Which Network Model Should I Use?

TOWARDS a Quantitative
Comparison of Spatial
Network Models in
Archaeology

Tim Evans, Physics Dept &
Complexity and Networks Group

The Problem

- Archaeology can be “Site Centric”
 - Regional and global interactions hard to consider
- Networks emphasise interactions

Given a set of sites and their locations
can we understand their interactions by
creating a network of edges?

Island Archipelagos as an Ideal Network

- Vertices = Major Population or Resource Sites
- Edges = Exchange between sites
 - physical trade of goods or transmission of culture
 - direct contact or island hopping links
- *Sea isolates communities → Natural Vertices*
- *Interactions controlled by physical limitations of ancient sea travel → Simple Links*
- *Coastal Sites often isolated like islands due to geography and difficulty of ancient land travel*

Different Spaces

- We shall consider our sites to lie in two-dimensional geographical space
- Alternative is to consider sites located in some artefact space
 - Frequency counts of objects found at a site define a vector in a large-dimensional space

Edge Models

An edge model generates a set of edges between a given set of edges

Note a different type of model defines 'Spatial Influence' e.g. Theissen Polygons, XTent

Deducing Interactions

- Geography controls interactions in models discussed here
 - As the crow flies
 - Accounting for geography by hand estimation
 - Accounting for geography computationally GIS

[Terrell 1977; Irwin 1983; Hage & Harary 1991; Broodbank 2000; Collar 2007; Bevan 2010]
- Artefact counts [Terrell 2010; Sindbæk 2007]
- Texts [Isaksen 2006; “Anskar’s Vita” Sindbæk 2008]

Examples

- PPA – Principal Point Analysis
- MDN – Maximum Distance Network
- Gravity Models
 - Doubly Self-Consistent Models
 - Rhill and Wilson
- Stochastic Models
 - ariadne

PPA - Proximal Point Analysis

- Equal sized sites
- Sites connect to **k** nearest neighbours
- Analyse graph
 - Often without directions on edges
 - Sometimes only local measures used *e.g. Degree*
 - Sometimes global measures used *e.g. ranking, centrality, betweenness*

Examples: Hage & Harary 1991; Terrell 1977; Irwin 1983; Broodbank 2000; Collar 2007

DPPA Example (Directed PPA)

Connect each site to its $k=2$ nearest neighbours

PPA Example

Ignore direction

- All edges equal
- Network now simply connected

MDN – Maximum Distance Network

Just connect each site to all sites lying within distance ***D***.

- Used as model for ad-hoc wireless models [e.g. Srinivasa & Haenggi 2010]
- Mathematical analysis possible as ***Random Geometric Graphs*** [e.g. Penrose 2003]
- Not much used in archaeology

MDN Example

Sites **D** or less apart are connected

Doubly Self-Consistent Gravity Model

- Edge from i to j is flow F_{ij}
$$F_{ij} = a_i A_i b_j B_j V(d_{ij}/D)$$
- Inputs
 - arrival and departure rates A_i and B_j
 - Distance cost function V and distance scale D .
- Solve by demanding self-consistent arrival and departure rates to fix a_i and b_j
$$\Rightarrow \sum_j F_{ij} = A_i \text{ and } \sum_i F_{ij} = B_j$$
- Equivalent to optimising a cost function

Doubly Self-Consistent Gravity Model Example

Distance scale D as before

Rihll and Wilson Gravity Model

- Edge from i to j is flow F_{ij}
$$F_{ij} = b_i B_i (A_j)^\alpha V(d_{ij}/D)$$
where α is an additional model parameter
- Self consistent departure rate fixes b_i
$$\Rightarrow \sum_j F_{ij} = D_i$$
- Departure rate D_i is either:-
 - (a) a fixed input (size of site), or
 - (b) set equal to arrival rate A_j
- Find A_j and interpret as importance of site

Riell and Wilson Gravity Model Example

Same ***D*** as before,
closest two sites have most connections

***Strongly
connected core***

Stochastic Model – ariadne

[Evans, Knappett and Rivers 2008-2012]

- Has intrinsic volatility set by `temperature' parameter
- Allows sites to vary in size in response to network connections
- Network will give an low value of a `cost' function
 - includes costs for sites and edges sizes and ascribes benefits to interactions

ariadne Example

Bigger ***D***, 3 other parameters, one solution has the closest two sites large with most connections

Comparing Networks

- Same arrangement of sites gives different networks
- How can we compare them?

Comparing networks (I)

Measure a quantity associated with vertices

- Integer valued quantities poor *e.g. degree*
- Avoid quantities defined for simple networks *e.g. average shortest path*

⇒ Work with quantities defined on weighted networks

e.g. PageRank, clustering, betweenness'

Comparing networks (II)

Measure similarity of each pair of vectors

- Pearson correlation coefficient if gaussian
- Rank values then compare ranks
(largest value 1st, smallest last, then use Kendal's tau or Spearman) if have outliers
- Other less traditional schemes

Comparing Networks (III)

Vertex	Page Rank DPPA	Page Rank ariadne	Rank DPPA	Rank ariadne
1	0.35	0.35	1.5	1.5
2	0.35	0.35	1.5	1.5
3	0.25	0.10	3	4
4	0.05	0.20	4	3

0.78

Example Networks

e.g. PageRank of vertices
(use averages for ties)

Vertex	PPA	DPPA	MDN	DCGM	RWGM	MC
1	1.5	1.5	2	2.5	1.5	1.5
2	1.5	1.5	2	2.5	1.5	1.5
3	3	3	2	2.5	3.5	4
4	4	4	4	2.5	3.5	3

PageRank Correlation Matrix (Kendal method)

	PPA	DPPA	MDN	DCGM	RWGM	MC
PPA	1	1	0.82	NA	0.94	0.78
DPPA	1	1	0.82	NA	0.94	0.78
MDN	0.82	0.82	1	NA	0.58	0.27
DCGM	NA	NA	NA	1	NA	NA
RWGM	0.94	0.94	0.58	NA	1	0.94
MC	0.78	0.78	0.27	NA	0.94	1

Test Data

- Use real data sets
 - 39 Minoan Aegean sites [Knappett et al]
 - 110 Geometric Greek Sites [Rihll & Wilson]
- Use artificial datasets
 - Random sprinkling of points constant probability density

Test Data

More realistic:-

- 40 sites
- In 4 groups centred at points of compass
- 10 points per group scattered around centre

Problem

Every model has some parameters

How do we choose values for different networks when we want to make a comparison?

e.g. is a $k=4$ PPA to be compared to a $D=100\text{km}$ MDN network?

Answer

Look for models with same `physical` characteristic:-

- Average Distance
 - usual definition does not apply to weighted networks
- Time scale to visit all nodes
 - use random walkers
- Others...?

PPA and MDN: Distance vs Degree

Distance
=
Physical

Degree
=
Model

Physical Characteristic

Measures of distance (time scales etc) still transformed (renormalised) from model to model

? Use network topological characteristics ?
e.g. choose parameter such that there is one distance/time is roughly proportional to geographical separation

MDN for Distance 61.0 vs 62.0

PPA $k_{out} = 2$

PPA $k_{out} = 3$

PPA $k_{out} = 4$

MDN

– first global cycle

D=82km

TOWARDS a Quantitative Comparison of Spatial Network Models in Archaeology

Still work on going, no good recommendation as yet but basic ideas are

- Measure (several) vertex properties in each model
- Similarity of model defined through similarity of vertex properties
e.g. Pearson, Kendal tau, Spearman coefficients
- Need criteria to fix parameters

Acknowledgements

- All work done with
 - **Carl Knappett (Toronto)**
 - **Ray Rivers (Imperial)**some work also with
 - Edmund Hunt (Imperial)** and
 - Eric Beales (Toronto)**
- Publications
 - <http://theory.ic.ac.uk/~time>
 - or google “Tim Evans archaeology”

Bibliography

- Bevan, A., Political Geography and Palatial Crete, *Journal of Mediterranean Archaeology*, **2010**, 23 .
- RIHLL, T.E. & A.G. WILSON, 1987. Spatial interaction and structural models in historical analysis: some possibilities and an example, *Histoire & Mesure* 2: 5-32.
- RIHLL, T.E. & A.G. WILSON, 1991. Modelling settlement structures in ancient Greece: new approaches to the polis, in J. Rich & A. Wallace-Hadrill (eds.), *City and country in the ancient world*: 59-95. London: Routledge.
- EVANS, T.S., KNAPPETT, C., & R.J. RIVERS, 2011. Interactions in Space for Archaeological Models, *Advances in Complex Systems* (to appear).
- EVANS, T., KNAPPETT, C., & R. RIVERS, 2009. Using statistical physics to understand relational space: a case study from Mediterranean prehistory, in D. Lane, S. van der Leeuw, D. Pumain & G. West (eds.), *Complexity perspectives in innovation and social change*: 451-79. Berlin: Springer Methodos Series.
- KNAPPETT, C., EVANS, T. & R. RIVERS, 2008. Modelling maritime interaction in the Aegean Bronze Age, *Antiquity* 82: 1009-24.
- Knappett, C.; Evans, T. & Rivers, R., The Theran eruption and Minoan palatial collapse: new interpretations gained from modelling the maritime network, *Antiquity*, **2011**, 85, 1008-1023
- Rivers, R.; Knappett, C. & Evans, T., Network Models and Archaeological Spaces in *Computational Approaches to Archaeological Spaces*, A. Bevan & Lake, M. (Eds.), *Left Coast Press*, **2011**
- Collar, A., Network Theory and Religious Innovation, *Mediterranean Historical Review*, *Routledge*, **2007**, 22, 149-162