Mathematical Optimization for Machine Learning and Decision

Yinyu Ye - October 18, 2022 – FOS, HKU

Mathematical Optimization

- Often consider the common analytic decision model of Machine Learning and Data/Decision/Management Science & Engineering:

  - Maximize or Minimize *f*(***x***)
    for all ***x*** ϵ some set X

- Decision variables represented by a vector/matrix ***x***,
- *f*(***x***): Objective function
- X: Constraint set formed by equations and inequalities
- If all functions are linear: Linear Programming

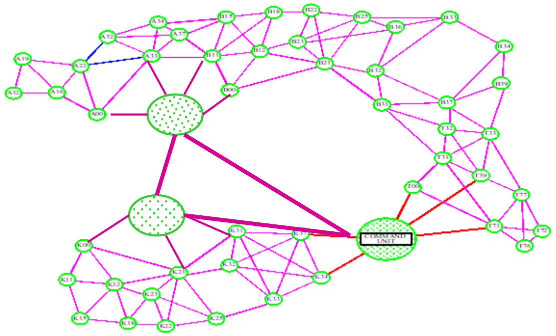
LP Giants won Nobel Prize…

Topic 1: Sensor Network Localization and Tracking via Semidefinite Optimization

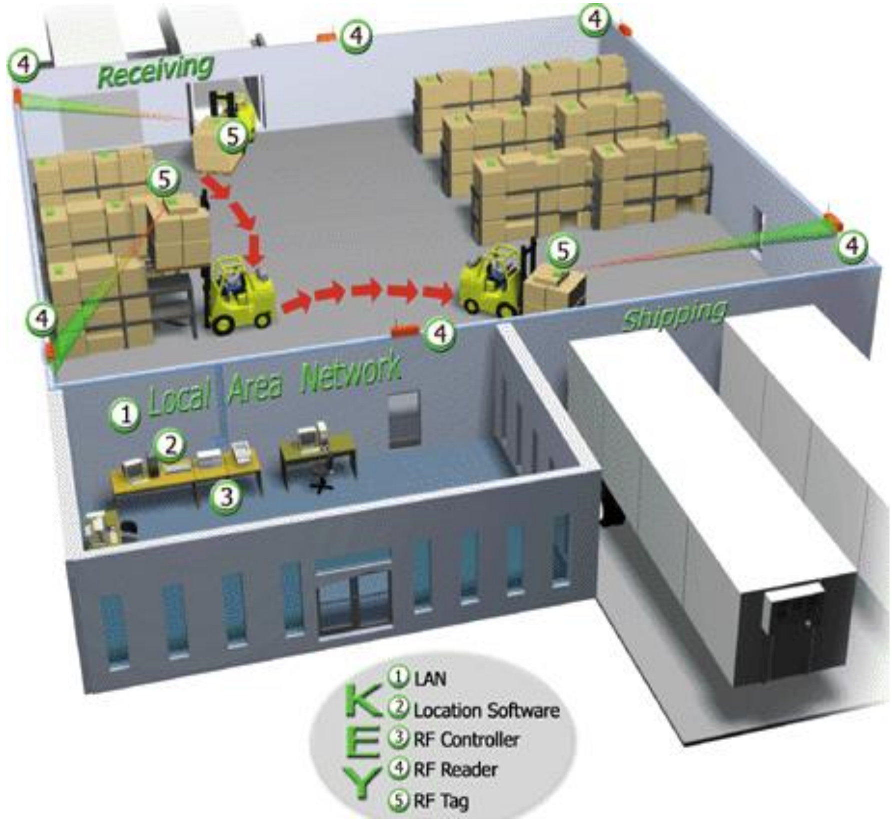
Ad Hoc Sensor Network Localization
Biswas, So, Y … 2005

- Localization
  - Given partial pair-wise measured distance values
  - Given some anchors' positions
  - Find locations of all other points that fit the measured distance values

This is also called graph realization on a fixed dimension Euclidean space



Good/Cargo RFID/Location in Supply Chain



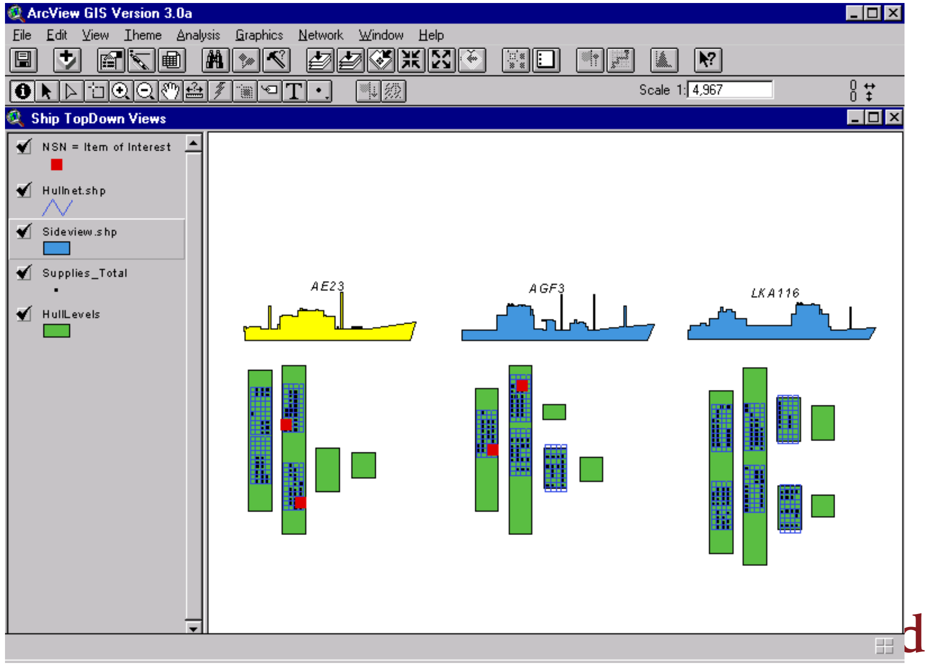
Accurate location:

Peer-to-peer Self-forming Mesh Network

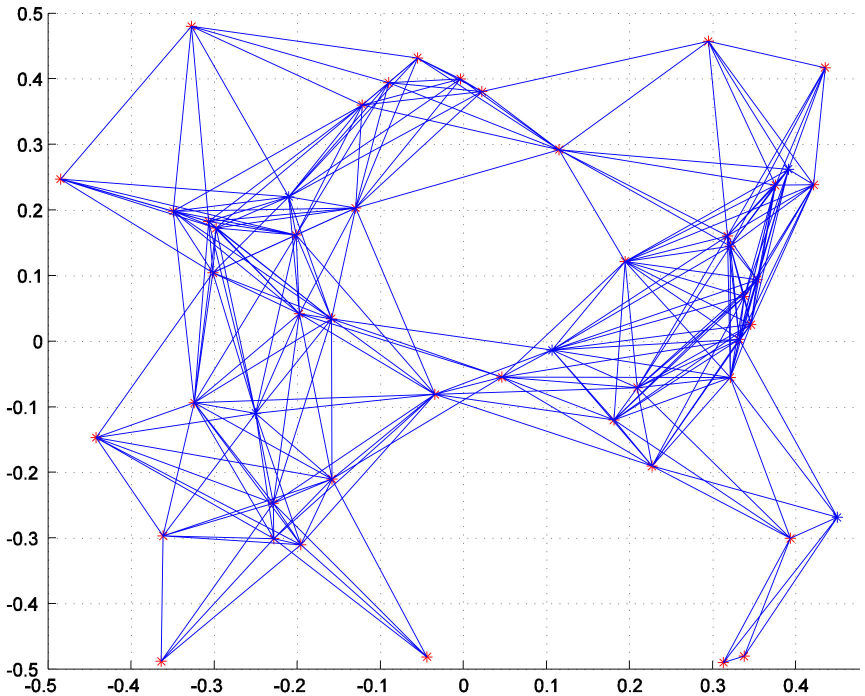
- No fixed infrastructure
- Communications using RF

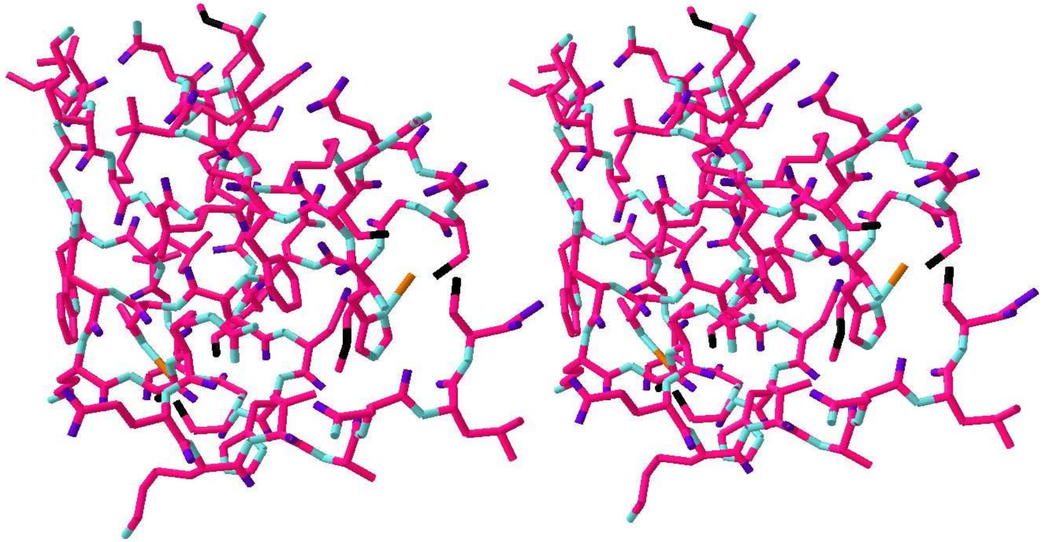
Sensor Integration

- Condition/History information
- Tamper Resistant Seals



A Unit Disk Sensor Network





System of Quadratic Equations or Quartic Least Squares

The problem can be formulated as follows:

$$\left\|x_i - x_j\right\|^2 = d_{ij}^2 \qquad (i, j) \in E_{ss}$$

$$\left\|a_k - x_j\right\|^2 = \bar{d}_{kj}^2 \qquad (k, j) \in E_{sa}$$

$$x_i \in R^d$$

- $\{\boldsymbol{a}_k\}$ are the positions of “anchors”.

Does it have a solution?

Is the solution unique?

In practice; one add objective to minimize noise errors

The above system is non-convex and generally intractable. To get something more tractable, we can consider a convex relaxation.

Semidefinite Programming Relaxation

<u>Step 1</u>: Linearization

$$\left\|x_i - x_j\right\|^2 = x_i^T x_i - 2x_i^T x_j + x_j^T x_j$$

$$\left\|a_k - x_j\right\|^2 = a_k^T a_k - 2a_k^T x_j + x_j^T x_j$$

Semidefinite Programming Relaxation

Step 1: Linearization

$$\|x_i - x_j\|^2 = \underbrace{x_i^T x_i}_{Y_{ii}} - 2\underbrace{x_i^T x_j}_{Y_{ij}} + \underbrace{x_j^T x_j}_{Y_{jj}}$$

$$\|a_k - x_j\|^2 = a_k^T a_k - 2a_k^T x_j + \underbrace{x_j^T x_j}_{Y_{jj}}$$

Tighten: $Y = X^T X$, $X=[x_1,\ldots,x_n]$

Step 2: Relax $\quad Y \geq X^T X \Leftrightarrow Z = \begin{bmatrix} I & X \\ X^T & Y \end{bmatrix} \geq PSD$

This is a conic linear program which is a convex optimization problem!

Biswas and Y 2004, So and Y 2005

The Biswas-Ye SDP Relaxation

More precisely, Use linear algebra trick and put things together:

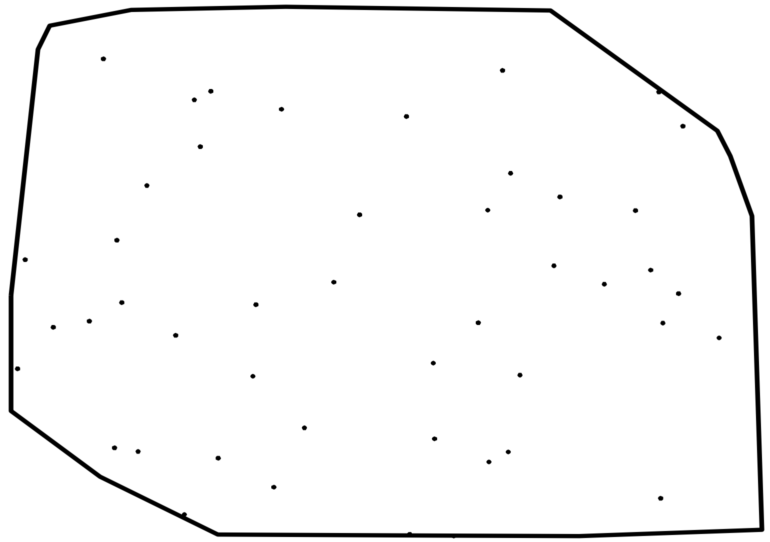
$$(0; e_i - e_j)(0; e_i - e_j)^T \bullet Z = d_{ij}^2 \qquad (i, j) \in E_{ss}$$
$$(a_k; -e_j)(a_k; -e_j)^T \bullet Z = \bar{d}_{kj}^2 \qquad (k, j) \in E_{sa}$$
$$Z \geq 0;\ Z_{1:d,1:d} = I_d$$

This is an instance of semidefinite programming/optimization (SDP), which can be solved (to any arbitrary accuracy) in polynomial time, were decision variables form a symmetric matrix and it has to be positive-semidefinite. This was a breakthrough optimization model, together with its interior-point methods, developed in early 90s, which led to the 2009 von Neumann Theory Price.

An important and interesting question is: when is the relaxation *exact*? SNL Localization Demo

Topic 2: Region Covering and Partition – Divide and Conquer (Carlsson et al. 2009)

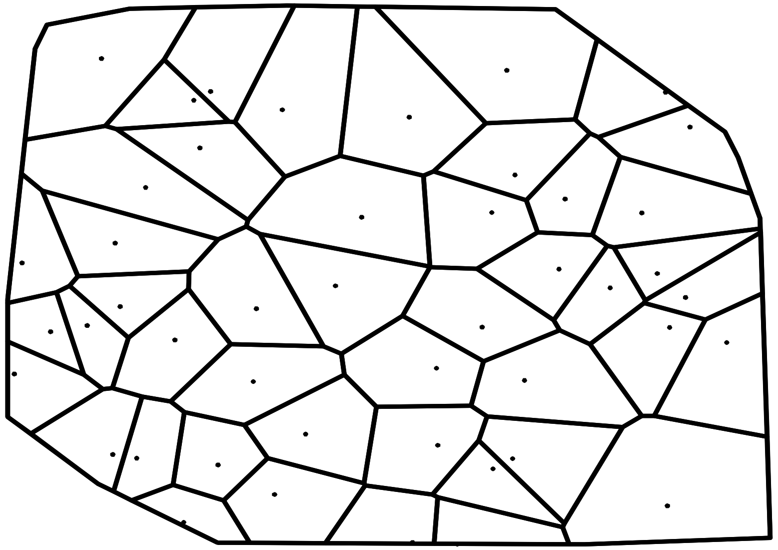
Problem Statement: Divide-Conquer



n points are scattered inside a convex polygon *P* (in 2D) with *m* vertices. Does there exist a partition of *P* into *n* sub-regions satisfying the following:

- Each sub-region is a convex polygon
- Each sub-region contains one point
- All sub-regions have equal area

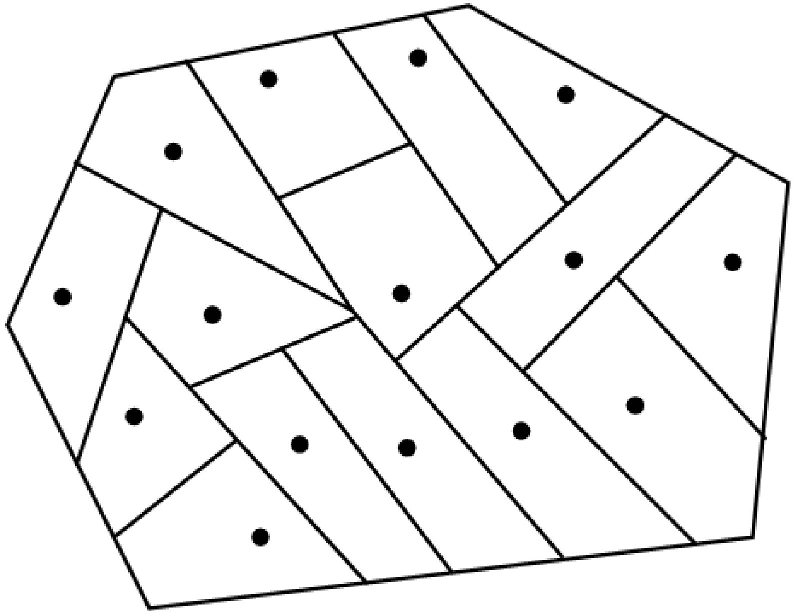
Related ML Problem: Voronoi Diagram

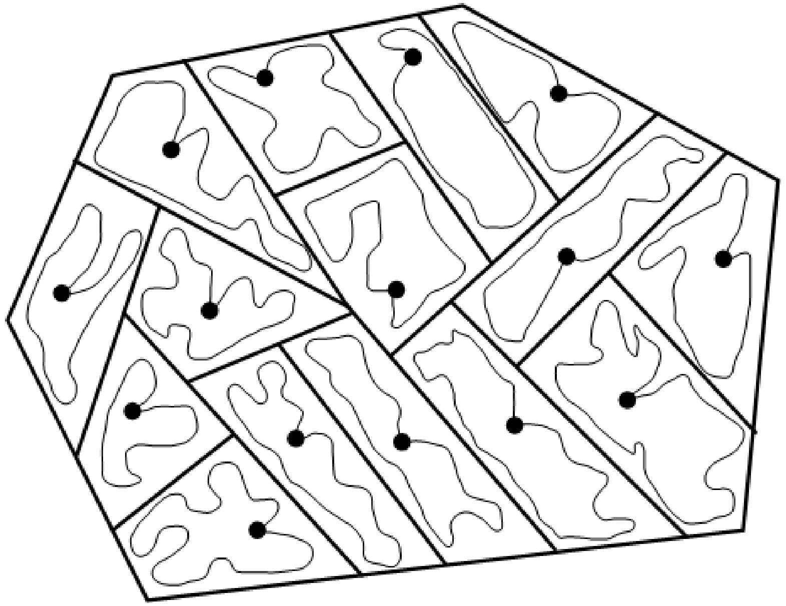


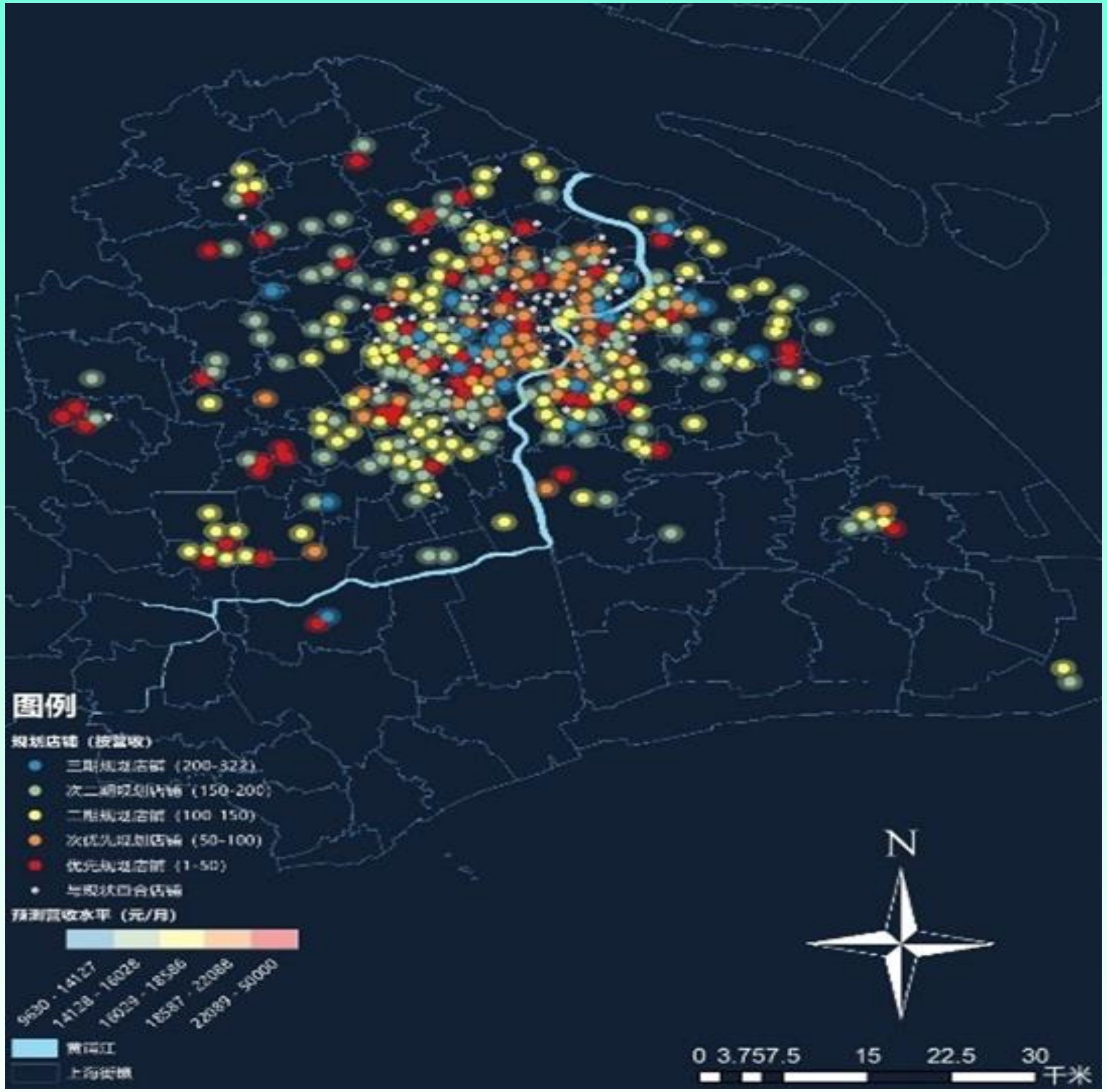
In the *Voronoi Diagram*, we satisfy the first two properties (each sub-region is convex and contains one point), but the sub-regions have different areas.

Our Result

Not only such an equitable partition always exists, but also we can find it exactly in running time $O(Nn \log N)$, where $N = m + n$.

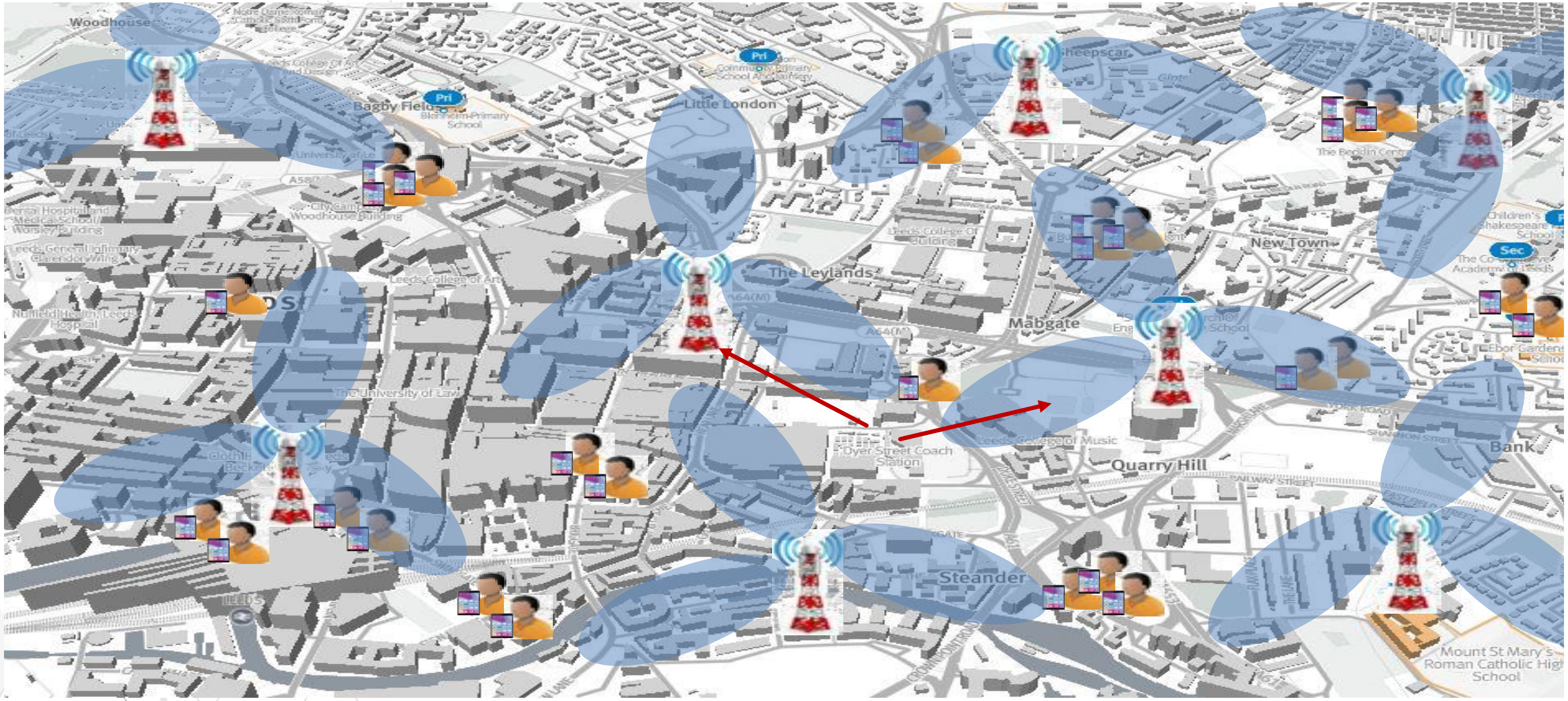






Communication APs of 5G

Wireless Tower: Adaptive Resource Allocation

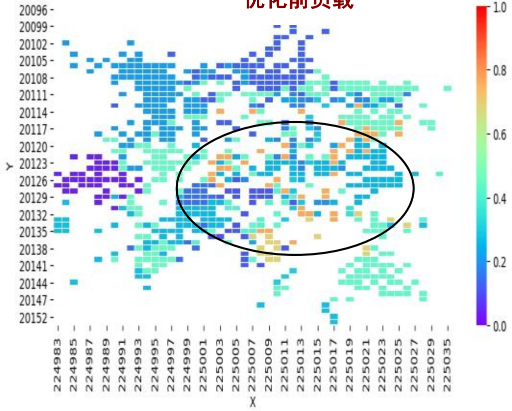


Test Result—Effectiveness

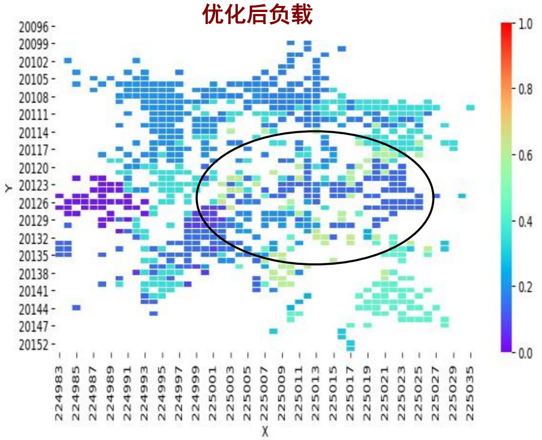
基于真实商用网络进行模型优化效果的测试验证验证统计结果：

	小区数	时段	区域平均负载	区域平均吞吐率(Mb/S)	高负载小区负载	高负载小区吞吐率(Mb/S)
优化前	27	中午及晚共6小时	31%	5.3	68%	2.3
优化后			30%	6.12(提升15%)	66%	2.8(提升22%)
优化前		晚7时话务高峰	37%	3.9	77%	1.6
优化后			33%	5.2(提升33%)	68%	2.1(提升32%)

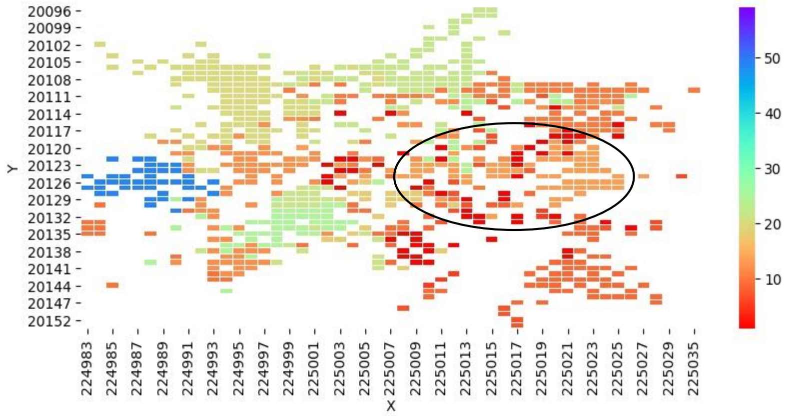
优化前负载



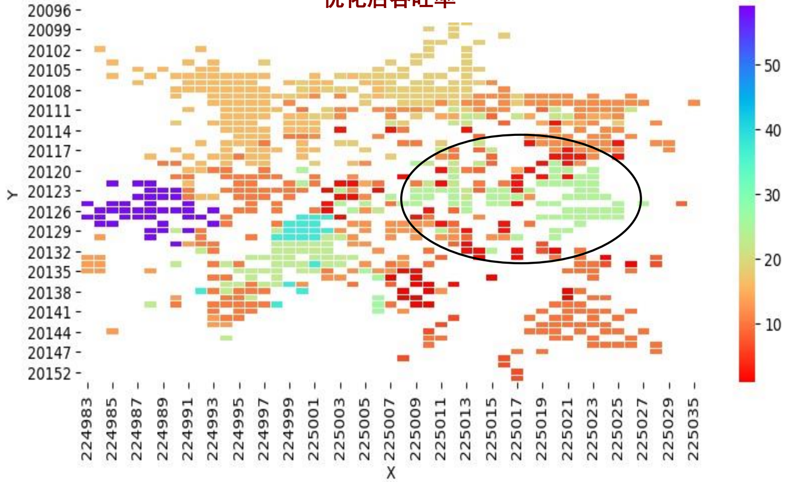
优化后负载



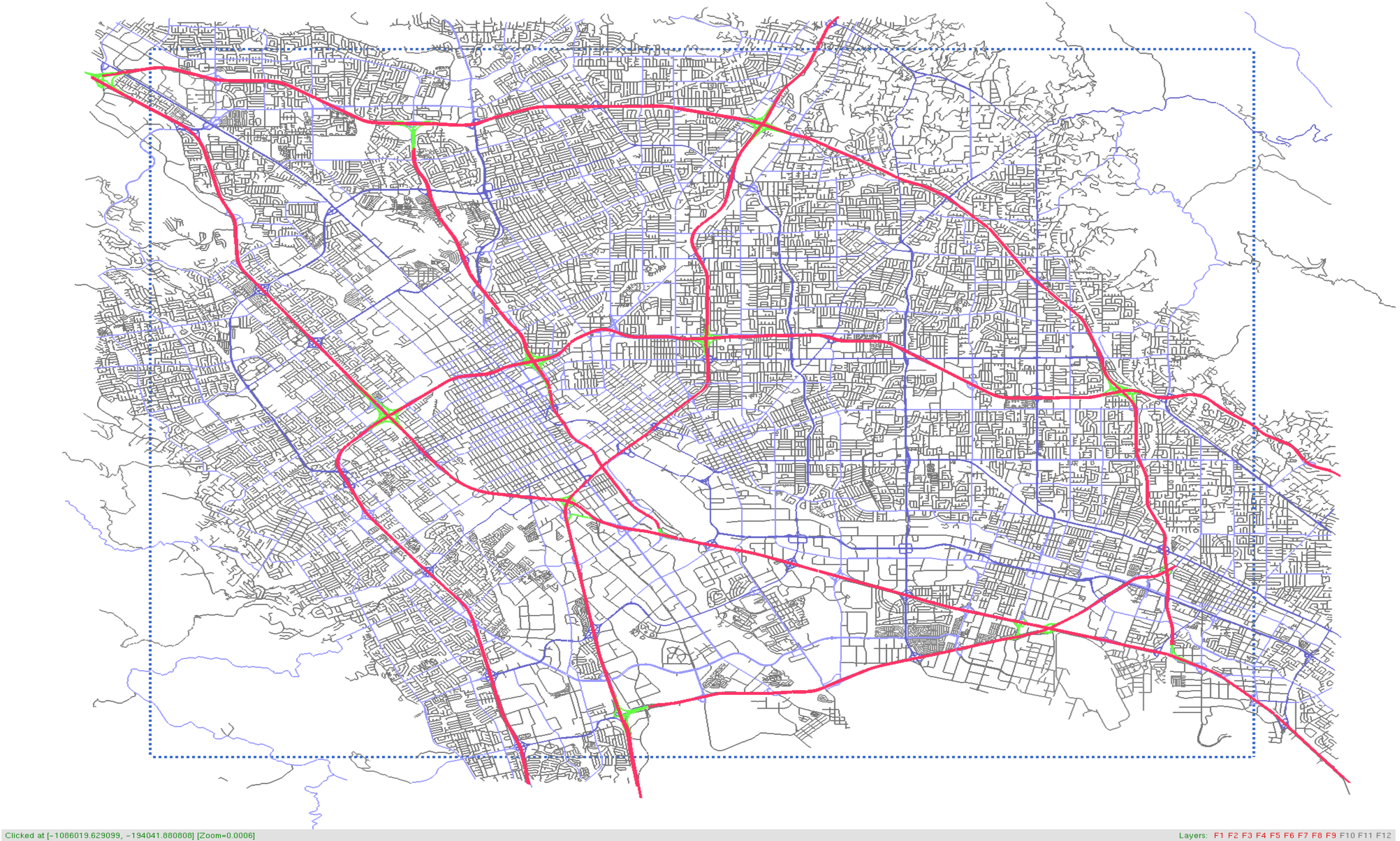
优化前吞吐率



优化后吞吐率

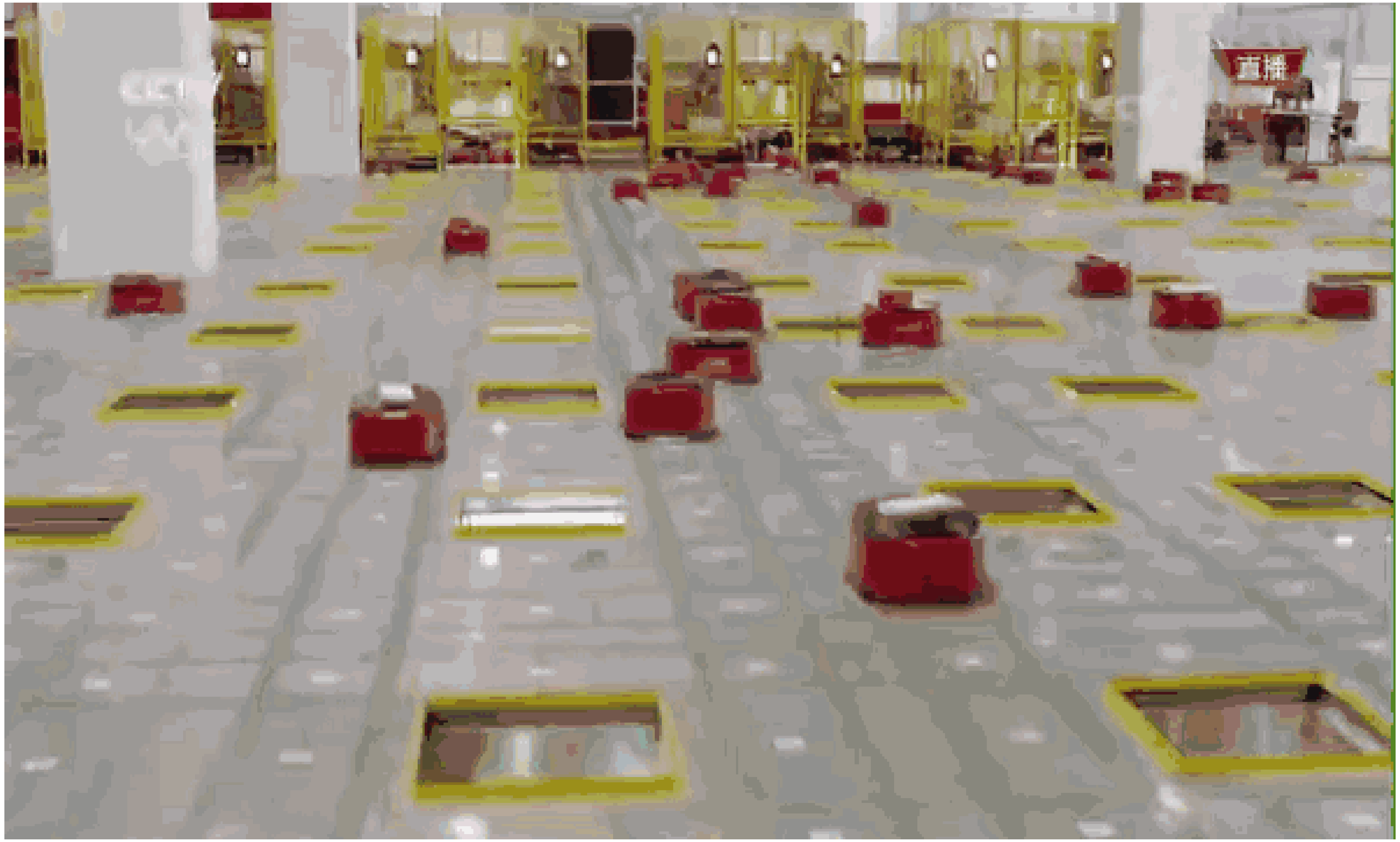


Application: Street View Application Map-Making



Application: Unmanned Warehouse

Real-Time Unmanned Vehicles Assignments and Routing in a Warehouse (Cardinal Operations)



The throughput is **3-4 more productive than managed by men.**

Topic 3: Energy Management System and PEV
Optimal Charging/Discharging (Nicole Taheri et al. 2011)

Plug-In Electric Vehicle Network

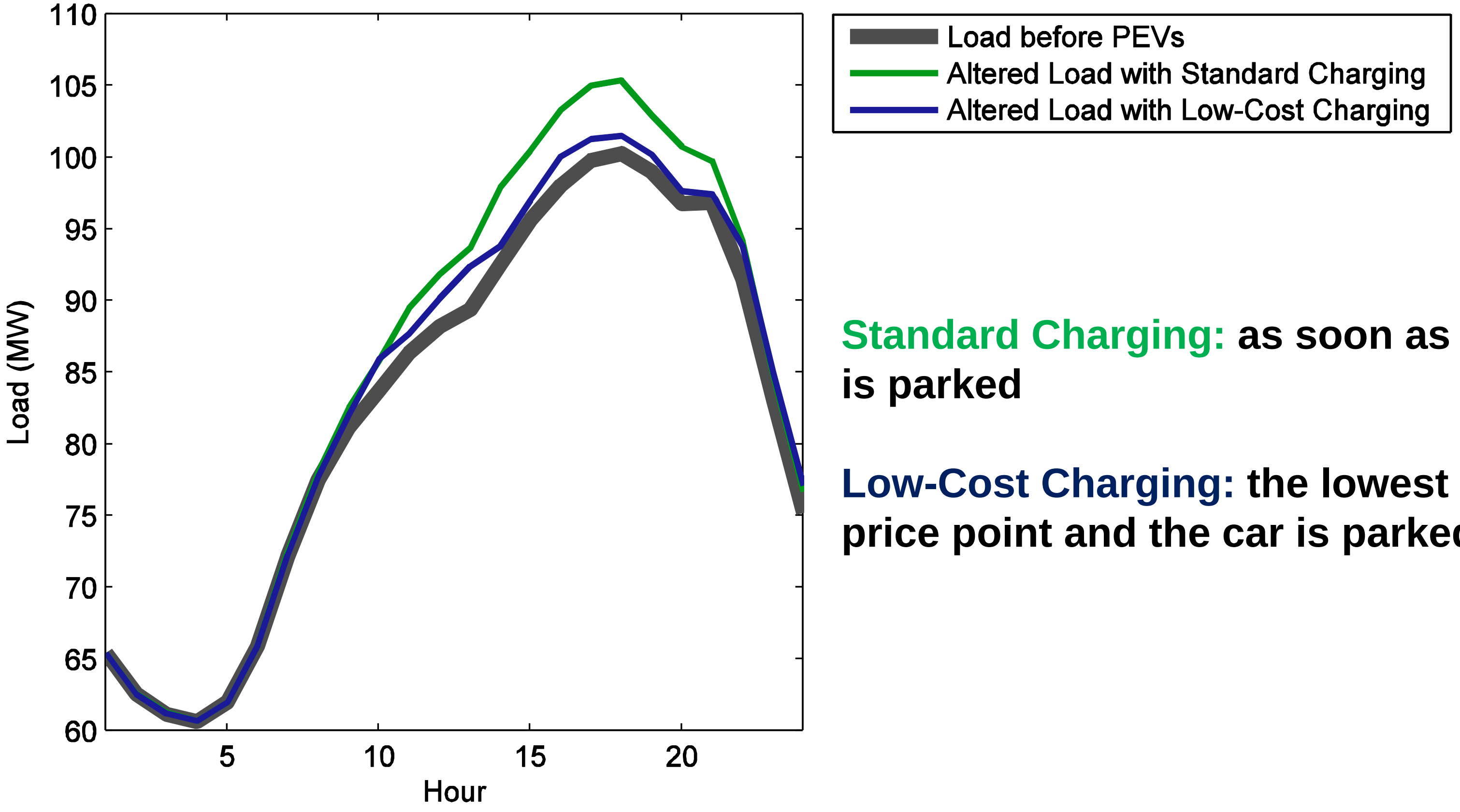
- Some estimates say there could be 100 million Plug-in Electric Vehicles (PEVs) on the road in the United States by 2030[1]
  - How will charging/discharging of PEVs add to the current load on the electricity grid?
  - Would smart management of these activities benefit both utility and consumer sides?

Motivated by these questions, we

- Construct a robust algorithm to dynamically assign low-cost, feasible, and satisfactory charging/discharging schedules for individual vehicles in a fleet
- Reduce the typical consumer cost of charging/discharging a PEV
- Lower the peak demand for electricity and benefit utility supplier to provide grid services

[1]EPRI PRISM Analysis, 2009

PEV Impact to the Electricity Grid



Standard Charging: as soon as the car is parked

Low-Cost Charging: the lowest PEG price point and the car is parked

A 30% penetration of Plug-in Electric Vehicles could impact the electricity grid.

Specific Problem Statement

The goal is to dynamically manage the charging/discharging of a fleet of PEVs so that:

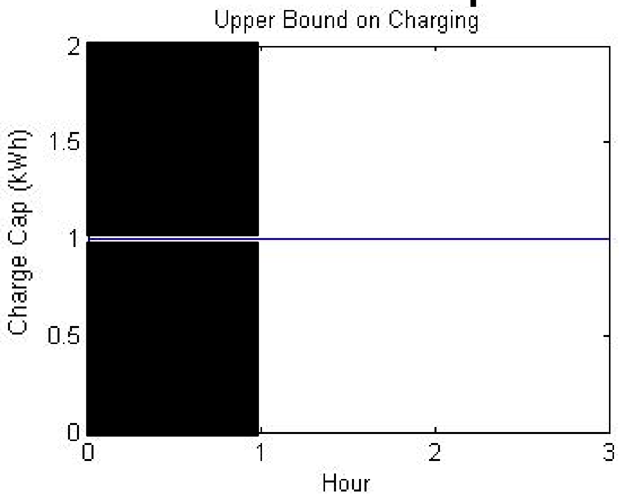
1. Every vehicle has enough energy in its battery to drive for a given period of time
2. The cost of charging is low
3. The peak electricity load does not increase and may even be reduced
4. The schedules are dynamic and robust to deal with uncertainty

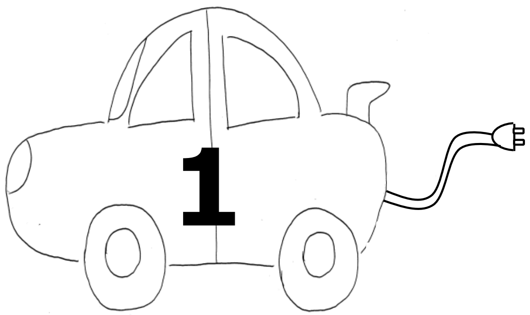
Using a linear program solution, one can make policy decisions about when to charge/discharge of every individual vehicle in a fleet based on:

- Energy demand / time of each vehicle in a period
- Electricity load capacity and scheduling obligation
- Publicly available electricity and gasoline prices
- Individual vehicle characteristics / types

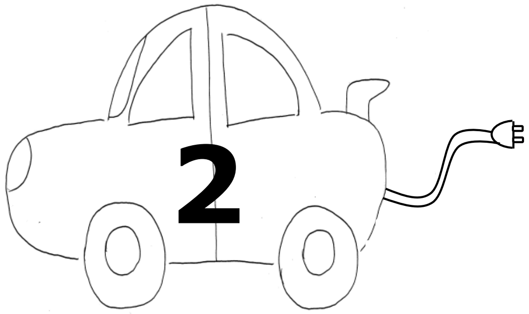
A Toy Example

There are two cars considered in a three hour period.





I need to charge1 kWh to drive in hour 3!

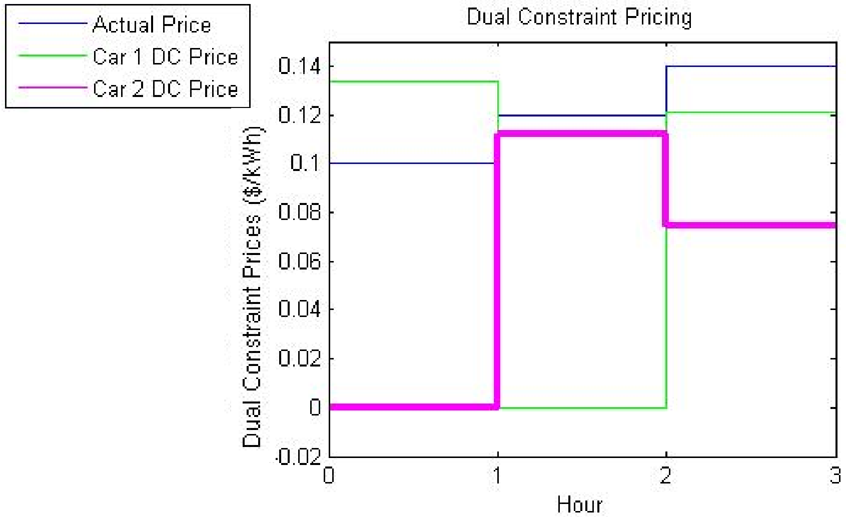


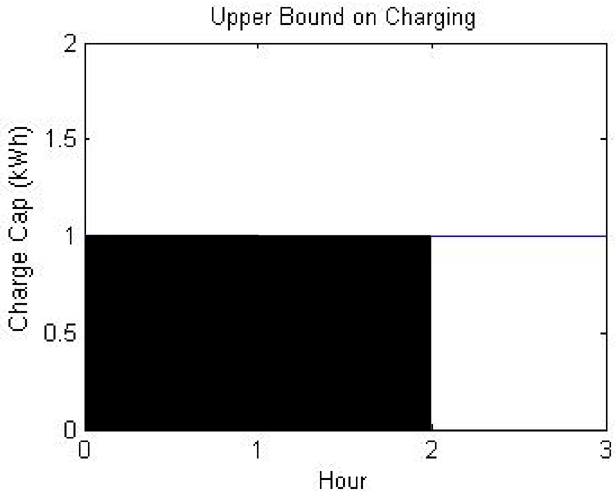
I need to charge 1 kWh to drive in hour 2!

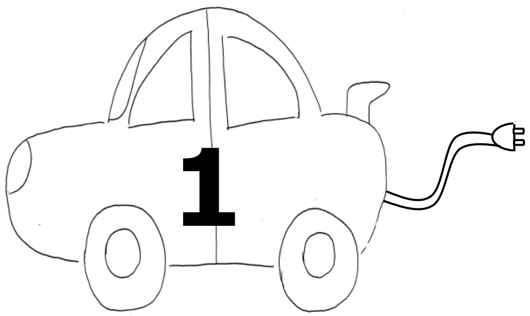
The problem can be tackled by **Linear Programming**

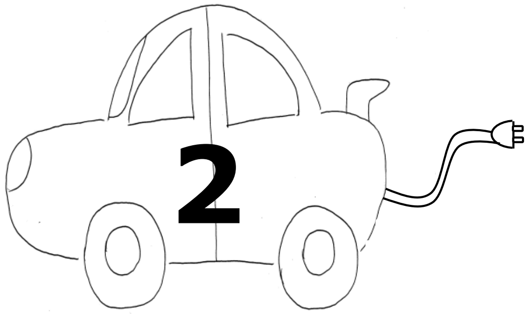
Customized Shadow Prices in the Example

There are two cars considered in a three hour period.









I need to charge1 kWh to drive in hour 3! **I need to charge 1 kWh to drive in hour 2!**

Test Based on Real Data

Vehicle Driving Behaviors

- Obtained from the 2009 National Household Transportation Survey (NHTS)
- Helpful discussions and filtered data from Morgan and Christine of EPRI
- The following results are based on data from urban California on a Monday

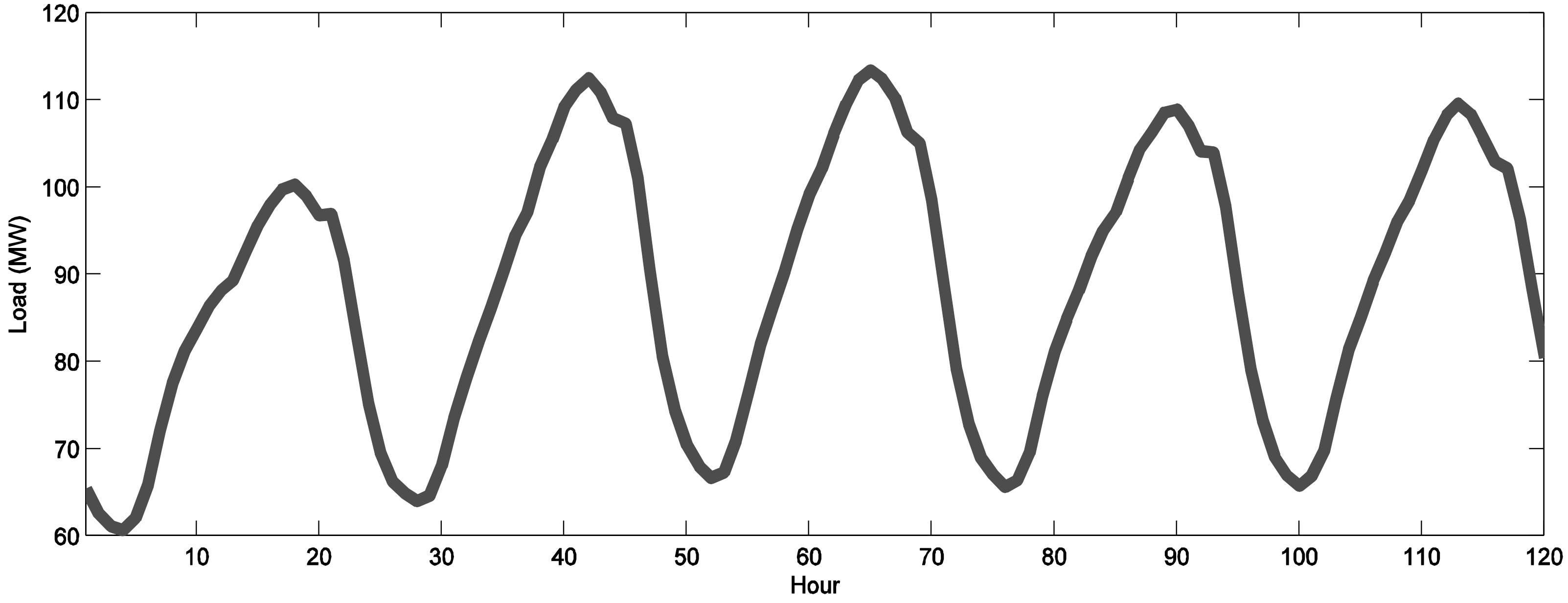
Electricity Load

- Obtained from CAISO OASIS (Open Access Same-Time Information System)
- Used the demand in the PG&E transmission access charge area for the week of August 22-28, 2011

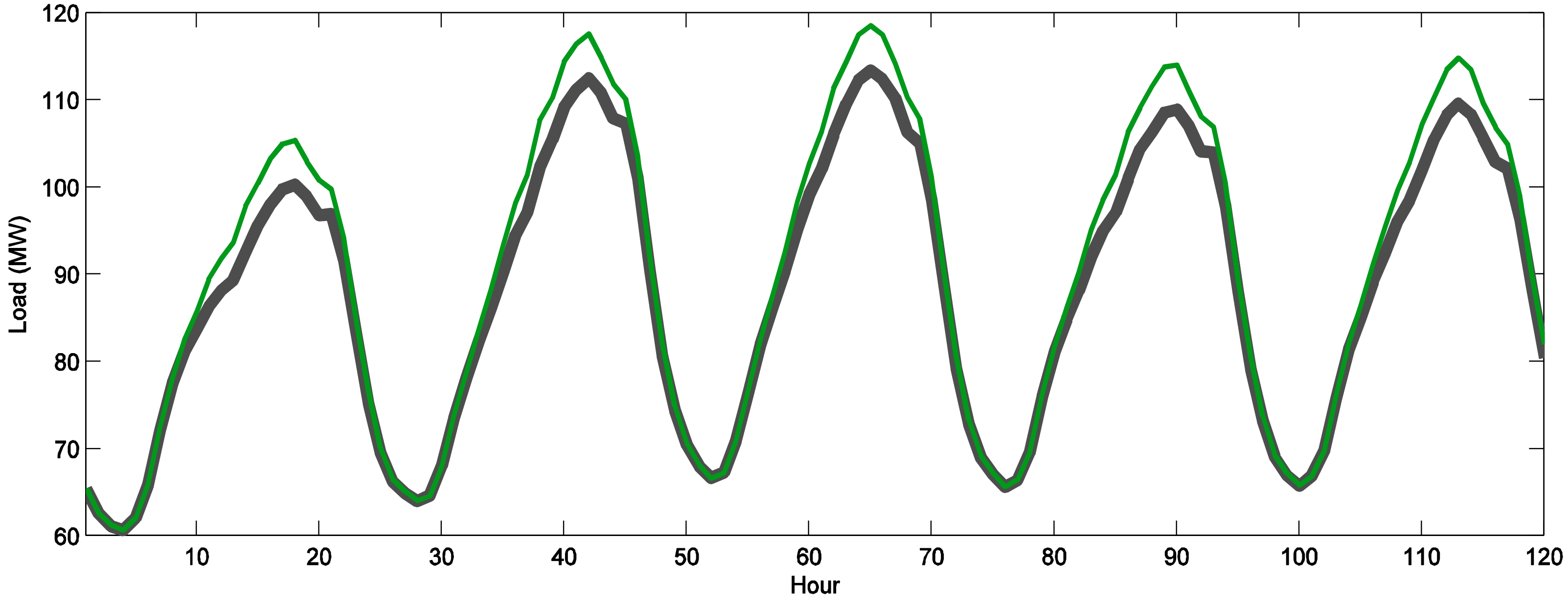
Electricity and Gasoline Prices

- Electricity prices: PG&E baseline summer time of use rates
- Gasoline prices: mean gas price in the zip code 94305 on August 2011

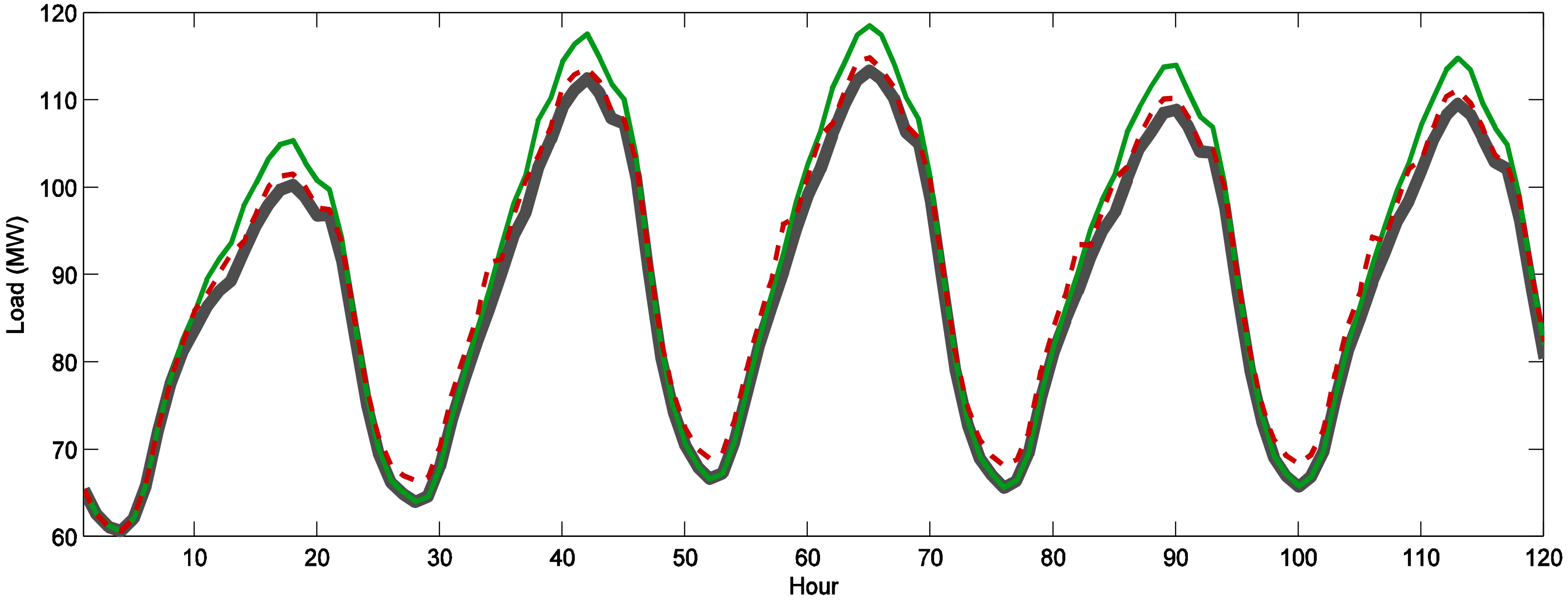
A Fleet of Vehicles, 5 Days



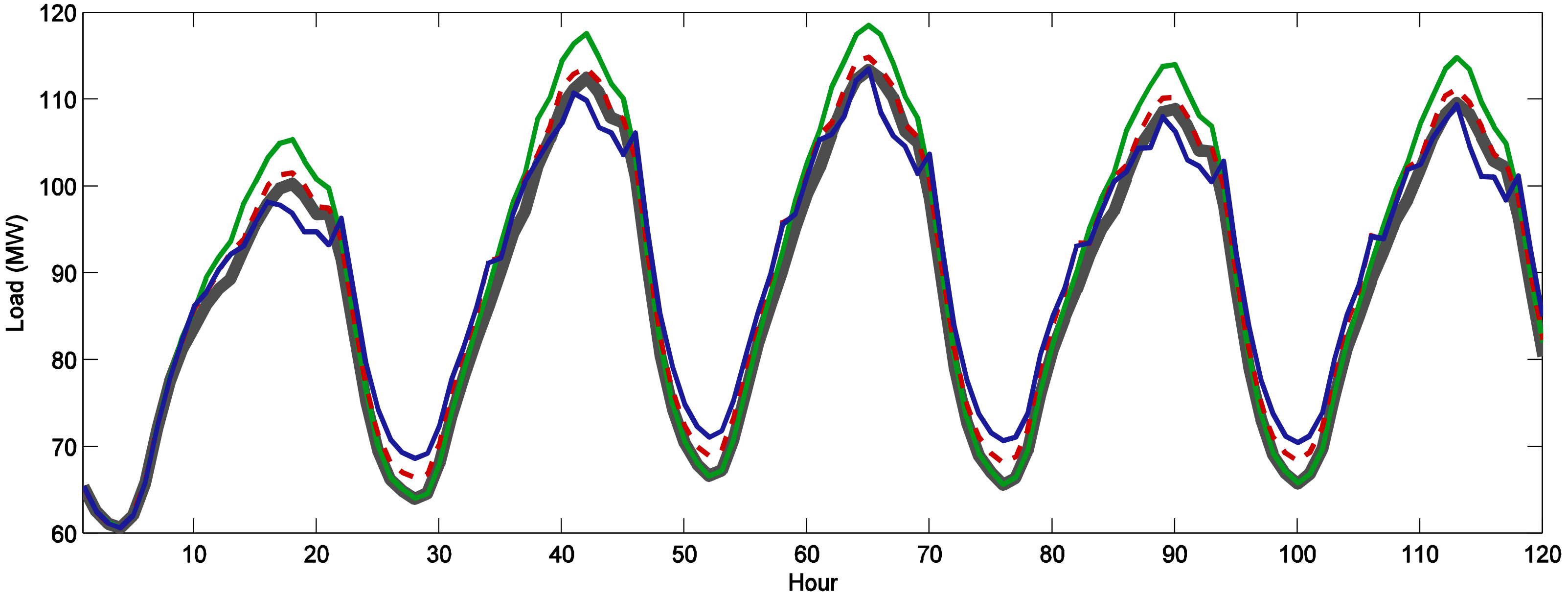
Standard Charging Policy



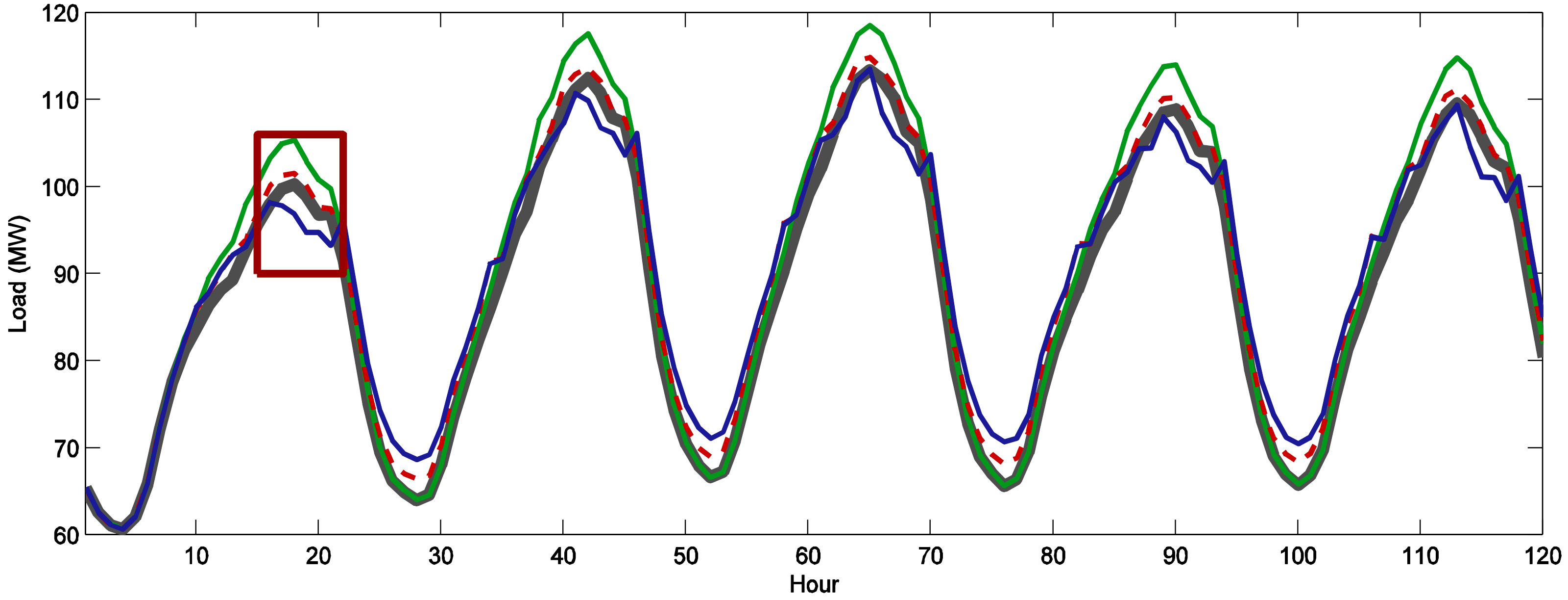
Lowest PGE Price Charging Policy



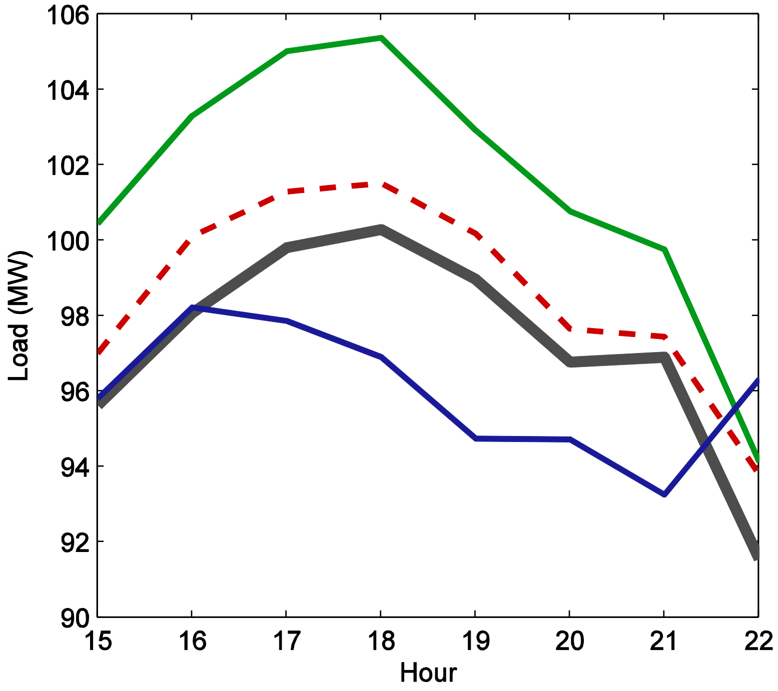
Linear Programming Charging Policy



Peak Reduction due to Discharging



Peak Reduction due to Discharging



	Standard	Low PGE	Linear Progr.
Total Fleet ($)	97,678	83,695	65,349
Mean Cost / Mile	0.068	0.044	0.0054
Increase in Peak	5.1%	1.4%	-0.25%

Beijing Public Transport Intelligent Urban Bus Operations Management with Mixed Fleet Types and Charging Schedule

北京公交 Beijing Public Transport

杉数科技 Cardinal Operations

Kickoff 2022.8

Background: Decision Intelligence in the case of Beijing Public Transport

北京公交 Beijing Public Transport
杉数科技 Cardinal Operations

北京市“十四五规划”目标

加快构建“综合、绿色、智能、安全”的立体化现代化城市交通系统

- 到2025年，中心城区绿色出行比例提高至 76.5%
- 全面推进智慧城市建设，重点发展智慧交通
- 围绕轨道交通优化地面公交线网，减少长距离、长时间运行线路，提高车辆利用率

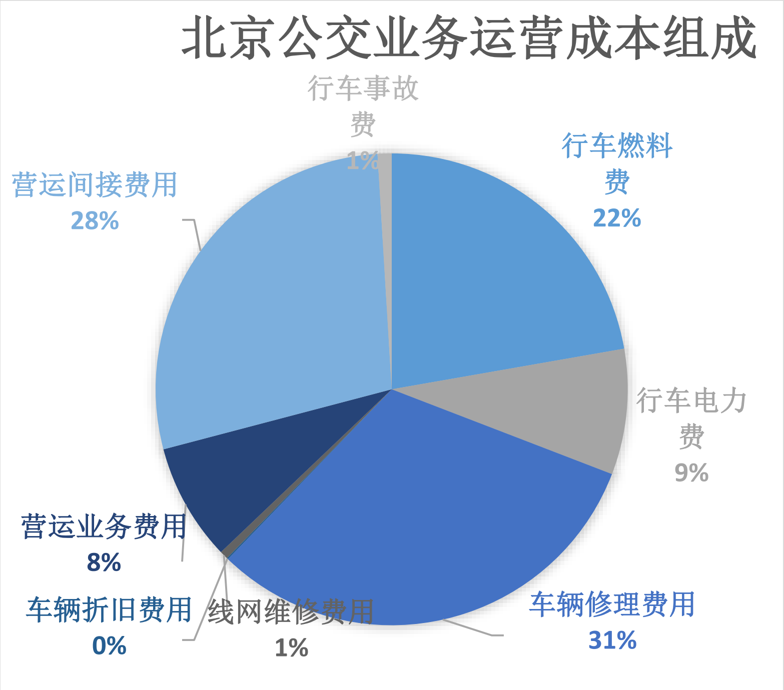
加快建立科学、高效的“城市智能运行决策管理体系”

- 高水平推动城市交通的数字转型和智慧升级，形成城市交通整体解决方案
- 加快建设公共交通网络化智能调度体系，让公交出行越来越可靠，时间有保证

More efficient and intelligent decision-making to achieve 14th Five-Year Plan goals

Beijing Public Transport Line 7 is selected as the Key Pilot Unit of the intelligent transformation of Beijing Public Transport

Intelligent Transformation
Empowered by Cardinal Operations



Beijing Public Transport's total operational costs reached 6.65 billion Yuan in 2020, of which fuels, electricity, maintenance, repair and other indirect costs accounted for over 90%. Preliminary analysis shows various potential use cases for optimization in cost reduction.

北京公交 Beijing Public Transport

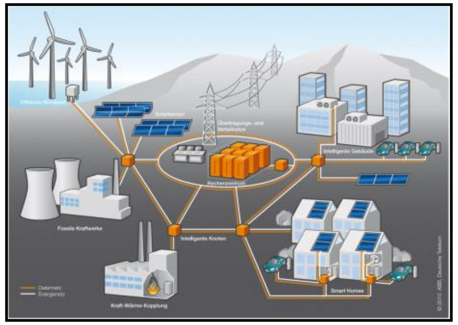
X

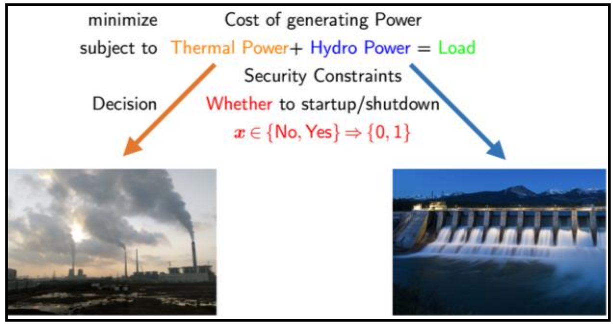
杉数科技 Cardinal Operations

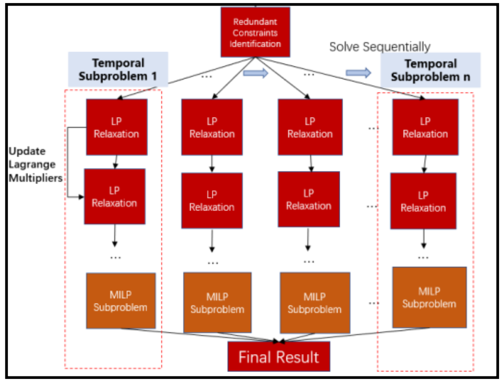
Beijing Public Transport, in partner with Cardinal Operations, aims to build an innovative integrated system for smart operations in urban public transportation operations, and explore larger markets in the future.

Topic 4: Unit Commitment and Power Grid Optimization

Algorithms and Applications, **Cardinal Operations** 2022







Unit Commitment Problem

- Electricity is generated from units (various generators)
- Transmitted *safely* and *stably* through power grids
- Consumed at minimum (reasonable) price

Optimization has its role to play

minimize Cost of electricity

subject to Safety and Stability

Adaptivity to various units

Unit commitment problem dispatches the units safely and stably at minimum cost

Case Study: Sichuan Thermal-Hydro Hybrid Model

- **A UC problem from real-life background (Sichuan Province)**
- **With 20 thermal and 230 hydro units**
- **Hydro units involve no decision (binary variables)**

Hardness

- **Costs are piecewise in generated power**
- **All the units are coupled by the Load balancing constraint**
- **A much larger and harder MILP model, *but***

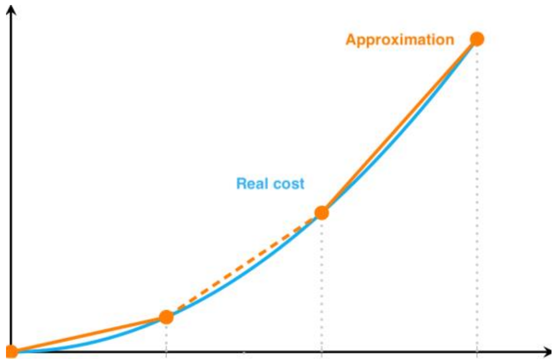
Better Modeling + Algorithm
Makes it Easier!

minimize	Cost of electricity
subject to	Thermal Power + Hydro Power = Load
	Other Constraints
Decision	Unit Operation Decision

Modeling Convex Piecewise Linear Objective

Better approximation needs **more** segments (often≥ 10)

Takeaway: better modeling matters quite a lot



Tackle Complicating Constraints by ADMM

Thermal or hydro alone are easy, but they are couple by a simple constraint

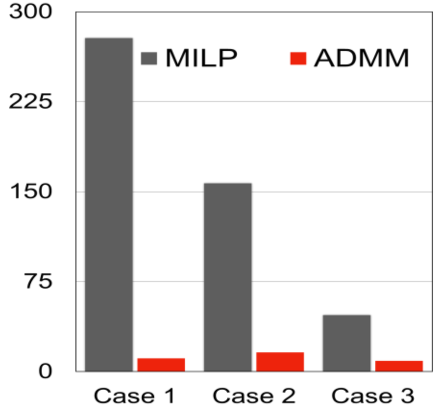
Decouple them: 20 times faster

minimize Cost of electricity

subject to Thermal Power+ Hydro Power = Load

Other Constraints

Decision Unit Operation Decision



Further Improvement: ML/AI and Probabilistic Branching (Divide-Conquer)

- **Power grid optimization is done everyday and there abundance of historical data**
- **Structure of the grid is stable and problems have similar structures**
- **Most variables can be predicted using problem parameters by Machine learning models**

Problem data → **Machine Learning** →

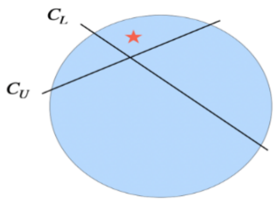
$\hat{x}_1 = 0.99$
$\hat{x}_2 = 0.92$
$\vdots$
$\hat{x}_n = 0.03$

→

$$\sum_{x_i \text{ near } 1} x_i \geq \xi \cdot \#\{x_i \text{ near } 1\}$$

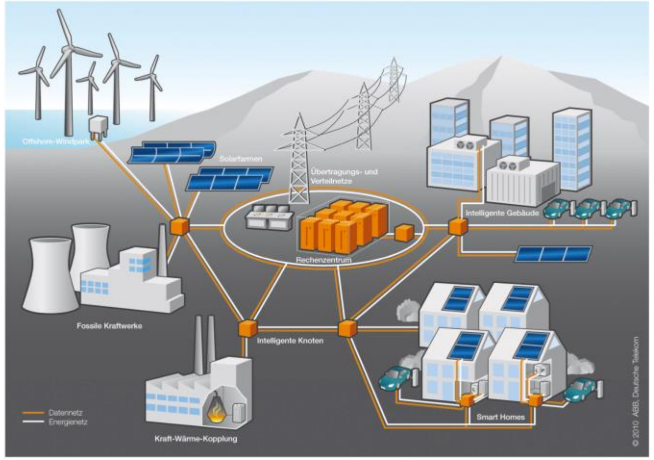
$$\sum_{x_i \text{ near } 0} x_i \leq \zeta \cdot \#\{x_i \text{ near } 0\}$$

→



Successively Implemented in a Much Larger Region

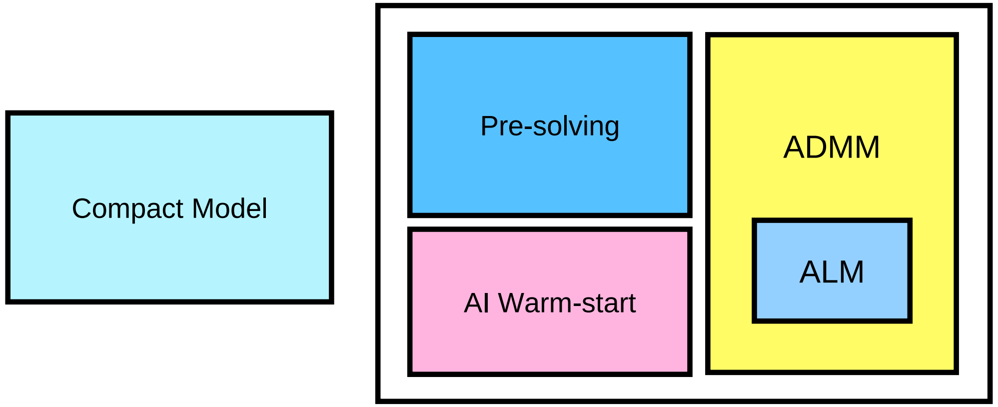
- **A much larger UC problem with security constraint**
- **With much more (millions of) constraints and variables**
- **More than 1000 units of Thermal, Hydro and New energy**
- **Consider interaction between regions and time periods**



Huge size + Various business logic + Complicated coupling constraints

- Intractable without **exploring structure**
- Accurate and succinct **model** helps
- Domain specific **algorithms** matter a lot
- **ML/AI** has a big role to play

Model, Algorithm and ML/AI together make it tractable



Topic 5: Managing engine maintenance in an aircraft fleet

杉数科技
Cardinal Operations

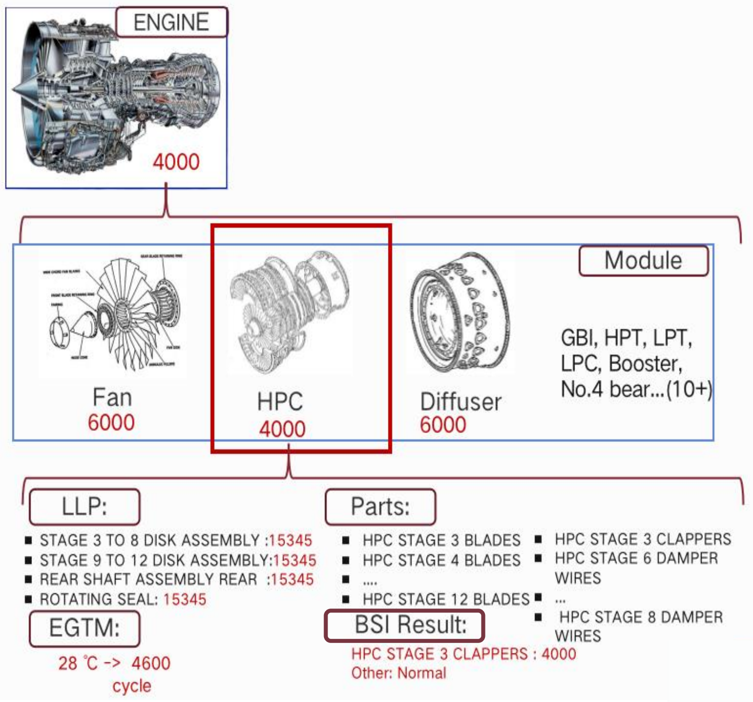
- Aircraft maintenance has been a major source of expense for airlines.
- The airlines make maintenance decisions for thousands of engines
- Engines must be sent to workshops (not airlines themselves)

Complexity of engine maintenance

- Difficult to monitor the health status (millions of parts and modules)
- Difficult to predict (deterioration suffers from stochasticity)
- Complex rules to make a valid work-scope on how to repair

Long-term planning issues

- Overhaul (vital maintenance) costs about 1-3 months in every ≈ 5 years
- Horizon > 15 years for avoid short-sighted plans
- Interest rate for costs



How to repair an aircraft engine?

杉数科技
Cardinal Operations

Steps to make a maintenance policy

- Monitor **health status** for each **module and part**
- Set a target: simple **repair** or **overhaul**
- Make a draft of maintenance **work-scope**
- Comply with manufacturer guidelines
- Negotiate with work shops to finalize
- …
- Send the engine to workshop (**shop-visit**)
- Wait for several months (**turn-around**)
- Put back to an aircraft (**on-wing**)

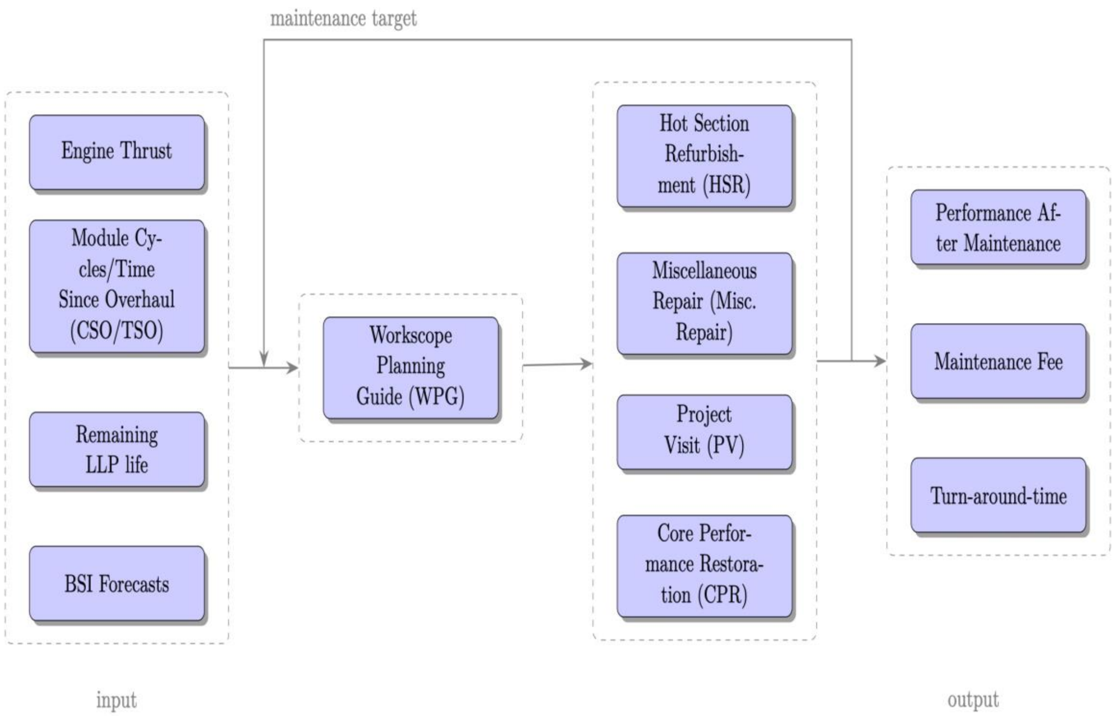
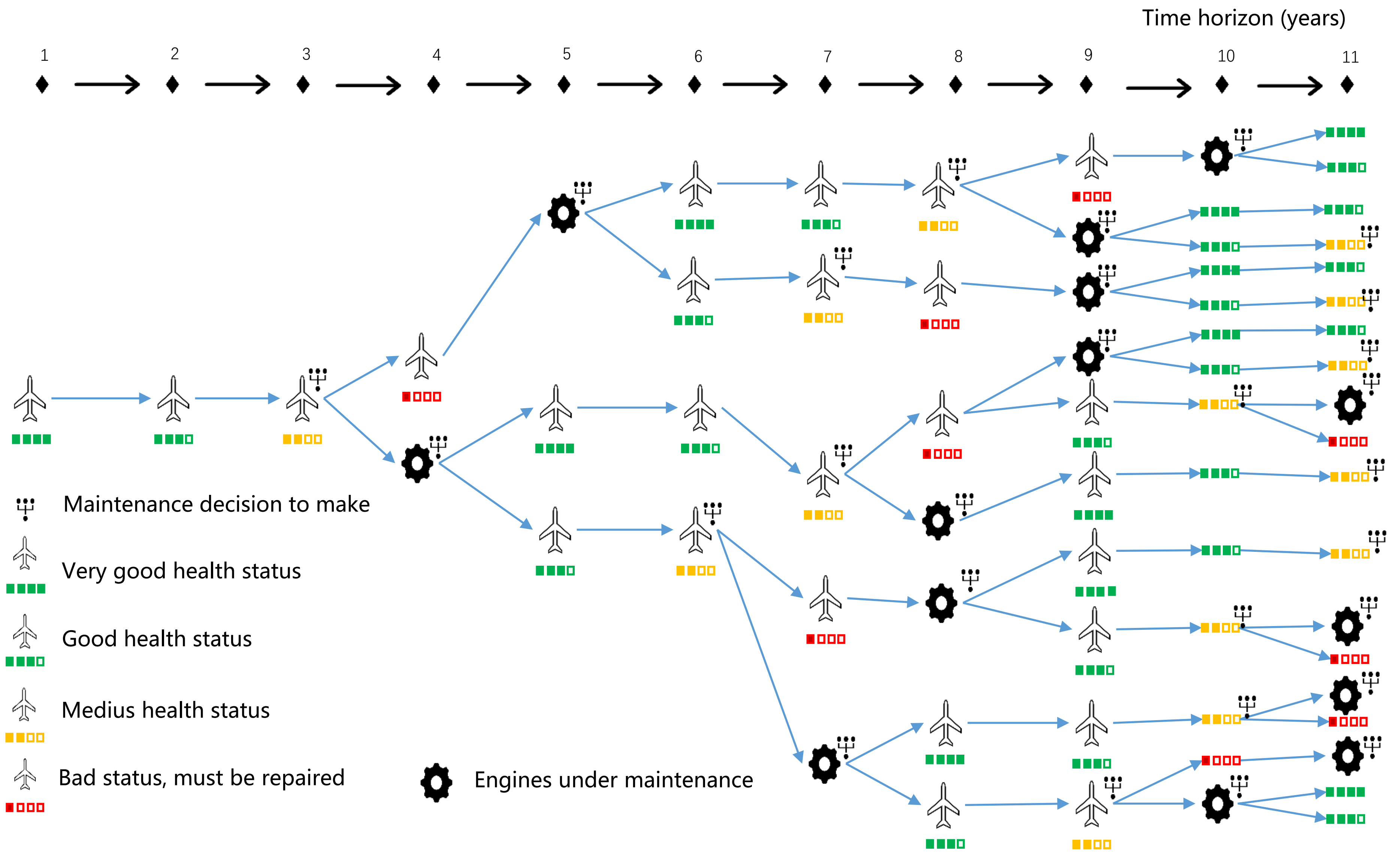


Fig. Steps to make a workscope for V2500 engines (on Airbus A320)

Life-cycle maintenance of a single engine

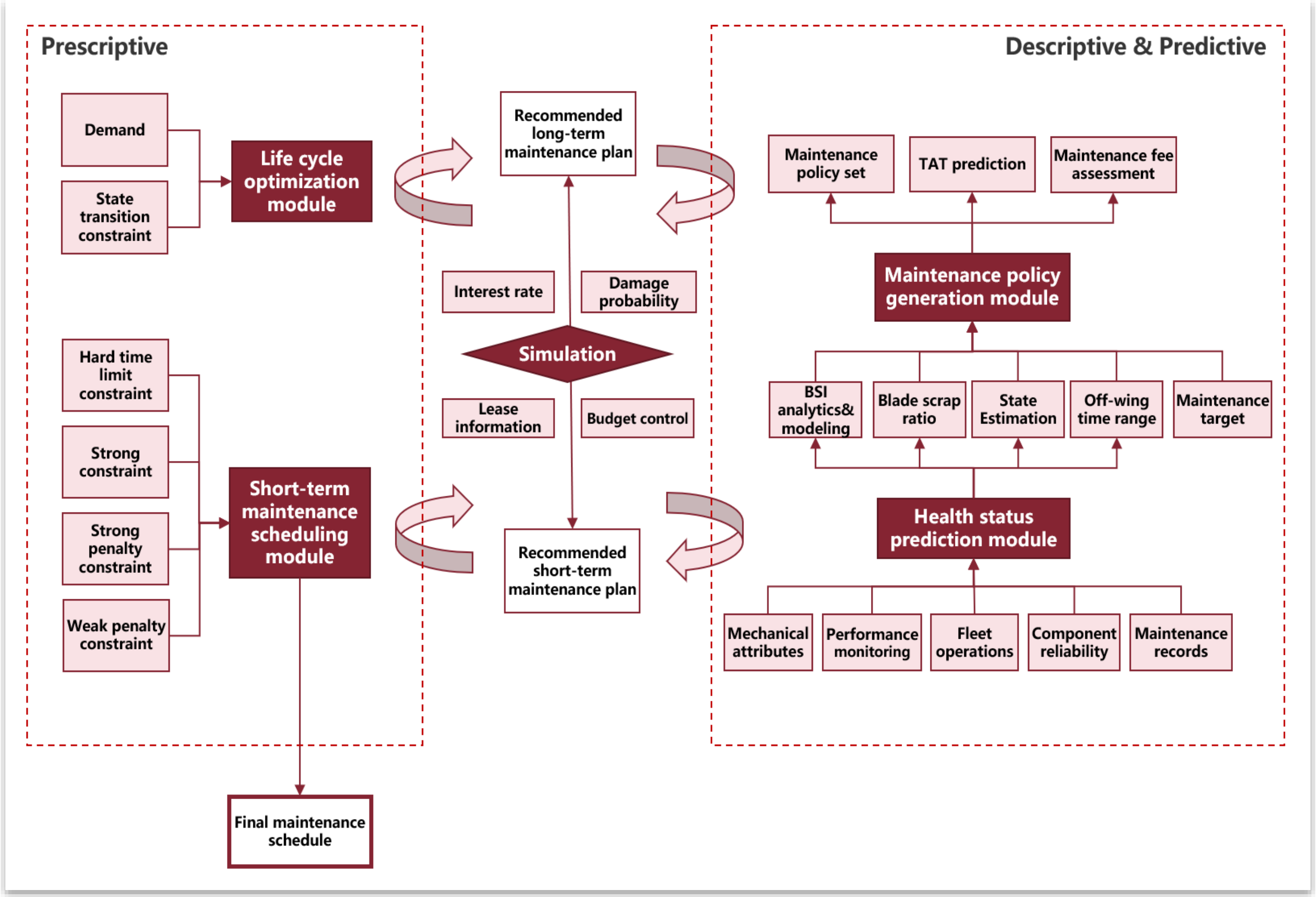
杉数科技 Cardinal Operations



An engine maintenance optimization system for CSAIR

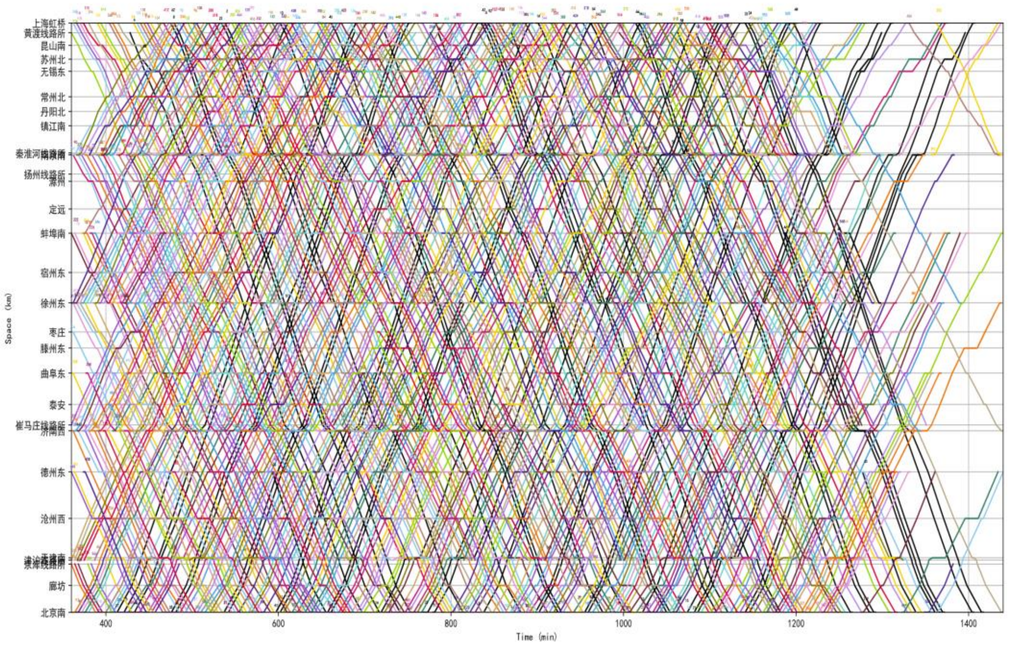
Project impact

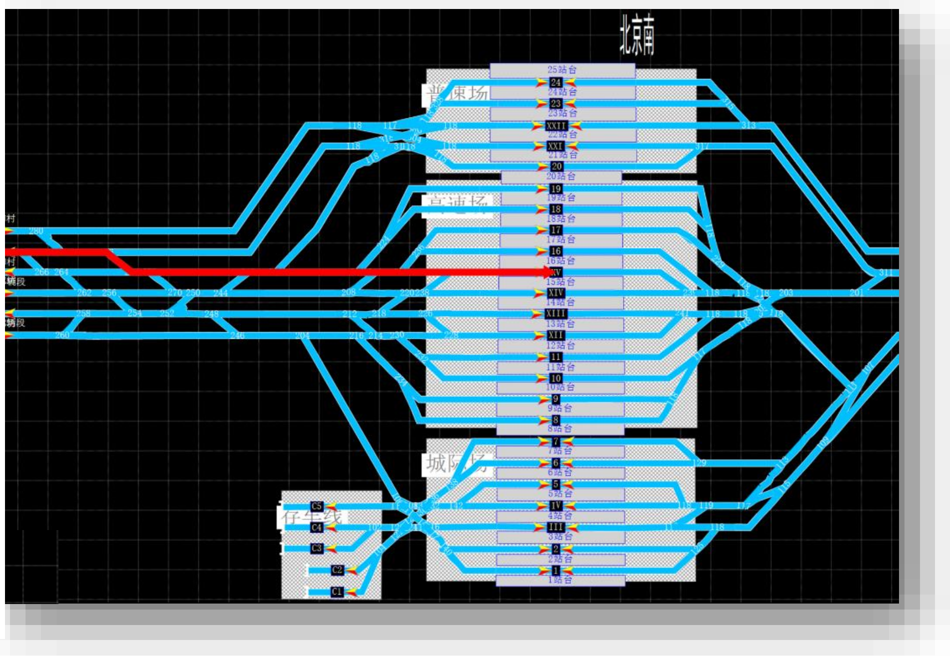
- Provide systematic tools for China Southern
- Save approximately 30% of the temporary rental days
- Reduce 95% of the “zero-spare” scenarios.
- Expect to save multi millions dollars per year



Topic 6: Beijing-Shanghai High-speed Railway Scheduling Optimization

Algorithms and Applications, **Cardinal Operations** 2022





Background

- China High-speed Railway has been committed to providing high-quality transportation services to passengers, and the formulation of train scheduling is a key link in the operation.
- **At present, train scheduling is based on human experience**, which becomes increasingly difficult to handle the growing network. Therefore, both industry and academia are seeking ways to **automate train scheduling**.
- The train scheduling problem can be divided into **Train Timetabling Problem (TTP)** and **Train Platforming Problem (TPP)**.

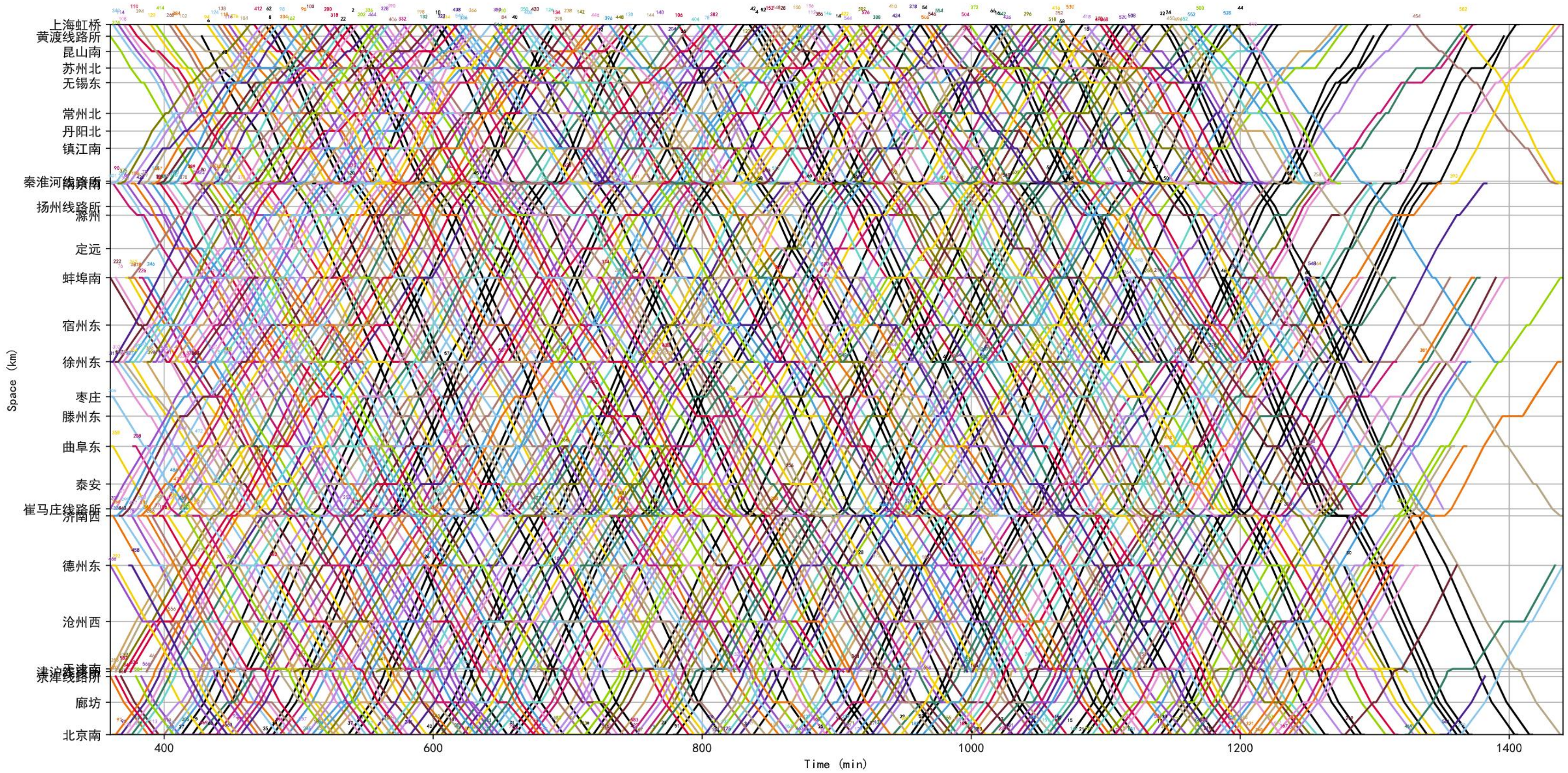
Optimization Model:

- **Objective:** maximize the number of trains placed in the train scheduling, thereby maximizing operating revenue；
- **Constraints:** describe the running behavior of trains and prevent train collisions；

- The project mainly solves **TTP for Beijing-Shanghai High-speed Railway** and **TPP at Beijingnan Railway Station**.
  - **Beijing-Shanghai High-speed Railway** is the busiest high-speed railway with the largest number of passengers in China. It is 1,318 km in total and passes 29 stations.
  - **Beijingnan Railway Station** is the largest railway station in Beijing, with the largest area and the largest number of trains.
  - Both problems are challenging scheduling tasks, which can be formulated as Mixed Integer Programming (MIP).

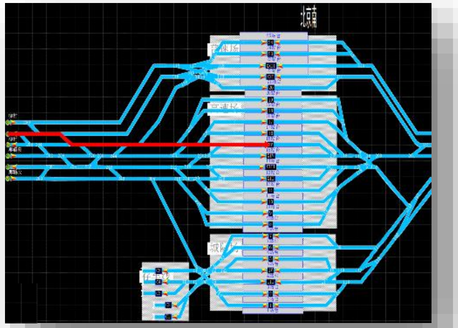
Numerical Results: TTP for Beijing-Shanghai

- We solve the TTP for Beijing-Shanghai high-speed railway using Cardinal Optimizer (COPT).
- COPT is the first fully independently developed mathematical programming solver in China with strong solving ability of MIP problem. It also has excellent performance in solving this problem.
- The result is presented in the following figure. We only need about **1000 seconds** to schedule 584 train in two directions.



Numerical Results: TPP at Beijingnan Station

- We solve the TPP at Beijingnan Railway Station using Cardinal Optimizer (COPT).
- Considering the connection pairs and ensuring the feasibility, we solve the model within **2 hours**, which is much less than manual scheduling.
- The result is presented in the following table, including time nodes about occupation at boundaries and tracks.



列车编号	前序车站	进入站界	进站路径	停靠站线	离开站界	出站路径	后序车站	进入站界时间	进入站线时间	离开站线时间	离开站界时间
361		站界:B10		站线:XIV	站界:B9	站线:10:XIV	廊坊		12:00:00	12:06:00	12:10:00
74	廊坊	站界:B8	站线:16:8	站线:8	站界:B7			11:57:00	12:02:00	12:17:00	
125		站界:B10		站线:11	站界:B9	站线:13:11	廊坊		12:06:00	12:13:00	12:17:00
114	廊坊	站界:B8	站线:7:17	站线:17	站界:B7			12:10:00	12:14:00	12:29:00	
251		站界:B10		站线:8	站界:B9	站线:16:8	廊坊		12:17:00	12:27:00	12:32:00
20	廊坊	站界:B8	站线:7:17	站线:17	站界:B7	站线:7:17		12:19:00	12:23:00	12:25:00	12:29:00
96	廊坊	站界:B8	站线:13:11	站线:11	站界:B7			12:25:00	12:29:00	12:44:00	
223		站界:B10		站线:17	站界:B9	站线:7:17	廊坊		12:29:00	12:44:00	12:48:00
8	廊坊	站界:B8	站线:8:16	站线:16	站界:B7			12:33:00	12:37:00	12:42:00	
23		站界:B10		站线:16	站界:B9	站线:8:16	廊坊		12:42:00	12:57:00	13:01:00
127		站界:B10		站线:11	站界:B9	站线:13:11	廊坊		12:44:00	12:49:00	12:53:00
572	廊坊	站界:B8	站线:5:19	站线:19	站界:B7			12:43:00	12:48:00	13:03:00	
124	廊坊	站界:B8	站线:6:18	站线:18	站界:B7			12:47:00	12:52:00	12:57:00	
102	廊坊	站界:B8	站线:15:9	站线:9	站界:B7			12:51:00	12:56:00	13:07:00	
225		站界:B10		站线:18	站界:B9	站线:6:18	廊坊		12:57:00	13:12:00	13:17:00
51		站界:B10		站线:17	站界:B9	站线:7:17	廊坊		12:59:00	13:01:00	13:05:00
116	廊坊	站界:B8	站线:13:11	站线:11	站界:B7			12:56:00	13:00:00	13:15:00	
169		站界:B10		站线:19	站界:B9	站线:5:19	廊坊		13:03:00	13:18:00	13:23:00
133		站界:B10		站线:9	站界:B9	站线:15:9	廊坊		13:07:00	13:22:00	13:27:00
161		站界:B10		站线:11	站界:B9	站线:13:11	廊坊		13:15:00	13:26:00	13:30:00
138	廊坊	站界:B8	站线:5:19	站线:19	站界:B7			13:13:00	13:18:00	13:33:00	
118	廊坊	站界:B8	站线:8:16	站线:16	站界:B7			13:27:00	13:31:00	13:36:00	
109		站界:B10		站线:19	站界:B9	站线:5:19	廊坊		13:33:00	13:41:00	13:46:00
100	廊坊	站界:B8	站线:8:16	站线:16	站界:B7			13:31:00	13:35:00	13:40:00	
229		站界:B10		站线:16	站界:B9	站线:8:16	廊坊		13:36:00	13:51:00	13:55:00
2	廊坊	站界:B8	站线:16:8	站线:8	站界:B7			13:34:00	13:39:00	13:47:00	
131		站界:B10		站线:16	站界:B9	站线:8:16	廊坊		13:40:00	13:55:00	13:59:00
3		站界:B10		站线:8	站界:B9	站线:16:8	廊坊		13:47:00	14:02:00	14:07:00
98	廊坊	站界:B8	站线:10:XIV	站线:XIV	站界:B7			13:43:00	13:47:00	14:02:00	
108	廊坊	站界:B8	站线:13:11	站线:11	站界:B7			13:47:00	13:51:00	14:06:00	

Takeaways

- **Realize the power of Data, Math and Science**
- **Utilize quantitative decision/computation models and efficient algorithms**
- **Always search for optimality!**

BusinessWeek

JANUARY 23, 2006

COVER STORY

Math Will Rock Your World

A generation ago, quants turned finance upside down. Now they're mapping out ad campaigns and building new businesses from mountains of personal data…