Research Statement

Linsen Chong

My research aims at developing practical tools from large-scale optimization, simulation, and data-driven techniques that improve the efficiency, robustness and reliability of complex systems, such as urban transportation systems. This task is challenging because of the scale, the occurrence of stochasticity, and the complex interactions between different components within a system. To overcome these challenges, my research involves: (i) developing tractable analytical models that infer critical information from data, and describe and analyze the behavior of a complex system, and (ii) developing practical optimization algorithms that embed these analytical models, data-driven models and simulation models to efficiently derive effective decisions.

My thesis research focuses on addressing challenging congestion problems of urban transportation systems. Currently, the efficiency of the urban transportation systems of many cities around the world does not reach its potential due to the traffic congestion problem, which causes significant delays, fuel overconsumption and high concentration of harmful emissions. I develop effective traffic management strategies that improve the efficiency of urban transportation systems using a variety of techniques that are suitable for addressing problems with intricate data and complex traffic patterns. These techniques include simulation-based optimization, derivative-free optimization, supervised learning, stationary and transient queueing theory, traffic control, as well as traffic flow theory. The effectiveness of these strategies has been shown in several large-scale urban transportation systems (Lausanne, New York City and San Diego). My research also offers important insights on the design of future urban transportation systems that should be adapted to new technologies (e.g., autonomous vehicles).

In my thesis research, I develop practical optimization algorithms for large-scale problems through the use of urban simulators. In this type of problem, known as the simulation-based optimization problem, the underlying complex system of interest is represented by a simulator and the objective function is simulation-based (i.e., the objective function has no available analytical form). This type of problem occurs when the underlying system is too complex to be modelled by any stand-alone analytical model, as is the case for many urban transportation optimization problems. Thus, a simulator is used as a realistic replication of the system. My task is to find the best solution (e.g., system operation strategy) via simulation with limited computational budgets (i.e., simulation runs).

Unlike traditional simulation-based optimization algorithms that treat simulators as a black-box and rely only on simulation information in the optimization procedure, my algorithms combine simulation information with structural information from problem-specific analytical traffic models as well as information from data-driven methods. The extra information accelerates the optimization process, which allows my algorithms to identify effective solutions with limited simulation runs. My algorithms advance the existing simulation-based optimization algorithms in addressing three areas of large-scale problems: stationary, dynamic and real-time problems. Most traditional algorithms are not capable of addressing these large-scale problems with limited computational budgets (i.e., simulation runs); my research fills in this gap.

The key to the computational efficiency of my algorithms lies in its use of problem specific analytical traffic models. They are macroscopic models that describe system behavior of the simulator. Thus,
although they do not include as detailed information (e.g., disaggregated, individual based) as the simulator, they are able to utilize the structural information of the underlying simulator, represent key phenomenon in the simulator and approximate the impact of decision variables on the objective function. In addition, the analytical models are designed to be tractable and computationally inexpensive to evaluate. The accuracy and tractability allow these models to be used to enhance the computational efficiency of many algorithms, such as standard optimization algorithms and estimation algorithms, in which analytical models, instead of simulators, need to be evaluated numerous times.

As noted, in my thesis research, I develop practical optimization algorithms to address three types of large-scale urban transportation problems: stationary, dynamic and real-time traffic management problems. For each type, I develop a corresponding optimization algorithm that embeds a novel problem-specific analytical model. However, my optimization algorithms are not limited for addressing optimization problems in transportation. Since by design they can incorporate any problem-specific analytical model, as long as an accurate and tractable model is available, my algorithms can be applied to any general large-scale optimization problem.

The rest of this statement summarizes my thesis research (i.e., the past and current research), and my future plans. Each section presents a type of large-scale optimization problem, the corresponding algorithm, and my contributions.

**Simulation-based optimization algorithm for stationary problems**

This work, published in Transportation Science (Osorio and Chong, 2015), proposes a computational efficient simulation-based optimization algorithm suitable to address high dimensional, generally constrained and stationary optimization problems. In the case study of this work, the new algorithm addresses a large-scale time-independent urban traffic signal control problem with a limited computational budget.

In this algorithm, optimization is performed on a surrogate analytical model, named as the stationary metamodel. This model consists of a stationary analytical traffic model that approximates the underlying simulator phenomenon (i.e., stationary spatial propagation of congestion) and a polynomial function of decision variables that can be used to improve the accuracy of the approximation during the optimization procedure. The stationary metamodel algorithm is used to address a traffic control problem of the city of Lausanne, Switzerland, with 99 decisions variables. This is considered a large-scale problem in the field of simulation-based optimization. Our results show that the metamodel algorithm can identify effective traffic management solutions with only 150 simulation runs. The solutions derived by the algorithm outperform those obtained by a traditional simulation-based optimization algorithm and a widely used commercial traffic signal timing software. This case study illustrates that the metamodel algorithm is suitable to address high dimensional complex transportation optimization problems and can identify effective solutions efficiently.

The stationary metamodel assumes that the system of interest has reached stationarity. Therefore, this algorithm is suitable to address a variety of optimization problems of stationary systems. The key to reaching efficiency is towards the use of an accurate and tractable stationary analytical model. In this study, the analytical model is developed based on the underlying properties of the urban simulator and does not depend on the objective function. Therefore, this work can easily be extended to address a
variety of large-scale problems, such as deriving management strategies that improve travel time reliability and strategies that reduce the fuel consumption of an urban transportation system.

**Simulation-based optimization algorithm for dynamic problems**

This work extends the capability of the metamodel algorithm for addressing dynamic simulation-based optimization urban transportation problems. This type of problem has time-dependent decision variables and constraints. To efficiently address these problems, the optimization algorithm needs to embed a dynamic analytical model that takes into account an extra dimension (i.e., time dimension) to describe the temporal impact of decision variables on system performances. Given the complexity of the dynamic phenomena of an urban transportation system, developing a tractable yet accurate analytical model is the major challenge.

In this work, published in Transportation Science (Chong and Osorio, forthcoming), I develop a new metamodel algorithm that includes a novel dynamic analytical traffic network model. This model, named as the dynamic metamodel, provides a detailed description of both spatial and temporal patterns of congestion, and is tractable for optimization purposes. To develop this model, I discretize the time span of interest into disjoint intervals, design interval-dependent parameters to represent time-dependent traffic conditions, and develop a novel transient queueing network model that incorporates these parameters to describe spatial and temporal congestion propagations. Through this key technique, the model complexity is only linear in the number of design intervals, which makes it computationally feasible for large-scale optimization.

In Chong and Osorio (forthcoming), I have applied the dynamic metamodel algorithm to solve a time-dependent large-scale traffic signal control problem for the city of Lausanne, showing that the proposed algorithm systematically identifies effective solutions with limited computational budgets. These solutions also outperform those derived by the stationary metamodel algorithm (i.e., Osorio and Chong (2015)), which illustrates the added value of incorporating a dynamic model to the metamodel optimization algorithm.

The dynamic metamodel algorithm can be applied to a variety of complex dynamic problems (e.g., scheduling problems). As long as a tractable dynamic analytical description of a problem is available, this algorithm can be used to address the problem efficiently.

**Simulation-based optimization algorithm for real-time problems**

In real-time simulation-based optimization problems, decisions need to be made for every short time interval. Therefore, for every interval, an optimization problem needs to be solved before the starting time of the next interval. This type of problem is very challenging since the optimization process needs to be carried out in real-time. My contribution in this field will be a novel optimization algorithm that allows this type of problem to be solved in real-time.

In this work, I addressed real-time traffic control problems where decisions need to be made every short interval of interest. To do this, I develop a novel optimization algorithm that consists of two methods: a model-driven optimization method and a data-driven machine learning method. The model-driven method is the metamodel simulation-based optimization algorithm I developed in Osorio and Chong (2015). The

Unlike most current real-time transportation optimization algorithms that use either model-driven methods or data-driven methods independently, the combination of two methods can potentially improve the performance of the online optimization framework. As shown in our experiments, when solving a traffic responsive control problem of Lausanne, the algorithm can derive signal plans that outperforms those derived by a framework with only the model-driven method and a framework with only the data-driven method. This research shows great potential for addressing a variety of online transportation problems: as more real-time data is becoming available yet these data is normally sparse in transportation problems (e.g., in an urban network, sensors are normally deployed on limited roads), the proposed framework that includes both problem-specific structural information (used in the model-driven method) as well as uses real-data as much as possible (in the data-driven method) is critical.

**Future work**

In the near future, my research focus will revolve around the analysis and the management of complex urban transportation systems. Currently, in the “big data” era, ubiquitous and accessible data collected by multiple sensors, such as GPS, road side units and mobile phones offers great opportunities to understand hidden mechanisms of urban transportation systems along with future challenges, as well as to find practical approaches to overcome them.

In this quest, I intend to develop practical models with techniques from statistical inference, applied probability, machine learning, operations research as well as traffic control theory and traffic control theory. In particular, I am interested in developing models that combine problem-specific analytical models and general purpose data-driven models, such that my models will benefit from the advantages of these two types of models (e.g., robustness, computationally efficiency, scalability, accuracy, etc.). I believe these models will provide guidance for the design and the implementation of better system management strategies, especially when future technologies (e.g., autonomous vehicles) emerge.

In the long run, I am broadly interested in addressing challenging large-scale practical problems using operations research methodologies, machine learning and statistical inference technologies. To address these problems, we first need to identify the mechanisms that lead to the problems in the first place using effective inference and machine learning techniques. Then, we need to use computationally efficient optimization algorithms from operations research to solve them. I believe the experience and skillsets I have gained for addressing complex urban transportation problems, especially the real-time traffic management problems, will help me achieve my goal.

**References:**
