Maximum entropy traffic matrix estimation

Keith Briggs and Geir Freysson

research.btexact.com/teralab/keithbriggs.html
research.btexact.com/teralab/geirfreysson.html

CRG presentation 2004 December 03 1600

typeset 2004 December 7 11:23 in pdfLaTeX on a linux system

The Problem

- Given:
  - *set of* N *nodes and* L *directed links*

The Problem

- Given:
  - *set of* N *nodes and* L *directed links*
  - *nodes labelled* $0, 1, \ldots, N-1$ *and links labelled* $0, 1, \ldots, L-1$

The Problem

- Given:
  - *set of* N *nodes and* L *directed links*
  - *nodes labelled* $0, 1, \ldots, N-1$ *and links labelled* $0, 1, \ldots, L-1$
  - *routing matrix* $A[s, d, l] =$ *fraction of traffic from source* s *to destination* d *passing through directed link* l

The Problem

- Given:
  - *set of N nodes and L directed links*
  - *nodes labelled $0, 1, \ldots, N-1$ and links labelled $0, 1, \ldots, L-1$*
  - *routing matrix $A[s, d, l] =$ fraction of traffic from source s to destination d passing through directed link l*
  - *$y[l] =$ traffic count on directed link l*

The Problem

- Given:
  - *set of* N *nodes and* L *directed links*
  - *nodes labelled* $0, 1, \ldots, N-1$ *and links labelled* $0, 1, \ldots, L-1$
  - *routing matrix* $A[s, d, l] =$ *fraction of traffic from source* s *to destination* d *passing through directed link* l
  - $y[l] =$ *traffic count on directed link* l
  - $\rho[n] =$ *source rate of node* n

The Problem

- Given:
  - *set of* N *nodes and* L *directed links*
  - *nodes labelled* $0, 1, \ldots, N-1$ *and links labelled* $0, 1, \ldots, L-1$
  - *routing matrix* $A[s, d, l] =$ *fraction of traffic from source* s *to destination* d *passing through directed link* l
  - $y[l] =$ *traffic count on directed link* l
  - $\rho[n] =$ *source rate of node* n

- To find:
  - $x[s, d] =$ *traffic from* s *to* d

The Problem

- Given:
  - *set of* N *nodes and* L *directed links*
  - *nodes labelled* $0, 1, \ldots, N-1$ *and links labelled* $0, 1, \ldots, L-1$
  - *routing matrix* $A[s, d, l] =$ *fraction of traffic from source* s *to destination* d *passing through directed link* l
  - $y[l] =$ *traffic count on directed link* l
  - $\rho[n] =$ *source rate of node* n
- To find:
  - $x[s, d] =$ *traffic from* s *to* d
- Assumptions:
  - *gravity model:* $g[s, d] = \rho[s]\rho[d]$

The Problem

- Given:
  - *set of N nodes and L directed links*
  - *nodes labelled $0, 1, \ldots, N-1$ and links labelled $0, 1, \ldots, L-1$*
  - *routing matrix $A[s, d, l]$ = fraction of traffic from source s to destination d passing through directed link l*
  - *$y[l]$ = traffic count on directed link l*
  - *$\rho[n]$ = source rate of node n*
- To find:
  - *$x[s, d]$ = traffic from s to d*
- Assumptions:
  - *gravity model: $g[s, d] = \rho[s]\rho[d]$*

The method

- $$\min_x \quad ||y - Ax||^2 + \lambda^2 \sum_{\substack{s,\,d \\ g[s,\,d] \neq 0}} x[s,d] \log(x[s,d]/g[s,d])$$

  $$\text{s.t.} \quad x[s,d] \geqslant 0$$

The method

- $$\min_{x} \quad ||y - Ax||^2 + \lambda^2 \sum_{\substack{s,\,d \\ g[s,\,d] \neq 0}} x[s,d] \log(x[s,d]/g[s,d])$$
  $$\text{s.t.} \quad x[s,d] \geqslant 0$$

- here $(Ax)_l = \sum_{s,d} A[s,d,l] x[s,d]$

The method

- $$\min_x \quad ||y - Ax||^2 + \lambda^2 \sum_{\substack{s,\,d \\ g[s,d] \neq 0}} x[s,d] \log(x[s,d]/g[s,d])$$
  $$\text{s.t.} \quad x[s,d] \geqslant 0$$
- here $(Ax)_l = \sum_{s,d} A[s,d,l]x[s,d]$
- KMB used a feasible sequential quadratic programming method for the optimization

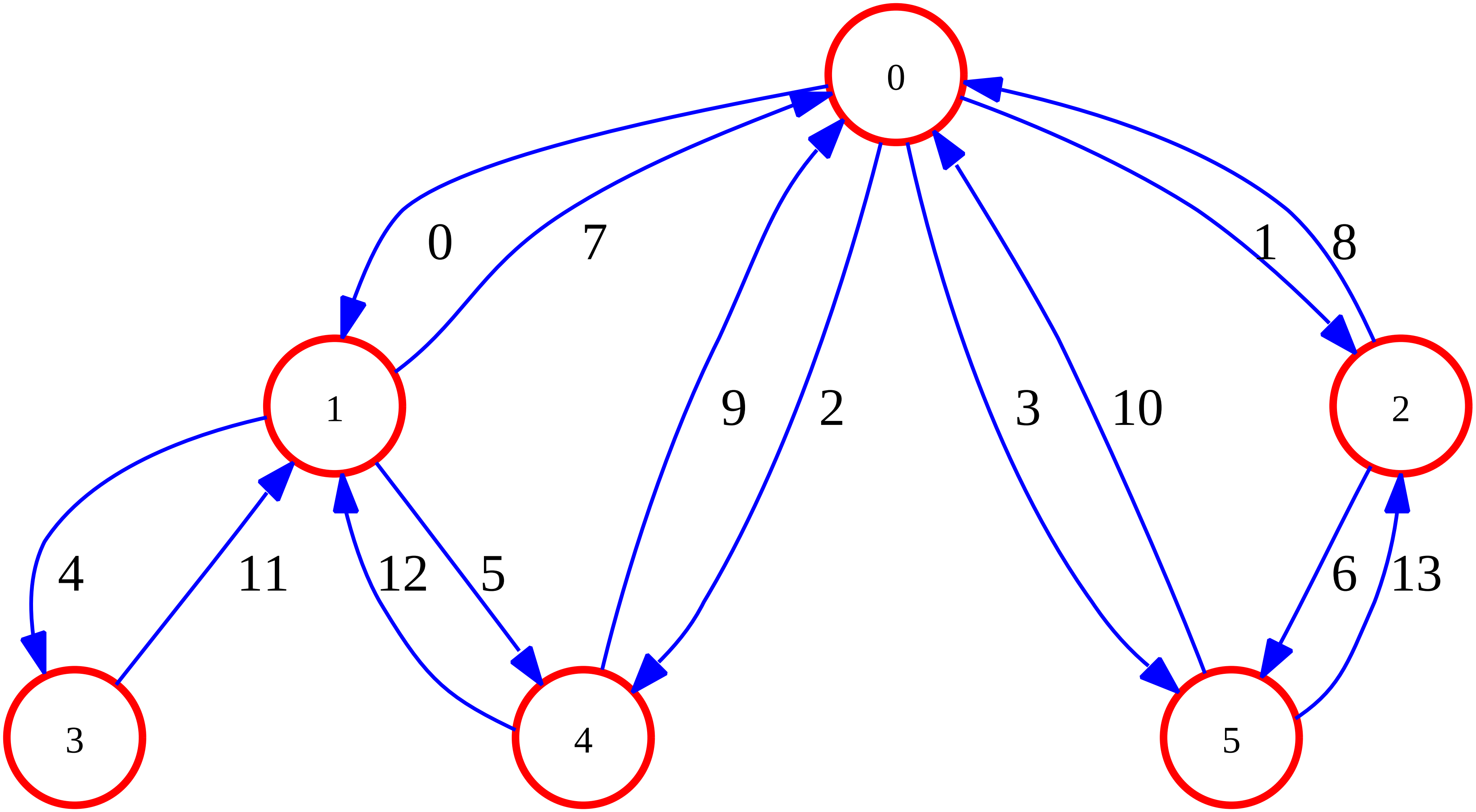
The method

- $$\min_{x} \; ||y - Ax||^2 + \lambda^2 \sum_{\substack{s,\,d \\ g[s,d] \neq 0}} x[s,d] \log(x[s,d]/g[s,d])$$
  s.t. $x[s,d] \geqslant 0$
- here $(Ax)_l = \sum_{s,d} A[s,d,l]x[s,d]$
- KMB used a feasible sequential quadratic programming method for the optimization
- Choosing the best value of λ is a problem

The method

- $$\min_x \quad ||y - Ax||^2 + \lambda^2 \sum_{\substack{s,\,d \\ g[s,d] \neq 0}} x[s,d] \log(x[s,d]/g[s,d])$$
  $$\text{s.t.} \quad x[s,d] \geqslant 0$$
- here $(Ax)_l = \sum_{s,d} A[s,d,l]x[s,d]$
- KMB used a feasible sequential quadratic programming method for the optimization
- Choosing the best value of λ is a problem
- Note: $x \log(x/g) = (x-g) + (x-g)^2/(2g) + O((x-g)^3)$

Example network



Example solution

- all routes shortest path

Example solution

- all routes shortest path
- traffic counts $y[n] = 5$

Example solution

- all routes shortest path
- traffic counts $y[n] = 5$
- source rates $\rho[i] = i+1$, for $i = 0, 1, \ldots, 5$

Example solution

- all routes shortest path
- traffic counts $y[n] = 5$
- source rates $\rho[i] = i+1$, for $i = 0, 1, \ldots, 5$
- with $\lambda = 1$, solution matrix x is

0.0000	0.5047	0.8727	0.3759	0.9086	0.9267
0.5047	0.0000	1.1974	1.0961	2.9307	1.2715
0.8727	1.1974	0.0000	0.8919	2.1554	5.1278
0.3759	1.0961	0.8919	0.0000	2.1829	0.9470
0.9086	2.9307	2.1554	2.1829	0.0000	2.2887
0.9267	1.2715	5.1278	0.9470	2.2887	0.0000

Example solution

- all routes shortest path
- traffic counts $y[n] = 5$
- source rates $\rho[i] = i+1$, for $i = 0, 1, \ldots, 5$
- with $\lambda = 1$, solution matrix x is

0.0000	0.5047	0.8727	0.3759	0.9086	0.9267
0.5047	0.0000	1.1974	1.0961	2.9307	1.2715
0.8727	1.1974	0.0000	0.8919	2.1554	5.1278
0.3759	1.0961	0.8919	0.0000	2.1829	0.9470
0.9086	2.9307	2.1554	2.1829	0.0000	2.2887
0.9267	1.2715	5.1278	0.9470	2.2887	0.0000

- rows are s, columns are d

Example solution

	0	1	2	3	4	5
0	0.000	0.505	0.873	0.376	0.909	0.927
1	0.505	0.000	1.197	1.096	2.931	1.272
2	0.873	1.197	0.000	0.892	2.155	5.128
3	0.376	1.096	0.892	0.000	2.183	0.947
4	0.909	2.931	2.155	2.183	0.000	2.289
5	0.927	1.272	5.128	0.947	2.289	0.000

ns-2 simulation

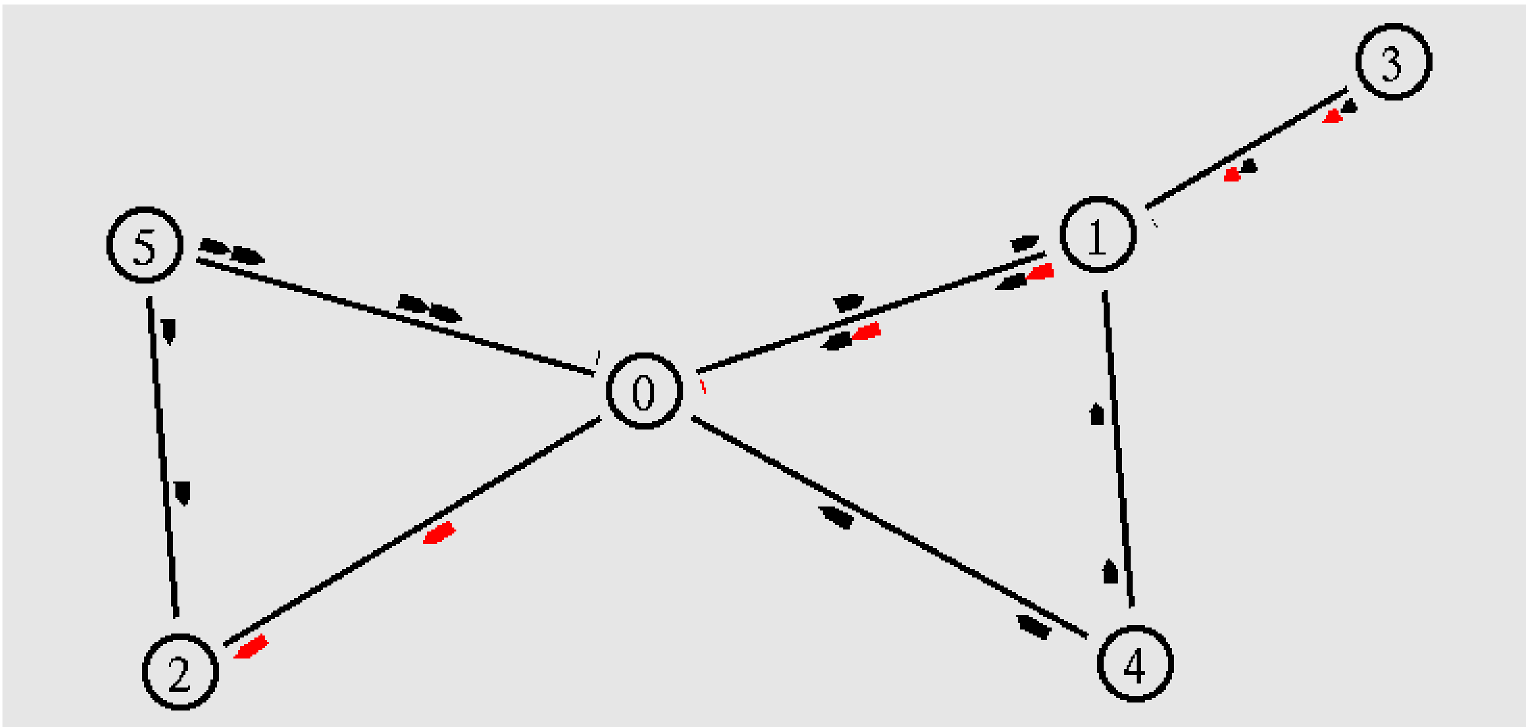
- Network simulated on ns-2

ns-2 simulation

- Network simulated on ns-2
- Shortest path currently used but others also available

ns-2 simulation

- Network simulated on ns-2
- Shortest path currently used but others also available
- Can measure all traffic on large network to test accuracy of maxent method



References

[zhang03] Y Zhang, M Roughan, C Lund & D Donoho *An information-theoretic approach to traffix matrix estimation* SIGCOMM03

[ns-2] www.isi.edu/nsnam/ns/