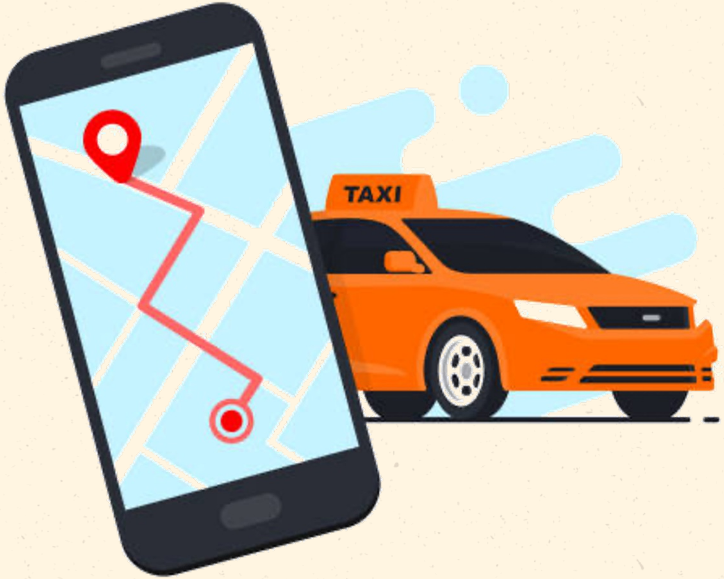
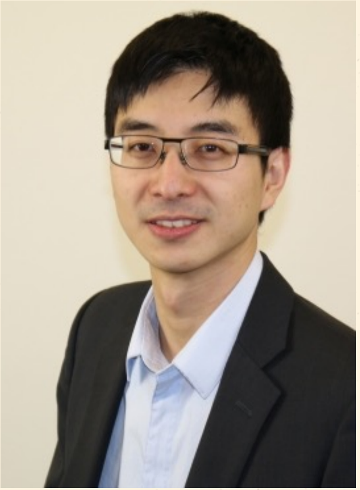
Dynamic Pricing and Matching for Online Marketplaces

Presenter: Sushil Mahavir Varma
5^{th} Year Ph.D. Student
Operations Research, Georgia Tech



Joint Work With...

Pornpawee Bumpensanti
Research Scientist
Amazon



He Wang
Assistant Professor
ISyE, Georgia Tech

Francisco Castro
Assistant Professor
Anderson, UCLA

Siva Theja Maguluri
Assistant Professor
ISyE, Georgia Tech

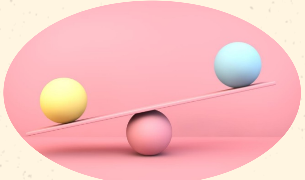
Gig Economy

Online Matching Platforms

$204 Billion in revenue in 2018 [Mastercard]
36% of the workers in the US join the gig economy [Gallup]

and many more….

Major Operational Challenges



Disparity of Supply and Demand

Unequal demand and supply agents in the market

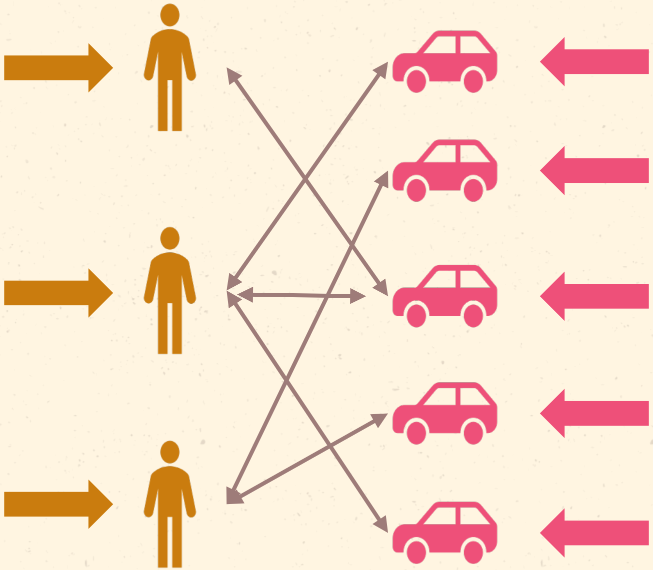
PRICING

Misalignment of Supply and Demand

Incompatible demand and supply agents in the market

MATCHING

Stochastic Network Viewpoint



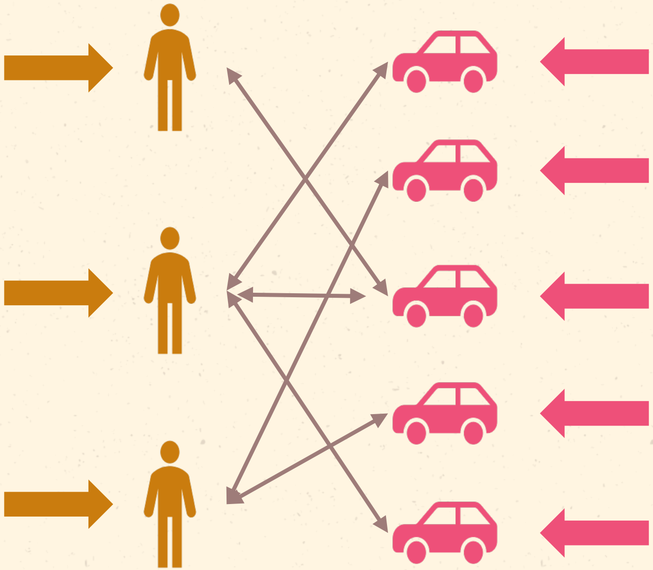
Type – Geographical location, normal/premium ride, etc.

Compatibility – Geographical proximity and matching preferences

Match – Disappear from the system instantaneously

Set **prices** and perform **matchings** that maximizes **profit** and minimizes the **delay**

Applications: Matching Networks



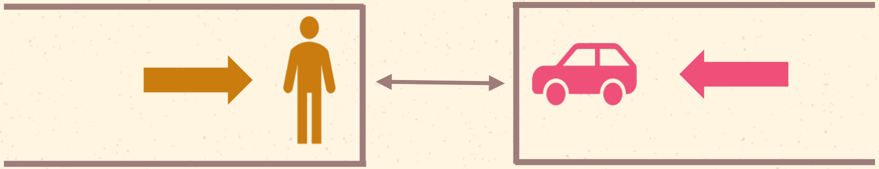
Ride-hailing

Payment Channel Networks

Quantum Networks

Set **prices** and perform **matchings** that maximizes **profit** and minimizes the **delay**

Technical Challenge: Simplest Case



$\lambda(q)$ $\mu(q)$

$\lambda > \mu$ – Transient
$\lambda < \mu$ – Transient
$\lambda = \mu$ – Null Recurrent

! Need **External Control** to make the system stable

Dynamic Pricing

Can be Analyzed in **Steady State**

Literature Survey

Many related models in the literature:

- **Bipartite Matching Models** [Adan, Weiss, 2012], [Caldeney et. Al. 2009], [Adan et. al. 2018], [Cadas et. al. 2019]
- **Matching Models** [Mairesse, Moyal, 2016], [Cadas et. al. 2020], [Moyal, Perry, 2017]
- **Matching Queues** [Gurvich, Ward, 2014]
- **Assemble to Order Systems** [Song, Zipkin, 2003], [Song, 1998], [Song et. al. 1999], [Song, 2002], [Song, Yao, 2002], [Plambeck, Ward, 2006], [Dogru et. al. 2010]
- **Other Related Models** [Anderson et. al.], [Akbarpour et. al. 2019]
- **Two-Sided Queues with few differences** [Hu, Zhou, 2018], [Nguyen, Stolyar, 2018], [Aveklouris et. al. 2021], [Ozkan, Ward, 2017], [Ozkan, 2020], [Blanchet, et. al. 2021]

Most models where the system is inherently unstable, only transience analysis have been done except

[Nguyen, Stolyar, 2018], [Blanchet, et. al. 2021].

We conduct more fine-tuned analysis

Table of contents

Part One

Dynamic Pricing and Matching for Two-Sided Queues

SMV, Bumpensanti, Maguluri, Wang

Operations Research 2022

Punchline: Near-optimal pricing and matching policy asymptotically

(with an $\eta^{1/3}$ ROC to the fluid upper bound)

Part Two

A Heavy Traffic Theory of Matching Queues

SMV, Maguluri

IFIP Performance 2021 (Best Paper)

Punchline: Phase transition for the limiting distribution of queue length, unlike classical queues

(in a heavy-traffic regime inspired by classical queues)

Model: Stochastic Matching Network

System operator sets the Price that determines arrival rates

Pricing

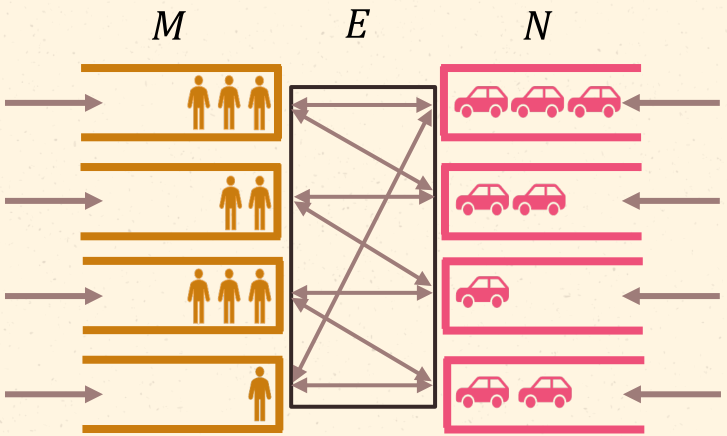
System operator decides to match certain pairs

Service

Arrivals

Poisson arrival with the prescribed arrival rates

Continuous Time Markov Chain



Objective

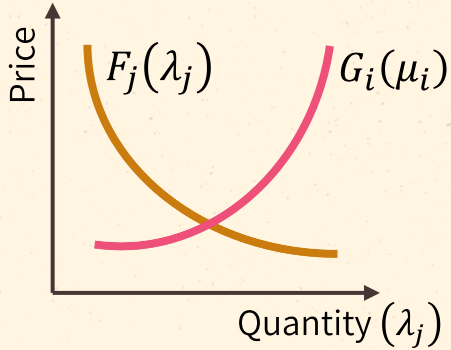
$$\max \mathbb{E}\left[\underbrace{\sum F_j\left(\lambda_j(\boldsymbol{q})\right)\lambda_j(\boldsymbol{q})}_{\text{Revenue}} - \underbrace{\sum G_i(\mu_i(\boldsymbol{q}))\mu_i(\boldsymbol{q})}_{\text{Cost}} \underbrace{- < \boldsymbol{s}, \boldsymbol{q} >}_{\text{Waiting Penalty}}\right]$$

Subject to:

- Feasible Matching
- Stable System

Notation

$F_j(\cdot)$ - Inverse demand curve
$G_i(\cdot)$ - Inverse supply curve
$\boldsymbol{q}$ – State of the System
$\boldsymbol{s}$ – Weight vector for queue lengths



Fluid Model

Replace Stochastic Quantities by their Deterministic Counterparts

$$\gamma^\star = \max \sum F_j(\lambda_j)\lambda_j - \sum G_i(\mu_i)\mu_i$$ **Revenue - Cost**

Can we achieve this bound?
In an asymptotic regime?

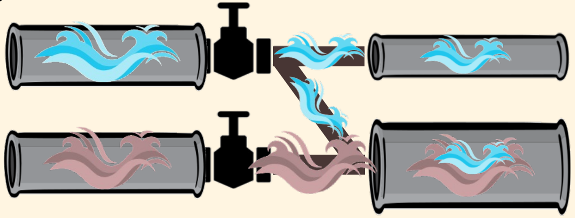
Subject to

$$\lambda_j = \sum_{i=1}^{n} \chi_{ij} \qquad \mu_i = \sum_{j=1}^{m} \chi_{ij}$$

Balance Equations to Match Customers and Servers

$$\chi_{ij} = 0 \ \forall (i,j) \notin E$$

Compatibility Constraint



Large Scale Regime

Scale the arrival rates by η and analyze the system as $\eta \to \infty$

Profit-Loss $(L^\eta = \eta\gamma^\star - R^\eta)$

[Fluid upper bound] − [profit under a given policy]

> Theorem [SV, Bumpensanti, Maguluri, Wang 2022]: Fluid Model Provides an Upper Bound on the Achievable Profit Under any Pricing and Matching Policy

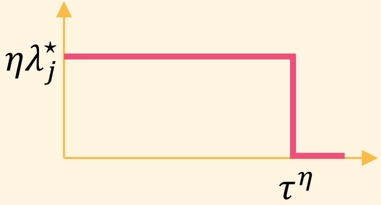
Main Result 1: Large Scale Regime

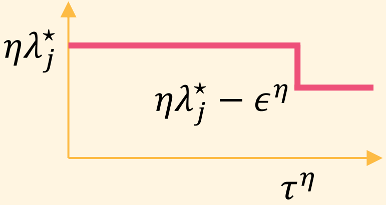
[SV, Bumpensanti, Maguluri, Wang 2022]

	Static Price	Two Price
Objective $L^\eta = \eta\gamma^\star - R^\eta$	$\lambda_j^\eta = \eta\lambda^\star \mathbb{I}_{\{q_j^{(c)} \leq \tau^\eta\}}$ $\mu_i^\eta = \eta\mu^\star \mathbb{I}_{\{q_i^{(s)} \leq \tau^\eta\}}$	$\lambda_j^\eta = \eta\lambda^\star - \epsilon^\eta \mathbb{I}_{\{q_j^{(c)} > \tau^\eta\}}$ $\mu_i^\eta = \eta\mu^\star - \epsilon^\eta \mathbb{I}_{\{q_i^{(s)} > \tau^\eta\}}$
Random Match Proportional to $\chi^\star$ Rescale if the queues are empty	$O(\eta^{1/2})$	$O(\eta^{1/3})$
Max-Weight Match to the type with the greatest number of compatible counterparts waiting	$O(\eta^{1/2})$	$O(\eta^{1/3})$

Lower Bound

$\Omega(\eta^{1/3})$





Notation

$(\lambda^\star, \mu^\star, \chi^\star)$ - Fluid Solution
$\eta\gamma^\star$- Fluid Optimal Value
R^η - Profit, Given Policy

- $\eta^{1/2}$ - CLT type of result, long line of literature using diffusion limits approach
- First $\eta^{1/3}$ type of result [Kim, Randhawa 2017]

Key Observations

Profit-Loss	Static Price	Two Price
Random	$O(\eta^{1/2})$	$O(\eta^{1/3})$
Max-Weight	$O(\eta^{1/2})$	$O(\eta^{1/3})$

Advantage of Dynamic Pricing

Two-Price Policy achieves lower profit-loss compared to Static Price Policy

Small amount of Dynamic Component

Two-Price Policy achieves optimal rate of convergence

Two-Price Policy is the Primary Driver

Two-Price policy coupled with naive matching policies result in optimal profit

Coming up: Advantage of Max-Weight in Large Market Regime

Intuition for $\eta^{1/3}$

$$L^\eta = \underbrace{\eta\gamma^\star - \mathbb{E}\left[\sum F_j^\eta\left(\lambda_j^\eta(\boldsymbol{q})\right)\lambda_j^\eta(\boldsymbol{q}) - \sum G_i^\eta\left(\mu_i^\eta(\boldsymbol{q})\right)\mu_i^\eta(\boldsymbol{q})\right]}_{\text{Revenue Loss}} + \underbrace{\mathbb{E}[< \boldsymbol{s}, \boldsymbol{q} >]}_{\text{Expected Queue Length}}$$

General Pricing Policy

ϵ Perturbation of the fluid policy

Taylor Series Expansion:

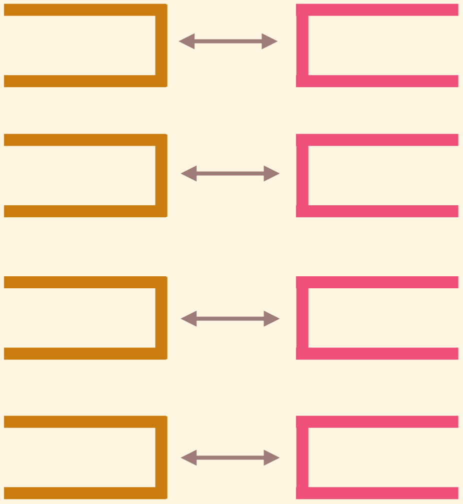
$$\eta P(x^\star + \epsilon) = \eta P(x^\star) + \underbrace{\eta\epsilon P'(x^\star)}_{0} + \eta\epsilon^2 P''(x^\star) + \cdots$$

Like a single server queue in HT

$$\underbrace{\eta\epsilon^2 \qquad\qquad 1/\epsilon}_{\text{Pick } \epsilon \sim \eta^{-1/3} \Rightarrow L^\eta \sim \eta^{1/3}}$$

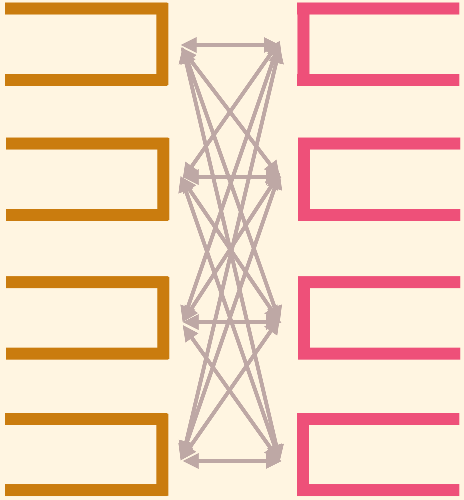
> Theorem [SV, Castro, Maguluri 2021]: For Pricing and Matching Policy such that
>
> $$\mathbb{E}[q] \leq \frac{1}{\delta} \Rightarrow P \leq \gamma^\star - \Omega(\delta^2)$$
>
> Two-Price + Max-Weight achieves this trade-off

Large Market Regime



$$L^{\eta} = \Omega\left(\eta^{1/3} n\right)$$

n independent matching queues



$$L^{\eta} = \Omega\left(\eta^{1/3} n^{1/3}\right)$$

One resource pooled matching queue with arrival rate $n\eta$

Goal

- Conditions on the graph such that it behaves like complete graph
- Policy that achieves resource pooling

Crp Condition

$$\sum_{j \in J} \lambda_j^{\star} < \sum_{i : \exists j \in J, (i,j) \in E} \mu_i^{\star} \quad \forall J \subset M$$

$$\sum_{i \in I} \mu_i^{\star} < \sum_{j : \exists i \in I, (i,j) \in E} \lambda_j^{\star} \quad \forall I \subset N$$

Hall's condition weighted graph

The above implies

- Graph is connected
- Fluid solution is in the interior of the "stability region"

Main Result 2: Large Market Regime

[SV, Bumpensanti, Maguluri, Wang 2022]

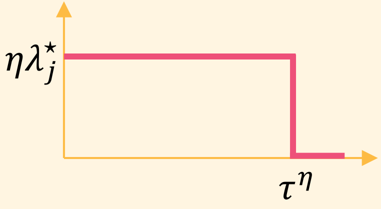
Objective $L^\eta = \eta\gamma^\star - R^\eta$	Static Price $\lambda_j^\eta = \eta\lambda^\star \mathbb{I}_{\{q_j^{(c)} \le \tau^\eta\}}$ $\mu_i^\eta = \eta\mu^\star \mathbb{I}_{\{q_i^{(s)} \le \tau^\eta\}}$	Two Price $\lambda_j^\eta = \eta\lambda^\star - \epsilon^\eta \mathbb{I}_{\{q_j^{(c)} > \tau^\eta\}}$ $\mu_i^\eta = \eta\mu^\star - \epsilon^\eta \mathbb{I}_{\{q_i^{(s)} > \tau^\eta\}}$
Random Match Proportional to $\chi^\star$ Rescale if the queues are empty	$O(\eta^{1/2})\Omega(n)$	$O(\eta^{1/3})\Omega(n)$
Max-Weight Match to the type with the greatest number of compatible counterparts waiting	$O(\eta^{1/2})O(n^{1/2})$	$O(\eta^{1/3})O(n^{1/3})$

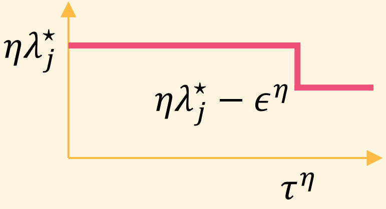
Notation

$(\lambda^\star, \mu^\star, \chi^\star)$ - Fluid Solution
$\eta\gamma^\star$- Fluid Optimal Value
R^η - Profit, Given Policy

Lower Bound

$$\Omega(\eta^{1/3})\Omega(n^{1/3})$$





Key Observations

Profit-Loss	Static Price	Two Price
Random	$O(\eta^{1/2})\Omega(n)$	$O(\eta^{1/3})\Omega(n)$
Max-Weight	$O(\eta^{1/2})O(n^{1/2})$	$O(\eta^{1/3})O(n^{1/3})$

Max-Weight is better than Random

Max-Weight exploits the underlying network structure

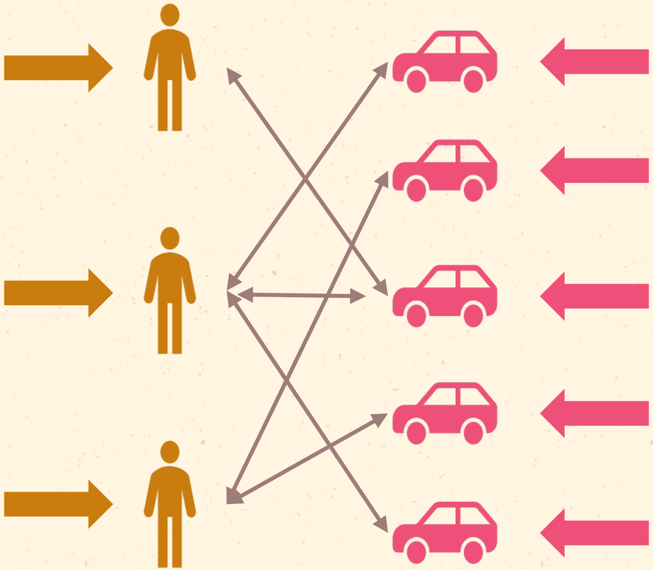
Max-Weight is optimal w.r.t. n

Max-Weight results in state space collapse – system behaves like a single-link two-sided queue

Two-Price + Max-Weight is optimal w.r.t. η and n

This illustrates the interplay of pricing and matching policy – right combination is important

Stochastic Network Viewpoint



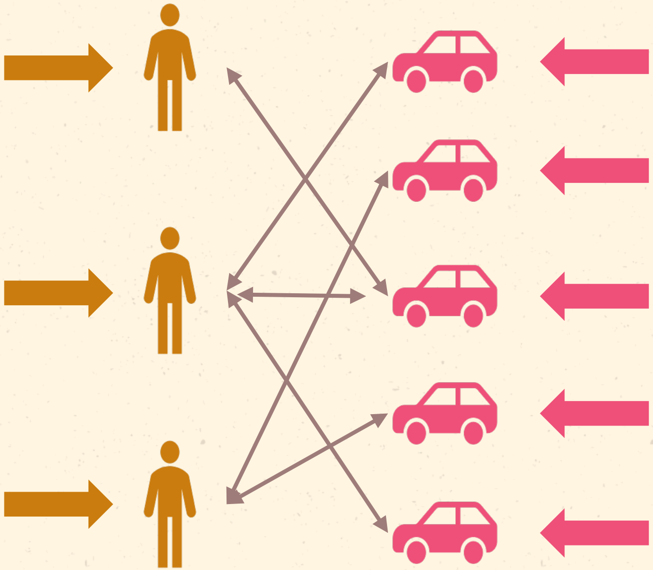
Type – Geographical location, normal/premium ride, etc.

Compatibility – Geographical proximity and matching preferences

Match – Disappear from the system instantaneously

Set **prices** and perform **matchings** that maximizes **profit** and minimizes the **delay**

Stochastic Network Viewpoint



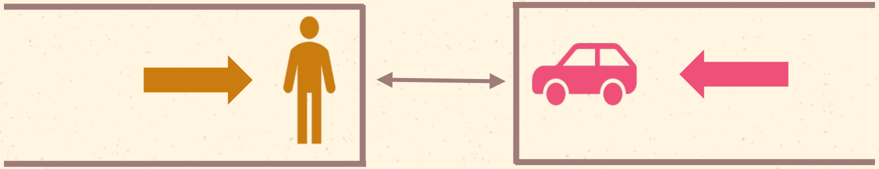
Type – Geographical location, normal/premium ride, etc.

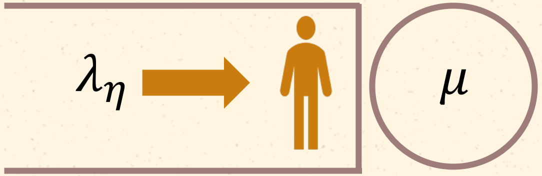
Compatibility – Geographical proximity and matching preferences

Match – Disappear from the system instantaneously

Analyze the entire **stationary distribution**, not just the **mean**

Matching Queue: Simplest Case





Difficult as even G/G/1 queue (light traffic) is still an open problem

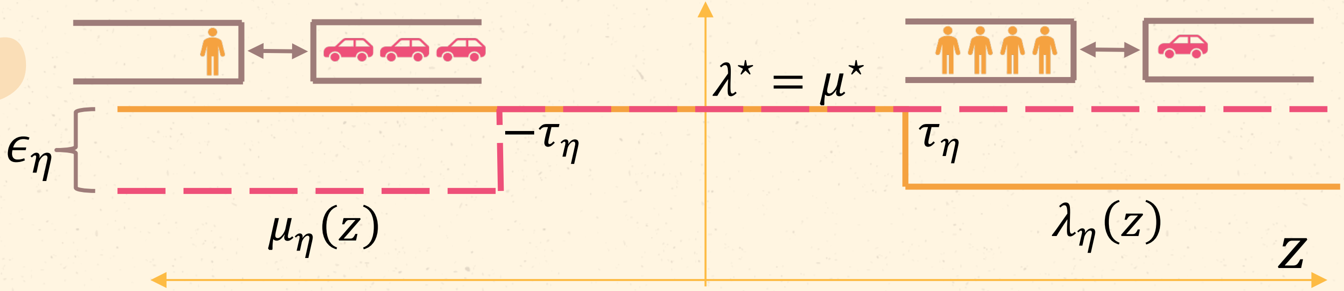
Consider an asymptotic regime: Heavy-Traffic

$$\lambda_\eta \to \mu$$

The system approaches null-recurrence

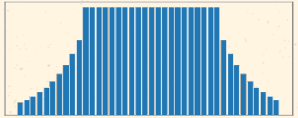
Analyze the entire **stationary distribution**, not just the **mean**

Phase Transition: Illustrative Example



$a_\eta^c, a_\eta^s \sim Bernoulli$

Birth-Death Process



ϵ_η: Magnitude Scaling Parameter

τ_η: Time Scaling Parameter

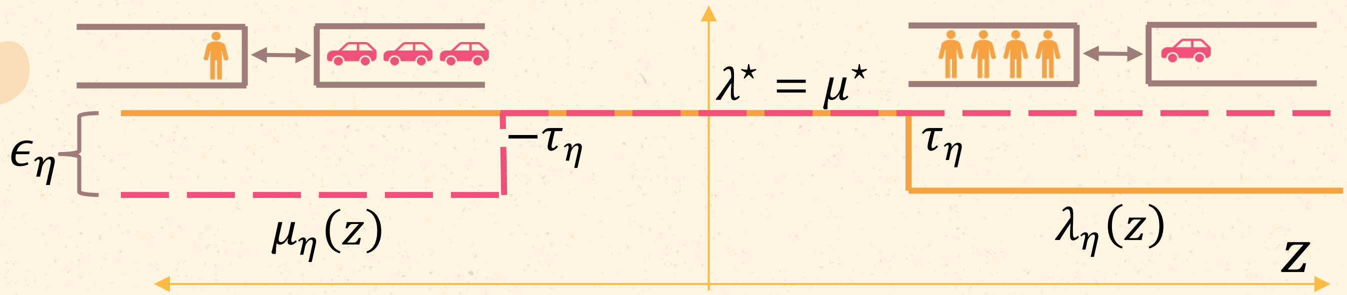
Heavy-Traffic is given by either $\epsilon \to 0$ and/or $\tau \to \infty$

Case I: $\epsilon_\eta \tau_\eta \to 0$

Case II: $\epsilon_\eta \tau_\eta \to (0, \infty)$

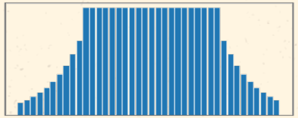
Case III: $\epsilon_\eta \tau_\eta \to \infty$

Phase Transition: Illustrative Example



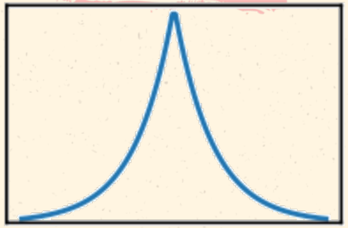
$a^c_\eta, a^s_\eta \sim Bernoulli$

Birth-Death Process



Case I: $\epsilon_\eta \tau_\eta \to 0$

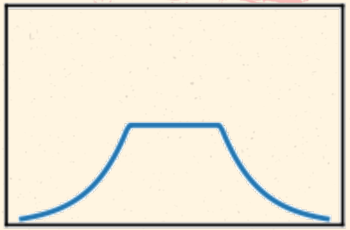
$\epsilon_\eta z_\eta \to Laplace$



Intuition: $\tau = c$ and $\epsilon \to 0$
(Similar to classical HT)

Case II: $\epsilon_\eta \tau_\eta \to (0, \infty)$

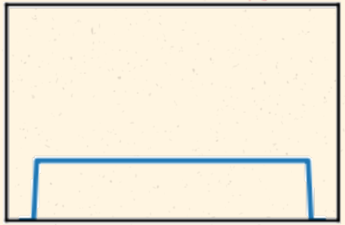
$\epsilon_\eta z_\eta, \dfrac{z_\eta}{\tau_\eta} \to Hybrid$



Laplace stitched with Uniform

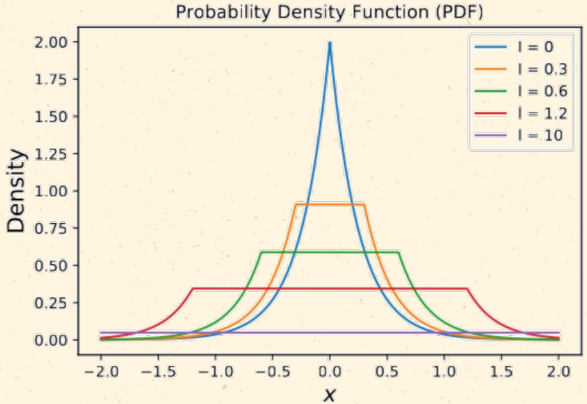
Case III: $\epsilon_\eta \tau_\eta \to \infty$

$\dfrac{z_\eta}{\tau_\eta} \to Uniform$



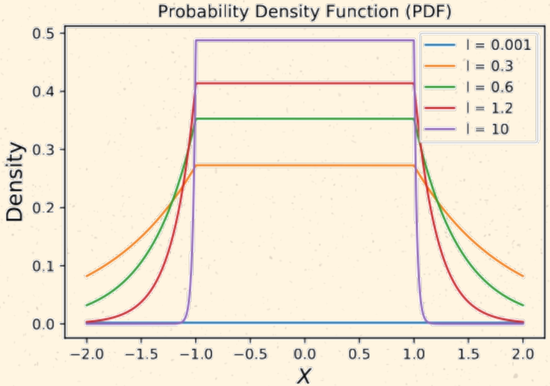
Intuition: $\epsilon = c$ and $\tau \to \infty$
(High drift outside the thresholds)

Phase Transition: Illustrative Example



ϵz

$\epsilon\tau \to l$



z/τ

Hybrid -> Laplace

$l \to 0$

Hybrid -> Uniform

$l \to \infty$

Main Result: Phase Transition [SV, Maguluri 2021]

Arrival Rates

$$\lambda(z) = \lambda^\star + \phi^c\left(\frac{z}{\tau_\eta}\right)\epsilon_\eta$$

$$\mu(z) = \mu^\star + \phi^s\left(\frac{z}{\tau_\eta}\right)\epsilon_\eta$$

Notation

ϵ_η: Magnitude Scaling Parameter

τ_η: Time Scaling Parameter

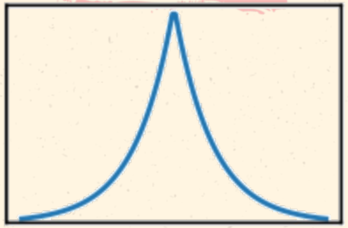
$\phi^c(\cdot), \phi^s(\cdot)$: General control curves

$a^c, a^s \sim General$

DTMC

Case I: $\epsilon\tau \to 0$

$\epsilon z \to Laplace$



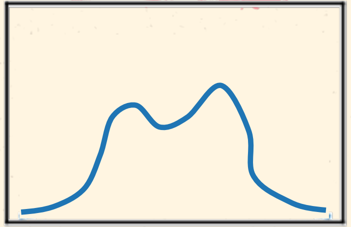
Intuition: $\tau = c$ and $\epsilon \to 0$

(Similar to classical HT)

$l \to 0$

Case II: $\epsilon\tau \to (0, \infty)$

$\epsilon z, \frac{z}{\tau} \to$ Gibbs

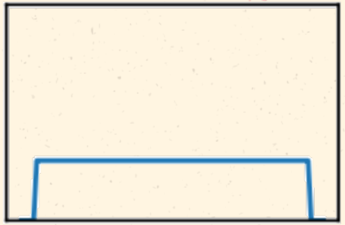


Explicitly depends on $\phi^c(\cdot), \phi^s(\cdot)$

$l \to \infty$

Case III: $\epsilon\tau \to \infty$

$\frac{z}{\tau} \to Uniform$



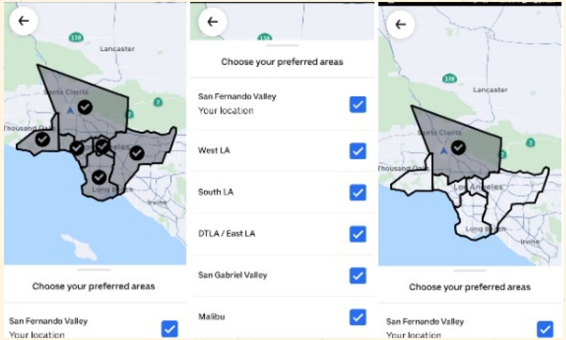
Intuition: $\epsilon = c$ and $\tau \to \infty$

(High drift outside the thresholds)

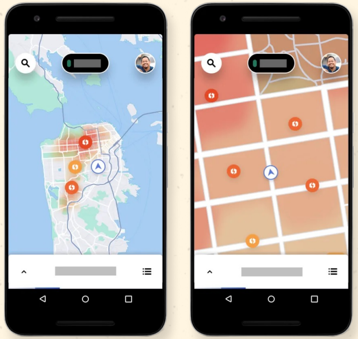
Other Lines of Work

Incorporating Strategic Servers

Area Preferences

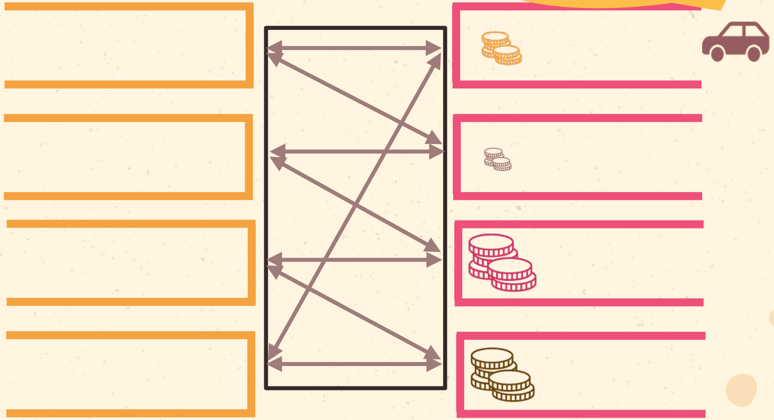


Surge Pricing



Model

What if I lie and join another queue



Set a Driver Destination

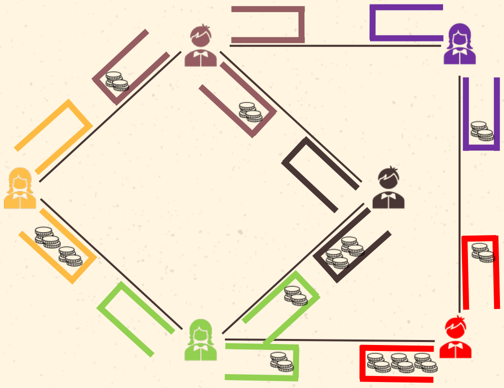
When you set a Driver Destination in your app, we'll try and match you with trip requests from riders going towards that destination.

Result [SV, Castro, Maguluri 2021]

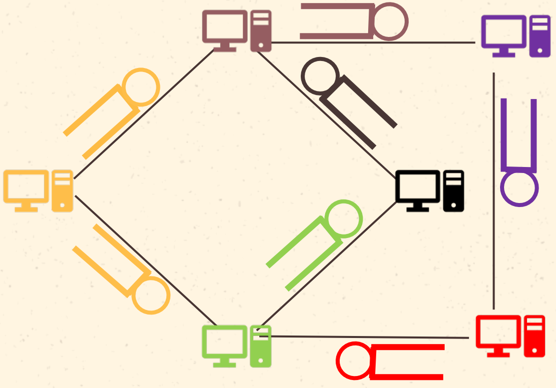
Incentive-compatible, near-optimal, pricing and matching policies for a wide variety of utility functions

Motivation and Model

Payment Channel Networks



Two-Sided Network



Classical Communication Network

- Each payment link in a payment processing network is a two-sided queue
- Analogous two-sided version of classical communication network
- The problem is to route transactions using the payment channels
- We propose a throughput optimal routing algorithm inspired by max-weight [SV, Maguluri 2021]

Matching queues: related papers

- **Matching Queues**

  SV, Maguluri, “A Heavy Traffic Theory of Matching Queues”
  Conf: IFIP Performance (Student Best Paper)

- **Stochastic Matching Network**

  SV, Bumpensanti, Maguluri, Wang, “Dynamic Pricing and Matching for Two-Sided Queues”
  Conf: SIGMETRICS, Jour: Operations Research

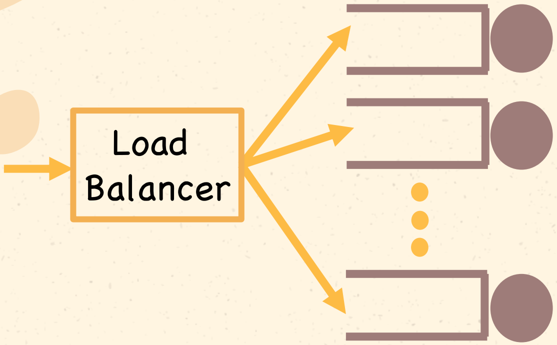
- **Strategic Agents**

  SV, Castro, Maguluri, “Near-Optimal Control in Ride-Hailing Platforms with Strategic Servers”
  Conf: SIGMETRICS

- **Payment Channel Networks**

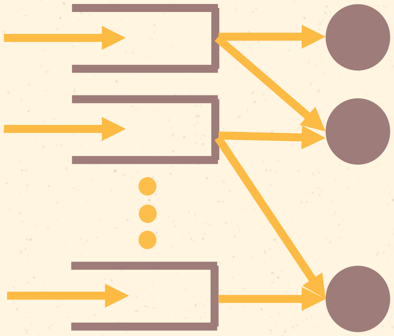
  SV, Maguluri, “Throughput Optimal Routing in Blockchain-Based Payment Systems”
  Jour: IEEE Transaction on Control of Network Systems

Stochastic Processing Networks



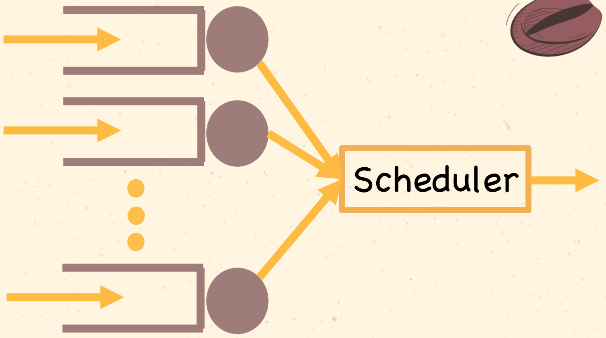
Power-of-d Choices Load Balancing in the Sub-Halfin Whitt Regime

SV, Castro, Maguluri 2022



Transportation Polytope and its Applications in Parallel Server Systems

SV, Maguluri 2021 **(INFORMS talk)**



Logarithmic Heavy Traffic Error Bounds in Generalized Switch and Load Balancing Systems

Lange, SV, Maguluri 2021, Journal of Applied Probability

Reinforcement Learning

Khodadadian, Jhunjhunwala, SV, Maguluri, On the Linear and Super-linear Convergence of Natural Policy Gradient Algorithm, Conf: IEEE CDC, Jour: System and Control Letters