Analytical formulation of the trip travel time distribution

Xiao Chen\textsuperscript{a}, Carolina Osorio\textsuperscript{b}

\textsuperscript{a}University of Coimbra
\textsuperscript{b}Massachusetts Institute of Technology

Abstract

This paper validates an analytical and tractable approximation of trip travel time variance for general topology networks. The main challenge in the derivation of such an approximation is to account for spatial between-link dependencies in an analytical yet also tractable manner. The approximation considered in this paper achieves tractability by extending Little's law for higher-order moments of finite capacity Markovian queueing networks. This paper presents a detailed validation of this approach. We compare the analytical approximations of path travel time variance to simulation-based estimates for two general topology queueing networks. Ongoing work, to be presented at the conference, uses this approach to address an urban traffic signal control problem that explicitly accounts travel time variability.

Keywords: Little's law; spillback; path travel time.

1. Introduction

Urban traffic management strategies are typically formulated such as to improve first-order performance metrics (e.g. expected trip travel times, expected link speeds). They have the potential to further enhance performance by accounting for higher-order distributional information, such as to improve, for instance, network reliability and network robustness. Enhancing the reliability of our networks is currently recognized as a critical goal in the US and in Europe [24, 25, 9]. The main challenge in addressing reliable or robust formulations of traditional transportation problems is the need to provide an analytical and tractable formulation of the trip travel time distribution, or of its first- and second-order moments. The complex between-link spatial-temporal dependency patterns makes analytical modeling of dependency, and hence accurate modeling of path metrics, a great challenge.

Numerous methods have been proposed for the approximation of expected trip travel time [23, 11, 3]. For a review, see Vlahogianni et al. [26]. The approximation of trip (or path) travel time variance has received less attention [6]. Nonetheless, empirical studies indicate the importance of path travel time variability in various travel...
choices such as departure time choice and route choice [28]. A recent stated preference (SP) survey indicates that for certain users the value of travel time variability is more than twice that of average travel time [4]. Another SP survey found that 54% of 564 morning commuters in Los Angeles (USA) considered travel time variability as either the most important or the second most important attribute in their commuting route choices [1].

Analytical approximations of path metrics are mostly derived based on simplifying, or even omitting, spatial temporal dependencies. The most common assumption is that of link independence [17], which typically underestimates path travel time variance for congested road networks [22]. Recent work that accounts for spatial between-link dependencies includes Westgate [27], Xing and Zhou [28], Charle et al. [6], Fu and Rilett [10], Rakha et al. [22].

Another approach to address reliable or robust traffic management problems is the use of stochastic traffic simulators. The latter can approximate the full distribution of the main network performance measures, making them suitable tools to design traffic management strategies that improve higher-order metrics. Recent work has designed a simulation-based optimization (SO) algorithm that allows designing traffic management strategies that improve both first-order (i.e., expectation) and second-order (i.e., variance) information of link travel times [8]. More importantly, the algorithm is computationally efficient, meaning that it can identify such strategies within few simulation runs.

In order to design a computationally efficient algorithm that embeds inefficient traffic simulators, information from other more efficient and tractable traffic models is used throughout the optimization process. In particular, information from the simulator is combined with information from an analytical differentiable traffic model. The latter combines ideas from traffic flow theory and queueing network theory [19, 18]. The role of the analytical model is to provide analytical problem-specific structural information to the generic SO algorithm. It is this combination of information from a simulation-based traffic model and an analytical traffic model that enables good short-term algorithmic performance.

We have extended this SO approach in order to address a signal control problem that aims at reducing both the expectation and the variance of link travel times [8]. The analytical traffic model used is a queueing network model, where each lane of an urban road is modeled as a set of finite space capacity M/M/1/K queues. Nonetheless, this work assumes between-link independence of travel times. The latter is a strong assumption that rarely holds for congested urban networks.

This paper considers an analytical traffic model that approximates the first- and second-order moments of trip travel time while accounting for the between-queue dependency. We validate the results using simulation-based estimates of trip travel time. As part of ongoing work we are embedding this analytical model within the SO framework of Osorio and Bierlaire [20] and using it to address a simulation-based signal control problem that improves trip travel time variability.

2. Methodology

Path travel time variance is defined as a function of the first- and second-order moments of path travel time. In queueing theory, Little’s law [14] describes the relationship between the expected number of users (e.g., travelers) in a network, \( E[L] \), and the expected time a user spends in the network \( E[W] \). The law is given by:

\[
E[L] = \lambda E[W],
\]

where \( \lambda \) is the arrival rate of users to the network.

Little’ law holds under very general conditions [15, 14]. Higher-order formulations of Little’s law have been proposed [13, 5, 12, 16]. For Poisson arrivals to the network, the higher-order law is given by:

\[
E[L(L - 1) \ldots (L - r + 1)] = \lambda^r E[W^r], r = 1, 2, ...\]

This higher-order law holds for a single M/ G/ 1 queue where: a) the arrival process is stationary; b) there is no overtaking (i.e., the first-in-first-out discipline holds); c) the time in the system of a user is independent of the arrival process of any arrivals after it. Keilson and Servi [13] proved that for a single M/ G/ 1 priority queue (infinite capacity and FIFO discipline) with two classes of customers, the second-order Little’s law holds for each class of
customers. More recently, Bertsimas and Nakazato [2] indicate that this law holds for certain queueing networks: overtaking-free queueing networks (such as tandem networks) with infinite space capacity queues.

To the best of our knowledge, the study of the relationship between higher-order moments of $L$ and of $W$ has not been considered for finite capacity queueing networks. These networks are of interest for urban transportation applications, since they account for the limited space capacity of roads. It is this limited space capacity that is at the origin of urban spillbacks and gridlocks. This paper presents an analytical and tractable extension of Little’s law for finite (space) capacity Markovian queueing networks. Our approach accounts for the spatial dependencies between adjacent queues. It approximates both the expectation and the variance of trip travel time in a general topology network.

We provide here a summary of the main ideas of the model. A detailed formulation is given in Chen and Osorio [7]. The proposed model maps an urban road network as a finite capacity queueing network. Each urban lane is represented by one or more $M/M/1/K$ queues. For a given set of queues that are adjacent (e.g., queues upstream and downstream of an urban intersection), we use a state-dependent formulation. We consider the joint state of all downstream queues, where a state specifies whether or not the downstream queue is full. We then propose state-dependent approximations for the arrival rate and the service rate of all adjacent queues.

3. Case Study

We consider two networks with different topologies. Network 1 consists of 8 queues and 3 different paths. The network topology is presented in Figure 1, the different paths information is shown in Table 1. We consider a set of scenarios with increasing levels of demand (i.e., increasing external arrival rates), in network 1, only queue 1 has external arrivals. The 5 scenarios are defined in Table 2, $i$ represents the index of each queue from 1 to 8, $\gamma_i$ means external arrival rate for queue $i$, $\mu_i$ means service rate at queue $i$, $k_i$ is the buffer size for queue $i$. Table 3 displays the transition probability matrix of the network, element $(i, j)$ of the matrix gives the probability of a user turning from queue $i$ to queue $j$.

Fig. 1. Network 1 topology

Table 1 Path information in network 1

| Path 1: queue 1 – queue 2 – queue 6 – queue 7 |
| Path 2: queue 1 – queue 3 – queue 4 – queue 6 – queue 7 |
| Path 3: queue 1 – queue 3 – queue 5 – queue 8 |

Table 2 Network and scenario specification for network 1

| $i$: | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| $\gamma_i$: | - | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| $\mu_i$: | 10 | 4 | 7 | 4 | 5 | 6 | 6 | 5 |
| $k_i$: | 8 | 8 | 8 | 8 | 8 | 8 | 8 | 8 |

Scenario | 1 | 2 | 3 | 4 | 5 |
$\gamma_1$: | 6 | 7 | 8 | 9 | 9.99 |
Network 2 consists of 10 queues and 5 different paths; its topology is displayed in Figure 2. Table 4 displays the 5 different paths information. We consider 5 scenarios with increasing levels of demand. These are defined in Table 5. In network 2, both queue 1 and queue 6 have external arrivals. Table 6 presents the transition probability matrix of the network.

| Q1 | Q2 | Q3 | Q4 | Q5 | Q6 | Q7 | Q8 |
|----|----|----|----|----|----|----|----|
| 0  | 0.3| 0.7| 0  | 0  | 0  | 0  | 0  |
| 0  | 0  | 0  | 0  | 1  | 0  | 0  | 0  |
| 0  | 0  | 0  | 0.4| 0.6| 0  | 0  | 0  |
| 0  | 0  | 0  | 0  | 1  | 0  | 0  | 0  |
| 0  | 0  | 0  | 0  | 0  | 1  | 0  | 0  |
| 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  |

Table. 4 Path information in network 2

| Path 1: queue 1 – queue 2 – queue 3 – queue 4 |
|---------------------------------------------|
| Path 2: queue 1 – queue 2 – queue 5 – queue 8 – queue 9 |
| Path 3: queue 1 – queue 2 – queue 5 – queue 10 |
| Path 4: queue 6 – queue 7 – queue 8 – queue 9 |
| Path 5: queue 6 – queue 7 – queue 10 |

Table. 5 Network and scenario specification for network 2

| i | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|----|
| $\gamma_i$ | 0 | 0 | 0 | 0 | - | 0 | 0 | 0 | 0 | 0  |
| $\mu_i$ | 10 | 10 | 6 | 6 | 6 | 10 | 9 | 9 | 9 | 9  |
| $k_i$ | 8 | 8 | 8 | 8 | 8 | 8 | 8 | 8 | 8 | 8  |

| Scenario | 1 | 2 | 3 | 4 | 5 |
|----------|---|---|---|---|---|
| $\gamma_1$ | 6 | 7 | 8 | 9 | 9.99 |
| $\gamma_6$ | 6 | 7 | 8 | 9 | 9.99 |
Table 6 Transition probability matrix for network 2

|     | Q1 | Q2 | Q3 | Q4 | Q5 | Q6 | Q7 | Q8 | Q9 | Q10 |
|-----|----|----|----|----|----|----|----|----|----|-----|
| Q1  | 0  | 1  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0   |
| Q2  | 0  | 0  | 0.4| 0  | 0.6| 0  | 0  | 0  | 0  | 0   |
| Q3  | 0  | 0  | 0  | 1  | 0  | 0  | 0  | 0  | 0  | 0   |
| Q4  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0   |
| Q5  | 0  | 0  | 0  | 0  | 0  | 0  | 0.5| 0  | 0  | 0.5 |
| Q6  | 0  | 0  | 0  | 0  | 0  | 0  | 1  | 0  | 0  | 0   |
| Q7  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0.5| 0  | 0.5 |
| Q8  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 1  | 0   |
| Q9  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0   |
| Q10 | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0   |

We validate the results derived from the proposed analytical approach to those estimated via stochastic simulation. We use a discrete-event simulation model of finite capacity Markovian networks [21]. For each network, the simulation estimates are obtained from 1000 replications, each with a total run time of 10,000 time units including a warm up period of 1,000 time units.

Figure 3 displays three plots. Each plot considers one of the 3 paths of network 1. Each plot considers the 5 demand scenarios (labeled 1 to 5 along the y-axis), and for each demand scenario the path travel time standard deviation is displayed. Three methods are compared: the simulation-based estimates (95% confidence intervals are plotted as solid lines), higher-order Little’s law approximations assuming between-link independence (represented by triangles), and higher-order Little’s law approximation accounting for between-link dependencies (this is the proposed approach, it is represented by stars).

All 3 plots of Figure 3 indicate that the accuracy of the traditional higher-order Little’s law decreases as demand increases. The proposed approach yields estimates that follow very similar trends to those of the simulation-based estimates.

The validation results for network 2 are displayed in Figure 4, there is one plot per path. The same as the plots for network 1, each plot considers 5 demand scenario for each path (5 paths in total). Three methods are compared. Just as for network 1, the accuracy of the traditional higher-order Little’s law decreases as congestion increases. For path 1 the traditional approach captures the trends better than the proposed approach. For all other paths, the proposed approach captures the same trends as the simulation-based estimates, and provides more accurate estimates of the trip travel time standard deviation. Overall for networks 1 and 2, the proposed approach yields good approximations for both low and high levels of congestion.
5. Conclusion

This work validates an analytical approximation of the standard deviation of path travel times, which accounts for between-link dependency, while also being scalable and tractable. The formulation is based on an approximate higher order Little’s law for finite capacity Markovian networks. The approximation is validated versus simulation-based estimates. The performance of the proposed model is also compared to that of a traditional higher-order Little’s law, which holds for infinite (space) capacity queues. As demand and congestion increase the proposed approach yields estimates that remain accurate, and have similar trends to the simulation-based estimates. The proposed approach yields good approximations for both low and high levels of congestion. Ongoing work uses this analytical model to address a simulation-based traffic signal control problem that improves trip travel time variability in congested urban networks.
Figure 4: Path travel time standard deviation for each of the 5 paths in network 2.
Acknowledgements

This work was partially funded by the Project EMSURE - Energy and Mobility for Sustainable Regions (CENTRO-07-0224-FEDER-002004).

References

[1] Abdel-Aty, A., K. R. and Jovanis, P. [1996]. Investigation of effect of travel time variability on route choice using repeated measurement stated preference data, 1493: 39–45.
[2] Bertsimas, D. and Nakazato, D. [1995]. The distributional Little’s law and its applications, Operations Research 43(2): 298–310.
[3] Bhaskar, A., Chung, E. and Dumont, A.-G. [2011]. Fusing loop detector and probe vehicle data to estimate travel time statistics on signalized urban networks, Computer-Aided Civil and Infrastructure Engineering 26(6): 433–450.
[4] Bonnaire, P. [2012]. Modeling Travel Time and Reliability on Urban Arterials for Recurrent Conditions, PhD thesis, UNIVERSITY OF SOUTH FLORIDA.
[5] Brumelle, S. L. [1972]. A generalization of L= λW to moments of queue length and waiting times, Operations Research 20(6): 1127–1136.
[6] Charle, W., Viti, F. and Tampere, C. [2010]. Estimating route travel time variability from link data by means of clustering, Proceedings of the 12th World Conference on Transport Research WCTR.
[7] Chen, X. and Osorio, C. [2014]. An analytical approximation of Little’s law for finite capacity Markovian networks, Technical report, Department of Civil and Environmental Engineering, Massachusetts Institute of Technology (MIT).
[8] Chen, X., Osorio, C. and Santos, B. F. [2012]. A simulation-based approach to reliable signal control, Proceedings of the International Symposium on Transportation Network Reliability (INSTR).
[9] Department of Transportation [2008]. Transportation vision for 2030, Technical report, U.S. Department of Transportation (DOT), Research and Innovative Technology Administration.
[10] Fu, L. and Rilett, L. R. [1998]. Expected shortest paths in dynamic and stochastic traffic networks, Transportation Research Part B: Methodological 32(7): 499–516.
[11] Geroliminis, N. and Skabardonis, A. [2011]. Identification and analysis of queue spillovers in city street networks, Intelligent Transportation Systems, IEEE Transactions on 12(4): 1107–1115.
[12] Haji, R. and Newell, G. F. [1971]. A relation between stationary queue and waiting time distributions, Journal of Applied Probability pp. 617–620.
[13] Keilson, J. and Servi, L. [1988]. A distributional form of Little’s law, Operations Research Letters 7(5): 223–227.
[14] Little, J. D. [1961]. A proof for the queuing formula: L= _W, Operations Research 9(3): 383–387.
[15] Little, J. D. C. [2011]. Little’s law as viewed on its 50th anniversary, Operations Research 59(3): 536–549.
[16] Marshall, K. T. and Wol , R. W. [1971]. Customer average and time average queue lengths and waiting times, Journal of Applied Probability pp. 535–542.
[17] Noland, R. B. and Polak, J. W. [2002]. Travel time variability: a review of theoretical and empirical issues, Transport Reviews 22(1): 39–54.
[18] Osorio, C. and Bierlaire, M. [2009a]. An analytic finite capacity queueing network model capturing the propagation of congestion and blocking, European Journal of Operational Research 196(3): 996–1007.
[19] Osorio, C. and Bierlaire, M. [2009b]. A surrogate model for traffic optimization of congested networks: an analytic queueing network approach, Technical Report 090825, Transport and Mobility Laboratory, ENAC, Ecole Polytechnique Federale de Lausanne. Available at: http://web.mit.edu/osoriooc/www/papers/osorBier09TechRepQgTraf.pdf.
[20] Osorio, C. and Bierlaire, M. [2013]. A simulation-based optimization framework for urban transportation problems, Operations Research 61(6): 1333–1345.
[21] Pirmin, M. [2007]. Simulation of finite capacity queueing networks, Transport and Mobility Laboratory, ENAC, Ecole Polytechnique Federale de Lausanne. Semester project.
[22] Rakha, H., El-Shwarby, I., Arafeh, M. and Dion, F. [2006]. Estimating path travel-time reliability, Intelligent Transportation Systems Conference, 2006. ITSC’06. IEEE, IEEE, pp. 236–241.
[23] Skabardonis, A. and Geroliminis, N. [2008]. Real-time monitoring and control on signalized arterials, Journal of Intelligent Transportation Systems 12(2): 64–74.
[24] Texas Transportation Institute [2012]. 2012 Urban mobility report, Technical report, Texas Transportation Institute (TTI), Texas A&M University System.
[25] Transport for London [2010]. Tra_c modelling guidelines. version 3.0, Technical report, Transport for London (TfL).
[26] Vlahogianni, E. I., Karlaftis, M. G. and Golias, J. C. [2014]. Short-term traffic forecasting: Where we are and where we’re going, Transportation Research Part C: Emerging Technologies.
[27] Westgate, B. S. [2013]. Vehicle travel time distribution estimation and map-matching via Markov chain monte carlo Methods, PhD thesis, Cornell University.
[28] Xing, T. and Zhou, X. [2011]. Finding the most reliable path with and without link travel time correlation: A Lagrangian substitution based approach, Transportation Research Part B: Methodological 45(10): 1660–1679.