Urban Traffic Analysis Using HSL Dataset

Kai Zhao	**University of Helsinki**
Pradeep Kumar Eranti	**Aalto University**
Tatiraju Venkata	**Aalto University**
Nautiyal Sudhanshu	**Aalto University**

April 2015

Outline

Introduction

1. Analyzing Urban Traffic with Network Analysis methods (What, Why).
2. Correlation between Bus Traffic and Urban Traffic (Why).
3. How to improve the urban transportation systems (How).

Conclusion

Introduction

Motivation:

Urban traffic analysis is important for
- Urban planing and traffic forecasting
- Smart city
- Better energy use

Introduction

We use Bus Delay as an identification of the Urban Traffic.

- More delay, more urban traffic
- Timetable time – actual arriving time

We use The number of Bus travelling between two bus stops as an indenticiation of bus traffic

- More buses, more bus traffic
- Number of buses between two bus stop within one hour

1. Analyzing Urban Traffic Delay with Network Analysis methods (What, why).

What does the city look like with traffic delay virtualization?

What is a key area causing the traffic delay?

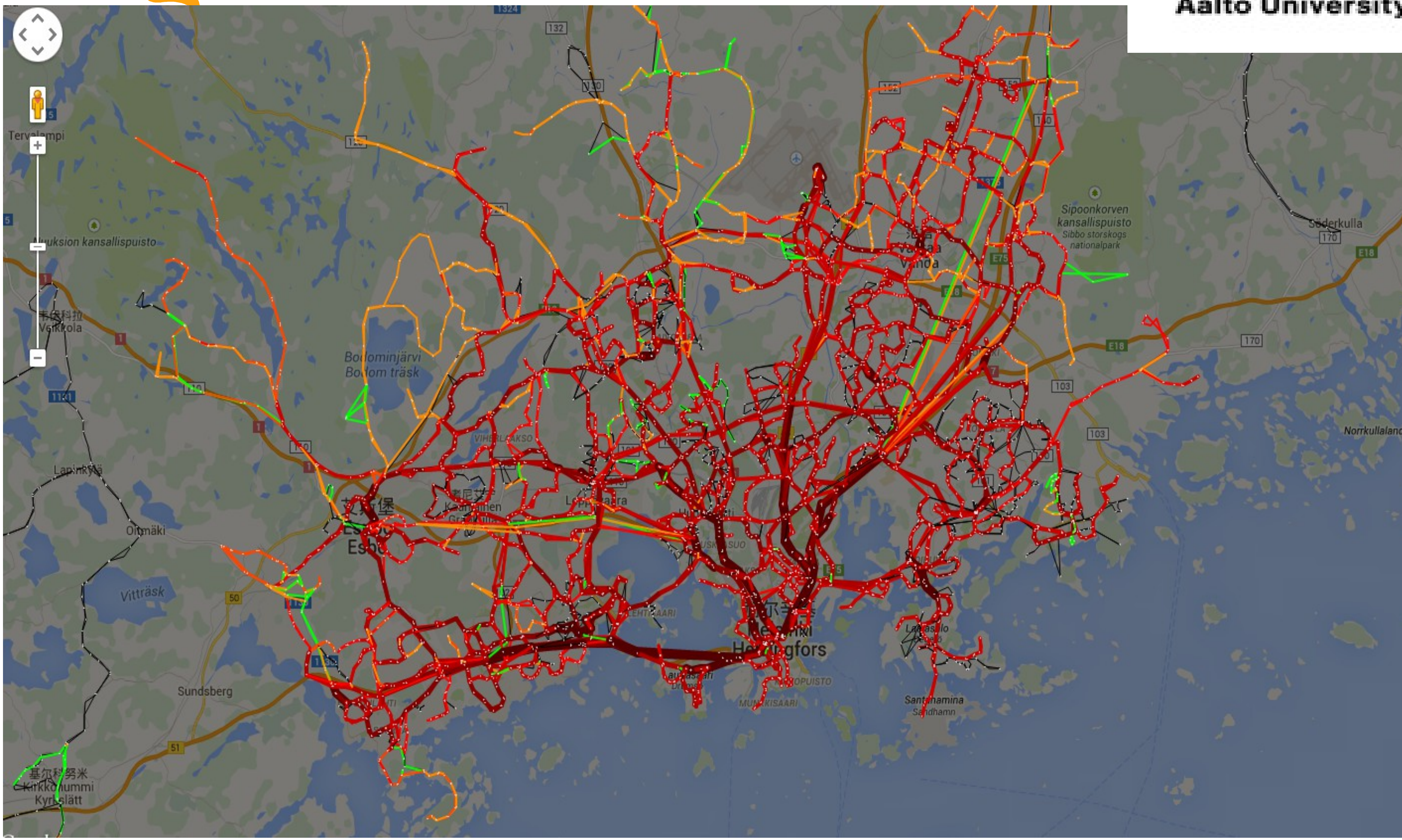
1. City is a Network

Each bus stop is a node

If there is a bus traveling between two bus stops within an hour, there is an edge between the two bus stops. (Temporal networks)

The edge weight is determined by the average bus delay over two bus stops.

1. Traffic virtualization



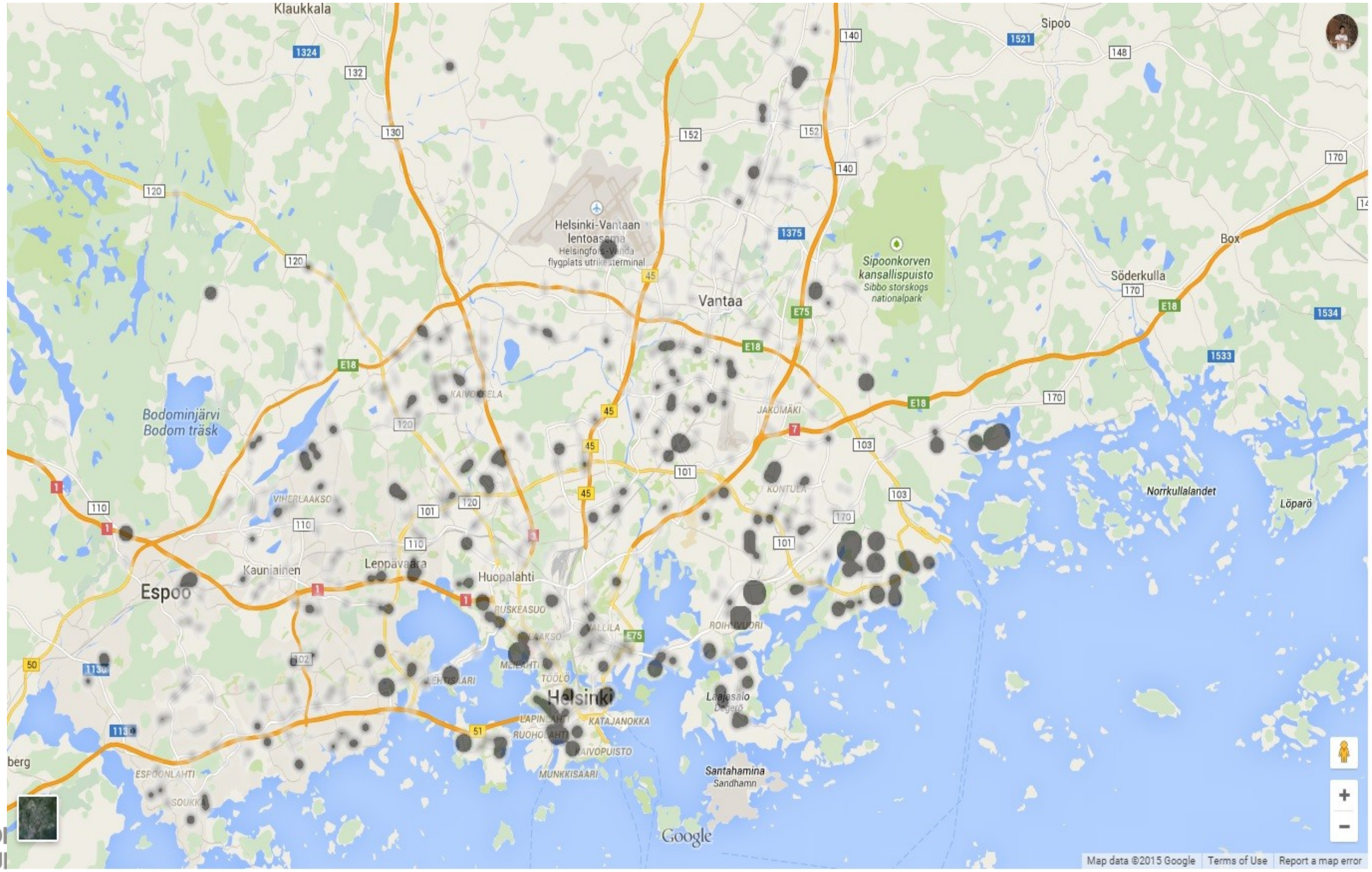
1. Key area causing the most traffic

We use **betweenness** centrality to quantify the important of the nodes (**Key Areas**).

Betweenness centrality quantifies the number of times a node acts as a bridge along the shortest path between two other nodes.

$$C_B(v) = \sum_{s \neq v \neq t \in V} \frac{\sigma_{st}(v)}{\sigma_{st}}$$

1. Key area causing the most traffic



1. Key area causing the city traffic

Most of the bus delay (**City Traffic**) happen during a certain **Hot Areas** during peak time.

- Railway station, metro station, shopping malls, workplace, education places, airports

2. Correlation between Urban Traffic and Bus Traffic (Why).

Why is there a bus delay? What cause it?

Did the bus traffic itself cause the urban traffic?

2. City is a Network

Bus delay (Urban traffic) over two bus stops Vs the number of buses (Bus traffic) traveling through that two bus stops.

We choose two peak time:
Monday between 8 am – 9 am
Monday at 4 pm – 5 pm

2. Akaike Weights Fitting distributions

Distribution	Probability density function (pdf)
Truncated Pareto	$Cx^{-\alpha}e^{-\lambda x}$
Log-normal	$\frac{1}{x\sigma\sqrt{2\pi}}exp[-\frac{(ln(x)-\mu)^2}{2\sigma^2}]$
Pareto	$(\alpha-1)x_{min}^{\alpha-1}x^{-\alpha}$
Exponential	$\lambda e^{-\lambda x}$

TABLE II

FITTED DISTRIBUTIONS

Akaike weights

Akaike's information criterion (AIC) is used in combination with Maximum likelihood estimation (MLE).
MLE finds an estimator that maximizes the likelihood function of one distribution. AIC is used to describe the best fitting one among all fitted distributions.

$$AIC = -2 log\left(L(\hat{\theta}|data)\right) + 2K.$$

Akaike weights

After determining the AIC value of each fitted distribution, we normalize these values as follows. First of all, we extract the difference between different AIC values called

$$\Delta_i = AIC_i - AIC_{min}.$$

Akaike weights

Then Akaike weights are calculated as follows,

$$W_i = \frac{exp(-\Delta_i/2)}{\sum_{r=1}^{R} exp(-\Delta_i/2)}.$$

Akaike Weights Fitting distributions

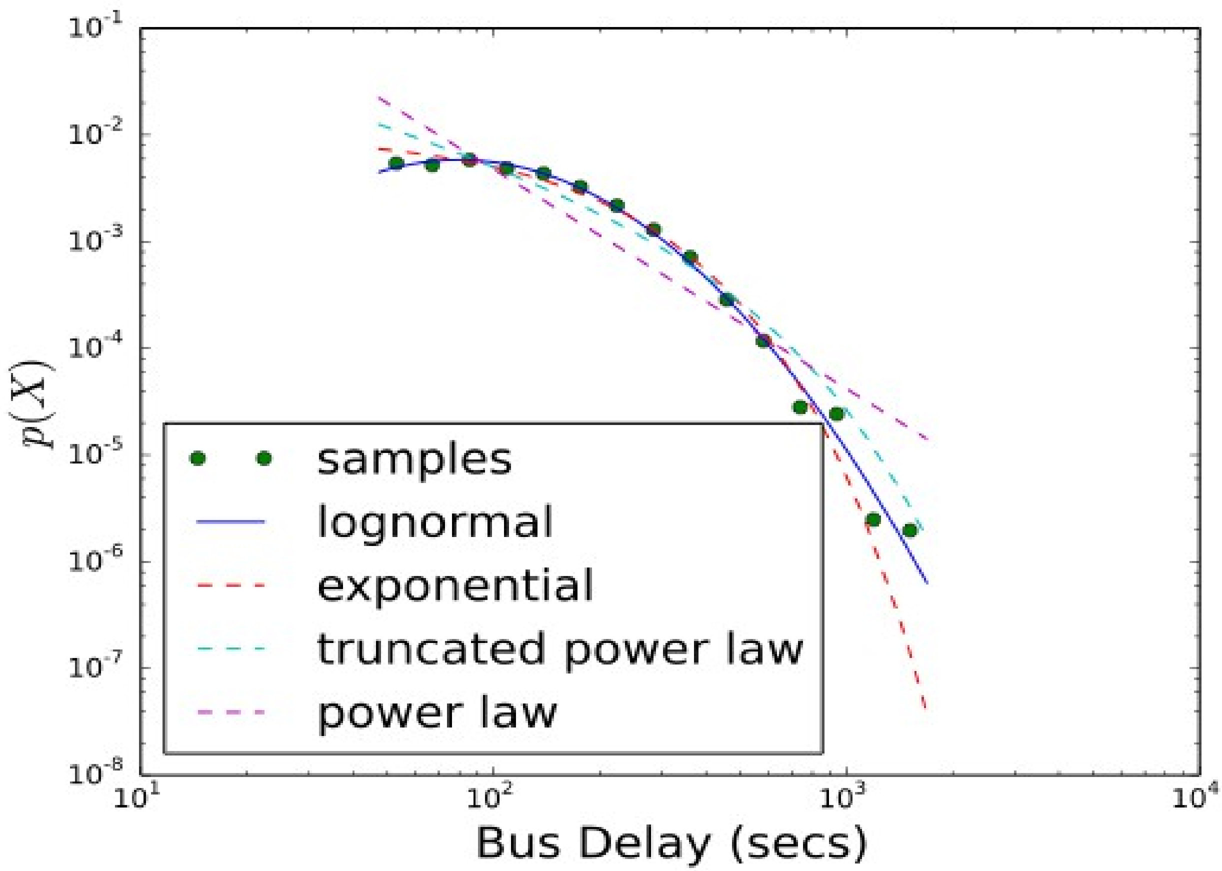
Distribution	Probability density function (pdf)
Truncated Pareto	$Cx^{-\alpha}e^{-\lambda x}$
Log-normal	$\frac{1}{x\sigma\sqrt{2\pi}}exp[-\frac{(ln(x)-\mu)^2}{2\sigma^2}]$
Pareto	$(\alpha-1)x_{min}^{\alpha-1}x^{-\alpha}$
Exponential	$\lambda e^{-\lambda x}$

TABLE II

FITTED DISTRIBUTIONS

2. Urban Traffic (Lognormal)

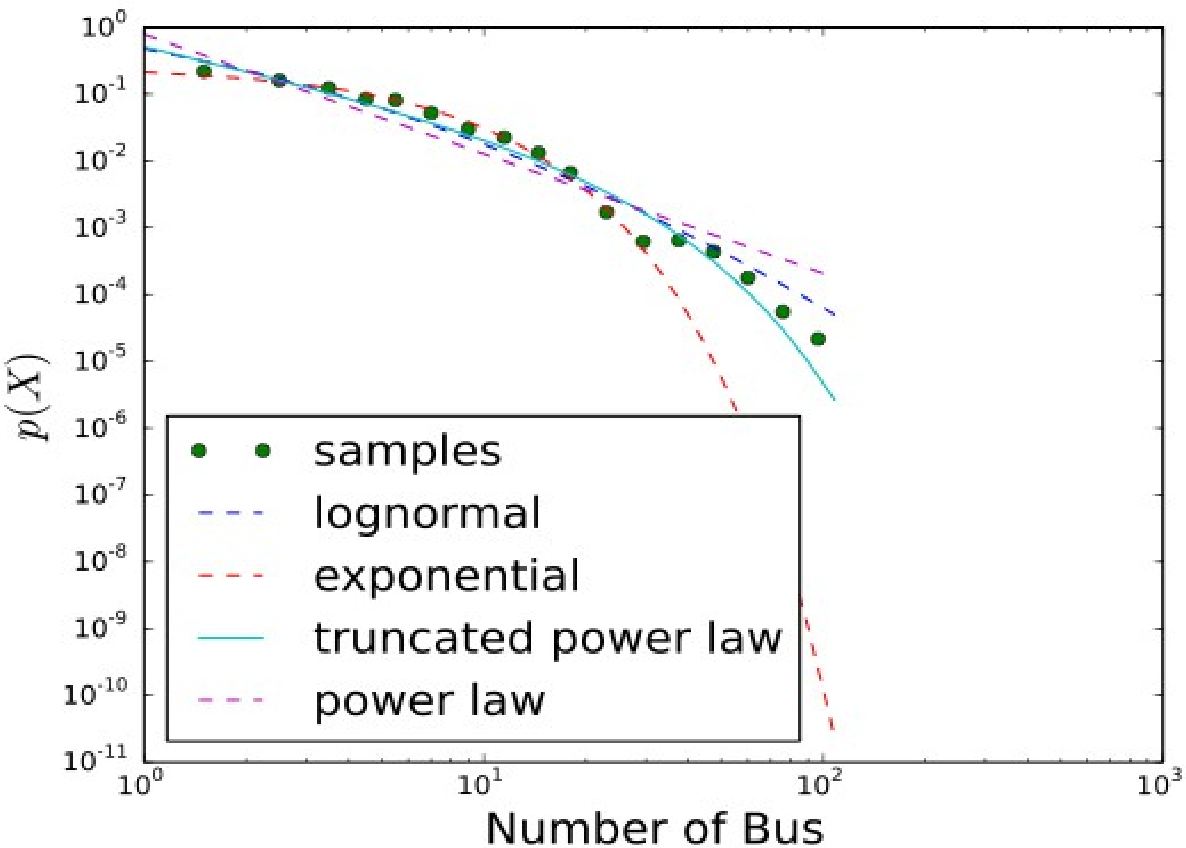
Lognormal Fit for Urban Traffic



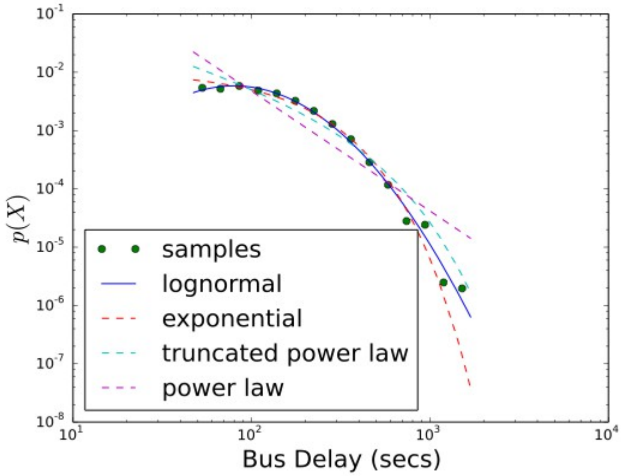
HELSINGIN YLIOPISTO
HELSINGFORS UNIVERSITET
UNIVERSITY OF HELSINKI

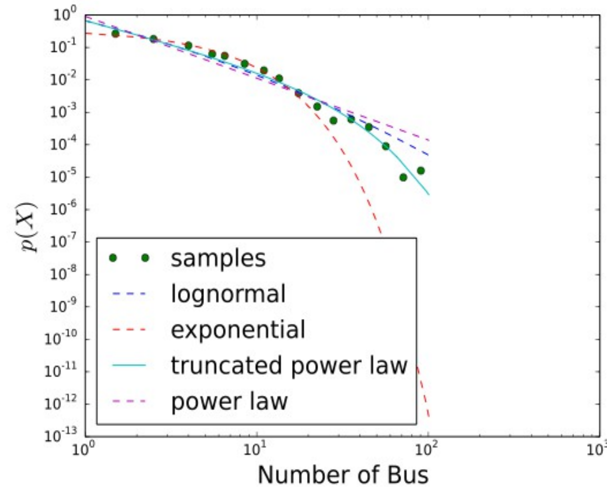
2. Bus Traffic (power-law)

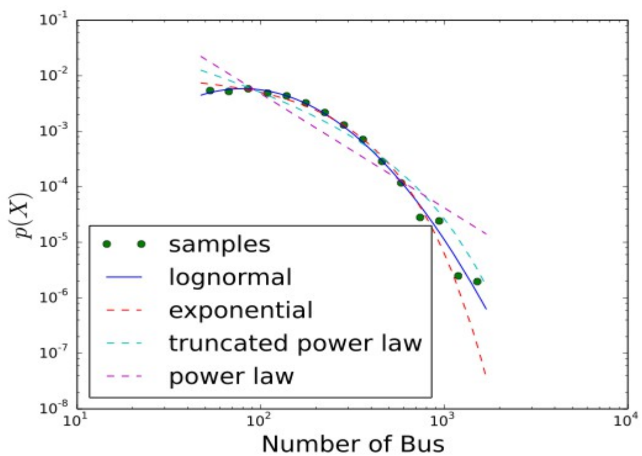
Power-law Fit for Bus Traffic

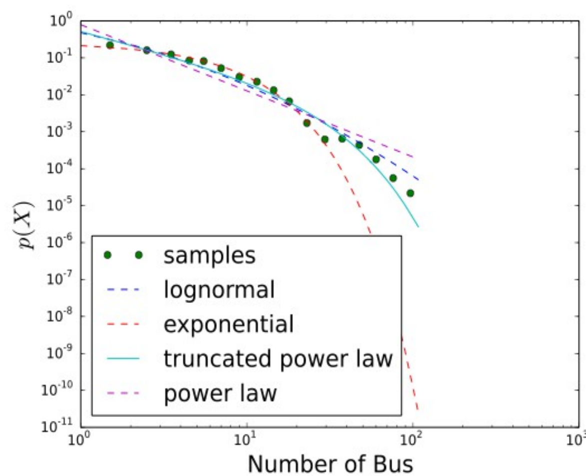


2. Bus Traffic is not Urban Traffic









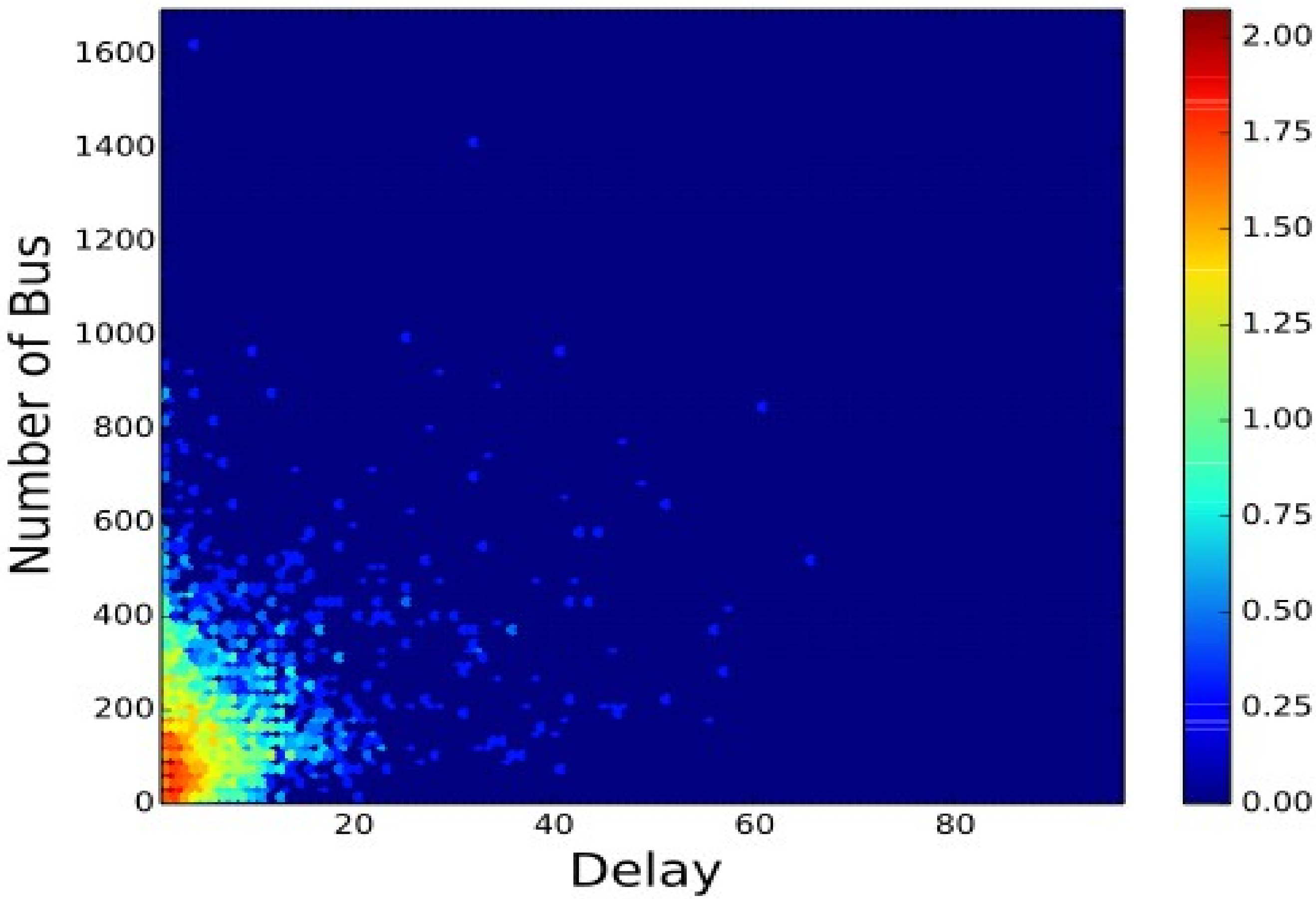
2. Pearson Correlation

The Pearson correlation is a measure of the linear correlation (dependence) between two variables X and Y.

$$\rho_{X,Y} = \frac{E[(X-\mu_X)(Y-\mu_Y)]}{\sigma_X \sigma_Y}$$

2. No correlation

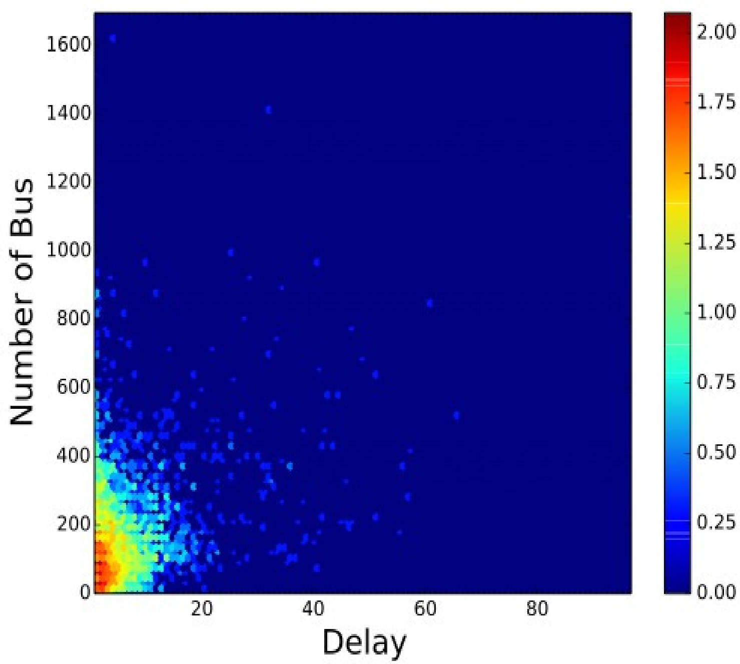
Between Bus Traffic and Urban Traffic

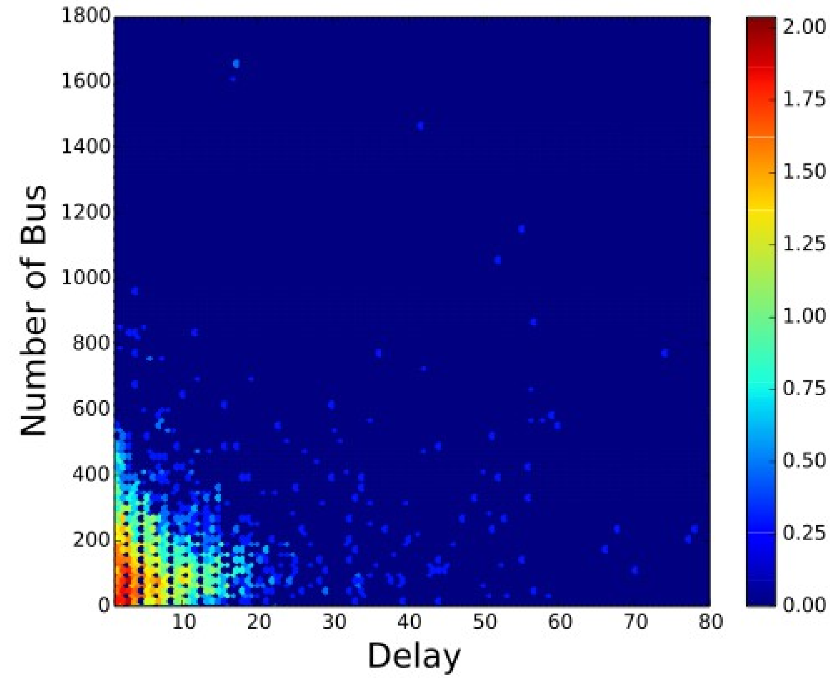


HELSINGIN Y
HELSINGFOR
UNIVERSITY

2. No correlation

Pearson value : 0.22403 and 0.13301





2. No correlation Betweer Bus Traffic and Urban Traffic

The bus delay is **not** caused by its own traffic, instead, it may be caused by:

- All other vehicles traffic
- The number of passengers getting on and getting off the bus

3. How to improve the urban transporation systems (How).

To improve the Urban Transportation systems:

- Better bus schedule plan in the **hot area**
  - Railway station, metro station, shopping malls, workplace, education places, airports

- Bus traffic itself is not an important cause of the urban traffic
  - Add more buses during the peak time
  - Reduce other vehicle usage, reduce picking-dropping time

Conclusion

We need better bus schedule plan in the hot areas

- Railway station, metro station, shopping malls, workplace, education places, airports

Bus traffic is not an important cause of the Urban Traffic

- Add more buses during the peak time

Aalto University

Thanks!