

ANASTASIOS GIOVANIDIS¹
QI LIAO²
SŁAWOMIR STAŃCZAK³

Measurement Based Self-Optimization in Random Access Communications⁴

¹Zuse Institute Berlin (ZIB), Discrete Optimization, Takustr. 7 D-14195, Berlin-Dahlem, Germany, giovanidis@zib.de

²Fraunhofer Institute for Telecommunications - Heinrich Hertz Institute (HHI), Einsteinufer 37 D-10587 Berlin, Germany, qi.liao@hhi.fraunhofer.de

³Fraunhofer Institute for Telecommunications - Heinrich Hertz Institute (HHI), Einsteinufer 37 D-10587 Berlin, Germany & TU Berlin Heinrich-Hertz-Lehrstuhl für Informationstheorie und theoretische Informationstechnik, Einsteinufer 25, D-10587 Berlin, Germany, slawomir.stanczak@hhi.fraunhofer.de

⁴This research has been partly supported by Alcatel-Lucent, Germany

Measurement Based Self-Optimization in Random Access Communications

Anastasios Giovanidis, Qi Liao, and Sławomir Stańczak

Abstract

The current work considers random access communications within a single cell and provides an algorithmic suggestion for the almost optimal coordination of user actions. In the scenario considered, an access effort is successful if (a) the signal is detected at the receiver and (b) no collision occurs. The first event is controlled by the user transmission power and the second by the choice of access (back-off) probability. These constitute the user action pair. The algorithm aims at exploiting information from measurements and user reports, in order to estimate current values of the system situation. Based on these, two optimization problems can be formulated and solved for the so called contention level and transmission power level at the base station side. The methodology to find the optimal values is based on minimization of a drift function. The two values are broadcast in order to help the users update their actions “almost optimally“. In this way the wireless cell can achieve self-optimization, without outside coordination, by relying on such intelligent information exchange and parameter estimation. Numerical results illustrate the great benefits of the suggested algorithm, compared to scenarios where the actions remain fixed, at a very low or even zero cost in power expenditure and delay.

Index Terms

Random Access Channel, Self Organizing Networks (SONs), Measurements, Collision Resolution, Drift Minimization, Power Control

I. INTRODUCTION

Random multiple access schemes have traditionally played an important role in wireless communication systems. Their use has been established especially in cases of bursty source traffic, where a multiplicity of users requires access to a central receiver. The most typical protocol is the celebrated ALOHA, first presented in the 1970's [1]. This has seen several modifications in the years to come, which aim especially at an increase in throughput performance. An interesting presentation of several related results can be found in [2]. Considering standardization, the IEEE 802.11 protocol for local area network communications has received large attention and extensive literature can be found related to its performance and optimization, see [3], [4], [5] and references therein. Random access standardization is also included in the 3rd Generation Partnership Project (3GPP) as an important element within the Long Term Evolution (LTE) systems [6].

The procedure is called *random access*, due to the fact that a number of sources aims at gaining access to a central entity, by randomly transmitting over the channel. In the general case, if more than one user transmit simultaneously, we say a *collision* occurs and all efforts are considered unsuccessful. After a failure, each source enters a *back-off* mode for a time interval chosen based on some probability distribution. This is a time when the source remains silent in order to avoid a new collision. This back-off time can be modeled in the slotted case by a per slot probability of transmission, less than 1. Using this technique, an increase in throughput is achieved at the cost of additional delay.

Random access is by nature a decentralized scheme, where sources (users) should individually decide on an action, in this case to transmit or not. The decision making can be facilitated by the availability of certain global information over the system state, which could be broadcast by the central receiver. In this way the delay-throughput tradeoff is enhanced at the cost of additional signaling, due to message passing, as well as further computations and processing at both ends of the communication channel. Furthermore, such protocols allow users to dynamically adapt to changes in system state. Such an approach can be seen as a self-optimization technique for the involved system. Global information is gathered at the receiver end and after certain processing, cell-specific information is passed to the users who update their actions. The current paradigm change in cellular wireless systems towards self-organizing functionality in the so called SONs (Self Organizing Networks), with as few intervention as possible coming from system operators, encourages this approach.

Important role in a self-organizing algorithm for random access plays the type of information measured at the central entity, its processing, as well as the messages broadcast to the users. Such an idea can be already found for the IEEE 802.11 protocol in order to take service differentiation into account in [5]. A description of how random access works in the 3GPP-LTE systems together with suggestions on the type of measurements and information extraction at the Base Station (BS) of an isolated cell is presented in [7]. On the other hand, a complete self-organizing algorithmic suggestion which is based on measurements and provides specific performance guarantees for a SON is something that, to the authors' knowledge, is not available in the current literature. This work aims at covering this gap by suggesting a protocol with specific information exchange between users and BS, that guarantees a near-optimal performance related to a certain measure. The algorithm is based on the optimization of a so-called *drift* function, which allows the system to reach an *almost optimal* steady state performance as $t \rightarrow \infty$. The approach here is completely novel, borrows however certain ideas from the existing literature, which are shortly presented in what follows together with our contributions.

A. Related Work and Contributions

Bianchi [3] has been the first to provide a precise performance analysis for a random access protocol, which uses exponential back-off times. His approach considers a *saturated system model*, where the number of users is kept fixed to N and all have a packet to send at each time slot. The results are based on the key approximation that the collision probability of a packet transmitted is constant and independent, which decouples the evolution of the system to N 1-dimensional Markov Chains. This is because the contention probability is not considered here a function of the back-off probabilities. The approach performs well for a large number of users.

A different approach, which departs from the above approximation has been suggested by Sharma et al. [5], where more general back-off strategies (generalized geometric) are considered in a slotted random access protocol. One of the major differences is that the system state is described by the current number of users per effort, while the collision probability is not independent per user, so that the back-off probability assigned to a user affects the collision probability of the others as well. The system is again saturated and the throughput is provided by obtaining the unique equilibrium point of the drift equations associated with the back-off process. The analysis is valid for a large number of users, for which the associated Markov Chain stays close to a so called typical state.

In the above works, the back-off probabilities are not adapted dynamically. First suggestions for controlling multiple access protocols can be found in Hajek and van Loon [8]. Gupta et al [4] have recently suggested a dynamic back-off adaptation mechanism, where contention is regulated by broadcasting a so called contention level to the users, who update in turn their access probability based on this and their current effort index. The contention level is updated at the BS by a Multiplicative-Increase-Decrease (MID) rule, after each success or failure.

Cross-layer optimization using the drift minimization is presented in the tutorial of Dimic et al [9]. Other channel-aware scheduling approaches, in conjunction with random access mechanisms such as ALOHA can be found in [10], and more recently in [11] and references therein.

The random access procedure for the LTE system model is described by Amirijoo et al [7]. The major difference from the previous models is that the miss-detection probability of a transmission influences the system state evolution together with the contention probability. This can be controlled by the transmission power of the user. A power ramping function is suggested to differentiate between users based on their current effort index. Furthermore, the authors suggest the utilization of user reports at the BS, so that the contention and miss-detection probability are estimated at each time slot. Based on such estimates the BS can broadcast relevant information so that the users can update their actions.

In the current work the problem formulation is based on a saturated system model, where the number of users remains constant within the cell. The LTE approach of the random access procedure as adopted, so that a transmission effort is successful when the signal is both detected and no collision occurs. For the state evolution of the system we make *the assumption that the collision probability is the same for all users, but is a function of the entire vector of back-off probabilities chosen*. This assumption is in some sense related to the approach of Bianchi but does not decompose the system, it rather describes its evolution as an N -dimensional Markov Chain. The transition probabilities are controlled for each user by the action pair of access (or transmission) probability and chosen power. The model is presented in detail in Section II.

We consider here as performance measure in Section III the expectation of some function of the system state, as time goes to infinity. The function chosen for the analysis is the sum of the current effort indices of all users. It is shown that the optimal action pair per user can be found by solving a dynamic program. An approximation to the optimal solution is found by choosing actions myopically

based on the drift minimization at each time slot.

The steps of the self-optimization algorithm are analytically presented in Section IV. The BS provides the users with two values, the current *contention level* and the current *power transmission level*, so that they can update their action pair. The first can be found as the solution of a reformulated drift minimization problem at the BS end. However certain unknown system quantities (the miss detection rate, contention rate as well as the number of users present within the cell) should be estimated and this is done based on measurements and user reports. The second level is updated based on a Multiplicative-Increase-Additive-Decrease (MIAD) rule.

Numerical results are shown in Section V. Our algorithm, which provides two possible values for the contention level and two actions (increase/decrease) for the power level, is compared to a fixed back-off and power level scenario, chosen such that a high performance is guaranteed for the system. Our suggestion outperforms such a scheme at a certain cost of higher power expenditure (which appears however only for the case of small user number) and additional delay. The costs are however trivial compared to the performance increase as illustrated in the simulation plots. Finally, Section VI concludes our work. Certain proofs of theorems presented in the main text can be found in the Appendix.

II. SYSTEM MODEL

We consider a fixed total number of N user terminals (UEs) with index n trying to obtain *random access* to a cell base station (BS) simultaneously over the wireless link. The time is slotted, with slot interval normalized to 1 and indexed by t . At each time slot all users belonging to the user set have the possibility to access the channel by transmission of a signal sequence. We use the following notation for the events of interest (all refer to time slot t which is omitted):

- PRN: a set of users with cardinality N is present within the cell
- TR($k \setminus N$): a user subset of cardinality k out of the total N transmits
- TR n : user n transmits
- MD n : miss-detection of user n at the BS
- COL: at least one pair of transmitting users collides
- SUn: user n has a successful transmission
- DR n : user n is dropped after M consecutive failures
- IDN: all N users within the set do not transmit and the slot remains idle

The complements of the above events are denoted by a C index, e.g. MD n^C for the event of detection of user n .

There are two criteria that determine the success of an attempt.

A) The received signal at the BS should have a Signal-to-Noise Ratio (SNR), which exceeds a predefined detection threshold γ_d . In case this does not happen, a miss-detection occurs and the user has to retry. The main reason, is that the user's transmitted signal experiences random fading over the link. Another one could be a wrong choice of transmission power due to estimation errors of the channel coefficient, in case such an estimate is available. The Detection Miss Probability (DMP) equals

$$\begin{aligned} DMP_n(p_n(t), \gamma_d) &:= \mathbb{P}[MDn|TRn] \\ &= \mathbb{P}[SNR_n(p_n(t), h_n(t)) \leq \gamma_d] \end{aligned} \quad (1)$$

and the probability is taken over the random quantity which can either be the entire channel fading coefficient or the fluctuation of the estimation error around a mean value. These two random variables are denoted both by h_n and they are independent identically distributed over the time slots t , whereas

the chosen transmission power is denoted by p_n . We attribute following behavior to the DMP functions under consideration

Property 1 *The detection miss probability $DMP_n(p_n(t), \gamma_d)$ is non-increasing with respect to the transmission power p_n and non-decreasing with respect to the threshold value γ_d .*

For the case $SNR_n = \frac{p_n(t) \cdot h_n(t)}{\sigma^2}$, where h_n denotes the channel fading and σ^2 is the zero mean noise variance, the property trivially holds. We note her that this is the case that mainly interests in the analysis and algorithmic suggestion to follow.

B) No collision of transmitted signals should occur. When more than one user tries to access the channel during the same time slot, there is a probability that the BS cannot separate the two signals, in which case all users that took part have to repeat the effort. In case no intelligent mechanism is available and the typical slotted ALOHA protocol is used [1], the probability of collision, under the condition that two or more users access simultaneously, *and their signals are detected*, equals 1.

In other protocols, such as those suggested in LTE standards [12], a pool of orthogonal sequences (e.g. Zadoff-Chu) is made available to all users. Each one chooses an element of this set randomly (uniform distribution) and the probability of collision is less than 1. For example, when N users are transmitting, if the sequence pool available at time t has $C(t) \geq N$ elements, the probability of collision for a transmitting user n can be found by solving the birthday problem [13] and equals

$$CP_n(N, t) \stackrel{\text{Remark 1}}{=} CP(N, t) \stackrel{\text{LTE}}{=} \mathbb{P}[COL | TR(N \setminus N) \& TRn \& MDn^C] = 1 - \frac{C(t)!}{C^N(t) (C(t) - N)!} \quad (2)$$

under the assumption that all signals have been detected. As a numerical example: If $N = 2$ users share a pool of $C(t) = 5$ sequences, the collision probability equals 0.2. In case $C(t) < N$ we set the probability always equal to 1, since at least two users will share the same sequence.

Remark 1 *In our model, the probability CP_n that a user n who transmits collides, equals the overall collision probability CP , for $N \geq 1$. In other words, we assume that when at least one pair of all N transmitting users takes part in a collision, all efforts are considered unsuccessful.*

Following properties are attributed to the collision probability expressions considered in this work (easily shown to hold for the example in (2))

Property 2 *The collision probability $CP(N, t)$ has positive differences with respect to the number of users N , that is*

$$CP(N + 1, t) - CP(N, t) > 0 \quad (3)$$

Furthermore, for no transmitting user $N = 0$, as well as a single user $N = 1$,

$$CP(0, t) = 0, \quad CP(1, t) = 0 \quad (4)$$

From the above, the probability of successful access to the BS, depends on the detection and the collision. Since the detection is related to the channel fading and the collision to the combined behavior of all the users and is conditional on the detection of user n 's signal, this probability takes as expression the following product

$$\begin{aligned}
SP_n(p_n(t), \gamma_d, N, t) &:= 1 - \mathbb{P}[SU_n^C | TR(N \setminus N) \& TRn] \\
&= 1 - \mathbb{P}[MDn | TRn] - \mathbb{P}[COL | TR(N \setminus N) \& TRn \& MDn^C] \\
&\stackrel{\text{Bayes, (1), (2)}}{=} (1 - DMP_n(p_n(t), \gamma_d)) \cdot (1 - CP(N, t))
\end{aligned} \tag{5}$$

There are two actions that user n can take for transmission at time slot t :

- The choice of the *transmission power level* $p_n(t)$, which influences the detection of the transmitted signal at the BS, as shown in (1).
- The choice of the *access (or transmission) probability* $b_n(t)$, which is the probability that the user transmits at the current slot.

$$b_n(t) := \mathbb{P}[TRn] \tag{6}$$

This influences the current number of simultaneously transmitting users in the cell and directly affects the collision probability in (2). The *back-off* probability simply equals $1 - b_n(t)$.

The set of actions for the entire system of N users at t is denoted by the $2N$ -dimensional vector $\mathbf{A}(t) := (\mathbf{b}(t), \mathbf{p}(t))$. The action space is denoted by \mathbb{A} and is the cartesian product $[0, 1]^N \times [0, P_1] \times \dots \times [0, P_N]$, where P_n is some individual user power constraint per slot.

It is clear from Property 1 that, the higher the chosen power, the higher the success probability in (5). The influence of the back-off probability is, however, a bit more complex. First of all, observe that in the definitions (1), (2) and (5) all probabilities are conditioned on the event TRn . If no back-off action is taken, then $b_n(t) = 1$ and TRn is always true. On the other hand, a different choice for the access probability, will have as a result that the actual success probability is given by the expression (5), multiplied by $b_n(t)$. The action of assigning $b_n(t) \leq 1$ to each user, displaces the UEs in time and the effect of collision is mitigated. Furthermore, since a lower than N set of users is taking part at the access of the medium per slot t , the collision probability will as well be reduced. This is explicitly stated in the following Proposition

Proposition 1 *Given an access probability b_n per user, the collision probability $CP^{(b)}$ equals*

$$CP^{(b)}(N, t) = \sum_{k=0}^{N-1} CP(k+1, t) \cdot \left(\sum_{l=1}^{L(N-1, k)} \prod_{i=1}^k b_{q_l^{k,i}} \prod_{j=1}^{N-1-k} (1 - b_{\hat{q}_l^{k,j}}) \right) \tag{7}$$

where $CP(k, t)$ is defined in (2), q_l^k is a subset of k -out-of- $(N-1)$ users with elements $q_l^{k,i}$, l is the index of the $L(N-1, k) = \binom{N-1}{k}$ different k -combinations and \hat{q}_l^k is the complementary user set with cardinality $N-1-k$. For the case where $b_n = b, \forall n$, the expression simplifies to

$$CP^{(b)}(N, t) = \sum_{k=0}^{N-1} CP(k+1, t) \cdot L(N-1, k) b^k (1-b)^{N-1-k}.$$

Furthermore, for $b_n \leq 1$, it holds

$$CP(N, t) \geq CP^{(b)}(N, t) \tag{8}$$

Each UE is constrained to at most M *access efforts* until success and the efforts are indexed by m . In case that, after M efforts, the user has not succeeded in obtaining access, the user is considered

discarded and replaced by a newcoming one, so that the total number in the system always remains the same. The same holds for each user that has a successful transmission. We say the system is *saturated*. The number of users, who are at transmission effort m at slot t is denoted by $X_m(t)$ and from the above

$$\sum_{m=1}^M X_m(t) = N, \quad \forall t \quad (9)$$

III. PERFORMANCE OPTIMIZATION

A. System States and Transition Probabilities

We define the state of user n at slot t as the current transmission effort $S_n(t) = m$, whereas the system state as the N -dimensional vector

$$\mathbf{S}(t) = (S_1(t), \dots, S_N(t)). \quad (10)$$

Altogether, there are M different user states and M^N different system states (e.g for a cell with 10 users and 5 maximum efforts, the number is approximately 10 million). The entire state space is denoted by \mathcal{S} . It is easy to see that the system state forms an N -dimensional Markov chain, since the access or not of the channel by the user, as well as the success or failure of the effort do not depend on the past state history. The dependency among the state evolution of the individual users is related to the collision probability $CP^{(b)}$, which is a function of the entire access probability vector.

We group the transitions for each user into (a) returning to state 1 in case of transmission and success, (b) moving to next effort in case of transmission and failure and (c) backing-off and remaining to the same state (only the time is used here as argument in the probability related functions for ease of notation)

- For $1 \leq m < M$:

$$\mathbb{P}[S_n(t+1) = 1 | S_n(t)] = b_n(t) (1 - DMP_n(t)) \cdot (1 - CP^{(b)}(t)) \quad (11)$$

$$\mathbb{P}[S_n(t+1) = S_n(t) + 1 | S_n(t)] = b_n(t) (DMP_n(t) + CP^{(b)}(t) - DMP_n(t) CP^{(b)}(t)) \quad (12)$$

$$\mathbb{P}[S_n(t+1) = S_n(t) | S_n(t)] = 1 - b_n(t) \quad (13)$$

- For $m = M$

$$\mathbb{P}[S_n(t+1) = 1 | S_n(t) = M] = b_n(t) \quad (14)$$

$$\mathbb{P}[S_n(t+1) = M | S_n(t) = M] = 1 - b_n(t) \quad (15)$$

A user in state M will either remain in the same state by backing off, or move to state 1 by transmitting, since - as mentioned - failure at the last effort removes the user from the system, who is replaced by a new one so that the total number remains always equal to N . Observe that the transition probabilities (14)-(15) for $m = M$, have the same expression as for the case $m < M$ given in (11)-(13), when DMP_n and $CP_n^{(b)}$ is set to 0.

Important is further to provide the *user dropping probability*

$$\begin{aligned} \mathbb{P}[DRn] &= \mathbb{P}[DRn | S_n(t) = M] \cdot \mathbb{P}[S_n(t) = M] \\ &= b_n(t) \cdot (DMP_n(t) + CP^{(b)}(t) - DMP_n(t) \cdot CP^{(b)}(t)) \cdot \mathbb{P}[S_n(t) = M] \end{aligned} \quad (16)$$

To calculate the steady state probabilities, the $M^N \times M^N$ transition probability matrix should be formed. Since the number of states is finite, and for each user (11) + (12) + (13) = 1 and (14) + (15) = 1, a steady state with probability sum equal to 1 always exists, although certain states may be transient and have zero probability.

B. Methodology and Approximation

Let us define a non-negative function of the system state $V(\mathbf{S}(t))$ and a performance measure, related to the steady state reached when $t \rightarrow \infty$

$$\mathcal{M}(V) := \lim_{t \rightarrow \infty} \mathbb{E}[V(\mathbf{S}(t)) | \mathbf{S}(0)]. \quad (17)$$

We will use the notion of *drift* to describe the expected change of the function V from time slot t to $t + 1$, given that the state of the system at t is $\mathbf{S}(t)$ and the action set taken is $\mathbf{A}(t)$. The latter is put as argument in the drift function, since it defines the system state transition probabilities $p_{s_t \rightarrow s_{t+1}}$.

$$D(V(\mathbf{S}(t)), \mathbf{A}(t)) := \mathbb{E}[V(\mathbf{S}(t+1)) - V(\mathbf{S}(t)) | \mathbf{S}(t)] \quad (18)$$

The importance of the drift is explained in the following propositions.

Proposition 2 *The performance measure can be written as an infinite sum of expected drifts over the discrete time axis, given the initial state $\mathbf{S}(0)$*

$$\mathcal{M}(V) = V(\mathbf{S}(0)) + \sum_{t=0}^{\infty} \mathbb{E}[D(V(\mathbf{S}(t)), \mathbf{A}(t)) | \mathbf{S}(0)]. \quad (19)$$

Proposition 3 *The problem of minimizing the system's performance measure, by choosing at each time slot t an appropriate action*

$$\begin{aligned} \mathbf{min} \quad & \mathcal{M}(V) \\ \mathbf{s.t.} \quad & \mathbf{A}(t) \in \mathbb{A}, \forall t \end{aligned} \quad (20)$$

can be solved as a dynamic program. The optimal solution satisfies Bellman's equation [14] for the cost-to-go function $J(\mathbf{S})$

$$J(\mathbf{S}) = \min_{\mathbf{A} \in \mathbb{A}} \left\{ D(V(\mathbf{S}), \mathbf{A}) + \sum_{\mathbf{S}' \in \mathcal{S}} p_{s \rightarrow s'} J(\mathbf{S}') \right\}, \forall \mathbf{S} \in \mathcal{S} \quad (21)$$

where the transition probabilities $p_{s \rightarrow s'}$ are functions of the actions chosen.

Proposition 4 *The solution of the drift minimization problem at each time slot t*

$$\begin{aligned} \mathbf{min} \quad & D(V(\mathbf{S}(t)), \mathbf{A}(t)) \\ \mathbf{s.t.} \quad & \mathbf{A}(t) \in \mathbb{A} \end{aligned} \quad (22)$$

is a suboptimal solution to the performance minimization problem in (20), which is called one-stage look-ahead (myopic), in the sense that the actions are chosen at each slot, considering only the transition to the next state and not the entire cost-to-go.

In the following paragraphs, the performance measure optimization will be approached by the drift minimization per time slot described above. Although it is a suboptimal solution for the original problem, it can however lead to tremendous improvement in the system performance and can furthermore allow for an implementation based on measurements and the ideas of self-optimization.

C. Performance Measure under Study and State-Dependent Actions

In the current work, the function V to be used is the sum of user states, which can be rewritten as the sum of cardinalities of users at a state, weighted by their effort index.

$$V(\mathbf{S}(t)) = \sum_{n=1}^N S_n(t) = \sum_{m=1}^M m \cdot X_m(t) \quad (23)$$

In this sense, a user who is currently at a higher effort, adds more to the function, than users in lower ones. Minimizing such a measure, aims at minimizing the expected number of users at higher efforts, as well as the drop probability, which by (16), depends on the steady state probability of a user being at state M .

We consider a homogeneous Markov chain by choosing the actions *state-dependent* rather than time-dependent. We write $\mathbf{A}(\mathbf{S}(t)) = (\mathbf{b}(\mathbf{S}(t)), \mathbf{p}(\mathbf{S}(t)))$. Under such an approach the probabilities that affect the state transitions can be modified as follows

$$\begin{aligned} DMP_n(p_n(t), \gamma_d) &\rightarrow DMP_n(p_n(S_n(t)), \gamma_d) \rightarrow DMP_m(t) \\ b_n(t) &\rightarrow b(S_n(t)) \rightarrow b_m(t) \\ CP^{(b)}(N, t) &\rightarrow CP^{(b)}(\mathbf{S}(t), N, t) \rightarrow CP^{(b)}(N, t). \end{aligned}$$

The collision probability remains $CP^{(b)}(N, t)$, since it has a global value and is time-dependent as a function of the number of users $X_m(t)$ being in a specific state. The drift expression can now be derived

$$\begin{aligned} D(V(\mathbf{S}(t)), \mathbf{A}) &= \sum_{n=1}^N \{1 \cdot \mathbb{P}[S_n(t+1) = 1 | S_n(t)] + \\ &\quad (S_n(t) + 1) \cdot \mathbb{P}[S_n(t+1) = S_n(t) + 1 | S_n(t)] + \\ &\quad S_n(t) \cdot \mathbb{P}[S_n(t+1) = S_n(t) | S_n(t)] - S_n(t)\} \\ &\stackrel{(11)-(13)}{=} \sum_{n=1}^N b_n(t) \cdot [1 - S_n(t) \cdot (1 - DMP_n(p_n(t), \gamma_d)) \cdot (1 - CP^{(b)}(N, t))] \\ &\stackrel{state\ dep.}{=} \sum_{m=1}^M X_m(t) \cdot b_m(t) \cdot [1 - m \cdot (1 - DMP_m(t)) \cdot (1 - CP^{(b)}(N, t))] \quad (24) \end{aligned}$$

IV. MEASUREMENT BASED SELF-OPTIMIZATION

In order for the system of users and base station to have near-optimal performance under the measure $\mathcal{M}(V)$, problem (22) should be solved per time slot at the base station side, where the drift is provided in (24) and the optimization variables are the state-dependent actions $\mathbf{A}(\mathbf{S}(t))$. These actions should then be transmitted as information to the user set. To do this, the objective function should be available at the base station, which includes following information:

- 1) The cardinality $X_m(t)$ of users at effort m in slot t , as well as the sum of present users in the cell $\sum_m X_m(t) = N(t)$.
- 2) The expression for DMP_m at effort m as a function of transmission power $p_m(t)$, or the actual value of this probability.

3) The expression for $CP^{(b)}$ as a function of N and $\mathbf{b}(t)$, or the actual value of this probability.

The concept of *self-optimization* in wireless networks is for the base station to adapt its functionality in an optimal way to changes that are not in its direct control. In our case, for random access communications, such changes could include a drastical increase or decrease in the number of users present in the cell, varying fading conditions or bursts which can not be modeled easily by a priori assumptions on statistical distributions and their moments, unpredictable user mobility etc. For all these reasons, closed form expressions or very precise information over 1. - 3. cannot be available.

To achieve the performance enhancement, however, the base station may approximate the objective function which resulted from a specific analytical model presented in the previous paragraphs, using *measurements* on channels and service quality, as well as *information, reported directly by the user set*. In this way, estimates of the information 1. - 3. can be used, towards a near optimal choice of actions. The aim is to achieve a considerable increase in performance, keeping the additional protocol overhead reduced. The general five steps of the suggested self-optimization algorithm in this work are given in Table I.

The subsections to follow, explain each step in detail. Before proceeding, we make certain assumptions, considering the action pair of access probabilities and transmission powers, so that the solution of the resulting problems is simplified and the required message passing between users and the BS is kept relatively low:

- The access probabilities are, as already mentioned, actions to be chosen optimally from the minimization of the drift at each time slot. The result should be broadcast over the channel in the future steps of the algorithm, so that the users can adapt their transmission in the optimal way. The approach in [4] is adopted here, where the per effort probability is set to

$$b_m(t) = \min \left\{ \frac{f(m)}{L(t)}, 1 \right\}, \quad \forall m \quad (25)$$

L is called *contention level* and $f(m)$ is some fixed function of the transmission effort. In this way, a simple variable L can simultaneously define the entire set of transmission probabilities. By choosing f monotone increasing in m , priority is given to users in higher efforts and the converse when f is monotone decreasing. Typical back-off protocols reduce by half the probability of accessing the channel after each failure, so in this case $f(m) = 2^{-m+1}$ and $b_1 = 1/L$. Other possible choices could be $f(m) = m^{-1}$ (which is mostly used in the current work and the simulations to follow) or $f(m) = m^{-0.5}$. In the following, the expression in (25) will sometimes be replaced by $b_m(t) = f(m)/L(t)$ and the constraint that the value is smaller or equal to 1 is taken into consideration in the constraint set of the minimization problem.

- We consider, furthermore, the transmission power to vary per effort as a ramping function. This is an approach often met in practice (for related approaches, the reader is referred to [7] and references therein). The power level for the first effort is given by P and for all efforts by the expression

$$p_m(t) = P(t) + (m - 1) \cdot \Delta p, \quad \forall m \quad (26)$$

where Δp is the ramping step with a fixed (tunable) value. In this way, exactly as the approach on the backoff probabilities previously, the entire vector of power actions can be defined by appropriate choice of the *power level* $P(t)$ per time slot.

A. Step 1: Measurements and User Reports

During the process of users trying to randomly access the channel, the BS is able to count the overall number of detected user efforts, as well as the overall number of successful efforts. Given a window of observation with length W , both these numbers depend on the time interval $[t - W + 1, t]$ and are denoted by $N_d(t), N_s(t)$ respectively. Furthermore, the users are considered to *report*, after each successful effort, the total number of trials required, in other words, their current state at success. In this way, the BS can keep track of the number of successes at effort m , within the observation window, denoted by $n_{s,m}(t), \forall m$. The reports over the success state can provide as well information over the overall number of transmissions of users being at some state m . As an example, if within the observation period two users report success at effort 3 and 2 respectively, the BS can estimate the number of transmissions at state $m = 1$ by 2, $m = 2$ by 2 and state $m = 3$ by 1, without considering users that have yet not declared success, or are dropped. We denote these estimates by $n_{t,m}(t), \forall m$ and their sum, which equals approximately the number of access efforts within the observation window, by $N_t(t) = \sum_{m=1}^M n_{t,m}$. Altogether, the set of gathered empirical information, updated per time slot, is summarized into

$$\mathcal{I}(t) := \{N_d(t), N_s(t), N_t(t), n_{s,m}(t), \forall m, n_{t,m}(t), \forall m\} \quad (27)$$

B. Step 2: Estimation of Unknowns in the Objective

Using the above counters, we can now approximate the unknowns in the expression (24) that were discussed in points 1.-3. at the beginning of this section.

- Considering the unknowns in 2. and 3. the actual overall contention probability $CP^{(b)}(N, t)$ from (2), (7) and per effort success probability $SP_m(t)$ from (5), can be replaced by contention and success rates, an idea which has already appeared in [7]. Observe that further information over the per effort DMP_m cannot be deduced from the above measurements. What can be calculated, instead, is an overall rate of miss-detection, without differentiating between efforts, which we denote by DMR . Then

$$CR(t) = 1 - \frac{N_s(t)}{N_d(t)} \quad (28)$$

$$SR_m(t) = \frac{n_{s,m}(t)}{n_{t,m}(t)}, \forall m \quad (29)$$

$$DMR(t) = 1 - \frac{N_d(t)}{N_t(t)}. \quad (30)$$

- Considering the number of users currently within the cell - discussed in point 1. - and their estimation, we proceed as follows. Instead of trying to find integer values, we consider instead rates of arrival. As rates of user arrivals, we understand here the number of users within the cell, divided by the length of the observation window. In case their number was a priori known somehow, this would equal the ratio $\frac{N}{W}$ for the whole system and $\frac{X_m}{W}$ for each one of the classes. Since we aim at minimizing the expression of the drift by appropriate choice of actions, dividing (24) by W will not change the optimal solution, just scale down the optimal value. As rate of user arrivals, we consider the ratio $\frac{N_s(t)}{W}$, which is the time dependent ratio of accepted users, divided by the observation window. The above is used under the assumption that only a very

small fraction of the users are dropped throughout the process, so that almost all users appearing within the cell, will eventually have at some point a success. To take dropped users as well into consideration, some further correcting term should be added, which could result from empirical observations.

The window is considered long enough, so that the resulting success rates per state, $SR_m(t)$ in (29), approach the actual success probability per effort. These can replace the entries in the one-step transition probability matrix, which have been provided in equations (11)-(13) and (14)-(15). For example in the case $M = 5$, we get (time dependency is omitted, considering that the values remain almost fixed for a certain long enough time interval)

$$\hat{\mathbf{P}}_5 := \begin{bmatrix} 1 - b_1(1 - SR_1) & b_1(1 - SR_1) & 0 & 0 & 0 \\ b_2SR_2 & 1 - b_2 & b_2(1 - SR_2) & 0 & 0 \\ b_3SR_3 & 0 & 1 - b_3 & b_3(1 - SR_3) & 0 \\ b_4SR_4 & 0 & 0 & 1 - b_4 & b_4(1 - SR_4) \\ b_5 & 0 & 0 & 0 & 1 - b_5 \end{bmatrix}$$

and the steady state probability distribution is found by solving the system $\pi = \pi \cdot \hat{\mathbf{P}}_M$, where π is the row vector of the unknown probabilities for the M states. The solution equals

$$\pi_1(t) = \frac{1}{1 + \sum_{i=2}^M \frac{b_1}{b_i}(1 - SR_1(t)) \cdot \dots \cdot (1 - SR_{i-1}(t))} \quad (31)$$

$$\pi_m(t) = \pi_1(t) \cdot \left(\frac{b_1}{b_m}(1 - SR_1(t)) \cdot \dots \cdot (1 - SR_{m-1}(t)) \right), \quad 2 \leq m \leq M \quad (32)$$

The ratios of the unknown backoff probabilities are involved in the expression above. From the previous discussion $b_1/b_m = f(1)/f(m)$, which are known, since the function f is chosen a priori. It is now easy to provide the estimation over the user arrivals per effort. That is:

$$\frac{X_m(t)}{W} \approx \pi_m(t) \cdot \frac{N_s(t)}{W} \quad (33)$$

and the π_m 's are the probabilities given in (31) and (32).

C. Step 3: Solving the Problem

It is possible now - after step 2 - to approximate the objective function and formulate the optimization problem in a way such that the optimal actions can be chosen per time slot. In our algorithmic suggestion, based on the measurements and estimated quantities described above, we break down the problem into two subproblems.

- **Backoff Probability Problem:** Making use of the estimated parameters, the objective function at the base station can be written as

$$\hat{D}(V(\mathbf{S}(t)), L(t)) := \frac{1}{L(t)} \cdot \left[\sum_{m=1}^M \pi_m \frac{N_s(t)}{W} f(m) \cdot (1 - m \cdot SR_m(t)) \right] \quad (34)$$

and the aim is to minimize the above convex function over the contention level L . Obviously, the optimization will have as a result either the maximum value of L , or the minimum one, depending on the sign of the term inside the square brackets (unless other more complicated constraints are taken into consideration). The lower bound is given by the constraint that all access probabilities should be less than or equal to 1

$$\frac{f(m)}{L(t)} \leq 1, \forall m \Rightarrow L(t) \geq \max \{f(m)\}.$$

To obtain an upper bound, we further provide a constraint on the probability that a time slot is idle (no user transmits). This probability should be less than or equal to \mathcal{A} - which is a design factor for the system.

$$\mathbb{P}[IDN] = \prod_{m=1}^M \left(1 - \frac{f(m)}{L(t)}\right)^{\frac{x_m(t)}{W}} \leq \mathcal{A} \Rightarrow \sum_{m=1}^M \pi_m \frac{N_s(t)}{W} \cdot \log \left(1 - \frac{f(m)}{L(t)}\right) \leq \log(\mathcal{A}).$$

The function on the left handside is increasing with L , thus the expression is an upper bound on L as well. Altogether, the solution is summarized in the following:

- if $\left[\sum_{m=1}^M \pi_m \frac{N_s(t)}{W} f(m) \cdot (1 - m \cdot SR_m(t))\right] \geq 0$ then the contention level is found by solving

$$\sum_{m=1}^M \pi_m \frac{N_s(t)}{W} \cdot \log \left(1 - \frac{f(m)}{L^*(t)}\right) = \log(\mathcal{A}) \quad (35)$$

- if $\left[\sum_{m=1}^M \pi_m \frac{N_s(t)}{W} f(m) \cdot (1 - m \cdot SR_m(t))\right] < 0$ then the contention level equals

$$L^*(t) = \max \{f(m)\}. \quad (36)$$

- **Power Control Problem:** In order to define the optimal transmission levels, a first approach would be to follow the above procedure and formulate again an optimization problem, given the backoff probabilities $f(m)/L^*(t)$ and the contention rates $CR(t)$ from (28). In order to fully determine the objective function based on (24), which can be denoted by $\tilde{D}(V(\mathbf{S}(t)), P(t))$, the closed form expression for the DMP_n as a function of power, with Property 1, should be available. This can be provided by some outage expression, it is however unlikely that it will always represent in the most accurate way the channel's fading behavior, hence the resulting power control will be unreliable. A different approach - which is adopted here - is to use a *Multiplicative-Increase-Additive-Decrease* (MIAD) control rule, as in the case of congestion control protocols in TCP [15]. In this way, the BS reacts on the change of the estimated DMR stepwise, by increasing or decreasing the power level $P(t)$ per time slot, depending on the current value $DMR(t)$. We set two levels of action, a high detection-miss level DMR^H and a low one DMR^L . The control loop then functions as follows. When the high level is exceeded, the power level is increased by multiplication with a tunable factor $1 + \delta_1$. This action increases rapidly the transmission power, since miss-detection is highly non-desirable. When the ratio falls under the low level DMR^L , which is considered satisfactory for system performance, the power is reduced in a conservative way, as an energy saving mechanism for the mobile users, by subtracting a constant tunable amount of δ_2 . A suggestion can be that δ_2 is set equal to the ramping step Δp in (26). The control loop is described by the following updates

$$P^*(t) = \begin{cases} P^*(t-1) \cdot (1 + \delta_1), & \text{if } DMR(t) > DMR^H \\ P^*(t-1) - \delta_2, & \text{if } DMR(t) < DMR^L \end{cases}. \quad (37)$$

D. Step 4 and 5: Broadcast of Information to the Users and Action Calculation

The last two steps of the algorithm involve the broadcasting of the action-related information to the users, which resulted from the previous analysis. This is given by the tuple of contention level and power level

$$\mathcal{J}(t) := \{L^*(t), P^*(t)\}. \quad (38)$$

The users that are present within the cell receive this information. Since each one of them is aware of his current individual state $S_n(t)$, and assuming that the expressions in (25) and (26) for the success probability and the power level per effort are known a priori to the mobile stations, each user can calculate its own action pair

$$A_n(S_n(t), \mathcal{J}(t)) = (b_n(t), p_n(t)) = \left(\frac{f(S_n(t))}{L^*(t)}, P^*(t) + S_n(t)\Delta p \right). \quad (39)$$

We have to note here that, in case this knowledge is not available at the mobiles, the BS should broadcast the entire vector of transmission powers and access probabilities calculated, and the mobile users should choose the actions related to their state.

The block diagram of the suggested algorithmic procedure is illustrated in Fig.1.

V. NUMERICAL RESULTS

The proposed algorithm has been implemented in a single cell scenario. The number of users within the cell at each time slot is considered fixed throughout each implementation period and users are randomly positioned, with a 2D uniform distribution. The algorithm is evaluated for the cases of $N = 1, 2, \dots, 14$ [users/time slot].

Considering the transmission scenario, each user randomly chooses at each attempt one sequence, out of a pool of 10 orthogonal sequences (we implement a simpler version of the LTE system model for collision avoidance), and transmits with a chosen backoff probability and transmission power. The signal experiences path loss due to the UE-BS distance. Fast fading is not considered here, since the effect of channel randomness is similar to that of random user positioning in a per effort scale. The channel is considered AWGN with noise mean equal to -133.2 dB.

An effort is successful when a sequence is detected and no two users among those taking part in the transmission collide, in the sense that no two users transmit the same chosen sequence simultaneously. A user is dropped when the effort fails at the maximum access effort $M = 5$. After a success or an event of dropping, users will be removed from the waiting-for-transmission list, and the same number of newly arriving users will be added, each given a random position on the plane.

Power and access probability for the users are computed per slot equal to the action pair in (39), for $f(m) = m^{-1}$. The suggested algorithm is further compared to a scenario, where access probabilities and target power are held fixed, while the ramping function for the transmission power is the same in both cases. To show the effect of the design factor \mathcal{A} , two different values are tested in the

suggested algorithm. The parameters of the two schemes are summarized in Table II and the performance comparisons are illustrated in Fig. 2 - 10.

The performance of the scheme and its comparison to the fixed scenario is better shown in the plots of the performance measure in Fig.2 and the dropped user rate in Fig.3. The performance measure is simply the actual value of the state function V in (23), for large numbers of time t , when the system is considered to have converged to a steady state. To make the comparison more "fair", the fixed scenario is chosen with access probability vector, such that the average occurrence of an idle slot is equal to the case where $\mathcal{A} = 0.25$. This is higher for the scenario where $\mathcal{A} = 0.5$ as shown in Fig.4. For both choices of \mathcal{A} the suggested algorithm outperforms the fixed scenario. The value of the performance function is lower - with increasing difference to the fixed case, for a higher number of users. Considering the dropping user rate, the dynamic scenario allows the system to remain stable - in the sense that the rate of dropped users does not tend to "explode" - for a higher value of N .

A more detailed comparison of the schemes is given in Fig.5-10. Specifically, Fig.5, 6 and 7 illustrate the system rates that influence the performance, meaning the detection-miss, contention and success rate respectively. Observe how the increase of the parameter \mathcal{A} improves all these, as well as the average effort for success, shown in Fig.8.

The last two figures provide the cost comparison for the two schemes, related to the average total number of time slots until success, as well as the expense in power for miss-detection avoidance. Higher benefits in the contention and success rate, naturally come at the cost of higher average delay up to success, as Fig.9 implies. On the other hand, the cost in average power expenditure per user equipment in Fig.10 is not necessarily higher in the dynamic scheme resulting by the algorithm, and can actually be less than in the fixed case, especially for a larger number of users in the cell.

Altogether, the dynamic scenario shows a significant increase in performance for the system at low costs in power and delay (sometimes even better), compared to a reasonable fixed scenario which aims at high performance. This is because the algorithm lets the system react to performance degradation, observed by the measured DMR and CR, by allowing the choice of two different actions for the power as well as the access probabilities. Obviously, availability of a larger action set would present even higher improvement, at the cost of significant signaling overhead from the BS to the UEs.

VI. CONCLUSIONS

We have suggested an algorithm for decentralized control of user back-off probabilities and transmission powers in random access communications. The algorithm is based on measurements and user reports at the base station side, which allow for an estimation of the number of users present within the cell, as well as the quantities of detection-miss and contention probability. By solving a drift minimization problem for the contention level and using closed loop updates for the transmission power level by a MIAD rule, the base station coordinates the actions chosen by the users, by broadcasting the information pair (L^*, P^*) .

The algorithmic steps, as well as the methodology of the drift minimization for a certain measure of interest referring to the steady state, provide a general suggestion to treat problems of self-organization in wireless networks.

Considering the specific scheme, a large variation of algorithms can be extracted, by choosing e.g. some different state function for the performance measure, or by introducing other kinds of user reports, which may provide more information to the central receiver, at the cost of increase in signaling. Furthermore, a larger action set can definitely provide a higher performance, compared to the proposed one - which introduces two possible values for the contention level (high/low) and two actions for the

power level (increase/decrease). Even in this scheme however, which is characterized by an “economy” of signaling and information exchange, the results - as illustrated by numerical examples - are extremely beneficial, especially as the user number in the cell increases.

APPENDIX

Proof: **[Proposition 1]** Similarly to the definition of CP in (2), the collision probability with non-certain transmission of users that are present in the cell, will take the expression

$$\begin{aligned}
CP^{(b)}(N, t) &:= \mathbb{P}[COL | PRN \& TRn \& MDn^C] \\
&= \sum_{\hat{k}=1}^N CP(\hat{k}, t) \cdot \mathbb{P}[TR(\hat{k} \setminus N) | TRn] \\
&= \sum_{k=0}^{N-1} CP(k+1, t) \cdot \mathbb{P}[TR(k \setminus N-1)]
\end{aligned} \tag{40}$$

where the second equality comes from the total probability theorem. The summation here starts with $\hat{k} = 1$, since the probability is conditional on the transmission of user n , hence at least one user will definitely access. By the index substitution $k := \hat{k} - 1 \geq 0$ in the third equality, we remove user n and consider the remaining $N - 1$ users.

For each k , we have to consider all k -out-of- $(N - 1)$ user samples without replacement. We denote each one by $q_l^k = \{q_l^{k,1}, \dots, q_l^{k,k}\}$ where $q_l^{k,i}$ are user indices belonging to the current k -sample. The total number is $L(N - 1, k) = \binom{N-1}{k}$ and for $k = 0$, $q_l^0 = \{\emptyset\}$. Furthermore, the complement of the index set q_l^k is the set of user indices not transmitting. This is denoted by \hat{q}_l^k and has size $N - 1 - k$. The probability of the index set q_l^k to transmit and \hat{q}_l^k not to, equals $\alpha_{k,l} := \prod_{i=1}^k b_{q_l^{k,i}} \cdot \prod_{j=1}^{N-1-k} (1 - b_{\hat{q}_l^{k,j}})$. Then

$$CP^{(b)}(N, t) = \sum_{k=0}^{N-1} CP(k+1, t) \cdot \left(\sum_{l=1}^{L(N-1,k)} \prod_{i=1}^k b_{q_l^{k,i}} \prod_{j=1}^{N-1-k} (1 - b_{\hat{q}_l^{k,j}}) \right) \tag{41}$$

Setting b fixed and equal for all users in the above, provides the second CP expression in the proposition.

To see why inequality (8) holds, observe that all possible index set pairs $\{q_l^k, \hat{q}_l^k\}$, $\forall k, l$, form a sample space for an experiment of N users being present within a cell and having a probability each to transmit or not. Hence $\sum_{k,l} \alpha_{k,l} = 1$ and the following holds

$$\begin{aligned}
CP^{(b)}(N, t) &\stackrel{(41)}{=} \sum_{k,l} CP(k+1, t) \cdot \alpha_{k,l} \\
&\stackrel{\text{Property 2}}{\leq} \sum_{k,l} CP(N, t) \cdot \alpha_{k,l} \\
&= CP(N, t)
\end{aligned}$$

Proof: **[Proposition 2]** Let $\mathcal{F}^{(t)} := \{\mathbf{S}(t), \dots, \mathbf{S}(0)\}$ be the information over the system realizations up to slot t . Obviously $\mathcal{F}^{(0)} \subseteq \mathcal{F}^{(t)}$ (formally $\mathcal{F}^{(0)}$ is a sub- σ -algebra of $\mathcal{F}^{(t)}$) and the tower ■

property for expectations [16, p.88] holds.

$$\begin{aligned}
\mathbb{E}[V(\mathbf{S}(t+1)) | \mathbf{S}(0)] &\stackrel{Tower}{=} \mathbb{E}[\mathbb{E}[V(\mathbf{S}(t+1)) | \mathcal{F}^{(t)}] | \mathcal{F}^{(0)}] \\
&\stackrel{Markov}{=} \mathbb{E}[\mathbb{E}[V(\mathbf{S}(t+1)) | \mathbf{S}(t)] | \mathbf{S}(0)] \\
&\stackrel{(18)}{=} \mathbb{E}[D(V(\mathbf{S}(t)), \mathbf{A}(t)) + V(\mathbf{S}(t)) | \mathbf{S}(0)] \\
&= \mathbb{E}[D(V(\mathbf{S}(t)), \mathbf{A}(t)) | \mathbf{S}(0)] + \mathbb{E}[V(\mathbf{S}(t)) | \mathbf{S}(0)]
\end{aligned}$$

and by repeating the process for $t, \dots, 0$ and taking limits for $t \rightarrow \infty$ at both sides we reach the result. \blacksquare

Proof: **[Proposition 3]** Consider the above series (19) up to a finite horizon $T+1$ and denote the related sum by $\mathcal{M}_T(V)$. The expected drift term for some $\tau \leq T$ equals

$$\mathbb{E}[D(V(\mathbf{S}(\tau)), \mathbf{A}(\tau)) | \mathbf{S}(0)] = \sum_{\mathbf{S}(1)} \dots \sum_{\mathbf{S}(\tau)} p_{s_0 \rightarrow s_1} \dots p_{s_{\tau-1} \rightarrow s_\tau} \cdot D(V(\mathbf{S}(\tau)), \mathbf{A}(\tau)) \quad (42)$$

where each summation is taken over the entire state space at time $1, \dots, \tau$. We denote by $p_{s_{\tau-1} \rightarrow s_\tau}$ the transition probability from state $\mathbf{S}(\tau-1)$ to $\mathbf{S}(\tau)$. It can be observed that $p_{s_{\tau-1} \rightarrow s_\tau}$, which can be controlled by the actions $\mathbf{A}(\tau-1)$ taken at time $\tau-1$, appear in all summands of $\mathcal{M}_T(V)$, for $\tau \leq \hat{t} \leq T$ and not for $0 \leq t \leq \tau-1$.

Following this observation, problem (20) can be rewritten as

$$\begin{aligned}
\mathbf{min} \quad & \min_{\mathbf{A}(T) \in \mathbb{A}} \mathcal{M}_T(V) \\
\mathbf{s.t.} \quad & \mathbf{A}(t) \in \mathbb{A}, t \in [0, \dots, T-1]
\end{aligned} \quad (43)$$

where the optimal choice of actions in T define the transition probabilities $p_{s_T \rightarrow s_{T+1}}$. These do not appear in other terms and can be found by minimizing the expression for the drift at T . The optimal cost-to-go equals

$$J(\mathbf{S}(T)) := \sum_{\mathbf{S}(T+1)} p_{s_T \rightarrow s_{T+1}}^* (V(\mathbf{S}(T+1)) - V(\mathbf{S}(T))) \quad (44)$$

The optimal cost-to-go is easily shown to satisfy the recursion, $\forall \mathbf{S}(\tau-1) \in \mathcal{S}$:

$$J(\mathbf{S}(\tau-1)) = \min_{\mathbf{A}(\tau-1) \in \mathbb{A}} \sum_{\mathbf{S}(\tau)} p_{s_{\tau-1} \rightarrow s_\tau} (V(\mathbf{S}(\tau)) - V(\mathbf{S}(\tau-1)) + J(\mathbf{S}(\tau))). \quad (45)$$

The expression remains the same, as the horizon $T \rightarrow \infty$ and taking also $\tau \rightarrow \infty$ we reach the result. \blacksquare

Proof: **[Proposition 4]** From the previous proposition, the optimal transition probabilities per time slot t are given by solving the problem

$$\min_{\mathbf{A}(t) \in \mathbb{A}} \sum_{\mathbf{S}(t+1)} p_{s_t \rightarrow s_{t+1}} (V(\mathbf{S}(t+1)) - V(\mathbf{S}(t)) + J(\mathbf{S}(t+1)))$$

Drift minimization simply does not take the term $J(\mathbf{S}(t+1))$ into account and considers only the transition $\mathbf{S}(t) \rightarrow \mathbf{S}(t+1)$. \blacksquare

REFERENCES

- [1] N. Abramson. The ALOHA system - Another alternative for Computer Communications. *in Proc. AFIPS Fall Joint Comput. Conf.*, 27, 1970.
- [2] A. Ephremides and B. Hajek. Information theory and communication networks: an unconsummated union. *IEEE Trans. on Inf. Theory*, 44, no. 6:2416–2434, Oct. 1998.
- [3] G. Bianchi. Performance analysis of the IEEE 802.11 distributed coordination function. *IEEE JSAC*, 18, issue:3:535–547, Mar. 2000.
- [4] P. Gupta, Y. Sankarasubramaniam, and A. Stolyar. Random-Access Scheduling with Service Differentiation in Wireless Networks. *INFOCOM 2005*, 3:1815 – 1825.
- [5] G. Sharma, A. Ganesh, and P. Key. Performance analysis of contention based medium access control protocols. *INFOCOM 2006*, pages 1–12, Apr. 2006.
- [6] 3GPP TS 36.300 (v 8.7.0) (May 2009) "Technical specification group radio access network; Evolved universal terrestrial radio access (E-UTRA) and evolved universal terrestrial radio access network (E-UTRAN); overall description" Release 8.
- [7] M. Amirijoo, P. Frenger, F. Gunnarsson, J. Moe, and K. Zetterberg. On Self-Optimization of the Random Access Procedure in 3G Long Term Evolution. *Proc. IEEE Integrated Network Management-Workshops, 2009.*, pages 177–184, Jun. 2009, New York, NY, USA.
- [8] B. Hajek and T. van Loon. Decentralized dynamic control of a multiaccess broadcast channel. *IEEE Trans. on Automatic Control*, AC-27, no. 3:559–569, June 1982.
- [9] G. Dimic, N. D. Sidiropoulos, and R. Zhang. Medium Access Control - Physical Cross-Layer Design. *IEEE Signal Processing Magazine*, 4, Sep. 2004.
- [10] L. Tong, Q. Zhao, and G. Mergen. Multipacket reception in random access wireless networks: From signal processing to optimal medium access control. *IEEE Comm. Magazine*, pages 108–112, Nov. 2001.
- [11] Y. Al-Harathi and S. Borst and P. Whiting. Distributed adaptive algorithms for optimal opportunistic medium access. *Mobile Netw Appl (Springer)*, 16, Issue 2:217–230, April 2011.
- [12] 3GPP TS 36.213 (V 8.7.0) Technical specification group radio access network; Evolved universal terrestrial radio access (E-UTRA); Physical layer procedures, Release 8, May 2009.
- [13] W. Feller. *An Introduction to Probability Theory and Its Applications, Volume I*. John Wiley & Sons, 1968.
- [14] M. L. Puterman. *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. Wiley & Sons, 2005.
- [15] Dah-Ming Chiu and Raj Jain. Analysis of the increase and decrease algorithms for congestion avoidance in computer networks. *Computer Networks and ISDN Systems*, 17, North Holland:1–14, 1989.
- [16] D. Williams. *Probability with Martingales*. Cambridge, 1991.

TABLES

TABLE I
GENERAL SELF-OPTIMIZATION ALGORITHM

STEP 1:	Gather empirical information \mathcal{I} at the base station (measurements and user reports).
STEP 2:	Estimate certain unknown factors (see 1. - 3. above) of the objective function (24).
STEP 3:	Solve the resulting approximate optimization problem in (22).
STEP 4:	Broadcast action-related information \mathcal{J} to the users.
STEP 5:	Calculate at the user side the required actions $A_n(t)$, $\forall n$, based on \mathcal{J} .

TABLE II
PARAMETER TABLE

Parameters	Value
Wireless Network	Single cell
User distribution	Uniform within cell
Number of users in cell	$\{1, 2, \dots, 14\}$
Sequence pool size	10
Fixed Tx Power	250 mW
Power ramping step Δp	20 mW
Maximum Tx Power	500 mW
Path loss PL	$128.1 + 37.6 \log(D \text{ km})$ dB
Noise	-133.2 dB
SNR threshold	8 dB
Maximum effort M	5
Fixed backoff probability	$[0.5, 0.4, 0.3, 0.2, 0.1]$
Number of slots	15000 slots
Window length W	200 slots
Backoff factor A	$\{0.25, 0.5\}$
Access Function $f(m)$	m^{-1}
Power control factor δ_1	2×10^{-4}
Power control factor δ_2	8 mW
DMR^H	3.5%
DMR^L	2.5%

FIGURES

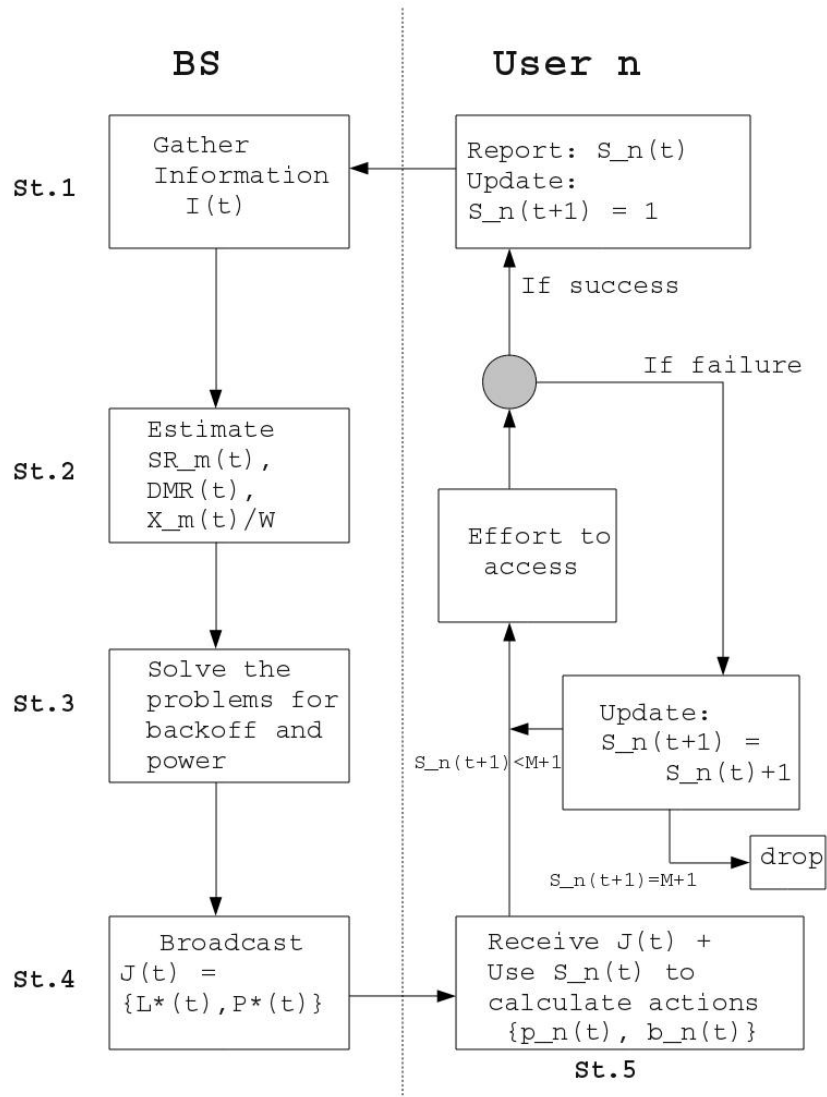


Fig. 1. Block Diagram of the suggested algorithm including the message passing between the BS and the mobile user

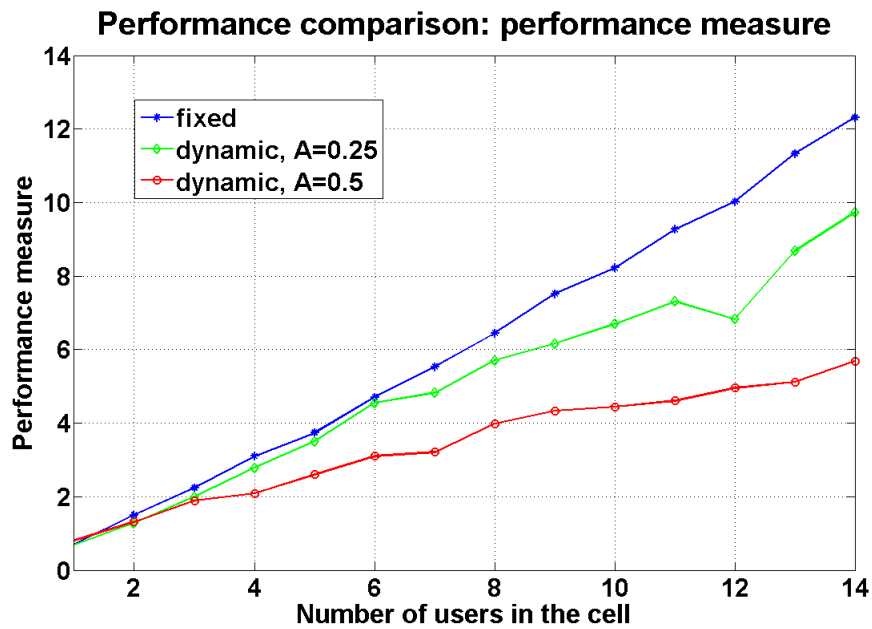


Fig. 2. Comparison of performance measure, equal to the chosen function V as $t \rightarrow \infty$. The measure improves with increasing idle probability bound \mathcal{A} .

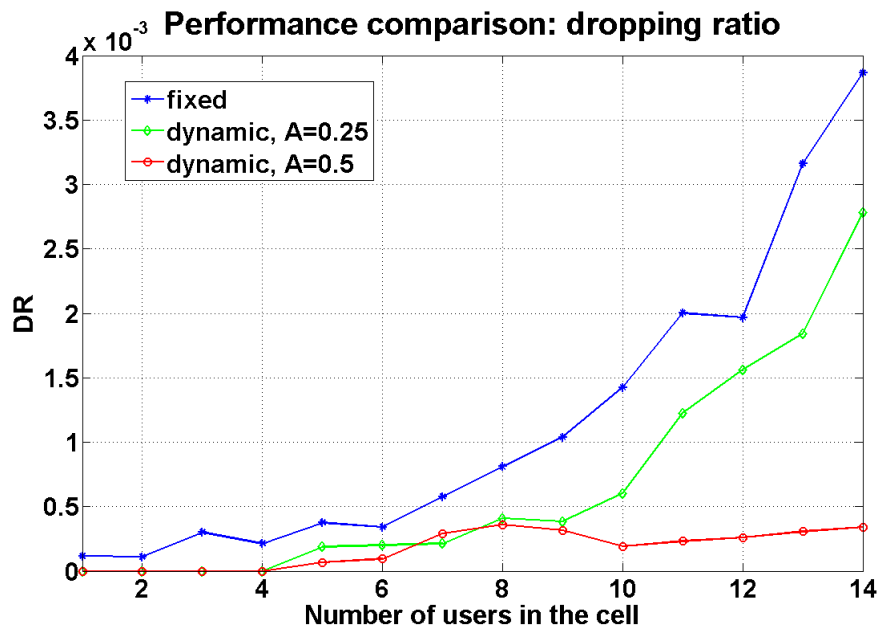


Fig. 3. Comparison of the average dropping rate (DR). The abrupt increase of the rate after a certain user number is an indicator that the system is not anymore stable for a further increase in the cell user number. Higher values of \mathcal{A} can increase the point when the instability appears, at the cost of delay.

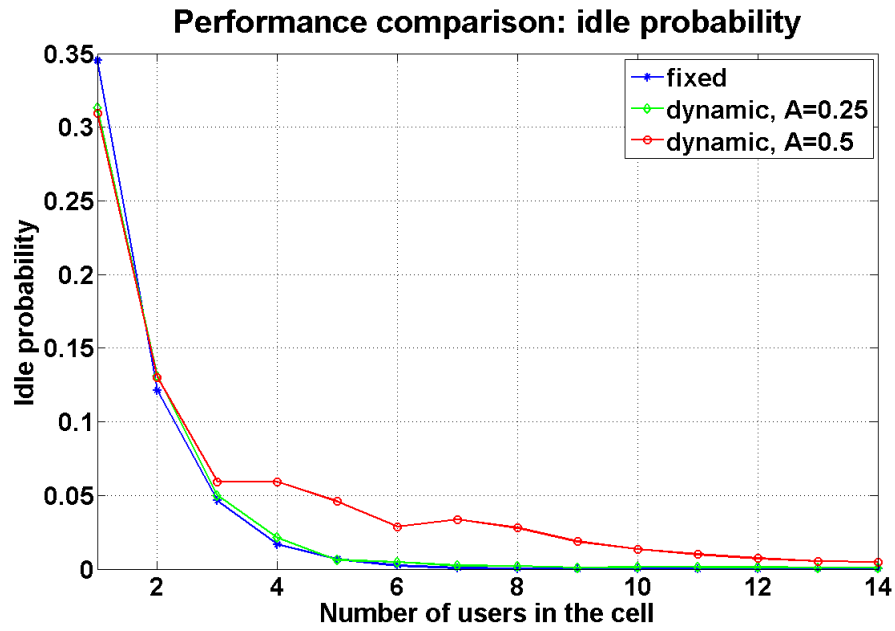


Fig. 4. Comparison of the average occurrence of idle slot per scheme. The dynamic scenario with $\mathcal{A} = 0.25$ follows the chosen fixed one.

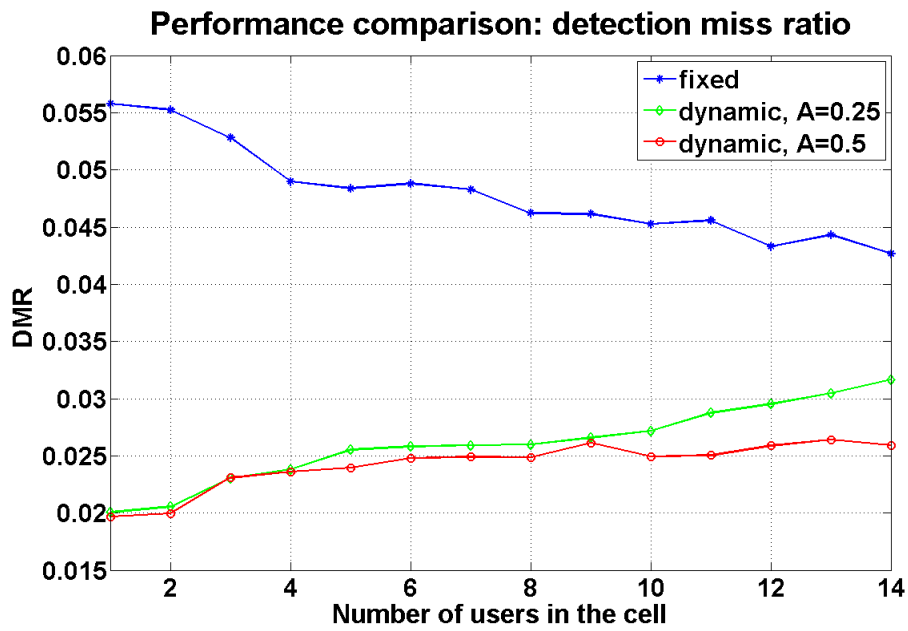


Fig. 5. Comparison of detection-miss rate (DMR) for the different schemes.

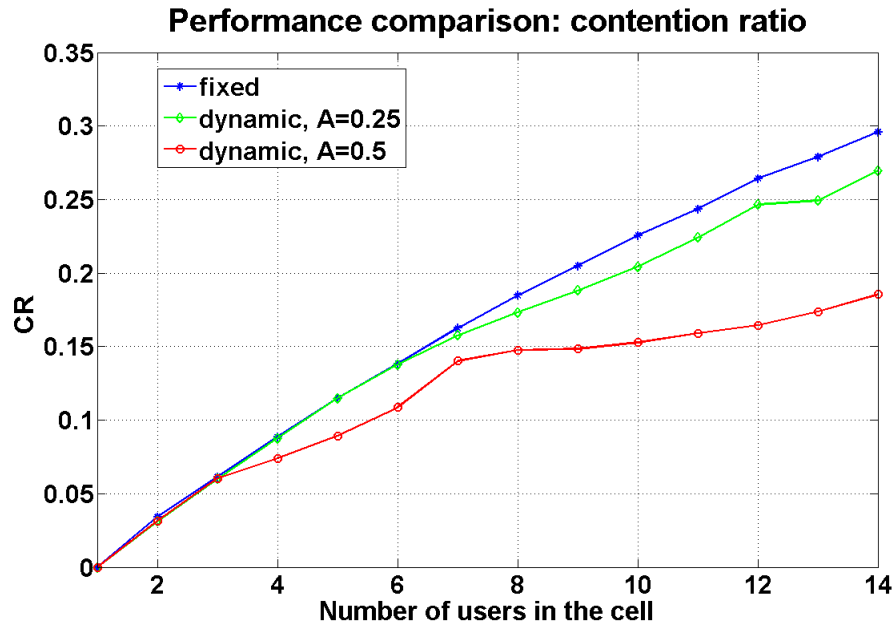


Fig. 6. Comparison of contention rate (CR) for the different schemes.

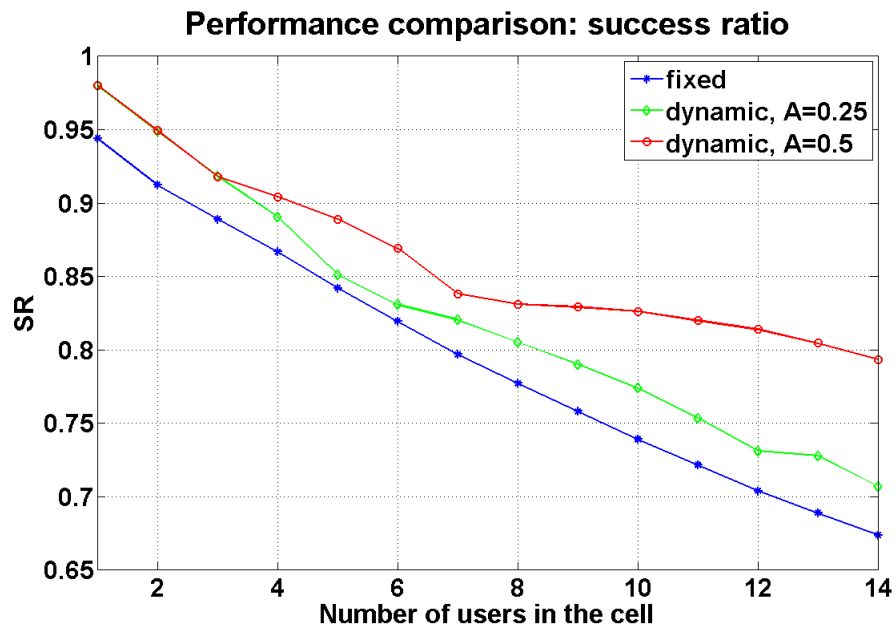


Fig. 7. Comparison of the overall success rate (SR) for the different schemes.

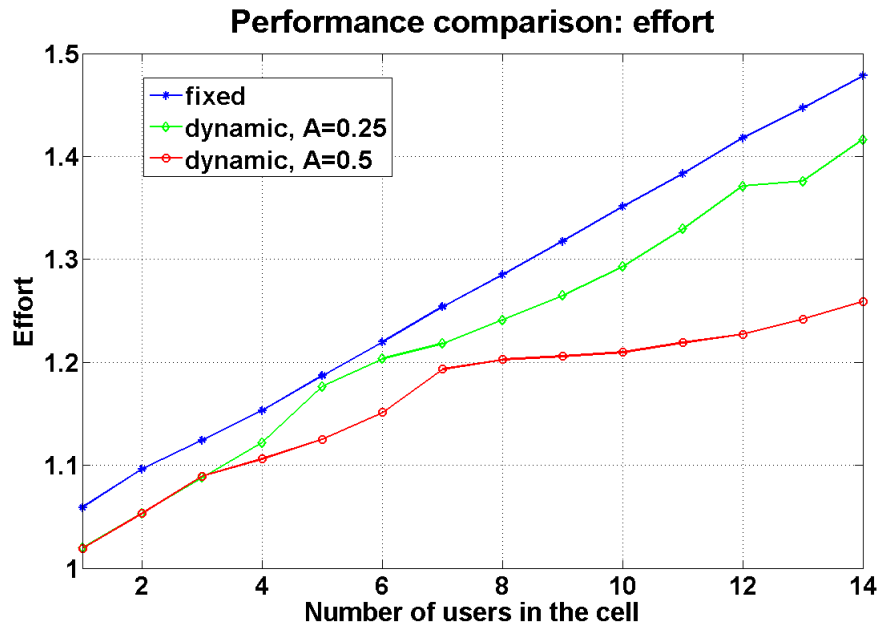


Fig. 8. Comparison of average number of efforts up to success (including the case of dropping).

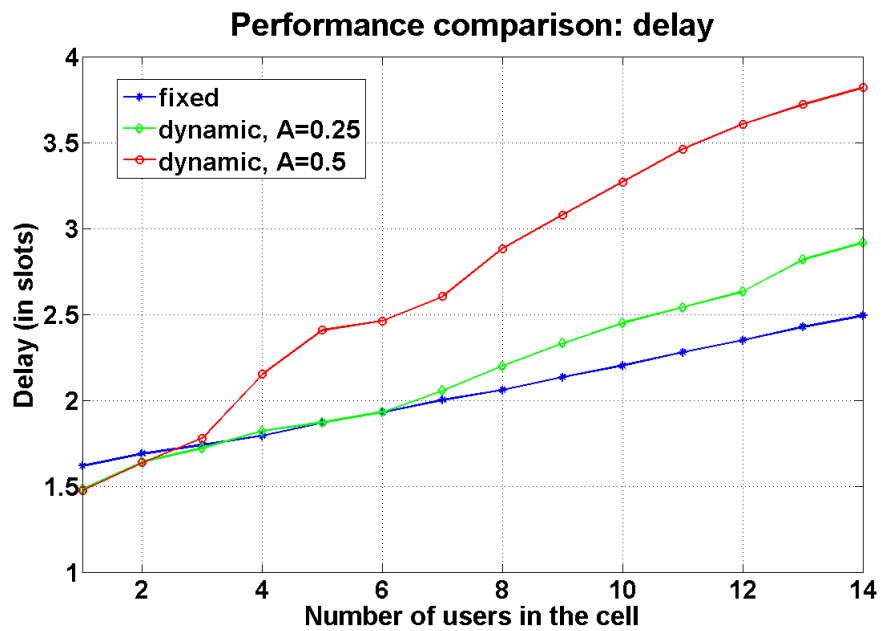


Fig. 9. Comparison of actual average delay up to success or dropping (alternatively: average time of user presence in the cell) in time slots.

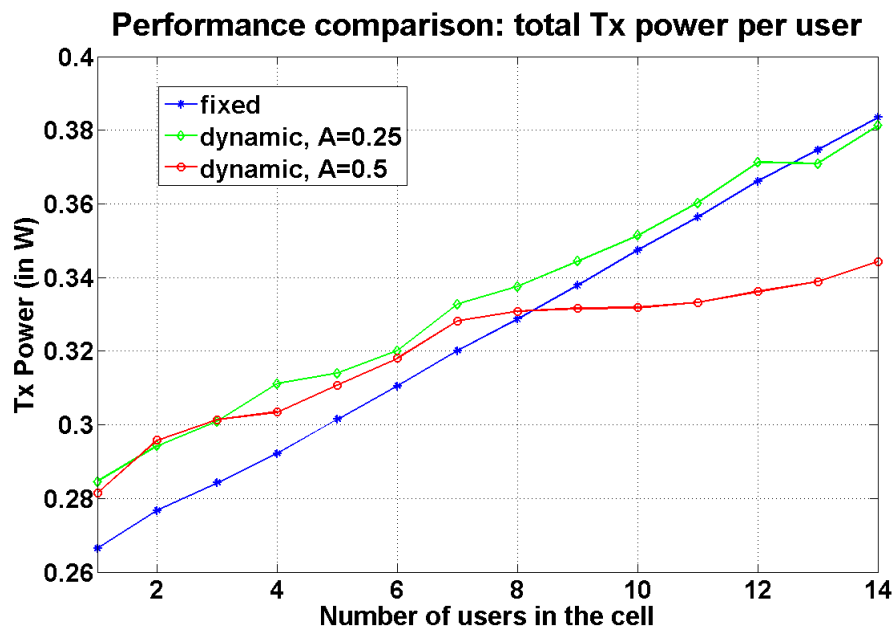


Fig. 10. Comparison of average total transmission power per user for the different schemes.