ZUber against ZLyft Apocalypse

Jérôme Thai[1] Chenyang Yuan[1] Alexandre Bayen[1]

[1]Department of Electrical Engineering & Computer Sciences
University of California at Berkeley

November 4, 2015

Outline

Denial-of-Service attacks in MaaS systems

Collaborative Consumption | Lyft | Uber

Uber Strikes Back, Claiming Lyft Drivers And Employees Canceled Nearly 13,000 Rides

Posted Aug 12, 2014 by Ryan Lawler (@ryanlawler), Contributor

CNN Money   Business   Markets   Tech   Media   Personal Finance   Small Biz   Luxury

Innovation Nation

Uber's dirty tricks quantified: Rival counts 5,560 canceled rides

Uber, Lyft Battle It Out In San Francisco With Ultra-Low Prices On Carpool Rides

By Salvador Rodriguez @sal19 s.rodriguez@ibtimes.com on January 26 2015 6:33 PM EST

Cyber-security concern in future Autonomous MaaS systems

Self-Driving Cars Compete With The IoT For The Title Of Most Hyped Technology; Big Data Out

FULL BIO >

Opinions expressed by Forbes Contributors are their own.

These technologies at the peak of the hype cycle also highlighted for me what's missing from this year's report. Given that the most hyped news out of Black Hat and Defcon conferences earlier this month were demonstrations of how to hack into cars (self-driving or not) and take control of them remotely, it is interesting that Gartner does not list any specific cybersecurity-related emerging technologies. It does mention, however, two general categories—"digital security" and "software-

WIRED Hackers Remotely Kill a Jeep on the Highway—With Me in It

BUSINESS DESIGN ENTERTAINMENT GEAR SCIENCE SECURITY

ANDY GREENBERG SECURITY 07.21.15 6:00 AM

HACKERS REMOTELY KILL A JEEP ON THE HIGHWAY—WITH ME IN IT

MENU FORTUNE SUBSCRIBE

NEWS POPULAR VIDEOS FORTUNE 500

Car hacking: how big is the threat to self-driving cars? OCTOBER 7, 2014

Car hacking: how big is the threat to self-driving cars?

Zombies

In computer science, a **Zombie** *is a computer that has been compromised remotely by a hacker to launch DoS attacks.*

Companies control the dispatch via

- Direct control with a dispatch center.
- Incentivization through hailing apps and surge pricing.

Assumption: attackers control a fraction of the vehicles via

- Spoofing of the hailing apps.
- Boosting customer demand with very low fares.

Zombies

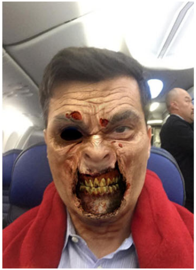
In computer science, a **Zombie** *is a computer that has been compromised remotely by a hacker to launch DoS attacks.*

Companies control the dispatch via

- Direct control with a dispatch center.
- Incentivization through hailing apps and surge pricing.

Assumption: attackers control a fraction of the vehicles via

- Spoofing of the hailing apps.
- Boosting customer demand with very low fares.



Objective

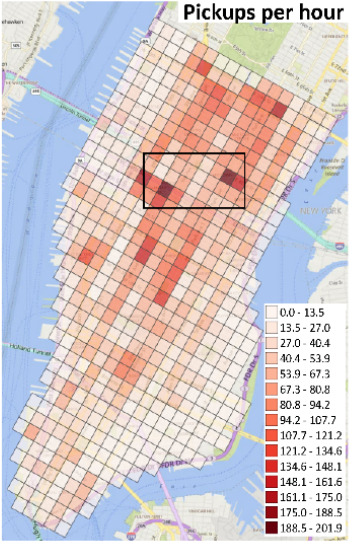
Quantifying the price of attacks for

- Depleting taxis in arbitrary locations.
- Minimize customer usage of the service.

Quantifying countermeasures via cost-benefit analysis

- Minimum price of attacks to protect the MaaS system.
- Adjusting cancellation fees.

Outline

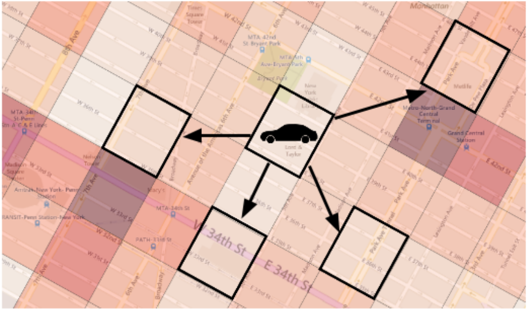


Tessellation:

- 531 squares the size of 2 city blocks
- 282,000 origin-destination pairs

From 75M taxi trips (2009-2015, weekdays, 5pm-7pm), learned:

- pickup rates
- routing distribution
- mean travel times



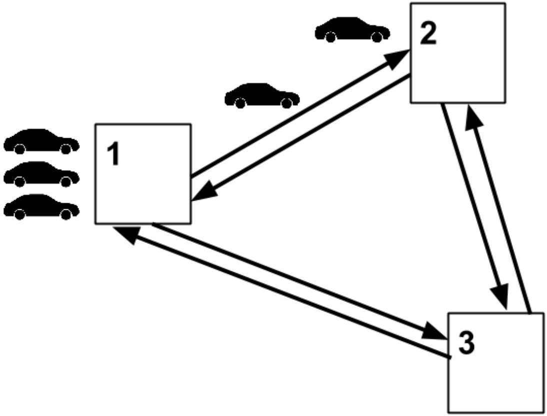
Learning of the demand

Dataset of 1B trips from Jan 2009 to Jun 2015. Chose trips:

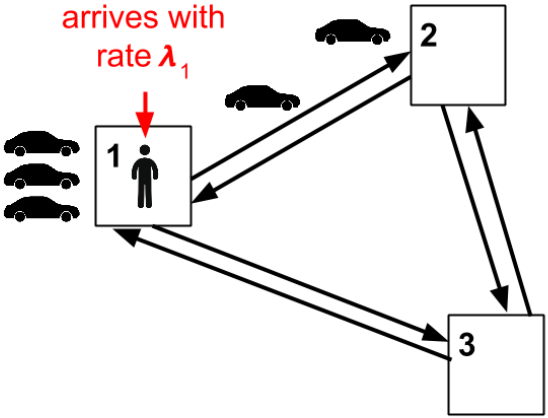
- Starting and ending in region
- Pickup between 5-7pm on all weekdays

Used Google's BigQuery to help infer the parameters for our model. Some high level statistics:

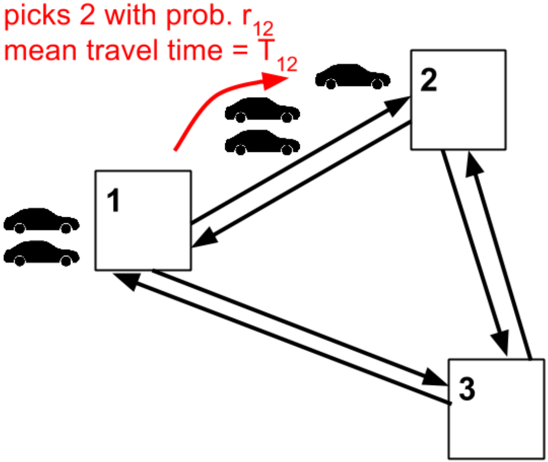
- Mean trip distance: 1.7 miles (standard dev: 1.2 miles)
- Mean travel time: 11 mins (standard dev: 5.5 mins)



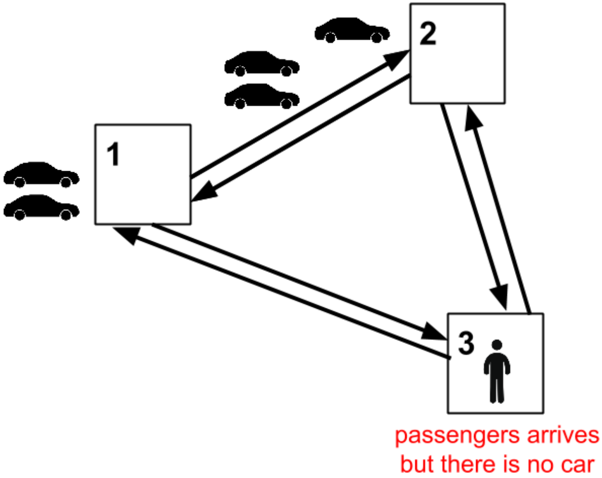
Example: network with three stations.



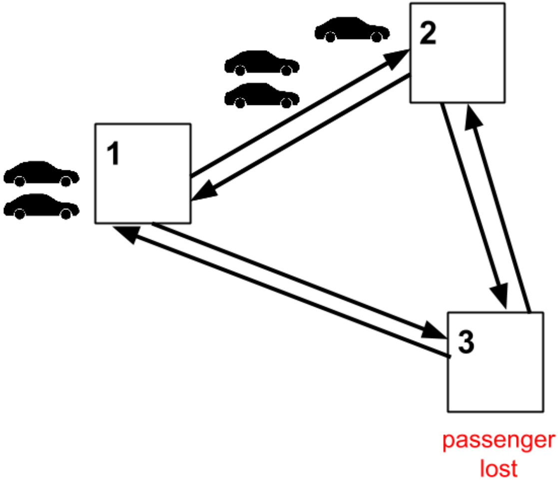
Customer arrives at station 1 with rate λ_1 and gets a car.



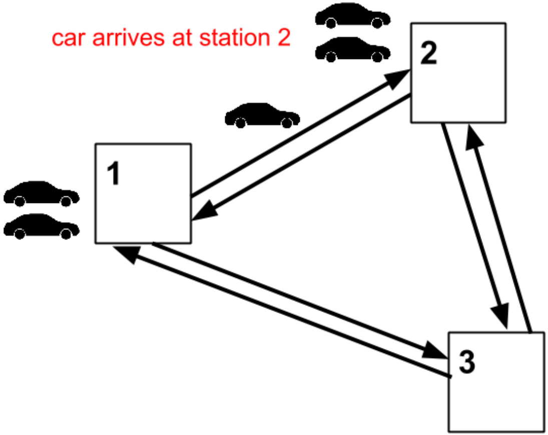
Picks up destination 2 (resp. 3) with probability r_{12} (resp. r_{13}).



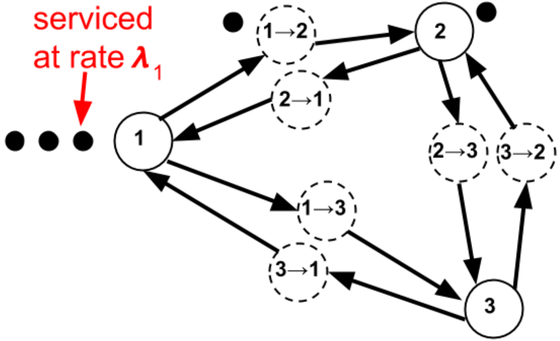
Customer arrives at station 3 with rate λ_3.



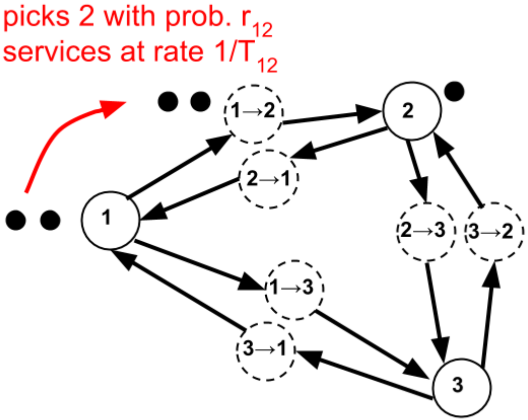
No car at station 3: passenger leaves the system.



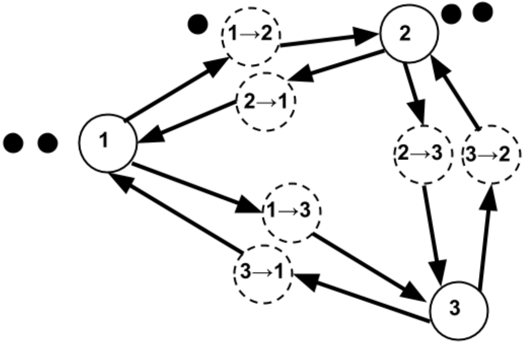
Car arrives at station 2.



Jackson network: station nodes + route nodes between pairs of stations.

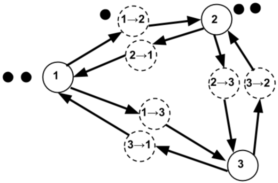


Car (packet) leaves station 1 to go to route node 1→2.



After spending T_{12} on route 1→2, arrives at station 2.

Casting into a Jackson network



- 1st car in line processed with rate ϕ_i (customer arrival rate at i)
- Routed to node $i \to j$ with probability α_{ij}
- Processed with rate $1/T_{ij}$ (T_{ij} = mean travel time from i to j)
- Routed to station j with probability 1
- Full specification

Service rate:	$\mu_i = \phi_i$	$\mu_{i\to j} = 1/T_{ij}$
Routing probabilities:	$p_{i,\,i\to j} = \alpha_{ij}$	$p_{i\to j,\,j} = 1$

Stationarity results

- In equilibrium, arrival rates π_i of cars at station i:

$$\pi_i = \sum_j p_{ji} \pi_j \quad \text{(balance equations)}$$

- $\gamma_i :=$ relative utilization $= \pi_i / \mu_i$ satisfies

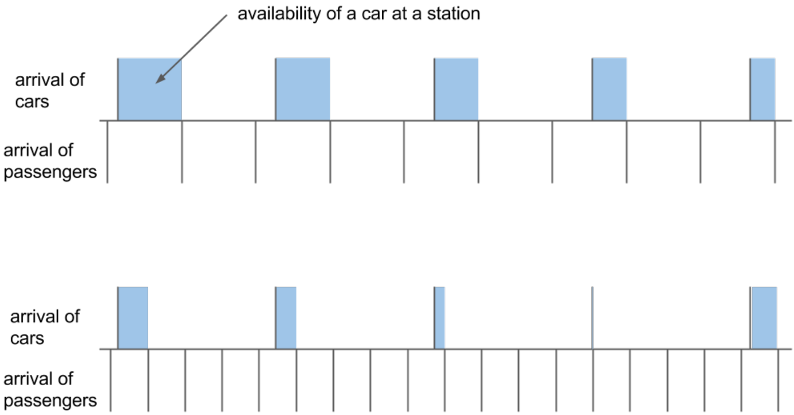
$$\gamma_i = \sum_j \frac{p_{ji} \mu_j}{\mu_i} \gamma_j$$

- $X_i :=$ number of vehicles in queue at station i (random variable)
- Availability at station i:

$$\Pr[X_i \geq 1] \propto \gamma_i = \pi_i / \mu_i$$

Intuition for Availability Proportional to Utilization

$\Pr[X_i \geq 1] \propto \gamma_i = \text{throughput/service rate} = \pi_i/\mu_i$



$\mu_i := 2\mu_i \quad \Longrightarrow \quad \gamma_i := \gamma_i/2$

$\Longrightarrow$ on average halve the probability that there is a car at the station.

Large fleet size

- Recall $\Pr[X_i \geq 1] \propto \gamma_i = \pi_i / \mu_i$
- Let α be the constant factor $\Pr[X_i \geq 1] = \alpha \gamma_i \leq 1$
- When the fleet size grows, stations with highest γ_i will be saturated

$$\Pr[X_i \geq 1] \approx 1 \quad \text{for } i \,:\, \gamma_i = \text{max}_j \gamma_j$$

- Hence, for large fleet sizes

$$\Pr[X_i \geq 1] \approx a_i := \gamma_i / \text{max}_j \gamma_j$$

- Limit a_i of $\Pr[X_i \geq 1]$ uniquely defined by

$$\gamma_i = \sum_j \frac{p_{ji} \mu_j}{\mu_i} \gamma_j \tag{1}$$

$$a_i := \gamma_i / \text{max}_j \gamma_j \tag{2}$$

Outline

	arrival rate	routing	authors	contribution
	ϕ_i	α_{ij}	George & Xia	Framework
	ψ_i	β_{ij}	Zhang & Pavone	Balancing
	ν_i	κ_{ij}	Thai, Yuan & Bayen	Cybersecurity

Three stochastic mechanisms: *Customers*, *Balancers*, and *Zombies*.

Stochastic control

- Generalized passenger arrival rate at station i

$$\lambda_i = \phi_i + \psi_i + \nu_i$$

- Upon arrival, prob. of a generalized passenger of being of each type

$$\Pr(\text{Customer}) = \phi_i/\lambda_i$$
$$\Pr(\text{Balancer}) = \psi_i/\lambda_i$$
$$\Pr(\text{Zombie}) = \nu_i/\lambda_i$$

- Generalized passenger routing

$$p_{ij} = \sum_{\text{type}} \Pr(i \to j \mid \text{type}) \Pr(\text{type})$$

$$p_{ij} = \frac{\phi_i\alpha_{ij} + \psi_i\beta_{ij} + \nu_i\kappa_{ij}}{\phi_i + \psi_i + \nu_i}$$

Combining Customers and Balancers

- Combined arrival rate and routing of Customers and Balancers

$$\varphi_i = \phi_i + \psi_i \qquad \delta_{ij} = \frac{\phi_i \alpha_{ij} + \psi_i \beta_{ij}}{\phi_i + \psi_i}$$

- Generalized passenger arrival rate and routing

$$\lambda_i = \varphi_i + \kappa_i \qquad p_{ij} = \frac{\varphi_i \delta_{ij} + \nu_i \kappa_{ij}}{\varphi_i + \nu_i}$$

	arrival rate	routing	
 + 	φ_i	δ_{ij}	Given
	ν_i	κ_{ij}	Control

Objective of attacks

Recall $a_i = \lim \Pr[X_i \geq 1]$ for large fleet size is well-defined.

Our objective is:

$$\min \sum_{i \in \mathcal{S}} w_i a_i$$

where the weights $w_i > 0$ are chosen such that:

- $w_i = \phi_i$ (customer arrival rate) to maximize the rate of customer loss
- $w_i = \sum_j \phi_i \alpha_{ij} T_{ij}$ to maximize customer travel time loss

We also add a l_2 regularization term: $\frac{\rho}{2} \sum_i {\nu_i}^2$ to have

- a strongly convex objective (numerical).
- discourage very large values of ν_i (physical).

Bound on attacks

Bound on the total rate of attacks:

$$\sum_i \nu_i \leq b$$

Reasons:

- Without it, easy to design strategy such that for any k

$$a_k = 1, \qquad a_i \to 0 \quad \forall\, i \neq k$$

- Issuing attacks has a cost, hence b is the budget for attacks.

Bound on the radius of attacks

Bound on the radius of attacks:

$$\kappa_{ij} = 0 \quad \text{if} \quad \text{dist}(i,j) \geq r$$

Reasons:

- Attacker has weaker control than customers and balancers.
- Attacks can be detected.

Define $\mathcal{E}$, pairs (i,j) of feasible attacks from station i to j. Then:

$$0 \leq \kappa_{ij} \leq \mathbf{1}_{\{(i,j)\in\mathcal{E}\}}$$

Problem Formulation

We fix an arbitrary $k = \text{argmax}_i\, a_i$, thus $a_k = 1$ and $a_i \leq 1$ for $i \neq k$.

$$\min_{\kappa_{ij},\nu_i,a_i} \sum_{i\neq k} w_i a_i + \frac{p}{2}\sum_i {\nu_i}^2 \qquad \text{Max. customer loss + reg.}$$

$$\text{s.t. } a_i = \sum_{j\in\mathcal{S}} a_j \frac{\delta_{ji}\varphi_j + \kappa_{ji}\nu_j}{\varphi_i + \nu_i} \qquad \text{Balance equations}$$

$$\mathbf{1}_{\{(i,j)\in\mathcal{E}\}} \geq \kappa_{ij} \geq 0,\ \sum_j \kappa_{ij} = 1 \qquad \text{Attacks within radius}$$

$$\nu_i \geq 0,\ \sum_i \nu_i \leq b \qquad \text{Attacks within budget}$$

Outline

Curse of dimensionality

- a_i is uniquely defined by

$$a_i = \sum_{j \in \mathcal{S}} a_j \frac{\delta_{ji}\varphi_j + \kappa_{ji}\nu_j}{\varphi_i + \nu_i}, \qquad a_k = 1$$

- Hence the objective $\sum_{i \neq k} w_i a_i + \frac{\rho}{2} \sum_i \nu_i{}^2$ is a function of ν_i, κ_{ij}
- Computing $\partial a_i / \partial \kappa_{kl}$ has N^2 complexity
- Hence gradient computation is N^4 $(N = 531)$
- We use block-coordinate descent

Block-coordinate descent

Recall:

ν_i	Zombie arrival rate at section i
κ_{ij}	Zombie routing probability from i to j
a_i	Availability at section i
w_i	Weights in objective function

Apply block-coordinate descent by fixing one of ν_i, κ_{ij}, and a_i

Type	Fix	Vary	Minimize	Solver Used
LP	ν_i	a_i, κ_{ij}	$\sum_i w_i a_i$	CPLEX
QP	a_i	κ_{ij}, ν_i	$\sum_i {\nu_i}^2$	CPLEX
QCQP	κ_{ij}	ν_i, a_i	$\sum_i w_i a_i + \frac{\rho}{2} \sum_i {\nu_i}^2$	Gradient descent

We repeat these steps in succession until convergence.

Interpretation of attacks

Each step of the block-coordinate descent can be interpreted as an attack strategy.

Attack Routing (fix ν_i, vary a_i, κ_{ij}): Fix attack rates on all stations, what is the best routing strategy for these attacks?

Min Attack (fix a_i, vary κ_{ij}, ν_i): Fix target availabilities, what is the best way to re-route the attacks

Attack Rate (fix κ_{ij}, vary ν_i, a_i): Fix the attack routing strategy, find the best attack rates that utilizes these strategies.

Solution of Min-Attack Problem

(Simplified) Problem formulation:

$$\min_{\kappa_{ij},\nu_i} \sum_i {\nu_i}^2 \qquad \ell_2 \text{ Regularization}$$

$$\text{s.t. } a_i(\varphi_i + \nu_i) = \sum_{j\in\mathcal{S}} a_j(\delta_{ji}\varphi_j + \kappa_{ji}\nu_j) \qquad \text{Balance equations}$$

Idea: define $x_{ij} := a_i\kappa_{ij}\nu_i$, then $a_i\nu_i = \sum_j x_{ij}$.

Then the constraints become linear flow constraints:

$$\min_{x_{ij}} \sum_i \frac{1}{2a_i^2}\left(\sum_j x_{ij}\right)^2$$

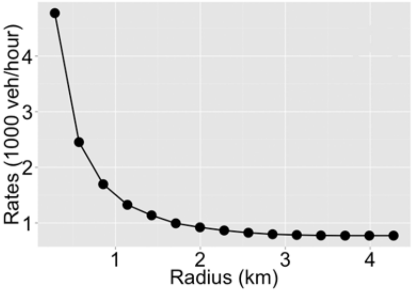
$$\text{s.t.} \sum_{j\neq i}(x_{ji} - x_{ij}) = s_i$$

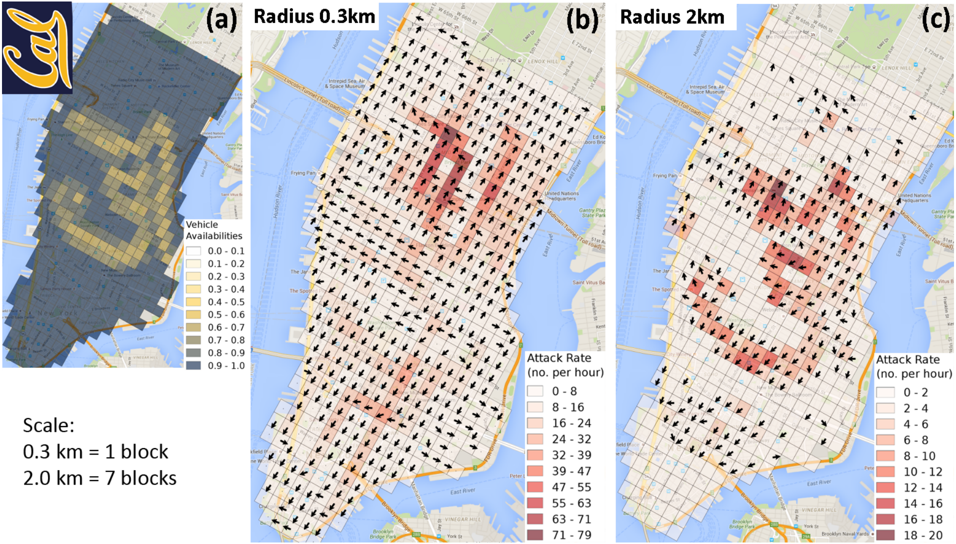
Replacing the quad. obj. by $\min \sum_{ij} T_{ij}x_{ij}$ gives standard Min-Cost Flow problem.

Outline

Arbitrary emptying the network

- Choose an arbitrary vector a_i, $i \in \mathcal{S}$ of availabilities on Manhattan
- Minimize the number of *Zombies* circulating to achieve a_i, $i \in \mathcal{S}$
- Constraint the radius of attacks



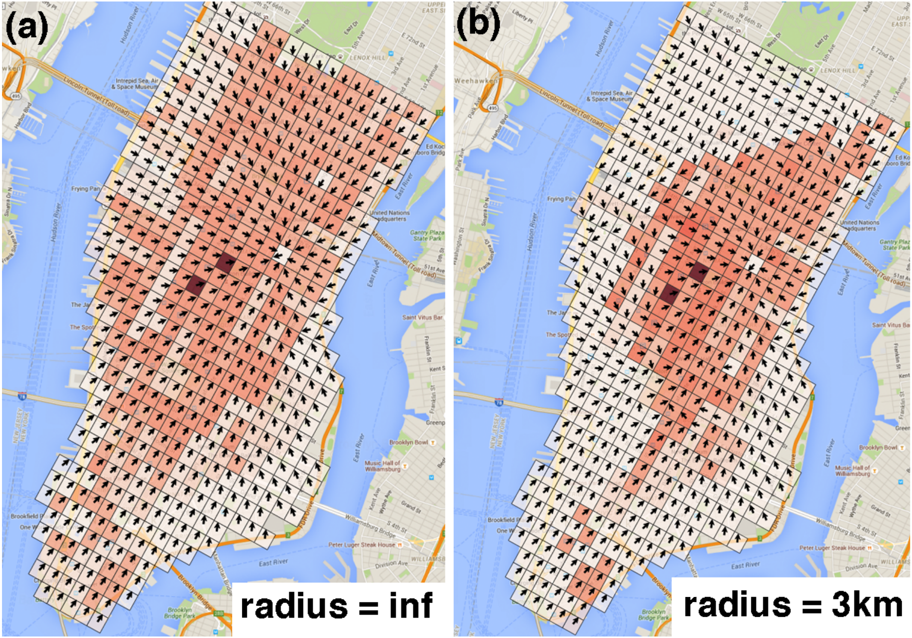


Scale:
0.3 km = 1 block
2.0 km = 7 blocks

Drawing the CAL logo on Manhattan

Maximizing passenger loss

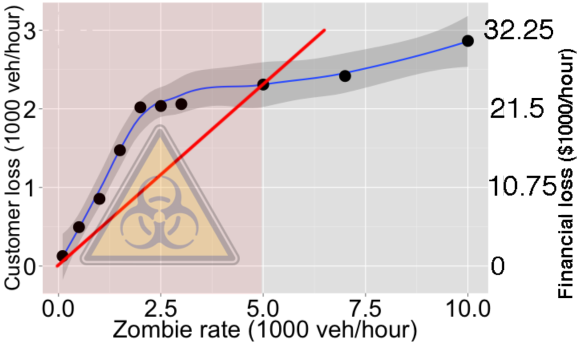
- No limit on the radius of attacks
- Set budget b of attacks to be from 100 to 10000 veh/hour.
- Represents from 0.8% to 44% of the total rate in the network.
- Start with uniform arrival rates and uniform routing probabilities.



Minimizing the availabilities

Simulate the transient state

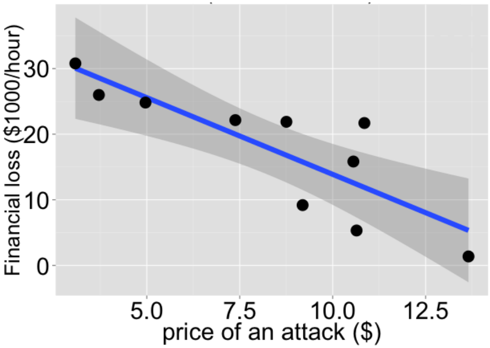
- Track passenger loss in a balanced MaaS system with 2500 taxis.
- Start injecting Zombies and track increased passenger loss for 1h.
- Figure shows the passenger loss incurred by the attacks:



- Right-axis: financial loss with (with average fare of $10.75).
- Red line: price of attack assuming a cost of $5/unit.

Cost-benefit analysis

- Each point: max financial loss for a given price of attacks.
- Cost of 1 unit of attack of $15: no economic incentive to attack.



Conclusions and future work

Direct extension:

- Attack-defender game.
- Robust dispatch and attacks.

Price of anarchy:

- From MaaS rivalr.
- From selfish behavior of taxi drivers.

Dynamical system

END