Online Collaborative Filtering with Nearly Optimal Dynamic Regret

Baruch Awerbuch*
Department of Computer Science
Johns Hopkins University
baruchaw@gmail.com

Thomas P. Hayes
Toyota Technological Institute at Chicago
hayest@tti-c.org

ABSTRACT

We consider a model for sequential online decision-making by many diverse agents. On each day, each agent makes a decision, and pays a penalty if it is a mistake. Obviously, it would be good for agents to avoid repeating the same mistakes made by other agents; however, difficulty may arise when some agents disagree over what constitutes a mistake, perhaps maliciously.

As a metric of success for this problem, we consider *dynamic regret*, *i.e.,* regret versus the off-line optimal sequence of decisions. Previous regret bounds usually use the much weaker notion of *static regret*, *i.e.,* regret versus the best single decision in hindsight. We assume there is a set of "honest" players whose valuations for the decisions at each time step are identical. No assumptions are made about the remaining players, and the algorithm assumes no information about which are the honest players.

We present an algorithm for this setting whose expected dynamic regret per honest player is optimal up to a multiplicative constant and an additive polylogarithmic term, assuming the number of options is bounded.

Categories and Subject Descriptors

I.2.11 [**Artificial Intelligence**]: Distributed artificial intelligence—*Multiagent systems*; G.3 [**Probability and Statistics**]: Probabilistic algorithms; I.2.6 [**Artificial Intelligence**]: Learning—*Knowledge acquisition*

General Terms

Algorithms, Theory

Keywords

Online decision making, Collaborative filtering, Regret, Learning, Multiagent systems

*Supported by NSF grants 0617883, 0515080, 0240551 and 0311795

SPAA'07, June 9–11, 2007, San Diego, California, USA.

1. INTRODUCTION

Informally, we would like to consider the following problem. Each day, different agents make irreversible decisions, and pay for their mistakes. It makes a lot of sense for the agents collectively to avoid making the same mistakes multiple times, rather than just individually. As von Bismarck is credited with saying, *"Fools learn from experience; wise men learn from the experience of others."*

It would be easy to leverage experience of others if everybody would have the same perception of the world. However, in reality these perceptions are different for a number of reasons, e.g., people may have different taste (for books, movies, food, etc.), sensors may experience different reception, eBay users may be dishonest, etc.

Consider the online learning where users of the systems learn to to trust and distrust other users based on their declared perceptions of reality. Eventually, it may be possible to learn all the users with identical views. However, in the meantime, many "mistakes" can be made. Consider a "community" set of users who have the same point of view. One example of a community is the set of honest users. Our goal is minimizing total number of mistakes of such a community, so that community experiences the same number of mistakes as if it were a single individual, i.e., as if members of the community would know exactly who is the community. The problem is that communities are not known in advance and need to be learned; the learning process involves making mistakes and learning from these mistakes. This is the framework pursued in [5], [6],[7],[9], [8], [4], [1].

The standard way in Learning Theory to measure number of mistakes is by introducing notion of regret against hindsight strategy. It is customary to limit the hindsight "insider" strategy to be completely static, as in [15] or [6]. In a centralized (single-agent) environment, achieving regret against dynamic strategies is obviously impossible. However, with multiple trustworthy agents, it may be possible in some scenarios to achieve low regret against dynamic optimal strategies.

Our goal is to study regret against dynamic strategies. We consider the setting as follows. We have a sequence of rounds, and a set of users making decisions on each round among a number (say, two) options. We assume that users take actions on each day in a certain order, and thus can learn from predecessors, provided that predecessors have the same view. The goal is to minimize total number of mistakes compared to dynamic strategy that at each round makes the best choice for the community of users. In this paper, we show that one can accomplish small regret in such a setting.

Day	1	2	3	4	5	6	7
Bad Agent 1	?,0	1,?	1,?	0,?	1,?	1,?	0,?
Good Agent 2	?,1	?,0	0,?	1,?	?,1	?,0	1,?
Bad Agent 3	?,0	?,0	0,?	0,?	?,1	?,0	?,0
Bad Agent 4	0,?	1,?	0,?	1,?	?,1	?,0	1,?
Good Agent 5	?,1	?,0	?,1	?,0	?,1	0,?	?,0
Good Agent 6	0,?	?,0	0,?	?,0	?,1	0,?	?,0

Figure 1: Example of an execution. In each entry, we give the costs for the choosing "left" and "right" as reported by that agent on the given round; the value for the unchosen option is given as "?" Agents 2, 5 and 6 are honest. Although Agent 4 is dishonest, this cannot be seen from the available information. The best choices are "left" on day 1, "right" on day 2, "left" on day 3, "right" on day 4, "left" on day 5, "right" on day 7. On day 6, both choices are the same. Notice that agents 5 and 6 "learn to distrust" agents 1 and 3 as the game goes on.

2. PROBLEM DEFINITION

Consider m decision options, which we will simply denote as $1, \ldots, m$, which are available to n agents i, $1 \leq i \leq n$, over a sequence of T rounds, $1 \leq t \leq T$, each round consisting of a single choice by each agent, performed sequentially.

The function $\gamma(i, k, t) \in \{0, 1\}$ determines the cost for agent i to choose option k at time t. At each time t, each Player i chooses an option $k = \pi(i, t)$, pays cost $\gamma(i, \pi(i, t), t)$ and *publishes* its incurred cost; the values of the $(k - 1)$ "unexplored" options remain unknown. Note that in case of an *adversarial* agent i, the cost $\gamma(i, k, t)$ can be determined in an adversarial manner, as a function of the current state of the algorithm and of the published strategy for the honest players.

We assume that there is a fixed schedule permutation of the players such that, in each round, the players make choices in the order given by that schedule. We index the players according to this schedule, so that Player j precedes Player i if and only if $j < i$. We also assume that each Player i knows the costs already incurred by Players j with $j < i$. In other words, execution is sequential, in a round-robin fashion determined by σ with n steps constituting a "round" during which costs do not change. Figure 1 illustrates one possible sequence of play with $n = 6$ and $T = 7$.

Let $\pi^*(i, t)$ be the decision minimizing such cost at time t, and let the regret of i at time t be

$$\rho(i, t) := \gamma(i, \pi(i, t), t) - \gamma(i, \pi^*(i, t), t)$$

Definition 2.1. *A subset $S^* \subset 2^n$ of players is* consistent *if their declared and undeclared costs for the same options are identical, namely, for all $i, j \in S$, $k \in \{0, 1\}$, $1 \leq t \leq T$, we have $\gamma(i, k, t) = \gamma(j, k, t)$.*

Our goal is to minimize, for each such set S^* the total regret per player in S^*, defined as

$$\rho(S^*) := \frac{1}{|S|} \cdot \sum_t \sum_{i \in S^*} \rho(i, t)$$

Remark: Notice that this definition of regret is an additive measure of how much worse our algorithm performs than the optimum *dynamic* strategy.

We assume that agents may be adversarial. However, we legislate the following "randomness verification act": random coins used by agent i, $\chi(i)$ must be publicly verifiable to be either a true coin, or at least as a coin with limited bias. Efficient peer-to-peer verification and generation of private random coins in a Byzantine environment have been implemented, for example, in [2, 3, 10, 11, 12, 13].

Remark: Note that in the case of an *adversarial* agent i, the cost $\gamma(i, j, t)$ can be determined in an adversarial manner, as a function of the current state of the algorithm and of the published strategy for the honest players. All players already know Player i's choice because all the randomness has been public. In fact, there is no private information in this model. By induction on the number of players who have already taken steps in the round, the internal states of all honest players are publicly known.

To gain a better understanding of our model, it may help to consider the performance of a natural "naive" approach. The first such approach is to "trust" everyone who we have never "caught lying". We will call such trusted players "advisors." When advisors give conflicting advice, we follow the advice of the majority. Unfortunately, this algorithm can lead to a total regret of $\Omega(|S|T)$ for a set of $|S|$ honest players, even when the honest players are in the majority. In particular, if the first X players are dishonest, and say that the better of the two options is bad, the next X honest players will believe them, and choose the other option, earning a regret of X for the round. Moreover, because none of these X honest players made the same choice as the dishonest players, none of the dishonest players is "caught," and so this scenario can happen as many as T times. When $|X| = \Omega(|S|)$, this gives an aggregate regret of $\Omega(T)$ *per honest player*, which is the worst possible.

Our main result is that, for players with identical tastes, the regret can be essentially divided by the number of such players.

Theorem 2.1. *For any set S of identical players out of n players in all, our algorithm (see Fig. 2) has expected total regret*

$$\mathbf{E}(\rho(S)) = O\left(m^2 \log^2 n + \frac{mT}{|S|}\right)$$

per player in S over T rounds, where m is the number of options.

Remark: If, in each round, one object is randomly selected to have cost 0, and all the rest have cost 1, then the expected per-player regret for any set S of players must be at least $(m - 1)T/2|S|$, since on average $(m - 1)/2$ wrong guesses will be made each round before finding the item of cost 0 (asssuming the players outside S cannot be relied on to report the correct values). The other term in our regret bound does not depend on T, and may be thought of as a "short-term penalty" due to initially not knowing which players are "trustworthy" (that is, in S) and which are not.

3. COMPARISON WITH EXISTING WORK

We stress a profound difference between this model and the easier problem of minimizing regret against the best static option; the latter model is very popular in machine learning community and theoretical computer: the comparison is against the *dynamic optimum*, rather than a static

Initialize: **for all** $1 \le k < i \le n$, set $w(i,k) \leftarrow 1$
At each time t, for each Player i, **do**

1. Let $W(i) \leftarrow \sum_{k<i} w(i,k)$
2. For each $j \in \{1,\ldots,m\}$, let $Q(i,j) \leftarrow |\{k < i\colon w(i,k) \text{ and } \gamma(k,j,t) = 0\}|$
3. **If** $\max_j Q(i,j) > \alpha W(i)$, **then** $\pi(i) \leftarrow \text{argmax}_j\, Q(i,j)$.
4. **Else** $\pi(i) \leftarrow$ random element of $\{1,\ldots,m\}$
5. Proceed with choice $\pi(i)$, accruing cost $\gamma(i,\pi(i),t)$.
6. $\forall k < i$, **if** $\pi(k) = \pi(i)$ and $\gamma(k,\pi(i),t) \neq \gamma(i,\pi(i),t)$, set $w(i,k) \leftarrow 0$

Figure 2: An online learning algorithm for Player i.

optimum as in [15] or [6]. We point out that previous work constitutes the following special cases of our problem.

Known taste types.

If all the agents knew exactly which agents shared their tastes, then they could simply ignore all the other agents. In each round t, the first m agents to act from each group with identical tastes, could be designated as "martyrs," whose job is simply to report the values $\gamma(i,1,t),\ldots,\gamma(i,m,t)$, thus sacrificing their own welfare for the benefit of the group. The rest of the group would then simply choose the best option.

The first mention of achieving low regret compared to the dynamic optimum is in [7], which presented a heuristic for this setting that is often more efficient than the trivial algorithm. However, in the worst case, its amortized regret *per honest player* is $\Omega(T)$. In contrast, Theorem 2.1 bounds the total amortized regret by a *poly-logarithmic* term, assuming $|S| > T$.

4. ALGORITHM

Our solution can be viewed as an extension of the "halving algorithm" of Barzdin and Freivalds [14], which is a special case of the celebrated weighted majority algorithm of Littlestone and Warmuth [15].

In our algorithm (see Figure 2) each agent i maintains a "trust vector" $w(i,j)$ that indicates whether i still "trusts" previous agent $j < i$. Initially, $w(i,j) = 1$, so each agent trusts all his predecessors. As the algorithm proceeds, if i ever catches j in a "lie," that is, i and j choose the same option k in round t, but $\gamma(i,k,t) \neq \gamma(j,k,t)$, then $w(i,j)$ is set to 0 from then on. The variable $W(i)$ denotes the total weight of trust towards all the previous agents, and $Q(i,k)$ is the total trust weight of agent i for previous agents "positively recommending" option k. By *positively recommending*, we mean that the agent chose option k and reported its cost as 0, *i.e.*, $\pi(j,t) = k$ and $\gamma(j,k,t) = 0$. If the relative weight of positive recommendations for an option k is at least a certain threshold value $\alpha/2m$, where α may be taken to be $1/4$, then Player i deterministically chooses whichever option has the greatest weight of positive recommendations. Otherwise, Player i chooses randomly, using a publicly verifiable random coin.

The intuitive motivation for this algorithm is that positive advice should be followed, since if it is true, it causes no regret, whereas if it is false, one catches the advisors lying, and therefore are not fooled twice. By further requiring that a positive fraction of the advisors give the same positive advice, we ensure that such lies can only occur a logarithmic number of times.

In contrast, negative recommendations, such as "Don't do what I did," are hard to deal with, since if such advice is followed, there is no way to know whether it was honest. Thus our algorithm chooses randomly when there are too many negative recommendations, just as if all prevous agents were known liars.

5. ANALYSIS OF THE ALGORITHM

We begin with some basic notation which will be used throughout the analysis. The letter t will always denote a time $1 \le t \le T$. Variables from our algorithm, when indexed by t, always refer to the value of the variable at the beginning of round t of the algorithm.

NOTATION 5.1. *For a player i and a time t, let $R_{i,t}$ denote Player i's regret due to his decision at time t. That is, $R_{i,t} = 1$ if Player i incurred cost 1 for his choice at time t, but could have incurred cost 0 for some other choice. Otherwise, $R_{i,t} = 0$. We will denote Player i' total regret by $R_i = \sum_{t=1}^{T} R_{i,t}$.*

DEFINITION 5.1. *Define, for every player i and time t, a random variable $\Phi(i,t)$, by*

$$\Phi(i,t) = \sum_{j<i} w_t(i,j) \ln(2 + W_t(j))$$

Note that, since $w_t(i,j)$ and $W_t(j)$ are non-increasing as a function of t, so is $\Phi(i,t)$. Finally, note that $\Phi(i,t) = 0$ if and only if $W_t(i) = 0$, that is, Player i has no advisors left. On the other hand, if $\Phi(i,t) \neq 0$, then

$$\ln(2) \le \Phi(i,t) \le W_t(i) \ln(i),$$

a fact we shall make frequent use of.

For the purposes of our analysis, let us assume that, at each time t, for each player i, a random option is determined in advance, regardless of whether the algorithm decides to make a random choice. When the algorithm makes a deterministic choice, the random choice is simply discarded. This viewpoint clearly has no effect on the choices made by our algorithm. Next we introduce a notion of when a player should be happy such a random sequence.

NOTATION 5.2. *Fix a Player i and a time t. Let N_{Good} denote the number of Good options for Player i at time t. We say that i is* Lucky at time t *if at least $2\alpha N_{\text{Good}} W/m$ of Player i's advisors at time t have, as their random selections, options which are Good for player i. Note that this definition depends only on the history of the game on rounds 1 to $t-1$, on the Good options for player i at time t, and the random choices for time t. In particular, we have said nothing about whether the advisors of i actually make deterministic or randomized decisions at time t.*

LEMMA 5.1. *Suppose $W_t(i) \geq 1$ and that Player i is Lucky at time t. Then, regardless of how the advisors play, at least one of the following holds:*

(i) $R_{i,t} = 0$ deterministically, or

(ii) $\mathbb{E}\,\Phi(i, t+1) \leq \left(1 - \frac{\alpha^2}{2m^2 \ln(i)}\right)\Phi(i,t)$, where the expectation is with respect to Player i's decision at time t. (Note that this includes the case when Player i deterministically makes a choice with $R_{i,t} = 1$.)

PROOF. If Player i chooses deterministically, then either we are in case (i), or at least $\alpha W_t(i)/m$ advisors said that the option i chose was Good, when it was actually Bad. In the latter case, Player i burns these advisors, each of whom had contributed at least $\ln(2)$ to $\Phi(i,t)$, and hence, deterministically,

$$\begin{aligned}\Phi(i,t+1) &\leq \Phi(i,t) - \frac{\alpha W_t(i)\ln(2)}{m}\\ &\leq \Phi(i,t)\left(1 - \frac{\alpha \ln(2)}{m \ln(i)}\right),\end{aligned}$$

in which case we are done.

Henceforth, assume Player i chooses randomly. This means that no option received more than $\alpha W_t(i)/m$ positive reviews by advisors of i. In particular, fewer than an $\alpha N_{\text{Good}}/m$ fraction of advisors gave positive recommendations for the N_{Good} options which are Good for Player i. On the other hand, since Player i is Lucky at time t, at least twice this many of Player i's advisors had random options which are Good for Player i. Hence, at least half of these advisors did one of:

1. Selected a Good option for Player i, but reported it as Bad, or
2. Deterministically selected a Bad option for Player i, and reported it as Good, or
3. Deterministically selected a Bad option for Player i, and reported it as Bad.

Let us consider the contribution of such an advisor, Player j, to $\Phi(i, t+1)$.

Cases (1) and (2). Player j gave a report which does not agree with Player i's valuation. Since Player i chooses randomly, Player j will be eliminated as an advisor with probability $1/m$. Thus

$$\begin{aligned}\mathbb{E}\left(w_{t+1}(i,j)\ln(2 + W(j,t+1))\right) &\leq \left(1 - \frac{1}{m}\right)\ln(2 + W(j,t))\\ &\leq \ln(2 + W(j,t)) - \frac{\ln(2)}{m}.\end{aligned}$$

Case (3). Player j deterministically chose a Bad option for player i and reported it as Bad. In this case, Player j burns at least an α/m fraction of her advisors. Noting that this implies $W_t(j) \geq 1$, we have

$$\begin{aligned}&\ln(2 + W_{t+1}(j)) - \ln(2 + W_t(j))\\ &\leq \ln\left(\frac{2 + W_t(j) - \lceil \alpha W_t(j)/m \rceil}{2 + W_t(j)}\right)\\ &= \ln\left(1 - \frac{\lceil \alpha W_t(j)/m \rceil}{2 + W_t(j)}\right)\\ &\leq \frac{-\lceil \alpha W_t(j)/m \rceil}{2 + W_t(j)}\\ &\leq \frac{-\alpha}{2m},\end{aligned}$$

where the last step follows easily by considering the two cases $W_t(j) = 1$ and $W_t(j) \geq 2$.

Recalling the definition of $\Phi(i,\cdot)$, we infer by linearity of expectation that

$$\begin{aligned}\mathbb{E}\,\Phi(i,t+1) &= \sum_{j<i} \mathbb{E}\left(w_{t+1}(i,j)\ln(2 + W_{t+1}(j))\right)\\ &\leq \Phi(i,t) - \alpha\beta W_t(i)\frac{\alpha}{2m}\\ &\leq \Phi(i,t)\left(1 - \frac{\alpha^2}{2m^2\ln(i)}\right),\end{aligned}$$

which completes the proof. □

COROLLARY 5.2. *Let i be any player. Let $R_{\text{lucky}}(i)$ denote the total regret of Player i on rounds when she is lucky. Then, for any $\delta > 0$,*

$$\Pr\left(R_{\text{lucky}}(i) \geq \frac{2m^2\ln(i)}{\alpha^2}\ln\left(\frac{(i-1)\ln(i)}{\delta\ln(2)}\right)\right) \leq \delta.$$

PROOF. By Lemma 5.1, on any lucky round for Player i, either $R_{i,t} = 0$ or

$$\mathbb{E}\,\Phi(i,t+1) \leq \left(1 - \frac{\alpha^2}{2m^2\ln(i)}\right)\Phi(i,t).$$

Moreover, if $\Phi(i,t) = 0$ then $R_{i,t} = 0$. Thus, it suffices to bound the probability that $\Phi(i,t)$ stays positive for at least $N = \frac{2m^2\ln(i)}{\alpha^2}\ln\left(\frac{(i-1)\ln(i)}{\delta\ln(2)}\right)$ rounds of regret. But, by Markov's inequality, this probability is bounded by

$$\begin{aligned}\Pr\left(R_{\text{lucky}}(i) \geq N\right) &\leq \Pr\left(\Phi(i,t_N) \neq 0\right)\\ &\leq \Pr\left(\Phi(i,t_N) \geq \ln(2)\right)\\ &\leq \frac{\mathbb{E}\,\Phi(i,t_N)}{\ln(2)}\\ &\leq \left(1 - \frac{\alpha^2}{2m^2\ln(i)}\right)^N \frac{\Phi(i,0)}{\ln(2)}\\ &\leq \delta \text{ by choice of } N.\end{aligned}$$

This completes the proof. □

LEMMA 5.3. *Suppose $\alpha = 1/4$. Let i be any player, and let A be a positive integer. Let X denote the number of rounds when $W(i) \geq A$ but player i is not Lucky. Then, for every $\delta > 0$,*

$$\Pr\left(X \geq T\mathrm{e}^{2-A/8m} + \ln(1/\delta)\right) \leq \delta.$$

PROOF. Fix a time t, and let η denote the number of Player i's advisors who choose a Good option. This is the sum of $W_t(i)$ independent $\{0,1\}$-valued indicator variables, each with expectation β, where $\beta \geq 1/m$ denotes the fraction of options which are Good for Player i. (Note that when $\beta = 0$, Player i is automatically Lucky.) Hence, by a multiplicative form of Chernoff's bound,

$$\Pr\left(\eta \leq \frac{1}{2}\beta W_t(i)\right) \leq \mathrm{e}^{-\beta W_t(i)/8} \leq \mathrm{e}^{-A/8m}, \qquad (1)$$

assuming that $W_t(i) \geq A$. By linearity of expectation, this implies that $\mathbb{E}\, X \leq T\mathrm{e}^{-A/8m}$. On the other hand, since (1) did not depend on the history prior to round t, X is stochastically dominated by the number of heads in a sequence of independent coin flips, each with probability $\mathrm{e}^{-A/8m}$ of heads. Applying a second version of Chernoff's bound now implies, for every $B \geq 0$,

$$\begin{aligned}
&\Pr\left(X \geq T\mathrm{e}^{2-A/8m} + \ln(1/\delta)\right) \\
&\leq \left(\frac{T}{T\mathrm{e}^{2-A/8m} + \ln(1/\delta)}\right)\mathrm{e}^{-(A/8m)(T\exp(2-A/8m)+\ln(1/\delta))} \\
&\leq \left(\frac{\mathrm{e}^{1-A/8m}T}{\mathrm{e}^{2-A/8m}T}\right)^{T\exp(2-A/8m)+\ln(1/\delta)} \\
&\leq \delta.
\end{aligned}$$

which completes the proof. □

LEMMA 5.4. *Let i be any player, let A be a non-negative integer, and suppose $\alpha = 1/4$. Then for every $\delta > 0$,*

$$\Pr\Big(W_T(i) \geq A \text{ and } R_i \geq T\mathrm{e}^{2-A/8m} + \ln(2/\delta) + 32m^2\ln(i)\ln\left(\frac{2(i-1)\ln(i)}{\delta\ln(2)}\right)\Big) \leq \delta.$$

PROOF. The desired conclusion follows by applying Corollary 5.2 and Lemma 5.3, with each contributing an error probability of $\delta/2$. □

COROLLARY 5.5. *Let S be a set of honest players with identical preferences. Then*

$$\mathbb{E}\, R = O(|S|m^2\log^2 n + mT),$$

where R denotes the combined total regret of players in S. Indeed, for every $\delta > 0$,

$$\Pr\Big(R > T\mathrm{e}^2(8m+1) + |S|\ln(2|S|/\delta) + 32|S|m^2\ln(n)\ln\left(\frac{2|S|n\ln(n)}{\delta\ln(2)}\right)\Big) \leq \delta.$$

PROOF. Note that since the players in S are honest and have identical preferences, they can never eliminate one another as advisors. Hence, for every non-negative integer A, the $(A+1)$'st player in S must have $W_T \geq A$. Apply Lemma 5.4 to each player in S, using error probability $\delta/|S|$. Summing up the results and simplifying gives the desired high-probability bound. □

Note that Theorem 2.1 is just the expectation part of Corollary 5.5.

6. REFERENCES

[1] Noga Alon, Baruch Awerbuch, Yossi Azar, and Boaz Patt-Shamir. Tell me who i am: An interactive recommendation system. In *ACM SPAA*, 2006.

[2] B. Awerbuch and C. Scheideler. Enforced Spreading: An improved protocol for provably secure distributed name service. Technical report, Johns Hopkins University, February 2004.

[3] B. Awerbuch and C. Scheideler. Group Spreading: A protocol for provably secure distributed name service. In *proc 31st ICALP*, 2004.

[4] Baruch Awerbuch, Yossi Azar, Zvi Lotker, Boaz Patt, and Mark Tuttle. Collaborate with strangers to find your preferences. In *proceedings of ACM-SPAA*, 2005.

[5] Baruch Awerbuch, David Holmer, Herbert Rubens, and Robert Kleinberg. Provably competitive adaptive routing. In *Infocom*. IEEE, 2005.

[6] Baruch Awerbuch and Robert Kleinberg. Collaborative learning. In *Proceedings ACM COLT (Computational Learning Theory)*, 2005.

[7] Baruch Awerbuch, Boaz Patt-Shamir, David Peleg, and Mark Tuttle. Collaboration of untrusting peers with changing interests. In *Proc. 5th ACM conference on Electronic Commerce*, pages 112–119. ACM Press, 2004.

[8] Baruch Awerbuch, Boaz Patt-Shamir, David Peleg, and Mark Tuttle. Adaptive collaboration in peer-to-peer systems. In *25th International Conference on Distributed Computing Systems (ICDCS)*. IEEE, June 2005.

[9] Baruch Awerbuch, Boaz Patt-Shamir, David Peleg, and Mark Tuttle. Improved recommendation systems. In *Proc. 16th ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1174–1183, January 2005.

[10] Baruch Awerbuch and Christian Scheideler. Robust distributed name service. In *Proc. of 3rd International Workshop on Peer-to-Peer Systems (IPTPS)*, feb 2004.

[11] Baruch Awerbuch and Christian Scheideler. Towards a scalable and robust dht. In *Proceedings of ACM SPAA*, 2006.

[12] Baruch Awerbuch and Christian Scheideler. Robust random number generation for peer-to-peer systems. In *In Proc. ACM OPODIS*, 2007.

[13] Baruch Awerbuch and Christian Scheideler. Towards scalable and robust overlay networks. In *Proceedings of IPTPS*, 2007.

[14] J. M. Barzdin and R. V. Freivalds. On the prediction of general recursive functions. *Soviet Mathematics Doklady*, 13:1224–1228, 1972.

[15] N. Littlestone and M. Warmuth. The weighted majority algorithm. *Information and Computation*, 108(2):212–260, 1994.