Does extra information harm or hinder: Probabilistic and state-dependent routing in networks with selfish routing

Ilze Ziedins

Joint work with:
Heti Afimeimounga, Lisa Chen, Bruce Calvert, Mark Holmes, Wiremu Solomon, Niffe Hermansson, Alex Wang, Anne Buijsrogge

The University of Auckland
Introduction
Auckland, Monday, 8.30 a.m., predicted traffic
(downloaded 6 July 2013)
Auckland, Monday 10 June, 8.30 a.m., actual traffic
**Wardrop or user equilibrium**

The journey times on all the routes actually used are equal, and less than those which would be experienced by a single vehicle on any unused route.

Wardrop, J.G. (1952)

Each user has an infinitesimal effect on the system.
Network with $R$ routes from $A$ to $B$.
Which route to take?
Probabilistic routing – user optimal/equilibrium policies

\( p_r = \) probability of taking route \( r \), with \( p_r \geq 0 \), \( \sum_r p_r = 1 \).
\( \mathbf{p} = (p_1, p_2, \ldots, p_R) \)

\( W_r(\mathbf{p}) = \) expected transit time via route \( r \in R \).

At a user/ Wardrop equilibrium, \( \mathbf{p}^* \), there exists \( c \) such that

\[
W_r(\mathbf{p}^*) = c \quad \text{if } p_r^* > 0
\]

\[
\geq c \quad \text{if } p_r^* = 0.
\]
State dependent routing – user optimal/equilibrium policies

Let \( D(r; n) \) be the probability of taking route \( r \) in state \( n \).

A decision policy \( D = \{D(r; n), r \in R, n \in S : \sum_r D(r; n) = 1\} \).

For a policy \( D \in \mathcal{D} \) and \( n \in S \),

\( z^D_r(n) \) is expected time to reach the destination for an arrival, if
- system is in state \( n \) immediately prior to their arrival,
- every other arrival follows policy \( D \),
- and they choose to take route \( r \).

A policy \( D^* \in \mathcal{D} \) is a user optimal policy or user/Wardrop equilibrium if for each \( n \in S \)

\[
D^*(r; n) > 0 \implies z^D^*_r(n) \leq z^D^*_s(n) \text{ for all } s \neq r, s \in R.
\]
Downs-Thomson network
Downs-Thomson network

\[ Q_1: \text{1 server, } \mu_1 \]
\[ Q_2: \text{\(\infty\) server, } \mu_2 \]

Two Poisson arrival streams
- dedicated users to queue 2 at rate \(\lambda_2\),
- general users at rate \(\lambda\).

General users choose route
- either probabilistic or state-dependent routing.

\( Q_1 \) single server queue (\(\cdot|M|1\)), exponential service times, mean \(1/\mu_1\).
\( Q_2 \) batch service \(\infty\) server queue, batch size \(N (\cdot|G(N)|\infty)\),
generally distributed service times with mean \(1/\mu_2\).
• Single server queue – private transportation (e.g. cars).
  – delay increases as load increases

• Batch service queue – public transportation (e.g. shuttle bus).
  – delay decreases as load increases
  – frequency of service increases as load increases

• This version of model first proposed by Calvert (1997) as queueing network version of transportation model that gives rise to the Downs Thomson paradox.

• Afimeimounga, Solomon, Z (2005, 2010)
Downs-Thomson network –
probabilistic routing
$Q_1$: 1 server, $\mu_1$

$Q_2$: $\infty$ server, $\mu_2$

$Q_1$ single server $\cdot |M|1$ queue.

Expected delay $W_1 = \frac{1}{\mu_1 - \lambda p}$

$Q_2$ batch service, batch size $N$, $\infty$ server $\cdot |G(N)|\infty$ queue.

Expected delay $W_2 = \frac{1}{\mu_2} + \frac{N-1}{2(\lambda_2 + \lambda(1-p))}$

Both $W_1$ and $W_2$ are increasing in $p$. 
$\mu_1 = 0.8$

$\lambda = 1, \lambda_2 = .1, \mu_2 = 1, N = 3$

$W_1 \ldots \ldots \ldots \ldots$
\[ \mu_1 = 0.8 \]
\[ \lambda = 1, \lambda_2 = 0.1, \mu_2 = 1, N = 3 \]

\[ W_1 \quad \text{-- -- -- -- -- --}, \quad W_2 \quad \text{—— ———} \]
\[ \mu_1 = 0.8 \]
\[ \lambda = 1, \lambda_2 = 0.1, \mu_2 = 1, N = 3 \]

\[ W_1 \quad \text{---} \quad \text{---} \quad \text{---}, \quad W_2 \quad \text{————} \]
\[ \mu_1 = 0.8 \]
\[ \lambda = 1, \lambda_2 = .1, \mu_2 = 1, N = 3 \]
\[ W_1 \ldots \ldots \ldots \ldots, W_2 \]
\[ \mu_1 = 0.8 \]
\[ \lambda = 1, \lambda_2 = 0.1, \mu_2 = 1, N = 3 \]
\[ W_1 - - - - - - - - , W_2 \]
\[ \mu_1 = 0.8 \]
\[ \lambda = 1, \lambda_2 = .1, \mu_2 = 1, N = 3 \]
\[ W_1 \text{ , } W_2 \]
\[ \mu_1 = 0.8, 0.95 \]
\[ \lambda = 1, \lambda_2 = 0.1, \mu_2 = 1, N = 3 \]
\[ W_1 \quad \cdots \quad \cdots \quad \cdots \quad W_2 \]
\[ \mu_1 = 0.8, 0.95 \]
\[ \lambda = 1, \lambda_2 = 0.1, \mu_2 = 1, N = 3 \]
\[ W_1 \text{---}, W_2 \underline{---} \]
\[ \mu_1 = 0.8, 0.95, 1.05 \]
\[ \lambda = 1, \lambda_2 = .1, \mu_2 = 1, N = 3 \]

W_1 \hspace{1cm} W_2
$\mu_1 = 0.8, 0.95, 1.05$

$\lambda = 1, \lambda_2 = .1, \mu_2 = 1, N = 3$

$W_1$ --- --- --- -- -- --, $W_2$ ———
\[ \lambda = 1, \lambda_2 = .1, \mu_2 = 1, N = 3 \]
\[ W = p^* W_1 + (1 - p^*) W_2 \]
\[ \lambda = 1, \lambda_2 = .1, \mu_2 = 1, N = 3 \]

\[ W = p^*W_1 + (1 - p^*)W_2 \]

**Downs-Thomson paradox**: delays for all users can increase if capacity of private transport (roading) is increased. Downs(1962), Thomson(1977)
$\lambda = 1, \lambda_2 = .1, \mu_2 = 1, N = 3$

\[ W = p^*W_1 + (1 - p^*)W_2 \]

\[ W = \min_p \{pW_1 + (1 - p)W_2\} \]
Multiple equilibria

\[ \lambda = 1, \lambda_T = 0.1, \mu_T = 3, N = 2 \]

\[ W(p^*_R), \text{ delay at user equilibrium} \]

\[ \ldots \ldots \text{ delay at socially optimal equilibrium.} \]
Summary

- Delay at user equilibria under probabilistic routing can be arbitrarily large ($\rightarrow \infty$ as $\lambda_2 \rightarrow 0$).

- There may be multiple equilibria (for $N = 2$ when queue 1 is $\vert M \vert 1$, more generally when queue 1 is $\vert G \vert 1$).

- Probabilistic routing can be viewed as users having limited information about the system – mean delays at queues only.

- What happens if users have information about instantaneous state of network?

- Challenge in answering this question is that expected delay depends not just on current state of network, but also on decisions made by future arrivals (similar to networks with overtaking).

Downs-Thomson network –
state dependent routing
State dependent policies

\[ X_1(t) = \text{number of customers in queue 1} \]
\[ \text{(including customer in service)} \]
\[ X_2(t) = \text{number of customers waiting for service in queue 2} \]
\[ \text{(not including those in service)} \]

State space \( S = \mathbb{Z}_+ \times \{0, 1, 2, \ldots, N - 1\} \).

Process \( X_D \) operating under decision policy \( D \) has transition rates:

\[ \begin{array}{ll}
  \mathbf{n} \rightarrow & \begin{cases} 
    \mathbf{n} - e_1 & \text{at rate } \mu_1 I_{\{n_1 > 0\}} \\
    \mathbf{n} + e_1 & \text{at rate } \lambda D(1; \mathbf{n}) \\
    (n_1, (n_2 + 1) \mod N) & \text{at rate } \lambda_2 + \lambda D(2; \mathbf{n})
  \end{cases}
\end{array} \]

A policy \( D^* \in \mathcal{D} \) is a user optimal policy or user equilibrium if

\[ D^*(1; \mathbf{n}) > 0 \iff z_1^{D^*}(\mathbf{n}) < z_2^{D^*}(\mathbf{n}) \quad \text{for all } \mathbf{n} \in S. \]
Points with $D^*(1; \mathbf{n}) = 1$ are indicated by $\bullet$.
Points with $D^*(2; \mathbf{n}) = 1$ are indicated by $\circ$.
Unique user optimal policy, $N = 10, \lambda = 1.5, \lambda_2 = 0.5, \mu_1 = 2, \mu_2 = 1$.

A policy $D \in \mathcal{D}$ is monotonic if for all $\mathbf{n} \in S$, $D$ satisfies both

(A) $D(2; \mathbf{n}) \leq D(2; \mathbf{n} + e_1)$ and

(B) $D(2; \mathbf{n}) \leq D(2; \mathbf{n} + e_2)$
Main results

- A user optimal policy exists and is unique, no randomization needed.
- The user optimal policy is monotonic.
- The user optimal policy is monotonic in the parameters $\lambda$, $\lambda_2$, $\mu_1$, $\mu_2$ in the following sense. Let $X^{(1)}$ and $X^{(2)}$ be two processes, with common batch size $N$ and user optimal policies $D^*(1)$, $D^*(2)$ respectively. If $\lambda^{(1)} \geq \lambda^{(2)}$, $\mu_1^{(1)} \leq \mu_1^{(2)}$, $\lambda_2^{(1)} \geq \lambda_2^{(2)}$ and $\mu_2^{(1)} \geq \mu_2^{(2)}$, then $D_1^*(1) \subset D_1^*(2)$.
- Proofs use coupling arguments.
- As part of the proofs show monotonicity of $z_2^D(n)$ in $\lambda$, $\lambda_2$, $\mu_1$, $\mu_2$; and in the decision policy.
\( \lambda = 1, \lambda_2 = .1, \mu_2 = 1, N = 3 \)

\[
W = p^* W_1 + (1 - p^*) W_2
\]

\[
W = \min_p \{pW_1 + (1 - p)W_2\}
\]
Expected transit times under user optimal policy for state-dependent routing (———–), and probabilistic routing (−−−−−−−−)

\[ \lambda = 1, \lambda_2 = 0.1, \mu_2 = 1, N = 3 \text{ for } 0 \leq \mu_1 \leq 3. \]
Expected transit times under user optimal policy for state-dependent routing (———–), and probabilistic routing (−−−−−−−) 

\( \lambda = 1, \lambda_2 = 0.1, \mu_2 = 1, N = 3 \) for \( 0 \leq \mu_1 \leq 3 \).
Summary

- For this network possible to prove user equilibrium exists, is unique and monotonic.

- Numerical results indicate state dependent routing can mitigate worst effects of probabilistic routing. That is, more information leads to improved performance.

- Improvement in performance may be of practical interest – webcams, GPS navigation ....

- Does it hold for other networks?

Extensions/variations
• Partial information
• Two batch-service queues
• Processor-sharing queues
• More complex network – Braess’s paradox
• Multiple queues
Partial information

Average Delay

Information about Queue 1
Information about Queue 2
Probabilistic Routing

\[ \lambda = 1, \lambda_2 = 0.1, \mu_2 = 1, N = 3 \text{ for } 0 \leq \mu_1 \leq 3 \]

(Buijsrogge, 2014)
Two batch-service queues

Expected transit times under user optimal policy for state-dependent routing (———), and probabilistic routing (−−−−−−−−−−−).

\[ \lambda = 4, \lambda_1 = 3, \lambda_2 = 1, \mu_2 = 2, N_1 = N_2 = 5 \] for \( 0 \leq \mu_1 \leq 6. \)

Chen, Holmes, Z(2012)
Two processor sharing queues with finite capacity $N_1, N_2$

Non-randomized user/Nash equilibrium policies, using policy iteration. Light blue – monotonic, Dark blue – non-monotonic, Dark red – periodic and monotonic, Light red – periodic and nonmonotonic

$\mu_1 = 1, N_1 = N_2 = 3.$

Chen (2012), Hermansson (2015)
Braess’s paradox – original network

A: 1 server, $\mu_1$

B: $\infty$ server, $\mu_2$

C: $\infty$ server, $\mu_2$

D: 1 server, $\mu_1$
Braess’s network – augmented with additional route

\[ \lambda \rightarrow \]

A: 1 server, \( \mu_1 \)

\[ \rightarrow \]

C: \( \infty \) server, \( \mu_2 \)

D: 1 server, \( \mu_1 \)

B: \( \infty \) server, \( \mu_2 \)

Take

- Route 1 \( A \rightarrow B \) with probability \( p_1 \).
- Route 2 \( C \rightarrow D \) with probability \( p_2 \).
- Route 3 \( A \rightarrow D \) with probability \( p_3 \).

Cohen and Kelly (1990)
Braess’s paradox – probabilistic routing

Expected transit time: \( \mu_1 = \mu_4 = 2.5, \mu_2 = \mu_3 = 0.5 \)

The arrival rate, \( \lambda \), varies from 0.10 to 4.75

(\( \lambda = 5.00 \) is upper bound on capacity of the system). Cohen and Kelly (1990)
Braess’s paradox – state-dependent routing

...... initial network, _________ augmented network,
● = simulation estimate

Expected transit time with state-dependent routing:
\[ \mu_1 = \mu_4 = 2.5, \mu_2 = \mu_3 = 0.5 \]

(Calvert, Solomon, Z (1997))
Some final comments

• Do user equilibria exist more generally under state dependent routing, and if yes, when are they unique?

• How to overcome poor performance at user equilibria?

• Does more information lead to shorter delays in general?
  Effects of partial information

• Convergence issues – effect of delayed information.

• Differing information and/or policies for different customer classes


