East Daley Capital
Using Mixed Integer Programming (MIP) to Model Midstream Energy Assets

Introduction

Many organizations use optimization to assist with upcoming decisions. However, optimization has more applications than just large-scale decision making. One such example is East Daley Capital (EDC) and its use of optimization to model complex historical datasets. EDC provides in-depth data and information about North American midstream energy assets - processing, storing, transporting and marketing of oil, natural gas and natural gas liquids.

With data from dozens of different sources available at different frequencies and qualities, Mixed Integer Programming (MIP) is a useful framework for filling gaps in the data. EDC's key thrust behind using a MIP model is to estimate the interconnectivity of wells, pipelines, and processing plants as these data are only known to the companies that own them. East Daley uses Gurobi to accelerate the development and execution of their internal models.

The Problem – Missing information

Complex systems often behave in predictable ways. However, when confronted with a mix of known and unknown information, it is often difficult to fill in the missing pieces. East Daley Capital gathers copious amounts of data on natural gas pipelines and the ecosystem of natural gas in order to understand the impact or potential effects on businesses.
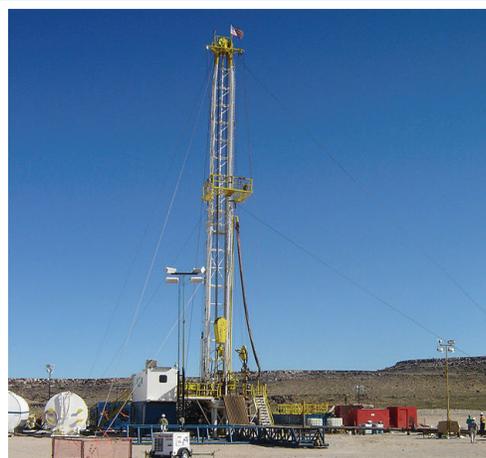
For example, to correctly value gas gathering and processing systems in Oklahoma's Anadarko Basin, an accurate throughput forecast is needed. In a perfect world, the specific gas wells connected to the system would be known and one could then forecast the system by aggregating forecasts for the individual wells. However, there is little to no public information detailing the wells that connect to each processing system.

Even though the exact information is not available, that doesn't mean there is no information. The company that operates the Gathering and Processing (G&P) system discloses stylized maps of the system, the Energy Information Administration (EIA) provides annual throughputs for gas processing plants, and the Oklahoma Tax Commission provides well production information.

This presents a litany of problems:

1. Data sets with a high number of mostly sparse dimensions. This is a problem because the information is sparse and it makes it difficult to use a rule based approach to solve the problem.

2. Dozens of data sources. Since the number and quality of data is always changing, it would be very challenging to maintain a "rule based" approach with so many data sets

3. Various rules for filling gaps.

In order to address these problems, East Daley could have chosen to refocus their efforts elsewhere or do consulting with limited scope. However, EDC decided to find a way to make the data more holistic and therefore, far more insightful.
.

The MIP Model

East Daley developed a mixed integer programming (MIP) model to fill in missing data with:

- >500K binary variables
- >1M constraints

The model considered the following information:

- Known data: Historical output of each well, historical throughput of some processing plants, and the partition of pipeline networks.

- Rules of thumb: When selecting pipelines to connect to wells, there are several factors to consider: 1. the possibility of sending gas to a nearby processing plant, 2. the fact that some companies are partners and others are competitors, and 3. the possibility of reusing existing gathering pipelines.

- Hunches: Impact of past services disruptions, such as known pipeline disruption, leads to a reduction in plant throughput of a specific plant.

East Daley Capital
Using Mixed Integer Programming (MIP) to Model Midstream Energy Assets

There is a hierarchy for rules of thumb and another one for hunches with the rules of them carrying more weight.

Therefore, given the historical data of wells, processing plants, pipeline networks, and the hierarchies of rules of thumb and hunches, the MIP model identifies which wells are "more likely" to be connected to pipelines that ship the volumes processed at plants, while satisfying the rules of thumb and hunches.

In this MIP model, everything is a soft constraint - even "known" data may not be 100% accurate. There is a hierarchy of assumptions that leads to a multi-objective model. For example, regulatory filings are more important than hunches.

The primary output of the MIP model is a partitioning of the wells, pipelines, and plants into systems. Results were then rolled up into higher-level models, such as a forecasting model and a macro-economic energy model.

Implementation

The MIP model was built by Levi DeLissa, Ryan Smith and Justin Carlson using the Gurobi C API. When solving the model with the Gurobi Optimizer, several Gurobi technology features were heavily used, including:

- Variable hints: Results from last run used as hints for the next one
- Multiple objectives: Very useful for handling soft constraints

The model ran for 24+ hours, yet good solutions were typically available within 4 to 12 hours.

Benefits

The MIP Model:

- Provides a holistic and objective view of the problem and consequently helps decision makers remove bias, since there is no need to argue about facts.
- Produces extremely useful insights for predicting total production, prices and fortunes of individual companies.
- Provides a simple framework for adding more data and rules into the model. When new facts become available it is generally straightforward to use this information and implement a new class of constraints.

> " In my opinion the ability to add new facts as they become available is one of the most beneficial aspects of this model. As we have collected more data over time, I have yet to encounter a dataset that couldn't be turned into a set of constraints.
>
> **Levi DeLissa,**
> **Data Integration & Optimization Engineer**

Why Gurobi?

East Daley chose Gurobi because of Gurobi's performance. Gurobi solved the problem faster than other solvers. EDC also found Gurobi support to be very responsive. Although EDC originally built its model using the Gurobi C API, they ultimately used the Python API because it provided much higher productivity. In particular, the overall flexibility in handling multiple objectives, allowing for variable hints, and generalized constraints is very helpful.

> " The features of the Gurobi Python API save me the headaches of trying to implement a multi-stage model myself.
>
> **Levi DeLissa,**
> **Data Integration & Optimization Engineer**

About East Daley Capital

East Daley Capital was founded in 2014. It has 28 employees and is in possession of the world's largest North American midstream asset level database. East Daley sells valuable information to investors, such as hedge funds, private equity firms, etc. East Daley studies many different energy markets. EDC brings greater transparency to the energy financial market by arming sophisticated investors with deeper, more accurate data that empowers their investment decisions.

For more information, visit Gurobi.com or contact us at info@gurobi.com or call +1 713-871-9341.

2