Task Routing for Prediction Tasks

Haoqi Zhang*, **Eric Horvitz**†, **Yiling Chen***, and **David C. Parkes***
*Harvard SEAS
Cambridge, MA 02138, USA
{hq, yiling, parkes}@eecs.harvard.edu
†Microsoft Research
Redmond, WA 98052, USA
horvitz@microsoft.edu

ABSTRACT
We study principles and methods for task routing that aim to harness people's abilities to jointly contribute to a task and to route tasks to others who can provide further contributions. In the particular context of prediction tasks, the goal is to efficiently obtain accurate probability assessments for an event of interest. We introduce *routing scoring rules* for promoting collaborative behavior, that bring truthfully contributing information and optimally routing tasks into a Perfect Bayesian Equilibrium under common knowledge about agents' abilities. However, for networks where agents only have local knowledge about other agents' abilities, optimal routing requires complex reasoning over the history and future routing decisions of users outside of local neighborhoods. Avoiding this, we introduce a class of *local routing rules* that isolate simple routing decisions in equilibrium, while still promoting effective routing decisions. We present simulation results that show that following routing decisions induced by local routing rules lead to efficient information aggregation.

INTRODUCTION
Organizations rely on a mix of expertise and on means for identifying and harnessing expertise for completing different kinds of tasks. The ability to leverage the expertise and interests of individuals effectively is crucial for the success of an organization. Accomplishing a task may require the expertise of multiple actors, and harnessing that expertise requires identifying who the experts are, as well as providing proper incentives (monetary or otherwise) for inducing contributions.

The 'burden' of identifying experts for any given task can be placed entirely on a centralized policy, or rely on individuals within the organization, based on their own judgments. Both approaches have flaws. In the former, an organization or system may not know which individuals have the required expertise. In the latter, while individuals may often know their own expertise, they may not know about available tasks that match their respective expertise.

In social networks and organizations, an individual's knowledge extends beyond their own expertise on tasks and topics to knowledge about the expertise of others. For example, a manager may know people in their organization who can best tackle a particular aspect of a project. Members within the same research group know whom within that group can best review a paper, or best contribute to answering a research question. Likewise, people using a social network like Facebook may know who among their friends can best answer a particular question, or otherwise provide valuable opinions on a topic of discussion. Even in situations where an individual cannot identify an expert who can contribute to a task, they may know others who do know of experts for handling that task (but aren't necessarily experts themselves), or otherwise be able to identify subsets of individuals among whom the requisite expertise is likely to exist (e.g., people who share a particular interest).

We have been exploring principles and methods for *task routing* that aim to harness people's abilities to both contribute to a solution and to route tasks to others, who they know can also solve and route. Task routing provides an interesting paradigm for problem solving, in which individuals become engaged with tasks based on others' assessments of their expertise. On the task level, effective task routing aims to take advantage of agents' knowledge about solving problems as well as agents' knowledge about other agents' abilities to contribute, and to use both types of knowledge to contribute to solving a task—and to properly reward participating agents for their contributions. On an organizational level, task routing may provide a means for allocating tasks to individuals effectively, where routing suggestions and decisions take into account not only an individual's expertise on the particular task, but also their ability and availability to contribute as a solver and/or router on other tasks.

We shall highlight key notions, challenges, and opportunities with task routing by focusing on the problem of efficiently obtaining accurate probability assessments about the occurrence of an event or truth of a proposition of interest. With this challenge, a question is passed along individuals within a network, where each individual in the path can update the posterior probability based on their information and also forward the task to a neighbor. We provide a formal model for this setting, and introduce *routing scoring rules* for incentivizing contributions under which truthful reporting of posterior probability assessments and optimal routing to other individuals is a Perfect Bayesian Equilibrium. We assume in this model that there is common knowledge of others' levels of expertise, and show that the incentive and optimality guarantees hold for general networks. However, while on cliques the optimal routing decisions are myopic, they require more generally finding the best chain of individuals moving forward, which is NP-hard.

Common knowledge is unlikely to hold for large social networks where one's information about others is limited to a local neighborhood (e.g., friends, and maybe friends of friends). To handle such cases, we also consider task routing where knowledge about others' abilities may only be locally shared. Unfortunately, optimal routing under the routing rules becomes even more computationally challenging, and require agents to perform complex reasoning over the history and future routing decisions of agents outside the local neighborhood. Avoiding this, we introduce a special class of *local routing rules* that isolate simple routing decisions in equilibrium, while still taking advantage of knowledge about the expertise of others to promote effective routing decisions. We achieve this by incentivizing agents to make routing decisions based on short, locally optimal paths, that can be easily computed using only the shared local knowledge. *Thus, we design incentive schemes that explicitly enable equilibrium behavior for which the inference required of agents is tractable.* This is analogous to the important role of strategyproofness in simplifying the strategic problem facing agents in mechanism design. Simulation results demonstrate that equilibrium routing strategies based on local routing rules lead to fast information aggregation.[1]

Related work

Leveraging individuals' abilities to both solve and route is a key component of the strategy of the team that placed first in the DARPA Red Balloon Challenge [4]. The task was to find ten balloons in the US, and the winning team introduced an incentive mechanism that promotes individuals to look for balloons and let their friends know about the task. The proposed mechanism aims to create information cascades that serve as a broadcast mechanism, but makes no consideration for identifying expertise or for the cost of communication. Our work focuses on the complementary perspective of quickly routing tasks to appropriate experts, while reducing communication costs by avoiding unnecessary broadcasting.

One can view routing scoring rules as an extension of market scoring rules [2] used in prediction markets, which provide proper incentives for individuals to improve probability estimates by contributing additional information. The major difference between task routing and a prediction market is in the 'burden' of identifying expertise: while prediction markets place the responsibility on individuals to find prediction tasks for which they have useful information, task routing incentivizes individuals to notify others with appropriate expertise who may otherwise be unaware of the task.

The problem of task routing is related to questions of decentralized search in networks [3], where the goal is to quickly deliver a message from a source to a target in a network. Kleinberg shows that for a particular grid-based model of a small-world network, individuals following local rules can quickly deliver the message with high probability. One can view the target in the model as representing a 'knowledge center,' where agents' knowledge about their neighbors' distance to the knowledge center enables effective routing. We remain interested in how such knowledge about the presence of knowledge centers and agents' distance to them can be incorporated into our model.

MODEL

We consider a task routing problem for which the goal is to efficiently obtain probability assessments for a question of interest from individuals within a social structure by appropriately rewarding them for (sequentially) refining the current prediction and routing tasks to other individuals who can provide further refinements. To formalize the setting, consider a single prediction task T, for which we would like to gather accurate probability assessments of the true state $\omega \in \Omega$. The probability assessment task can be for any state of the world that will be revealed later in time, e.g., "Will it snow next Tuesday in Boston?" or "Will the Celtics win the NBA championship this year?" We consider discrete state spaces, and assume without loss of generality a binary state space, such that $\Omega = \{Y, N\}$.

Consider a game with n players, where each player is represented by a node on the *routing graph* $G = (V, E)$. Edges in the graph may be directed or undirected, and indicate whether a particular player can route the task to another player. The task is initially assigned to a source player (named player 1, with later players numbered sequentially) who is either determined exogenously by the system or endogenously as the individual who posed the task. The source player is asked to update the probability of state Y from the prior probability p^0 to some probability p^1, and in addition, to route the task to a neighbor on the routing graph. The selected neighbor is then asked to update the assessment p^1 and route to one of its neighbors, and so on, until the game ends after a pre-specified number of rounds R (player R does not route). We assume players receiving the task are provided with a list of participants so far, as well as information about the number of rounds that remain, and can thus use this knowledge in their routing decisions. *Our goal is to design incentive mechanisms that will induce each player to truthfully update probability assessments, and to route the task to other players that can best refine the prediction.*

We model players' knowledge about the task as follows: The true state of the world is drawn with respect to p^0, which is common knowledge to all players. While no player observes the true state directly, each player may receive additional information about the true state by privately observing the outcome of coin flips drawn from the true state. To capture differences in the level of knowledge between players, we allow players to observe different numbers of coin flips, where players observing more coin flips are a priori more knowledgeable. Formally, we represent player i's signal c_i as a random bit vector of length l_i, where bit c_{ik} is a random variable over the outcome of the k-th coin flip observed by player i. We assume bits of signal are *conditionally independent* (i.e. independent conditioned on the true state) and drawn from the same distribution (known to all players), such that $\Pr(c_{ij} = x|\omega) = \Pr(c_{km} = x|\omega)$ for all players i, k, bits j, m, and realizations x. Each bit of signal is assumed to be *informative*, that is, $\Pr(c_{ij} = x|\omega = Y) \neq \Pr(c_{ij} = x|\omega = N)$ for all i, j. We also assume that bits of signal are *distinct*, that is, $\Pr(\omega = o|c_{ij} = x) \neq \Pr(\omega = o|c_{ij} = y)$ for all i, j, o and $x \neq y$. We assume the realization of each player's signal is private, but that players may know about

[1]Some proofs are omitted or condensed in this version of the paper.

how many coin flips other players observe.

With conditionally independent signals, each player can properly update the posterior probability without having to know the signals of previous players or their length, as long as previous updates were done truthfully [1]. This is useful practically in that players do not have to keep track of nor communicate their signals, and can simply update the posterior probability that sufficiently summarizes all information collected thus far.

ROUTING SCORING RULES

With rational, self-interested players, with no intrinsic value for solving or routing a particular task, enabling effective task routing requires mechanisms that will incentivize players to both truthfully report posterior probabilities and to route tasks to individuals who can best refine the predictions of the tasks. In the forecasting literature, *proper scoring rules* [5] are mechanisms that incentivize a forecaster to truthfully reveal his subjective probability of an event. A well-known proper scoring rule is the *logarithmic scoring rule*, under which a player reporting probability q for state Y is rewarded $\ln(q)$ when the true state is Y and $\ln(1-q)$ when the true state is N. The *logarithmic market scoring rule* [2] extends the logarithmic scoring rule to settings where we wish to aggregate multiple people's information. Given a sequence of players, player i reporting p^i is rewarded $s_i - s_{i-1}$, where s_i denotes the score of player i as computed by the logarithmic scoring rule. Note that a player's score is positive if and only if he improves the prediction by assigning a higher probability to the true state than the player before him.

Building on the logarithmic market scoring rule, we introduce *routing scoring rules* that incentivizes accurate predictions and effective routing decisions. We first consider the *myopic routing scoring rule* (MRSR), which rewards player i in the routing game with the payment

$$(1-\alpha)s_i - s_{i-1} + \alpha s_{i+1},$$

where $\alpha \in (0,1)$ is a constant set by the system. Note that player i's payment is based on the incremental value he provides for refining the prediction, as measured by his report and the report of the player he routes the task to. For player 1, s_0 denotes the score computed with respect to the prior p^0. The final player R does not route, and is paid $s_R - s_{R-1}$.

LEMMA 1. *The total payment from the system in the routing game with MRSR is $s_R - s_0 + \alpha(s_R - s_1)$.*

The lemma follows from taking telescoping sums, and states that the center needs to only pay for the difference between the final assessment and the initial assessment, since each player is only paid for the additional information they provide and their routing decision. Intuitively speaking, this scoring rule aims to reward a player for submitting accurate probability assessments and *myopically* routing to the player who can most accurately refine the probability assessment.

We can extend the MRSR to reward players' routing decisions based on the accuracy of information after $k > 1$ more players have provided their information. The *k-step routing scoring rule* (kRSR) pays player i

$$(1-\alpha)s_i - s_{i-1} + \alpha s_{i+k}.$$

This routing scoring rule rewards a player based on his report, as well as the eventual consequence of his routing decision k steps into the future. Unlike MRSR, it rewards players for routing to players who may not have information themselves but are able to route to others who do. When player i is paid based on player R's score (that is, $i + k = R$), we call kRSR the *path-rewarding routing scoring rule* (PRSR). Note that the last player still does not route, and is paid $s_R - s_{R-1}$.

The choice of routing rule affects players' routing decisions in equilibrium, which in turn affects how much information is aggregated. To see the connection between a player's score and the amount of information aggregated thus far, note that the expected score is strictly increasing in the total number of coin flips collected:

LEMMA 2. *Let S' and S'' denote two possible sequences of players through the first k rounds of the routing game. Assume all players truthfully update posterior probabilities, and let $E[s_k|S]$ denote the expected score of the k-th player in S taken with respect to the signal distribution of players in S. $E[s_k|S'] > E[s_k|S'']$ holds if and only if $\sum_{i\in u(S')} l_i > \sum_{j\in u(S'')} l_j$, where $u(S)$ is the (unique) set of players in S.*

The proof relies on KL-divergence. Intuitively speaking, additional bits of information can only improve the accuracy of the prediction in expectation. Since the logarithmic scoring rule rewards accuracy, collecting more coin flips will also lead to having higher scores in expectation.

CASE OF COMMON KNOWLEDGE

Having introduced routing scoring rules of interest, we consider how to use them in the routing game. We first consider the case where players are equally well-informed of other players' expertise. This captures situations where players know (of) everyone and their expertise, as is typical in a small firm or a research group. We model this by assuming that the number of coin flips l_i observed by any player i is common knowledge to all players. While players know each other's relative levels of expertise perfectly, the actual signal realizations are still assumed private.[2]

Clique topology

We first consider the routing game on a *clique*, where each player can route the task to any other player in the social graph. Given the common knowledge assumption, the clique topology implies that myopically routing the task to the most knowledgeable individual who has yet to be routed the task will allow us to aggregate as much information as possible, as quickly as possible. We show that the myopic routing scoring rule induces this behavior in equilibrium:

THEOREM 1. *Assume players' signal lengths are common knowledge, and that players are risk neutral. Let $S_{>i}$ denote the set of players who have yet to receive the task after i rounds. Under the myopic routing scoring rule, it is a Perfect Bayesian Equilibrium (PBE) for each player i to truthfully update the posterior probability, and to route the task to player $i+1 \in \text{argmax}_{m\in S_{>i}} l_m$, with the belief that all other players update the posterior probability truthfully.*

[2] By taking appropriate expectations, our results easily extend to settings where players are equally well-informed but are uncertain about the number of coin flips other players observe.

PROOF. (sketch) We show that no player wishes to deviate from the equilibrium strategy, given the belief that all other players report truthfully. Consider player i. We let $r_{<i}$ denote the reported information of players before i, such that $P(\omega|r_{<i})$ represents the posterior belief immediately before i submits its information. To prove the theorem, we first show that player i should honestly report its realized signal r_i and update the posterior beliefs to $P(w|r_i, r_{<i})$ by establishing that (a) truthful reporting maximizes s_i, and that (b) for any player m who may be routed the task, truthful reporting by player i maximizes the score s_m. For (a), note that, since s_i is based on the logarithmic scoring rule, truthful reporting maximizes the expectation of s_i. For (b), we show that the expected score of s_m (from the perspective of player i) is strictly greater when player i honestly reports its signal r_i than any alternative signal $\hat{r}_i$. The argument relies on KL-divergence, and is omitted in the interest of space.

It is left to show that player i maximizes s_{i+1} by routing to the player in $S_{>i}$ with the most coin flips. This follows directly from Lemma 2. □

General networks

We now turn to consider the routing game on general networks. Edges in the routing graph define paths through which the task can be routed, e.g., only managers can route tasks between teams, only professors can route questions to other professors, and only friends can route to friends.

The goal remains to aggregate as much information as possible with R rounds, which, in expectation, is equivalent to collecting as many coin flips as possible. We can state the computational problem of finding the optimal route in terms of collecting coin flips:

PROBLEM 1. *Consider the routing graph $G = (V, E)$, where nodes are assigned non-negative weights w_i (coin flips) that are collected upon visiting node i. Given a starting node s, find a path of length at most k such that the sum of weights of the set of nodes on the path is maximized.*

Note that players can route to other players who have received the task before (e.g., the path need not be *simple*), but no additional information is collected in subsequent visits. Immediately, we see that myopic routing will not always find the optimal solution to this problem, as routing to the neighbor with the most coin flips does not consider the chain of coin flips that may come through the path thereafter. In fact, the problem is NP-hard for variable path length k:

LEMMA 3. *Problem 1 is NP-hard.*

PROOF. Consider a reduction from the Hamiltonian Path problem. Let all nodes have weight 1, and set $k = |V|$. The solution path has total weight $|V|$ if and only if all nodes are visited within k steps, that is, a Hamiltonian Path exists. □

While the problem is NP-hard for a variable path length k, for small constant k the optimal path may be tractable to compute via exhaustive search. But even if players can compute the optimal path, we still need to find incentives that induce players to honestly report their information and to route along the optimal path. The path-rewarding routing scoring rule does just that.

THEOREM 2. *Assume players' signal lengths are common knowledge, and that players are risk neutral. Let $S_{>i}$ denote the set of players who have yet to receive the task after i rounds. Let Q_i denote the solution to problem 1 for which $k = R - i$, $s = i$, and $w_m = l_m$ if $m \in S_{>i}$ and 0 otherwise. Under the path-rewarding routing scoring rule, it is a PBE for each player i to truthfully update the posterior probability, and to route the task to the next player in the path provided by Q_i, with the belief that all other players follow this strategy.*

The common knowledge assumption allows players to compute the optimal path moving forward from their point on, so that the optimal path can be chained together through the individual routing decisions of each player. Since PRSR rewards each agent's routing decision based on the final score, it is in each agent's interest to maximize the number of coin flips collected along the entire routing path.

CASE OF LOCAL KNOWLEDGE

While people may know one another's expertise in small organizations, the common knowledge assumption becomes unreasonable for larger organizations and social networks. Any given individual will not necessarily know everyone else, and may only have summary information about the expertise and connectivity of individuals outside of a locally defined neighborhood. Some individuals may be better informed than others, e.g., by being connected to more people, having a better sense of the network topology, etc.

Although it is unreasonable to assume common knowledge in this setting, it may be reasonable to assume that within a certain locality (e.g., a group of friends, a particular department, or a specific research area), individuals are equally well-informed of others' expertise within this locality. For example, all friends of a particular person are aware of his expertise, and friends of his friends may also be aware. Similarly, it may be reasonable to assume that people know a local portion of the routing graph, e.g. individuals know not only their friends but also their friends' friends.

Formally, we say that a routing graph satisfies the *local knowledge assumption within m-hops* if, for all nodes (players) i, (a) l_i is common knowledge to all players connected to i via some path of length at most m, and (b) i knows all paths of length at most m connecting i to players j. For example, 1-hop local knowledge assumes all friends of a particular person know the person's level of expertise, and 2-hop local knowledge extends this shared knowledge to his friends of friends. We make no further assumptions on players' knowledge of other players beyond m hops, e.g., some may know of people that are more hops away, or have a better sense of the distribution of expertise in the routing graph.

Let us consider the problem of computing the optimal routing path under the local knowledge assumption. While players can compute routing decisions based on local information, routing optimally would require players to consider the expertise and routing abilities of players down the chain that are outside of their local neighborhood. For example, a player may need to use the history of routing decisions to infer why certain people were not routed the task (but could have been), based on which to update its beliefs of how information is distributed over the network. This is necessary because optimal routing requires players to consider in their routing decisions the uncertainty over the value generated by the routing decisions of subsequent players whom are be-

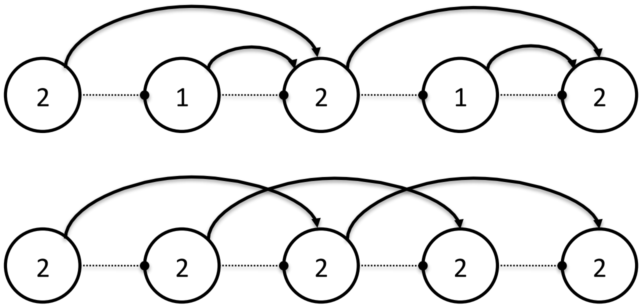

Figure 1: Illustration of the 2-1-2-1 and 2-step routing rules. Arrows depict dependencies in routing payments.

yond their locality. Not only is such reasoning complex and impractical, any equilibrium to induce optimal routing is likely to be fragile as it requires players to form intricate beliefs about other players' beliefs.

An attempt to avoid such issues may suggest incentivizing players based on a m-step routing rule whenever the local knowledge assumption holds for m-hops. The problem with this suggestion is that a player still has to consider routing decisions of players outside its locality because maximizing its payoff requires considering the routing decisions of the chain of players within its locality. For example, consider the two-step routing rule as illustrated in Figure 1 (bottom). Note that for any player, the score two steps forward will depend in part on the routing decision of the next player. But since the next player is also paid by the two-step routing rule, it's routing decision will depend not only on the knowledge of the next player, but also their routing decision. Since each player has to consider the routing decision of the next player, each player has to reason about the potential routing decisions of all players down the chain—just to compute the expected score after two steps.

To avoid such impracticality, we present a class of *local routing rules* under which players' strategies in equilibrium rely only on computations based on local information, but nevertheless take advantage of the available local knowledge. Let us first consider the case of two-hop local knowledge, for which we have the following local routing rule:

DEFINITION 1. *The* 2-1-2-1 routing rule *pays odd players based on the two-step routing scoring rule, and even players based on the myopic (one-step) routing scoring rule.*

The 2-1-2-1 routing rule incentivizes players to compute locally optimal paths of length two (see top of Figure 1), where by paying each pair of players in the routing sequence by the same amount for their routing decisions allows the odd player to be forward-looking and the even player to finish the path (of length two). Since even players are paid based on the MRSR, they will route to the available player with the most number of coin flips. Under two-hop local knowledge, an odd player knows the number of coin flips that can be collected from the next even player and also from the odd player that is then routed the task, and can thus compute the best local path without regard to routing decisions beyond its locality. Note that players still need to take into account which other players have already participated, but no other inference based on history is necessary.

Expanding on the idea, we construct a class of such local routing rules (MRSR, 2-1-2-1, 3-2-1-3-2-1, ...) that incentivizes players to compute locally optimal paths for m-hop local knowledge:

DEFINITION 2. *The m-hop local routing rule pays player i based on the $[m-(i-1) \bmod m]$-step routing scoring rule.*

We can characterize the equilibrium behavior as follows:

THEOREM 3. *Assume that players are risk neutral and m-hop local knowledge holds. Let $S_{>i}$ denote the set of players who have yet to receive the task after i rounds. Let Q_i denote the solution to problem 1 for which $k = m-(i-1) \bmod m$, $s = i$, and $w_j = l_j$ if $j \in S_{>i}$ and 0 otherwise. Under the m-hop local routing rule, it is a PBE for each player i to truthfully update the posterior probability, and to route the task to the next player in the path provided by Q_i, with the belief that all other players follow this strategy.*

PROOF. (sketch) Using similar arguments as the proof sketch for Theorem 1, we can show that players should truthfully update the posterior based on their signals. To show player i should route based on Q_i, we first note that Q_i is directly computable given m-hop local knowledge. Since Q_i maximizes the number of coin flips collected in the next k steps, applying Lemma 2 proves the point, and the theorem. □

SIMULATION

The equilibrium strategies induced by local routing rules can be viewed as approximations to computing the optimal route. As we do not have closed-form analyses about the performance of routing decisions based on the local routing rules, we can explore their performance empirically. In this section, we demonstrate via simulations that routing decisions based on local rules can quickly aggregate information as a task is routed through the network.

We consider connected random graphs with 100 nodes and average degree $d \in \{4, 10\}$, generated using the Watts-Strogatz model [6]. By varying the re-wiring probability β, the model allows us to generate graphs that interpolate between a regular lattice ($\beta = 0$) and a $G(n,p)$ random graph ($\beta = 1$), with small-world networks emerging at intermediate values of β. We associate each node with a number of coin flips, which is drawn independently either discretely from U[1,10], or from a skewed distribution where the value is 1 with probability 0.9 and 46 with probability 0.1. Note that the distributions have equal mean (5.5), but that the skewed distribution more closely resembles a setting in which there are few experts. For graphs generated in this manner, we simulate player strategies under local routing rules (MRSR, and m-hop with $m = 2$, $m = 3$) by computing local paths in the manner noted in Theorem 3, where revisited nodes are treated as having no value. As a baseline, we consider a random routing rule that routes to a random neighbor, and whenever possible, to a random neighbor who has yet to be assigned the task. Note that the expected performance of the baseline is bounded by 5.5 coin flips per round, as we would expect from randomly picking unvisited nodes in the graph.

Table 1 shows the average number of coin flips collected after 10 steps by players following local routing rules on graphs with varying β, average degree, and coin flip distribution over 100 trials (standard errors are small and hence

β	Dist.	$d = 4$			$d = 10$		
		MRSR	m=2	m=3	MRSR	m=2	m=3
.03	U	69	71	72	83	84	85
0.1	U	71	72	75	85	86	87
1.0	U	76	78	80	89	89	90
.03	S	80	87	104	150	183	227
0.1	S	88	109	146	181	226	259
1.0	S	120	155	183	227	258	278

Table 1: Comparison of routing performance after 10 steps on connected Watts-Strogatz graphs based on uniform (U) and skewed (S) coin flip distributions with fixed mean (5.5).

not reported). We see that routing rules are particularly effective in cases where there are few experts (S), and when the graph has a sufficiently high connectivity (higher d and β) that there exist paths through which experts can be routed the task. But even in cases with uniformly distributed coin flips (U) and low average degree ($d = 4$), we see that local routing rules collect significantly more coin flips than the upper bound of 55 we would expect from randomly choosing nodes, and despite connectivity constraints include many high valued nodes in its path (recall the max per node is 10).

The difference in routing performance among local routing rules is rather small for uniformly distributed values, but is more significant when the distribution is skewed. In this case, effective routing may require finding short paths to experts who are not neighbors. That said, this difference shrinks for graphs with higher degree, as high value nodes become more easily reachable (recall that as graphs approach cliques, myopic is optimal). Figure 2 shows the average number of coin flips collected by local routing rules as we progress through the routing game on graphs with $\beta = 0.1$, $d = 10$, and skewed coin flip distributions. We see that for $m \geq 2$ the performance under the routing rules are essentially the same, suggesting that we can sometimes achieve near-optimal performance globally with just two-hop local knowledge. Note that for all local routing rules the rate of information aggregation eventually slows down, which denotes the point at which virtually all experts have been routed the task.

CONCLUSION

Our analysis of task routing for prediction tasks introduces routing scoring rules, which in equilibrium bring agents to truthfully contribute information, and to route tasks based on simple computations that nevertheless lead to effective information aggregation. Future work on task routing for prediction tasks should consider other information structures and specialized network topologies, and may also incorporate intrinsic values for solving or routing, as well as costs for acquiring additional bits of signal. More generally, future work on task routing should continue to consider specific task level issues, but also organizational issues relating to distributing streams of tasks in a way that takes into account people's solving and routing abilities over the space of tasks, as well as their general availability and the need for 'load-balancing.' The numerous potential applications in the task routing space point to a continued need to strike a balance between adopting a principled approach and appealing to practicality.

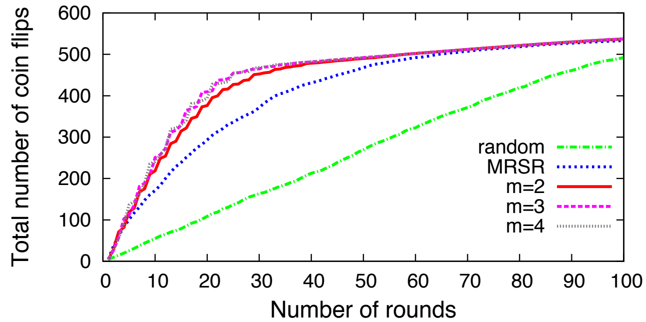

Figure 2: Comparison of routing performance among local routing rules for graphs with $\beta = 0.1$, $d = 10$, and skewed coin flip distributions. Values are averaged over 100 trials.

Acknowledgments

The first author acknowledges support from the Department of Defense (DoD) through the National Defense Science & Engineering Graduate Fellowship (NDSEG) Program. In addition, research was sponsored by the Army Research Laboratory and was accomplished under Cooperative Agreement Number W911NF-09-2-0053. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Army Research Laboratory or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation here on.

REFERENCES

[1] Y. Chen, D. M. Reeves, D. M. Pennock, R. D. Hanson, L. Fortnow, and R. Gonen. Bluffing and strategic reticence in prediction markets. In *WINE*, 2007.

[2] R. Hanson. Logarithmic market scoring rules for modular combinatorial information aggregation. *Journal of Prediction Markets*, 1(1):3–15, February 2007.

[3] J. Kleinberg. Complex networks and decentralized search algorithms. In *Proc. International Congress of Mathematicians*, 2006.

[4] G. Pickard, I. Rahwan, W. Pan, M. Cebrian, R. Crane, A. Madan, and A. Pentland. Time Critical Social Mobilization: The DARPA Network Challenge Winning Strategy. *ArXiv e-prints*, Aug. 2010.

[5] R. Selten. Axiomatic characterization of the quadratic scoring rule. *Experimental Economics*, 1(1):43–61, June 1998.

[6] D. J. Watts and S. H. Strogatz. Collective dynamics of "small-world" networks. *Nature*, 393(6684):440–442, June 1998.