A complex cocktail of networks and reality

Simon Dobson

School of Computer Science
— *and* —
St Andrews Institute for Data-Intensive Research

`simon.dobson@st-andrews.ac.uk`

University of St Andrews | FOUNDED 1413

Introduction

What this talk is about

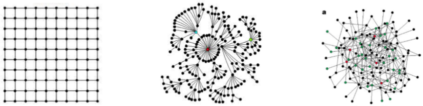
- This is the story of a paper
- . . . and what it tells us about the interface between network science, data science, and scientific applications

My aim

- Using real-world data to drive abstract study
- The limitations we hit as a result
- How we might address these limitations

Complex networks

Networks and processes



Living between regularity and randomness [1]

- Heterogeneous degree distribution, fragile notion of "neighbourhood"
- Evaluate processes at each node, affecting behaviour of neighbours, often with a stochastic component
- Canonical example is the SIR model of disease propagation

[1] A.-L. Barabási and R. Albert. Emergence of scaling in random networks. *Science*, 286(5439):509–512, 1999. URL doi://10.1126/science.286.5439.509

Multiplexes

A *multiplex* (or more correctly a *multilayer network*) is a collection of two (or more) networks [2]

- Nodes in the different networks are *coupled*
- Study properties of the individual networks or of the ensemble
- One network may be “less wide” than the other, and so offer “shortcuts” for processes

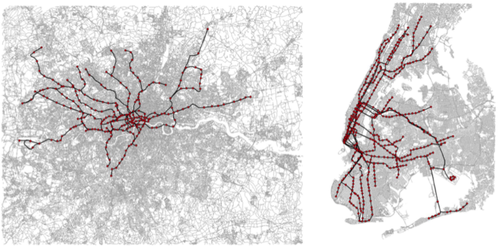
[2] M. Kivelä, A. Arenas, M. Barthélemy, J. Gleeson, Y. Moreno, and M. Porter. Multilayer networks. *Journal of Complex Networks*, 2(3):203–271, 2014. URL doi://10.1093/comnet/cnu016

Studying the real world

The problem: Urban transportation

Coupled transport networks [3]

- Street and tube/subway form a multilayer network
- How does the addition of the tube affect travel times?
- How does this change as the tube speeds up?



[3] M. Batty. *The new science of cities*. MIT Press, 2013. ISBN 978-0-292-01952-1

Framing the problem

Our study [4]

- Simplify to treat as a purely topological problem
- This study doesn't include explicit consideration of flows (although later work does)
- Study the *betweenness centrality* of nodes as the relative speeds of the two networks changes
- How does *outreach* change?

[4] E. Strano, S. Shai, S. Dobson, and M. Barthélemy. Multiplex networks in metropolitan areas: generic features and local effects. *Journal of the Royal Society Interface*, 12(111), October 2015. URL doi://10.1098/rsif.2015.0651

Setting up the study

Acquired street and tube data from Open Street Map

- London: $N_s = |V_s| = 325K$ street nodes, $N_t = |V_t| = 263$ tube nodes
- New York: $N_s = 68K$, $N_t = 454$

Data hygiene

- Needed substantial manual cleaning
- Streets don't meet, tubes don't come up where they should, …
- Some slight simplifications, *i.e.*, tubes always emerge at street junctions

Methodology

Travel costs

- $\tau_s(i,j)$ the travel cost (in time units) between $i,j \in V_s$ using only street edges
- $\tau_m(i,j)$ the travel cost using the multiplex (street and tube)
- $0 \leq \beta \leq 1$ the ratio of speed between street and tube (tube is faster for smaller β)

Shortest paths

- $\sigma_{i,j}$ the number of shortest paths between $i,j \in V_s$ using only the street network
- Similarly define $\sigma^m_{i,j}$ the number of shortest paths using the mulitiplex

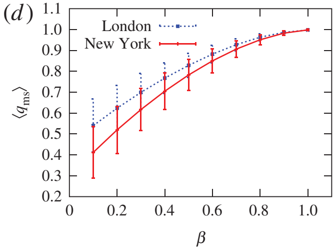
How much does the tube affect travel costs?

Metric

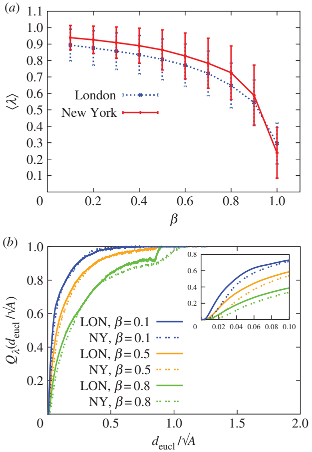
- Ratio of travel costs from a node $i \in V_s$ to all other nodes using the multiplex *vs* using the streets only

$$q_{ms}(i) = \frac{1}{N_s - 1} \sum_{j \in V_s} \frac{\tau_m(i,j)}{\tau_s(i,j)}$$

- Halving β reduces $\langle q_{ms} \rangle$ by about 20%



Interdependence

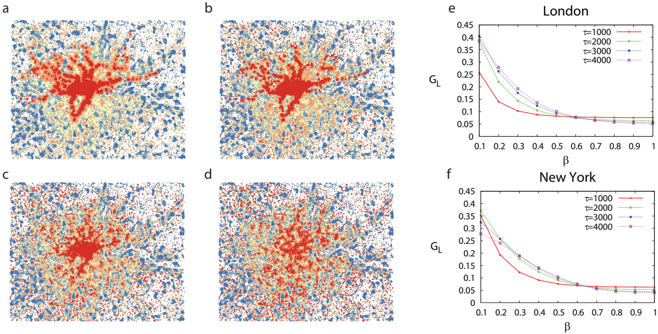


Ratios of shortest paths

- Already known in the planning literature as *inter-modal connectivity*
- $\lambda(i,j) = \frac{\sigma^m_{i,j}}{\sigma_{i,j}}$
- Large tube influence
- For $\beta = 0.8$, $\langle\lambda\rangle = 0.7$: 70% of journeys use the tube

Outreach

- *Spatial outreach* of a node i is the average Euclidean distance to all nodes reachable with a travel cost τ
- $L_\tau(i) = \frac{1}{N(\tau)} \sum_{j \in \{k | \tau_m(i,k) \leq \tau\}} d(i,j)$
- How "commutable" is a city



So far so good. . .

Analysis of real cities

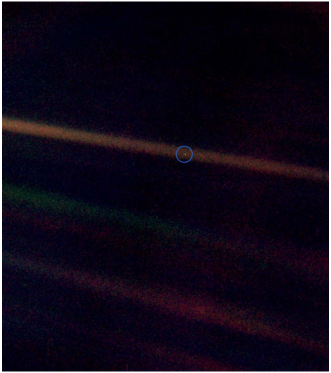
- Mathematically and computationally tractable
- Use publicly-available data (with a little work)

Implications for urban planning

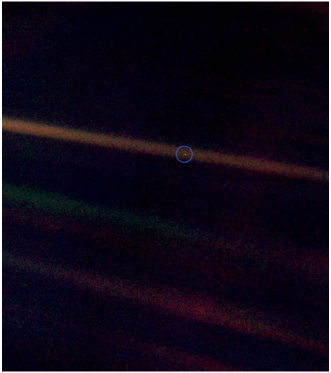
- Impact of planning decisions
- Provides a quantitative underpinning for ideas that urban planners already use informally

Reality intrudes

But there's a problem with reality. . .



But there's a problem with reality. . .



- There's only one of it

(Carl Sagan's Pale Blue Dot, taken on Valentine's Day 1976.)

Universality

Want to know that ideas work everywhere

- Our paper was originally titled *“Multiplex networks in metropolitan areas: universal features and local effects”*

Universality

Want to know that ideas work everywhere

- Our paper was originally titled *"Multiplex networks in metropolitan areas: universal features and local effects"*
- . . . and the referees asked, "how do you know?"

Universality

Want to know that ideas work everywhere

- Our paper was originally titled *“Multiplex networks in metropolitan areas: universal features and local effects”*
- . . . and the referees asked, “how do you know?”
- . . . which of course is a fair question: how *could* you know?

Universality

Want to know that ideas work everywhere

- Our paper was originally titled *"Multiplex networks in metropolitan areas: universal features and local effects"*
- ...and the referees asked, "how do you know?"
- ...which of course is a fair question: how *could* you know?
- ...so we changed the title of the paper ☺

Repetition and reality

The usual network science approach

- Characterise a network topology (*e.g.*, $\langle k \rangle$)
- Generate thousands of random instances
- Compute summary statistics over processes on each different instance

This doesn't work for the topologies of real cities (and other networks)

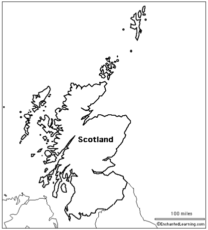
- A limited number of instances
- Topologies that *really are* complex
- Multiplexes with complex coupling constraints and probabilities

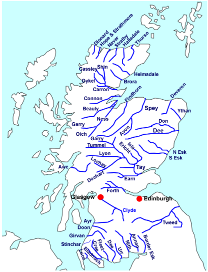
Faking it

A challenge

Can we synthesise realistic “really complex” networks?

- Complex dependencies between node degrees, spatial locations, and couplings between networks
- It’s know that degree couplings have a major impact on epeidemic processes, for example [5]





[5] S. Shai and S. Dobson. Effect of resource constraints on intersimilar coupled networks. *Physical Review E*, 86 (6), December 2012. URL doi://10.1103/PhysRevE.86.066120

A possible approach

Study the topologies independently

- Of each component network in the multiplex
- Degree distribution, degree correlations between adjacent nodes, . . .

Two ways to form the multiplex

1. Identify and couple corresponding nodes
2. Identify a growth process – the dynamics *of* the multiplex

Benefits

Aim for universaility

- Even in problems that are sensitive to the detailed mathematical structure
- Project forward *and* back
- Easier “what if?” analysis

Canonical example is the Black Death

- Why was it so deadly?
- Interactions between topology, survival times, disease dynamics
- Which factors were the most critical?

Conclusion

- Realistic to study urban-scale networks using network science
- Data publicly-available, but needs care and cleaning
- Multiplex networks of interest are too complex to synthesise currently
- ...but it would be nice to learn how to, and would enormously increase the phenomena we could study
- Including the effects of past and projected changes to actual locations

Thank you

Topology! The stratosphere of human thought! In the twenty-fourth century it might possibly be of use to someone. . .

Alexander Solzhenitsyn, The First Circle

References

A.-L. Barabási and R. Albert. Emergence of scaling in random networks. *Science*, 286(5439):509–512, 1999. URL `doi://10.1126/science.286.5439.509`.

M. Batty. *The new science of cities*. MIT Press, 2013. ISBN 978-0-292-01952-1.

M. Kivelä, A. Arenas, M. Barthélemy, J. Gleeson, Y. Moreno, and M. Porter. Multilayer networks. *Journal of Complex Networks*, 2(3):203–271, 2014. URL `doi://10.1093/comnet/cnu016`.

S. Shai and S. Dobson. Effect of resource constraints on intersimilar coupled networks. *Physical Review E*, 86(6), December 2012. URL `doi://10.1103/PhysRevE.86.066120`.

E. Strano, S. Shai, S. Dobson, and M. Barthélemy. Multiplex networks in metropolitan areas: generic features and local effects. *Journal of the Royal Society Interface*, 12(111), October 2015. URL `doi://10.1098/rsif.2015.0651`.