum_{i=1}^{k} \binom{k}{i} \frac{(-1)^{i+1}}{1 - q^i} \ .\Box$$

Let us come back to the structure $SLTC_{k,p}$. In this case all transmission goes independently through lines $\{(i,1), (i,2), (i,3), \ldots\}$, where $i \in [k]$. Hence the random variable $L_k^{(S)}$ denoting the number of the last layer that received information, follows the same distribution as $\max\{Z_1, \ldots, Z_k\}$, where Z_i's are independent and $Z_i \sim Geo(1-p)$. Therefore, from Lemma 1 we get

$$\mathrm{E}[L_k^{(S)}] = \sum_{i=1}^{k} \binom{k}{i} \frac{(-1)^{i+1}}{1 - p^i} \ .$$

From this formula we get the following equalities:

1. $\mathrm{E}[L_1^{(S)}] = \frac{1}{1-p}$,
2. $\mathrm{E}[L_2^{(S)}] = \frac{3}{2}\frac{1}{1-p} - \frac{1}{2(1+p)}$,
3. $\mathrm{E}[L_3^{(S)}] = \frac{11}{6}\frac{1}{1-p} - \frac{3}{2(1+p)} + \frac{2+p}{3(1+p+p^2)}$.

We will establish now an alternative formula for $\mathrm{E}[L_n^{(S)}]$ for $n > 2$ and for fixed $p < 1$ which allow us to obtain good approximations for arbitrary values of the parameter p.

Theorem 1. *Suppose that $n > 2$ and $0 < p < 1$. Then*

$$\mathrm{E}[L_n^{(S)}] = \frac{1}{2} + \frac{\mathrm{H_n}}{\log(1/p)} + \frac{2}{\log(1/p)} \sum_{k=1}^{\infty} \Re\left[\mathrm{B}\left(n+1, \frac{2k\pi\mathrm{i}}{\log(p)}\right)\right] .$$

Proof. Let $f(z) = \frac{1}{1-p^{-z}}$ be a complex function for $z \in \mathbb{C}$. It is analytic except for $\mathfrak{z}_k = \frac{2k\pi i}{\log(p)}$, $k \in \mathbb{Z}$. If $n > 2$ then using S. O. Rice method (e.g. [13]) we get

$$\sum_{k=1}^{n} \binom{n}{k} (-1)^k \frac{1}{1-q^k} = -\sum_{k=-\infty}^{\infty} \operatorname*{Res}_{z=\mathfrak{z}_k} \mathrm{B}(n+1, z) f(z) .$$

It is easy to check that

$$\operatorname*{Res}_{z=\mathfrak{z}_k} \mathrm{B}(n+1, z) f(z) = \begin{cases} \frac{1}{2} - \frac{\mathrm{H_n}}{\log(p)} & : k = 0 \ , \\ \mathrm{B}(n+1, \mathfrak{z}_k) \frac{1}{\log(p)} & : k \neq 0 \ , \end{cases}$$

from which we get the required formula. □

Directly from the definition of the Beta function we deduce that $\mathrm{B}(n+1, x) = (1+x/n)^{-1}\mathrm{B}(n, x)$. Using this recurrence several times we get

$$\mathrm{B}(n+1, x) = \frac{1}{x} \prod_{a=1}^{n} \frac{1}{1+\frac{x}{a}} .$$

Therefore if $n > 2$ then

$$\left|\sum_{k=1}^{\infty} \mathrm{B}(n+1, \mathfrak{z}_k)\right| \leq \sum_{k=1}^{\infty} \frac{1}{|\mathfrak{z}_k|} \prod_{a=1}^{n} \frac{1}{|1+\mathfrak{z}_k/a|} \leq \sum_{k=1}^{\infty} \frac{1}{|\mathfrak{z}_k|} \frac{1}{|\mathfrak{z}_k|} \frac{1}{|\mathfrak{z}_k/2|} = \sum_{k=1}^{\infty} \frac{2}{|\mathfrak{z}_k|^3} .$$

Let

$$r_n = \frac{2}{\log(1/p)} \sum_{k=1}^{\infty} \Re\left[\mathrm{B}\left(n+1, \frac{2k\pi\mathrm{i}}{\log(p)}\right)\right] .$$

Therefore for $n > 2$

$$|r_n| \leq \frac{4}{\log(1/p)} \sum_{k=1}^{\infty} \frac{|\log p|^3}{(2\pi k)^3} < 0.02 |\log p|^2 .$$

Hence we get the following two corollaries:

Corollary 1. *Suppose that $n > 2$ and $0 < p < 1$. Then*

$$\mathrm{E}[L_k^{(S)}] = \frac{1}{2} + \frac{\mathrm{H_k}}{\log(1/p)} + r_k$$

where $|r_k| < 0.02|\log(p)|^2$.

Corollary 2. *If $p \to 1$ then*

$$\mathrm{E}[L_k^{(S)}] = \frac{\mathrm{H_k}}{1-p} + \frac{1}{2}(1 - \mathrm{H_k}) + O(1-p) .$$

For $n > 2$ the last corollary follows directly from the previous one and for $n \in \{1, 2\}$ it follows from the explicate formulas for $\mathrm{E}[L_k^{(S)}]$. The last corollary can be also deduced directly from Lemma 1 and the classical equality $\sum_{i=1}^{k} \binom{k}{i} \frac{(-1)^{i+1}}{i} = H_k$.

Let us formulate a special case of the previous corollary:

Corollary 3. *If* $p = \frac{1}{2}$ *then*

$$\mathrm{E}[L_k^{(S)}] = 1/2 + \frac{1}{\ln 2}\mathrm{H_k} + r_k \ ,$$

where $|r_k| < 0.01$.

Notice that $1/\ln 2 = 1.4427...$. More precise calculus shows that in this case ($p = \frac{1}{2}$) we have $|r_k| < 0.0021$ for all $k > 2$.

3 Fully Dependent Lines

Let $T((x,i),(y,i+1)) = p$. We call the structure $FLTC_{k,p} = (\{\mathcal{X}_i\}_{i \geq 1}, T)$ a Full Long Transmission Channell of width k. Such structure represents for example settings with directorial antennas with wide range of broadcasting or can model the sensor networks deployed in long channels (e.g. in the mines). By the analogy to previous sections, the random variable $L_k^{(F)}$ denotes the number of the last layer with an active node.

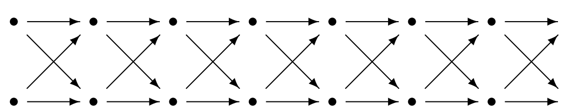

Fig. 2. Symmetric long transmission channel

We may look on the routing in the transmission channel in this case as on the Markov chain with states $\{0, 1, \ldots, k\}$, where a state i denotes the event that precisely i nodes are active. The probability of transitions are given by the formula

$$p_{a,b} = \binom{k}{b}(1 - q^a)^b (q^a)^{k-b} \ .$$

where, as usual, $q = 1 - p$. We shall show upper and lower bounds for the expected length $\mathrm{E}[L_k^{(F)}]$ of the transmission range in this model.

Theorem 2.

$$\mathrm{E}[L_k^{(F)}] \leq \frac{1}{(2 - q^k)^k} \left(\frac{1}{q}\right)^{k^2} .$$

Proof. As before, let $q = 1 - p$ be the probability that that the signal are not transmitted between two particular vertices (x, i) and $(y, i + 1)$ for $i > 0$ and $x, y \in [k]$. Let $N_{x,i}^{in}$ denotes an event that none of links entering the node on the level x in the time i is active. i.e. $N_{x,i}^{in} = \bigcap_{y \in [k]} \{X((y, i-1), (x, i)) = 0\}$. Similarly, $N_{x,i}^{out}$ denotes the event that none of links departing from x is active, i.e. $N_{x,i}^{out} = \bigcap_{y\in[k]}\{X((x,i),(y,i+1))=0\}$. Let $N_{x,i} = N_{x,i}^{in} \cup N_{x,i}^{out}$. Noting that the the events that all links entering the x are not active and the event that all links outcoming from x are not active are independent, one can easy see that $Pr[N_{x,i}] = q^k(2-q^k)$. Finally let us define $N_i = \bigcap_{x\in[k]} N_{x,i}$ - event, that in the i-the layer each node is cut-off from next or previous layer. Since, $N_{x,i}$ and $N_{y,i}$ are independent for $x \neq y$, thus $Pr[N_i] = q^{k^2}(2-q^k)^k$.

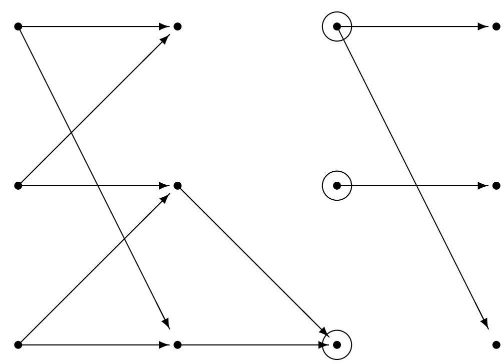

Fig. 3. Possible realization of N_3 event

Observe that if the i-th layer is cut-off then $L_k^{(F)} \le i$, i.e. $\{L_k^{(F)} > t\} \subset \bigcap_{i<t} {N_i}'$ and that ${N_i}'$ and ${N_j}'$ are independent if $|i-j| \ge 2$. Thus

$$\Pr[L_k^{(F)} > 2t] \le \Pr\left[\bigcap_{i<2t} {N_i}'\right] \le \Pr\left[\bigcap_{i<2t,2|i} {N_i}'\right] = \Pr[{N_1}']^t \le (1-q^{k^2}(2-q^k)^k)^t.$$

From the above formula one can easily get that $2 \cdot L_k$ is stochastically dominated by geometric distribution with the success parameter $q^{k^2}(2-q^k)^k$. In particular

$$\mathrm{E}[L_k^{(F)}] \le \frac{1}{2}\mathrm{E}[Geo(q^{k^2}(2-q^k)^k)] = \frac{1}{2(2-q^k)^k}\left(\frac{1}{q}\right)^{k^2}$$

as stated in the theorem. $\square$

We shall use in next considerations the following simple but usefull lemma:

Lemma 2. *Let $(X(n))$ be a Markov chain with a state space S, let $A \subseteq S$ and let $r > 0$ be such that*

$$\Pr[X(1) \in A | X(0) \in A] > r.$$

Let $W = \min\{n : X(n) \notin A\}$. Then

$$\mathrm{E}[W_n | X(0) \in A] \ge \frac{1}{1-r}\ .$$

Theorem 3.

$$\mathrm{E}[L_k^{(F)}] > \frac{1}{\sqrt{2}}\frac{\sqrt{k}}{2^k}\left(\frac{1}{q}\right)^{\frac{1}{4}k(k+2)}\left(1+O(q^{k/2})\right)$$

Proof. We shall use in the proof the following inequality $\binom{n}{k} \le 2^n/\sqrt{n/2}$ which holds for all $n \ge 1$ and all k. Let us fix $k \ge 1$ and let $k^* = \lceil \frac{n}{2} \rceil$. Then for a fixed a we have

$$\sum_{b<k^*} p_{a,b} = \sum_{b<k^*} \binom{k}{b} q^{a(k-b)}(1-q^a)^b < \sum_{b<k^*} \binom{k}{b} q^{a(k-b)} <$$

$$\sqrt{2}\frac{2^k}{\sqrt{k}} \sum_{b<k^*} q^{a(k-b)} = \sqrt{2}\frac{2^k}{\sqrt{k}} q^{ka} \sum_{b<k^*} \left(\frac{1}{q^a}\right)^b <$$

$$\sqrt{2}\frac{2^k}{\sqrt{k}} q^{ka} \frac{\left(\frac{1}{q^a}\right)^{k^*} - 1}{\frac{1}{q^a} - 1} < \sqrt{2}\frac{2^k}{\sqrt{k}} q^{ka} \frac{\left(\frac{1}{q}\right)^{ak^*}}{\frac{1}{q^a} - 1} = \sqrt{2}\frac{2^k}{\sqrt{k}} \frac{q^{a(1+k-k^*)}}{1-q^a} .$$

Therefore, if $a \ge k^*$ then

$$\sum_{b<k^*} p_{a,b} < \sqrt{2}\frac{2^k}{\sqrt{k}} \frac{q^{a(1+k-k^*)}}{1-q^a} \le \sqrt{2}\frac{2^k}{\sqrt{k}} \frac{q^{k^*(1+k-k^*)}}{1-q^{k^*}} .$$

It is easy to check that $k^*(1+k-k^*) \ge \frac{1}{4}k(k+2)$ and $k^* \ge \frac{k-1}{2}$for each $k \ge 1$. Hence

$$\sum_{b<k^*} p_{a,b} < \sqrt{2}\frac{2^k}{\sqrt{k}} \frac{q^{\frac{1}{4}k(k+2)}}{1-q^{(k-1)/2}}.$$

Let $A = \{k^*, k^*+1, \ldots, k\}$. From the above equation we get

$$\Pr[X(1) \in A | X(0) \in A] > 1 - \sqrt{2}\frac{2^k}{\sqrt{k}} \frac{q^{\frac{1}{4}k(k+2)}}{1-q^{(k-1)/2}} ,$$

hence from Lemma 2 we get

$$\mathrm{E}[L_k^{(F)}] > \frac{1}{\sqrt{2}} \frac{\sqrt{k}}{2^k} \frac{1-q^{(k-1)/2}}{q^{\frac{1}{4}k(k+2)}} = \frac{1}{\sqrt{2}} \frac{\sqrt{k}}{2^k} \frac{1}{q^{\frac{1}{4}k(k+2)}} \left(1 + O(q^{k/2})\right) .$$

□

Let us look once again at the Markov chain considered in this section. Notice that state 0 is absorbing one and that due to our assumptions state k is the initial state. Let W_a be the expected time of a run from the state a to the absorbing state 0. Notice that $\mathrm{E}[L_k^{(F)}] = W_k$. Let $Q = (p_{a,b})_{a,b \in [n]}$ be the fundamental matrix of this Markov chain. Then

$$(W_a)_{a=1,\ldots,n} = (I-Q)^{-1}[1,\ldots,1]^T .$$

where I is the identity matrix of dimension $k \times k$. This matrix can be computed explicitly using symbolic calculation tools for small values of k. From this one can easily get

Theorem 4. *For all $k \le 6$ and $p \to 1$*

$$\mathrm{E}[L_k^{(F)}] = \frac{1}{2^k-1}\left(\frac{1}{q}\right)^{k^2} + \left(\frac{k}{2^k-1}\right)^2 \left(\frac{1}{q}\right)^{k^2-1} + O\left(\frac{1}{q}\right)^{k^2-2} . \tag{1}$$

These results are obviously consistent with bounds for $\mathrm{E}[L_k^{(F)}]$ from Theorems 2 and 3. Moreover, confirmed by numerical expermiments, we state the following conjecture.

Conjecture 1. *Equation 1 is true for all $k \ge 1$.*

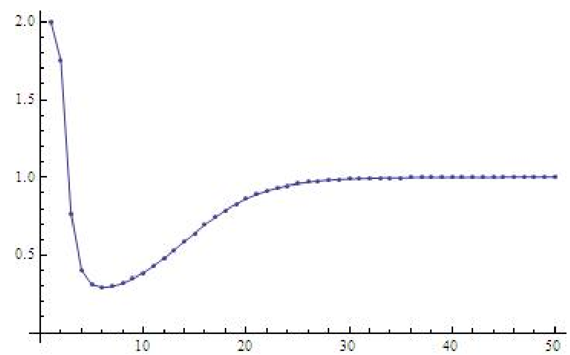

Fig. 4. The plot of a sequence $\mathrm{E}[L_k^{(F)}]/2^{k(k-1)}$ for fully dependent lines and $p = \frac{1}{2}$

3.1 Special case $p = \frac{1}{2}$

The hypothesis formulated by use of equation (1) deals with the case when $p \to 1$. In order to check the reliability of the fully dependent lines for interferences in a realistic but strongly difficult conditions we calculate the numbers $\mathrm{E}[L_k^{(F)}]$ for $p = \frac{1}{2}$ and for all $k \leq 50$. The results for $k \leq 7$ are summarized in the following table:

k	1	2	3	4	5	6	7
$\mathrm{E}[L_k^{(F)}]$	2	7	48.6806	1647.78	326 751	3.12466×10^8	1.30655×10^{12}

At Fig. 4 we show the sequence $\mathrm{E}[L_k]/2^{k(k-1)}$ for $p = \frac{1}{2}$ for $k \leq 50$. We can read from it that for all values from this region and this figure suggests that $\lim_{k\to\infty} \mathrm{E}[L_k^{(F)}]/2^{k(k-1)} = 1$ for $p = \frac{1}{2}$. We state the following conjecture:

Conjecture 2. *If* $p = \frac{1}{2}$ *then* $\mathrm{E}[L_k^{(F)}] \sim 2^{k(k-1)}$.

3.2 Forking Propagation Strategy

In this subsection we consider an intermediate model that can be regarded as a hybrid of the simple long transmission channel and the fully long transmission channel. The idea is simple. There are k lines. Each station broadcasts to its successor in the same line and except that to r neighboring nodes above and below (if exist). Thus, each node may have at most $l = 2r + 1$ successors. Note however that for example, nodes in the first line broadcast just to $r + 1$ nodes. More formally, let $T((x, i), (y, i + 1)) = p$ if and only if $|x - y| \leq r$. Otherwise $T(e)$ is set to 0. We call the structure a l-Forking Long Transmission Channel $ForkLTC_{l,k,p} = (\{\mathcal{X}_i\}_{i\geq 1}, T)$. We assume that $2r + 1 = l \leq k$. Such model is a natural representation for devices with directorial antennas with moderate angle of transmission or limited range allowing to reach only closer nodes. Similarly as in previous sections $L_k^{(\mathrm{Fork},l)}$ denotes the number of the last layer with an active node.

Theorem 5. *Let* $k \geq l$.

$$\mathrm{E}[L_k^{(Fork,l)}] \leq \frac{1}{2^k}\left(\frac{1}{q}\right)^{lk - \frac{1}{2}l^2 + l - \frac{1}{4}}$$

The proof of this theorem is very similar to the proof of the Theorem 2.

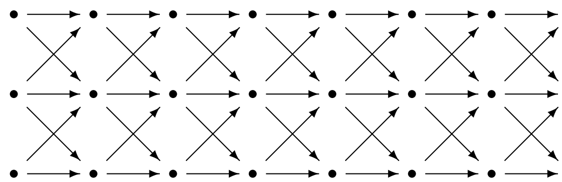

Fig. 5. 3-forking long transmission channel for for $k = 3$

4 Random Walk of Two Processes on k Fully Dependent Lines

In this section we consider the model wherein k transmission paths are passed through a fully dependent lines. Let us fix a number $n > 0$ and let $\mathcal{X}_1 = \{(0,1)\}$, $X_i = \{0,1\} \times \{i\}$ for $i = 2, \ldots, n-1$ and $\mathcal{X}_n = \{(0,n)\}$. We define k transmission paths as a $p_l = (x_1^{(l)}, x_2^{(l)}, \ldots, x_n^{(l)})$ such that $x_i^{(l)} \in_R \mathcal{X}_i$ for $l \in [k]$. That is, hops in consecutive layers are chosen uniformly at random. To model such a system we simply assume that $T((x,y)) = 0$ for every pair (x,y) that is **not** a subsequence of any transmission path. Otherwise, $T((x,y)) = p$. Let us consider the structure $MTC_n = ((\mathcal{X}_i)_{i\in[k]}, T)$. Similarly as in previous subsections let $L_k^{(2,i)}$ denotes the last layer with an active node.

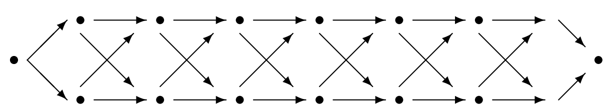

Fig. 6. Symmetric long transmission channel

Theorem 6. *Let $n \geq 2$. Then there are 4^{n-1} pairs of routes from the root $(0,1)$ to the sink $(0,n)$. The expected number of common hops $C_{n,k}^{(2,i)}$ in a random pair of routes of length n is $\frac{1}{4}n + \frac{1}{2}$.*

Proof. When two packages are transmitted through the structure MTC_n then at each step there are four possible positions for the packages: both of them are at the lower strip (we denote this situation by 00), both of them are at the upper strip (we denote this situation by 11) and there are two other possibilities, marked by 01 and 10 when they are at different strips. Let $\mathcal{L}_{xy}$ denote the combinatorial class of all paths starting from xy. We weight these classes by the length of path and we add one additional norm measuring the number of hops through the same edge of the graph. We consider the simple generating functions $L_{xy}(z,u)$ for this class, where z is a variable marking the length of the path and u is used for counting the number of common hops. Then we have:

$$\begin{aligned}
L00 &= 1 + zuL00 + zL01 + zuL11 + zL10, \\
L10 &= zL10 + zL00 + zL11 + zL01, \\
L01 &= zL01 + zL00 + zL11 + zL10, \\
L11 &= zuL11 + zuL00 + zL10 + zL01
\end{aligned}$$

After solving this system of linear equations, extracting $[z^n]L00(z,1)$ and $[z^n]\frac{\partial L00(z,u)}{\partial u}|_{u=1}$ (see e.g. [14]) we obtain required results. $\square$

Theorem 7. $\mathrm{E}[L_2^{(2,i)}] = \frac{11}{8}\frac{1}{1-p} - \frac{3}{8}\frac{1}{1+p}$

Proof. The result is a consequence of Theorem 6 and formulas for L_1 and L_2 from Sec. 2.1. □

Let us now discuss impact of extending the approach described above to the model with $k \geq 2$ fully dependent lines. More exactly, we investigate expectation of $L_k^{(2,i)}$ i.e. the maximal number of layers that is reached by at least of two independently sent processes on the structure of fully depend channel. We show that increasing the parameter k does not help much.

Theorem 8.

$$\frac{11}{8}\frac{1}{1-p} - \frac{3}{8}\frac{1}{1+p} = \mathrm{E}[L_2^{(2,i)}] \leq \mathrm{E}[L_3^{(2,i)}] \leq \ldots \leq \mathrm{E}[L_2^{(S)}] = \frac{3}{2}\frac{1}{1-p} - \frac{1}{2(1+p)}.$$

Proof. First, let us note that probability of having an active node in the n-th layer depends only on the number of common edges chosen by both processes before reaching the n-th layer. Moreover, one can easily see that we can assume that the first process is streamed through first (upper) strip and the second (called a *free process*) is independent - it does not change distribution of number of common edges. Indeed, we just cyclically shift both nodes chosen in consecutive layers. In such representation, the number of common edges in the system with $k+1$ nodes in each layer is stochastically dominated by number of common edges in the system with k nodes in each layer. More precisely, one easy see that the free process can be identified with the sequence of numbers from $[k]$ (or $[k+1]$ respectively) representing the chosen nodes in consecutive layers. The free process uses the same edge as the first process in the i-th step if and only if the free process is represented by two consecutive 1's in the i-th and the $i+1$-th positions. Now, applying a standard coupling argument we get the stochastic domination. □

Note that for practical values ($p > 1/2$) Theorem 8 means that extending parameter k does not improve substantially the reliability of the considered system.

5 Summary and Open Problems

We showed in this paper a very big difference between transmission by independent k lines and fully dependent k lines in a hostile environment: in the first case the expected length of the transmission is $\mathrm{H_k}/\log(1/p)$ and in the second one we get a range of order $2^{-k}(1/(1-p))^{k^2}$. We also check basic properties of a number of other intermediate models.

We believe that presented analysis may be used as a building block for careful investigation of other, more complex systems. From practical point of view, it seems to be very important to analyze the case wherein different connections may fail with different probabilities (e.g according to different physical distances between nodes). However, such analysis as well as creating a realistic model seem to be hard. Our result can also be applied to evaluate reliability some other protocols – including schemes providing anonymous communication (e.g [11])

There are a lot of problems connected with models investigated in this paper such as: does the conjectures 1 and 2 are true?; what is the variance of considered random variables?; what are the limit distributions of adequately standardized considered variables?

References

1. Arnold, B., Balakrishnan, N., Nagaraja, H.: A First Course in Order Statistics. John Wiley & Sons, New York (1992)
2. Cai, Z., Lu, M., Wang, X.: Distributed initialization algorithms for single-hop ad hoc networks with minislotted carrier sensing. IEEE Trans. Parallel Distrib. Syst. **14**(5) (2003) 516–528
3. Bollobás B., Riordan. O.: Cambridge University Press, 2006
4. Cichoń, J., Kutyłowski, M., Zawada, M.: Initialization for ad hoc radio networks with carrier sensing and collision detection. In: ADHOC-NOW, LNCS 4104, 308-320, Springer Verlag (2006) 308–320
5. Cichoń, J., Kutyłowski, M., Zawada, M.: Initialization for ad hoc radio networks with carrier sensing and collision detection. In: ADHOC-NOW, LNCS 4104, 308-320, Springer Verlag (2006) 308–320
6. Nakano, K., Olariu, S.: Randomized initialization protocols for ad hoc networks. IEEE Trans. Parallel Distrib. Syst. **11**(7) (2000) 749–759
7. Stevic, S.: Asymptotic behavior of a sequence defined by iteration. Matematicki Vesnik **48**(3-4) (1966) 99–105
8. Gumbel, E.J.: Statistics of Extremes Columbia University Press
9. Athreya, K.B., Ney P.E.: Branching Processes Dover Publications (2004)
10. Rackoff C., Simon D. R.: Cryptographic Defense Against Traffic Analysis STOC 1993, 672–681
11. Iwanik J., Klonowski M., Kutyłowski M.: DUO–Onions and Hydra–Onions – failure and adversary resistant onion protocols Communications and Multimedia Security 2005, Springer Verlag 2005
12. Valiant L. G.: A scheme for fast parallel communication SIAM Journall of Computing 11:350–361, 1982
13. Knuth D. E.: Art of Computer Programing. Vol I.
14. Flajolet P., Sedgewick R.: Analytic Combinatorics Cambridge University Press, 2008
15. Chatzigiannakis I., Kirousis L., Stratiotis T.: Probabilistic Protocols for Fair Communication in Wireless Sensor Networks. Algosensors 2008, LNCS 5389, 100–110.
16. Chatzigiannakis I., Dimitriou T., Nikoletseas S.E., Spirakis P.: A probabilistic algorithm for efficient and robust data propagation in wireless sensor networks. Ad Hoc Networks (ADHOC) 4(5):621-635 (2006)