The Consultant

A Tale of Two Stakeholder Collaborations in Transportation

I'm thinking for this we need a slightly different skill set.

Philip J. Coulson

1 Gridwise

My collaboration with *Gridwise*, the developer of a gig driver assistant app, began in June 2021. This collaboration came about through a connection to my co-advisor, Norman Sadeh. From the standpoint of academic research, collaborating with Gridwise constituted a unique opportunity:

- *Gig platforms* have the data and resources to conduct large-scale studies such as that of Ong et al. [14]. Yet, such research is primarily motivated by optimising a given platform's own performance in a competitive market, and thus it rarely considers cross-platform dynamics.
- *Gig drivers* seek to optimise for individual goals and offer insight into multiple platforms [16, 19], but information asymmetry limits their visibility into market conditions.
- *Gig assistants* developed by third parties, such as Gridwise, integrate data from many individual drivers and platforms at scale, providing a uniquely comprehensive base of data.

The closest point of comparison in the literature is Calacci et al. [4]'s FairFare tool for estimating platform take rate, which was developed in collaboration with a driver union. However, the size of their dataset is smaller and much less generalisable (76 000 trips over 18 months, likely from a single city, compared to over 100 000 trips per month per city from Gridwise).

1.1 Project Goals

At our initial meeting with Gridwise's executive team, we intended to find a project that involving machine learning, optimisation, and game theory, with potential research directions including spatiotemporal pricing [12], platform choice [9], and driver modelling. However, through discussions with Gridwise, we quickly identified driver scheduling as a major pain point, since inefficiencies in scheduling cause drivers to deadhead in their vehicles without fulfilling requests. Thus, working time and location recommendations offer the potential to improve drivers' quality of life. However, we also learnt that drivers resist having their behaviour prescribed to them; an example of this can be seen in drivers' resistance to surge-chasing [11]. We therefore concluded that schedule recommendations should be presented to drivers as a set of options, not forcefully. In October 2021, we reached an agreement on the goal of the project: to develop a schedule recommendation algorithm that took as input historical and user-specific data (including schedule constraints) to suggest schedules of working times, locations, and platforms.

1.2 First Attempts

By March 2022, Gridwise had supplied a multi-year dataset of trips from Los Angeles, and I initially used it to build a linear regression model of earnings for individual trips. As part of this, I invested a considerable amount of work into engineering relevant features: platforms, start and end times, start and end neighbourhoods, special events, and so on. From exploring the data, I found that earnings could be surprisingly variable, and that tips and bonuses — which are not necessarily predictable a priori — constitute a large portion of drivers' earnings.

However, trip-level predictions were not the end goal of the project. To evaluate deadhead time, I needed to consider the expected number and duration of all trips taken by drivers over an (e.g. hourly) shift. In doing so, optimising for driving locations emerged as an insurmountable technical challenge. Every trip repositions a driver to a potentially distant location where patterns of supply and demand may differ. Drivers also have different preferences and constraints for where to start and finish each day's driving activity. These factors resulted in a combinatorial explosion in the number of potential routes to optimise over. Furthermore, Gridwise's dataset only captured trips made by drivers, not repositioning actions taken during deadheading, which made it infeasible to accurately condition drivers' earnings on their locations. Eventually, I simplified the prediction task to generate estimated earnings for each hourly timeslot during a week.

Another challenge was the extreme imbalance in the data. This was not a data quality issue, but rather an inherent property of the environment. Most drivers are relatively casual and may only log several dozen trips over years, while a handful are highly active drivers who may log thousands of trips in the same timeframe. Meanwhile, we also needed to rely on relatively recent data, since market dynamics shift over time. Both issues hindered my ability to personalise predictions and recommendations for individual drivers. My prediction model initially estimated both market-level and individual earnings. When I filtered historical trips to individual drivers, there was either insufficient data (as drivers tend to stick to the same routines) or insufficient generalisability between drivers (as highly active drivers would be oversampled). Following my suggestion in Section 3.8.1, we considered generating estimates based on data from clusters of similar drivers by location or platform, but this was ultimately never implemented.

Once I completed exploratory data analysis and an initial prediction model, I developed the optimisation problem from Section 3.4.1: to maximise predicted earnings subject to drivers' availability and goals. I also created a web-based prototype of the tool from Section 3.4.2, and demoed it in May 2022. Gridwise was excited by our prototype, but informed us that optimising for platforms was more important than optimising for locations. I added various options that would allow drivers to select the constraints they wanted to activate. I considered learning these constraints from data, but we reasoned that we could maximise perceived control by eliciting them through a user interface. Based on my observation of variance in the predictions, I also introduced a range-based uncertainty display in the predicted earnings, using the 10th and 90th percentiles (Section 3.4.3).

1.3 Best Laid Plans

User-centred design was a key consideration from the beginning, and the goal was always to carry out a user study like Chapter 3 to assess and improve the effectiveness of the tool. Our evaluation plan was threefold: (1) conducting a formative study to understand what inputs and outputs drivers needed from the tool (Section 3.3); (2) working with Gridwise's product and UX teams to convert the tool into a deployable form, as a static schedule or an interactive webpage; and (3) deploying the tool with a sample of drivers to receive feedback. Gridwise recruited a sample of 43 drivers for focus groups by August 2022. We explored two options for the pilot study. First, Gridwise could host the tool while interfacing with my optimisation backend, which would require API interaction, but would not be human subjects research because Gridwise would anonymise all data before sending it to us. Second, we could host both the tool and the optimisation backend, which would require no implementation effort by Gridwise but would require Institutional Review Board (IRB) approval. After repeated exchanges, I received approval from our IRB in October 2022 for the first option, which they did not consider to be human subjects research.

Under our initial plan, Gridwise would draft the survey for the user study, and I would provide feedback. In the end, we drafted three survey instruments collaboratively: Survey 1 (Intake), Survey 2 (Pre-Survey, after recommendation but before use), and Survey 3 (End of Day survey, after use); see Section 3.5. Compared to the survey questions in Appendix A.1.1, our draft focused more on the usefulness of the recommendations and their impacts on drivers' outcomes, rather than on the impact of uncertainty displays. We also identified the best modes of delivery: a web-based survey and email-based schedule recommendations. This was sufficient for a one-time study, although Gridwise recognised that a continuous model update and deployment process would be needed for production. I delivered an API for my tool to Gridwise along with the survey instruments, and we targeted early 2023 for the start of the study.

Gridwise reassessed their product priorities at this point. In February 2023, we were informed that allocating development and logistical resources to our collaboration would be difficult, even with the steps that we took to minimise the effort Gridwise would need to expend. In particular, we were informed that the project timeline had become desynchronised by the two months I had spent communicating with the IRB. Two options were possible at this point: either I could shoulder the majority of the effort, or I could complete the study as an intern for Gridwise. From the perspective of collaboration effort and intellectual property, we preferred the first option.

1.4 Bouncing Back

Despite the roadblocks we encountered, we committed to moving this project forward to at least produce a publication. We considered either improving the algorithmic components (for an AI venue) or refining the user study (for a human-computer interaction venue). Given that a sophisticated algorithm would not be practically meaningful if users did not find it useful, I chose the latter direction. This choice led to the simplification of my earnings prediction model to use a historical average, rather than a more sophisticated machine learning model.

The next several months were divided between multiple tasks. On the one hand, I reviewed the literature on human-AI interaction to identify a concrete contribution (which heavily informed Section 3.2). During this process, I chose to focus the study on trust, specifically based on the model of Solberg et al. [17]; my survey instruments were changed to incorporate the HCT [13] and TiA [10]. On the other hand, I also designed a revised interview protocol for the IRB, including changes to the recruitment procedure, sample size, and compensation structure. Our limited research budget required me to make the survey instruments and the overall study as concise as possible. We also had to clearly separate Gridwise's role (distributing recruitment emails) from our own (as the principal investigators). At this point, Gridwise did not have time to meet with us, and we were not actively generating any new results that could further the collaboration.

By June 2023, the revised protocol was nearly finalised. As with our initial plan, we first sought to conduct a pilot study, so as to improve the interface design of the tool before deploying it in a full-scale user study. Two Master's students who expressed interest in the project worked with me to design the Figma prototype in Section 3.3.1. With a prototype in hand, we asked for help from Gridwise to recruit a small sample of drivers. Gridwise's growth team met with us in late June to begin collaborating on participant outreach and recruitment. The growth team took an interest in our research and became our main point of contact for the remainder of the project.

1.5 All Roads Lead...

We received IRB approval in July 2023. Afterwards, we iterated on the wording of email communications with the growth team, which led to further changes to the protocol. Concurrently, I expanded the demo website from May 2022 to serve as a website for the entire study. Once the pilot was ready, Gridwise distributed recruitment messages to an initial wave of 500 active drivers in Los Angeles, but only two participants came from this sample. This led me to realise that recruitment and retention would be significantly more difficult for gig drivers than other populations (e.g. workers on Amazon Mechanical Turk). A second round of outreach resulted in two more participants; a fifth participant dropped out.

Given these difficulties in recruiting, I was eager to begin the full user study as quickly as possible. In early August, we communicated to the Gridwise growth team that we needed a new sample of historical trip data — as well as current data to assess the impact of the tool's recommendations on drivers' earnings. Unfortunately, however, no pipeline for exporting the current data existed. It was too much of a risk for us to ask Gridwise to create such a pipeline. The best solution that we could devise involved manual, once-daily uploads of a CSV file that I programmatically retrieved. However, this still made it infeasible for the tool to provide live feedback to drivers. These limitations partially motivated me to shift the study from focusing on the impact of the tool towards focusing on the impact of its design and of environmental uncertainty.

The full user study launched with several email campaigns from Gridwise that exhausted the pool of active Gridwise users in Los Angeles. These campaigns had a conversion rate of 1% — significantly lower than we had hoped. Almost half of the email recipients had opened our email but did not participate; we hypothesised that the length of the email copy, the time investment, or the compensation could be potential causes. In early September 2023, we decided to expand the study to include the three other cities: New York, Chicago, and Houston. We also relaxed the pool of participants from active users of Gridwise to include all active drivers (who may or may not be actively using Gridwise). I excluded the previous participants from my analysis.

Over three weeks in these four cities, Gridwise released several email campaigns each week to capture drivers who were available at different times. I was able to recruit the 51 participants for the user study described in Section 3.5.1. Even then, ensuring that participants followed up across all 7 days of the longitudinal study was challenging. Before the study began, I had relaxed the participation requirement from 7 consecutive days to 7 out of 14 days. Initial observations of attrition led me to implement daily reminder emails to prompt participants who had not completed all required surveys. These measures led to a good, but still suboptimal, retention rate of $^2/_3$.

Recruiting participants for the follow-up interview (Section 3.5.3) was the most difficult task. I needed participants to freshly recall their experiences in the user study, but it took several days for a handful of participants to follow up with their availability. The low response rate even led me to follow up with participants who did not complete the rest of the user study, although I was unsuccessful. Furthermore, several participants repeatedly rescheduled interviews — one claimed their phone was stolen — and ultimately did not attend. Nevertheless, the responses of the seven attendees still led me to a variety of valuable insights (Section 3.7).

Working with a third Master's student, I submitted a paper based on this project. On the next day, we shared our findings with Gridwise's growth and executive teams. They found our results to be interesting and also counterintuitive in places (especially regarding the effects of uncertainty on trust and retention). Although we hoped to continue working on the algorithmic portion of the project, our active collaboration with Gridwise ultimately ended here.

2 Path Master, Econolite, and Strongsville

My collaboration with stakeholders in traffic signal control (TSC) came through *Traffic21*, our transportation research institute. Initially, we planned to collaborate with the *Traffic Management Centre (TMC) of Cranberry Township*, a nearby municipality with both challenging traffic logistics and modern infrastructure. They had recently deployed the Edaptive algorithm of *Econolite*, a leading traffic technology company whose controllers are used for ⅔ of traffic signals in the United States and ⅓ of traffic signals in Canada. During our first meeting in September 2021, the Cranberry TMC connected us with two other stakeholders: Econolite and *Path Master*, a traffic technology distributor. Again, these groups of stakeholders offered unique opportunities:

- *Traffic management centres* are intimately familiar with traffic conditions in the municipalities that they manage, and are the first line of response to public concerns about traffic issues. As such, they have deep expertise in traffic problems within a specific deployment context.
- *Traffic technology companies* are constantly seeking to innovate their software and hardware platforms, so that they can deliver next-generation systems to a variety of municipalities. As such, they have broad expertise in traffic solutions that apply across deployment contexts.

- *Traffic technology distributors* and *traffic engineering consultants* are typically responsible for distribution and deployment across a smaller but still diverse range of municipalities. As such, they are well-versed in connecting general traffic solutions to specific traffic problems.

2.1 Project Goals

Our initial proposal to the Cranberry TMC involved two project ideas: a quantitative evaluation of Edaptive's performance, and the automated integration of an incident detection and signal plan recommendation system into Edaptive. The latter project was a particular pain point for the TMC: Yao and Qian [18] had developed a signal plan recommendation system that detects traffic anomalies from crowdsourced data, and notifies traffic engineers of potential traffic flow issues in advance. However, the TMC would need to manually activate the recommended signal plans, and they wanted these capabilities to be automatically integrated into Edaptive.

In subsequent conversations with Traffic21, we recognised that we would be better positioned to make contributions through a direct collaboration with Econolite — in particular, to apply my research interest of multi-agent reinforcement learning (MARL) to TSC. In October 2021, we began collaborating directly with Econolite, and we finalised the necessary agreements over the following weeks. During the intervening time, I familiarised myself with common practice in the field, including the use of traffic simulation in VISSIM and SUMO. Due to its balance of realism and efficiency, I settled on SUMO as a simulator for training MARL algorithms. However, Econolite also connected us with PTV, the developer of VISSIM (soon after, Econolite acquired PTV). VISSIM provides more advanced simulation capabilities than SUMO and would be more strongly favoured for evaluation by stakeholders.

Generally speaking, the goal that emerged from our early conversations with Econolite was to outperform the state of the art in TSC, particularly the link-pivoting heuristics of Day and Bullock [7], while also taking uncertainty in detection into account. My exploration of uncertainty and other key considerations in deploying RL for TSC as an end-to-end system eventually led to Chen et al. [6], the 2022 review paper that informed the structure of this thesis.

As the collaboration continued over the following months, my priority was to use SUMO to model a deployment site that I could use to evaluate MARL algorithms. The ultimate goal was to collaborate with the deployment site to bring MARL into the field, at least as part of a small-scale test with several intersections. Econolite and PTV were unable to engage with us for several months. However, in the meantime, my literature review made me aware of MPLight [5] and other state-of-the-art MARL algorithms for TSC. I identified the lack of heterogeneity and constraint enforcement as concrete limitations in these algorithms; these concepts eventually developed into Chapters 4 and 6. At the same time, I experimented with using denoising techniques to account for uncertainty in RL, using the SUMO-based codebases of sumo-rl [1] and RESCO [2].

When Econolite re-engaged with us in March 2022, we discussed potential deployment sites and evaluation criteria (e.g. percentage on green as opposed to queue length; see Section 2.2.1). Cranberry Township was not particularly viable as a deployment site due to their recent implementation of Edaptive; it would have been difficult to isolate inherent limitations in Edaptive from yet-unresolved implementation issues. We did not identify any promising deployment sites until our first joint meeting with Econolite and Path Master in April 2022. Path Master suggested several deployment sites that were close to us in Pennsylvania and Ohio. We sought a deployment site that: (1) had heavy cross traffic and was not just a single arterial corridor; (2) had modern detection and communication infrastructure; and (3) had an existing deployment of Edaptive. All of these considerations led us to the city of Strongsville, Ohio, which has a road network similar in size and traffic to that of Cranberry Township.

2.2 First Attempts

During the summer, we worked with Path Master in contacting the *TMC of Strongsville*. Our pace of communication for this project was much slower than with Gridwise, but more in line with that of typical stakeholder collaborations. Eventually, our first meeting with the Strongsville TMC took place in July 2022. During this meeting, we framed RL-based TSC as not just a research project, but an exciting technology with as-of-yet unrealised deployment benefits. I found it important, if challenging, to place less emphasis on our technical expertise than on our capacity to address Strongsville's needs. Strongsville expressed interest in the project and committed to sharing data with us. They also suggested potential directions to explore: driver perception (which also influenced Chapter 4) and preemption for transit and emergency vehicles. Nevertheless, they also had some reservations about how RL could disrupt their traffic patterns.

In August 2022, we met with PTV about licencing the VISSIM software. We also received a VISSIM simulation of the Strongsville road network that they had converted from OpenStreetMap data. However, the simulation did not contain any vehicles, traffic signals, routes, or any of the other essential components of a traffic simulation. A brief lapse in communication occurred, during which I designed an RL training environment for VISSIM based on Alegre et al. [1]'s SUMO wrapper. This was to ensure that we could use VISSIM to demonstrate the performance of the trained RL policy — but not necessarily for training, due to its slower speed compared to SUMO.

Through working with Strongsville, Path Master provided me with access to Centracs, the city's cloud-based traffic monitoring system. This was a more efficient alternative to Path Master manually exporting count data. Although only 36 of the 57 intersections in Strongsville's road network had data that I could access through Centracs, I considered this to be a sufficient sample. By November 2022, I gained limited access to hourly detector count data through Centracs, which I began to import into the VISSIM simulation. To associate individual detectors with their corresponding lanes in the simulation, Path Master shared spreadsheets of detector assignments. However, it was unclear how I should separate lane-level count data into turning movements (e.g. dedicated right-turn lanes do not exist at most of Strongsville's intersections). Although Path Master indicated that it would be possible to set up detection zones for turning movements, this would be a significant burden on their part. My initial approach involved a constant 70%/30% split. However, after repeated iteration to improve the accuracy of the simulation, this evolved into the quadratic integer program from Section 5.3.2. I also injected a constant 10% of heavy vehicle traffic, but revised this to reflect ratios for different vehicle classes based on ODOT data.

Path Master also shared signal plan documents that identified phases, timing constraints, and coordinated signal plans at each intersection. This allowed me to replicate them in the simulation, and also to define my RL formulation more precisely (Section 6.5.1). I discovered inconsistencies between these signal plans and the base time-of-day signal plans actually implemented in

Strongsville (for instance, the base patterns in Centracs signal some phases for longer than their maximum green durations in the documents). My final simulation merges these two data sources.

2.3 Meet the New Boss

While I developed the simulation of Strongsville, we began planning a site visit to better understand traffic conditions in the city. We also felt that the Strongsville TMC, and their traffic engineering consultants at *TMS Engineers*, would be better positioned to answer our questions through an in-person discussion. This site visit took place on November 22, 2022, with me and my co-advisor Fei Fang in attendance, along with several of our collaborators from Path Master. For the site visit, I had prepared an initial version of the VISSIM simulation that I could present, but I also wanted to verify my modelling assumptions. However, rather than asking these stakeholders for their input, we approached the site visit from the perspective of helping Strongsville with their pain points, with our research inputs as a byproduct.

Our site visit occurred shortly after several traffic issues in Strongsville: a camera detector failure, followed by a controller failure, and a pedestrian being struck by a car. The city also informed us of their routine traffic challenges, which highlighted the importance of this project:

- **Holiday traffic**. Traffic in the city reaches a peak in the two months following Thanksgiving. Much of this traffic originates from the I-71 offramp on SR 82, which conflicts with traffic from adjacent malls on SR 82. During other occasions, such as July 4 celebrations, the city does not use programmed signal plans and instead uses the police to direct traffic. We discussed learning RL-based control policies for these scenarios from human feedback.
- **US 42 and SR 82**. The central intersection in the city is a daily challenge (Figure 1a), and it also affects adjacent intersections through downstream blockage. This intersection was problematic for many years until a more recent re-tuning of its signal plan. However, during rush hour traffic, it is still common for vehicles to queue for more than one cycle.
- **Detection logistics**. The city's detection infrastructure plays a crucial role in signal plan performance, as demonstrated by the recent camera detector failure. However, some stakeholders had entrenched preferences for camera detectors or loop detectors, due to tradeoffs between reliability and accuracy. These conflicts were hurdles to planning and maintenance.

As before, we wanted to determine the feasibility of a pilot test. The Strongsville TMC outlined several clear requirements: the signal plan design would need to be interpretable for the purpose of regulatory compliance, and support traceability in the case of critical incidents; the performance of the system would need to represent a notable improvement to justify its cost to the city council; and the possible responses of drivers would need to be accounted for (e.g. impatient drivers not respecting yellow or even red signals). We committed to not altering any of Strongsville's signal plans until we had confidently verified the performance of RL policies in simulations. This led us to revise our project plan to include offline evaluation, which would entail receiving live detector inputs from Strongsville, and converting them into a real-time simulation to assess the performance of RL policies. For online evaluation, we planned to first deploy on a smaller corridor of 2–3 intersections to build trust before deploying to the rest of the road network.

Upon reviewing our simulation, Strongsville's traffic engineer found it to be very inaccurate. Instead of queues that caused downstream blockage, there were only a handful of vehicles on the roads, even at peak hours. I partially attributed these discrepancies to VISSIM's probabilistic routing, as well as to traffic being aggregated at the hourly level by Centracs. (The engineers advised us to reduce the size of our time segments to 15 minutes.) During the evening rush hour, Strongsville's TMC staff demonstrated their video monitoring system, including an offline Centracs dashboard with detailed information about controller statuses and more granular detector data. We witnessed the traffic challenges that the TMC spoke of on the camera feeds, and shortly afterwards in person when Path Master gave us a tour of several key intersections. We also learnt about the configuration of Strongsville's detection and control infrastructure (Figure 1b).

(a) Queued traffic at the intersection of Pearl Road (US 42) and Royalton Road (SR 82), during the evening rush hour at 5 pm.

(b) Inside a signal controller cabinet installed at the intersection, including an Econolite Cobalt controller on the second row.

Figure 1: Photographs from our site visit in Strongsville, Ohio on November 22, 2022.

On the one hand, our site visit of Strongsville was an extremely informative experience that helped shape this project. On the other hand, it also significantly raised the bar of trust that we would need to achieve before any form of physical deployment would be possible. I shifted away from developing methods to address uncertainty in RL-based TSC, as it was not the most pressing challenge for Strongsville. Instead, my work split into two parallel streams: improving the quality of the simulation in collaboration with Path Master, and improving the performance of RL algorithms on the simulation in collaboration with Econolite. In December 2022, our contact at Econolite left the company and was replaced by another person. This second contact also left the company by April 2023. A third contact joined the project in July 2023, and has been our contact since then. A stakeholder from PTV left the collaboration due to limited involvement in July 2023.

2.4 Adapt or Die

To improve the simulation, I discovered that VISSIM's vehicle routing algorithm caused some of the generated vehicles to disappear in the middle of roadways, as there was not enough space for the vehicles to complete necessary lane changes (see Section 4.4.2). This motivated the "fringe route" regularisation in Section 5.3.2. To iterate more quickly on the simulation, I decided to switch back to SUMO as my main traffic simulator. Since SUMO does not support probabilistic vehicle routing, I initially used the simulator's default demand generation pipeline (see Section 5.2.1). The linear program relaxation used by SUMO's pipeline had systematic issues with generating insufficient levels of traffic compared to the counts, which ultimately led me to create my own pipeline for Chapter 5.

But discrepancies remained between simulated and actual traffic conditions, suggesting a more fundamental issue in the detector data. I sought access to Strongsville's offline Centracs dashboard, where I could generate reports of traffic volumes for each minute. VPN access was eventually granted by January 2023. However, the minute-by-minute data was less comprehensive than I had hoped, and the resulting traffic volumes were also similar to the hourly data. Nevertheless, through this system, I was also able to view Strongsville's live camera detector feeds. By July 2023, I had performed some preliminary manual counting, which suggested that the camera detector counts fell very short of actual traffic volumes in Strongsville.

Path Master suggested using advance and downstream radar detectors to obtain more accurate counts. Radar detectors capture free-flow traffic instead of queueing vehicles at the stop bar, but are not available at all intersections in Strongsville. I tried (1) to interpolate between radar and camera detector counts, and (2) to rescale the camera detector counts to match my manual counts. Neither of these heuristics were technically rigorous, and they still led to inconsistencies in relative volumes. We considered involving Econolite's field staff to validate the detector configuration, but this would have been too costly for the city. Thus, my simulation work stalled for some time.

2.5 Failed Experiments

To improve the performance of RL policies, I first needed to establish a baseline for performance. One issue was that, at the time, I had no way of replicating Edaptive's closed-source controller in my SUMO simulation. We decided to use the base time-of-day pattern chosen by Edaptive, although we recognised that this would increasingly diverge from reality as the simulation progressed. By May 2023, I had evaluated the MPLight algorithm on the Strongsville road network [5]. My evaluation first used data from around the time of our site visit in November 2022, and then newer data from May 2023. To communicate the benefits of RL, I produced animated visualisations of traffic at the intersection of US 42 and SR 82 under different control strategies at 9 am, 12 pm, and 5 pm.

MPLight reduced queue lengths from the base time-of-day pattern by as much as 25%. These results held across different iterations of the simulation, and also as I added vehicles of different classes to the simulation. However, our collaborators suggested that this good performance was mainly due to the policy not following the same signal plan constraints as Edaptive. Over several meetings in October and November 2023, I learnt about Edaptive's optimisation routine and clarified the nature of these signal plan constraints. This led me to recognise that the sumo-rl environment [1] was not suitable due to its lack of support for cycle-based actions and constraints. I developed the event-based environment introduced in Section 6.5.1 over a month.

We discussed several solutions for implementing signal plan constraints with Econolite:

- **Bilevel signal plan optimisation**. From Econolite, we had learnt that Edaptive changes between time-of-day plans infrequently, but performs incremental adjustments to cycles, splits, and offsets more frequently. I considered designing a hierarchical RL algorithm to replace this entire optimisation loop based on the options framework [3], but ultimately we opted to develop RL algorithms that could be used by Edaptive as optimisation subroutines.
- **Hyperparameter optimisation**. A set of parameters controls the optimisation procedure of Edaptive; for instance, it only adjusts cycle lengths (1) when the volume-to-capacity ratio exceeds a certain limit and (2) up to a certain time interval. Currently, traffic engineers set these parameters based on their intuition, but Econolite suggested that we could use RL to suggest values as a minimally intrusive application. However, it was unclear how we could evaluate this without access to Edaptive, and how much improvement we could expect.
- **Phase sequence optimisation**. In Section 6.3, I described how Edaptive always executes phases in the same order. However, Edaptive supports altering the ordering to e.g. convert leading turn phases to lagging turn phases. Part of my thesis proposal involved optimising this ordering using a branch-and-bound procedure [8, 15], in which information from a small set of RL trajectories is used as a heuristic to select phases. However, we did not proceed with this idea, as only a small proportion of customers alter their phase sequences.
- **Interpretability methods**. To make RL-based signal policies more transparent for stakeholders, I considered representing them as decision trees where actions correspond to incremental adjustments of a base time-of-day pattern. For these policies, I would also train models to estimate the counterfactual impacts of these actions on signal performance metrics. This ultimately led to the work in Chapter 7. However, our collaborators considered RL policies that directly output cycle-offset-split plans to be clearer. Econolite also had existing models to estimate counterfactual signal performance metrics.

The final set of techniques that I use to enforce these constraints emerged during the following months: imitation learning (Section 6.4.2) and action projection (Section 6.4.3) in December 2023, and action masking (Section 6.4.1) in January 2024. By March, I discarded the remaining solutions and showed that action masking could improve queue lengths over time-of-day plans by 15%.

Nevertheless, due to the scale of the Strongsville simulation, iterating on RL training was slow and involved. Meanwhile, I continued to pursue other algorithmic developments, including HYDRAVIPER in Chapter 7, using benchmark environments. Although I had invested significant effort into meaningful exploration, we felt that we had reached a plateau; we needed a deeper understanding of factors that would benefit or degrade performance. The Strongsville TMC also desired further improvements to the simulation before they could trust my RL evaluation results.

2.6 New Life

Starting from March 2024, an undergraduate student joined the project to improve my simulation of Strongsville. We identified two likely sources of error: one in the detector data from Strongsville, and one in the count-to-route demand modelling pipeline from SUMO.

Our initial focus was on the detector data. Path Master was surprised by the level of error that we reported based on manual counts; they were concerned about the extent to which these counts may impact Edaptive's performance. The first error correction approach we envisioned involved building a denoising model, but we quickly recognised that we did not have the capacity to manually annotate camera footage on the scale required to produce a training dataset. By April, we began investigating object detection and tracking algorithms, which ultimately led to the method in Section 5.3.1 by September. My initial method for capturing camera footage was using a screen recorder over a VPN connection to Strongsville's offline Centracs dashboard. However, this was not scalable or reliable: the footage I captured suffered from low framerates and pixelation, and I often lost recordings due to the instability of the VPN connection.

In June 2024, Path Master shared the IP addresses of the camera detectors' RTSP feeds. Since they did not install all of the camera detectors in Strongsville, I could access only a subset of these feeds. I developed a script that ran locally on Strongsville's server to automatically capture and record the available feeds into hour-long videos. After verifying that it incurred minimal overhead, I ran this script for 24 hours on September 6–7, 2024. On this dataset, the accuracy of our vehicle tracking method was still hampered by detector overlays. As described in Section 5.3.1, I applied video filters to remove these overlays and to smooth the footage. However, Path Master also reconfigured the streams to remove the overlays, and I recorded a second video sample on December 14, 2024. Our new results gave us confidence in the accuracy of the detector data, and we felt that our vehicle tracking method had enough utility to be deployed as a standalone solution.

Our focus then turned to the demand modelling pipeline in February 2025. This started with an overhaul of the route optimisation method to use the quadratic integer program in Section 5.3.2. My experience with error levels among detector counts from the different sources led me to convert equality constraints to inequality constraints based on bounds estimated from data. The additional regularisation terms were also inspired by the intuition of Path Master. To further leverage Path Master's knowledge, I developed the LLM agent in Section 5.3.3. All of this work culminated in a presentation of the updated simulation to the Strongsville TMC in May 2025. Strongsville's engineers recognised the significant improvement in the simulation, although they had further feedback on several locations where the simulation did not capture realistic traffic patterns. This may have been due to the fact that I used detector data from September 2024 (i.e. a non-holiday period).

Finally, I returned to the implementation of constraints for RL-based TSC, which I expanded into Chapter 6. I corrected implementation errors in my event-based environment, which would likely not have persisted had I invested additional engineering effort back in December 2023.

At this point, we were nearing the end of my PhD. We remained enthusiastic about a possible field test. However, considering the typical pace of stakeholder collaborations, we recognised that the necessary work would take at least half a year, especially since Strongsville expected a mature product rather than a research project in progress. It would likely be difficult to recruit another student to complete the work, given the lack of active funding. Nevertheless, Path Master, the Strongsville TMC, and Econolite in particular all remain interested in advancing the collaboration.

Bibliography

[1] Lucas N. Alegre, Ana L.C. Bazzan, and Bruno C. da Silva. 2021. Quantifying the impact of non-stationarity in reinforcement learning-based traffic signal control. *PeerJ Computer Science* 7 (2021), e575.

[2] James Ault and Guni Sharon. 2021. Reinforcement learning benchmarks for traffic signal control. In *Proceedings of the 35th Conference on Neural Information Processing Systems, Datasets and Benchmarks Track (NeurIPS '21)*. NeurIPS, Virtual, 1–11.

[3] Pierre-Luc Bacon, Jean Harb, and Doina Precup. 2017. The Option-Critic Architecture. In *Proceedings of the 31st AAAI Conference on Artificial Intelligence (AAAI '17)*. AAAI, San Francisco, USA, 1726–1734.

[4] Dana Calacci, Varun Nagaraj Rao, Samantha Dalal, Catherine Di, Kok-Wei Pua, Andrew Schwartz, Danny Spitzberg, and Andrés Monroy-Hernández. 2025. FAIRFARE: A Tool for Crowdsourcing Rideshare Data to Empower Labor Organizers. arXiv:2502.11273

[5] Chacha Chen, Hua Wei, Nan Xu, Guanjie Zheng, Ming Yang, Yuanhao Xiong, Kai Xu, and Zhenhui Li. 2020. Toward A Thousand Lights: Decentralized Deep Reinforcement Learning for Large-Scale Traffic Signal Control. In *Proceedings of the 34th AAAI Conference on Artificial Intelligence (AAAI '20)*. AAAI, New York, USA, 3414–3421.

[6] Rex Chen, Fei Fang, and Norman Sadeh. 2022. The Real Deal: A Review of Challenges and Opportunities in Moving Reinforcement Learning-Based Traffic Signal Control Systems Towards Reality. In *Proceedings of the 12th International Workshop on Agents in Traffic and Transportation (ATT '22)*. CEUR, Vienna, Austria, 1–21.

[7] Christopher M. Day and Darcy M. Bullock. 2011. *Optimization of Offsets and Cycle Length Using High Resolution Signal Event Data*. Working Paper SPR-3409. Joint Transportation Research Program. 36 pages.

[8] Marc Etheve, Zacharie Alès, Côme Bissuel, Olivier Juan, and Safia Kedad-Sidhoum. 2020. Reinforcement Learning for Variable Selection in a Branch and Bound Algorithm. In *Proceedings of the 2020 International Conference on Integration of Constraint Programming, Artificial Intelligence, and Operations Research (CPAIOR '20)*. Springer, Virtual, 176–185.

[9] Xiaotong Guo, Andreas Haupt, Hai Wang, Rida Qadri, and Jinhua Zhao. 2023. Understanding multi-homing and switching by platform drivers. *Transportation Research Part C: Emerging Technologies* 154 (2023), 104233.

[10] Jiun-Yin Jian, Ann M. Bisantz, and Colin G. Drury. 2000. Foundations for an Empirically

Determined Scale of Trust in Automated Systems. *International Journal of Cognitive Ergonomics* 4, 1 (2000), 53–71.

[11] Min Kyung Lee, Danyel Kusbit, Evan Metsky, and Laura Dabbish. 2015. Working with Machines: The Impact of Algorithmic and Data-Driven Management on Human Workers. In *Proceedings of the 2015 CHI Conference on Human Factors in Computing Systems (CHI '15)*. ACM, Seoul, South Korea, 1–13.

[12] Hongyao Ma, Fei Fang, and David C. Parkes. 2022. Spatio-Temporal Pricing for Ridesharing Platforms. *Operations Research* 70, 2 (2022), 1025–1041.

[13] Maria Madsen and Shirley Gregor. 2000. Measuring Human-Computer Trust. In *Proceedings of the 11th Australasian Conference on Information Systems (ACIS '00)*. AAIS, Brisbane, Australia, 6–8.

[14] Hao Yi Ong, Daniel Freund, and Davide Crapis. 2021. Driver Positioning and Incentive Budgeting with an Escrow Mechanism for Ride-Sharing Platforms. *INFORMS Journal on Applied Analytics* 51, 5 (2021), 373–390.

[15] Christopher W. F. Parsonson, Alexandre Laterre, and Thomas D. Barrett. 2023. Reinforcement Learning for Branch-and-Bound Optimisation Using Retrospective Trajectories. In *Proceedings of the 37th AAAI Conference on Artificial Intelligence (AAAI '23)*. AAAI, Washington, DC, USA, 4061–4069.

[16] Varun Nagaraj Rao, Samantha Dalal, Eesha Agarwal, Dana Calacci, and Andrés Monroy-Hernández. 2025. Rideshare Transparency: Translating Gig Worker Insights on AI Platform Design to Policy. In *Proceedings of the 28th ACM Conference on Computer-Supported Cooperative Work and Social Computing (CSCW '25)*. ACM, Bergen, Norway, 1–49.

[17] Elizabeth Solberg, Magnhild Kaarstad, Maren H. Rø Eitrheim, Rossella Bisio, Kine Reegard, and Marten Bloch. 2022. A Conceptual Model of Trust, Perceived Risk, and Reliance on AI Decision Aids. *Group & publisher Management* 47, 2 (2022), 187–222.

[18] Weiran Yao and Sean Qian. 2020. Learning to Recommend Signal Plans under Incidents with Real-Time Traffic Prediction. *Transportation Research Record* 2674, 6 (2020), 45–59.

[19] Angie Zhang, Alexander Boltz, Chun-Wei Wang, and Min Kyung Lee. 2022. Algorithmic Management Reimagined For Workers and By Workers: Centering Worker Well-Being in Gig Work. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems (CHI '22)*. ACM, New Orleans, USA, 1–20.