DISPATCH ROUTING OPTIMIZATION
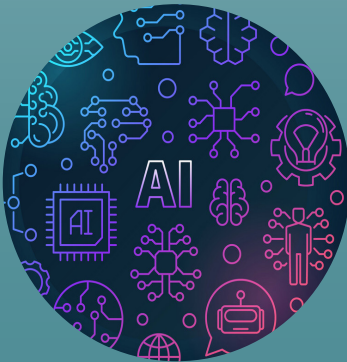## MOSAIC DATA SCIENCE | WHITE PAPER

MACHINE LEARNING

Mosaic data scientists collaborate with customers, digging deep into the data to inform design and deployment of custom ML tools that make a difference.

ARTIFICIAL INTELLIGENCE

Mosaic integrates powerful AI tools into clients' existing technology stack to solve complex business challenges.

BUSINESS ANALYTICS

Mosaic helps corporations of all shapes and sizes take advantage of their data, transforming their decision-making processes.

INTRODUCTION

The vehicle routing problem (VRP) is a combinatorial optimization and integer programming problem which asks "What is the optimal set of routes for a fleet of vehicles to traverse in order to deliver to a given set of customers?" It generalizes the popular traveling salesman problem.

It first appeared in a paper by George Dantzig and John Ramser in 1959, in which the first algorithmic approach was written and applied to petrol deliveries. The context is that of delivering goods located at a central depot to customers who have placed orders for such goods. The objective of the VRP algorithm is to minimize total route cost. ii
This case has gained popularity in the commercial market with plenty of companies needing to move plenty of product. Software providers and consultancies have jumped at the opportunity to design algorithmic approaches for minimizing transportation costs and maximizing profits. As technologies make it economical for companies to acquire and store data, organizations can integrate previously unknown variables such as weather, real-time asset locations, inventory levels, transactions, etc.

With this collection of variables there comes a need for data mining companies and data science consultants to analyze this information in real-time and provide insights and analytics to influence decisions.
For this specific white paper, Mosaic will focus on the use case of In Home Routing for a hypothetical big- box retailer. So, we will be looking at solving a piece of the vehicle routing problem, specifically for In Home service technicians fixing appliances purchased through said retailer.

PROBLEM STATEMENT

This big-box retailer has a successful in-home appliance business, where a customer buys an item from the brick and mortar store, and purchases the In Home repair services attached with the appliance all at the same time. The company plans to invest multiple millions of dollars into scheduling and routing software for the In Home repair technicians. This software is built around an algorithm and can be biased to produce profitable routing solutions given the correct inputs.

One of the issues this hypothetical retailer faces is not having identified the correct set of inputs or relative weighting of each in order to incorporate into the algorithm. Therefore, the solution has been configured to bias solutions towards either a reduction in transit time, reduction in overtime, or a reduction in overall headcount.
A way to test out the effectiveness of the routing / scheduling software would be to test it in two separate markets where the retailer operates. The two markets identified could be the Greater Boston area, consisting of 3 different states, Massachusetts, Rhode Island, and New Hampshire, and the Pacific Northwest consisting of 3 different states; Oregon, Washington, and Idaho.

Let's say for the sake of argument after months of testing it has been difficult to measure success due to conflicting metrics and completely different results in both markets. One market might actually see a measurable increase in transit times, while others might see only a nominal improvement. The retailer has made modifications to the core algorithm throughout the experiment and has been able to shift results. This causes other connected metrics to also shift. For example, the algorithm can be shifted to bias towards fewer miles driven; this decreases transit time, but may also increase overtime as techs that are closest may have to work extra hours. Focus on overtime control results in longer drive times as the system drives techs further based on their total workload for the week. There appears to be a decline in the completion rate in both markets, which decreases customer satisfaction, while increasing overall costs and reducing available capacity for future customers.

In order to properly quantify the solution and identify the configuration of algorithm parameters, these competing metrics must be dealt with in terms of profitability.

This retailer now needs to develop a profitability model for their routing / scheduling software. With a wealth of analytics consulting companies to choose from, the retailer can either figure out this model in-house or contract the work out.

The profitability model needs to provide:

1. The data drivers (and relative weightings) that must be added to the routing optimization algorithm to produce the most profitable results
2. A methodology for measuring route quality in terms of profitability that can then be used to evaluate current routing software and potential replacements

The following criteria must be met:

1. Clear identification of factors that influence profitability of the retailers' solution
2. A model that can be implemented to provide appropriate weightings of the aforementioned factors when considering:

   • *Which technicians should be assigned to specific service events*
   • *Which possible appointments (dates/time windows) should be offered to end customers*

3. A dashboard (or equivalent reporting) to measure the expected profitability of planned routes
4. A methodology established for evolving the model as needs change

VARIABLES TO CONSIDER

Before we get to the analytic approach let's define Likely Candidate Factors, as they play a large role. The 'likely candidate factors' are a brief overview of potential factors that can influence the profitability of a set of routes.

• Technician-specific factors
  i. Each technician has their own hourly labor rate, the majority start their day and end their day at home, resulting in thousands of unique start/end locations.
  ii. Some techs work more efficiently than others.

• Customer-specific factors
  i. Different product plans offer different Service Level Agreements
     i.[a] A Business-to-Business customer still in warranty period is different than a Direct-to-Consumer with an expired warranty
  ii. Integrate with Customer Lifetime Value Model if available

• Event-specific factors
  i. Certain product types might have a different margin other product types
  ii. Duration of the visit
  iii. Repair types can by vary product type and likelihood the tech will have the skills/parts to complete the job in a single trip

HOW IS AN ANALYTICS CONSULTING FIRM SUPPOSED TO SOLVE THIS?

In the following sections, Mosaic will attempt to lay out a framework of predictive analysis to help construct the profitability model.

Mosaic typically begins these engagements with an onsite kickoff meeting to meet stakeholders, become familiar with data sources and systems, clearly define the problem, and learn the applicable business processes.
For the above use case, Mosaic would want to accomplish the following objectives:

1. Develop a profitability framework for in-home scheduling
   i. Cost drivers
   ii. Revenue sources
   iii. Service level effects
   iv. Drivers of uncertainty
   v. Profitability metrics

2. Deep-dive on current and candidate scheduling tools
   i. Scope and fidelity
   ii. Interfaces
   iii. Inputs and sources
   iv. Outputs and consumers
   v. Analysis of schedules
   vi. Process for running software
   vii. Related business processes

3. Deep-dive on requirements for profitability dashboard/reports
   i. Users
   ii. Use cases
   iii. Information requirements
   iv. Interactivity requirements
   v. Integration requirements
   vi. Technical requirements (software, hardware, accessibility, etc.)

4. Initial data exploration to determine data sources, assess data quality, and identify gaps

5. Establish processes for access to relevant data and the proposed solution software

PHASE 1: ACTUAL PROFITABILITY MODEL

This phase of the project begins with an exploratory analysis of the data provided by the retailer. Based on the profitability model framework, objectives and metrics agreed to during the onsite kickoff, Mosaic's data scientists would begin with an analysis of the drivers of profitability with the objective of developing a model that can estimate the profitability of a fixed output of the routing optimization software after the fact – once all technician activities for a given time period have been observed and recorded.

Then we can analyze the effects of key cost, revenue, and uncertainty driver on schedule profitability. Correlated effects between pairs or small groups of factors will be identified. Factors covered by the analysis will include technician factors, customer factors, and event-specific factors.

PHASE 2: PREDICTIVE PROFITABILITY MODEL

Based on the insights from Phase 1, Mosaic would develop a predictive model that would generate an expected profitability, including associated measures of uncertainty, from a candidate routing solution. This model may use statistical, simulation, machine learning, optimization, or any other relevant analytical techniques with the goal of maximizing precision and accuracy of the profitability forecasts. The post-analysis profitability model developed in the first phase of the project would be used to benchmark model accuracy. The model development process during this phase would be iterative, incorporating insights from previous model versions to evolve modeling approaches and input variables in order to steadily improve accuracy metrics.

A primary focus during this phase of the project would be on the impact of factors that drive uncertainty. Uncertainty would be driven by factors that influence a given technician's ability to complete a set of planned maintenance visits as scheduled. These factors would likely include many of the factors that are part of the phase 1 profitability model – technician factors, customer factors, and event factors.

PHASE 3: SCHEDULER OPTIMIZATION MODEL

The final phase of model development would translate insights from the predictive profitability model into a model for determining the configuration of inputs to the scheduling software that maximize expected profit in the resulting schedule. The inputs/parameters controlled by the optimization model would be selected from the full list of potential inputs to the scheduling tool. The inputs/parameters would be prioritized based on insights from the previous model development exercises and from exploratory analysis focused on determining how inputs to the scheduling tool drive the profitability of the output schedules. The predictive profitability model developed in the previous phase would be used to evaluate the profitability of results returned by the vendor solution and to evaluate the performance of the optimization model.

Specific optimization techniques would be selected based on effectiveness but could include deterministic optimization models, simulation-based models, or stochastic search models. If feasible, the optimization model may leverage multiple runs of the solution tool with selected inputs to dynamically guide the search for optimal inputs. The model would be designed such that it can adjust to changes in the profitability models – e.g., changes in the relative weighting between direct profit from maintenance activities and inferred profitability of improved service levels.

Based on stakeholder requirements, the model could be designed to output a single optimal schedule from the scheduling tool based on a pre-selected profitability objective function or designed to generate a small number of alternate schedules covering a range of profitability objectives. In the latter case, the multiple options would be presented to a human decision maker (dispatcher or scheduler) to make the final selection of the schedule that best meets the current, potentially dynamic objectives.

PHASE 4: DASHBOARD/REPORT DEVELOPMENT

The fourth phase of the project would focus on dashboard development and could be initiated in parallel with the other project phases. Initial requirements would be established during the kickoff. The dashboard could include a summary of the profitability metrics used during the optimization, the inputs generated by the optimization model, and a summary of the expected profitability of the schedule generated from the optimized inputs. Additional features could include:

- Comparison of expected profitability between the optimal schedule and schedules generated from alternate input configurations and parameterizations
- Reporting of actual profitability of previously executed schedules (e.g., from the previous day)
- Drill-down interactivity allowing users to explore recommended schedules
- Integrations with business process tools allowing users to initiate distribution of a selected schedule directly from the dashboard

Requirements and templates would be updated throughout the model development process to ensure that relevant information from the models are incorporated appropriately. For example, the dashboard design will need to account for whether the optimization model returns a single input set and associated schedule or a small number of candidates from which the user can select.

EXECUTION

Mosaic would follow and does follow the CRISP-DM process for analytics projects. This flexible process framework emphasizes the iterative nature of analytics projects and the firm rooting of all analytics activities in a deep understanding of business objectives and constraints. All models and dashboards/reporting functionality would be thoroughly tested in order to ensure robustness, reliability, and full alignment with established business objectives and constraints.

CONCLUSION

By utilizing predictive analysis techniques, analytics firms such as Mosaic can help solve the problem of quantifying the impact of implementing a new scheduling tool prior to actual implementation. Once this model is in place the retailer can begin to look at defining their routing areas, integrating with a preexisting customer lifetime value model, and further improving the routing engine, helping realize return on investment.

[I] Vehicle routing problem. (2016, September 26). In Wikipedia, The Free Encyclopedia. Retrieved 16:53, September 26, 2016, from https://en.wikipedia.org/w/index.php?title=Vehicle_routing_problem&oldid=741301514

[II] Vehicle routing problem. (2016, September 26). In Wikipedia, The Free Encyclopedia. Retrieved 16:53, September 26, 2016, from https://en.wikipedia.org/w/index.php?title=Vehicle_routing_problem&oldid=74130151

FOR MORE INFORMATION

Want to learn more?
Please contact info@mosaicdatascience.com

mosaic
D A T A   S C I E N C E

ABOUT MOSAIC DATA SCIENCE

We provide innovative machine learning, AI and analytics consulting across organizations.

Mosaic is a leading data science consulting company focused on helping organizations build and deploy actionable analytics solutions. Our customers are as varied as the techniques we use — some just starting their first predictive analytics project; others with deep in-house machine learning expertise.

HOW WE WORK WITH YOU

We work in a highly collaborative partnership with our customers to ensure you get only the best results to consistently drive business value.

MACHINE LEARNING
We design and deploy predictive algorithms to solve the most challenging problems facing businesses today

ARTIFICIAL INTELLIGENCE
We bring a wealth of knowledge on how to tune AI models to deliver the maximum business value

BUSINESS ANALYTICS
We leverage technical expertise and experience across a swath of industries, bringing fresh approaches to challenging problems