International Journal of Networking and Computing – www.ijnc.org
ISSN 2185-2839 (print) ISSN 2185-2847 (online)
Volume 4, Number 1, pages 111–130, January 2014

k-Selection Protocols from Energetic Complexity Perspective*

Marcin Kardas, Marek Klonowski, Kamil Wolny
Wrocław University of Technology,
Wrocław, Poland
Email: {firstname.lastname}@pwr.wroc.pl

and

Dominik Pająk
INRIA,
Bordeaux, France
Email: dominik.pajak@labri.fr

Received: July 26, 2013
Revised: October 29, 2013
Accepted: December 2, 2013
Communicated by Akihiro Fujiwara

Abstract

In this paper we discuss energetic complexity aspects of k-Selection protocols for a single-hop radio network (that is equivalent to Multiple Access Channel model). The aim is to grant each of k activated stations exclusive access to the communication channel. We consider both deterministic as well as randomized model. Our main goal is to investigate relations between minimal time of execution (time complexity) and energy consumption (energetic complexity). We present lower bound for energetic complexity for some classes of protocols for k-Selection (both deteministic and randomized). We also present and analyse several randomized protocols efficient in terms of both time and energetic complexity.

1 Introduction

This paper is devoted to energetic efficiency of protocols solving k-Selection problem. Let us recall that the problem is to grant each of k (out of n) activated stations exclusive access to the communication channel.

It was originally formulated for MAC (Multiple Access Channel). However, this problem can also be stated in an equivalent form for a single-hop radio network. In such a system, for practical reasons, energy consumption is of critical importance. Indeed, while discussing radio networks we often have in mind small battery-supplied sensing devices, that cannot be easily re-charged.

The problem is discussed in various settings. In all of cases there are n stations and some k of them are *activated* and want to broadcast their messages to all other stations. The message

*Partially supported by grant number NN206 369739 of the Polish Ministry of Science and Higher Education (the first three authors) and by ANR Project Displexity (the forth author).

is successfully transmitted only if exactly one station transmits at a given time. In the case of simultaneous transmission of two or more stations a *collision* occurs and no message is delivered to any recipient. If the collision is distinguishable from the background noise we call the model *with collision detection* (CD). Otherwise, the model is described as *no-collision detection* (noCD). The core of the problem is that the subset of activated stations is **unknown in advance** (except that its cardinality is constrained) and stations have to communicate via a very restricted communication channel.

Remarks about the model of energy usage In this paper we concentrate on energetic complexity of k-Selection protocols understood as the maximal energetic effort over all stations. The motivation for such setting comes from sensing networks, wherein it is not feasible to re-charge batteries after deployment. The energy usage of a particular station is the number of rounds when the station transmits, whether the message is delivered or not. In many applications it is required that all (or almost all) devices have to be working for proper acting of network. That is, if the battery of at least a single station runs out, we can assume that the network does not work properly. Therefore, a lifespan of the system is determined by the most loaded station, which motivated us to consider the maximal energy usage. Such approach is used in literature, however one can also find papers where authors consider the average energetic effort of stations instead. It should be noted that in most of the cases in the analysis of the algorithms, finding or even estimating the maximal effort over all stations is technically more challenging. Indeed, in almost all models discussed so far the stations are ubiquitous at the beginning of the protocol thus their effort is represented by random variables from exactly the same distribution that makes the analysis much simpler.

Similarly, there are two common approaches to energetic expense of a station in listening mode. The first one is to take into account both transmitting and listening rounds. In particular, it is the case when all stations are located close to each other. The second approach assumes that energetic cost of listening is dramatically smaller than transmitting and can be treated as negligible. Note also that in the case of some of considered classes of protocols, both approaches are equivalent. Indeed, for example in oblivious algorithms discussed in Section 7 receiving any transmission does not influence the execution and stations can be switched-off instead of being in listening mode.

1.1 Previous work

The k-Selection problem is a classic issue in distributed computing. In recent years it has gained additional interests motivated by expansion of radio (sensor) networks technologies. It is hard to enumerate all important literature related to this topic, thus we mention only the most fundamental papers we are aware of. Komlos and Greenberg considered the oblivious model (acting of stations in one round does not depend on previous rounds) with collision detection. They showed in [19] that k-Selection can be deterministically completed in time $O(k \log(\frac{n}{k}))$. This result can be adapted to the model without collision detection. Moreover the lower bound for the time complexity $\Omega(k \log(\frac{n}{k}))$ which was obtained in [9] holds also for model without collision detection. In [9] the superimposed codes method as well as selective families approach were used. Hayes presented in [13] an adaptive solution which satisfies the same time complexity as for the oblivious model. In [12] the lower bound $\Omega(k \log_k n)$ for the family of adaptive (acting of stations in one round **may** depend on previous rounds) deterministic protocols was proved. In a similar model, Martel [23] showed an interesting randomized approach for finding a maximal value among n stations, which succeeds in the expected time $O(k+\log n)$. Kowalski noted in [20] that Martel's expected time complexity can be improved to $O(k + \log\log n)$ by using the Willard algorithm as a subprocedure. Martel algorithm can be easily adapted to k-Selection problem. Then the time complexity is $O(k)$, because only active stations transmit messages. Another important, recent paper is [2]. The randomized, adaptive solution presented by Anta and Mosteiro guarantees that all of k stations successfully transmit a message in time $(e+1+\xi)k+O(\log^2(\frac{1}{\epsilon}))$ with a reasonably small error probability ϵ taken from a short interval and a fixed negligible constant ξ. In [22] authors analyzed a problem (related to k-Selection) of learning a subset of m stations out of k active ones. Work of Nakano and Olariu [25] can be easily adapted to obtain an algorithm solving w.h.p. (with high probability) the k-Selection problem in

$O(k)$ expected time and $O(\log\log k)$ expected energy usage over all stations. Some comprehensive reference to related work can be found also in [14].

Energetic efficiency of algorithms for radio networks is considered in several papers, devoted to initialization protocols [6], size approximation problem [15], alerts for weak devices [17] or routing aspects [26]. However, to the best of our knowledge, except [25] there are no results about energetic complexity issues of the k-Selection protocols or any other protocol that can be applied to our problem in a straightforward manner.

1.2 Organization of this paper

Our paper presents different results about k-Selection problem in various models.

Section 2 presents the model in details and formulates the problem we investigate.

Section 3 contains a lower bound for the time/energy trade-off for so-called *uniform* randomized algorithms.

Sections 4 introduces a new Monte Carlo algorithm for solving k-Selection problem in $O(k^{1+\epsilon})$ rounds (for any given $\epsilon > 0$) to solve k-Selection problem w.h.p., in such a manner so that no station is awake for more than $O(1)$ rounds.

Section 5 presents a Las Vegas approach for solving k-selection problem. The algorithm requires in expectancy $O\big(k^{\max\{\alpha(\lfloor\frac{1}{\alpha}\rfloor+1),2-\alpha\lfloor\frac{1}{\alpha}\rfloor\}}\big)$ rounds. The expected number of the maximal energy usage is constant. This protocol is an extension of the Massey algorithm.

Section 6 contains analysis of energetic complexity of some known, optimal (in terms of execution time) protocols. We show that some well-known algorithms are not efficient in terms of energy usage.

Section 7 presents a lower bound for all deterministic oblivious algorithms.

Section 8 contains simulations supporting analytic results and giving some insight into practical perspective of investigated protocols. We give evidence that some of our protocols are practical in terms of time of execution.

To the best of our knowledge, in principle all presented results are original except stated otherwise.

2 Model

We consider a single-hop radio network with n stations. The set of all stations is denoted by V. In the case of deterministic algorithms we assume that each station has a unique label from the set $\{1,\ldots,n\}$. Time is assumed to be slotted into rounds. We assume that stations are fully synchronized as if they had access to a global clock. At the beginning of the protocol execution a subset of k stations is activated and each of them has a message that has to be transmitted. Using terminology from [20] we consider *static* k-Selection — all stations start the execution of their algorithm in the same round.

Stations communicate via a single channel. In our paper we assume the collision detection capability of the channel, i.e., the background noise that is received if no station transmits is distinguishable from the noise generated by two or more stations transmitting in the same round. Thus, in each round the communication channel can be in one of three possible states — SILENCE, SINGLE transmission or COLLISION. Note however, that all of the protocols except the GMO algorithm described in Section 5 do **not** use CD capability .

We consider both deterministic as well as randomized algorithms. In the latter case we assume that stations are indistinguishable and have access to a perfect source of random bits. Moreover, sources of different stations are stochastically independent. Note that in the case of deterministic model we assume that stations have unique identifiers. Otherwise solving k-Selection problem would be impossible due to inherent symmetry.

Problem statement Suppose there are n stations and some k of them are activated. Active stations perform (synchronously and in parallel) some common algorithm $\mathcal{A}$, which determines whether to transmit in a given round or remain silent. We say that the algorithm $\mathcal{A}$ solves k-Selection problem if and only if after the execution each active station successfully transmitted its message at least once. That is, we require that for each active station there is a round in which it has exclusive access to the communication channel.

In the deterministic model, each out of n stations has a unique (except trivial 1-Selection case) transmission pattern assigned in advance, so the behavior of a given station is completely determined by its label. Therefore, the deterministic algorithm solving k-Selection problem has to grant exclusive access in case of any k-element subset of labels, denoting activated stations. For $k = \Omega(n)$ there is a trivial solution in which station with label i transmits exactly in the i-th round. The solution is asymptotically optimal, because any algorithm needs at least k rounds. Thus, the most interesting case is for $k \ll n$.

On the other hand, in case of the randomized model the total number of stations n is unimportant and only the number k of active ones matters. In this model the stations are indistinguishable. Therefore, we break the symmetry by means of randomization. We say that randomized algorithm $\mathcal{A}$ solves with high probability (w.h.p.) the k-Selection problem if and only if the probability that all active stations successfully transmit their messages is at least $1 - O(\frac{1}{k})$.

Energy metrics One of the main practical problems in radio networks is the fact that all devices have limited energy resources and moreover in some realistic cases it is very hard to replace their batteries. Thus, the level of energy usage may really matter. In this paper we use the measure of energetic complexity defined as follows. We define $\mathcal{E}_v$, an *energetic effort* of a station $v \in V$, as the number of rounds when v transmitted. Note that both successful as well as unsuccessful (due to collisions) transmissions count. The energetic complexity of the algorithm is defined as $\mathbb{E}\left[\max_{v \in V} \mathcal{E}_v\right]$ for the worst case over all subsets of activated stations. Note that this value is well defined also for deterministic algorithms. Let us stress that $\max_{v \in V} \mathbb{E}\left[\mathcal{E}_v\right] \leq \mathbb{E}\left[\max_{v \in V} \mathcal{E}_v\right]$, and usually the inequality is strict. That is, we look for the expected energetic effort over all stations. Let us note that such measure has been used among others in [17, 21]. On the other hand in some remarkable papers different metrics have been used (e.g., [3, 4, 14]).

Let us note that energetic complexity is closely related to *message complexity* (e.g., [27], Sec. 2). Nevertheless both notions are substantially different. We believe that energetic complexity described above is more adequate for distributed system of battery-supplied devices if only short messages are transmitted.

3 Lower Bound for Uniform Algorithms

In this section we present a lower bound for randomized uniform k-Selection protocols.

3.1 Uniform Algorithms

Definition 1. *Algorithm $\mathcal{A}$ solving k-Selection is called* ***uniform*** *if, and only if, in round i every station that has not yet transmitted successfully, transmits independently with probability p_i (the same for all active stations).*

In particular for k active stations the probability that exactly j stations transmit in the i-th round is[1]

$$\binom{k}{j} (p_i)^j (1 - p_i)^{k-j} \ .$$

Note that p_i may depend on the state of the communication channel in previous rounds.

In general, p_i can be even chosen randomly from some distribution during the execution of the protocol (finally all stations have to use, however, the same value p_i). Due to simplicity and

[1] We assume that the binomial coefficient is 0 for $k < j$ or $j < 0$.

robustness, uniform algorithms are commonly used. For example, algorithms proposed by Martel in [23] and by Anta and Mosteiro in [2] are uniform ones. Such protocols are discussed also in the general MAC settings and are related to ACK-based protocols in other cases (cf. [7]).

3.2 Lower Bound

Before we introduce the key technical lemma let us recall that *selection resolution* (see e.g., [30]) is the problem of obtaining one SINGLE in possibly small number of rounds. More precisely, there are k stations that want to transmit, and the protocol is successfully completed if exactly one station transmits in a round. This problem is in fact equivalent to leader election in a Multiple Access Channel. Let us stress, however, that 1-Selection is a trivial problem that is **not** an instance of a leader election problem. That is, in 1-Selection **only one** station is active and does not have to compete for the channel.

The lemma below shows a relation between time of execution and expected number of collisions.

Lemma 1. *Let $k > 1$ be the number of active stations. If uniform algorithm $\mathcal{A}$ solves selection resolution in the expected time t, then the expected number of rounds with COLLISION during the execution of $\mathcal{A}$ is $\Omega(\frac{1}{t^2})$.*

Proof. Algorithm is uniform, thus in the i-th round each station transmits independently with the same probability p_i. Note however, that in every execution the probabilities $\{p_i\}_{i\geq 1}$ may differ and depend for example on the state of the channel in previous rounds. Let P_i be the random variable denoting the probability of transmission used by stations in the i-th round and let T denote run time of the algorithm and $\mathbb{E}[T] = t$. Algorithm works until first SINGLE appears, thus $P_i = 0$ for every $i > T$. Let B denote the event that there exists i such that $P_i \geq \frac{1}{2kt}$ and let $\bar{B}$ be its complement. We want to show that $P[B] \geq \frac{1}{2}$. Note that if for some i, $P_i < \frac{1}{2kt}$, then

$$P[\text{SINGLE in round } i] = kP_i\left(1-P_i\right)^{k-1} \leq kP_i < \frac{1}{2t},$$

$$P[\text{SILENCE or COLLISION in round } i] > 1 - \frac{1}{2t}.$$

We want to bound the conditional expectation $\mathbb{E}[T|\bar{B}]$. The conditional expectation is well defined, if $P[\bar{B}] > 0$. But if $P[\bar{B}] = 0$ the statement $P[B] \geq \frac{1}{2}$ holds trivially:

$$\mathbb{E}[T|\bar{B}] = \sum_{t'\geq 1} t' P[T = t'|\bar{B}] = \sum_{t'\geq 1} P[T \geq t'|\bar{B}] > \sum_{t'\geq 1}\left(1-\frac{1}{2t}\right)^{t'} = 2t.$$

This implies that $\mathbb{E}[T] = \mathbb{E}[T|\bar{B}]P[\bar{B}] + \mathbb{E}[T|B]P[B] > 2tP[\bar{B}]$. Since $\mathbb{E}[T] = t$, thus $P[\bar{B}] < \frac{1}{2}$, and $P[B] \geq \frac{1}{2}$. Therefore, with probability more than $\frac{1}{2}$, during the execution of the algorithm there exists a slot i_0 with probability of transmission $P_{i_0} \geq \frac{1}{2kt}$.

Now we want to bound probability P_c of COLLISION occurrence in round $i = i_0$. It is clear that $P_c = 1 - (1-P_i)^{k-1}\left(P_i k + (1-P_i)\right)$. Let us prove the following auxiliary lemma

Lemma 2. *For any $n > 2$ and $0 \leq x \leq 1$*

$$(1-x)^n \leq 1 - nx + \frac{1}{2}n(n-1)x^2\ .$$

Proof. It is enough to prove that $f(x) = (1-x)^n - 1 + nx - \frac{1}{2}n(n-1)x^2 \leq 0$ for all $x \in [0,1]$. Note that $f'(x) = n(1-(1-x)^{n-1} - (n-1)x)$. From the Bernoulli inequality we instantly get that $f'(x) \leq 0$. Since $f(0) = 0$ and $f(x)$ is non-increasing on $[0,1]$ the proof is completed. □

Assume that $P_i = \frac{1}{2kt}$. Then for $k > 1$,

$$\begin{aligned} P_c &\geq 1 - \left(1 - \frac{1}{2kt}\right)^{k-1}\left(\frac{1}{2t} + \frac{2kt-1}{2kt}\right) \\ &\geq 1 - \left(1 - \frac{k-1}{2kt} + \frac{(k-1)(k-2)}{8k^2t^2}\right)\left(\frac{k-1}{2kt} + 1\right) \\ &= \frac{(k-1)^2}{4k^2t^2} - \frac{(k-1)(k-2)}{8k^2t^2}\left(\frac{k-1}{2kt} + 1\right) \\ &\geq \frac{(k-1)^2}{8k^2t^2}\left(1 - \frac{k-1}{2kt}\right) \geq \frac{1}{64t^2}. \end{aligned}$$

Note that the second inequality is a consequence of Lemma 2. Since $k > 1$, thus $\frac{k-1}{k} \geq \frac{1}{2}$. We can also assume $t \geq 1$, because any algorithm requires at least one step to solve the *selection resolution*. We proved that if $P_i = \frac{1}{2kt}$, then $P[\text{COLLISION in round } i] \geq \frac{1}{64t^2}$. Obviously, if $P_i \geq \frac{1}{2kt}$, then also $P[\text{COLLISION in round } i] \geq \frac{1}{64t^2}$, because the probability of transmission for each station increases. It follows that with probability at least $\frac{1}{2}$ during any execution of the algorithm there exists a round i_0, where probability of COLLISION is at least $\frac{1}{64t^2}$. This implies that the expected number of COLLISIONs in the algorithm $\mathcal{A}$ is at least $\frac{1}{128t^2} = \Omega(\frac{1}{t^2})$. □

Theorem 1. *Any uniform k-Selection algorithm with expected time of execution $O(k\,\text{polylog}(k))$ has energetic complexity $\Omega\left(\frac{\log k}{\log\log k}\right)$.*

Proof. Let us consider any k-Selection algorithm with expected time of execution $O(k\,\text{polylog}(k))$. We show that the expected number of COLLISIONs during the execution is $\Omega\left(\frac{k}{\text{polylog}(k)}\right)$.

By the i-th *era* we understand the number of rounds between the $(i-1)$-st and i-th successful transmissions for $1 < i \leq k$ (including the round with the i-th transmission). The 1st era is just the number of rounds before the first successful transmission. Let T_i be a random variable denoting duration of i-th era and let T denote the run time of the algorithm. Moreover, let the station that transmitted successfully in i-th era be called i-th transmitter. Clearly $T = \sum_{i=1}^{k} T_i$, thus $\mathbb{E}[T] = \sum_{i=1}^{k} \mathbb{E}[T_i]$. Since $\mathbb{E}[T] = O(k\,\text{polylog}(k))$, there are $\Omega(k)$ eras with the expected duration being $O(\text{polylog}(k))$. From Lemma 1, we know that if era has the expected run time t, the expected number of COLLISIONs is $\Omega\left(\frac{1}{t^2}\right)$. Finally, there are $\Omega(k)$ eras with the expected number of COLLISIONs being $\Omega\left(\frac{1}{\text{polylog}(k)}\right)$. Thus the expected number of COLLISIONs during the execution of the algorithm is $\Omega\left(\frac{k}{\text{polylog}(k)}\right)$. Similarly, during the first $k - \sqrt{k}$ eras the expected number of COLLISIONs is $\Omega\left(\frac{k-\sqrt{k}}{\text{polylog}(k)}\right) = \Omega\left(\frac{k}{\text{polylog}(k)}\right)$.

Since the protocol is uniform, each active station is equally likely to transmit in a round with COLLISION. This can be represented in terms of balls and bins model. More precisely, stations are represented by bins. If COLLISION occurs we throw **one** ball to the bin randomly chosen from bins representing active stations. Clearly the number of balls in the most loaded bin is a lower bound for the number of transmissions of station with maximal number of transmissions[2].

Let us consider a group of the last $\sqrt{k}$ transmitters. All those transmitters are exposed to $\Omega\left(\frac{k}{\text{polylog}(k)}\right)$ COLLISIONs (in expectation). If there is $\Omega\left(\frac{k}{\text{polylog}(k)}\right)$ balls then with high probability $\Omega\left(\frac{\sqrt{k}}{\text{polylog}(k)}\right)$ balls are placed in bins representing the last $\sqrt{k}$ transmitters.

From [28] we have that in case with $m = \frac{\sqrt{k}}{\text{polylog}(k)}$ balls and $n = \sqrt{k}$ bins, the maximum load is $\Omega\left(\frac{\log k}{\log\log k}\right)$ with probability at least $1 - 1/k$. Thus the expected maximum number of transmissions over last $\sqrt{k}$ transmitters is $\Omega\left(\frac{\log k}{\log\log k}\right)$, which ends the proof. □

[2]Note, that each collision affects always more than one station. For simplicity we use, however, only one ball.

3.3 Uniform versus Non-uniform Algorithms

The result presented in the previous subsection implies that there is no uniform k-Selection algorithm working in linear time with maximum energy usage being $o\big(\frac{\log k}{\log\log k}\big)$. However, there are non-uniform algorithms that are more efficient in terms of energy consumption. For example, the initialization algorithm by Nakano and Olariu [25] can be modified in a straightforward manner to obtain a k-Selection algorithm with linear time of execution and no station being awake for more than $O(\log\log k)$ rounds w.h.p. Thus the number of transmissions of each station is $O(\log\log k)$ as well. This is another argument showing that both uniform and non-uniform classes of algorithms are substantially different.

4 Energy Efficient Algorithm

In this section we present a k-Selection algorithm with constant energy consumption and moderate time of execution. The protocol requires $O(k^{1+\epsilon})$ rounds for an arbitrary $\epsilon > 0$, after which w.h.p. all k active stations successfully transmit their messages. More importantly, the energy usage of each station can be bounded by a constant dependent only on ϵ. Therefore, the energetic complexity is $O(1)$, what was the main design goal. Last but not least, the algorithm is flexible in the sense that one can obtain various energy-time trade-offs depending on the needs of particular system.

Our construction is based on the protocol described by Nakano and Olariu in [25]. The algorithm consists of $3 + \lfloor \log_2(1 + \frac{1}{\epsilon}) \rfloor$ iterations. In each iteration, stations that have yet not transmitted successfully try to transmit their messages in one out of $\lceil 2k^{1+\epsilon} \rceil$ rounds. The choice is independent from other stations and uniform over all rounds of a particular iteration. The pseudocode of the protocol is shown in Algorithm 1.

Algorithm 1 Energy Efficient k-Selection

```
1:  max_iter ← 3 + ⌊log_2(1 + 1/ε)⌋                ▷ number of iterations
2:  rounds ← ⌈2k^{1+ε}⌉                            ▷ number of rounds per iteration
3:  iter ← 1
4:  status ← COLLISION
5:  while iter ≤ max_iter and status ≠ SINGLE do
6:      iter ← iter + 1
7:      i ← uniform({1, ..., rounds})              ▷ when to transmit
8:      for round ← 1 to rounds do
9:          if round = i then
10:             status ← transmit(packet)          ▷ try to transmit
```

4.1 Complexity Analysis

It should be clear that the energy usage of any station is at most $max_{iter} = 3 + \lfloor \log_2(1 + \frac{1}{\epsilon}) \rfloor < 3 + \log_2(e)\frac{1}{\epsilon} = O(1)$. Similarly, one can see that the total time of the protocol is $max_{iter} \cdot rounds = O(k^{1+\epsilon})$. The presented algorithm is of Monte Carlo type, which means that with a certain probability, after its execution some stations may fail to transmit. We show that the probability of failure is $O(\frac{1}{k})$.

Lemma 3. *Assume that n stations transmit uniformly and independently in one out of m rounds. For $l \geq 1$:*

- *if $\frac{n(n-1)}{6m} \geq l\log(n)$, then with probability exceeding $1 - \frac{1}{n^l}$, fewer than $\frac{2n(n-1)}{m}$ stations fail to transmit successfully,*
- *if $\frac{n(n-1)}{6m} < l\log(n)$, then with probability exceeding $1 - \frac{1}{n^l}$, fewer than $20\log(n)$ stations fail to transmit successfully.*

Proof. Note that this lemma is a modification of the result of Nakano and Olariu from [25]. Using Corollary 4.2 in [25], if $\frac{n(n-1)}{6m} \geq \log(nf(n))$ for some positive real-valued function $f(n)$, then with probability at least $1 - \frac{1}{nf(n)}$, fewer than $\frac{2n(n-1)}{m}$ stations fail to transmit. Thus, it is sufficient to take $f(n) = n^{l-1}$ to prove the first case. The second case is proved by a simple application of the Lemma 4.3 from [25]. □

Theorem 2. *For any given $\epsilon > 0$, after execution of Algorithm 1 by k stations, all of them transmit successfully with probability at least $1 - O(\frac{1}{k})$.*

Proof. Let us consider what happens after the first iteration of Algorithm 1: there are $n = k$ participating stations and $m = 2k^{1+\epsilon}$ rounds. If $\epsilon \geq 1$, then for sufficiently large k we have $\frac{n(n-1)}{6m} < \log(n)$. Therefore, with probability at least $1 - \frac{1}{k}$, after the first iteration there are at most $20\log(k)$ remaining stations, for which with probability exceeding $1 - \frac{1}{k}$, two additional rounds are sufficient for successful transmission of all remaining stations [25].

On the other hand, if $0 < \epsilon < 1$, then for sufficiently large k we have $\frac{n(n-1)}{6m} > \log(n)$, thus with probability exceeding $1 - \frac{1}{k}$, the first iteration ends with fewer than $\frac{2n(n-1)}{m} < \frac{2k^2}{2k^{1+\epsilon}} = k^{1-\epsilon}$ remaining stations. Inductively, if the i-th iteration starts with at most $k^{1-(2^{i-1}-1)\epsilon}$ stations and $(2^i - 1)\epsilon < 1$, then by Lemma 3 with probability at least $1 - \frac{1}{k}$, after the i-th iteration fewer than $k^{1-(2^i-1)\epsilon}$ stations pass to the $(i+1)$-st iteration. Thus, until $i \geq \log_2(1 + \frac{1}{\epsilon})$, with probability $1 - O(\frac{1}{k})$ after the i-th iteration there are fewer than $k^{1-(2^i-1)\epsilon}$ stations that still need to transmit. After iteration $\lfloor \log_2(1 + \frac{1}{\epsilon}) \rfloor$ we use the second case of Lemma 3, thus w.h.p. the next iteration ends with $O(\log k)$ stations. Therefore, again, two additional rounds are sufficient for successful transmission of each station with probability at least $1 - \frac{1}{k}$. □

5 GMO: Generalized Massey Algorithm

In previous section we introduced an energy efficient algorithm solving the k-Selection problem. It is of Monte Carlo type, i.e., the run time and the maximum energy usage are deterministic, but there is some low yet positive probability of failure[3]. The GMO algorithm presented in this section is of Las Vegas type. That is, the time and energy complexities are random variables but the execution is always successful. The algorithm solves k-Selection problem effectively in polynomial time (with respect to k) such that the expected maximum energy usage is constant.

Algorithm 2 GMO: Generalized Massey Algorithm

```
1:  c ← 0
2:  status ← COLLISION
3:  repeat
4:      if c = 0 then
5:          i ← uniform({0, ..., m − 1})              ▷ round number to transmit in
6:      else
7:          i ← m                                      ▷ sentinel value
8:      for round ← 0 to m − 1 do                      ▷ rounds of current stage
9:          if round = i then
10:             status ← transmit(packet)              ▷ try to transmit
11:         if round ≤ i then
12:             c ← c + collision(round)               ▷ increase c by 1 in case of collision
13:     c ← c − 1
14: until status = SINGLE
```

[3]One could overcome the issue by repeating the execution until the first success. However, the solution is not time efficient, as even two colliding stations enforce a new execution.

In order to ensure that the maximal number of transmissions by a single station is low, we propose an algorithm based on "load balancing" approach. That is, we deviate from uniform-broadcasting paradigm and force stations to broadcast in different rounds. This obviously has impact on time complexity. Due to special construction this side-effect is strictly limited. More precisely, instead of single independent rounds, the algorithm is performed in *stages* consisting of m rounds. Each station that participates in a given stage transmits in exactly one out of m rounds, in order to gain exclusive access to the channel. The parameter m may depend on k.

At the beginning all k stations are active. Stations which are active in a given stage choose independently at random one out of m rounds in which they try to transmit. If there is a collision in a round the algorithm is recursively repeated by only these stations that transmitted in this round – that is, if there is a COLLISION in one round, then transmitting stations are thinned down in consecutive stages. The recursion is stopped when there is no collision, i.e., there is either SILENCE or SINGLE transmission.

It should be noted that the algorithm described above we call GMO since it can be seen as a generalization of Massey algorithm (see [24]). That is, the Massey algorithm is GMO with parameter $m = 2$. In practical settings GMO algorithm can be realized similarly to original Massey algorithm. In particular, stations need to keep a local counter c based on feedback from the channel to determine whether to participate in current stage (see Algorithm 2). Therefore, the execution of the algorithm resembles a depth-first search performed on a dynamically created tree – every time a collision is encountered, a new branch is added, which must be taken into account by subsequent (in DFS sense) nodes. The pseudocode of the algorithm is presented in Algorithm 2.

5.1 Energy Usage and Run Time Analysis

We present here a formal analysis of the GMO properties. The key observation is that each execution of the algorithm can be represented by some random labeled tree. Namely, we consider the so-called *trie* (see [18, 29])) data structure. Let us recall that trie, or a prefix tree, is an ordered tree whose primary purpose is to store a set of strings over some fixed alphabet Σ. Edges of the trie are labeled by symbols from Σ and allow one to search for any words or prefixes, which are represented by leafs or inner nodes, respectively. Figure 1a shows an example of a trie (for detailed introduction to tries and its relevance to various algorithms, see [18]).

For a fixed m-element alphabet $\Sigma_m = \{r_0, r_1, \ldots, r_{m-1}\}$ and a sequence $S = (s_1, s_2, \ldots, s_k) \in (\Sigma_m^*)^k$ of k words over Σ_m, we can construct the corresponding trie $\mathcal{T}_S$. We start with an empty root node, which represents the $S = \emptyset$ case ($k = 0$). If S consists of exactly one word, then it is placed in the root and the construction is done. Otherwise, for $|S| \geq 2$, we partition S into S_{r_0}, S_{r_1},...,$S_{r_{m-1}}$ by considering for each word its first symbol, which is then trimmed. Each non-empty set is used to recursively construct subtries attached to the root. The trimmed symbols are used as labels for new edges (see Figure 1a). As the construction is deterministic the correspondence between the sequence of words S and its trie $\mathcal{T}_S$ is one-to-one.

We show that for a given m, the execution of the GMO algorithm solving k-Selection problem can be represented by some trie of k words over m-element alphabet. Moreover, the mapping is a bijection. We proceed as follows. Every time a station participates in some stage, it chooses one round out of m possible. We encode the choice by a symbol from Σ_m (for example, we can assume that m_i represents choosing ith round). The consecutive choices of the station are concatenated into a single word. In this way, we can represent any particular pass of the GMO with k stations as a sequence of k words $S = (s_1, s_2, \ldots, s_k)$. Obviously, there exists a unique trie $\mathcal{T}_S$ which represents the execution.

We notice that the level of a specified leaf in the trie $\mathcal{T}_S$ built from the words of S is equal to the energy used by some station and as a result, height of the trie is the maximum energy used by a station. Similarly, the number of stages of the algorithm is equal to the total number of inner nodes in the trie (cf. Figures 1b and 1c). Of course, to make use of the above correspondence, we need to consider tries constructed in a random fashion. Fortunately, the properties of random tries have been extensively studied. Clement et al. [8] consider, among other parameters, height and size expectation of tries embedded in a very general framework of dynamical sources. We make use of the

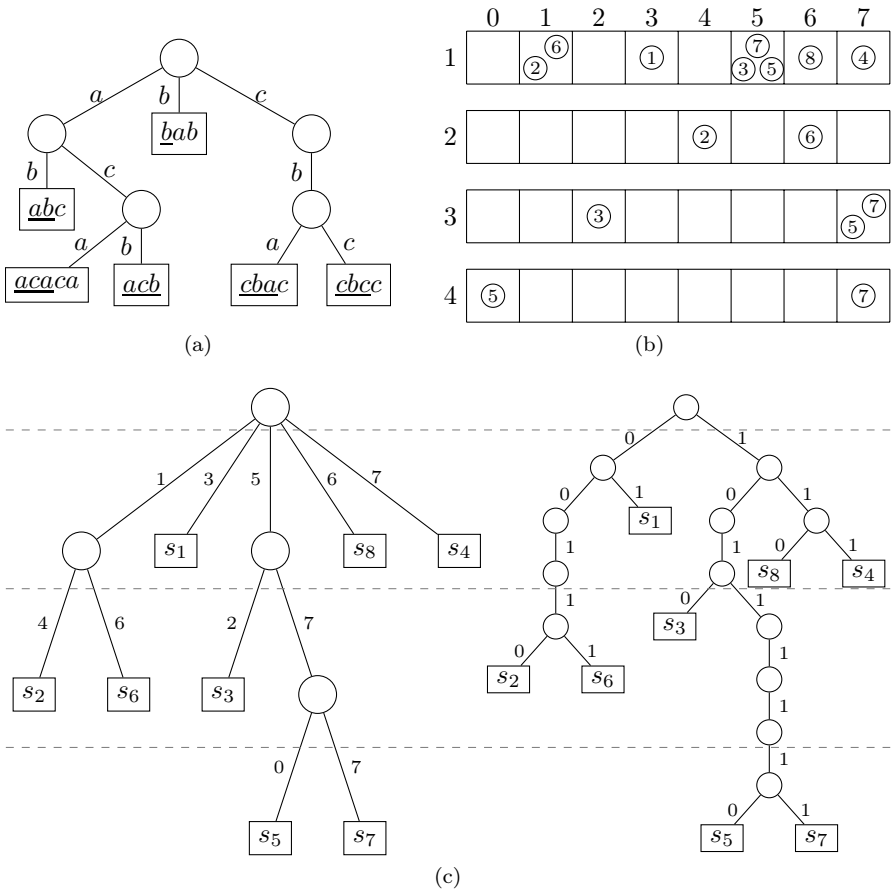

Figure 1: (a) Exemplary trie, built from set of words $\{abc, acaca, acb, bab, cbac, cbcc\}$ over alphabet $\{a, b, c\}$. (b) Example run of GMO algorithm with $k = 8$ stations working for 4 stages, each consisting of $m = 8$ rounds. Numbers in circles represent stations. (c) Trie representation of the same pass of algorithm and its binary equivalent.

following fact presenting more classical results for symmetric tries[4] (proofs and further references can be found in [11] and [18]):

Fact 1. *Let $\mathcal{T}_{k,m}$ denote a random symmetric m-ary trie built from k words. Then*

1. *the expected value of binary trie's height is $\mathbb{E}[H(\mathcal{T}_{k,2})] = 2\log_2 k + O(\log\log k)$,*
2. *the expected value of m-ary trie's size, for constant m, is $\mathbb{E}[S(\mathcal{T}_{k,m})] = \frac{k}{\ln m}(1 + \delta(k)) + O(1)$, where $\delta(k)$ is a function with small module.*

While properties of random tries have been well studied, to the best of our knowledge there are no results regarding the case when alphabet cardinality m may depend on the number of words k. Therefore, we first show how to reduce an m-ary trie to a binary one, and then we make use of Fact 1. This allows us to prove the following theorems.

Theorem 3. *The expected maximum energy usage of GMO performed by k stations with stages of length m satisfies*

$$\frac{2\log_2 k + \delta(k)}{\lceil \log_2 m \rceil} \leq \mathbb{E}[\mathcal{E}_{max}] \leq \frac{2\log_2 k + \delta(k)}{\lfloor \log_2 m \rfloor} + 1,$$

[4]The symmetric trie is a trie constructed from random words, where each letter is sampled uniformly and independently.

where $\delta(k)$ is some $O(\log\log k)$ function.

Proof. We will use coupling argument in order to show correspondence between m-ary and binary tries. Let us assume that station i contains some random real number x_i sampled uniformly (and independently from other stations) from $[0,1)$. The number is used by ith station as a source of randomness. That is, we represent x_i in base m and for consecutive calls to uniform() function by the ith station, the next digit of x_i is returned. For a given k-element set of stations we run simultaneously two stochastically dependent copies of the algorithm – the first one with m rounds in a stage and the second one with exactly two rounds per stage. In addition, we require that in both passes corresponding stations begin with the same random number. Obviously, the numbers represented in base m are equivalent to the words from Σ_m^*. For any sequence of words we have unique tries, both in m-ary and binary case. Let us assume that $m = 2^d$ for some integer $d \in \mathbb{N}^+$. Then, for tries $\mathcal{T}_{k,m}$ and $\mathcal{T}_{k,2}$ representing respectively the first and the second run, we have:

$$H(\mathcal{T}_{k,m}) = \left\lfloor \frac{H(\mathcal{T}_{k,2}) - 1}{d} \right\rfloor + 1, \tag{1}$$

where $H(\cdot)$ denotes height of trie. Indeed, for $k < 2$ the equality is trivial to be checked, and for $k \geq 2$ there are at least two numbers (i.e., words of digits) in $\mathcal{T}_{k,m}$ having first $H(\mathcal{T}_{k,m}) - 1$ digits the same, so there are two words in $\mathcal{T}_{k,2}$ with at least $d \cdot (H(\mathcal{T}_{k,m}) - 1)$ bits of common prefix. Conversely, there are two words in $\mathcal{T}_{k,2}$ with the longest common prefix of length $H(\mathcal{T}_{k,2}) - 1$. Thus, the first $d \cdot \left\lfloor \frac{H(\mathcal{T}_{k,2})-1}{d} \right\rfloor$ bits encode prefix of length $\left\lfloor \frac{H(\mathcal{T}_{k,2})-1}{d} \right\rfloor$ digits of at least two words in $\mathcal{T}_{k,m}$. Therefore, we have

$$\left\lfloor \frac{H(\mathcal{T}_{k,2}) - 1}{d} \right\rfloor \leq H(\mathcal{T}_{k,m}) - 1 \leq \frac{H(\mathcal{T}_{k,2}) - 1}{d}.$$

By (1) and floor function property $x - 1 < \lfloor x \rfloor \leq x$, we have

$$\frac{\mathbb{E}\left[H(\mathcal{T}_{k,2})\right] - 1}{d} < \mathbb{E}\left[H(\mathcal{T}_{k,m})\right] \leq \frac{\mathbb{E}\left[H(\mathcal{T}_{k,2})\right] - 1}{d} + 1.$$

To get rid of $m = 2^d$ constraint, we make use of the fact that the expected value of height of trie is non-increasing with respect to the cardinality of alphabet. Thus,

$$\mathbb{E}\left[H(\mathcal{T}_{k,2^{\lceil \log_2 m \rceil}})\right] \leq \mathbb{E}\left[H(\mathcal{T}_{k,m})\right] \leq \mathbb{E}\left[H(\mathcal{T}_{k,2^{\lfloor \log_2 m \rfloor}})\right]$$

and both, lower and upper bounds apply to the already considered case. We use asymptotic formula for height from Fact 1, which ends the proof. □

Theorem 4. *For $m = k^\alpha, \alpha > 0$, GMO solves k-Selection with the expected number of rounds*

$$\mathbb{E}\left[T_k\right] = O\big(k^{\max\{\alpha(\lfloor \frac{1}{\alpha} \rfloor + 1), 2 - \alpha\lfloor \frac{1}{\alpha} \rfloor\}}\big).$$

Proof. In the GMO there is a one-to-one correspondence between stages and collisions of stations. That is, each collision results in additional stage (extra m rounds) and, conversely, each stage (except the first) is caused by exactly one collision. Therefore it suffices to count the number of collisions. If two or more stations collide in a given round, then all of them have the same energy usage at the moment of the collision. We partition collision rounds into disjoint levels by considering energy usage of stations at the beginning of a round (the levels correspond to the levels of execution trie). Let X_r be a random variable denoting the number of collisions in r-th level, so the total number of collisions is $X = \sum_{r=0}^{\infty} X_r$ and $T_k = mX$. There are m^r possible rounds on the r-th level, represented by words of length r over Σ_m. For a word $\sigma \in \Sigma_m^r$ let Y_σ be a random indicator variable denoting whether in round represented by σ there is collision ($Y_\sigma = 1$) or not[5] ($Y_\sigma = 0$). We have $X_r = \sum_\sigma Y_\sigma$ where the sum is taken over all words $\sigma \in \Sigma_m^r$ of length r. Moreover, as rounds are chosen uniformly, for a given r, all Y's have the same distribution denoted by $\mathcal{Y}_r$. If some stations

[5]Particularly, in a given execution a round represented by σ may not physically exist.

collide on the r-th level of the execution trie, they must have collided with each other on all of the previous levels. At a given level all colliding stations sampled the same round number out of m possible. Therefore,

$$\begin{aligned}\mathbb{E}[Y_\sigma] &= \Pr(Y_\sigma = 1) = \Pr(\text{at least two stations sampled } \sigma)\\ &= 1 - \Pr(\text{none of stations sampled } \sigma) - \Pr(\text{exactly one station sample } \sigma)\\ &= 1 - \Big(1 - \frac{1}{m^r}\Big)^k - \frac{k}{m^r}\Big(1 - \frac{1}{m^r}\Big)^{k-1}.\end{aligned}$$

By linearity of expectation,

$$\begin{aligned}\mathbb{E}[X_r] &= \sum_{\sigma \in \Sigma_m^r} \mathbb{E}[Y_\sigma] = m^r \mathbb{E}[\mathcal{Y}_r]\\ &= m^r\Big(1 - \Big(1 - \frac{1}{m^r}\Big)^k - \frac{k}{m^r}\Big(1 - \frac{1}{m^r}\Big)^{k-1}\Big)\\ &= k^{\alpha r} - k^{\alpha r}\Big(1 - \frac{1}{k^{\alpha r}}\Big)^k - k\Big(1 - \frac{1}{k^{\alpha r}}\Big)^{k-1}.\end{aligned}$$

Hence

$$\begin{aligned}\mathbb{E}[X] &= \sum_{r=0}^{\infty}\Big(k^{\alpha r} - k^{\alpha r}\Big(1 - \frac{1}{k^{\alpha r}}\Big)^k - k\Big(1 - \frac{1}{k^{\alpha r}}\Big)^{k-1}\Big)\\ &= \sum_{r=0}^{\lfloor \frac{1}{\alpha}\rfloor}\Big(k^{\alpha r} - k^{\alpha r}\Big(1 - \frac{1}{k^{\alpha r}}\Big)^k - k\Big(1 - \frac{1}{k^{\alpha r}}\Big)^{k-1}\Big)\\ &\quad + \sum_{r=\lfloor \frac{1}{\alpha}\rfloor+1}^{\infty}\Big(k^{\alpha r} - k^{\alpha r}\Big(1 - \frac{1}{k^{\alpha r}}\Big)^k - k\Big(1 - \frac{1}{k^{\alpha r}}\Big)^{k-1}\Big)\\ &\le \sum_{r=0}^{\lfloor \frac{1}{\alpha}\rfloor} k^{\alpha r} + \sum_{r=\lfloor \frac{1}{\alpha}\rfloor+1}^{\infty}\Big(k^{\alpha r} - k^{\alpha r}\Big(1 - k^{1-\alpha r}\Big) - k\Big(1 - k^{1-\alpha r}\Big)\Big)\\ &= O(k^{\alpha\lfloor \frac{1}{\alpha}\rfloor}) + O(k^{2-\alpha(\lfloor \frac{1}{\alpha}\rfloor+1)}) = O(k^{\alpha\lfloor \frac{1}{\alpha}\rfloor} + k^{2-\alpha(\lfloor \frac{1}{\alpha}\rfloor+1)}).\end{aligned}$$

Multiplying the result by $m = k^\alpha$ ends the proof. □

Remark 1. *If m is a constant, i.e, it does not depend on k, then by Fact 1 the expected run time of GMO is*

$$\mathbb{E}[T_{k,m}] = \frac{km}{\ln m}(1 + \delta(k)) + O(1),$$

where $\delta(k)$ is a function with small module.

Finally, to summarize the analysis of the GMO we state the following corollary:

Corollary 1. *For $m = k^\alpha$ with $\alpha > 0$ the GMO solves k-Selection problem in the expected number of $\mathbb{E}[T] = O\big(k^{\max\{\alpha(\lfloor \frac{1}{\alpha}\rfloor+1),\,2-\alpha\lfloor \frac{1}{\alpha}\rfloor\}}\big)$ rounds with the expected maximum energy usage $\mathbb{E}[\mathcal{E}_{max}] \le \frac{2}{\alpha} + 1 + o(1)$.*

6 Energetic Complexity of Martel Algorithm

One of the motivations for constructing more efficient algorithms in terms of energy usage was the fact that common algorithms solving k-Selection neglected the problem of energy expenditure. Moreover, our study shows that many time-efficient algorithms perform poorly in means of maximum energy usage. As an example, we present in this section an analysis of the Martel algorithm that has optimal expected run time in the considered model.

In simple terms, the k-Selection algorithm by Martel [23] exploits the idea that the problem may be solved by running k times a selection resolution algorithm. That is, we start with k stations, which compete together in order to elect the leader. The selected station successfully broadcasts its message. Next, the execution continues recursively with $k-1$ stations until each one becomes a leader. It is easily seen that by listening to the channel the stations are aware of the number of stations that are still active. The knowledge is then used in a leader election subprocedure, where for m active stations they try to transmit with the probability $p = \frac{1}{m}$. Such p guarantees the optimal probability of obtaining the single transmission. However, the approach proved not to be efficient in terms of the energy usage.

Theorem 5. *Algorithm solving k-Selection proposed by Martel has expected energetic complexity* $\Omega(\log k)$.

Proof. Let us denote the time between the $(i-1)$-st and i-th occurrence of SINGLE in Martel algorithm, as i-th era. The first era is the time until the first SINGLE appears. Let X_i be a random variable denoting duration of the i-th era. We need to show that in Martel algorithm, for all $i > 1$, $\mathbb{E}[X_i] = O(1)$. But this fact is proved by Martel in [23],(Lemmas 2.1, 2.2 and 2.3). If for some i, $\mathbb{E}[X_i] = O(1)$, then from Lemma 1, the expected number of COLLISIONs in the i-th era is $\Omega(1)$. Consider energy consumption of the last transmitting station (i.e., the k-th one) denoted as $\mathcal{E}_{v_{last}}$. Station v_{last} has a chance to participate in COLLISION in each era. If the i-th era's expected number of COLLISIONs is δ, then, since the algorithm is uniform, each active station in this era has equal chance to participate in COLLISION. Expected energy consumption of each active station in this round is at least $\frac{2\delta}{k-i+1}$. Thus

$$\mathbb{E}[\mathcal{E}_{v_{last}}] = \Omega\left(\sum_{i=2}^{k} \frac{1}{k-i+1}\right).$$

We note that $\sum_{i=2}^{k} \frac{1}{k-i+1} = H_{k-1}$ is $(k-1)$-st harmonic number. Let us recall that harmonic number $H_n = \log n + \gamma + O(\frac{1}{n})$, where $\gamma = 0.57721\ldots$ is the Euler-Mascheroni constant. Finally, $\mathbb{E}[\mathcal{E}_{v_{last}}] = \Omega(\log k)$. It is clear that $\mathbb{E}[\mathcal{E}_{v_{last}}] \leq \mathbb{E}[\mathcal{E}_{max}]$. □

7 Lower Bounds for Deterministic Oblivious Algorithms

In this section we investigate oblivious, deterministic k-Selection protocols. This means that schedule of transmissions for each station is defined before execution of the algorithm. That is, each station knows if it shall transmit in each round before the algorithm is started. In particular, decision about transmission does not depend on the state of the communication channel in previous rounds. Thus the algorithm can be viewed as an assignment of binary vectors to stations. Since the protocol is executed in an environment common to all stations, i.e., single hop network, the message is successfully received only if exactly one station is transmitting at a given time.

For every station $v \in V$ we denote by $w(v)$ the binary vector in the following way. Let $w(v)_i$ denote i-th position in the vector $w(v)$. If station v is transmitting in round i, then $w(v)_i = 1$, otherwise $w(v)_i = 0$.
Let $\mathcal{C} = \{c_1, c_2, \ldots, c_n\}$ be a set of binary words (a code) of length t. The number of vectors n is the size of the code. Given k words $c_{i_1}, c_{i_2}, \ldots, c_{i_k}$, we define the sum of vectors $c_{i_1} \vee c_{i_2} \vee \cdots \vee c_{i_k}$ as a bitwise Boolean sum.

Definition 2. *We say that binary vector v covers vector v' if for every i*

$$v_i \geq v'_i \ .$$

In other words, v covers v' if for each coordinate with value 1 in v', the corresponding coordinate in v is also 1. Below we recall the definition of *superimposed codes* introduced by Kautz and Singleton in [16].

Definition 3. *Let r be a positive integer. We say that set of binary words $\mathcal{C} = \{c_1, c_2, \ldots, c_n\}$ is r-superimposed code if for any distinct words $c_{i_0}, c_{i_1}, c_{i_2}, \ldots, c_{i_r}$, the word c_{i_0} is **not** covered by $c_{i_1} \vee c_{i_2} \vee \cdots \vee c_{i_r}$.*

From the definition of the k-Selection protocol we have following fact:

Fact 2. *Algorithm solves the k-Selection problem if and only if the corresponding set of vectors is a $(k-1)$-superimposed code.*

Indeed, there is a $1-1$ correspondence between the superimposed codes and oblivious k-Selection algorithms pointed in [5]. In [10] Erdős, Frankl and Füredi proved a theorem about families of sets which has direct application in superimposed codes.

Fact 3 (see [10, Proposition 2.1])**.** *Let $f_k(t, \varepsilon)$ be the maximum size of the k-superimposed code of length t, where each codeword has exactly ε ones, then $f_k(t, \varepsilon) \leq \binom{t}{\lceil \frac{\varepsilon}{k} \rceil} / \binom{\varepsilon - 1}{\lceil \frac{\varepsilon}{k} \rceil - 1}$.*

Lower bound on length of k-superimposed codes implies lower bound on time complexity of any oblivious, deterministic k-Selection algorithms. Using techniques similar to those in [10] we can bound the size of any k-superimposed code with restricted number of ones in codewords. In effect we can bound the time complexity of any oblivious k-Selection algorithm with energetic complexity $\mathcal{E}_{max}$.
Let us recall the following, well-known fact:

Fact 4. *The binomial coefficient $\binom{n}{k}$ satisfies*

$$\frac{n^k}{k^k} \leq \binom{n}{k} \leq \frac{n^k}{k!}.$$

Theorem 6. *Run time t of any deterministic, oblivious algorithm solving k-Selection with energetic complexity $\mathcal{E}_{max}$ satisfies*

$$t \in \Omega\left(\mathcal{E}_{max}\left(\frac{n}{(k-1)^2}\right)^{\frac{1}{\lceil \frac{\mathcal{E}_{max}}{k-1} \rceil}}\right).$$

Proof. We assume that $k, t, \mathcal{E}_{max}$ are functions of n, and n tends to infinity. First, we want to prove that the relation $n \leq \sum_{\mathcal{E}=1}^{\mathcal{E}_{max}} f_{k-1}(t, \mathcal{E})$, must hold for every deterministic, oblivious algorithm solving k-Selection with run time t, and the maximum energy consumption $\mathcal{E}_{max}$. We can partition vectors into groups of the same Hamming weight $W_1, W_2, \ldots$, i.e., $w(v) \in W_i$ if $h(w(v)) = i$, where $h(w)$ is Hamming weight of the vector w. Clearly, the set $W = \{w(v) : v \in V\}$ is a $(k-1)$-superimposed code, because algorithm solves k-Selection. Thus, each set W_i is also $(k-1)$-superimposed. From the definition of the function f_k, $|W_i| \leq f_{k-1}(t, i)$. On the other hand $n = |W| = |\bigcup W_i| = \sum_{i=1}^{\mathcal{E}_{max}} |W_i|$. Thus, $n \leq \sum_{\mathcal{E}=1}^{\mathcal{E}_{max}} f_{k-1}(t, \mathcal{E})$. From Fact 3 we obtain

$$f_{k-1}(t, \mathcal{E}) \leq \binom{t}{\left\lceil \frac{\mathcal{E}}{k-1} \right\rceil} \Big/ \binom{\mathcal{E} - 1}{\left\lceil \frac{\mathcal{E}}{k-1} \right\rceil - 1}.$$

Using an identity for the binomial coefficient and applying Fact 4 we have:

$$\begin{aligned}
\binom{\mathcal{E} - 1}{\left\lceil \frac{\mathcal{E}}{k-1} \right\rceil - 1} &= \frac{\mathcal{E} - \left\lceil \frac{\mathcal{E}}{k-1} \right\rceil + 1}{\mathcal{E}} \binom{\mathcal{E}}{\left\lceil \frac{\mathcal{E}}{k-1} \right\rceil - 1} \\
&\geq \left(1 - \frac{1}{k-1}\right) \left(\frac{\mathcal{E}}{\left\lceil \frac{\mathcal{E}}{k-1} \right\rceil - 1}\right)^{\left\lceil \frac{\mathcal{E}}{k-1} \right\rceil - 1} \\
&\geq \left(1 - \frac{1}{k-1}\right) (k-1)^{\left\lceil \frac{\mathcal{E}}{k-1} \right\rceil - 1}.
\end{aligned}$$

Again we apply Fact 4 directly to the sum $\sum_{\mathcal{E}=1}^{\mathcal{E}_{max}} f_{k-1}(t,\mathcal{E})$ and we make use of the inequality obtained above. It is easy to see that the following observations are satisfied:

$$\begin{aligned} n &\leq \sum_{\mathcal{E}=1}^{\mathcal{E}_{max}} \frac{\binom{t}{\lceil \frac{\mathcal{E}}{k-1} \rceil}}{\left(1-\frac{1}{k-1}\right)(k-1)^{\lceil \frac{\mathcal{E}}{k-1} \rceil - 1}} \\ &\leq \left(1+\frac{1}{k-2}\right) \sum_{\mathcal{E}=1}^{\mathcal{E}_{max}} \frac{t^{\lceil \frac{\mathcal{E}}{k-1} \rceil}}{\left\lceil \frac{\mathcal{E}}{k-1} \right\rceil ! \, (k-1)^{\lceil \frac{\mathcal{E}}{k-1} \rceil - 1}} \\ &= \left(1+\frac{1}{k-2}\right)(k-1) \sum_{\mathcal{E}=1}^{\mathcal{E}_{max}} \frac{\left(\frac{t}{k-1}\right)^{\lceil \frac{\mathcal{E}}{k-1} \rceil}}{\left\lceil \frac{\mathcal{E}}{k-1} \right\rceil !}. \end{aligned}$$

Since $\mathcal{E}$ occurs in the above sum only in term $\left\lceil \frac{\mathcal{E}}{k-1} \right\rceil$, we have the same $(k-1)$ summands. Thus,

$$\begin{aligned} n &\leq \left(1+\frac{1}{k-2}\right)(k-1)^2 \sum_{s=1}^{\lceil \frac{\mathcal{E}_{max}}{k-1} \rceil} \frac{\left(\frac{t}{k-1}\right)^s}{s!} \\ &\leq \left(1+\frac{1}{k-2}\right)(k-1)^2 \frac{\Gamma\left(\left\lceil \frac{\mathcal{E}_{max}}{k-1} \right\rceil + 1, \frac{t}{k-1}\right)}{\left\lceil \frac{\mathcal{E}_{max}}{k-1} \right\rceil ! e^{-\frac{t}{k-1}}}, \end{aligned}$$

where $\Gamma(s,x) \stackrel{df}{=} \int_x^\infty t^{s-1} e^{-t} \, \mathrm{d}t$ is the incomplete gamma function. For $s \in \mathbb{N}_+$ the function $\Gamma(s,x)$ has the following expansion $\Gamma(s,x) = (s-1)! \, e^{-x} \sum_{k=0}^{s-1} \frac{x^k}{k!}$. From [1] we know that $\frac{\Gamma(s,x)}{x^{s-1}e^{-x}} \to 1$ as $x \to \infty$. But it is proved in [10] that $t = \Omega(k \log n)$, even without energy restriction. Thus $\frac{t}{k} = \Omega(\log n)$. From asymptotic behavior of $\Gamma(s,x)$, we know that:

$$\Gamma\left(\left\lceil \frac{\mathcal{E}_{max}}{k-1} \right\rceil + 1, \frac{t}{k-1}\right) = O\left(\left(\frac{t}{k-1}\right)^{\lceil \frac{\mathcal{E}_{max}}{k-1} \rceil} e^{-\frac{t}{k-1}}\right).$$

From the fact that $n! \geq (\frac{n}{e})^n$ after some simplifications we get:

$$\begin{aligned} \frac{\left(\frac{t}{k-1}\right)^{\lceil \frac{\mathcal{E}_{max}}{k-1} \rceil} e^{-\frac{t}{k-1}}}{\left\lceil \frac{\mathcal{E}_{max}}{k-1} \right\rceil ! e^{-\frac{t}{k-1}}} &\leq \frac{\left(\frac{t}{k-1}\right)^{\lceil \frac{\mathcal{E}_{max}}{k-1} \rceil}}{\left(\frac{\lceil \frac{\mathcal{E}_{max}}{k-1} \rceil}{e}\right)^{\lceil \frac{\mathcal{E}_{max}}{k-1} \rceil}} \\ &\leq e^{\lceil \frac{\mathcal{E}_{max}}{k-1} \rceil} \left(\frac{t}{\mathcal{E}_{max}}\right)^{\lceil \frac{\mathcal{E}_{max}}{k-1} \rceil}. \end{aligned}$$

From calculations above we obtain the following facts

$$\begin{aligned} n &= O\left((k-1)^2 \left(\frac{et}{\mathcal{E}_{max}}\right)^{\lceil \frac{\mathcal{E}_{max}}{k-1} \rceil}\right), \\ (k-1)^2 &\left(\frac{et}{\mathcal{E}_{max}}\right)^{\lceil \frac{\mathcal{E}_{max}}{k-1} \rceil} = \Omega(n), \end{aligned}$$

and finally $t = \Omega\left(\mathcal{E}_{max} \left(\frac{n}{(k-1)^2}\right)^{\frac{1}{\lceil \frac{\mathcal{E}_{max}}{k-1} \rceil}}\right)$. □

The above theorem yields a spectrum of time–energy complexity trade-offs for oblivious, deterministic k-Selection algorithms. For example, it implies the following corollary.

Table 1: Simulation results of Protocol 1 for $k = 10$ stations.

ϵ	1	1/2	1/4	1/128
max_{iter}	4	4	5	10
$time$	800	256	180	210
$iter_1$	0.44242	1.32249	2.24018	3.55332
$iter_2$	0.00258	0.038901	0.153025	0.548462
$iter_3$	0	0.000628	0.005062	0.038523
$iter_4$	0	0.00002	0.000134	0.002036
$iter_5$			0.000002	0.000096
$iter_6$				0.000008
$iter_7$				0
$iter_8$				0
$iter_9$				0
$iter_{10}$				0
$failed$	0	10	1	0

Corollary 2. *The run time t of any deterministic, oblivious algorithm solving k-Selection for $k = O\left(n^{1/4}\right)$ with energetic complexity $\mathcal{E}_{max} = O\left(k\frac{\log n}{\alpha \log\log n}\right)$ satisfies, for any $\alpha > 0$,*

$$t = \Omega\left(k\frac{\log^{1+\frac{\alpha}{2}} n}{\alpha \log\log n}\right).$$

8 Experimental Results

In addition to the analysis of the protocols presented for random model, we show empirical results obtained by means of computer simulations. We have evaluated the performance of Protocol 1 for networks consisting of $k = 10$ and $k = 10^4$ activated stations. The results allow us to speculate on tightness of the analysis, as well as to see how the protocol behaves in the case of a small number of activated stations. We have also run simulations of Martel algorithm to compare the difference between maximum energy usage and the energetic effort of the last station.

8.1 Energy Efficient Protocol

Table 1 shows results of simulations of Protocol 1 solving 10-Selection problem for different values of the ϵ parameter. The $time = max_{iter} \cdot \lceil 2k^{1+\epsilon} \rceil$ is a total number of rounds needed by the protocol to complete. The number of stations left activated after consecutive iterations, $iter_i$, was obtained by averaging outcomes of 10^6 simulation runs. The last row shows how many (out of 10^6) runs ended with failure, which is a case when after max_{iter} iterations there are some stations that were unable to broadcast their messages.

Results in Table 2 were obtained in a similar manner as for Table 1, but for $k = 10^4$ stations and 10^5 simulation runs. It can be seen that the Protocol 1 behaves much better for larger number of stations, as one could expect based on the results of the analysis.

8.2 GMO Algorithm

We simulated GMO for $k = 10^4$ stations and $m = k^\alpha$ rounds per stage to compare our theoretical result of Theorem 4. Figure 2 shows the exponent of run time. One can notice that for $0 < \alpha < 1$, w.h.p. levels up to $\lceil \frac{1}{\alpha} \rceil - 1$ are full as there are significantly less rounds than stations. Therefore, we consider a simple modification of GMO, such that stations skip the first $\lceil \frac{1}{\alpha} \rceil - 1$ levels. That is,

Table 2: Simulation results of Protocol 1 for $k = 10^4$ stations.

ϵ	1	1/2	1/4	1/128
max_{iter}	4	4	5	10
$time$	800000000	8000000	1000000	214930
$iter_1$	0.4965	49.8927	487.765	3720.12
$iter_2$	0	0.0013	1.1899	591.196
$iter_3$	0	0	0	16.0588
$iter_4$	0	0	0	0.01214
$iter_5$			0	0
$iter_6$				0
$iter_7$				0
$iter_8$				0
$iter_9$				0
$iter_{10}$				0
$failed$	0	0	0	0

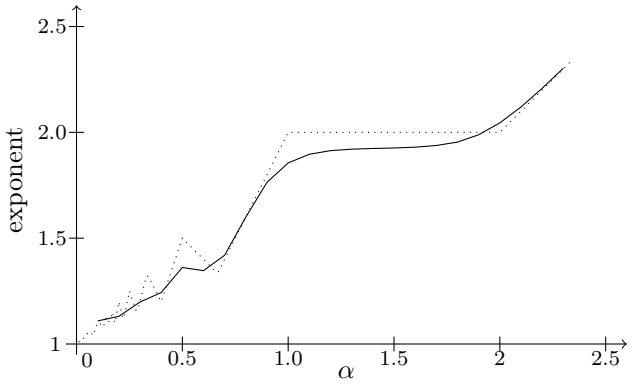

Figure 2: Run time of GMO for $m = k^\alpha$ rounds per stage. Solid line shows the exponent $\log_k(T_{k,m})$ of the average run time for $k = 10^4$ stations and 10^4 simulations. Dotted line shows the theoretical exponent $\max\{\alpha(\lfloor\frac{1}{\alpha}\rfloor + 1), 2 - \alpha\lfloor\frac{1}{\alpha}\rfloor\}$ established in Theorem 4.

instead of transmitting, stations may assume that there was collision. The result of this modification is presented in Figure 3.

8.3 Martel Algorithm

In Section 6 we proved $\Omega(\log k)$ lower bound on energy usage of the last station in the Martel algorithm. While this result obviously translates to the lower bound of the energetic complexity of the algorithm, one could ask how big is the difference between maximum energy usage and the energetic effort of the last station. Figure 4 shows results of 10^5 simulations for different number of stations (logarithmic scale).

9 Conclusions and Further Research

In our paper we presented several results about energetic aspects of k-Selection protocols in a single-hop radio network. We believe that presented approach can be applied to more realistic scenarios. In particular, it is clear that some results can be easily applied for dynamic counterparts of k-Selection

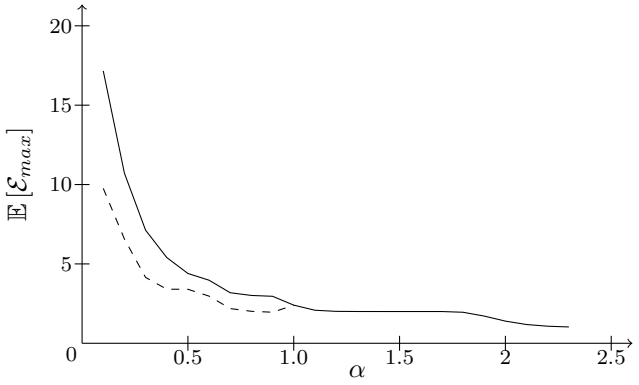

Figure 3: The maximum energy usage of GMO for $m = k^\alpha$ rounds per stage. Solid line shows the average for $k = 10^4$ stations and 10^4 simulations. Dashed line shows the maximum energy usage of modified algorithm, where we skip first $\lceil \frac{1}{\alpha} \rceil - 1$ levels.

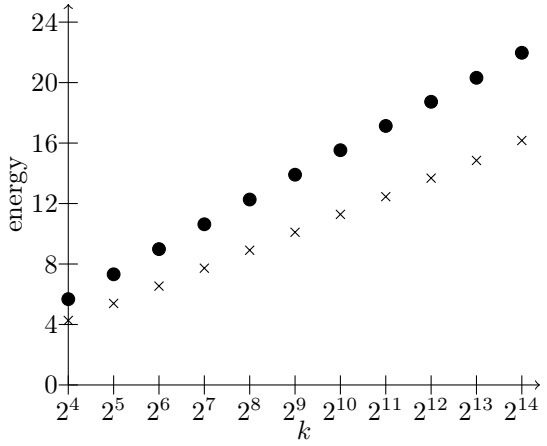

Figure 4: Energy usage in Martel algorithm. • is an average maximum energy usage and × is an average energy usage of the last station.

problem (described, e.g., in [20]) at least for some models.

We believe that most interesting and most challenging task is to find a general relation between energy consumption and time necessary for completion of k-Selection in randomized model, especially when we take into account transmission as well as listening. We tried to obtain such result, without effects, using information theory approach techniques.

References

[1] Milton Abramowitz and Irene A. Stegun. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. Dover, New York, ninth dover printing, tenth gpo printing edition, 1964.

[2] Antonio Fernández Anta and Miguel A. Mosteiro. Contention resolution in multiple-access channels: k-selection in radio networks. In My T. Thai and Sartaj Sahni, editors, *COCOON*, volume 6196 of *Lecture Notes in Computer Science*, pages 378–388. Springer, 2010.

[3] Amol Bakshi and Viktor K. Prasanna. Energy-efficient communication in multi-channel single-hop sensor networks. In *ICPADS*, pages 403–410. IEEE Computer Society, 2004.

[4] Paul Balister, Béla Bollobás, Animashree Anandkumar, and Alan S. Willsky. Energy-latency tradeoff for in-network function computation in random networks. In *INFOCOM*, pages 1575–1583. IEEE, 2011.

[5] Annalisa De Bonis and Ugo Vaccaro. Constructions of generalized superimposed codes with applications to group testing and conflict resolution in multiple access channels. *Theor. Comput. Sci.*, 306(1-3):223–243, 2003.

[6] Jacir Luiz Bordim, JiangTao Cui, Naohiro Ishii, and Koji Nakano. Doubly-logarithmic energy-efficient initialization protocols for single-hop radio networks. In *IPDPS*. IEEE Computer Society, 2002.

[7] Bogdan S. Chlebus, Dariusz R. Kowalski, and Mariusz A. Rokicki. Adversarial queuing on the multiple access channel. *ACM Transactions on Algorithms*, 8(1):5, 2012.

[8] J. Clément, Ph. Flajolet, and B. Vallée. Dynamical sources in information theory: A general analysis of trie structures. *Algorithmica*, 29:307–369, 2001.

[9] Andrea E. F. Clementi, Angelo Monti, and Riccardo Silvestri. Selective families, superimposed codes, and broadcasting on unknown radio networks. In S. Rao Kosaraju, editor, *SODA*, pages 709–718. ACM/SIAM, 2001.

[10] P. Erdős, P. Frankl, and Z. Füredi. Families of finite sets in which no set is covered by the union of r others. *Israel Journal of Mathematics*, 51:79–89, 1985.

[11] Philippe Flajolet. On the performance evaluation of extendible hashing and trie searching. *Acta Informatica*, 20:345–369, 1983.

[12] Albert G. Greenberg and Shmuel Winograd. A lower bound on the time needed in the worst case to resolve conflicts deterministically in multiple access channels. *Journal of the ACM*, 32(3):589–596, July 1985.

[13] J.F. Hayes. An adaptive technique for local distribution. *IEEE Trans. Commun.*, pages 1178–1186, 1978.

[14] Stephan Holzer, Yvonne Anne Pignolet, Jasmin Smula, and Roger Wattenhofer. Monitoring churn in wireless networks. *Theor. Comput. Sci.*, 453:29–43, 2012.

[15] Tomasz Jurdzinski, Miroslaw Kutylowski, and Jan Zatopianski. Energy-efficient size approximation of radio networks with no collision detection. In Oscar H. Ibarra and Louxin Zhang, editors, *COCOON*, volume 2387 of *Lecture Notes in Computer Science*, pages 279–289. Springer, 2002.

[16] W. Kautz and R. Singleton. Nonrandom binary superimposed codes. *Information Theory, IEEE Transactions on*, 10(4):363 – 377, oct. 1964.

[17] Marek Klonowski, Miroslaw Kutylowski, and Jan Zatopianski. Energy efficient alert in single-hop networks of extremely weak devices. In Shlomi Dolev, editor, *ALGOSENSORS*, volume 5804 of *Lecture Notes in Computer Science*, pages 139–150. Springer, 2009.

[18] Donald E. Knuth. *The art of computer programming. Volume 3: Sorting and Searching*. The art of computer programming. Addison-Wesley, Reading, Massachusetts, 3rd edition, 1997.

[19] János Komlós and Albert G. Greenberg. Correction to 'an asymptotically nonadaptive algorithm for conflict resolution in multiple-access channels'. *IEEE Transactions on Information Theory*, 31(6):841, 1985.

[20] Dariusz R. Kowalski. On selection problem in radio networks. In Marcos Kawazoe Aguilera and James Aspnes, editors, *PODC*, pages 158–166. ACM, 2005.

[21] Miroslaw Kutylowski and Wojciech Rutkowski. Adversary immune leader election in ad hoc radio networks. In *ESA*, pages 397–408, 2003.

[22] Gianluca De Marco and Dariusz R. Kowalski. Towards power-sensitive communication on a multiple-access channel. In *ICDCS*, pages 728–735. IEEE Computer Society, 2010.

[23] Charles U. Martel. Maximum finding on a multiple access broadcast network. *Inf. Process. Lett.*, 52(1):7–15, 1994.

[24] Jim L. Massey. Collision-resolution algorithms and random-access communications. Technical report, April 1980.

[25] Koji Nakano and Stephan Olariu. Energy-efficient initialization protocols for radio networks with no collision detection. In *ICPP*, pages 263–270, 2000.

[26] Koji Nakano and Stephan Olariu. Energy-efficient randomized routing in radio networks. In *In Proc. of DIALM*, pages 35–44, 2000.

[27] David Peleg. *Distributed Computing: A Locality-Sensitive Approach*. SIAM Monographs on Discrete Mathematics and Applications, 2000.

[28] Martin Raab and Angelika Steger. "balls into bins" - a simple and tight analysis. In Michael Luby, José D. P. Rolim, and Maria J. Serna, editors, *RANDOM*, volume 1518 of *Lecture Notes in Computer Science*, pages 159–170. Springer, 1998.

[29] Wojciech Szpankowski. *Average Case Analysis of Algorithms on Sequences*. John Wiley & Sons, Inc., New York, NY, USA, 2001.

[30] Dan E Willard. Log-logarithmic selection resolution protocols in a multiple access channel. *SIAM J. Comput.*, 15(2):468–477, 1986.