Market-aware Proactive Skill Posting

Ashiqur R. KhudaBukhsh, Jong Woo Hong, and Jaime G. Carbonell

Carnegie Mellon University
{akhudabu, jongwooh, jgc}@cs.cmu.edu

Abstract Referral networks consist of a network of experts, human or automated agent, with differential expertise across topics and can redirect tasks to appropriate colleagues based on their topic-conditioned skills. *Proactive skill posting* is a setting in referral networks, where agents are allowed a one-time local-network-advertisement of a subset of their skills. Heretofore, while advertising expertise, experts only considered their own skills and reported their strongest skills. However, in practice, tasks can have varying difficulty levels and reporting skills that are uncommon or rare may give an expert relative advantage over others, and the network as a whole better ability to solve problems. This work introduces *market-aware proactive skill posting* where experts report a subset of their skills that give them competitive advantages over their peers. Our proposed algorithm in this new setting, proactive-DIEL$_\Delta$, outperforms the previous state-of-the-art, proactive-DIEL$_t$ during the early learning phase, while retaining important properties such as tolerance to noisy self-skill estimates, and robustness to evolving networks and strategic lying.

Keywords: active learning; referral networks; proactive skill posting

1 Introduction

A referral network [1] consists of experts, human or autonomous agents, where each expert (teacher, worker, agent) can redirect difficult tasks to appropriate expert colleagues. Such networks draw inspiration from real-world examples of networks of physicians or consultancy firms. The *learning-to-refer* challenge involves estimating topic-conditioned expertise of colleagues in a referral network in an active learning framework.

In this paper, we focus on *proactive skill posting* [2], a setting where experts perform a one-time (only at the beginning of simulation or when they join a network) local-network advertisement to network-connected colleagues of a subset of their skills, focusing on the expert top skills. In the real world, such skill advertisements are common as experts often tell their colleagues about tasks they are good at and often forge links with colleagues via social networks. However, such (potentially noisy) priors are private information, and experts may strategically lie or unknowingly overestimate or underestimate their own skills to attract more referrals, i.e., profit from more business. Hence, a key component of algorithms designed for this setting is a mechanism to elicit truthful reporting of skills to make the system incentive compatible. However, when human experts select the set of topics to advertise, they not only consider their true absolute expertise at those particular topics, but also take their relative advantage over others into account. For instance, in a network of physicians, an accomplished brain surgeon would want to publish the fact that she is skilled at brain surgery, even though her success rate at diagnosing the common cold could be much higher. In a similar vein, in this work, we introduce the notion of *market-aware skill posting*, i.e. experts posting skills on topics they have relative advantage over others and propose proactive-$\mathtt{DIEL}_\Delta$ for this setting.

Our key contributions are the following: First, we introduce *market-aware proactive skill posting* in referral networks, previously not addressed in the proactive skill posting literature. Second, we perform extensive empirical evaluations on existing data sets comparing against proactive-$\mathtt{DIEL}_t$, the known state-of-the-art, and demonstrate that our newly-introduced algorithm, proactive-$\mathtt{DIEL}_\Delta$, outperforms proactive-$\mathtt{DIEL}_t$ in terms of early learning-phase advantage. We construct additional data sets with larger variance in task-difficulty and show that the performance gap between proactive-$\mathtt{DIEL}_\Delta$ and proactive-$\mathtt{DIEL}_t$ widens in the presence of tasks with varying difficulty levels. Finally, we show that like its predecessors, proactive-$\mathtt{DIEL}_\Delta$ is robust to strategic lying, evolving networks, and noisy self-skill estimates.

Related work: The referral framework draws inspiration from referral chaining, first proposed in [3] and subsequently extended in [4, 5] (for further relevant literature, see, e.g., [1]). Recent research in referral networks has focused on three broad directions: identifying key algorithms and evaluating relative performance on uninformed prior settings [1], designing algorithms immune to strategic lying that incorporate partially available (potentially noisy) priors [2, 6], and robustness to practical factors such as evolving networks, capacity constraints [1] and time-varying expertise [7].

Our work on *proactive skill posting* is related to the bandit literature with *side-information* [8, 9] in the sense that algorithms in this setting do not start from scratch, but have leg-up based on task-relevant information. However, a key difference is that, instead of obtaining that side data from observed trials [9] or shape of the reward distribution [10], the *side-information* in our case is obtained in a decentralized manner through advertisement of skills by the experts themselves, who may in fact willfully misreport to attract more business. This ties our work broadly to the vast literature in adversarial machine learning [11] and truthful mechanism design [12, 13, 14]. For further relevant literature, see, e.g., [2, 6].

2 Referral Networks

We summarize the basic notation, definitions, and assumptions and provide necessary background for *market-aware proactive skill posting*.

Referral network: Represented by a graph (V, E) of size k in which each vertex v_i corresponds to an expert e_i ($1 \leq i \leq k$) and each bidirectional edge $\langle v_i, v_j \rangle$ indicates a *referral link* which implies e_i and e_j can co-refer problem instances.

Subnetwork: of an expert e_i: The set of experts linked to an expert e_i by a referral link.

Referral scenario: Set of m instances $(q_1, \ldots, q_m)$ belonging to n topics $(t_1, \ldots, t_n)$ addressed by the k experts $(e_1, \ldots, e_k)$ connected through a referral network (V, E).
Expertise: Expertise of an expert/question pair $\langle e_i, q_j \rangle$ is the probability with which e_i can solve q_j.
Referral mechanism: Following previous proactive skill posting literature [2, 6], for a per-instance query budget Q, we kept fixed to $Q = 2$ across all our current experiments. The referral mechanism consists of the following steps.

1. A user issues an *initial query* q_j to a randomly chosen *initial expert* e_i.
2. The initial expert e_i examines the instance and solves it if possible. This depends on the *expertise* of e_i wrt. q_j.
3. If not, a *referral query* is issued by e_i to a *referred expert* e_j within her subnetwork, with a query budget of $Q-1$. *Learning-to-refer* involves improving the estimate of who is most likely to solve the problem.
4. If the referred expert succeeds, she sends the solution to the initial expert, who sends it to the user.

A detailed description of our assumptions can be found in [1, 2]. Some of the important assumptions are: the network connectivity depends on (cosine) similarity between the topical expertise (guided by the observation that experts with similar expertise are more likely to know each other), and the distribution of topical-expertise across experts can be characterized by a mixture of Gaussian distributions. Note that, in our model, it is still possible that experts with very little overlap in skills are connected for reasons beyond similar expertise (e.g., same geolocation, common acquaintances, etc.), making them prime candidates for referrals. Also, an expert pair with substantial overlap in expertise areas may still have specific topics where one expert is stronger than the other, making referrals mutually beneficial. For topical-expertise distribution, a mixture of two Gaussians was used to represent the expertise of experts with specific training for the given topic (higher mean, lower variance), contrasted with the lower-level expertise (lower mean, higher variance) of the layman population.

We now present necessary background for *market-aware proactive skill posting* and describe what distinguishes it from traditional *proactive skill posting.*
Advertising unit: a tuple $\langle e_i, e_j, t_k, \mu_{t_k} \rangle$, where e_i is the *target expert*, e_j is the *advertising expert*, t_k is the topic and μ_{t_k} is e_j's (advertised) topical expertise. Similar to rewards in our uninformative prior setting, an advertising unit is also locally visible, i.e., only the *target expert* gets to see the advertised prior for a given unit.
Advertising budget: In practice, experts have limited time to socialize with different colleagues and get to know each other's experience. We incorporate this phenomenon through the notion of budget and assume each expert is allocated a budget of B advertising units, where B is twice the size of that expert's subnetwork. Effectively means that each expert reports her top two skills to everyone in her subnetwork.
Explicit and implicit bid: A topic that is advertised in an advertising protocol is an explicit bid. Similarly, a topic that is not advertised, for which an upper

bound can be assumed, is an implicit bid. Top two skills are differently defined in traditional proactive skill posting and market-aware skill posting. We describe the difference in the advertising protocol next.

Advertising protocol: a one-time advertisement that happens right at the beginning of the simulation or when an expert joins the network. The advertising expert e_j reports to each target expert e_i in her subnetwork the two tuples $\langle e_i, e_j, t_{best}, \mu_{t_{best}} \rangle$ and $\langle e_i, e_j, t_{secondBest}, \mu_{t_{secondBest}} \rangle$, i.e., the top two topics in terms of the advertising expert's topic means.

Now, we describe the primary distinction between traditional *proactive skill posting* and *market-aware proactive skill posting*. In traditional proactive skill posting, for a given expert, $\mu_{t_{best}}$ is simply her maximum topical expertise. In *market-aware proactive skill posting*, we propose that every expert has access to an estimate of $\overline{\mu_{t_k}}$ (average network skill on each topic t_k) and reports the skills with her largest relative advantage μ_{Δ} (where for a given expert/topic pair $\langle e_j, t_k \rangle$, $\mu_{\Delta_{e_j,t_k}} = \mu_{e_j,t_k} - \overline{\mu_{t_k}}$).

Next, we illustrate the difference described above with the following example. Consider a referral network of N experts and across five different topics, $t_1, \ldots, t_5$, the average network expertise are respectively, 0.1, 0.3, 0.8, 0.9, 0.4. Now, consider an expert e whose expertise on the aforementioned five topics are 0.4, 0.3, 0.7, 0.65, 0.5, respectively. In traditional proactive skill posting, for every colleague e_i of e, e will have the following two advertising units: $\langle e_i, e, t_3, 0.7 \rangle$ and $\langle e_i, e, t_4, 0.65 \rangle$, reporting her skills on t_3 and t_4, the two topics she has highest expertise in an absolute scale. However, notice that e is unlikely to have any substantial relative advantage over other in those two topics as her expertise on those two topics is sufficiently lower than the network expertise. Also, t_1 is the hardest topic where average expertise of the network is 0.1 and e is relatively stronger in t_1 with $\mu_{\Delta_{e,t_1}} = 0.3$. Hence, in a market-aware skill posting setting, e will report her skills in t_1 and t_5, the two topics where she has relative advantage with the two advertisement units $\langle e_i, e, t_1, 0.4 \rangle$ and $\langle e_i, e, t_5, 0.5 \rangle$.

3 Distributed Referral Learning

Considering a single expert and a given topic, learning-to-refer is an action selection problem, and each expert maintains an action selection thread for each topic in parallel. If we think in the context of multi-armed bandit (MAB), an action or arm corresponds to a referral choice, i.e., picking an appropriate expert from the subnetwork. In order to describe an action selection thread, we first name the topic T and expert e. Let $q_1, \ldots, q_N$ be the first N *referral queries* belonging to topic T issued by expert e to any of her K colleagues in her subnetwork denoted by $e_1, \ldots, e_K$. For each colleague e_i, e maintains a reward vector $\mathbf{r}_{i,n_i}$ where $\mathbf{r}_{i,n_i} = (r_{i,1}, \ldots, r_{i,n_i})$, i.e., the sequence of rewards observed from expert e_i on issued n_i referred queries. Understandably, $N = \sum_{i=1}^{K} n_i$. Let $m(e_i)$ and $s(e_i)$ denote the sample mean and sample standard deviation of $\mathbf{r}_{i,n_i}$.

DIEL: First proposed in [15], Interval Estimation Learning (`IEL`) has been extensively used in stochastic optimization and action selection problems. Action

Algorithm 1: DIEL(e, T)

Initialization: $\forall i, n_i \leftarrow 2$, $\mathbf{r}_{i,n_i} \leftarrow (0,1)$
Loop: Select expert e_i who maximizes

$$score(e_i) = m(e_i) + \frac{s(e_i)}{\sqrt{n_i}}$$

Observe reward r
Update $\mathbf{r}_{i,n_i}$ with r, $n_i \leftarrow n_i + 1$

selection using Distributed Interval Estimation Learning (DIEL) works in the following way [2]. As described in Algorithm 1, at each step, DIEL [2] selects the expert e_i with highest $m(e_i) + \frac{s(e_i)}{\sqrt{n_i}}$. The intuition is that high mean selects for best performance, and high variance selects for unexplored expert capability on topic, thus optimizing for amortized performance, as variance decreases over time, and best mean is selected reliably among the top candidates. DIEL operates in an uninformed prior setting, and every action is initialized with two rewards of 0 and 1, allowing us to initialize the mean and variance and making all experts equally likely to get picked in the beginning.

We now describe proactive-DIEL$_\Delta$, our proposed new algorithm and proactive-DIEL$_t$, our baseline.

3.1 Initialization

proactive-DIEL$_t$ initialization: Rather than DIEL sets $reward(e_i, t_k, e_j)$ for each i, j and k with a pair $(0, 1)$ in order to initialize mean and variance, proactive-DIEL$_t$ initializes $reward(e_i, t_k, e_j)$ for each advertisement unit $\langle e_i, e_j, t_k, \mu_{t_k}\rangle$ with two rewards of μ_{t_k} (explicit bid).

To initialize topics for which no advertisement units are available (implicit bid), we initialize the rewards as if the expert's skill was the same as on her second best topic, that is, with two rewards of $\mu_{t_{secondBest}}$, effectively being an upper bound on the actual value.

proactive-DIEL$_\Delta$ initialization: A similar prior-bounding technique can be used in proactive-DIEL$_\Delta$ with the following modification. Recall that, each expert has knowledge about $\overline{\mu_{t_k}}, \forall k$ – the average network expertise across all topics. Let t^{Δ}_{best} and $t^{\Delta}_{secondBest}$ be the two explicit bids for an expert e with $\mu_{t^{\Delta}_{secondBest}} - \overline{\mu}_{t^{\Delta}_{secondBest}} \leq \mu_{t^{\Delta}_{best}} - \overline{\mu}_{t^{\Delta}_{best}}$, and $t^{\Delta}_{implicit}$ be any implicit bid. The following inequality holds,

$$\mu_{t^{\Delta}_{implicit}} - \overline{\mu}_{t^{\Delta}_{implicit}} \leq \mu_{t^{\Delta}_{secondBest}} - \overline{\mu}_{t^{\Delta}_{secondBest}} \tag{1}$$

since the relative advantage of any implicit bid must be less than or equal to the relative advantage of $t^{\Delta}_{secondBest}$. Rearranging equation 1, we get

$$\mu_{t^{\Delta}_{implicit}} \leq \mu_{t^{\Delta}_{secondBest}} + \overline{\mu}_{t^{\Delta}_{implicit}} - \overline{\mu}_{t^{\Delta}_{secondBest}} \tag{2}$$

All terms of the right-hand side of the equation 2 are known, and every implicit bid is initialized with two rewards of $\mu_{t^{\Delta}_{secondBest}} + \overline{\mu}_{t^{\Delta}_{implicit}} - \overline{\mu}_{t^{\Delta}_{secondBest}}$.

3.2 Penalty on Distrust

proactive-$\texttt{DIEL}_{\Delta}$ follows the same penalty mechanism, *Penalty on Distrust*, as proactive-$\texttt{DIEL}_t$[1]. In this approach, in addition to assigning a binary reward depending on the task outcome, we assign a penalty. For a given instance, if the reward is r and the penalty p, the effective reward is $r - p$. The assigned penalty incorporates a factor we may call *distrust*, as it estimates a likelihood the expert is lying, given our current observations. Further details can be found in [6].

4 Experimental Setup

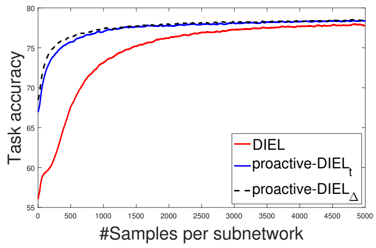
Baselines: We used DIEL (the parameter-free version first presented in [2]), the known best-performing algorithm in uninformed prior setting, and proactive-$\texttt{DIEL}_t$, the best-performing algorithm in proactive skill setting as our baselines.
Data set: Our test set for performance evaluation is the same data set, $\mathcal{D}$, used in [1, 2][2], which consists of 1000 *referral scenarios*. Each *referral scenario* consists of a network of 100 experts and 10 topics.

In addition to $\mathcal{D}$, we constructed two data sets inducing larger variance in task-difficulty. Recall that, for topical-expertise distribution, we consider a mixture of two Gaussians (with parameters $\lambda = \{w_i^t, \mu_i^t, \sigma_i^t\}$ $i = 1, 2.$). One of them ($\mathcal{N}(\mu_2^t, \sigma_2^t)$) has a greater mean ($\mu_2^t > \mu_1^t$), smaller variance ($\sigma_2^t < \sigma_1^t$) and lower mixture weight ($w_2^t << w_1^t$). For a given topic t_i, we modify the topical expertise for all experts in the following way:
$\mu_{t_i,e_j} = d_{factor}\ \mu_{t_i,e_j} \forall j$, where $d_{factor} \sim U[C, 1]$. The multiplicative factor ensures that the initial property of being sampled from a mixture of Gaussians holds. Different values for the parameter C allows us to vary the difficulty level of a task. We generated two additional data sets using $C = 0.25$ (denoted as $\mathcal{D}_{0.25}$) and $C = 0.5$ (denoted as $\mathcal{D}_{0.5}$).
Performance Measure: Following previous proactive skill posting literature [2, 6], we use two performance measures – overall task accuracy of our multi-expert system and *ICFactor*, an empirical measure for evaluating Bayesian-Nash incentive compatibility (a weaker form of incentive compatibility where being truthful is weakly better than lying). If a network receives n tasks of which m tasks are solved (either by the *initial expert* or the *referred expert*), the overall task accuracy is $\frac{m}{n}$. As an empirical measure for evaluating Bayesian-Nash incentive compatibility, we use *ICFactor* (described in [2]); an *ICFactor* value greater than 1 implies truthfulness in expectation, i.e., truthful reporting fetched more referrals than strategic lying. For evolving networks and noisy skill estimates, we used the same experimental setting as described in [6].

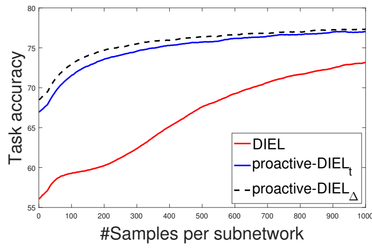
5 Results

Figure 1 and Table 1 summarize main experimental results of our proposed new algorithm, proactive-$\texttt{DIEL}_{\Delta}$. The results highlight the following. First, Figure 1(a) demonstrates that we can use similar prior-bounding technique to

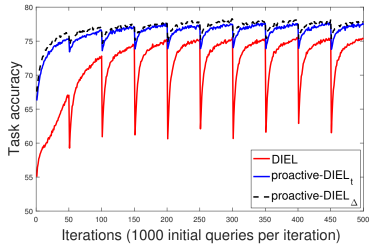
[1] The subscript t stands for trust.
[2] The data set can be downloaded from https://www.cs.cmu.edu/~akhudabu/referral-networks.html



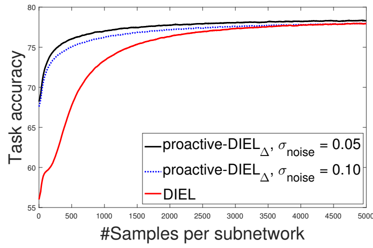
(a) Overall performance



(b) Early learning advantage



(c) Performance on evolving networks

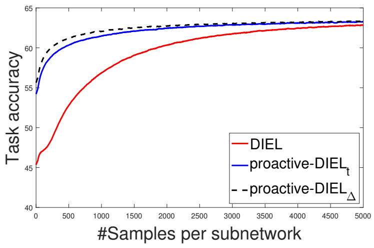


(d) Noisy skill estimates

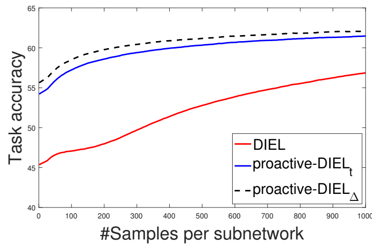
Figure 1. Performance comparison on data set $\mathcal{D}$.

initialize proactive-DIEL$_\Delta$ and a reward adjustment mechanism described in [6] and obtain improved cold-start performance than proactive-DIEL$_t$, the known state-of-the-art. Particularly, as shown in Figure 1(b), the advantage during the early phase of learning is superior which widens as the tasks get harder (see, Figure 2). A paired t-test reveals that during the early learning phase (1000 samples or less per subnetwork), proactive-DIEL$_\Delta$ is better than both the baselines with p-value less than 0.0001. Second, in a real-world setting, it is easy to imagine situations where new entrants will join a network while old members leave. In a situation, where at regular interval a sizable chunk of the network composition changes, Figure 1(c) shows that the early-phase learning advantage translates into better adaptation to evolving networks. Third, our proposed algorithm is tolerant to noisy self-skill estimates (shown in Figure 1(d)), thus absolutely accurate estimates of own skill, which is an impractical assumption, is not particularly necessary for the algorithm to succeed. Table 1 lists *ICFactor* for all possible strategy combinations of reporting an expert's top two skills. We found that, across all three data sets, proactive-DIEL$_\Delta$ exhibited robustness to strategic lying.

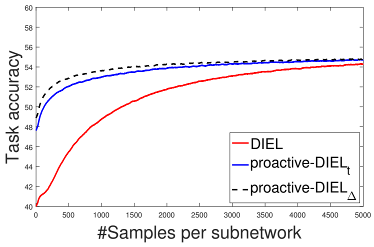
We conclude our results section with a summary of our main results and an outlook to future research directions. Our results demonstrate: (1) Market-aware skill posting gives larger early-learning phase advantage than previously known state-of-the-art, (2) our proposed algorithm, proactive-DIEL$_\Delta$ is robust to noise in self-skill estimates and strategic lying, and (3) the performance gap between



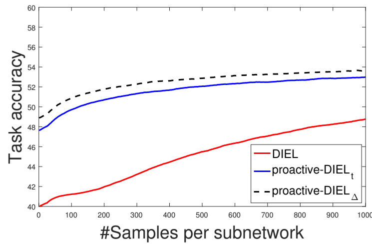
(a) Overall performance ($\mathcal{D}_{0.5}$)



(b) Early learning advantage ($\mathcal{D}_{0.5}$)



(c) Overall performance ($\mathcal{D}_{0.25}$)



(d) Early learning advantage ($\mathcal{D}_{0.25}$)

Figure 2. Performance comparison on data sets $\mathcal{D}_{0.25}$ and $\mathcal{D}_{0.50}$.

Table 1. Empirical analysis of Bayesian-Nash incentive-compatibility on three data sets, $\mathcal{D}$, $\mathcal{D}_{0.5}$ and $\mathcal{D}_{0.25}$. Each row represents a specific combination of strategies an expert can use to report her best and second-best skills. All values indicate that being truthful is no worse than lying.

$\mu_{t^\Delta_{best}}$	$\mu_{t^\Delta_{secondBest}}$	$\mathcal{D}$	$\mathcal{D}_{0.5}$	$\mathcal{D}_{0.25}$
Truthful	Overbid	1.0174	1.0006	1.0026
Overbid	Truthful	1.1892	1.4393	1.7843
Overbid	Overbid	1.2159	1.7993	1.7620
Truthful	Underbid	1.0976	1.0617	1.0860
Underbid	Truthful	1.1301	1.1586	1.1333
Underbid	Underbid	1.3672	1.1900	1.1588
Underbid	Overbid	1.1213	1.0982	1.1404
Overbid	Underbid	1.2835	1.4303	1.4616

proactive-DIEL$_\Delta$ and proactive-DIEL$_t$ widens when we use larger variance in task difficulty (4) the early learning-phase advantage is particularly useful in evolving networks. Future research directions may include: (1) extending market-aware skill posting to other MAB algorithms (e.g., ϵ-Greedy, Q-Learning) (2) in presence of a larger amount of noise, or when $\overline{\mu_{t_k}}$ is not known, regularizing the advertised priors relative to the subnetwork and (3) designing proactive skill posting algorithms robust to time-varying expertise.

Bibliography

[1] KhudaBukhsh, A.R., Carbonell, J.G., Jansen, P.J.: Robust learning in expert networks: a comparative analysis. Journal of Intelligent Information Systems (2018)

[2] KhudaBukhsh, A.R., Carbonell, J.G., Jansen, P.J.: Proactive Skill Posting in Referral Networks. In: Australasian Joint Conference on Artificial Intelligence, Springer (2016) 585–596

[3] Kautz, H., Selman, B., Milewski, A.: Agent amplified communication. (1996) 3–9

[4] Yolum, P., Singh, M.P.: Dynamic communities in referral networks. Web Intelligence and Agent Systems **1**(2) (2003) 105–116

[5] Yu, B.: Emergence and evolution of agent-based referral networks. PhD thesis, North Carolina State University (2002)

[6] KhudaBukhsh, A.R., Carbonell, J.G., Jansen, P.J.: Incentive compatible proactive skill posting in referral networks. In: European Conference on Multi-Agent Systems. Springer. (2017)

[7] KhudaBukhsh, A.R., Carbonell, J.G.: Expertise drift in referral networks. In: Proceedings of the 17th International Conference on Autonomous Agents and MultiAgent Systems, International Foundation for Autonomous Agents and Multiagent Systems (2018) 425–433

[8] Langford, J., Strehl, A., Wortman, J.: Exploration scavenging. In: Proceedings of the 25th international conference on Machine learning, ACM (2008) 528–535

[9] Shivaswamy, P., Joachims, T.: Multi-armed bandit problems with history. In: Artificial Intelligence and Statistics. (2012) 1046–1054

[10] Bouneffouf, D., Feraud, R.: Multi-armed bandit problem with known trend. Neurocomputing **205** (2016) 16–21

[11] Huang, L., Joseph, A.D., Nelson, B., Rubinstein, B.I., Tygar, J.: Adversarial machine learning. In: Proceedings of the 4th ACM workshop on Security and artificial intelligence, ACM (2011) 43–58

[12] Babaioff, M., Sharma, Y., Slivkins, A.: Characterizing truthful multi-armed bandit mechanisms. SIAM Journal on Computing **43**(1) (2014) 194–230

[13] Tran-Thanh, L., Stein, S., Rogers, A., Jennings, N.R.: Efficient crowdsourcing of unknown experts using multi-armed bandits. In: European Conference on Artificial Intelligence. (2012) 768–773

[14] Biswas, A., Jain, S., Mandal, D., Narahari, Y.: A truthful budget feasible multi-armed bandit mechanism for crowdsourcing time critical tasks. In: Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems, International Foundation for Autonomous Agents and Multiagent Systems (2015) 1101–1109

[15] Kaelbling, L.P.: Learning in embedded systems. MIT press (1993)