Interdependency analysis of junctions for congestion mitigation in Transportation Infrastructures

George Stergiopoulos
INFOSEC Laboratory, Athens Univ. of Economics and Business
76 Patission Ave.
Athens, Greece GR-10434
geostergiop@aueb.gr

Evangelos Valvis
INFOSEC Laboratory, Athens Univ. of Economics and Business
76 Patission Ave.
Athens, Greece GR-10434

Foivos Anagnou-Misyris
INFOSEC Laboratory, Athens Univ. of Economics and Business
76 Patission Ave.
Athens, Greece GR-10434

Nick Bozovic
Athens Univ. of Economics and Business
76 Patission Ave.
Athens, Greece GR-10434
nbozovic@aueb.gr

Dimitris Gritzalis
INFOSEC Laboratory, Athens Univ. of Economics and Business
76 Patission Ave.
Athens, Greece GR-10434
dgrit@aueb.gr

ABSTRACT

The resilience of the Transportation road infrastructure network is of major importance, since failures such as prolonged road congestion in specific parts of the infrastructure often initiate major cascading effects that block transportation and/or disrupt services of other infrastructures over wide areas. Existing traffic flow analysis methods lack the ability to understand cascading effect of congestions and how to improve overall resilience in greater areas. Dependency risk graphs have been proposed as a tool for analyzing such cascading failures using infrastructure dependency chains. In this paper, we propose a risk-based interdependency analysis methodology capable to detect large-scale traffic congestions between interconnected junctions of the road network and provide mitigation solutions to increase traffic flow resilience. Dependency risk chains of junctions provide important information about which junctions are affected when other major junctions are congested in the road transportation network. Targeted mitigation mechanisms for traffic congestion can be proposed and the causes of bottlenecks can be analyzed to introduce road constructions or reparations with the best possible results in relieving traffic. We applied the proposed methodology on data collected by the UK government using cyber-physical traffic sensors over the course of 6 years. Our tool analyzed the UK major/A road transportation network, detected n-order junction dependencies and automatically proposed specific mitigation solutions to increase the overall resilience of the road infrastructure network. Simulation results indicate that detected mitigation options, if applied, can increase overall congestion resilience in wider areas of the network up to 12% by lowering likelihood of congestion.

SIGMETRICS'17, Urbana-Champaign, Illinois, USA

DOI: 10.1145/nnnnnnn.nnnnnnn

CCS CONCEPTS

•Mathematics of computing → Graph algorithms; •Information systems → Network data models; Data analytics; Data cleaning; **•Theory of computation → Network flows; •Computer systems organization → Embedded and cyber-physical systems;**

KEYWORDS

traffic congestion mitigation, graph analysis, risk assessment, cyber-physical systems, Transportation infrastructure

ACM Reference format:
George Stergiopoulos, Evangelos Valvis, Foivos Anagnou-Misyris, Nick Bozovic, and Dimitris Gritzalis. 2017. Interdependency analysis of junctions for congestion mitigation in Transportation Infrastructures. In *Proceedings of ACM SIGMETRICS International Workshop, Urbana-Champaign, Illinois, USA, 2017 (SIGMETRICS'17),* 6 pages.
DOI: 10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

The U.S. department of Homeland security identifies Highways and Motor Carriers as one of the seven key subsectors of the Transportation Systems Sector [3]. Intelligent transportation systems (ITS) aim at dealing with congestion in urban areas. Efficient use of existing road network systems results in reduced traffic congestion, emissions, energy consumption and cost. ITS technologies combined with Advanced Traveler Information Systems (ATIS), travel advisories, Variable Message Signs (VMS) and others attempt to relieve congestion and decrease travel time by selecting routes, departure times, and even the mode of transportation [17].

Traffic flow prediction models help both road users and transportation authorities improve traffic resilience, alleviate congestion or reduce traffic incidences in advance, by predicting the future states of traffic flow [9].

Dependency modeling, simulation and analysis of infrastructures have been studied extensively by researchers. Several methodologies and tools that focus on dependency analysis, estimate the impact [7], [19] or the risk derived from the dependencies within a critical infrastructure or among interdependent infrastructures [12], [14], [15], [13], [28]. The risk usually depends on two factors; the *likelihood* (or probability) of a negative event occurring and the *impact* (consequences) of that negative event. Such impact usually results in a complete non-operation or a partial malfunction in an infrastructure due to dependencies in infrastructure networks, usually called a *failure*.

1.1 Contribution

In this paper we utilize a previous time-based dependency analysis methodology for Critical Infrastructure dependency modeling [14], [15], [13] to analyze road network traffic flows. We apply the proposed methodology in a dataset with UK road traffic time flows, published by the UK Department of Transport. The dataset includes all major "A" roads and motorways, including journey times and traffic flows over a 6-year period for every 15-min intervals. Our two major contributions are:

(1) A methodology able to model road networks as dependency graphs and calculates the dependency risk between interconnected road junctions using assessment of traffic flow data. The methodology is able to assess the risk of congestion in junctions and calculate risk dependency paths (i.e. how congestion in one junction affects other connected junctions through roads). Road junctions are represented as graph vertices, with roads linking one junction to another as edges between the vertices. The methodology uses the min-max algorithm and statistical dynamic averages to calculate likelihood and impact of congestion.
(2) An analysis of UK's major/A-road transportation network, detecting n-order junction dependencies and automatically proposing junctions for specific mitigation solutions that increase the overall resilience of the entire network. The overall risk of cascading congestions was computed, which indicates the probability of impact transmission in a path through the connected edges. Specific junctions that initiate cascading traffic congestions were detected and simulation results indicated that applying traffic control on them may increase overall traffic flow resilience in wider areas while decrease risk of congetions up to 12%. We cross-referenced results with reports to check their validity.

2 RELATED WORK

Traffic flow simulation models designed to characterize the behavior of the complex traffic flow system have become an essential tool in traffic flow analysis and experimentation. Traffic flow models may be categorized according to [9]:

(1) Detail (submicro, microscopic, mesoscopic, macroscopic)
(2) Independence scale (continuous, discrete, mixed)
(3) Process representation (deterministic, stochastic)
(4) Application area (stretches, links, junctions)
(5) Type (traffic management, design, optimization)

A *short-term prediction* model should be precise in providing the vicinity measurements of traffic flow in the subsequent instants in the event of a traffic accident. On the other side, *long-term predictions* allow users to have a global insight of the traffic at any time, although the prediction will unlikely be accurate in the case of an atypical event (e.g. an accident). Specifications have ranged from Kalman filtering [18], [31], exponential filtering [20], non-parametric statistical methods [6], [23], spectral and cross-spectral analyses [24], [25] and sequential learning [5]. Nevertheless, a large amount of literature has been concerned with predictions from pure time-series models ranging from ARIMA [32], [16], [29] to dynamic generalized linear models [33], [34]. Daily mobility patterns have been used as a clustering stage in [17] to discover typicalities within the traffic flow data registered by road sensors, which permits building prediction models for each of such discovered patterns. Bayesian network framework is introduced to model the correlation structure of highway networks in the context of traffic forecast, as an optimization problem to identify optimal dependency structure [21].

Our methodology can be considered as a cyber-physical, deterministic Long-term optimization model that utilizes risk assessment, statistical methods and graph theory to promote decision making for increasing long-term resilience against congestion. To our knowledge, there is currently no solution able to calculate the risk of road dependencies between major "A" road junction in nation-wide networks and determine how the transportation network reacts to congestion mitigation measures; let alone quantify it.

3 BUILDING BLOCKS

The proposed methodology is based on a previous formed method, the multi risk dependency analysis [14], [15] and [13]. Furthermore, it extends CIDA, a CI dependency analysis tool [26].

3.1 Dependency analysis methodology

Dependency analysis assesses the risk of *n*th-order dependencies by applying results of organization-level risk assessments performed by critical infrastructure owners and operators. A dependency can be defined as a "one-directional reliance of an asset, system, network or collection thereof (within or across sectors) on an input, interaction or other requirement from other sources in order to function properly" [27]. The dependencies are modeled using a directional graph $G = (N, E)$ where N is a set of nodes and E is a set of edges. In this work, N is the set of the components of the modeled infrastructure (i.e., the junctions of road network) and E is the set of the links among the components (i.e., the roads among the junctions). The graph is directional to represent dependencies from one component to other components within the critical infrastructure. An edge from a node N_i to node N_j, i.e., $N_i \rightarrow N_j$, depicts the dependency relationship between them. The disruption transferred through this dependency can be described by the values impact $I_{i,j}$ and likelihood $L_{i,j}$. The combination of these two values indicates the dependency risk $R_{i,j}$ to component N_i due to its dependence on component N_j, which is denoted by the edge $N_i \rightarrow N_j$. The dependency risk is quantified as an integer scaled $[0 \ldots 5]$, with 0 representing no risk and 5 severe risk. This value as associated with each edge, refers to the level of cascade derived risk for the receiver due to the dependency. The results of traffic analysis are used as input to this method.

If $CI_{Y_0} \rightarrow CI_{Y_1} \rightarrow \ldots \rightarrow CI_{Y_n}$ is a chain of dependencies, $L_{Y_0,\ldots,Y_n}$ is the likelihood of the nth-order cascading effect and I_{Y_{n-1},Y_n} is the impact of the $CI_{Y_{n-1}} \rightarrow CI_{Y_n}$ dependency, then the cascading risk exhibited by CI_{Y_n} due to the nth-order dependency is computed as:

$$R_{Y_0,\ldots,Y_n} = L_{Y_0,\ldots,Y_n} \cdot I_{Y_{n-1},Y_n} \equiv \prod_{i=0}^{n-1} L_{Y_i,Y_{i+1}} \cdot I_{Y_{n-1},Y_n} \quad (1)$$

The cumulative dependency risk considers the overall risk exhibited. The cumulative dependency risk, denoted as $DR_{Y_0,Y_1,\ldots,Y_n}$ is defined as the overall risk produced by an nth-order dependency:

$$DR_{Y_0,\ldots,Y_n} = \sum_{i=1}^{n} R_{Y_0,\ldots,Y_i} \equiv \sum_{i=1}^{n} \left(\prod_{j=1}^{i} L_{Y_{j-1},Y_j} \right) \cdot I_{Y_{i-1},Y_i} \quad (2)$$

Eq. (2) computes the overall dependency risk as the sum of the dependency risks of the affected nodes in the chain due to a failure realized in the source node of the dependency chain. The risk computation employs a risk matrix that combines the likelihood and incoming impact values of each vertex in the chain. Interested readers are referred to [14] for additional details about dependency risk estimation.

The nth-order dependency risk is then calculated as the cumulative impacts on the affected nodes in the dependency chain.

3.1.1 The "Likelihood of Congestion" formal metric for road anticipating road congestion. Each relationship is assigned with a likelihood value, which declares, how likely the junction described by the current relationship is, to be congested. Intuitively, this value is a probability, based on which we can make predictions about each junction's state, at different times. In order for this to be calculated, we firstly need to check whether each relationship is proportionally fair to the other relationships describing the same junction. Generally, when talking about proportional fairness, we are referring to a system, in which two or more competitive entities are battling for resource control, and how we can maintain balance between them [11]. For example, in a computer network, our goal is to maximise the total throughput while at the same time allowing all the users to experience at least a minimal level of service.

The above is calculated according to [11] as follows :

A vector or rates x_r is proportionally fair if it is feasible, that is $x_r \geq 0$, and if for any other feasible vector x_r^*, the aggregate of proportional changes is zero or negative :

$$\sum_{r \in \mathbb{R}} \frac{x_r^* - x_r}{x_r} \leq 0 \quad (3)$$

In our case, x_r is the flow of the currently examined relationship, and x_r^* is the vector containing all the flow values for all other relationships referring to the same junction as x_r . We mark each of our relationships as "good" iff they satisfy (3), and "bad" otherwise. Note that he do not have to check for feasibility, since both vectors contain traffic flow values, which are always positive.

After marking all the relationships as either "good" or "bad", the likelihood value for a relationship R is calculated as follows :

$$L_R = \frac{Number\ of\ times\ R\ appears\ marked\ as\ \text{"bad"}}{Total\ number\ of\ times\ R\ appears} \quad (4)$$

3.1.2 The "Impact" metric for congested roads. Aside from the likelihood value, each relationship is assigned with an Impact value as well. As the name suggests, this metric declares how severe a possible congestion will be on that edge. Impact thresholds are calculated based on averages of best and worst case data entries. The metric is described in levels, ranging from 1 to 5, with 5 being the highest. The impact level for each relationship of a junction is calculated as follows :

(1) We create a $[1, 5]$ scale from our domain of flows (flow values for a junction).
(2) For each relationship, we rescale its flow value into the correct impact level in the range $[1, 5]$.

We used the well-known linear scaling algorithm to achieve the above. Its result depicts the number of units of the original interval which are equal to 1 unit of the new interval.

The approach described above is a dynamic one. There is no common scale for all the junctions. This happens, because in order to calculate the impact level for each junction, we rescale based on its own maximum and minimum flow, so each produced impact level describes how impactful a flow value is for the currently examined junction only. Thus, for each junction, the scale describing its impact levels adapts to its own characteristics (its min and max flow values).

3.2 Critical Infrastructure Dependency Analysis (CIDA) tool

CIDA [26] is a Critical Infrastructure Dependency Analysis tool that uses the Neo4j graph database [2]. According to recent empirical [4][10][22] studies, Neo4j outperforms other similar systems in load time. CIDA supports decision makers to analyze and assess dependency risk paths for complex and large-scale scenarios of interdependent CIs. CIDA takes as input the nodes and edges of a graph.

Empirical research [30] has revealed that cascading effects beyond fifth-order dependencies rarely affect critical infrastructures, fifth-order was the upper limit on the junction dependencies that were evaluated. Also, no cyclical paths were allowed as our purpose was modeling and examination of discrete, driver-intended real-world routes.

4 UK ROAD TRAFFIC CONGESTION ANALYSIS AND MITIGATION

First we describe the data set used in our analysis and then we describe the proposed methodology in detail.

4.1 Data set provider

The public data set was provided by the UK Government's Ministry of Transportation [1]. It is licensed under the Open Government License and provides average journey time, speed and traffic flow information for all 15-minute intervals since April 2009, on all motorways and 'A' roads managed by the UK's Highway Agency,

Table 1: Dataset cleaning problems.

Error ID	Error Description
1	Data quality 1 entries missing the Flow variable.
2	Data quality 1 entries missing Link Description variable.
3	Files repeating themselves after being over.

known as the Strategic Road Network. In the data set, journey times and speeds are estimated using a combination of cyber-physical sensors, including Automatic Number Plate Recognition (ANPR) cameras, in-vehicle Global Positioning Systems (GPS) and inductive loops built into the road surface. Journey times are derived from real vehicle observations and imputed using adjacent time periods or the same time period on different days.

4.1.1 Data validation. In ensuring the quality of the dataset used in experiments, and in addition to all already mentioned concerns and steps taken, the question of dataset balance arose. As our dataset mainly involves measurements over time the first step was to transform the data into a balanced panel by removing or extrapolating tuples to ensure that all nodes have measurements corresponding to the same intervals. It is important to note that since the task at hand is not that of a mainstay recognition classifier but rather a reliability analysis of dependencies that arise in the transportation infrastructure, the notion of a biased dataset is not as relevant as in a, for example an image classifier. It is our belief that, since our dataset is made of real traffic data spanning several years, any potential imbalance of classes is a real world representation and should not be tampered with. In addition, the amount of data was big enough to exclude unrealistic biases in data. Thus, instead of rebalancing the classes we opted to report all results of the experiments in f_1 score, $f_1 = 2 \times \frac{(precision \times recall)}{(precision + recall)}$, as that is a much better indicator of the overall performance of the prototype, by representing the impact of the correct identification of classes even when naturally unbalanced (i.e. A1 roads near London are congested much more often than on ruler areas).

Dataset entries are ranked by the UK ministry of transportation using a [1-5] scale; 1 being the high-quality sensor records and 5 the worst/potentially erroneous entries. We decided to not take into account the data set entries having data quality less than 2, in order to increase the consistency and integrity of our final results. The initial thought was that errors may exist only in those entries. Unfortunately, that was not the case. During the early stages of development, we came across several issues: Some entries from the data set having data quality 1 but were missing critical information or variables, and also the file's structure itself had some inconsistencies that was needed to be taken care of. For reference, we compiled a small table presenting the most important issues faced during data validation:

5 RESULTS

The CIDA tool generated all the dependency chains derived from the UK road dataset graph database. The process of path generation for existing graph databases had significant requirements in time and memory. For example, considering records for a specific road link for a 30-day, 24-hour month, there exist 720 edges for each junction link in the respective graph (30 days * 24 hours). The number of paths examined was about 5GB of text data, requiring thus too much computational time for the generation and a remarkable amount of temporary memory for the storage. To solve this, we utilized sub-graphs. A sub-graph could be described as an instance of a graph, which consists only of the desired properties of the initial graph.

Table 2: Top dependency chains per day - Dataset Statistics.

MOST OCCURRING JUNCTION	Average Risk M25 J17
M60 J15	2,654851485
MOST OCCURRING JUNCTION TOP 5% PATHS - TOP 50% RISK	Average Risk A36 - For all records
A36	4,40421875
MOST OCCURRING JUNCTION TOP 1% PATHS - TOP 64% RISK	Average Risk A607 - For all records
A607	3,766568627
COVARIANCE of 1st Junction with overall Risk	Multiple Regression R
0,297588933	0,502507459

A unique file was created for each timeslot of the month which was then aggregated for data analysis to examine critical congestion dependencies. We performed a thorough analysis of the entire dataset concerning all available historical real-world data for years 2009 through 2015. Preliminary analysis was performed on two different granularity levels for calculating dependency chains of junctions: average flow metrics per day, and average metrics per hour. This allowed us to: (a) compare an output dataset in the order of many gigabytes with a more useful one and see if detailed results agree with daily averages and (b) to have enough granularity to answer specific questions about day-to-day traffic jams and pinpoint potential extraordinary factors that affect average daily flows in the entire UK network. As an example, Figure 2 depicts the top 20 worst dependency chains detected for UK major road junctions from all analyzed years.

By parsing the entire dataset concerning the top ten paths from each day, statistical analysis showed that some junctions clearly appear more than others in congested dependencies and seem to greatly affect the entire UK network (or specific parts of it) in terms of traffic jams. Table 2 below depicts metrics calculated by analyzing the bulk of all dependency chains.

We calculated the Covariance of all likelihood and impact values produced by the presented methodology to test unintended likelihood-impact links. A perfect methodology should provide a low covariance rating, since likelihood and impact have distinct purposes; namely how often a congestion occurs and how big the impact (i.e. reduction of average speed) is. Fortunately, the covariance of the two variables was measured to be indeed low; around 14%. We consider this to be a robust result, since both variables utilize traffic flows for their calculation but, on the other hand, calculate different aspects of road traffic.

Multiple Regression (Multiple R) metrics measure the correlation between observed risk values of dependency paths and nodes presented in them. A value of 1 (100%) means a perfect positive relationship and a value of zero means no relationship at all. In fields that attempt to predict human behavior (such as psychology), it is entirely expected that R-squared values will be low, typically lower than 50%. Humans are harder to predict than, say, physical processes. Still, flow congestion in the transportation systems incorporated both the human factor and every-day patterns, for example some hours are clearly worse than others; say when people go to work.

Furthermore, the R-squared value is not high but the depicted results have statistically significant predictors. This can aid us to draw important conclusions about how changes in the predictor values are associated with changes in the response value. Regardless of the R-squared, the significant coefficients still represent the mean change in the response for one unit of change in the predictor while holding other predictors in the model constant. Obviously, this type of information can be extremely valuable.

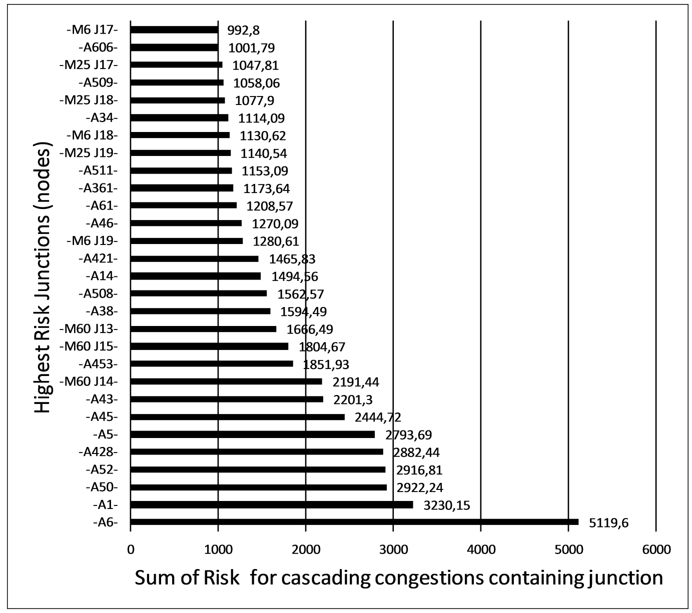

Figure 1: UK's Worst Risk Junction Hotspots.

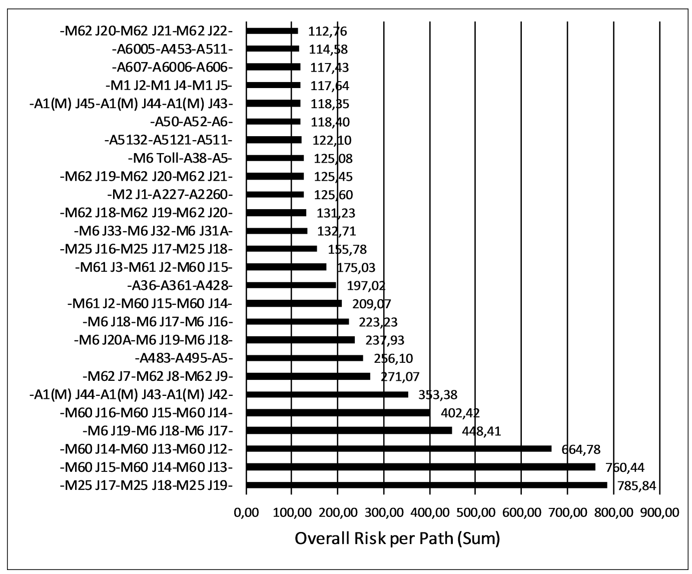

Figure 2: Top 20 Worst Dependency Paths(Depth=3)

5.1 Evaluation and comparison of results

We also utilized CIDA's results to detect the worst junction to affect traffic in the UK (and corresponding roads they connect). We cross-referenced our findings with a wide technical report published by INRIX [8]. Inrix reported the same TOP roads and corresponding junction detected by the CIDA tool. Namely, the worst major/"A" road traffic bottlenecks were located at London's M25, between Junctions 15 and 16 and between J21 (M1) and J21A (A405). CIDA provided the same results but also extended this detection also between Junctions 14-15 and 17 to 21/21A. INRIX also reports A38 N (M) junction with M6 (J6) which was also detected by CIDA, although in different top 10 positions (CIDA: 6-7th position, INRIX: 4th). Generally, INRIX and CIDA's top worst traffic hotspots were almost identical, with other entries being M60, M5, A50 etc., something that provides good evidence that CIDA depicts correct results and is also able to extend the INRIX list. Table **??** depicts CIDA's top worst traffic junctions and corresponding roads for the UK between 2009-2016.

CIDA though went further in detecting other junctions affected by traffic congestions in these roads/junctions, creating the entire dependency path table for the worst junctions and how traffic cascades into adjacent roads. We simulated applying mitigation works on junction to alleviate traffic as a reduction in the overall likelihood of a junction to initiate cascading traffic congestions. Applying traffic control in this way on suggested top worst initiating junctions showed an increase overall traffic flow resilience in wider areas of the network (and, consequently, a decrease in risk of congestion) up to 12%. We believe that this is a normal result, since Covariance analysis on initiating junctions (i.e. how much do congestions cascade in other parts of the road network) was found to be relatively high (30%), since this metric concerns the entire UK major network (covariance could reach as high as 70% for specific dependency junction chains).

Data calculated are enormous so only a small sample is given in Fig. 2.

REFERENCES

[1] Data.Gov.Uk. https://data.gov.uk/. (????). Accessed: 2017-05-30.
[2] Neo4j: The World's Leading Graph Database. (????). http://www.neo4j.org/
[3] U.S. Department of Homeland Security, Critical Infrastructure Sectors. https://www.dhs.gov/transportation-systems-sector. (????). Accessed: 2017-05-30.
[4] Shalini Batra and Charu Tyagi. 2012. Comparative analysis of relational and graph databases. *International Journal of Soft Computing and Engineering (IJSCE)* 2, 2 (2012), 509–512.
[5] Haibo Chen and Susan Grant-Muller. 2001. Use of sequential learning for short-term traffic flow forecasting. *Transportation Research Part C: Emerging Technologies* 9, 5 (2001), 319–336.
[6] Gary A Davis and Nancy L Nihan. 1991. Nonparametric regression and short-term freeway traffic forecasting. *Journal of Transportation Engineering* 117, 2

(1991), 178–188.
[7] Luisa Franchina, Marco Carbonelli, Laura Gratta, Maria Crisci, and Daniele Perucchini. 2011. An impact-based approach for the analysis of cascading effects in critical infrastructures. *International journal of critical infrastructures* 7, 1 (2011), 73–90. http://inderscience.metapress.com/index/U065X74041X23Q55.pdf
[8] INRIX Research Graham Cookson. 2016. Europe's Traffic Hotspots: Measuring the impact of congestion in Europe. *INRIX Company* (2016). http://www2.inrix.com/traffic-hotspots-research-2016?utm_medium=referral&utm_source=inrix&utm_campaign=roadway-analytics&utm_term=research-report&utm_content=press-release
[9] Serge P Hoogendoorn and Piet HL Bovy. 2001. State-of-the-art of vehicular traffic flow modelling. *Proceedings of the Institution of Mechanical Engineers, Part I: Journal of Systems and Control Engineering* 215, 4 (2001), 283–303.
[10] Salim Jouili and Valentin Vansteenberghe. 2013. An empirical comparison of graph databases. In *Social Computing (SocialCom), 2013 International Conference on*. IEEE, 708–715. DOI: http://dx.doi.org/10.1109/SocialCom.2013.106
[11] Frank P Kelly, Aman K Maulloo, and David KH Tan. 1998. Rate control for communication networks: shadow prices, proportional fairness and stability. *Journal of the Operational Research society* 49, 3 (1998), 237–252.
[12] Gerd H Kjølle, Ingrid Bouwer Utne, and Oddbjørn Gjerde. 2012. Risk analysis of critical infrastructures emphasizing electricity supply and interdependencies. *Reliability Engineering & System Safety* 105 (2012), 80–89. http://www.sciencedirect.com/science/article/pii/S095183201200021X
[13] Panayiotis Kotzanikolaou, Marianthi Theoharidou, and Dimitris Gritzalis. 2011. Interdependencies between critical infrastructures: Analyzing the risk of cascading effects. In *International Workshop on Critical Information Infrastructures Security*. Springer, 104–115.
[14] Panayiotis Kotzanikolaou, Marianthi Theoharidou, and Dimitris Gritzalis. 2013. Assessing n-order dependencies between critical infrastructures. *International Journal of Critical Infrastructures 6* 9, 1-2 (2013), 93–110.
[15] Panayiotis Kotzanikolaou, Marianthi Theoharidou, and Dimitris Gritzalis. 2013. Cascading effects of common-cause failures in critical infrastructures. In *International Conference on Critical Infrastructure Protection*. Springer, 171–182.
[16] S Vasantha Kumar and Lelitha Vanajakshi. 2015. Short-term traffic flow prediction using seasonal ARIMA model with limited input data. *European Transport Research Review* 7, 3 (2015), 1–9.
[17] Ibai Laña, Javier Del Ser, and Ignacio Iñaki Olabarrieta. 2016. Understanding daily mobility patterns in urban road networks using traffic flow analytics. In *Network Operations and Management Symposium (NOMS), 2016 IEEE/IFIP*. IEEE, 1157–1162.
[18] Iwao Okutani and Yorgos J Stephanedes. 1984. Dynamic prediction of traffic volume through Kalman filtering theory. *Transportation Research Part B: Methodological* 18, 1 (1984), 1–11.
[19] Benoit Robert. 2004. A method for the study of cascading effects within lifeline networks. *International journal of critical infrastructures* 1, 1 (2004), 86–99.
[20] Paul Ross. 1982. *Exponential filtering of traffic data.* Number 869.
[21] Samitha Samaranayake, Sébastien Blandin, and Alexandre Bayen. 2011. Learning the dependency structure of highway networks for traffic forecast. In *2011 50th IEEE Conference on Decision and Control and European Control Conference*. IEEE, 5983–5988.
[22] Bin Shao, Haixun Wang, and Yanghua Xiao. 2012. Managing and mining large graphs: systems and implementations. In *Proceedings of the 2012 ACM SIGMOD International Conference on Management of Data*. ACM, 589–592.
[23] Brian Lee Smith, Billy M Williams, and R Keith Oswald. 2000. Parametric and nonparametric traffic volume forecasting. In *Transportation Research Board 79th Annual Meeting*.
[24] Anthony Stathopoulos and M Karlaftis. 2001. Temporal and spatial variations of real-time traffic data in urban areas. *Transportation Research Record: Journal of the Transportation Research Board* 1768 (2001), 135–140.
[25] Anthony Stathopoulos and Matthew G Karlaftis. 2001. Spectral and cross-spectral analysis of urban traffic flows. In *Intelligent Transportation Systems, 2001. Proceedings. 2001 IEEE*. IEEE, 820–825.
[26] G Stergiopoulos, P Kotzanikolaou, M Theocharidou, and D Gritzalis. 2014. CIDA: Critical Infrastructure Dependency Analysis Tool. *Information Security and Critical Infrastructure Protection Laboratory, Department of Informatics, Athens University of Economics and Business, Athens, Greece (http://github.com/geostergiop/CIDA)* (2014). http://github.com/geostergiop/CIDA
[27] Her Majestyfis Treasury. 2011. National infrastructure plan 2011. *Infrastructure UK, UK Treasury Department (HM Treasury), http://cdn. hm-treasury. gov. uk/national_infrastructure_plan291111. pdf* (2011).
[28] Ingrid Bouwer Utne, Per Hokstad, and Jørn Vatn. 2011. A method for risk modeling of interdependencies in critical infrastructures. *Reliability Engineering & System Safety* 96, 6 (2011), 671–678. http://www.sciencedirect.com/science/article/pii/S0951832010002620
[29] Mascha Van Der Voort, Mark Dougherty, and Susan Watson. 1996. Combining Kohonen maps with ARIMA time series models to forecast traffic flow. *Transportation Research Part C: Emerging Technologies* 4, 5 (1996), 307–318.
[30] Michel Van Eeten, Albert Nieuwenhuijs, Eric Luiijf, Marieke Klaver, and Edite Cruz. 2011. The state and the threat of cascading failure across critical infrastructures: the implications of empirical evidence from media incident reports. *Public Administration* 89, 2 (2011), 381–400.
[31] Joe Whittaker, Simon Garside, and Karel Lindveld. 1997. Tracking and predicting a network traffic process. *International Journal of Forecasting* 13, 1 (1997), 51–61.
[32] Billy M Williams and Lester A Hoel. 2003. Modeling and forecasting vehicular traffic flow as a seasonal ARIMA process: Theoretical basis and empirical results. *Journal of transportation engineering* 129, 6 (2003), 664–672.
[33] HM Zhang. 2000. Recursive prediction of traffic conditions with neural network models. *Journal of Transportation Engineering* 126, 6 (2000), 472–481.
[34] Hongjun Zhang, Stephen Ritchie, and Zhen-Ping Lo. 1997. Macroscopic modeling of freeway traffic using an artificial neural network. *Transportation Research Record: Journal of the Transportation Research Board* 1588 (1997), 110–119.