Human Environment and Transport Inspectorate
Ministry of Infrastructure and Water Management

Universiteit Leiden

Understanding dynamics of truck co-driving networks

Gerrit Jan de Bruin [1,2], Cor J. Veenman [1,3], H. Jaap van den Herik [1], and Frank W. Takes [1]

[1]Leiden University

[2]Ministry of Infrastructure and Water Management, The Netherlands

[3]Data Science Department, TNO

IDlab ILT 14 October 2019

Data

Human Environment and Transport Inspectorate
Ministry of Infrastructure and Water Management

Universiteit Leiden

- 900,000 trucks
- 18,000,000 measurements
  - License plate (identifier)
  - Time stamp
  - Location
  - Speed
  - Number of axes
  - Mass
- 17 measurement locations
- 2 years of measurements

Data

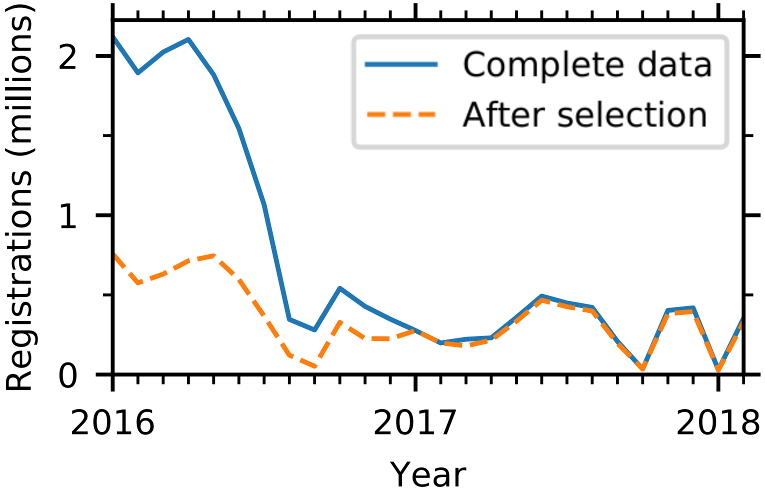
Human Environment and Transport Inspectorate
Ministry of Infrastructure and Water Management

Universiteit Leiden

- 900,000 trucks
- $18,000,000 \rightarrow 9,000,000$ measurements
  - License plate (identifier)
  - Time stamp
  - Location
  - Speed
  - Number of axes
  - Mass

  For Dutch trucks:
  - Maximum allowed mass
  - Truck place of origin
- $17 \rightarrow 8$ measurement locations
- 2 years of measurements

Measurement over time



Intro truck co-driving network

Human Environment and Transport Inspectorate
Ministry of Infrastructure and Water Management

Universiteit Leiden

When are trucks related?

Intro truck co-driving network

Human Environment and Transport Inspectorate
Ministry of Infrastructure and Water Management

Universiteit Leiden

When are trucks related?

- When trucks show similar behaviour.
- When trucks have shared attributes (i.e., same company, same sector, etc.).
- When they frequently show up together!

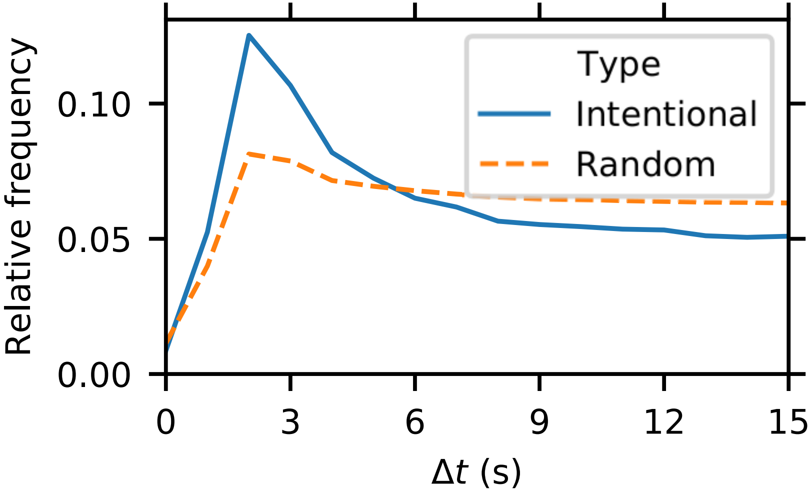
Intro truck co-driving network

A *co-driving event* is when two trucks appear at most 8 seconds[1] apart at the same location.

We define *intentional co-driving behaviour* between two trucks when there are at least two[2] co-driving events between two trucks, separated by more than 2 hours[3].

[1] This value is determined empirically.
[2] Idem.
[3] Idem.

Time gap between trucks



Intro network theory

$G = (V, E)$

- G: graph $\leftrightarrow$ network
- V: vertex $\leftrightarrow$ node $\leftrightarrow$ objects $\leftrightarrow$ actors $\leftrightarrow$ entities
- E: edge $\leftrightarrow$ link $\leftrightarrow$ ties $\leftrightarrow$ connections
- $\Gamma(u)$: neighbourhood of node u
- s: strength ($s_u = |\Gamma(u)|$)

Intro network theory

Human Environment and Transport Inspectorate
Ministry of Infrastructure and Water Management

Universiteit Leiden

$G = (V, E)$

- G: graph $\leftrightarrow$ network
- V: vertex $\leftrightarrow$ node $\leftrightarrow$ objects $\leftrightarrow$ actors $\leftrightarrow$ entities
- E: edge $\leftrightarrow$ link $\leftrightarrow$ ties $\leftrightarrow$ connections
- $\Gamma(u)$: neighbourhood of node u
- s: strength ($s_u = |\Gamma(u)|$)

Properties of most real-world networks:

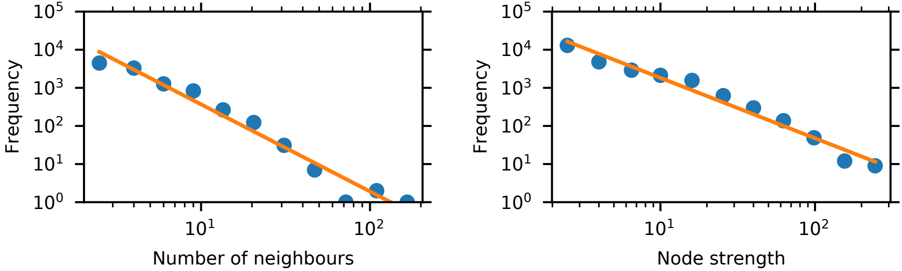
- Large GC (Giant Component):
  Most nodes in the network are connected
- Low average path
- Power-law degree distribution

Network Statistics

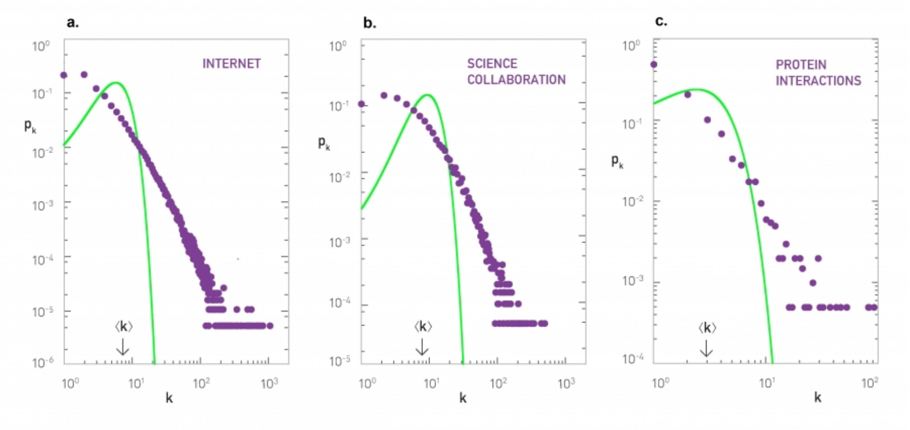
Table: Statistics of the truck co-driving network.

Metric	Full Network
Number of nodes	25,553
Number of links	73,059
Number of node pairs connected	27,986
Fraction nodes in giant component	62%
Fraction links in giant component	79%
Density	2.2×10^{-4}
Average shortest path length	7.8
Diameter	24

Degree and weight distribution



Degree and weight distribution



Previous work

Human Environment and Transport Inspectorate
Ministry of Infrastructure and Water Management

Universiteit Leiden

Understanding Behavioral Patterns in Truck Co-Driving Networks

Gerrit Jan de Bruin[1,2,3], Cor J. Veenman[1,4], H. Jaap van den Herik[2], and Frank W. Takes[1]

[1] Leiden Institute of Advanced Computer Science (LIACS), Leiden University
g.j.de.bruin@liacs.leidenuniv.nl
[2] Leiden Centre of Data Science (LCDS), Leiden University
[3] Human Environment and Transport Inspectorate (ILT), Netherlands Ministry of Infrastructure and Water Management
[4] Data Science Department, TNO

Abstract. This paper examines the co-driving behavior of truck drivers using network analysis. From a unique spatiotemporal dataset encompassing more than 10 million measurements of trucks passing 17 different highway locations in the Netherlands, we extract a so-called *co-driving network*. In this network, nodes are truck drivers and edges represent pairs of trucks that are systematically driving together. The obtained co-driving network structure has various properties common to real-world networks, such as a dominant giant component and a power law degree distribution. Moreover, network distance metrics and community detection reveal that the network has a highly modular structure. We furthermore propose a method for understanding the network community structure through attribute assortativity. Results indicate that co-driving links are mostly established based on geographical aspects: truck drivers from the same country or the same region in the Netherlands are more inclined to drive together. The resulting improved understanding of co-driving behavior has important implications for society and the environment, as trucks coordinating their driving behavior together help reduce traffic congestion and optimize fuel usage.

Keywords: co-driving networks, infrastructure networks, network analysis, community detection, assortativity

1 Introduction

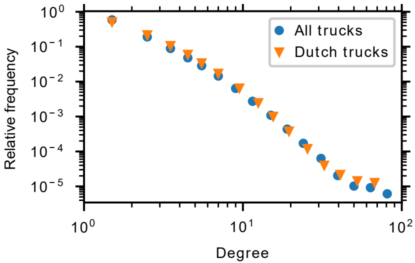
Techniques from the field of network science are used in a broad range of domains to extract knowledge from the network structure of real-world systems [1,12]. In this paper we use a network approach to investigate the factors that stimulate co-driving behavior of truck drivers. We furthermore look at what patterns can be found in groups of truck drivers who systematically drive together. To do so, we use network community detection [5] as well as various metrics related to assortativity (also known as mixing patterns, see [10]).

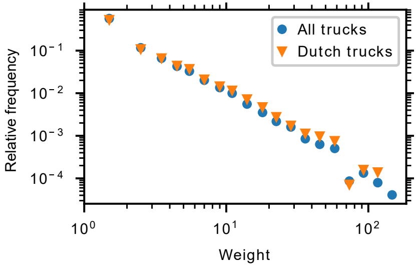
Making sense of communities using the attributes available in the data.

Available at:

`gerritjandebruin.nl/CN2018.pdf`

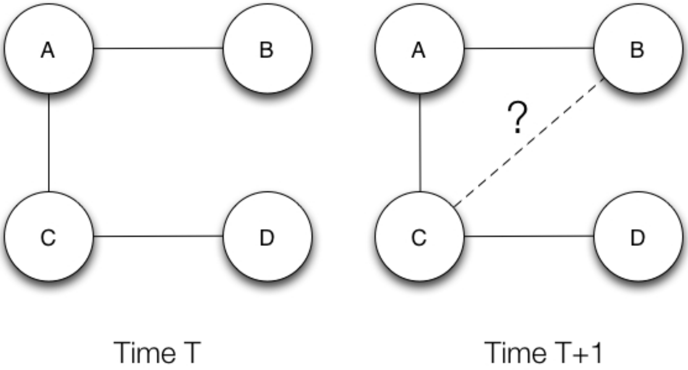
Degree and Weight Distribution





But what kind of links are present in this network?

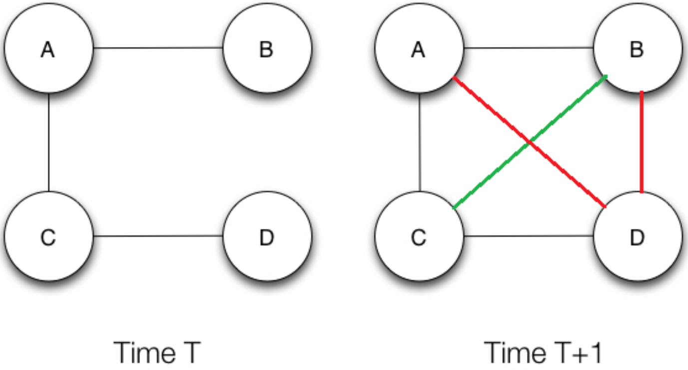
Link prediction



Source:

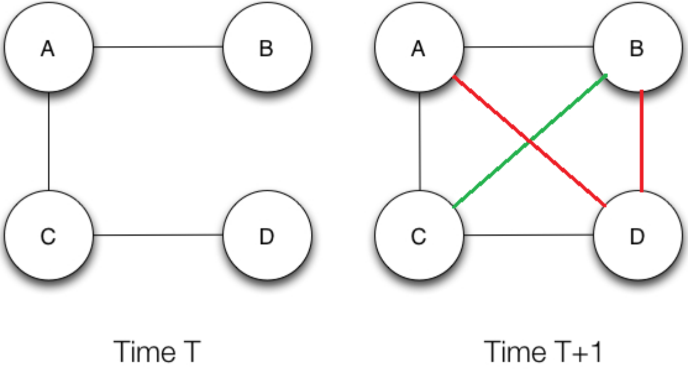
https://franktakes.nl

Link prediction



Source: `https://franktakes.nl`

Link prediction



Source: `https://franktakes.nl`

(A, D) 0

(B, C) 1

(B, D) 0

Random forest

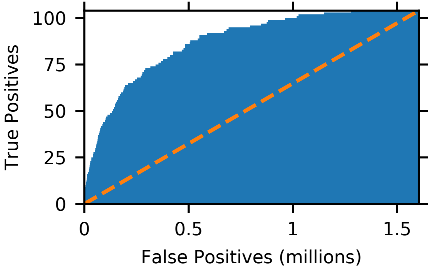


Figure: ROC curve of the random forest link prediction classifier. The AUROC is 0.84.

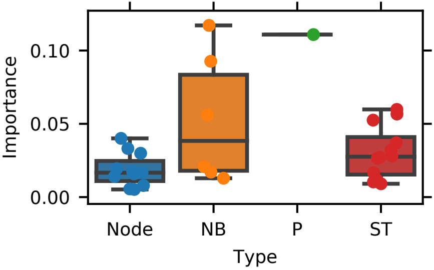


Figure: Gini feature importance. NB, P and ST are the neighbourhood, path and spatiotemporal features resp.

Features in link prediction

Human Environment and Transport Inspectorate
Ministry of Infrastructure and Water Management

Universiteit Leiden

Table 1: Features for truck pair (a, b) used in the link prediction model. The rightmost column lists the feature importance calculated using Gini importance as provided by the random forest classifier.

Index	Feature	Type	Importance
X_1	$truck\_country\,(a) = truck\_country\,(b)$	node	0.005
X_2	$truck\_ax\,(a) + truck\_ax\,(b)$	node	0.006
X_3	$\lvert truck\_ax\,(a) - truck\_ax\,(b)\rvert$	node	0.008
X_4	$truck\_len\,(a) + truck\_len\,(b)$	node	0.017
X_5	$\lvert truck\_len\,(a) - truck\_len\,(b)\rvert$	node	0.040
X_6	$truck\_mass\,(a) + truck\_mass\,(b)$	node	0.016
X_7	$\lvert truck\_mass\,(a) - truck\_mass\,(b)\rvert$	node	0.030
X_8	$driving\_hours\,(a) + driving\_hours\,(b)$	node	0.016
X_9	$\lvert driving\_hours\,(a) - driving\_hours\,(b)\rvert$	node	0.030
X_{10}	$weekend\_driver\,(a) + weekend\_driver\,(b)$	node	0.014
X_{11}	$\lvert weekend\_driver\,(a) - weekend\_driver\,(b)\rvert$	node	0.019
X_{12}–X_{19}	$last\_week_l\,(a+b)$ for $l = 1, \ldots, 8$	spatio-temporal	0–0.027
X_{20}–X_{27}	$last\_month_l\,(a+b)$ for $l = 1, \ldots, 8$	spatio-temporal	0–0.057
X_{28}–X_{45}	$last\_year_l\,(a+b)$ for $l = 1, \ldots, 8$	spatio-temporal	0.010–0.060
X_{46}	$\lvert\Gamma(a)\rvert + \lvert\Gamma(b)\rvert$	neighbourhood	0.117
X_{47}	$\Big\lvert \lvert\Gamma(a)\rvert - \lvert\Gamma(b)\rvert \Big\rvert$	neighbourhood	0.013
X_{48}	$\lvert\Gamma(a) \cup \Gamma(b)\rvert$	neighbourhood	0.093
X_{49}	$\lvert\Gamma(a) \cap \Gamma(b)\rvert$	neighbourhood	0.021
X_{50}	$s_a + s_b$	neighbourhood	0.056
X_{51}	$\lvert s_a - s_b\rvert$	neighbourhood	0.017
X_{52}	shortest path length in G	path	0.111

Thank you!

The paper, which will be published to Springer-Verlag, is available at:
`https://gerritjandebruin.nl/CN2019.pdf`.

This presentation is available at:
`https://gerritjandebruin.nl/ILT2019.pdf`.

Want to know more about network science?
Start at `http://networksciencebook.com`.

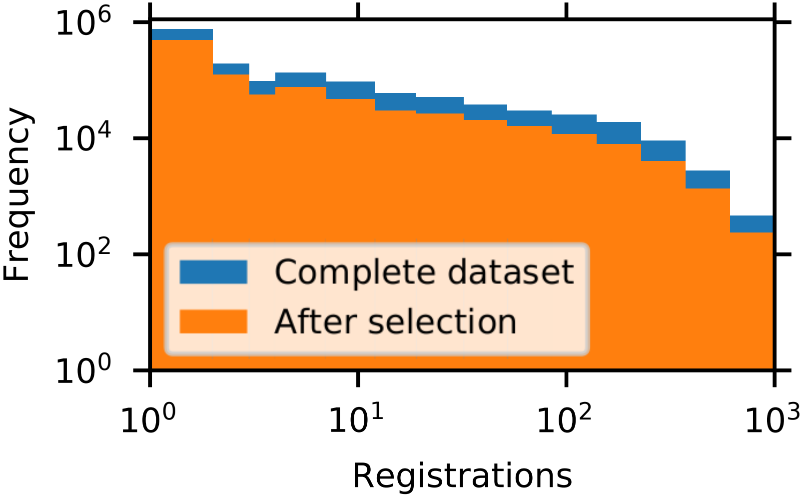
Measurement locations

Human Environment and Transport Inspectorate
Ministry of Infrastructure and Water Management

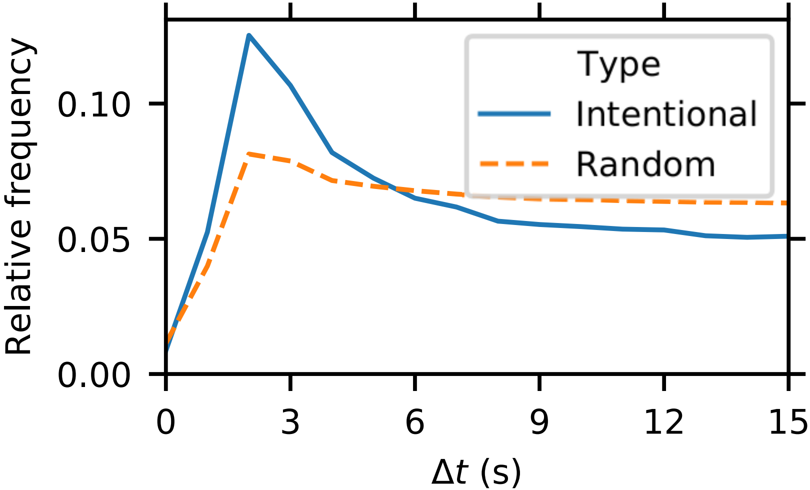
Universiteit Leiden

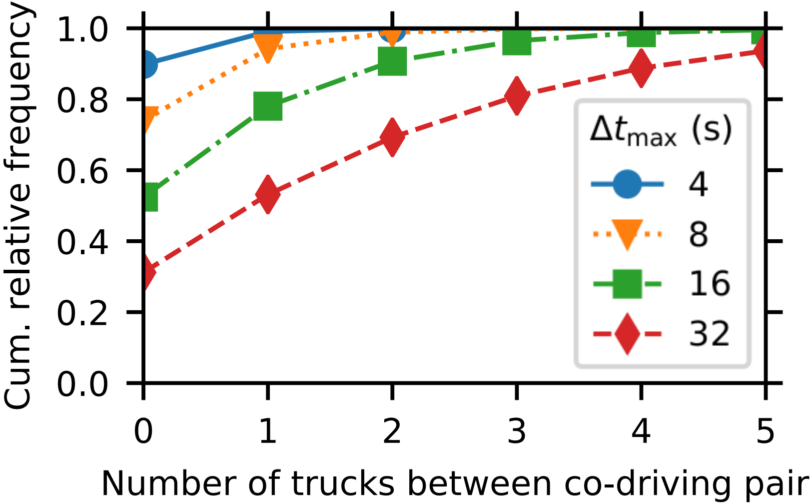
https://goo.gl/rNoitG

Number of registrations per truck



Time gap between trucks




Cum. relative frequency
1.0
0.8
0.6
0.4
0.2
0.0
0
1
2
3
4
5
Number of trucks between co-driving pair
$\Delta t_{\max}$ (s)
4
8
16
32

Human Environment and Transport Inspectorate
Ministry of Infrastructure and Water Management

Universiteit Leiden