ECAI 2016
G.A. Kaminka et al. (Eds.)
© 2016 The Authors and IOS Press.
This article is published online with Open Access by IOS Press and distributed under the terms of the Creative Commons Attribution Non-Commercial License 4.0 (CC BY-NC 4.0).
doi:10.3233/978-1-61499-672-9-1620

Distributed Learning in Expert Referral Networks

Ashiqur R. KhudaBukhsh and **Peter J. Jansen** and **Jaime G. Carbonell**[1]

Abstract. Human experts or autonomous agents in a referral network must decide whether to accept a task or refer to a more appropriate expert, and if so to whom. In order for the referral network to improve over time, the experts must learn to estimate the topical expertise of other experts. This paper extends concepts from Reinforcement Learning and Active Learning to referral networks, to learn how to refer at the network level, based on the proposed distributed interval estimation learning (DIEL) algorithm. Diverse Monte Carlo simulations reveal that DIEL improves network performance significantly over both greedy and Q-learning baselines [3], approaching optimal given enough data.

1 INTRODUCTION

Consider a network of experts with differing expertise, where any expert may receive a problem (aka a task or a query) and must decide whether to work on it or to refer the problem, and if so to which other expert. How can a network, or its individual experts, learn how to refer tasks effectively?

This paper proposes a new Distributed Active Learning approach in referral networks. Our referral model assumes an initial sparse topology of a referral graph where each expert knows a handful of colleagues so that $E \sim O(V)$ (E and V denote the number of edges and vertexes in the network, respectively). Learning consists of each expert improving its estimates of the ability of colleagues to solve different classes of problems. We address learning to refer comparing overall network performance contrasting an exploitation-centered (greedy optimization) with a balanced exploration-exploitation tradeoff (amortized optimization), showing that the former outperforms at first, and the latter overtakes as the network learns more effectively over time.

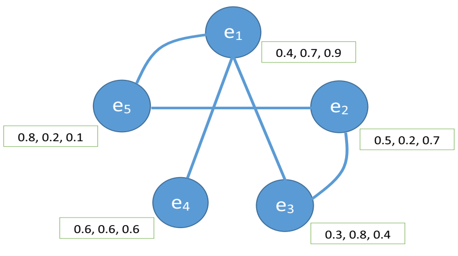


Figure 1. A referral network with five experts.

To illustrate the problem, consider the extremely simple graph, representing a five-expert network, shown in Figure 1. The nodes of the graph are the experts, and the edges indicate that the experts 'know' each other, that is, they can send or receive referrals and communicate results. We assume 3 different topics (subdomains) can be distinguished – call them t_1, t_2, and t_3 – and the figures in brackets indicate an expert's expertise in each of these.

In the example, with a query belonging to t_2, if there was no referral, the client may consult first e_2 and then possibly e_5, leading to a probability of getting the correct answer of $0.2 + (1 - 0.2) \times 0.2 = 0.36$. With referrals, an expert handles a problem she knows how to answer, and otherwise if she had knowledge of all the other experts' expertise she could ask e_2 who would refer to e_3 for the best skill in t_2, leading to a solution probability of $0.2 + (1 - 0.2) \times 0.8 = 0.84$.

Our referral mechanism consists of the following steps: 1) A user issues an *initial query* to an *initial expert*. 2) If the initial expert is able to solve it, she returns the solution; 3) if not, she selects a *referred expert* within her subnetwork, who solves or refers in turn. *Learning-to-refer* means improving the estimate of who is most likely to solve the problem.

Our primary contribution is the distributed learning-to-refer framework and a distributed learning algorithm. To learn referrals, we borrowed ideas from Reinforcement Learning [5] and Active Learning [1, 7], up till now rarely applied to referral networks, and compared performance under various conditions and learning algorithms. We extended Interval Estimation learning [4, 6] to DIEL, a distributed interval learning methods, and compare DIEL with well-known algorithms: ϵ-greedy Q-learning and double Q-learning, and with an upper-bound where every expert has access to an oracle that knows the true topic-mean and thus can refer optimally.

2 REFERRAL NETWORK

We first present our notation and assumptions.

- A set of m instances $(q_1, q_2, \ldots, q_m)$ belonging to n topics $(topic_1, topic_2, \ldots, topic_n)$ are to be addressed by k experts
- The experts are connected through a *referral network*, a graph (V, E) where each vertex v_i denotes an expert e_i and each edge $\langle v_i, v_j \rangle$ indicates a *referral link*. The probability of an edge is: $P(ReferralLink(v_i, v_j)) = \tau + c\ Sim(e_i, e_j)$. We used cosine similarity of topic means for *Sim*. The *subnetwork* of each expert e_i is the set of of all experts $\{e_j\}$ that e_i can refer to.
- The *expertise* for an expert-instance pair, $\langle e_i, q_j \rangle$, is the probability that she can successfully solve the problem, i.e., $Expertise(e_i, q_j) = P(solve(e_i, q_j))$.
  Topic-wise distributional assumption: We take the expertise distribution for a given topic t to be a mixture of two truncated 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 higher mean ($\mu_2^t > \mu_1^t$), smaller variance ($\sigma_2^t < \sigma_1^t$) and lower mixture weight ($w_2^t << w_1^t$).
  Instance-wise distributional assumption: We model the expertise of a given expert on instances under a topic by a truncated Gaussian distribution with small variance. i.e.,
  $Expertise(e_i, q_j) \sim \mathcal{N}(\mu_{topic_p, e_i}, \sigma_{topic_p, e_i})$,
  $\forall q_j \in topic_p, \forall p, i : \sigma_{topic_p, e_i} \leq 0.2$.

[1] Carnegie Mellon University, USA, email:{akhudabu, pjj, jgc}@cs.cmu.edu

3 DISTRIBUTED REFERRAL LEARNING

Action selection using Interval Estimation Learning [2] (IEL) estimates a the upper confidence interval for the mean reward by

$$UI(a) = m(a) + t_{\frac{\alpha}{2}}^{(n-1)} \frac{s(a)}{\sqrt{n}} \quad (1)$$

where m(a) is the mean observed reward for a, $s(a)$ is the sample standard deviation of the reward, n is the number of observed samples from a, and $t_{\frac{\alpha}{2}}^{(n-1)}$ is is the critical value for the Student's t-distribution ($n-1$ degrees of freedom, $\frac{\alpha}{2}$ confidence level) (in our case, the action is the selection of a referred expert among possible choices in the subnetwork). Next, IEL selects the action with the highest upper confidence interval.

Algorithm 1 performs a single referral (per-task query budget Q = 2). The function $expR_h(e', topic)$ estimates e''s topical expertise. DIEL (Distributed Interval Estimation Learning), and DMT (Distributed Mean-Tracking) differ in h, DIEL estimating reward by equation (1) and DMT by using the sample-mean.

Input: A set of k experts $e_1, e_2, ..., e_k$. A set of n topics $topic_1, topic_2, ..., topic_n$. A $k \times k$ referral network.
Initialize rewards.
for *iter* $\leftarrow$ 1 ***to*** *maxIter* **do**
  Assign instance q to an initial expert e randomly
  if e fails to solve q **then**
    topic $\leftarrow$ *getTopic*(q)
    expectedReward $\leftarrow$ 0
    bestExpert $\leftarrow$ 0
    for each expert e' in the subnetwork of e **do**
      if $expR_h(e', topic) \geq$ *expectedReward* **then**
        bestExpert $\leftarrow e'$
        expectedReward $\leftarrow expR_h(e', topic)$
      end
    end
  end
  referredExpert $\leftarrow$ *bestExpert*
  if *referredExpert* solves q **then**
    update(*reward*(e,*topic*,*referredExpert*),1)
  else
    update(*reward*(e,*topic*,*referredExpert*),0)
  end
end

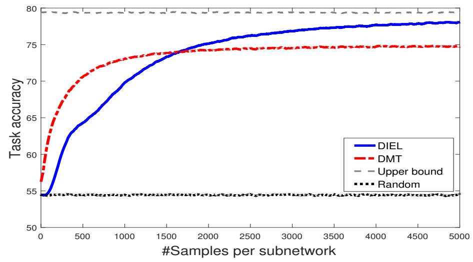
Algorithm 1: DISTRIBUTED REFERRAL LEARNING, Q = 2

4 EXPERIMENTAL SETUP AND RESULTS

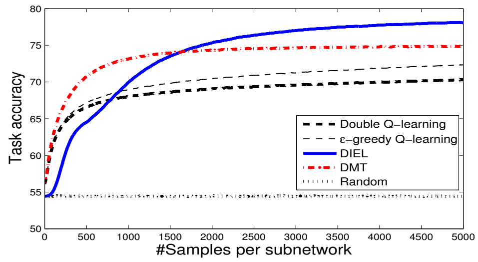
Parameter	Description	Distribution
τ	$P(ReferralLink(v_i, v_j))$	Uniform(0.01, 0.1)
c	$= \tau + c\,Sim(e_i, e_j)$.	Uniform(0.1,0.2)
μ_1	Truncated mixture of two	Uniform(0,b)
μ_2	Gaussians for topics	Uniform(b,1)
		$b \in \{0.1, 0.2, 0.3, 0.4, 0.5\}$
σ_1		Uniform(0.2,0.4)
σ_2		Uniform(0.05,0.15)
w_2		$\mathcal{N}(0.03, 0.01)$, $w_2 \geq 0$

Table 1. Parameters for our synthetic data set.

We evaluated the performance of our learning algorithms on on 1000 scenarios, each with 100 experts, 10 topics and a referral network, whose parameters are in Table 1. Our measure of performance is the overall task accuracy of the multi-expert network. As an upper bound we considered the performance achieved by a network where every expert has access to an oracle that knows the true topic-mean (i.e., $mean(Expertise(e_i, q) : q \in topic_p)\ \forall i, p$) of every expert-topic pair. Our baseline is a strategy where a task can be queried to maximum two randomly chosen experts.



(a) Comparison with baselines and upper bound.



(b) Comparison with Q-learning and DQ-learning.

Figure 2. Performance comparison of referral algorithms.

On every simulation, DIEL and DMT outperformed the baseline by a substantial margin. Over time DIEL clearly outperforms DMT and approaches the topical (oracle) upper bound. Figure 2(b) shows that both DIEL and DMT outperform Double Q-learning and ϵ-greedy Q-learning (optimized using a rough parameter sweep) for the duration of the experiment, although unlike DMT, Double Q-learning and ϵ-greedy Q-learning continue to improve.

REFERENCES

[1] P. Donmez, J. G. Carbonell, and P. N. Bennett, 'Dual Strategy Active Learning', *Machine Learning ECML 2007*, 116–127, (2007).

[2] P. Donmez, J. G. Carbonell, and J. Schneider, 'Efficiently learning the accuracy of labeling sources for selective sampling', *Proc. of KDD 2009*, 259, (2009).

[3] H. V. Hasselt, 'Double q-learning', in *Advances in Neural Information Processing Systems*, pp. 2613–2621, (2010).

[4] L. P. Kaelbling, *Learning in embedded systems*, MIT press, 1993.

[5] L. P. Kaelbling, M. L. Littman, and A. P. Moore, 'Reinforcement Learning: A Survey', *Journal of Artificial Intelligence Research*, **4**, 237–285, (1996).

[6] Andrew W Moore, Jeff Schneider, et al., 'Memory-based stochastic optimization', *Advances in Neural Information Processing Systems*, 1066–1072, (1996).

[7] M. Saar-Tsechansky and F. Provost, 'Active learning for class probability estimation and ranking', in *Proc. of IJCAI-01*, pp. 911–920, (2001).