Optimization and Operations Research in Mitigation of a Pandemic

Yinyu Ye

Department of Management Science and Engineering
Institute of Computational and Mathematical Engineering
Stanford University
Stanford, CA 94305, U.S.A

May 2020

Thanks to Cottle, Glynn, Iglehart, Luenberger, and Saunders for their help in preparing the presentation

Contents

Operations Research: An interdisciplinary-science-based and data-driven strategy/policy/decision making process/system under complex/uncertain/dynamic environments.

- Inventory and Risk Pooling of Medical Equipment in a Pandemic
- Reopen Economy -- Algorithm-Based Supply-Chain-Network Management
- Monte Carlo/Computer Simulation Methods for Drastic and Rare Scenario Analyses
- New Norm: Operation/Optimization under Social Distance Constraints
- Indoor GPS and Tracking by Sensor Network Localization for Contact-Tracing
- Dynamic and Equitable Region Partitioning for Hospital/Health-Care Services
- Efficient Public Goods Allocation under Tight Capacity Restriction via Market Equilibrium Mechanisms/Platforms
- Identifications and Protective Measures for High-Risk Groups in a Pandemic
- Machine-Learning (e.g., Logistic Regression) of Multiple Social Features to Reduce Pandemic Fatality

Inventory and Risk Pooling of Medical Equipment in a Pandemic

Managing Uncertainty

The War of Ventilators: Inventory Management in Uncertain Environments – Decentralization vs. Centralization

I want 30,000

You ask too much!

Governors Fight Back Against Coronavirus Chaos: ‘It’s Like Being on eBay With 50 Other States’ ***(New York Times, March 31, 2020)***

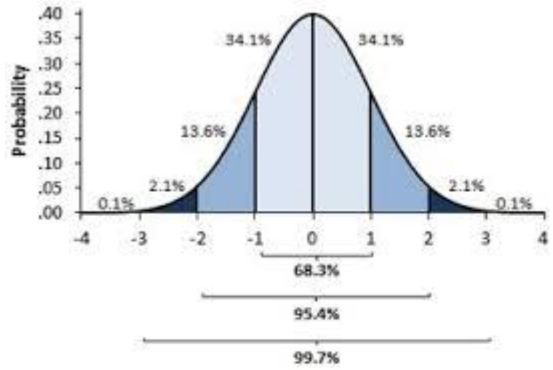
There is a “bizarre situation” in which every state buys its own ventilators, pitting them against each other in bidding war. The federal government, Fema, should have been the purchasing agent: buy everything and then allocate by need to the states. However, Fema gets involved and starts bidding and even drives up the price.

Beat Uncertainty Through Safety Stock: How Many to Order

The order quantity q^* satisfies:

$$\text{Prob}(D \leq q^*) = 0.99$$

$$\Rightarrow \quad F(q^*) = 0.99$$



Consider $D \sim N(10000, 9000^2)$

The needed quantity of ventilators is

$$q^* = \mu + \boxed{F_{norm}^{-1}(0.99) * \sigma}$$ ← Safety Stock

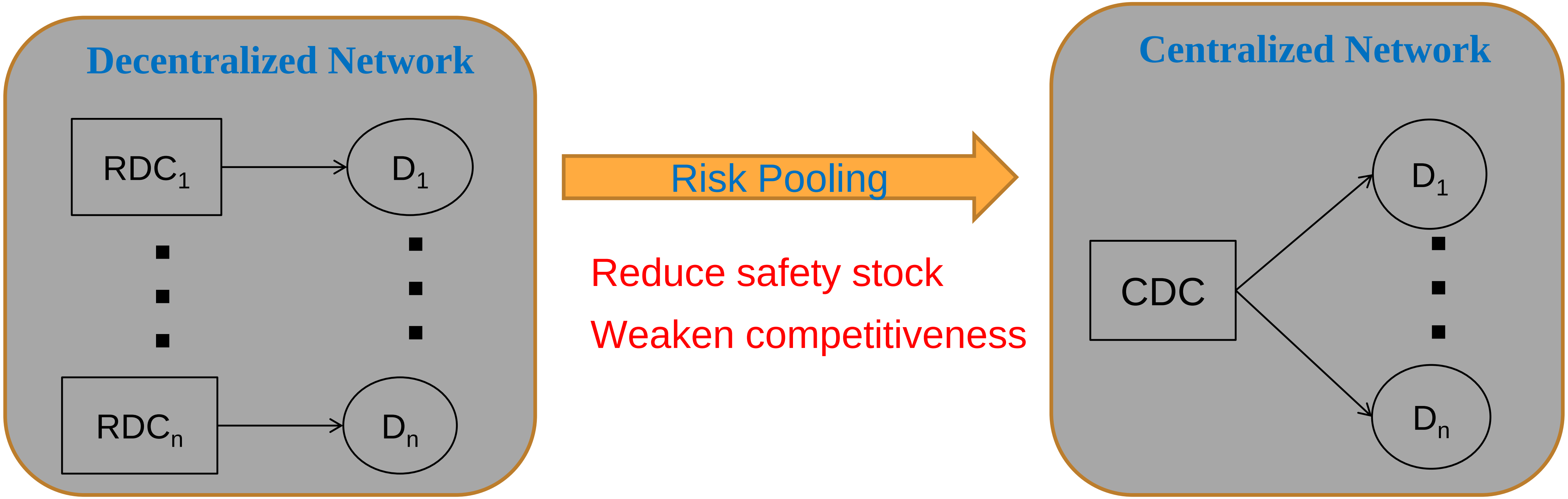
$$= 10000 + 2.3263 * 9000$$

$$= 30937$$

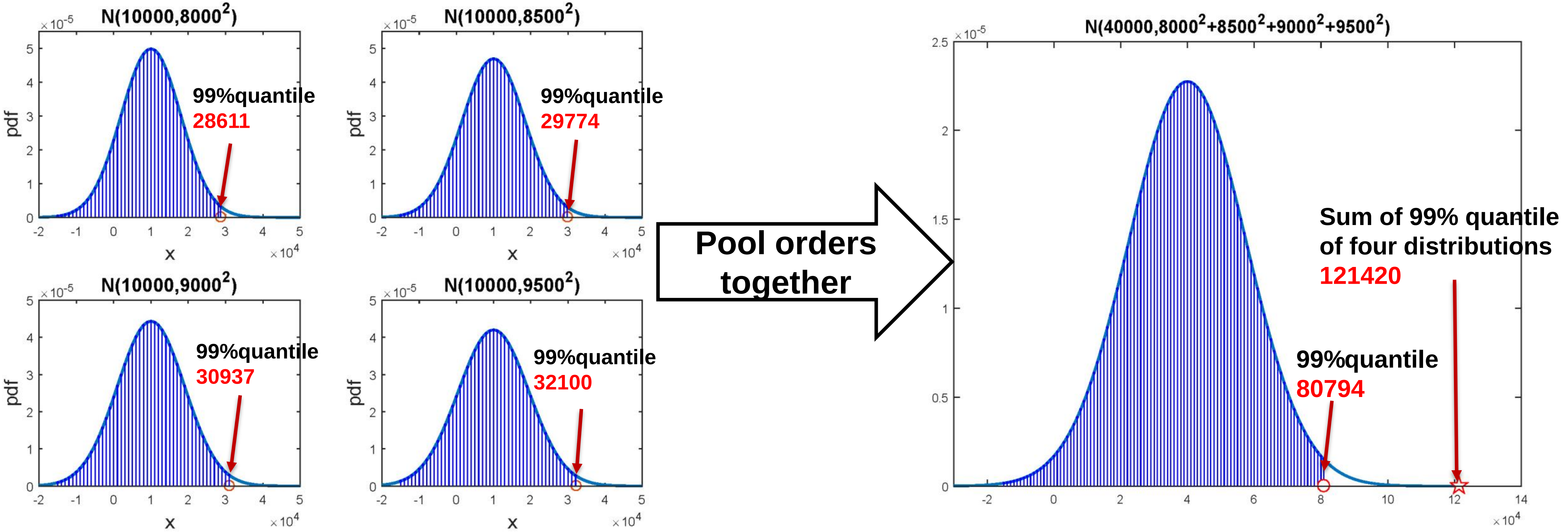
Inventory Network Management: Risk Pooling via Centralization

How many ventilators to order for COVID-19?

- The number of patients is random and follows a certain distribution $F(\cdot)$
- Desired guarantee level : 99%



Inventory Network Management:
Centralization vs. Decentralization



***Even though they may be correlated, the variance can still be reduced due to time delays…**

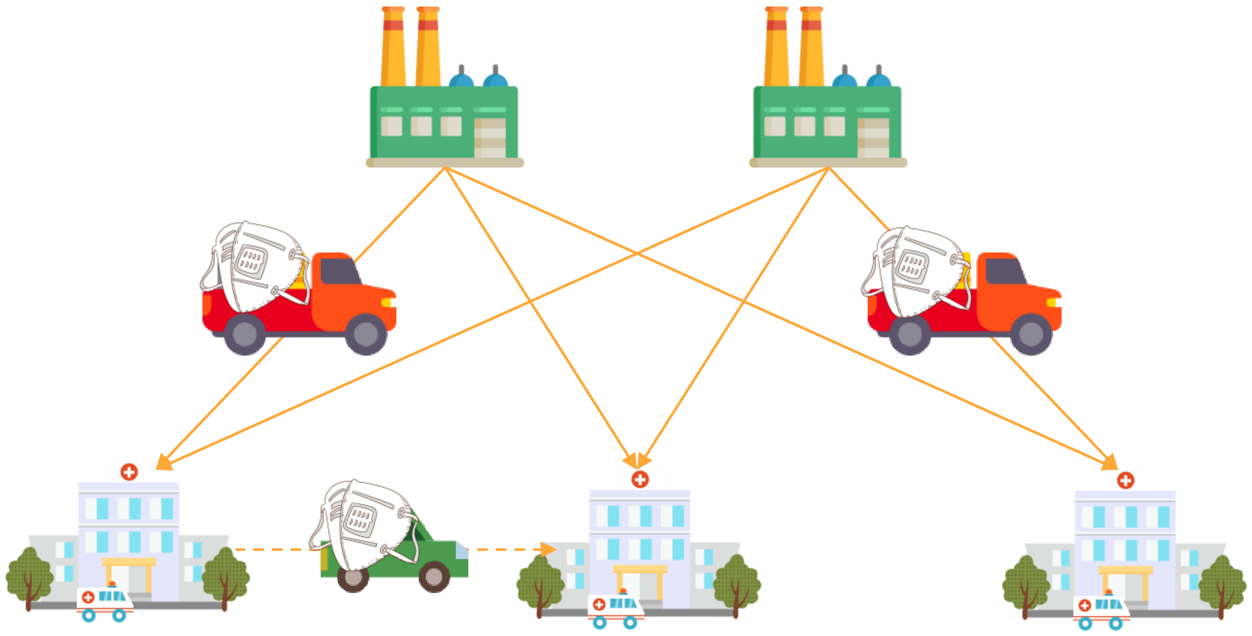
New York and 6 Other States Form Consortium to Buy Vital COVID-19 Supplies to Fight 'Totally Inefficient' Process *(Time, MAY 3, 2020)*

Reopen Economy -- Algorithm-Based Supply-Chain-Network Management Machine Decisioning

Complex Supply-Chain Network Management

Forecasting + Inventory Replenishment + Production Scheduling + Vehicle Routing

1. Make centralized inventory replenishment and efficiently allocate inventory to RDCs facing the Challenge of Coronavirus
2. During the pandemic, there is violent change of supply-chain network and resources; the solution should take in emergent changes in business and data and make **Robust, Flexible, and Agile** responses.



Supply-Chain Challenges

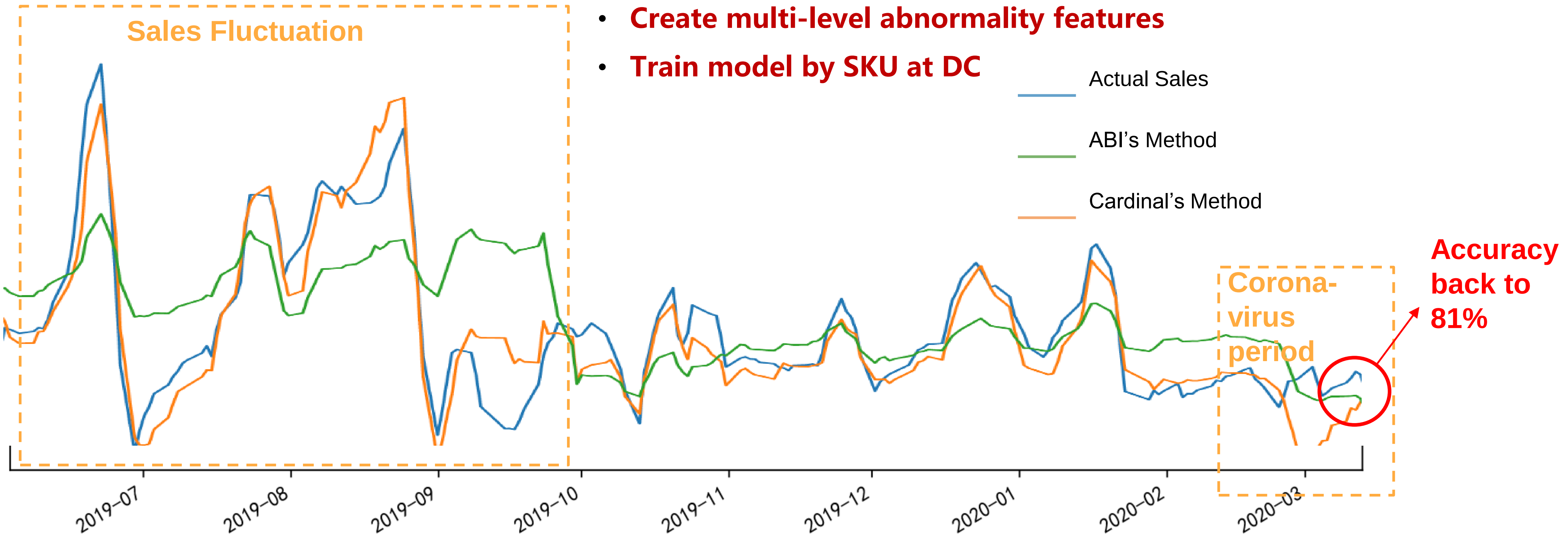
- **Lowered demand and volatility**
- **Interruption of factory, warehouse and transportation resources**
- **Change production cycle time and production leadtime**
- **Shortage in critical materials**
- **……**

Forecasting: Model Adjustment to Pandemic Environment

A Case from Cardinal Operations

10% improvement in forecasting accuracy

- **Takes in supply-chain network and resource changes**
- **Create multi-level abnormality features**
- **Train model by SKU at DC**



- **Forecast accuracy improves to 80% on average**
- **Achieve a stable 10% improvement**
- During Coronavirus, the new model quickly learns March trend and improves April accuracy to 81%!

Robust: 90%Quantile-Based Dynamic DOI Calculation in Pandemic

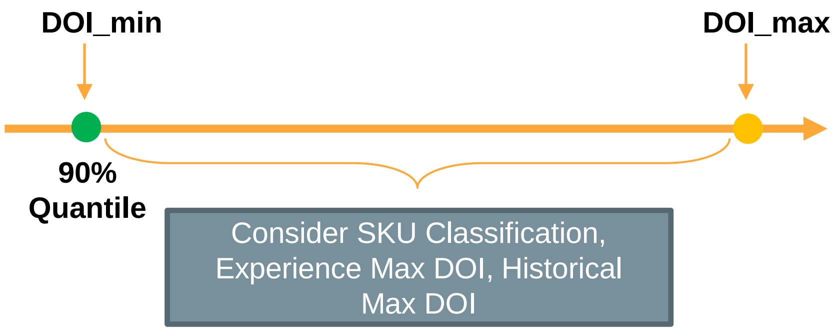
DOI (Days of Inventory) is calculated at SKU-level and updated daily.

AS-IS

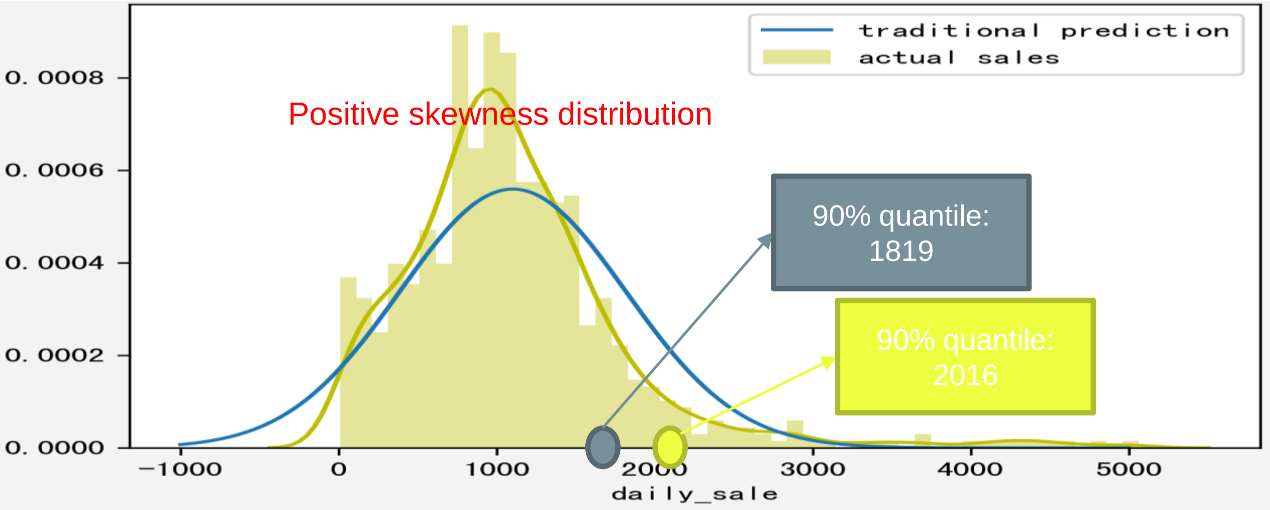
- Refer to traditional safety stock calculation
- Forecast errors are assumed to be Gaussian i.i.d
- DOI_min is monthly/quarterly updated

$$S_s = K\sqrt{E(LT)\,\sigma_d^2 + E(d)^2\sigma_{LT}^2}$$

TO-BE



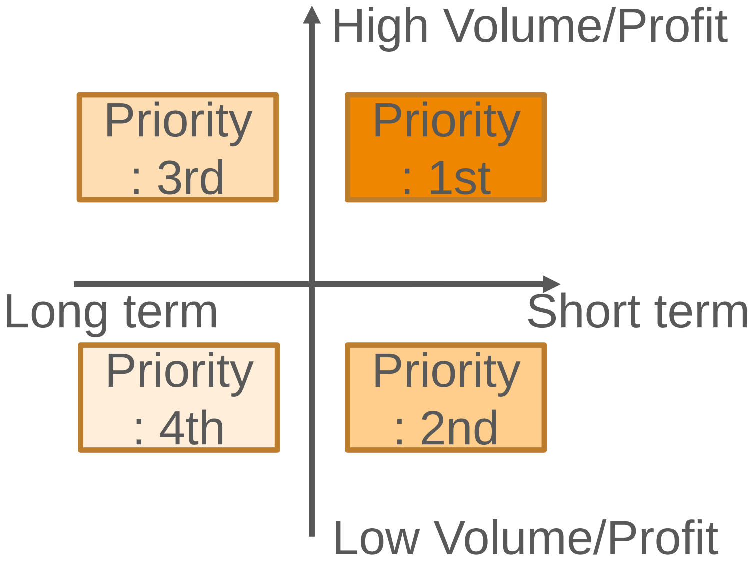
- Use 90% quantile forecast to calculate DOI_min
- DOI_min is daily updated along with demand forecast at SKU level



Simulation Results:

- Achieve better Out-of-Stock performance (-3.39%) with lower inventory level (-8.8%)
- Quantile-based dynamic DOI can quickly respond to the market changes, so it performs especially better in the fluctuating period.

Flexible: Possible Material Substitution



Alternative BOM (Bill of Materials)

Alternative Product

Priority Re-setting

Flexible Working Time

Agile I: Adjustments to the Regular Optimization Framework

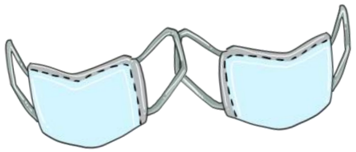
Objective Functions

min Transportation Cost
\+ Risk of out-of-stock
\+ Risk of overstock
\+ Violation of DC special request replsmt.
\+ Usage of Hub DC to transfer
\+ Volume of obsolescent inventory
\+ ……

Decision Variables and Business Constraints

1. **Optimize replenishment quantity**
   DOI_min/max (risk of out-of-stock/overstock)
2. **Factory constraints**
   inventory / production plan / push target
3. **DC constraints**
   capacity / hub DC
4. **Transportation capacity**
5. **Other special constraints**
   Half/full/even pallet, co-shipment

→ **Base Constraints** *(2 days ahead of plan date)*



6. **Shipping constraints**
   SKU and truck type/ weight / MOQ
7. **Stakeholders' feedback**
   Fix/increment/deduction of replenishment request

→ **Truck Arrangement & Feedback** *(1 day ahead of plan date)*

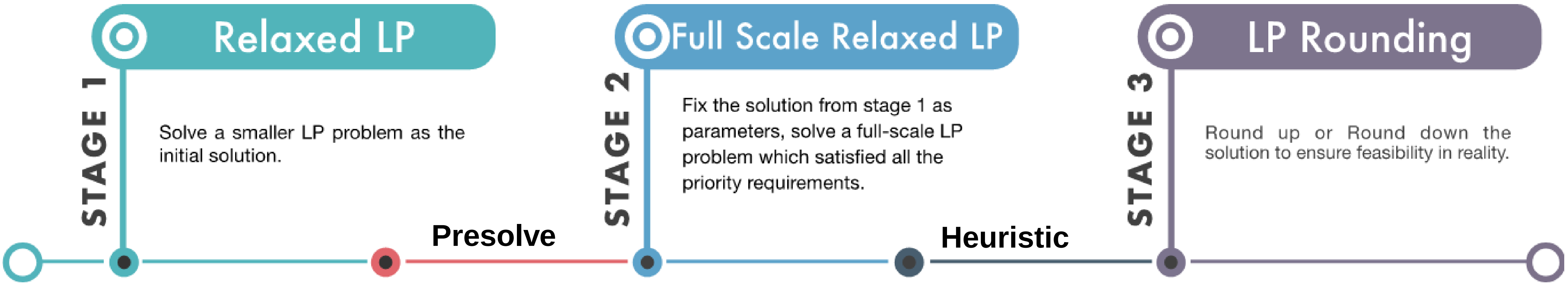
8. **D-0 adjustment**
   urgent request / tail volume

→ **Execution/Urgent** *(Real-time adjustment)*

COPT Performance
- Solving time: $\leq 10\ min$
- Number of decision variables: $\geq 10^5$
- Number of constraints: $\geq 10^5$

Integer Linear Programming Solver

Traditional ILP solver is not fast enough, so…



Performance

- Modeling and solving time: $\leq$ $3\ hours$
- Number of decision variables: $\geq 10^7$
- Number of constraints: $\geq 10^7$

Advantages

- Flexible: rolling on daily basis
- Emergency response
- Root cause analysis

The Results for Budweiser Korea

OOS (out of stock)

- Average out-of-stock rate for 2020 April is <u>0.37%</u>, which is 46% improvement from client's <u>0.69%</u>.

MUC (misallocation unit cost)

- Average misallocation unit cost for 2020 April is <u>15.98</u>, which reduces 3.32 kwr per 10L from client's <u>19.3.</u>

Abnormality Analysis

- <u>Production halts, inventory is insufficient to cover sales.</u>

If production is further improved, OOS further reduces to 0.16% (77% improvement); MUC further reduces to 11kwr per 10L.

- <u>Transportation capacity reduces, or brewery shutdowns, and there is misallocation of resources.</u>

If these return to normal, MUC further reduces to 1.2kwr per 10L (85% improvement)

Agile II: CDC+RDC Case from Cardinal Operations

Based on a major logistic company's warehouse network and logistics system, provide its customer (product owners) efficient inventory network management solution.

Network design
- Position of RDCs, coverage cities of each RDCs

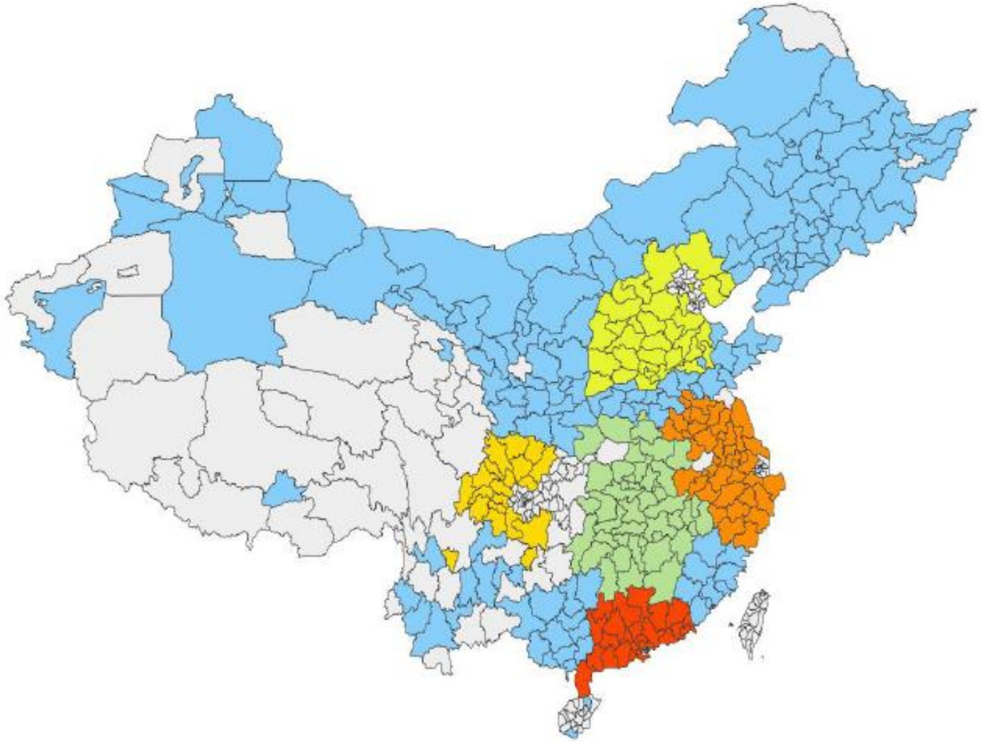
Assortment
- Which SKUs forward to RDCs
- Which SKUs only hold at CDC

Forecast
- Use machine learning to dynamically predict the demand distribution of each SKU

Inventory
- Dynamic safety stock calculation for RDCs and CDC

Demand fulfillment
- Optimize demand fulfillment for real-time order pools

The Simulation Results



Solution Example

Improvement

No.	Category	Cost saving %	Cost saving
1	Clothing	24.40%	¥1,384,240
2	Clothing	14.01%	¥1,016,667
3	Clothing	17.92%	¥2,752,222
4	Clothing	18.73%	¥2,305,534
5	Clothing	24.28%	¥1,985,278
6	Baby care	6.20%	¥590,467
7	Clothing	14.40%	¥473,323
8	3C	16.21%	¥11,678,263
9	Clothing	12.26%	¥412,173

Monte Carlo/Computer Simulation Methods for Drastic and Rare Scenario Analyses

How to Evaluate a Drastic but Rare Scenario/Event Outcome

- **Sample Testing**
  - Random sampling, Focused or Importance sampling
- **Robust Decisioning and Optimization**
- **Computer-Simulation-Based Stress Tests**
  - Simulation allows us to quickly and inexpensively acquire knowledge concerning a problem that is usually gained through experience
  - Monte Carlo simulation is an important and flexible tool for modeling situations in which uncertainty is a key factor
  - Analyze how a health-care system fares in drastic and rare scenarios based on rare event simulation

Design of a Simulation System for Financial/Banking Networks

- Estimate the distribution of losses for a banking network
- Test the losses incurred by the initial shock and the losses resulting from the contagion process

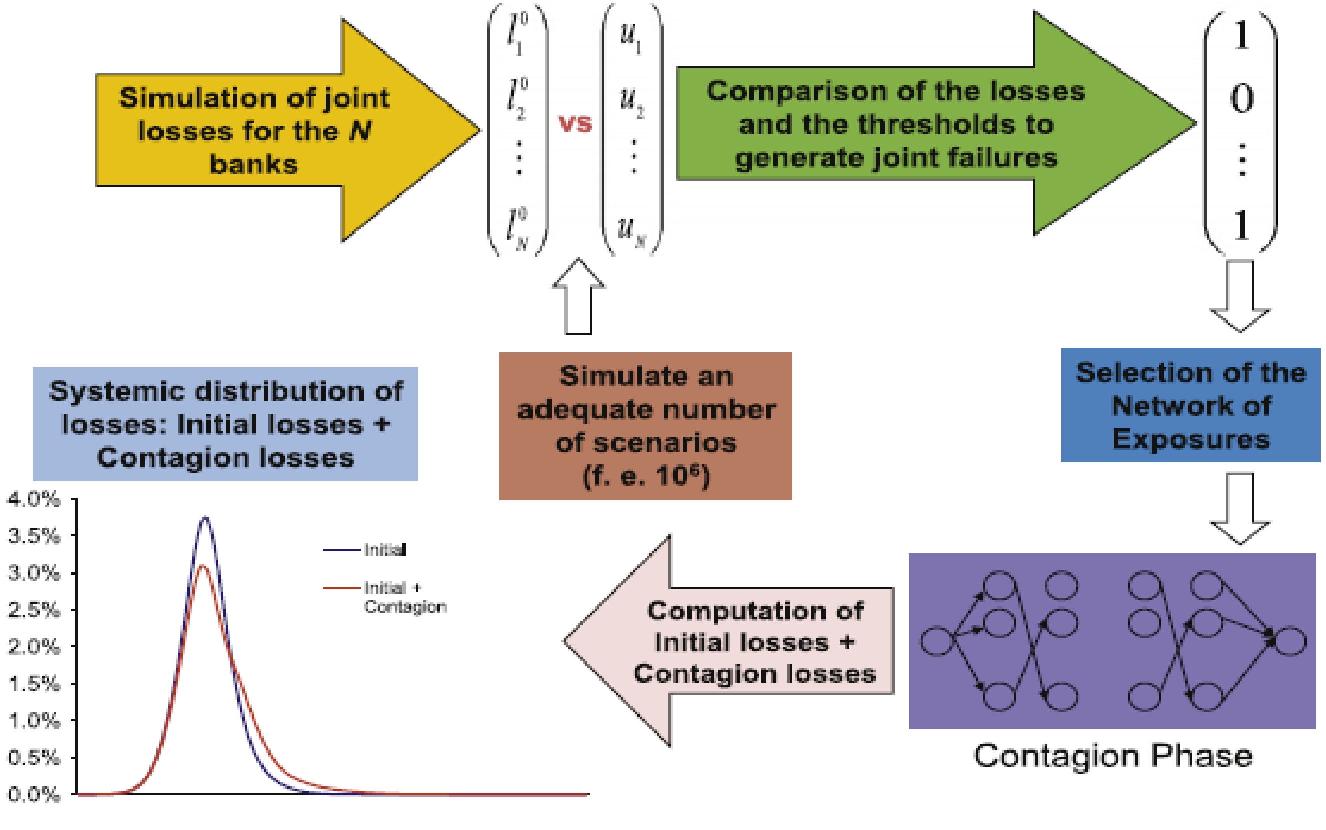


Fig. 1. The simulation algorithm.

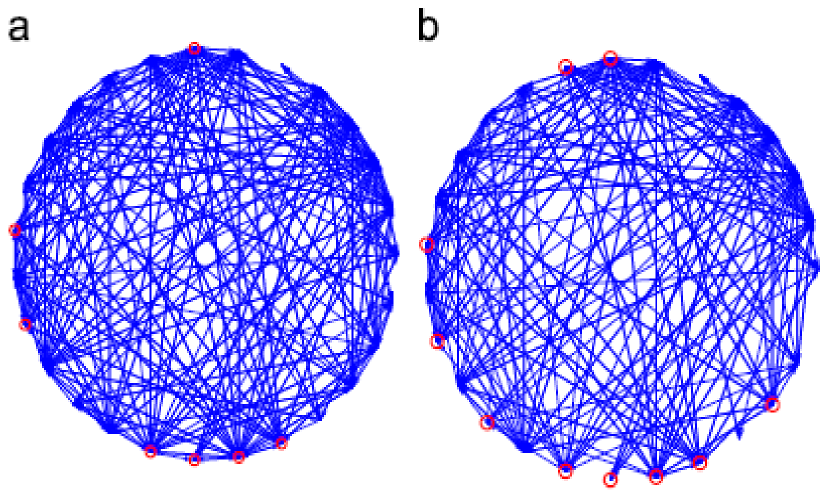
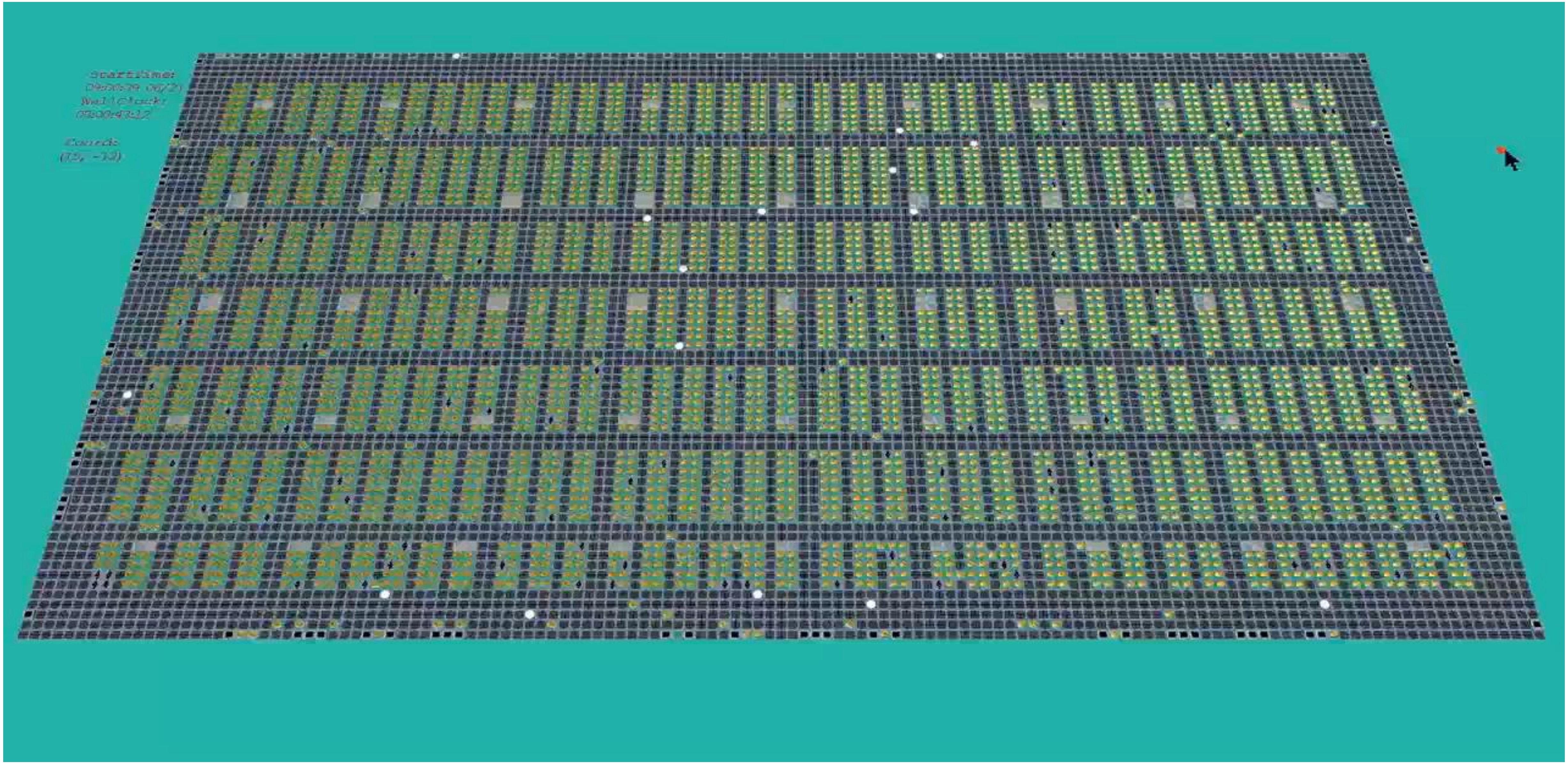


Fig. 2. The interbank exposures network on the 21st (a) and on the 27th (b) of December 2007.

Figures come from:
Serafín MJ et al., 2010. Systemic risk, financial contagion and financial fragility. *Journal of Economic Dynamics & Control*. 34: 2358–2374)

Simulation System for a Fully Unmanned Warehouse



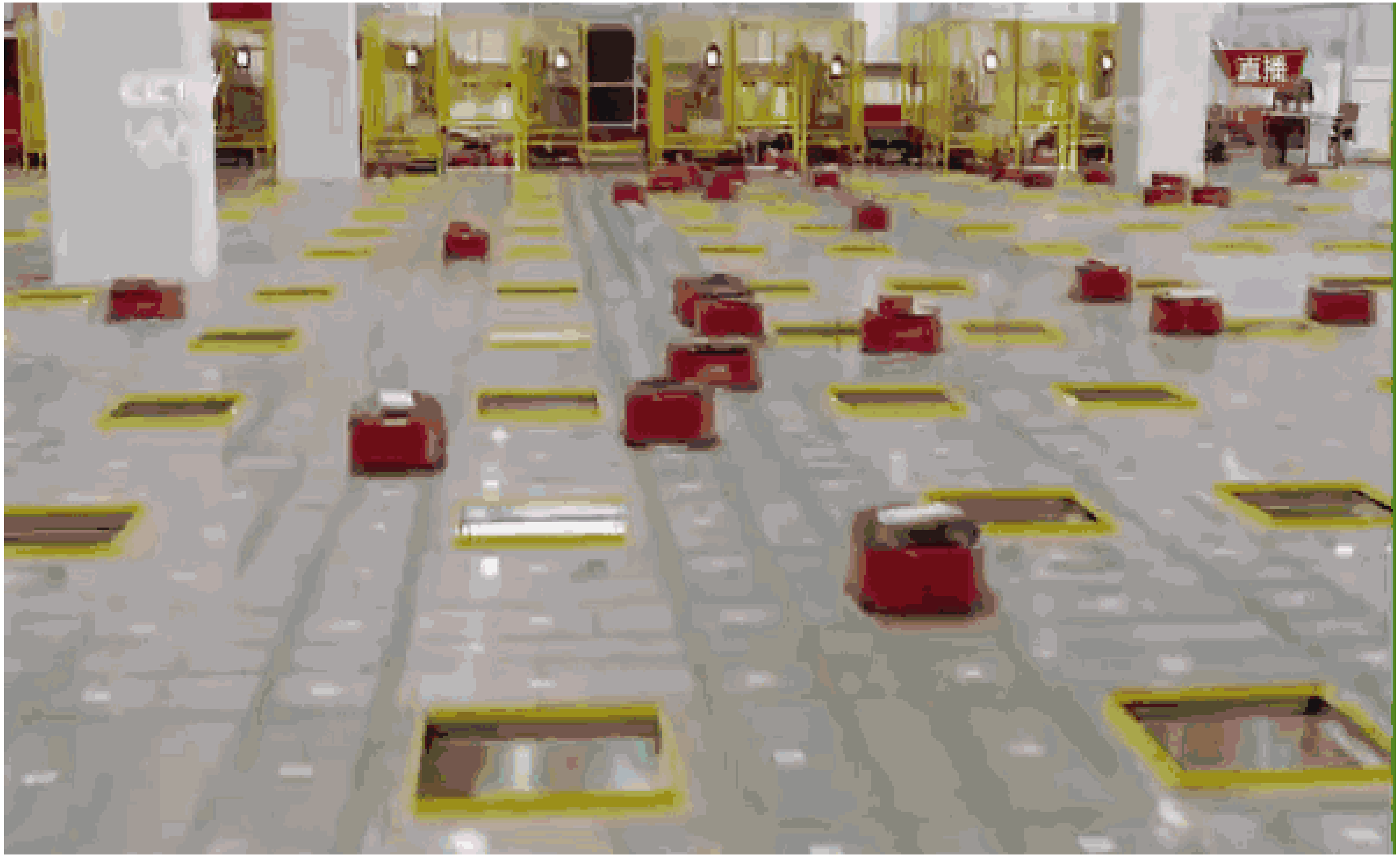
Stress Testing for 11/11 of an Unmanned Warehouse

A case from Cardinal Operations on 2017 Data

On 11/11, by adding 14 AGVs and 3 Workstations + Optimization, the system managed to boost the productivity by 283%.

2500 m^2 | 10000 orders/day | 335257 Pieces (Inventory)

Daily	11/11
AGVs：36	AGVs：50
Workstations：4	Workstations：7
Pieces/h per Workstation：600	Pieces/h per Workstation：1200
Outbound Quantity：20000	Outbound Quantity：56545



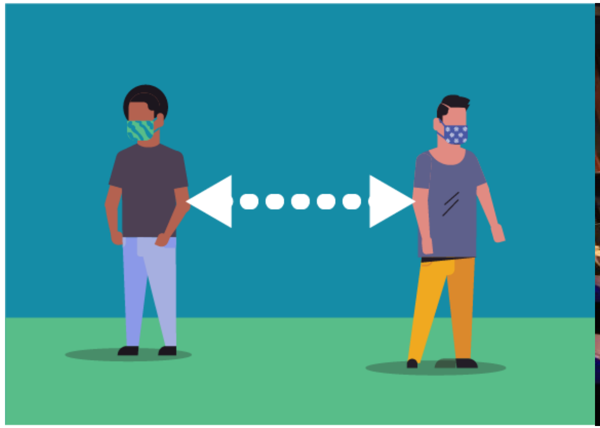
Beijing 8-minute show at closing ceremony of PyeongChang Winter Olympics

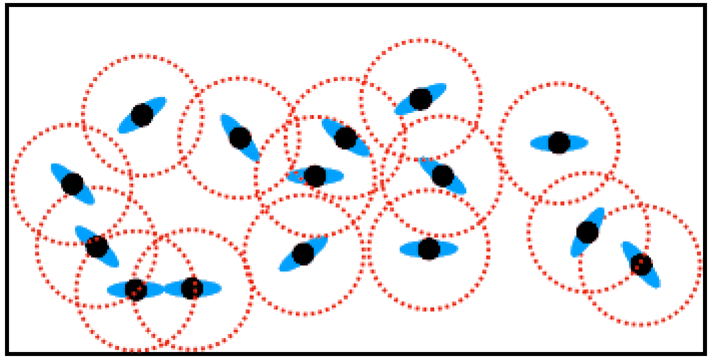
—— China’s First Unmanned Warehouse of JD.com

New Norm: Operation/Optimization under Social Distance Constraints

Semidefinite Programming Solution

Social Distancing: Mathematical Implication and Solution



Accommodate people in finite space with sufficient distance from each other

Math Representation of SD

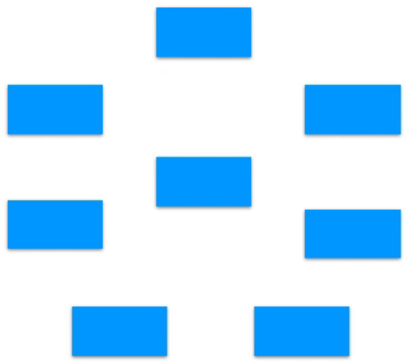
$$||x_i - x_j||_2 \geq 6$$

nonconvex in position variables x.

Two Scenarios

People accommodated Discretely

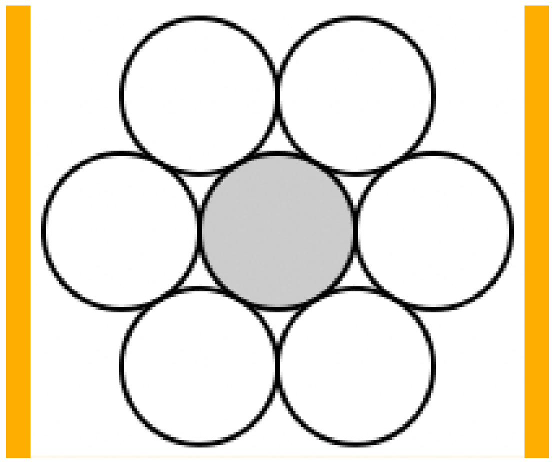
- Indoors: theater, restaurant, school, etc.
- Combinatorial/discrete optimization
- The Max-Independent Set Model



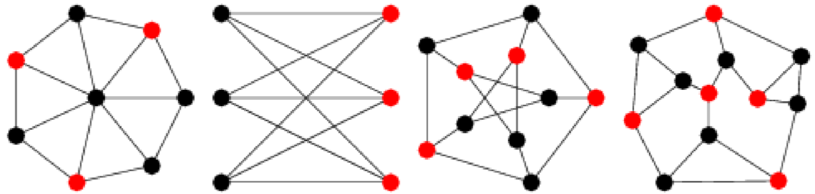
Tables at a restaurant

People accommodated Continuously

- Outdoors: beach, square, etc.
- Non-convex continuous optimization
- The Kissing Problem



Max-Independent Set Problem



Given seats in a theater, find an arrangement of maximum seats such that no two seats are within the unsafe distance

Given a graph G, find a subset of vertices of maximum cardinality such that no two vertices in the subset are directly connected

IP Formulation of Max-independent Set Problem

$$\begin{aligned} \max_{x} \quad & \textstyle\sum_i x_i \\ \text{s.t.} \quad & x_i + x_j \le 1 \quad (i,j) \in E \\ & x_i \in \{0,1\} \quad \forall i \end{aligned}$$

Whether a vertex appears in the set

SDP Relaxation

$$\begin{aligned} \max_{X \in \mathbb{S}^{n\times n}} \quad & J \bullet X \\ \text{s.t.} \quad & I_n \bullet X = 1 \\ & X_{i,j} = 0 \quad (i,j) \in E \\ & X \succeq 0 \end{aligned}$$

Max-independent set problem is NP Hard,

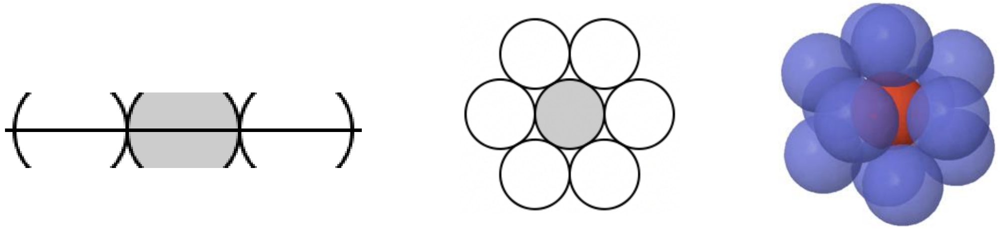
but approximation is possible on planar graphs

(Chiba, Norishige , T. Nishizeki , and N. Saito, 1982)

SDP relaxation can be applied to find upper bound

(Lovasz, and L, 1979)

Kissing Problem



Given a unit sphere, find the maximum number of nonoverlapping unit spheres in d dimension that "kiss" the center sphere

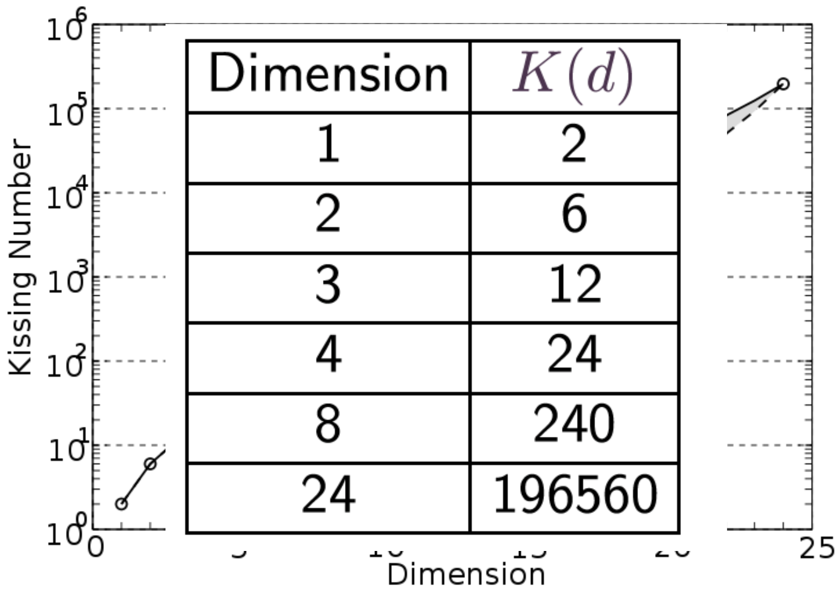
Measure distance by Euclidean Norm

$$\text{dist}(x_i, x_j) = \|x_i - x_j\|_2$$

Need for safe distance δ

$$\|x_i - x_j\|_2 \geqslant \delta$$

- No closed form solution for dimension *d* (Kucherenko, et al, 2007)
- Can be formulated and relaxed as SDP feasibility problem for a given number of spheres
- Upper bounds can be provided
- For Quadratic Optimization see Luo et al. 2010



Dimension	$K(d)$
1	2
2	6
3	12
4	24
8	240
24	196560

The quadratic constraint is non-convex and results in hardness

New Extension: Humanized Arrangement?

What if we allow families/friends to sit together ?

- Potentially more people accommodated
- Independent Set fails to capture the extension
- Max-Independent Set Problem with Clusters
- Can be formulated as 0-1 integer programming

IP Formulation for Seat Assignment

$$\max_{\mathbf{x}} \quad \sum_{i=1}^{m} \sum_{p=1}^{n} x_{ip}$$

subject to

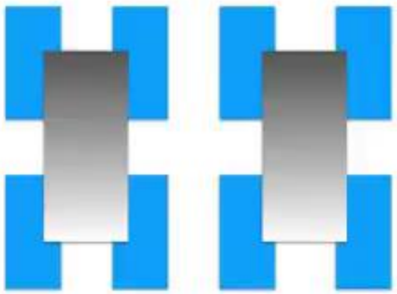
$$\sum_{i=1}^{m} x_{ip} \leqslant 1 \qquad \forall p$$

$$\sum_{p=1}^{n} x_{ip} \leqslant 1 \qquad \forall i$$

$$x_{ip} + x_{jq} \leqslant 1 \qquad \forall \text{ strangers } i, j, \text{ close seats } p, q$$

$$x_{ip} \in \{0, 1\} \qquad \forall i, p$$

Whether person i is assigned to seat p



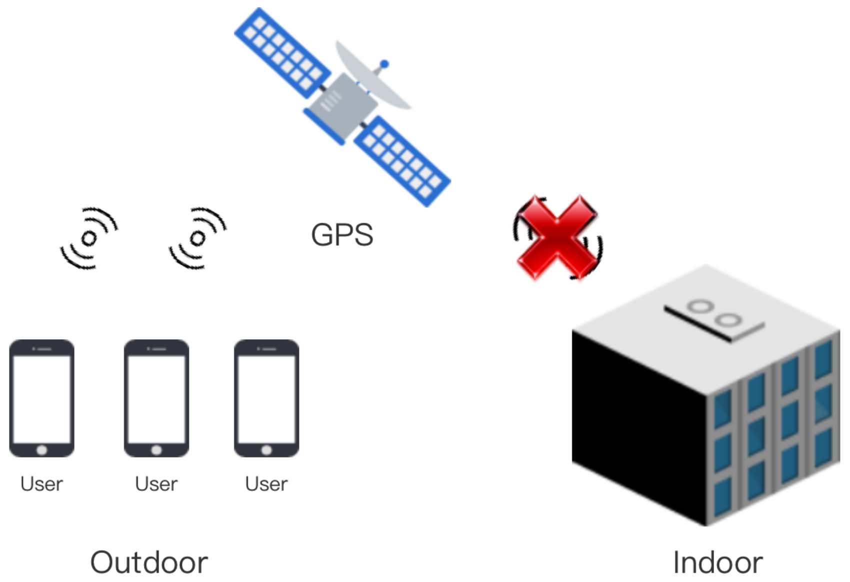
Seats at a restaurant

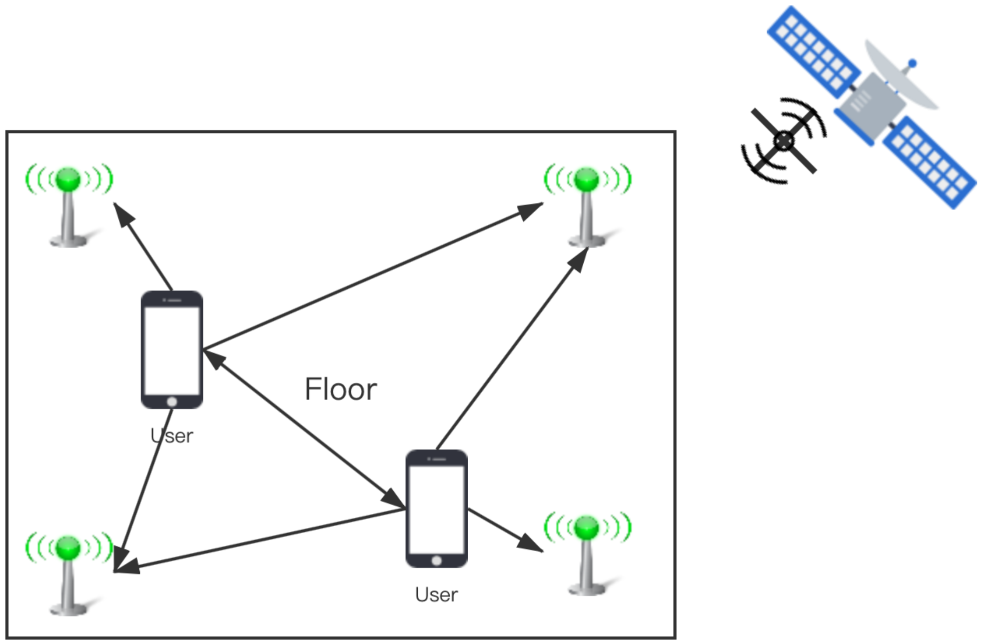
Intuitive and Heuristic Approach?

Indoor GPS and Tracking by Sensor Network Localization for Contact-Tracing

Indoor Trajectory Tracking Sensor Localization Problem

Identify trajectory during pandemic





Outdoors: Using GPS

Indoors: Using Indoor Signal Anchors

Sensor Network Localization (SNL)

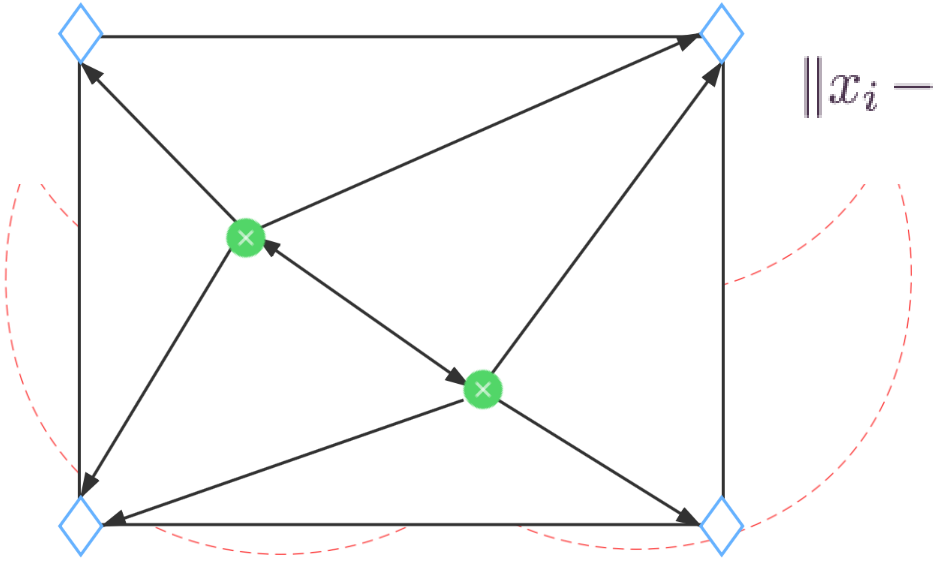
Given m anchor points $a_1, \ldots, a_m \in \mathbb{R}^d$ whose locations are known and n sensors points $x_1, \ldots, x_n \in \mathbb{R}^d$ whoses locations we wish to determine. Furthermore, we are given the Euclidean distance $\bar{d}_{ij}$ between a_k and x_j for some k, j and d_{ij} between x_i and x_j for some i, j. The Sensor Localization Problem is to find a realization of $x_1, \ldots, x_n$ such that

Distance between anchor and sensor

$$\|a_k - x_j\|^2 = \bar{d}_{kj}, \bar{d}_{kj} \text{ specified}$$

Distance between sensor and sensor

$$\|x_i - x_j\|^2 = d_{ij}, d_{ij} \text{ specified}$$



- Hard to track even for $d = 1$
- Can be formulated and relaxed as SDP feasibility problem (Biswas and Y 2004; So and Y, 2007)

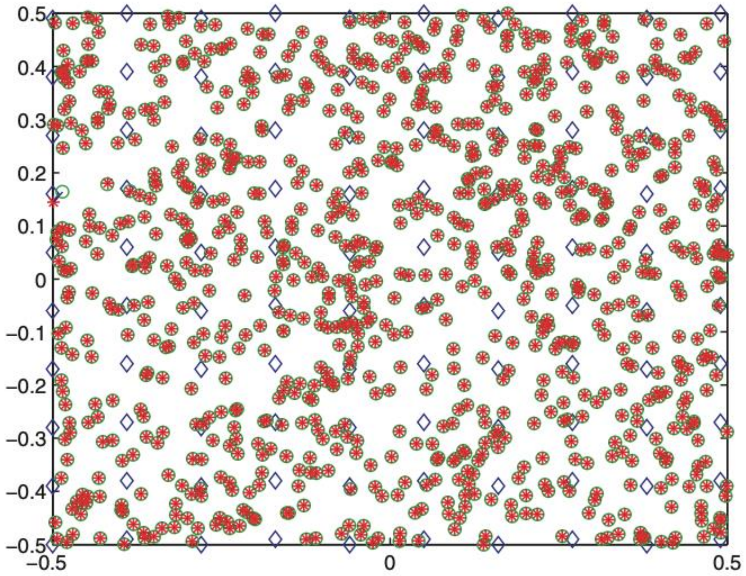
SNL Solution by Semidefinite Programming

Let $X = [x_1, \ldots, x_n] \in \mathbb{R}^{d \times n}$

$$
\begin{aligned}
\min_{X,Y} \quad & 0 & \\
\text{s.t.} \quad & e_{ij}^\top Y e_{ij} = d_{ij}^2 & , \bar{d}_{kj} \text{ specified} \\
& (a_k; e_j)^\top \begin{pmatrix} I_d & X \\ X^\top & Y \end{pmatrix} (a_k; e_j) = \bar{d}_{kj}^2 & , d_{ij} \text{ specified} \\
& Y = X^\top X &
\end{aligned}
$$

Relax $Y \succeq X^\top X$ and let $Z = \begin{pmatrix} I_d & X \\ X^\top & Y \end{pmatrix} \succeq 0$

$$
\begin{aligned}
\min_{Z} \quad & 0 & \\
\text{s.t.} \quad & Z_{1:d,1:d} = I_d & \\
& (\mathbf{0}; e_{ij})(\mathbf{0}; e_{ij})^\top \bullet Z = d_{ij}^2 & , d_{ij} \text{ specified} \\
& (a_k; e_j)(a_k; e_j)^\top \bullet Z = \bar{d}_{kj}^2 & , \bar{d}_{kj} \text{ specified} \\
& Z \succeq 0 &
\end{aligned}
$$

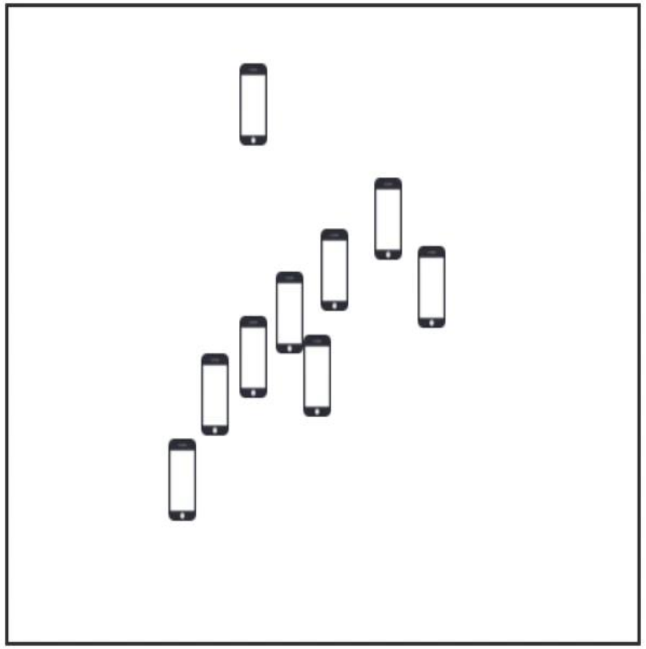


- Relaxation is tight for uniquely localizable graph
- Solution is too slow
- Can be acceleration by edge-based SDP (Wang et al. 2008)

Real-time Sensor Localization Problem

(Naber & Ye 2020, Wang & Ding 2008)

- Work under milder conditions
- A real-time version of sensor localization problem
- Retrieve moving trajectory and predict
- A combination of ESDP for tracking and Gradient Method for error minimization



Edge-based Relaxation

$$\begin{array}{lll} \min\limits_{Z} & 0 & \\ \text{s.t.} & Z_{(1,2)} = I & \\ & (\mathbf{0}; e_{ij})(\mathbf{0}; e_{ij})^\top \bullet Z = d_{ij}^2 & , d_{ij} \text{ specified} \\ & (a_k; e_j)(a_k; e_j)^\top \bullet Z = \bar{d}_{kj}^2 & , \bar{d}_{kj} \text{ specified} \\ & Z_{(1,2,i,j)} \succeq 0 & d_{ij} \text{ specified} \end{array}$$

Objects move subject to linear differential equation $\frac{dX(t)}{dt} = AX(t) + Zt + C$

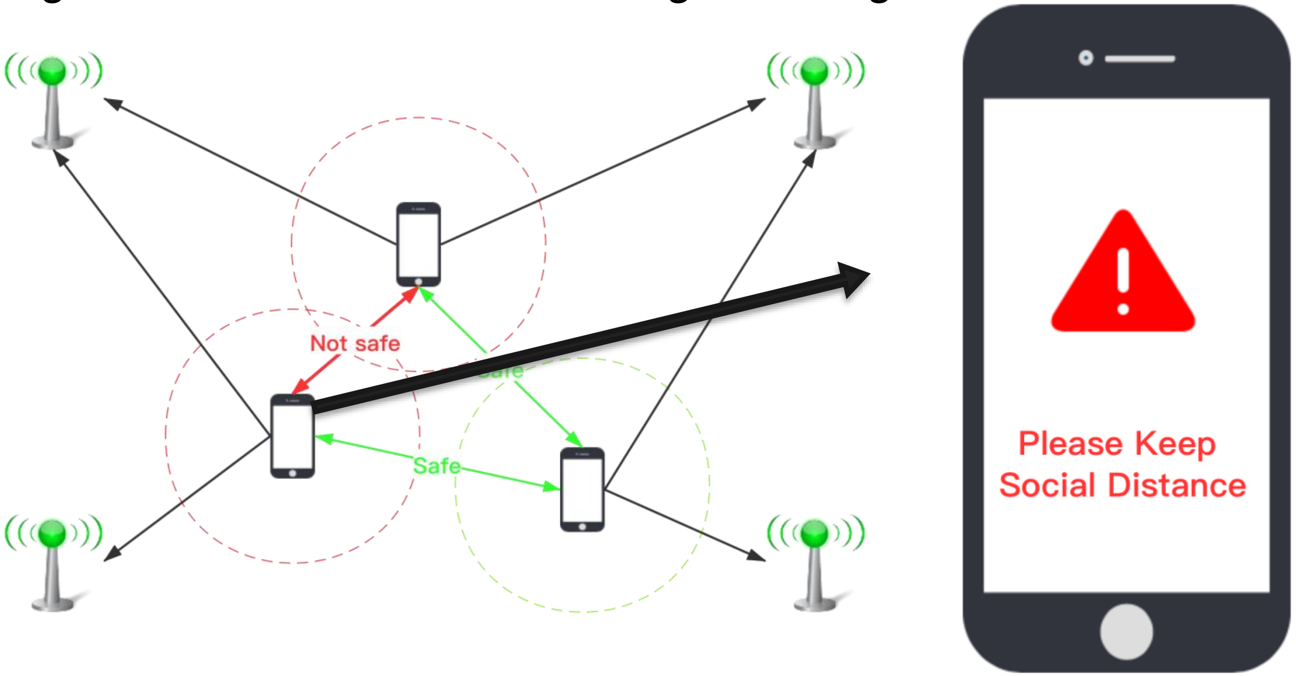
$$\begin{array}{lll} \min_\gamma & \sum_{i=1}^n \eta_i \gamma_i^2 & \\ \text{s.t.} & \gamma_i \geq \frac{X(t_i) - X(t_{i-1})}{t_i - t_{i-1}} - AX(t_i) - Zt_i - C & , \forall i \\ & \gamma_i \geq -\frac{X(t_i) - X(t_{i-1})}{t_i - t_{i-1}} + AX(t_i) + Zt_i + C & , \forall i \\ & A, C, Z \in \Lambda & \end{array}$$

A least-squares problem

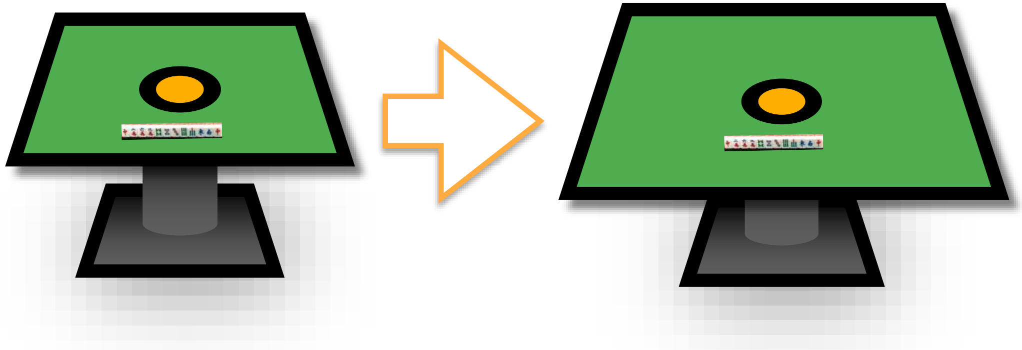
Simple Distance Checking and Enforcing I

Primitive distancing enforcement

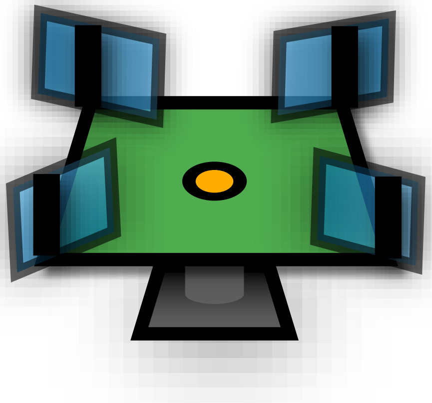
High-tech Solution for Distancing Alarming



Card-Play Table Redesign

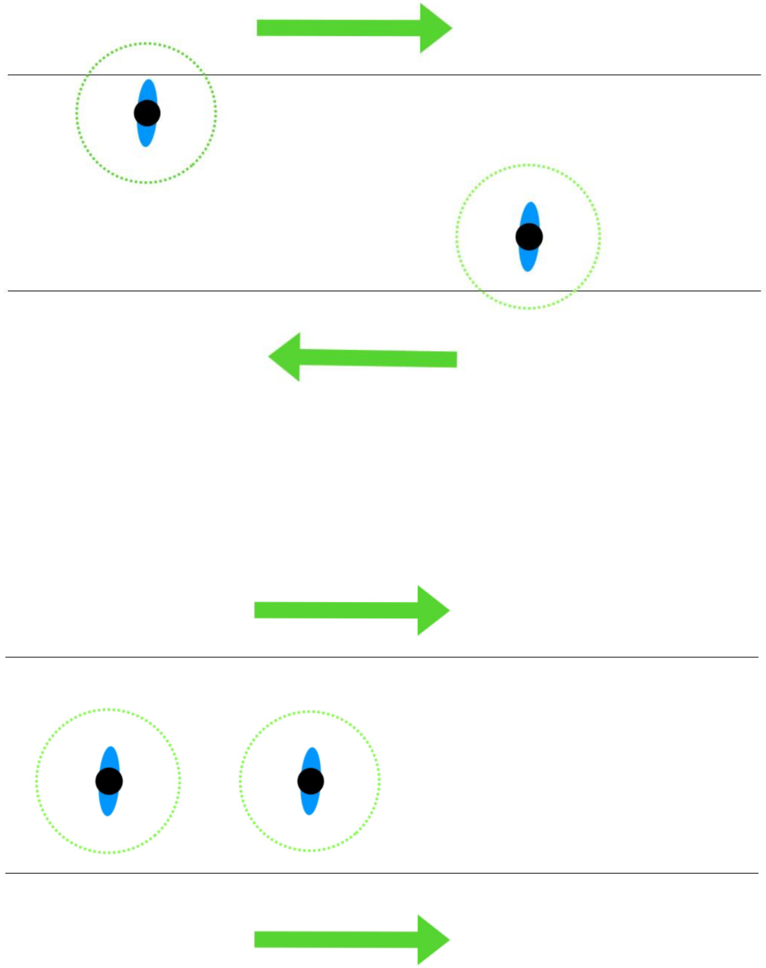


OR



Simple Distance Checking and Enforcing II

One-Way or Two-Way for pedestrian environments?

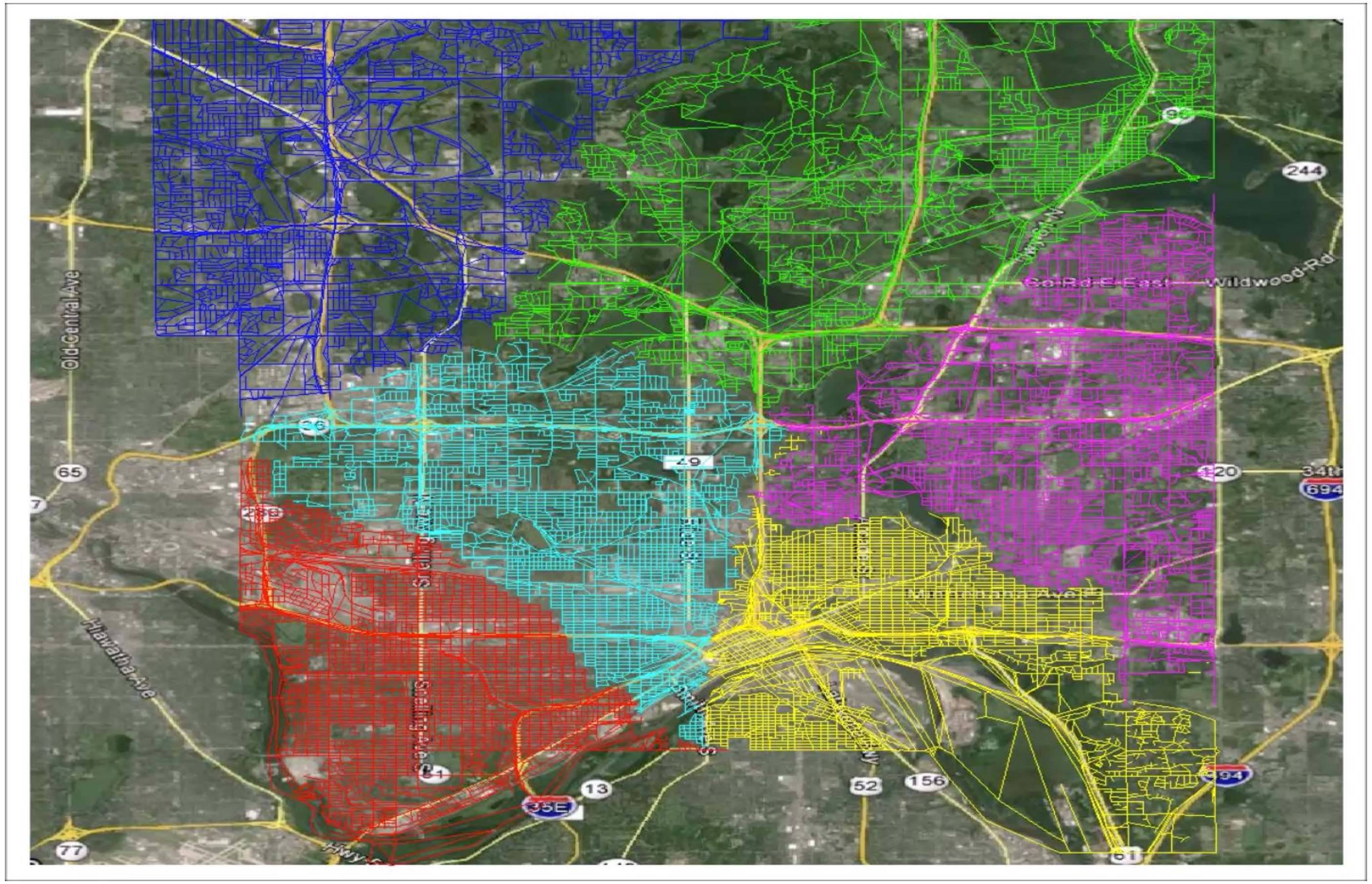


Dynamic and Equitable Region Partitioning for Hospital Services

Computational Geometry Solution

Dynamic Hospital Service Region Partitioning

Computational Geometry Theory/Algorithm



Input Data:

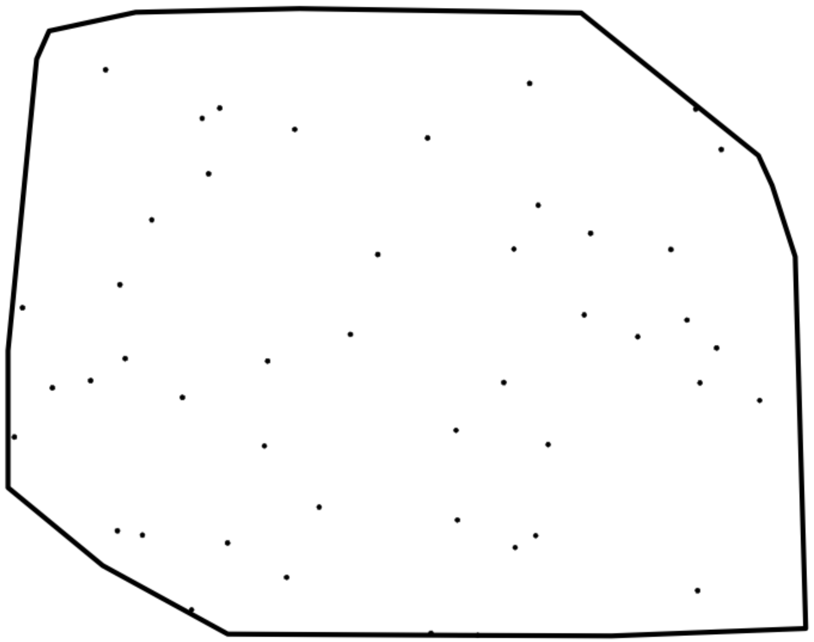
- Hospital location and capacity
- Pandemic density distribution prediction

Planning:

Partition the city into multiple regions such that

- Each region has a hospital nearby
- Each hospital will not be overrun
- Can be easily adjusted by input data change

Plane-Geometry Problem Statement and Theorem



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

– Travel convenience

Each sub-region contains one point

– Service center for the region

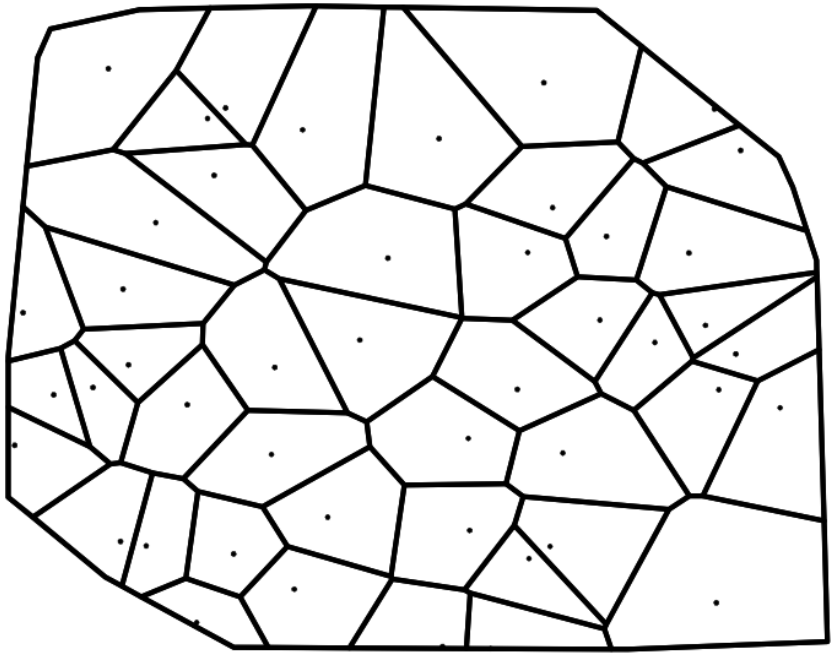
All sub-regions have equal area

– Load balance

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

Armbruster, Carlsson and Y, 2012

Related Problem: Voronoi Diagram

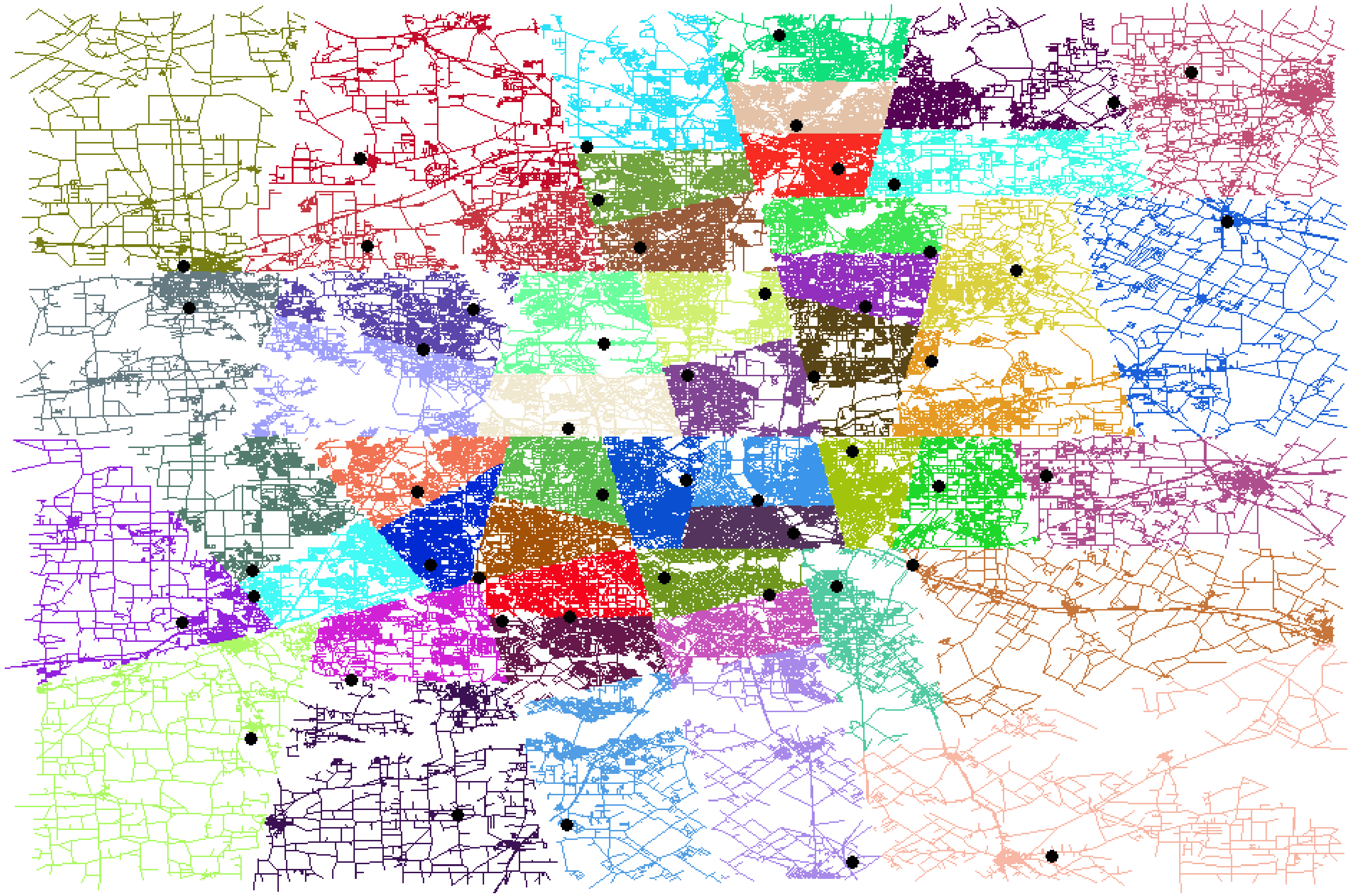


Voronoi Diagram: draw a middle perpendicular line between every two points.

The partition satisfies the first two properties (each sub-region is convex and contains one point), but the sub-regions have different areas.

In practice one can adjust the boundary to achieve the third property

Equitable Partition with Nonuniform Density Partition



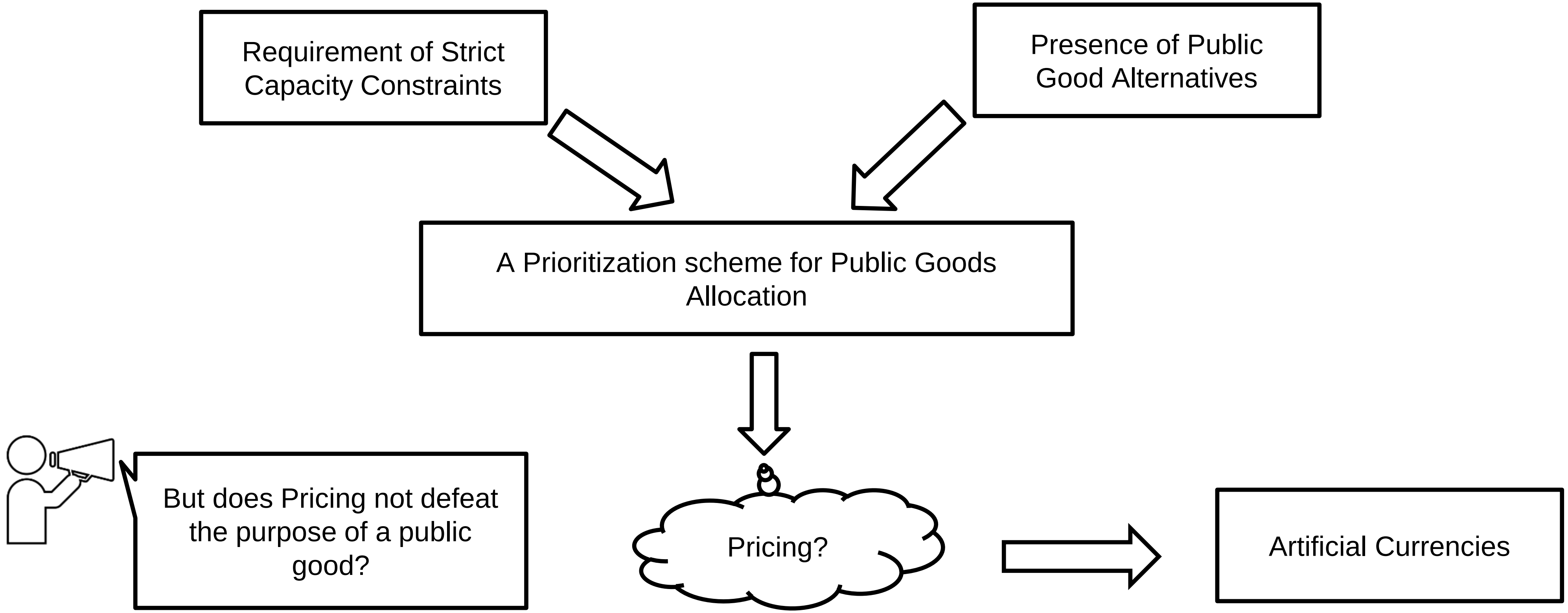
Efficient Public Goods Allocation under Tight Capacity Restriction via Market Equilibrium Platforms

The Need for Efficient Public Good Allocation-Market Equilibrium Model

Either open: An overcrowded open beach

Or closed: a completely empty beach generating no value to society

A Fisher-Market-Based Mechanism Design



How can we design a non-monetary market mechanism and still guarantee a socially efficient allocation that is desirable for all consumers?

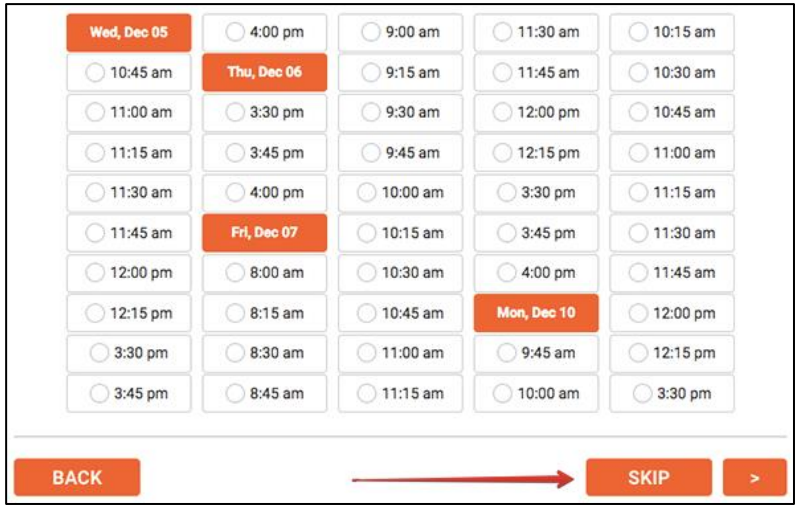
An overview of our Solution

Step 1: Create Schedule for Use of Public Good, e.g. Beach

Step 2: Prices assigned for different times of use

Step 3: Transfer Electronic Coupons

Step 4: Users Purchase time of use permits



Enforcement: Park rangers/ entrance booths to check for permits

Pathway to Setting Prices

- **Addressing Step 2: How can we set appropriate prices for different times of use given that the mechanism does not have complete information on everyone's utilities or constraints?**

Phase 1: Initialize Prices with some prior knowledge of utilities and constraints

Phase 2: Mechanism Learns Utilities and Constraints of Consumers

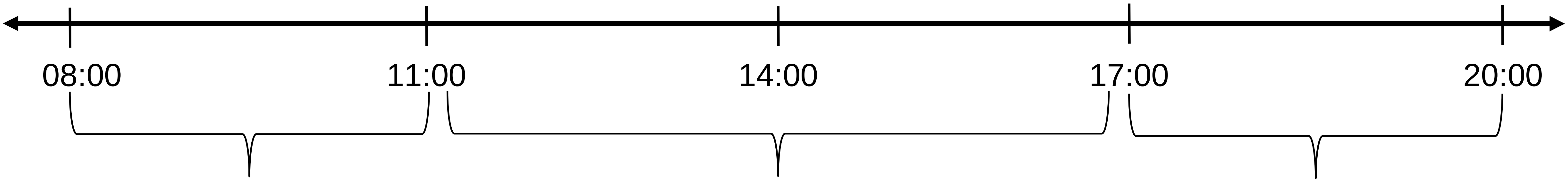
Phase 3: Iteratively Update Prices based on learnt information of consumer behavior

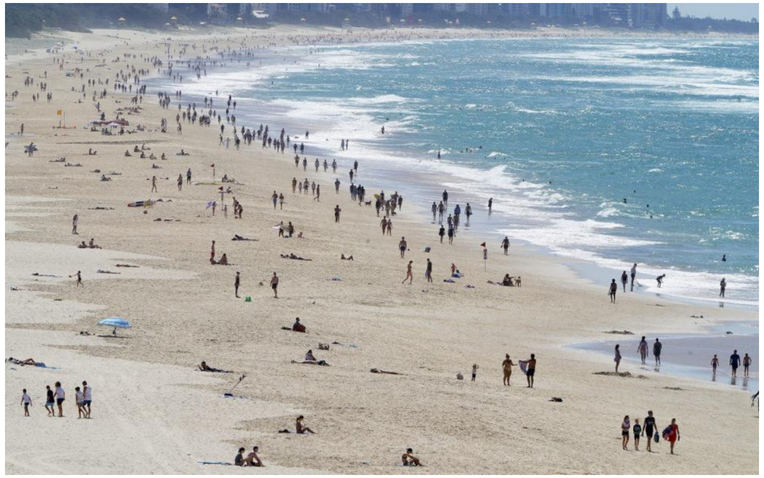
Through our proposed pricing scheme that generalizes the Fisher Market framework, customers will purchase permits in a "controlled" manner resulting in neither overcrowded or underused public resources

Ideal Outcome

To achieve an Intermediate between the two extreme scenarios, open or closed, through "Time of Use" goods

Create different time zones and people purchase permits to use the beach at one time-zone so that the population density on beach can be upper limited





The Fisher-Market with Budget and Physical Constraints
Centralized vs Decentralized Decision Making

- In the social optimization problem

Choose: $\lambda_i = \sum_t r_{it}$

- **Main Result 1: KKT Equivalence of Social and Individual Optimization Problems**

The dual of capacity constraints is the equilibrium price vector

- How do we obtain the perturbation parameter

Fixed-Point Iterations

Individual Optimization Problem:

$$\max_{\mathbf{x}_i} \sum_j u_{ij} x_{ij}$$
$$\text{s.t.} \quad \mathbf{p}^T \mathbf{x}_i \leq w_i$$
$$A_t^T \mathbf{x}_i \leq 1, \forall t \in T_i$$
$$\mathbf{x}_i \geq \mathbf{0}$$

$$\Updownarrow$$

Social Optimization Problem:

Budget Perturbation

$$\max_{\mathbf{x}_i} \sum_i (w_i + \lambda_i) \log(\sum_j u_{ij} x_{ij})$$
$$\text{s.t.} \quad \sum_i x_{ij} \leq \bar{s}_j, \forall j \in [M]$$
$$A_t^T \mathbf{x}_i \leq 1, \forall t \in T_i, \forall i \in [N]$$
$$x_{ij} \geq 0, \forall i, j$$

r_{it}: Dual Variable of Physical Constraint

Jalota, Pavone and Y, 2020

Test and "Time of Use" for School/Class?

- Create pods/clusters. Children/students form groups of about 12 and stay together all day, avoiding contact with other pods. **Test every group as ONE example for possible infection (if positive then test individuals in the group) before schools open.**

- Partial days. Keep classed smaller by splitting the children into groups. For example, one group comes to school one day and the other group the next day. The rest of the work is done at home. The split could also be morning/afternoon.

- Split classes into either Zoom or In-Person. Also create pods for university in-person classes for course discussions and team projects.

Identifications and Protective Measures for High-Risk Groups in Pandemics

Protective Measures and Statistical Learning

The purpose of protective measures is to

‘Flatten the curve’
which means:

- Relieve each day’s medical pressure
- Not to eradicate the virus
- But reduce the rate of death

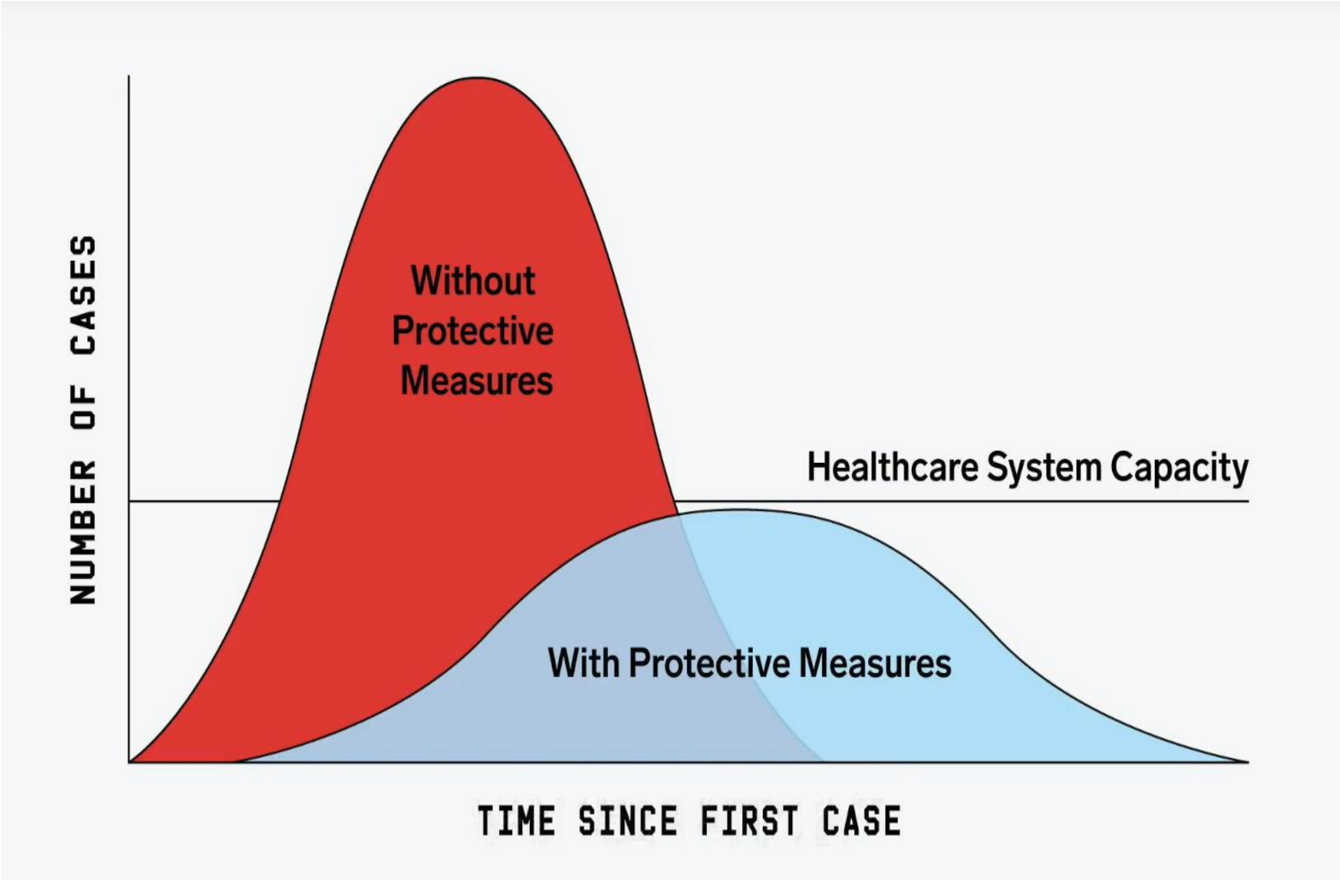
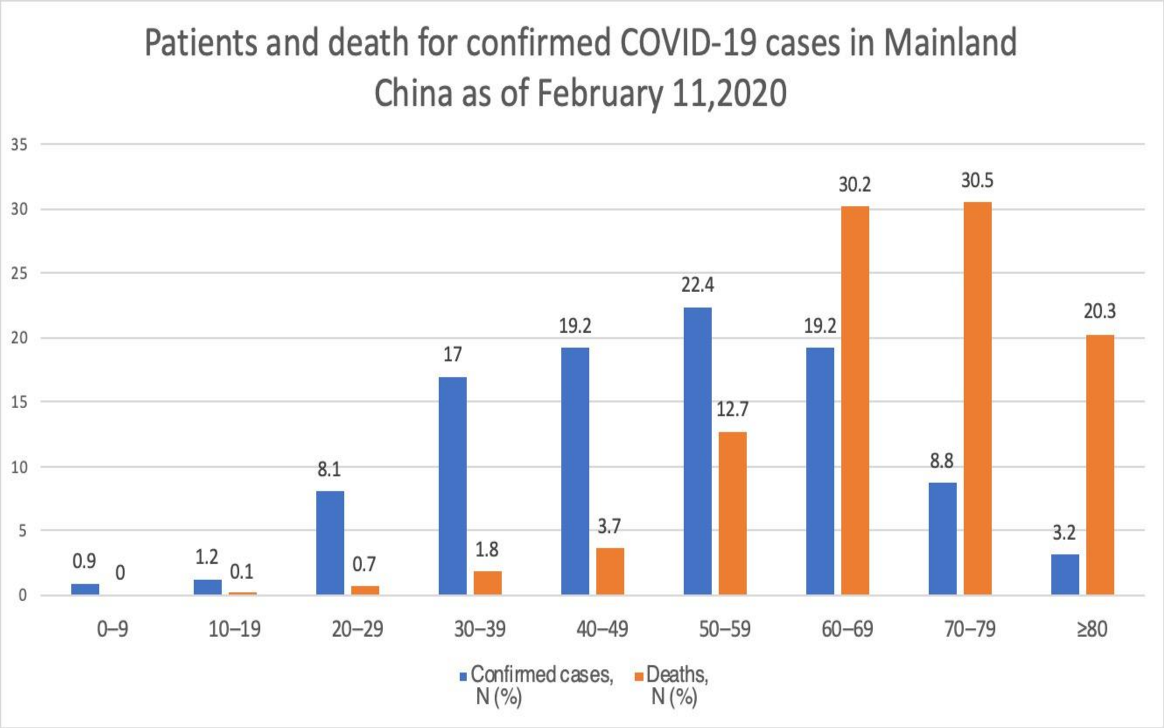


Figure: The expected result with protective measures

What would the measures be?
we need some characteristics of the virus pandemic

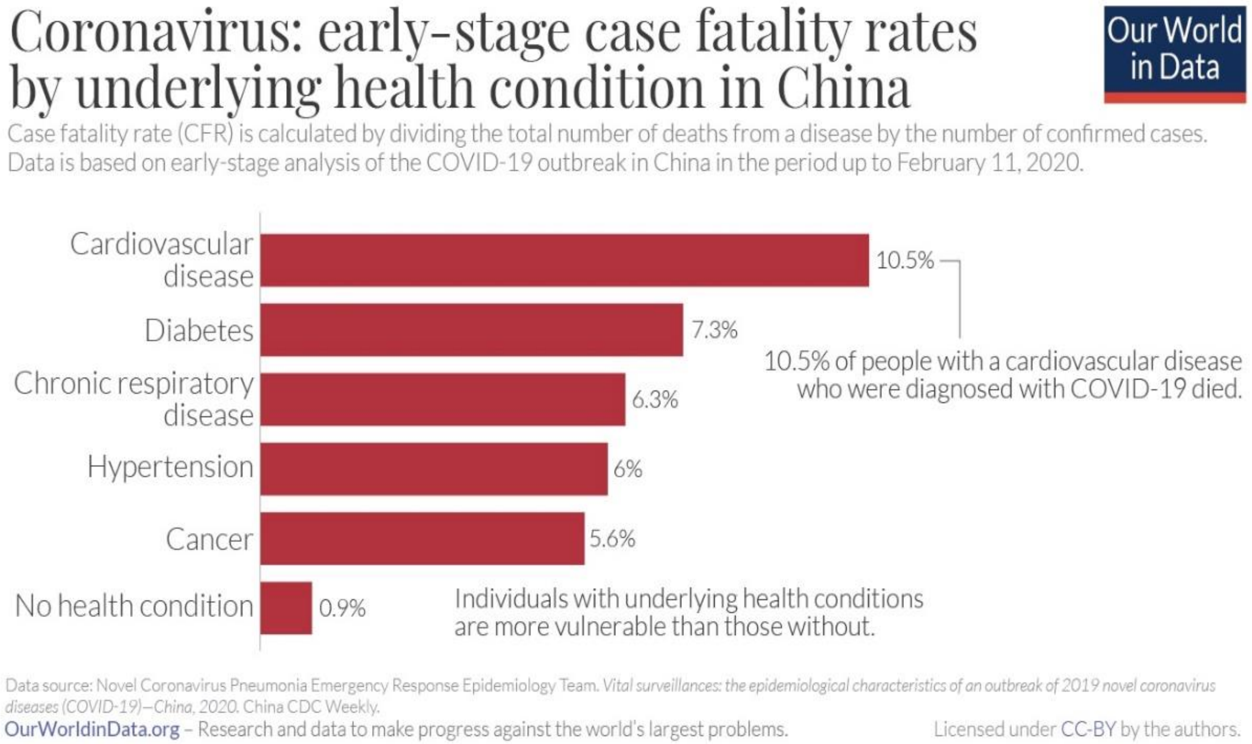
Death Risk Among Different Age-Groups

Age Group



Age	Percentage of confirmed cases(%)	Percentage of deaths(%)
Under 50	≈50	≈5
Over 50	≈50	≈95

Underline Health Conditions



Death Risk with Diseases	Cardiovascular	Diabetes	Chronic respiratory	Hypertension	Cancer
Ratio to normal	≈12	≈8	≈7	≈7	≈6

Calculate the Death Probabilities using the Bayes Formula
Statistical Learning

- Let A and C denote the event of $death$ and $survival$ respectively, after being confirmed.
- $B/\bar{B}$ denotes the patient's age is over/under 50 years old.
- $P(A)$ denotes the case fatality rate which is around 7.7% according to official data.

Bayes formula:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|C)P(C)}$$

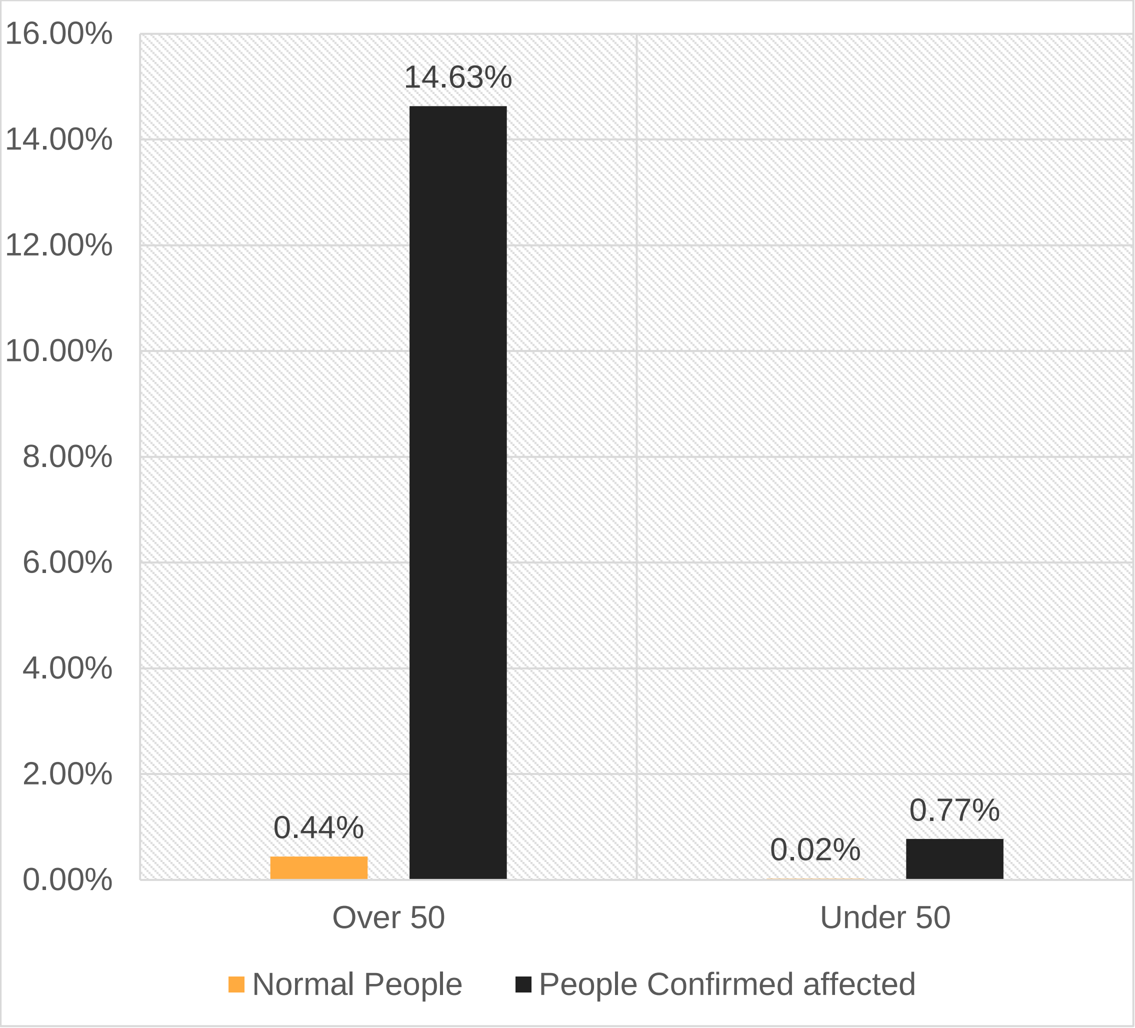
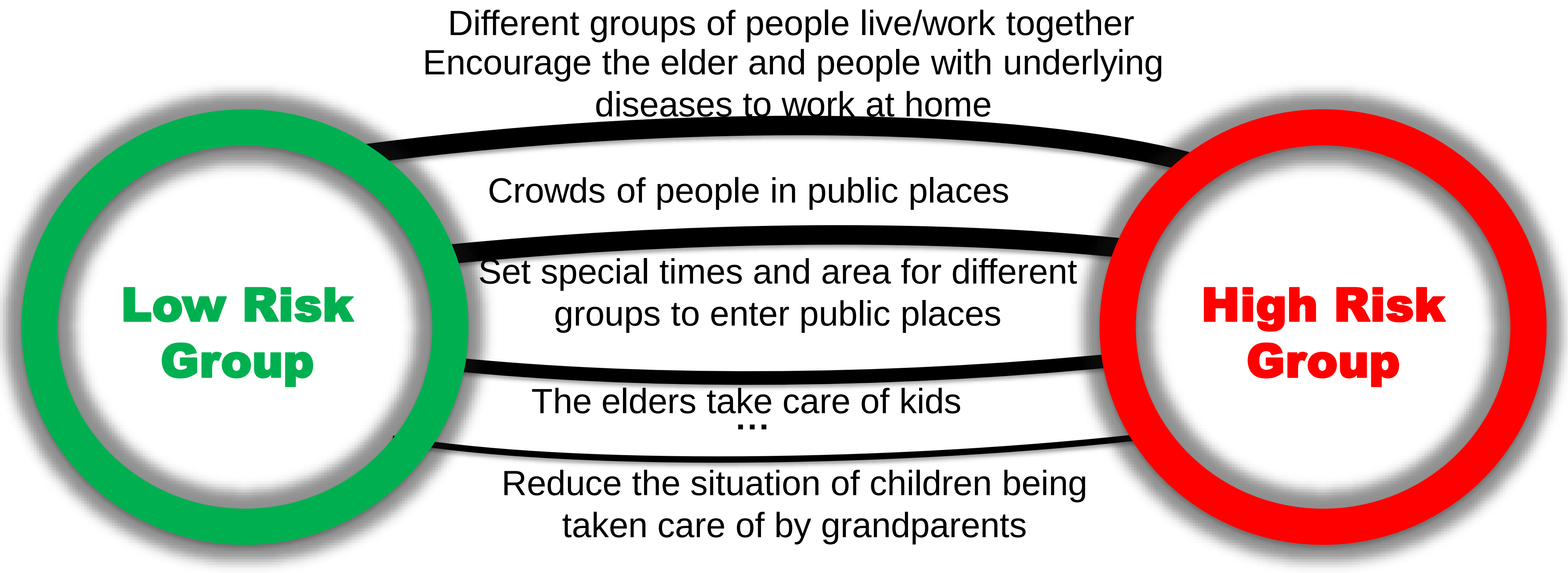


Figure: The result of Bayes formula

Protective Measures and Edge-Cut



- **We need to do something to cut the 'thicker edges' which indicate greater contact reductions?**
- **Once there is a vaccine (or more than one), prioritize who gets it first if supplies are limited?**
- **Should people all get the same vaccine or so some characteristics indicate that certain people would do better with a different one?**

More Public and Social Policies?

Encourage young graduates to become Uber/Didi drivers and replace the grandparents to pick up children or use of School Buses.

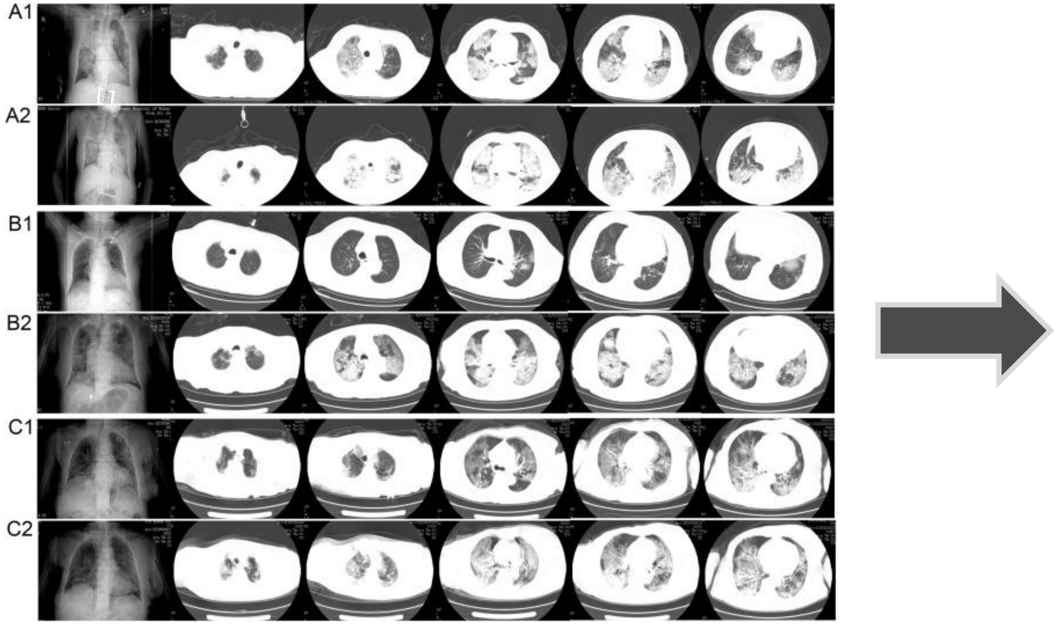
Arrange for kindergarten, primary school, middle school and university to resume classes as soon as possible.

Each person reduces the number of his/her public dance partners …

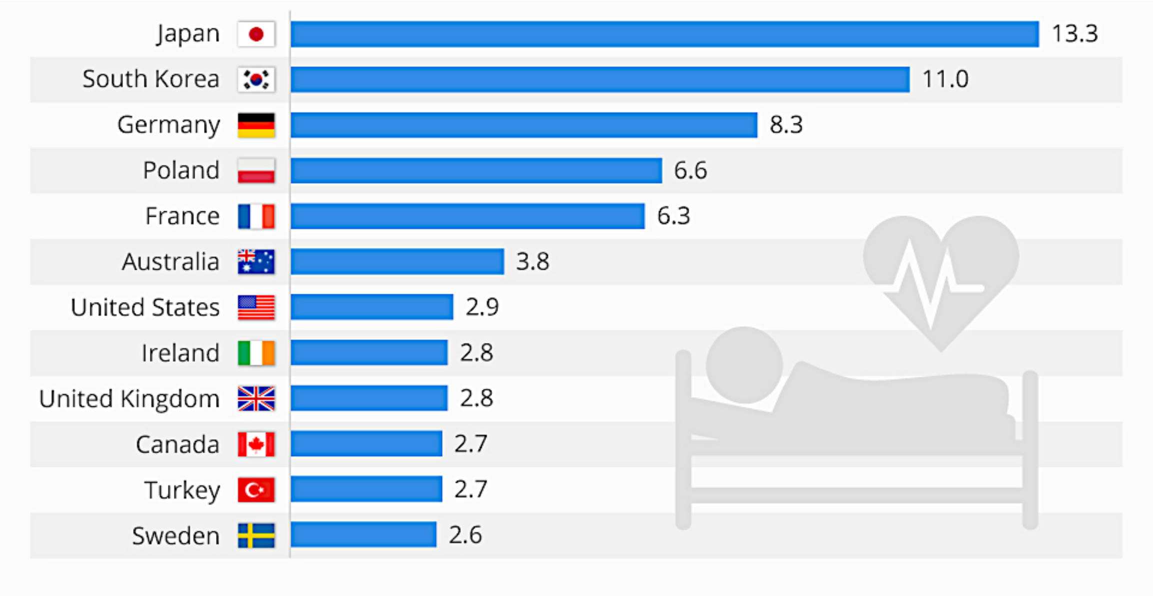
Machine-Learning (e.g., Logistic Regression) with Multiple Social Features to Reduce Pandemic Fatality

Logistic Regression with Social/Behavior Features

Machine learning can do a great favor in pandemic prevention. Lots of researchers predict the risk of death of a patient, given his clinical and pathological characteristics.



Chest CT scans of patients.

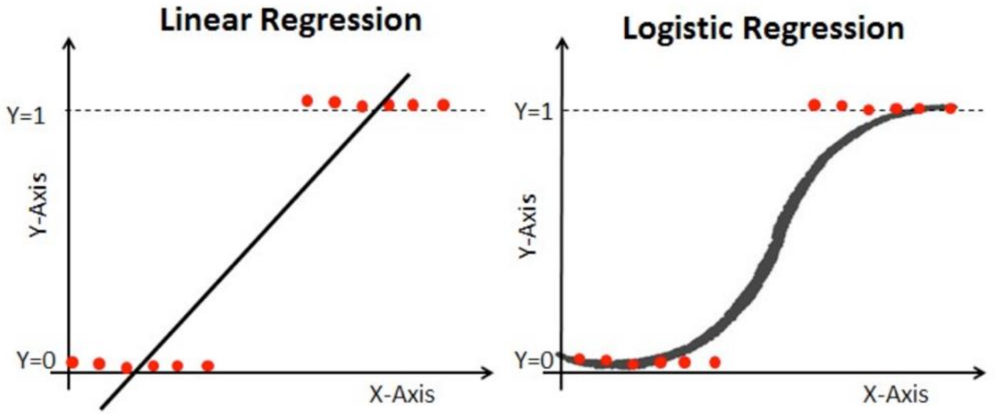


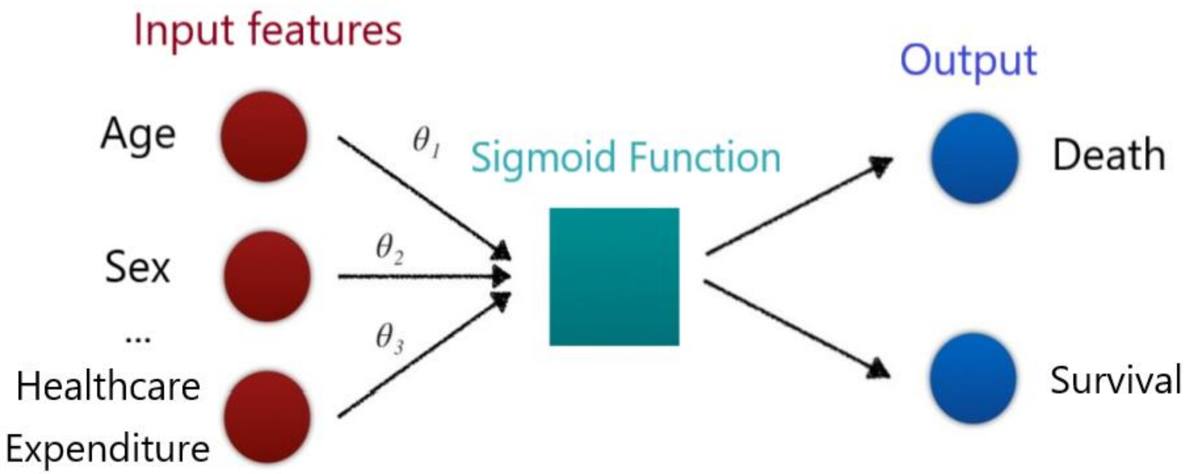
The Countries with the most hospital beds (per 1,000 of the population in selected countries).

We mainly focus on some social features such as the quality of regional healthcare systems, etc. Apply logistic regression and see what insight we can gain from it.

Logistic Regression Model

Logistic regression is a simple algorithm that can be used for binary classification tasks.





Sigmoid function: $$sigmoid(x) = \frac{e^x}{1 + e^x}$$

Actual label

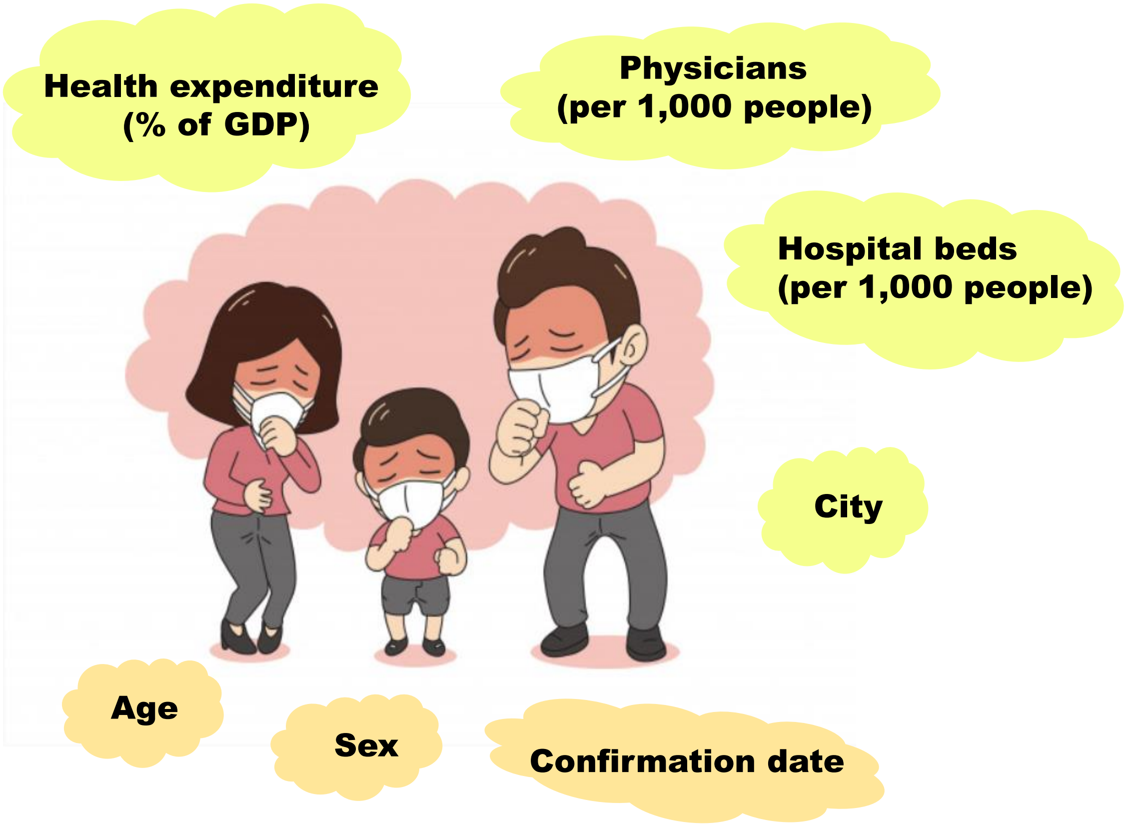
Cost function: $$Cost(y, \hat{y}) = -[y \log \hat{y} + (1 - y) \log(1 - \hat{y})]$$

$$J(\theta) = \frac{1}{m} \sum_{i=1}^{m} Cost(y_i, \hat{y}(\theta)_i)$$

Predicted label

Implementing Logistic Regression:

A preliminary case study from Cardinal Operations



Data Collection
- Defining the problem and assembling a dataset.
- The main dataset comes from nCoV-2019.

Data Preparation
- Wrangle data and prepare it for training.
- Deal with missing values, normalization…

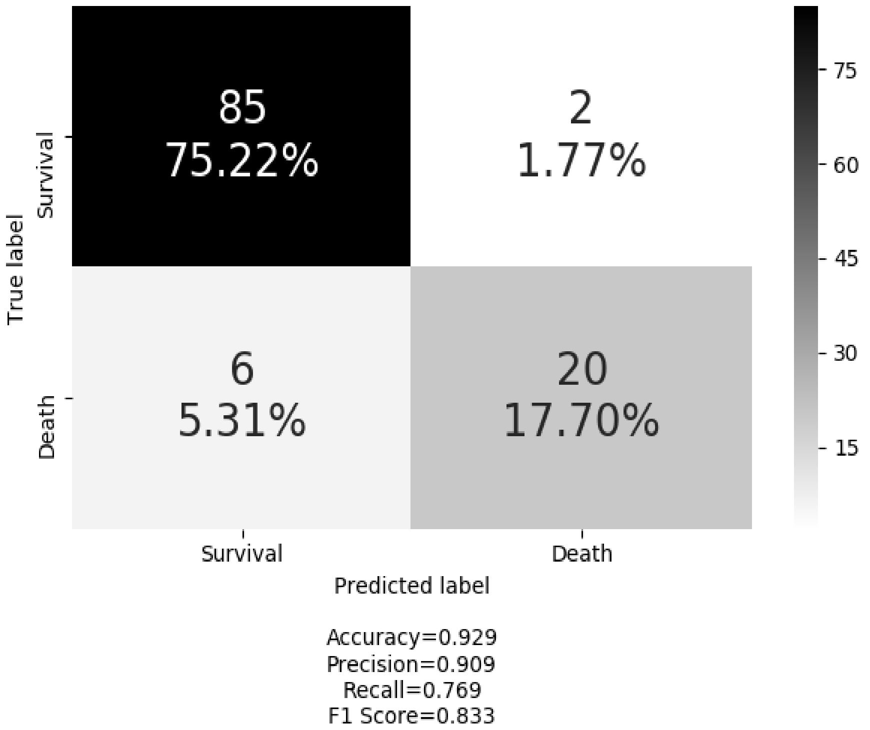
Train the Model
- Train logistic model with L2 regularization.

Evaluate the Model
- Check AUC, accuracy, confusion matrix etc.

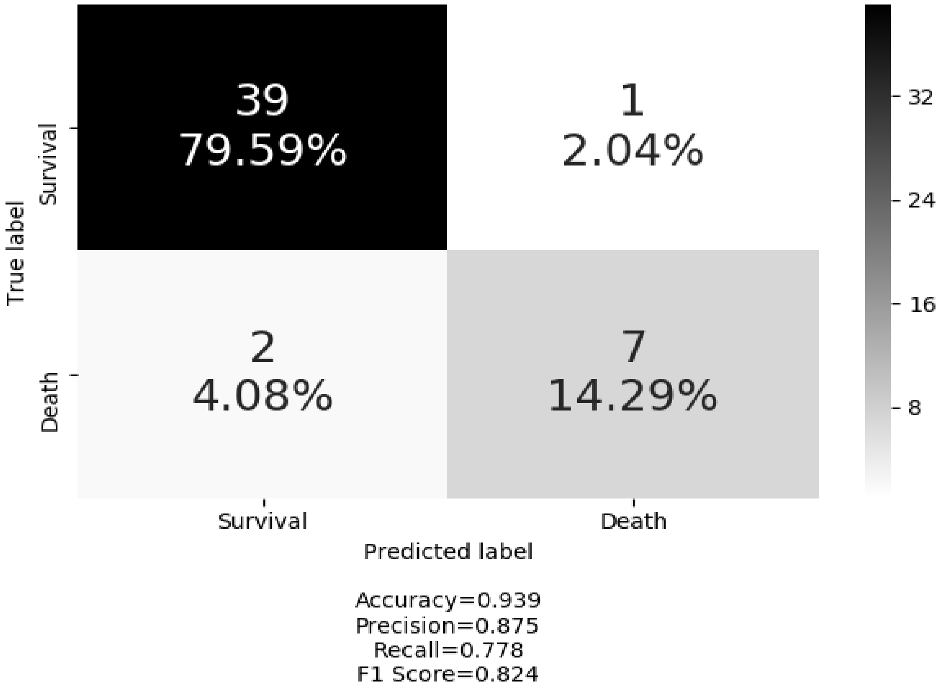
Gain insight
- What policies might protect people and prevent them from death?

Outcome and Performance I

Our model is able to achieve 94% accuracy and 87% AUC on test set.



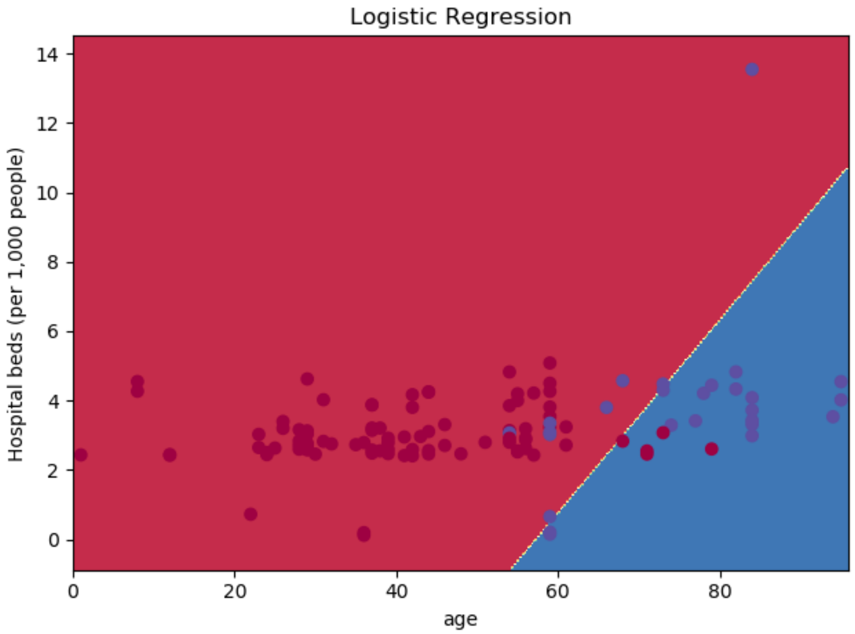
Metrics on the training set



Metrics on the test set

Outcome and Performance II

Next figures shows the decision boundaries and what helps to differentiate probabilities into $survival$ class and $death$ class. Patients in countries with more well-developed public healthcare systems is less at risk of death.



Decision boundary on Age and $Hospital\ beds\ (per\ 1{,}000\ people)$

Decision boundary on Age and $Physicians\ (per\ 1{,}000\ people)$

Countries should try to design and develop their healthcare facilities in accordance with their needs to defeating COVID-19.

Summaries for Mitigation of a Pandemic and Economy Reopening

- **Strategies: The goal is not to eradicate virus but to reduce the ultimate fatality rate by frequent sample testing, Transparent reporting and data bases, Vulnerable group identification, Health-care quality improvement…**
- **Policies: Differentiate high and low risk groups (such as work at home or office), Reward earlier retirement, Encourage college graduates to do social work, Social Distancing…**
- **Operations: Centralized and mobilized task-force, Centralized inventory management of materials/equipment, Algorithm-based planning and dispatching, Efficient public spaces/goods allocation…**
- **Methodologies: Based on Sciences/Technologies such as Mathematical Optimization, Computer Monte Carlo Simulation, Statistical and Machine Learning, Mechanism Designs, High-Tech Solutions…**
- **Reopen Economy: Low-risk groups return to work, Data-driven decision making rather than vision-driven, Short-term action rather than long-term planning, Lean and agile supply-chain, Machine Decisioning…**