Sequential Decision Making for
Improving Efficiency in Urban Environments

Pradeep Varakantham

School of Information Systems, Singapore Management University
pradeepv@smu.edu.sg

Abstract

Rapid "urbanization" (more than 50% of world's population now resides in cities) coupled with the natural lack of coordination in usage of common resources (ex: bikes, ambulances, taxis, traffic personnel, attractions) has a detrimental effect on a wide variety of response (ex: waiting times, response time for emergency needs) and coverage metrics (ex: predictability of traffic/security patrols) in cities of today. Motivated by the need to improve response and coverage metrics in urban environments, my research group is focussed on building intelligent agent systems that make sequential decisions to continuously match available supply of resources to an uncertain demand for resources. Our broad approach to generating these sequential decision strategies is through a combination of data analytics (to obtain a model) and multistage optimization (planning/scheduling) under uncertainty (to solve the model). While we perform data analytics, our contributions are focussed on multi-stage optimization under uncertainty. We exploit key properties of urban environments, namely homogeneity and anonymity, limited influence of individual entities, abstraction and near decomposability to solve "multi-stage optimization under uncertainty" effectively and efficiently.

1 Problems of Interest and Significance

Many decision problems in urban environments can be characterised as requiring a match between limited resource supply and an unpredictable demand for resources. Given below are a few practical real world urban decision problems of interest to us:

- Taxi fleets: Resource supply corresponds to the available taxis and demand corresponds to customers needing taxis. The goal in this problem is to increase revenues for taxis (or reduce wait times for customers) by continuously matching available taxis to customer demand or proxies for customer demand (ex: taxi stands).
- Bike sharing systems: Resource supply corresponds to available bikes at base stations and demand corresponds to customers needing bikes. The goal in this problem is to reduce lost demand due to unavailability of bikes at base stations. We are focussed on lost demand, as it can lead to customers employing private vehicles, which in turn will lead to increased carbon emissions and traffic congestion. A similar problem is relevant to car sharing systems as well.
- Emergency response: Resource supply corresponds to ambulances or fire trucks at base stations and demand corresponds to emergency events. The goal in this problem is to reduce response time for emergency events by dynamically moving the "right" ambulances to the "right" base stations.
- Traffic patrol and Security: Resource supply corresponds to traffic or security personnel at base locations and demand corresponds to potential for traffic violations or security incidents. The goal in this problem is to prevent traffic violations and security incidents by reducing predictability in patrols of traffic/security personnel without sacrificing on coverage of "important" locations.
- Theme parks: Resource supply corresponds to attractions and demand corresponds to patrons visiting the attractions. The goal in this problem is to reduce wait times by providing decision support to patrons on visiting the "right" attractions at the "right" times.

We now situate these urban decision problems in the context of existing work in Artificial Intelligence and Operations Research on general resource allocation problems. While there are other factors (offline/online, objectives etc.), we categorise using the following three criterion to precisely highlight differences between existing work and our work : (a) Scale of problems; (b) Cooperative/Competitive nature of decision makers (ones doing matching) or supply or demand; and (c) Deterministic or Non-deterministic (Stochastic and Dynamic) nature of the environment. Figure 1 provides this categorisation identifying specific research threads in a category using names of models/representations/frameworks. We first describe the four categories associated with existing research:

1. **Deterministic and Cooperative Problems**: In this category, (Distributed) Constraint Satisfaction [Yokoo and Hirayama, 2000] and (Distributed) Constraint Optimization [Pragnesh Jay Modi and Yokoo, 2005] models have been employed to represent problems where values (can represent demand) have to be assigned to variables (resources) so as to
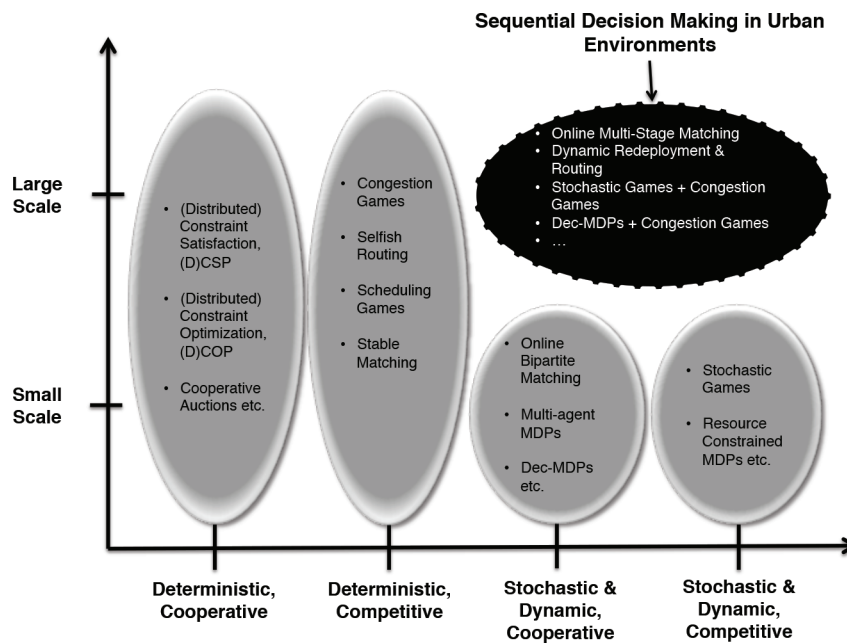
Figure 1: Characterisation of existing research and our work (dotted oval with black fill). Key difference of our work is the focus on large scale problems where there is stochasticity and dynamism.

either maximize satisfaction of constraints or minimize cost of constraint violations. Another representation in this category is the Cooperative auctions [Lagoudakis *et al.*, 2004] framework, where a centralised authority decides the outcome of an auction for either demand/resources among cooperative entities.

2. **Deterministic and Competitive Problems**: Game Theory provides natural models for strategic decision making in the presence of competing entities. In this category, we specifically consider game theoretic frameworks of relevance to resource allocation, namely congestion games [Rosenthal, 1973], selfish routing [Roughgarden and Tardos, 2002] and scheduling games. Existing methods for solving these frameworks provide equilibrium strategies on allocating resources to individual players. Stable matching [Gale and Shapley, 1962] is another framework that has received significant interest in representing problems where there are two sides (students and universities, employers and employees etc.) and entities on one side have to be matched to the entities on the other side. The objective is to provide a stable matching, so that they players on neither side have an incentive to deviate from their match.

3. **Stochastic/Dynamic, Cooperative and Small Scale Problems**: Markov Decision Problems (MDPs) [Puterman, 1994] and its extensions to multiple agents, decentralised decision making, i.e., MMDPs [Guestrin *et al.*, 2001], Dec-MDPs [Becker *et al.*, 2004] respectively are leading frameworks for this category of problems. The goal in this frameworks is to compute policies that provide sequential decisions to optimise expected value (over the uncertainty).

4. **Stochastic/Dynamic, Competitive and Small Scale Problems**: Stochastic Games [Shapley, 1953] is the most relevant framework for representing this category of problems

in view of representing stochasticity/dynamism and competitive nature of players effectively. Resource constrained MDPs [Dolgov and Durfee, 2006] is a specialized framework of stochastic games, where reward interactions between agents occur due to resources.

Urban decision problems of interest are in the fifth category and are specified in the black oval of Figure 1, thus providing a clear differentiation with existing work. Specifically, we are interested in large scale (even societal scale) problems where there is both stochasticity and dynamism, irrespective of the nature of the entities (cooperative or competitive) involved. To provide an intuitive estimate of the scale, we considered urban decision problems where there are 10000 taxis serving thousands of customers, 300 base stations carrying close to 6000 bicycles serving thousands of customers, a theme park with around 20 attractions serving tens of thousands of customers on any particular day, emergency response systems that serve two large urban cities in Asia and so on. The practical relevance coupled with the significant computational complexity involved in solving them make the urban decision problems not only empirically significant but also technically challenging and relevant.

2 Solution Methods

Our solution methods for these urban decision problems are based on exploiting basic properties of the urban environments, namely:

(a) **Homogeneity and Anonymity**: Typically in urban environments, there is homogeneity in supply (ex: 90% of taxis in Singapore are identical and have same fare structure) and demand components (ex: customers going from a source

to a destination are identical from the perspective of taxis) and more over there is anonymity in interactions between supply and demand (ex: assigning any one of the two taxis at a taxi stand to a near by customer typically have identical match value).

(b) **Limited Influence of Individual Entities (Supply/Demand)**: While there are typically a large number of entities involved in urban environments, the impact of each of them on the overall outcome is typically very small.

(c) **Abstraction**: Urban decision problems where there is supply demand matching are amenable to abstraction. That is to say, we initially abstract a group of supply components (depending on specific domain properties) into an abstract supply component and create an abstract problem. The sequential decision strategy to match supply and demand can initially be computed for this abstracted problem and then incrementally improved by reducing the abstraction in the problem. We have successfully demonstrated this in the context of bike sharing systems as explained later.

(d) **Decomposability**: In many of the urban decision problems, it is easy to identify multiple parts of the overall problem that are nearly decomposable. For instance, in bike sharing systems, the problem of moving bikes between stations so as to reduce lost demand and the problem of finding routes for trucks that move bikes are nearly decomposable.

We now describe a few of our major contributions (that exploit the above properties) in the context of the categorisation in Figure 1.

2.1 Large scale, Stochastic/Dynamic and Cooperative

We have multiple key contributions in this space and we describe our most recent ones. Our work on online spatio-temporal matching [Lowalekar *et al.*, 2016] provides strategies for continuously matching available supply (ex: taxis) to demand (ex: customers) while considering the impact of the match on potential demand that will arrive in the next time points of interest. The goal is to maximize expected number of jobs/revenue or minimize wait time for customers. We employ Stochastic Optimization with Sample Average Approximation, where potential future demand scenarios are generated from the data. Specifically, we exploit anonymity in matches and decomposability across demand scenarios to provide a scalable mechanism for online sequential matching. We evaluate our approach on two large real world taxi data sets in comparison to the standard greedy (myopic) approach typically employed in taxi applications (ex: Uber, Ola, Lyft, Grab etc.).

Our second work of relevance is on dynamic repositioning of bikes using trucks [Ghosh *et al.*, 2015] to reduce lost demand in bike sharing systems. Due to uncoordinated pickups and drop-offs of bikes between stations, it is fairly common to observe full or empty base stations. Such situations, specifically empty base stations result in loss of demand, which can result in usage of private vehicles by customers and which in turn has a bearing on carbon emissions and traffic congestion. Therefore, we compute sequential decision making strategies to continuously reposition bikes in relevant base stations so as to reduce lost demand. Specifically, we exploit abstraction and near decomposability (between repositioning of bikes and routing of trucks) to provide a scalable approach that generates high quality solutions offline. We evaluate our approach on two large real world bike sharing data sets.

Our third work [Varakantham *et al.*, 2014] of relevance to this category is associated with the general Decentralised MDP (Dec-MDP) model, which is used to represent many multi-agent sequential decision making problems under uncertainty. Specifically, this research is motivated by the need for coordinating traffic or security personnel (supply), to improve coverage and reduce predictability in patrols so as to reduce security incidents or traffic violations (demand). Specifically, we provide a new model and optimization approaches that are better equipped to exploit homogeneity and anonymity . Our approaches are able to generate solutions efficiently for multi-agent problems with hundreds of agents and we demonstrate superior performance to existing approaches, specifically on large scale problems.

Finally, in our work on emergency response [Saisubramanian *et al.*, 2015], we compute dynamic movement strategies for ambulances or fire trucks to move between base stations so as to reduce response times for emergency events. We exploit homogeneity in ambulances, anonymity in matches and decomposability in emergency request graph to provide a scalable approach. We evaluated our approach on two real world emergency response data sets from asian cities and demonstrated improvement over current practice and current best approach.

2.2 Large scale, Stochastic/Dynamic and Competitive

Even in this category, our first key contribution [Varakantham *et al.*, 2012] is in the context of taxi fleets. However, in this work, the the goal is to provide decision support to selfish taxi drivers (when they have no customer on board) on moving between different zones so as to increase chances of finding customers. This is in contrast to our work on online matching [Lowalekar *et al.*, 2016], where taxi driver has clear incentive to follow the decision provided (as he/she profits from using the application). We provide a new model that is a combination of the Stochastic Games and Congestion Games models to represent these problems of interest in a concise way. The key challenge in this work is to provide decisions where individual taxi drivers do not have an incentive to deviate from the provided strategy. By exploiting anonymity in agent interactions and augmenting the well known Fictitious play approach, we provide a scalable mechanism for equilibrium strategy computation that was better in comparison to greedy strategies typically employed by taxi drivers. We demonstrated these results on a simulation validated on a large taxi data set in Singapore.

Our second contribution [Ghosh *et al.*, 2016] extends on our earlier work in bike sharing [Ghosh *et al.*, 2015] to consider cities where there is a significant variance in demands at base stations. Specifically, our goal is to compute robust repositioning strategies by assuming an adversarial environment that aims to increase lost demand. We exploit abstraction and fictitious play to provide a scalable approach. We evaluated

our algorithm on a new real world bike sharing data set where there is significant variance in demands.

3 Results

Our approaches have been evaluated on simulations that are validated using large real world data sets. Few of the key concrete results are as follows:

1. In taxi fleets where taxi drivers employ applications (ex: Uber, Ola etc.) to obtain customers, we have demonstrated the limitations of myopic reasoning adopted in those applications. More importantly, we show that online sequential decision making strategies that anticipate future demand yield up to 90% of optimal solutions in comparison to 60% with myopic approaches.

2. In taxi fleets where taxi drivers operate individually and in their own selfish interest, we demonstrated an increase of both average and minimum revenue for taxi drivers, along with an increase in availability of taxis to customers by following equilibrium strategies. Concretely, we demonstrated a revenue increase of 40 SGD per day in expectation for each taxi driver. These results are based on a 2 year dataset of a major taxi company in Singapore.

3. We were able to reduce the key performance indicator for emergency response systems, namely the α-quantile response time ($\alpha = 0.8$) by at least 2 minutes on two real world datasets from asian cities.

4. On bike sharing data sets, we demonstrated a reduction of 22% and 45% in lost demand on two real world bike sharing datasets over current practice (repositioning at the end of the day). We also demonstrated a reduction of 10% and 42% in lost demand over an online myopic approach.

5. We improved scalability of decentralised power supply restoration [Agrawal *et al.*, 2015] by at least 30 fold by exploiting near decomposability amongst regions.

References

[Agrawal *et al.*, 2015] Pritee Agrawal, Akshat Kumar, and Pradeep Varakantham. Near-optimal decentralised power supply restoration in smart grids. In *International Joint Conference on Autonomous Agents and Multi-Agent Systems (AAMAS)*, 2015.

[Becker *et al.*, 2004] Raphen Becker, Shlomo Zilberstein, Victor Lesser, and Claudia Goldman. Solving transition independent decentralized Markov decision processes. *Journal of Artificial Intelligence Research*, 22:423–455, 2004.

[Dolgov and Durfee, 2006] Dmitri Dolgov and Edmund Durfee. Resource allocation among agents with MDP-induced preferences. *Journal of Artificial Intelligence Research*, 27:505–549, 2006.

[Gale and Shapley, 1962] D. Gale and L. S. Shapley. College admissions and the stability of marriage. *American Mathematical Monthly*, 69:9–14, 1962.

[Ghosh *et al.*, 2015] Supriyo Ghosh, Pradeep Varakantham, Yossiri Adulyasak, and Patrick Jaillet. Dynamic redeployment to counter congestion or starvation in vehicle sharing systems. In *International Conference on Automated Planning and Scheduling (ICAPS)*, 2015.

[Ghosh *et al.*, 2016] Supriyo Ghosh, Michael Trick, and Pradeep Varakantham. Robust repositioning to counter unpredictable demand in bike sharing systems. In *International Joint Conference on Artificial Intelligence (IJCAI)*, 2016.

[Guestrin *et al.*, 2001] C. Guestrin, D. Koller, and R. Parr. Multiagent planning with factored MDPs. In *Proc. of the Neural Information Processing Systems*, pages 1523–1530, 2001.

[Lagoudakis *et al.*, 2004] M. Lagoudakis, P. Keskinocak, A. Kleywegt, and S. Koenig. Auctions with performance guarantees for multi-robot task allocation. In *International Conference on Intelligent Robots and Systems*, pages 1957–1962, 2004.

[Lowalekar *et al.*, 2016] Meghna Lowalekar, Pradeep Varakantham, and Patrick Jaillet. Online spatio-temporal matching in stochastic and dynamic domains. In *AAAI Conference on Artificial Intelligence (AAAI)*, 2016.

[Pragnesh Jay Modi and Yokoo, 2005] Milind Tambe Pragnesh Jay Modi, Wei-Min Shen and Makoto Yokoo. ADOPT: Asynchronous Distributed Constraint Optimization with Quality Guarantees. *Artificial Intelligence*, 161(1):149–180, 2005.

[Puterman, 1994] Martin L. Puterman. *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. John Wiley & Sons, Inc., New York, NY, USA, 1st edition, 1994.

[Rosenthal, 1973] Robert W Rosenthal. A class of games possessing pure-strategy Nash equilibria. *International Journal of Game Theory*, 2(1):65–67, 1973.

[Roughgarden and Tardos, 2002] Tim Roughgarden and Éva Tardos. How bad is selfish routing? *Journal of the ACM*, 49(2):236–259, 2002.

[Saisubramanian *et al.*, 2015] Sandhya Saisubramanian, Pradeep Varakantham, and Lau H. Chuin. Risk based optimization for improving emergency medical systems. In *AAAI Conference on Artificial Intelligence (AAAI)*, 2015.

[Shapley, 1953] L. S. Shapley. Stochastic games. *PNAS*, 39(10):1095–1100, 1953.

[Varakantham *et al.*, 2012] Pradeep Varakantham, Shih-Fen Cheng, Geoffrey Gordon, and Asrar Ahmed. Decision support for agent populations in uncertain and congested environments. In *Proceedings of the AAAI Conference on Artificial Intelligence (AAAI)*, pages 1471–1477, 2012.

[Varakantham *et al.*, 2014] Pradeep Varakantham, Yossiri Adulyasak, and Patrick Jaillet. Decentralized stochastic planning with anonymity in interactions. In *Proceedings of the AAAI Conference on Artificial Intelligence (AAAI)*, pages 2505–2512, 2014.

[Yokoo and Hirayama, 2000] Makoto Yokoo and Katsutoshi Hirayama. Algorithms for Distributed Constraint Satisfaction: A Review. *Autonomous Agents and Multi-Agent Systems*, 3(2):198–212, 2000.