4-2013

Modeling Social Information Learning among Taxi Drivers

Siyuan LIU
Carnegie Mellon University

Ramayya KRISHNAN
Carnegie Mellon University

Emma BRUNSKILL
Carnegie Mellon University

Lionel NI
Hong Kong University of Science and Technology

Citation

LIU, Siyuan; KRISHNAN, Ramayya; BRUNSKILL, Emma; and NI, Lionel. Modeling Social Information Learning among Taxi Drivers. (2013). *Advances in Knowledge Discovery and Data Mining: 17th Pacific-Asia Conference, PAKDD 2013, Gold Coast, Australia, April 14-17, 2013, Proceedings, Part II*. 7819, 73-84.
Available at: https://ink.library.smu.edu.sg/sis_research/3471

Modeling Social Information Learning among Taxi Drivers

Siyuan Liu[1], Ramayya Krishnan[1], Emma Brunskill[1], Lionel M. Ni[2]

[1] Carnegie Mellon University
[2] Hong Kong University of Science and Technology

Abstract. When a taxi driver of an unoccupied taxi is seeking passengers on a road unknown to him or her in a large city, what should the driver do? Alternatives include cruising around the road or waiting for a time period at the roadside in the hopes of finding a passenger or just leaving for another road enroute to a destination he knows (e.g., hotel taxi rank)? This is an interesting problem that arises everyday in many cities worldwide. There could be different answers to the question poised above, but one fundamental problem is how the driver learns about the likelihood of finding passengers on a road that is new to him (as in he has not picked up or dropped off passengers there before). Our observation from large scale taxi drivers behavior data is that a driver not only learns from his own experience but through interactions with other drivers. In this paper, we first formally define this problem as **Socialized Information Learning (SIL)**, second we propose a framework including a series of models to study how a taxi driver gathers and learns information in an uncertain environment through the use of his social network. Finally, the large scale real life data and empirical experiments confirm that our models are much more effective, efficient and scalable that prior work on this problem.

1 INTRODUCTION

The study of how a person gathers information and makes decisions has a long and varied literature. In the previous research, collective intelligence [13, 19, 23, 27], intelligent agent [7, 10, 12, 21, 22], transfer learning [6, 25, 26] and evidence-based reasoning [4, 8, 14] and other methods are proposed to investigate an agent's learning theory. But due to the new challenges raised in dynamic uncertain environment [11, 17, 18], prior work on this topic is either inefficient or inaccurate. Now consider the following problem.

In our application context, there are 3,187 taxi drivers, and among them there are 25.2% new drivers (less than one year driving experience), 47.1% normal drivers (one to two years driving experience) and 27.7% experienced drivers (greater than two years driving experience). When a driver comes to road to pick up passengers, there are two actions (action set) to choose: *waiting* at the current location or *cruising* to other locations. The knowledge of the taxi driver can be described as the histogram of the waiting time before picking up a passenger, the number of picking-ups, and the revenue (income) at the given time and location. Given a situation that the drivers come to an unknown road, in a survey of 1,000 taxi drivers, we study how the drivers follow their own experienced knowledge and the socialized knowledge (the knowledge obtained from other drivers) to accordingly make actions. We find that different drivers have very different learning preferences. The new drivers prefer to follow the socialized knowledge, but the experienced drivers prefer to follow their experienced knowledge. The standard deviation is larger in new drivers than the experienced drivers. Hence, different drivers would take various knowledge sources.

In our dataset, we also have the communication records to indicate the socialization of taxi drivers. When a new driver was assigned to different group compositions, we studied the income changes (one week). We take 70% as the threshold to define a group composition, e.g., *New* means more than 70% of the drivers in the group are new drivers. We also tested other thresholds larger than 50%, and the results are very similar. It is very interesting that even the new driver has less social closeness to the *Experienced* group than the *New* group, but not only individual income but also total income in the *Experienced* group have the greater increases than the *New* group. Hence the more socializations may not give the more accurate knowledge and better income. For example, the experienced drivers could give the new driver more accurate knowledge than other drivers.

In a word, the problem is that how the drivers socialize with each other to construct the accurate knowledge in an uncertain environment, which we define as **Socialized Information Learning**. To tackle the problem, we have two steps to be accomplished, 1) how to retrieve an accurate knowledge set for an individual driver in an uncertain environment; 2) how to utilize a social network to learn the information. In this paper, we first propose an **Individual Information Model** to describe taxi drivers' information collection, considering the features of the required places and the similarity between the required places and the experienced places. Second, we introduce the social network structure with a probability weighting function into the model to describe the non-linear socialized information learning, called **Socialized Information Model**.

To summarize the contributions of our work, first, we are the first to discover the **Socialized Information Learning** problem in taxi drivers, and we define it as a new uncertain information learning problem; second, we propose a framework including a series of novel models to solve the socialized information learning problem (not only model taxi drivers' behaviors by themselves, but also their social behaviors via the group information) and investigate it in the dynamic field; third, we employ large scale real life datasets to test our models, and the empirical results show that our models outperform the state-of-the-art in terms of effectiveness, efficiency and scalability.

The paper is structured as follows.In Section 2, we formally define the socialized information learning problem. We propose a series of models to solve the new socialized information learning problem in Section 3, and the empirical experiment results are illustrated and analyzed in Section 4. The related work to our study is surveyed in Section 5. Finally, we conclude our work and give directions to the future work.

2 PROBLEM DEFINITION

2.1 Definitions

Definition 1. (Agent) *Agent is defined as an entity that is capable of perceiving knowledge and accordingly do action.*

In our work, an agent is a taxi driver.

Definition 2. (Agent group) *Agent group is defined as a set of agents that are capable of perceiving knowledge from each other and accordingly do action.*

In our work, an agent group is a predefined group of taxi drivers by a taxi company. We define the knowledge of a taxi driver as experienced knowledge and socialized knowledge as below.

Definition 3. (Experienced knowledge) *The experienced knowledge (EK) is defined as a set of information collected from the agent's own experience, that is, historical records.*

In our work, the experienced knowledge is a set of the CDFs of waiting-time, picking-up and revenue distributions at given locations and times, from a given taxis historical GPS logs and business records. For the road without a given taxi driver's experienced knowledge is called an unknown road.

Definition 4. (Socialized knowledge) *The socialized knowledge (SK) is defined as a set of information collected from other agents' information in the same group, that is, other agents' historical records.*

In our work, the socialized knowledge is a set of CDFs of waiting-time, picking-up and income distributions at given locations and times, from a given taxis group members GPS logs and business records at the same given locations and times.

Definition 5. (Action) *Action of an agent is defined as a selection of a mutual exclusion set.*

Action of a taxi driver is defined as cruising or waiting for a passenger. At a given location and time, a taxi driver can select an action (make a decision) of cruising to other locations until picking up a passenger or waiting for a time period at the given location until picking up a passenger.

Definition 6. (Socialization) *Socialization of an agent is defined as a communication between two agents.*

Socialization of a taxi driver is a call between two taxi drivers in the same group. Each socialization is recorded a vector: $(i, j, t_s, t_e, \iota_s, \iota_e)$, where i is the caller taxi ID, j is the callee taxi ID, t_s is the call start time, t_e is the call end time, ι_s is the calling start location (longitude and latitude) and ι_s is the calling end location.

Definition 7. (Socialization closeness) *Socialization closeness between two agent is defined as a function of communications between the two agents.*

The socialization closeness between two taxi drivers is defined as:

Definition 8. (Socialization closeness of taxi drivers) *Given two taxi driver i and j in a group, a time interval t, i has the communication attribute set $S_i^t = \{s_i^{t,1}, ..., s_i^{t,f}, ..., s_i^{t,m}\}$, where $1 \leq f \leq m$ and m is the number of attributes. The socialization closeness within the time interval t, $\gamma_{i,j}^t$ is*

$$\gamma_{i,j}^t = \frac{1}{m}\sum_{f=1}^{m} w_f s_{i,j}^{t,f} \tag{1}$$

where w_f^t is the weight of an attribute f in the time interval t. The socialization closeness set is $\Gamma^t = \{..., \gamma_{i,j}^t, ...\}$, where $i, j \in \mathbb{N}$, $i, j = 1, 2, ..., n$, and n is the number of taxi drivers in a given group.

In our study, we take the mean of three attributes (m=2), the number of calls and the call duration as socialization closeness values. Our solution can be extended to other cases where $m > 2$ and other functions. The default time interval is set as a minute, which is set by the communication service company. In this work, we equally take the weights in 8, which is predefined by users.

For the input communication data, we can construct a social network, *G* = (*V*, *E*), where *V* is a set of nodes (taxi drivers), and *E* is a set of edges (socialization with socialization closeness as the weight on the edge).

Definition 9. (Decision knowledge) *The decision knowledge (DK) is defined as the information taken by an agent to make an action.*

For a taxi driver, the decision knowledge is based on a set of CDFs of waiting-time, picking-up and income distributions at given locations and times to make a certain action.

2.2 Socialized Information Learning

The formal definition of **Socialized Information Learning** (SIL) is as below.

Given: a set of agents Q, a set of experienced knowledge E, a set of socialized knowledge S, and a social network G with socialization closeness Γ.

Goal: a set of decision knowledge D.

Specifically, given a taxi driver with experienced knowledge and socialized knowledge, under a social network, to make a good action to pick up a passenger in an unknown road, what is decision knowledge to support the given taxi driver's action?

In our work, the decision knowledge utilized by a taxi driver to decide the next move is calculated by a decision function as below.

Definition 10. (Decision function of a Taxi Driver) $P[v^t(\iota) \,|n^t(\iota) \geq \varpi] \geq \omega$*, where ι is a location index, t is a time index, $v^t(\iota)$ is the revenue, $n^t(\iota)$ is the number of passengers, ϖ and ω are thresholds.*

The above function means if given a probability of a certain number of passenger is greater than a given threshold (ϖ), the conditional probability of revenue is greater than ω, the driver is going to wait for passengers at the given location, otherwise, cruise to other locations.

3 SOCIALIZED INFORMATION LEARNING FRAMEWORK

3.1 Individual Information Model

The basic idea of **Individual Information Model** is as follows. First, based on the utilities of the road and buildings, we label each grid and cluster the grids into different clusters. Second, given a location, we evaluate the similarity between the given location and the taxi's experienced locations. Third, we weight the similar experienced knowledge and the socialized knowledge at the given location and time, and finally get the decision knowledge. To make the following expression clear, we take the revenue as a knowledge example to illustrate the model.

Definition 11. *Given a physical location ι =(x,y) and the report set Φ, the revenue spectrum $V_{\Phi}^{(t)}(\iota)$ is the set of all the reported revenues (of the given taxi) at time t in location ι in Φ, i.e.,*

$$V_{\Phi}^{(t)}(\iota) = \{v | \exists \phi_m^{(t)} \in \Phi, (x_m^{(t)}, y_m^{(t)}, v_m^{(t)}) = (x, y, v)\}$$

The revenue spectrum at all time instances is also written as $V_{\Phi}(\iota) = \cup_{t \in [0,+\infty]} V_{\Phi}^{(t)}(\iota)$. Since the time is discrete, Φ contains a finite number of reports and thus the revenue spectrum is finite as well.

Definition 12. *The location revenue $v_{\Phi}^{(t)}(\iota)$ is defined as the average of the revenue spectrum in location ι at t,*

$$v_{\Phi}^{(t)}(\iota) = \frac{1}{|V_{\Phi}^{(t)}(\iota)|} \sum_{v_m \in V_{\Phi}^{(t)}(\iota)} v_m \tag{2}$$

In the real data, the revue spectrum is very sparse and lossy, hence we employ a moving average technique to reconstruct a sufficient spectrum as below.

Definition 13. *The experienced revenue knowledge $\mathrm{H}^{(t)}(\iota)$ of a given location $\iota = (x, y)$ is defined as an exponential moving average of the complementary cumulative distribution function (CCDF) of the instant revenue over the revenue spectrum, i.e.,*

$$\mathrm{H}^{(t)}(\iota) = \alpha_\iota \cdot \mathrm{H}^{(t-\tau_\iota)}(\iota) + (1 - \alpha_\iota) \cdot (1 - \mathrm{P}(v \leq v_{\Phi}^{(t)}(\iota))) \tag{3}$$

where $\mathrm{P}(v \leq v^{(t)}(\iota))$, $v \in V_{\Phi}(\iota)$ represents the probability that the revenue at ι is less than or equal to the location instant revenue $v_{\Phi}^{(t)}(\iota)$. α_ι and τ_ι are two parameters to capture the dynamism of ι.

The parameter α_ι is a smoothing factor of exponential moving average in H. It is used to capture the degree of dynamism of the location dynamics. In general, a smaller α_ι indicates a higher dynamism and vice versa. The parameter τ_ι is the interval between two moving averages. It reflects the periodic property of the location dynamics.

Different locations will present distinctive dynamic behaviors, resulting in various settings. In order to systematically study the speed distribution, we apply Fourier Transformation (FT). FT can transform the function from time domain to frequency domain, revealing inherent periodic property of original function as well as the amplitude of the corresponding frequency. Specifically, given the revenue distribution function over time $v^{(t)}(\iota)$ at a location ι, its FT can be calculated by,

$$\hat{f}(\xi) = \int_{-\infty}^{+\infty} v^{(t)}(l) e^{-2\pi i t \xi} dt, \tag{4}$$

where t is the variable.

To calculate the $\mathrm{H}^{(t)}(\iota)$, we maintain the location history information for six months. As the computation is carried out at a data center in a centralized manner, the computational and storage cost is acceptable.

Definition 14. *The similarity between the given location and the taxi's experienced locations is defined by the linear correlation coefficient.*

Given the similar experienced revenue knowledge $\mathrm{H}^{(t)}(\iota)$ and the socialized revenue knowledge $\mathrm{H}^{'(t)}(\iota)$ from other drivers, we have the decision revenue knowledge as below.

$$\hat{\mathrm{H}}^{(t)}(\iota) = \beta \mathrm{H}^{(t)}(\iota) + (1-\beta)\mathrm{H}^{'(t)}(\iota), \tag{5}$$

where β is a parameter capturing the weight of a taxi driver following the own experienced revenue knowledge, which is determined by users (taxi drivers). In our work, we learned this parameter from the real life data, which is elaborated in Section 4. The method to retrieve the socialized revenue knowledge $\mathrm{H}^{'(t)}(\iota)$ is proposed in the next section.

3.2 Socialized Information Model

Given a taxi driver, there exists knowledge limitation (hard to possess the knowledge of the whole city), hence the taxi driver may consult other taxi drivers via a social network G. In this subsection, we propose a **Socialized Information Model** to learn knowledge from other drivers in a social network.

Definition 15. *(***Socialization probability***) The probability $\mathrm{p}_{i,j}$ of a given taxi driver i socializing with a taxi driver j is the percentage of socialization closeness between the two drivers over the total socialization closeness among other drivers being socialized by the driver in a given time period.*

Hence the socialized knowledge for a given taxi driver can be described by a probability-based weighting function over a set of knowledge. Under the expected utility theory, a taxi driver weights probabilities of learning information linearly. However, the evidence suggests that the taxi drivers weight probabilities in a non-linear manner. An example is given as following.

$$\mathrm{H}_i^{'(t)}(\iota) = \left(\frac{1}{n} \cdot \sum_{k=1}^{n} [\mathrm{p}_{i,k}\mathrm{H}_k^{(t)}(\iota)]^m\right)^{\frac{1}{m}}, \tag{6}$$

where m is a parameter in generalized mean, determining the appropriate mean (in our work, we select as 2). $\mathrm{p}_{i,k}\mathrm{H}_k^{(t)}(\iota)$ in Eq. 6 means the socialized knowledge of driver

i from driver j, under the probability of driver i being able to access the experienced knowledge of driver k.

Unfortunately, this method does not work well under the following two cases: 1) overweight small probabilities and underweight large ones (S1); 2) do not choose stochastically dominated options when such dominance is obvious (S2). Hence we utilize a probability weighting function to conduct a non-linear weighting of socialized knowledge.

The particular probability weighting function is

$$w(p) = \frac{p^{\lambda}}{[p^{\lambda} + (1-p)^{\lambda}]^{\frac{1}{\lambda}}}, \tag{7}$$

where $0.5 < \lambda \leq 1$, $w(p)$ is a weighted probability and p is $p_{i,j}$ in Eq. 6.

After weighting $p_{i,j}$, we utilize Eq. 6 to compute the socialized knowledge.

4 EMPIRICAL EXPERIMENTS

4.1 Experiment Setup

Datasets description: We collected one year taxi operation data records, including taxi tracking records, taximeter records and communication records, in a large city in China. The scale of the whole dataset is almost 1 Terabytes. Taxi tracking records provide taxis' group information and traces, including location and time information; taximeter records provide taxis' revenue, waiting time and picking-up with location-time logs; communication records provide taxis' social information, e.g., social closeness. We employ six months data as training data and the other six months data as test data. We also collected one month traffic surveillance video data in the city. The traffic data can provide us the traffic flow, traffic lights, taxis' picking-up and dropping-off information, which we employ as the ground truth of taxis' behaviors in the city.

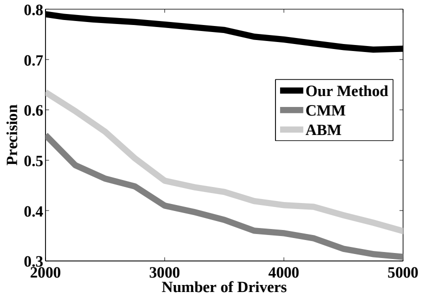
Experiment environment: A server with four Intel Core Quad CPUs (Q9550 2.83 GHz) and 32 GB main memory.

Baseline methods: We compare our method with two baseline methods: one is the classic method in collective intelligence, called CMM [13], which is popular and generate many latest approaches; the other is the representative method in agent modeling in intelligent agent and social system, called ABM [21]. The parameter settings of the above two methods in our work follow their parameters in their papers [13, 21].

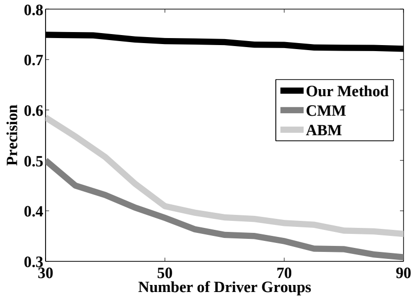
Evaluation metrics: In our experiments, we employ time cost, scalability to evaluate efficiency, and precision, recall, and F1 to evaluate effectiveness.

4.2 Parameter Learning

Given a taxi driver, we learn the parameters β in Eq. 5 from the driver's historical behaviors as follows. When the driver comes to an unknown road, if the driver makes no call to other drivers to consult the given road's information, we assume the driver follow the own experienced knowledge; else if the driver makes calls to other drivers who have



(a) Different Drivers



(b) Different Groups

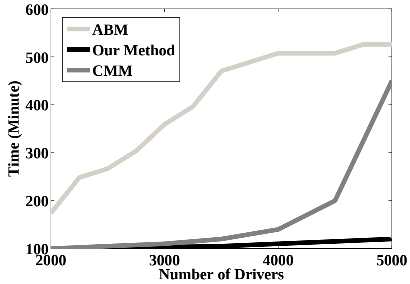
Fig. 1. Precision in Different Drivers and Groups.

the experienced knowledge of the given road, and the given driver accordingly makes an action, we assume the driver follow the socialized knowledge. Based on the records in the historical data, we can have the percentage of a given driver's follow behavior. In our study, we take the percentage of following the own experienced knowledge as β. For different drivers, β is different and updating along the new records coming to the dataset. The parameter updating is intuitive which is not elaborated in the paper. In the following experiments, we utilize the percentage number as the parameter.

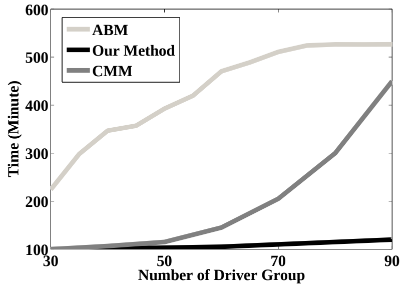
4.3 Effectiveness Evaluation

Table 1. Our Method's Effectiveness in Different Driver Categories

DC	**N**	**P**	**R**	**F**
New	1260	67.8%	57.8%	62.3%
Normal	2355	71.7%	67.3%	69.4%
Experienced	1385	80.2%	78.7%	79.4%



(a) Different Drivers



(b) Different Groups

Fig. 2. Efficiency in Different Drivers and Groups.

Table 2. Baseline Method's Effectiveness in Different Driver Categories

DC	N	P	R	F
New	1260	31.2%	17.8%	22.7%
Normal	2355	39.7%	27.3%	32.3%
Experienced	1385	45.9%	36.5%	40.8%

We evaluate effectiveness by *precision*, *recall*, and *F1*. *Precision* is the fraction of retrieved results that are relevant to the search, that is, the number of waiting/ cruising actions resulting from our model over the number of all waiting/ cruising actions made by taxis resulting from our model. *Recall* is the fraction of retrieved results that are relevant to the query that are successfully retrieved, that is, the number of waiting/ cruising actions resulting from our model over the number of all waiting cruising actions made by taxis.

To evaluate the effectiveness of our method, we design two categories of experiments.

Category 1: effectiveness in different driver categories. The results of our method are listed in Table 1. DC is the driver category, N is the number of drivers in the category, P is *Precision*, R is *recall*, and F is *F1*. The results of ABM are listed in Table 2. In our experiment, CMM returns much worse results than ABM.

Category 2: effectiveness in different drivers, groups and time series. We conducted *precision*, *recall*, and *F1* tests in different drivers, groups and time series. Due to page limitation, we only show the *precision* results. In Figure 1 (a) and (b), we test the accuracy in one month data. The results show that our method returns much more accurate results than the baseline methods, not only in different drivers scenario, but also in different groups. In a conclusion, our method's accuracy is much better than baseline methods, and the accuracy also shows great scalability.

4.4 Efficiency Evaluation

We conducted efficiency tests in different drivers, groups and time series. The efficiency is measured by the time cost in a knowledge learning process. In Figure 2 (a) and (b), we test the efficiency in one month data. The results show that our method costs much less time than the baseline methods, not only in different drivers scenario, but also in different groups. In a conclusion, our method's efficiency is much more better than baseline methods, and the efficiency also shows great scalability.

5 RELATED WORK

In [9], they proposed an approach to the problem of driving an autonomous vehicle in normal traffic. In [1], they discussed the spatial dispersion problem. But the work in this category either does not consider the social structure to retrieve the accurate information, or does not consider the dynamics in the learning process. In organized learning theory, this category work assumes that the sum of individual knowledge does not equate to organizational knowledge [2, 3, 5, 15, 16, 24]. In [4], they studied the distinction between individual knowledge and organizational knowledge, and prove the assumption. In [21], Ronald *et al.* demonstrated the design and implementation of an agent-based model of social activity generation. Szuba *et al.* [23] attempted to formally analyze the problem of individual existence of a being versus its existence in a social structure through the evaluation of collective intelligence efficiency. Heylighen *et al.* [13] argued that the obstacles created by individual cognitive limits and the difficulty of coordination could be overcome by using a collective mental map (CMM). Deng *et al.* [7] explored the use of active learning techniques to design more efficient trials. Rettinger *et al.* [20] studied the learning of trust and distrust in social interaction among autonomous, mentally-opaque agents. Wang *et al.* [25] presented an algorithm for finding the structural similarity between two domains, to enable transfer learning at a structured knowledge level. Cao *et al.* [6] proposed an adaptive transfer learning algorithm to adapt the transfer learning schemes by automatically estimating the similarity between a source and a target task. Zhu *et al.* [27] turned the co-training algorithm into a human collaboration policy. Unfortunately current work can not work well in our socialized information learning because of the challenges from dynamic updating along the time and large scale socializations.

6 CONCLUSION AND FUTURE WORK

In this paper, we model the social information learning among taxi drivers and employ large scale real life data and empirical experiments to confirm our models in terms of much better effectiveness, efficiency and scalability than the state-of-the-art. Our models could be relevant to other domains, e.g., studying animal behavior, or where people go to sell things. We leave taxi driver decision model as the future work. How to model taxi driver's decision based on their collected information is a very interesting but challenging topic. Our current work can make such future work on the accurate and updated information.

7 Acknowledgment

This research is supported by the Singapore National Research Foundation under its International Research Centre @ Singapore Funding Initiative and administered by the IDM Programme Office, and Huawei Corp. grant YBCB2009041-27, and University Transportation Center grant (DTRT12-G-UTC11) from the US Department of Transportation.

References

1. S. Alpern and D. Reyniers. Spatial dispersion as a dynamic coordination problem. *Theory and decision*, 53(1), 2002.
2. L. Argote and E. Miron-Spektor. Organizational learning: From experience to knowledge. *Organization Science*, 22, 2011.
3. C. Argyris and D. Schn. *Organizational learning: A theory of action perspective*. Addison Wesley, 1978.
4. G. Bhatt. Information dynamics, learning and knowledge creation in organizations. *Learning Organization, The*, 7(2), 2000.
5. N. Bontis and V. M. Coss. Managing an organizational learning system by aligning stocks and flows. *Journal of Management Studies*, 39, 2002.
6. B. Cao, J. Pan, Y. Zhang, D. Yeung, and Q. Yang. Adaptive transfer learning. In *Proc. of AAAI*, 2010.
7. K. Deng, J. Pineau, and S. Murphy. Active learning for developing personalized treatment. In *Proc. of UAI*, 2011.
8. K. Devlin. A framework for modeling evidence-based, context-influenced reasoning. In *Proc. of CONTEXT*, 2003.
9. J. Forbes, T. Huang, K. Kanazawa, and S. Russell. The batmobile: Towards a bayesian automated taxi. In *Proc. of IJCAI*, 1995.
10. W. Fu, L. Song, and E. P. Xing. Dynamic mixed membership blockmodel for evolving networks. In *Proc. of ICML*, 2009.
11. Y. Ge, H. Xiong, C. Liu, and Z.-H. Zhou. A taxi driving fraud detection system. In *Proc. of ICDM*, 2011.
12. R. Glaubius, T. Tidwell, C. Gill, and W. Smart. Real-time scheduling via reinforcement learning. In *Proc. of UAI*, 2010.
13. F. Heylighen. Collective intelligence and its implementation on the web: Algorithms to develop a collective mental map. *Comput. Math. Organ. Theory*, 5(3):253–280, Oct. 1999.

14. J. Kay, W. T. Niu, and D. J. Carmichael. Oncor: ontology- and evidence-based context reasoner. In *Proc. of IUI*, 2007.
15. S. J. Lee and Z. Popović. Learning behavior styles with inverse reinforcement learning. *ACM Trans. Graph.*, 29, July 2010.
16. P. Li, J. X. Yu, H. Liu, J. He, and X. Du. Ranking individuals and groups by influence propagation. In *Proc. of the* 15^{th} *PAKDD*, 2011.
17. S. Liu, C. Liu, Q. Luo, L. Ni, and R. Krishnan. Calibrating large scale vehicle trajectory data. In *Proc. of IEEE MDM*, 2012.
18. S. Liu, Y. Liu, L. M. Ni, J. Fan, and M. Li. Towards mobility-based clustering. In *Proc. of ACM SIGKDD*, 2010.
19. M. Nickel, V. Tresp, and H.-P. Kriegel. A three-way model for collective learning on multi-relational data. In *Proc. of ICML*, 2011.
20. A. Rettinger, M. Nickles, and V. Tresp. Statistical relational learning of trust. *Mach. Learn.*, 82(2), 2011.
21. N. Ronald, V. Dignum, C. Jonker, T. Arentze, and H. Timmermans. On the engineering of agent-based simulations of social activities with social networks. *Inf. Softw. Technol.*, 54(6), 2012.
22. S. Russell. Learning agents for uncertain environments (extended abstract). In *Proc. of COLT*, 1998.
23. T. Szuba, P. Polanski, P. Schab, and P. Wielicki. Transactions on computational collective intelligence. chapter On efficiency of collective intelligence phenomena, pages 50–73. 2011.
24. G. F. Templeton, B. R. Lewis, and C. A. Snyder. Development of a measure for the organizational learning construct. *J. Manage. Inf. Syst.*, 19(2), 2002.
25. H. Wang and Q. Yang. Transfer learning by structural analogy. In *Proc. of AAAI*, 2011.
26. L. Xiao, D. Zhou, and M. Wu. Hierarchical classification via orthogonal transfer. In *Proc. of ICML*, 2011.
27. X. Zhu, B. Gibson, and T. Rogers. Co-training as a human collaboration policy. In *Proc. of AAAI*, 2011.