Understanding links between hotels and activities in a tourism destination using a growing network model

Juan M Hernández

Institute of Tourism and Sustainable Economic Development (TIDES)
Department of Quantitative Methods in Economics
University of Las Palmas de Gran Canaria

Introduction

- A social network is set a nodes (agents) which are linked through edges.

Karate club network, 1970

Women and events network, 1930

- Nowadays, large-sized networks are analyzed
- More sofisticated statistical analysis is employed

Political Blogs in US election, 2005

Introduction

- Tourism is a complex system and susceptible to be analyzed using complex network methodology.
- Previous contributions of complex network methodology in tourism research:
  - Limited in size ($N < 10^2$)
  - Applications:
    - ★ Analysis of topological characteristics [Grama and Baggio, *Ann. Tourism Res.*, 2014]
    - ★ Communities detection [Baggio and Sainaghi, *Tourism Manag.*, 2016]
    - ★ Movement patterns in destination [Smallwood et al., *Tourism Manag.*, 2012]
    - ★ Node attributes influencing relationships [Khalilzadeh, *Ann. Tourism Res.*, 2018]

Introduction. Empirical facts

Our project is inspired by the observations of the tourist activity developed in the southern area of the island of Gran Canaria, Spain...,

Maspalomas (Gran Canaria), 1965

Maspalomas (Gran Canaria), 2010

but also may be applied to other destinations!

Cancun, Mexico, 1970

Cancun, Mexico, 2010

Introduction

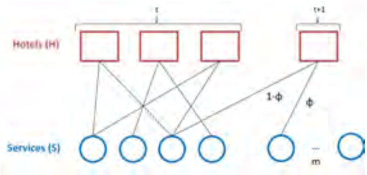
- Our **objective** is to build a growing network to represent the interrelationships between services in a tourist destination.
- A tourist who visit two different services makes the interrelationship.
- The model makes use of:
  - Evolving bipartite network models [Ramasco et al., *Phys. Rev. E* , 70, 03610666, 2004; Zhang et al., *Phys. A*, 392, pp. 6100–6106, 2013]
  - Studies on within destination movements of tourists [Lew and McKercher, *Ann. Tourism Res.*, 33, pp. 403–423, 2006; McKercher and Lau, *Tourism Geograph.*, 10, pp. 355–374, 2008]
    - ⋆ Tourist visit primary attractions even in they are located far from the hotel.
    - ⋆ Other attractions are substitutable. Distance influences.

Model 1. Assumptions

- Assume two categories of nodes, lodgings (H) and services (S).
- A link between a lodging $i \in H$ and service $j \in S$ appears if a representative tourist of lodging i visits service j during his/her staying in the destination.
- Assume that links are unweighted, undirected and permanent.
- Assume that every lodging includes exactly c links to services ($c \geq 1$).

Model 1. Growing rules

- At any time $t > t_0$, one new lodging and m new services are created in the destination.
- A representative tourist of a new hotel visit c services:
  - $\phi \in [0, 1]$ of them at random.
  - $1 - \phi$ by linear preferential attachment according to service's degree $s_j(t)$.



Representation of the growing rule of the supply network of a tourist destination

- After some calculations, the evolution of s_j can be described by

$$\frac{\partial s_j}{\partial t} = \frac{(1-\phi) m s_j + c\phi}{mt},$$

- Assuming t sufficiently large, we have

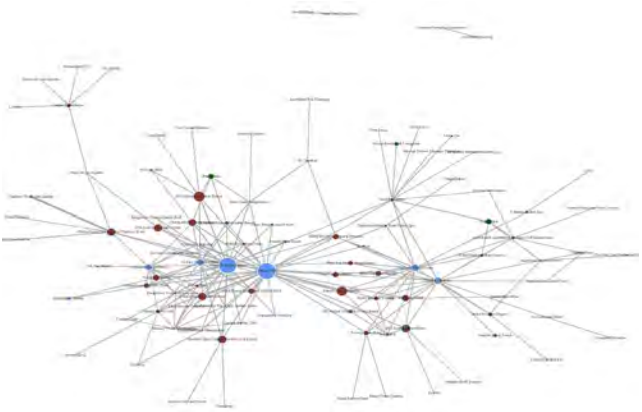
$$p(s) \simeq m(c\phi)^{\frac{1}{1-\phi}} \left(c\phi + m(1-\phi)s\right)^{-\frac{1}{1-\phi}-1}.$$

Study site and data

- The model parameters were adjusted to empirical data.
- The data are extracted from the tourist activity developed in Gran Canaria from 2005 to 2016.
- The data was collected from the user opinions published in the web-site trypadvisor.com.
- It includes some relevant biases:
  - Lodgings and services are limited to those published in the web-site.
  - The sample is also limited to those registered users.

Table 1. Basic statistics. $H \equiv$ Logings; $S \equiv$ Services; $L \equiv$ Links; $< c > \equiv$ Mean degree of lodgings; $< s > \equiv$ Mean degree of services; $\rho \equiv$ Density.

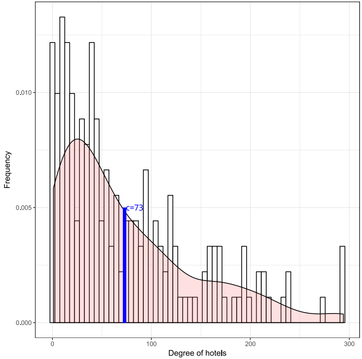
H	S	L	$< c >$	$< s >$	ρ
182	1,496	13,359	73.4	8.9	$9.5\ 10^{-3}$



Representation of the lodging-services network. Lodgings are colored in brown, services in blue and green. The link lodging-service indicates that at least 15 opinions of the service was made by tourists hosted in the lodging. Node's degree is represented by the ball size. Edges' thickness illustrate the number of opinions.

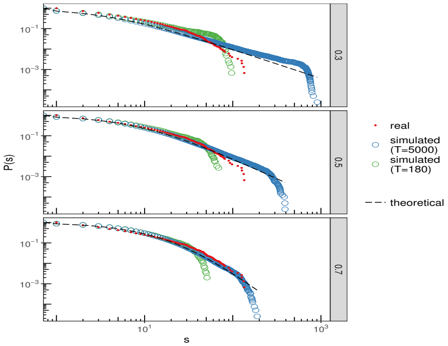
Model 1. Simulations and comparison with real data

- The degree distribution of hotels failed by assumption



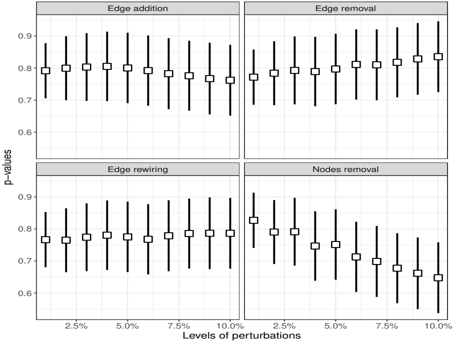
Lodgings degree distribution in the empirical sample of Maspalomas (Gran Canaria). The lodging's degree in the model is $c = 73$.

Model 1. Simulations and comparison with real data



Comparison among the cumulative degree distribution of services (s) for the empirical sample of recommendations in Maspalomas (red points), the theoretical predictions (dashed lines), the simulations of the model for a time horizon $T = 5000$ (blue circles) and for a time horizon $T = 180$ (green circles). Every graph assumes a specific value of the percentage of services chosen at random. The other parameters are determined according to the real data.

Model 1. Perturbation analysis



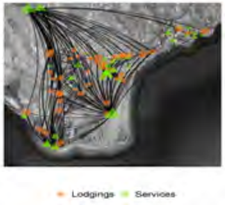
Results of the Kolmogorov-Smirnov test for the degree distribution of the perturbed empirical networks and the theoretical distribution with $\phi = 0.5$. Each graph shows the results for a specific type of perturbation. The x-axis shows the level of perturbation (1%,2%,...,10%), while the graph shows the mean p-value $\pm$ standard deviation for each level of perturbation.

Model 2. Motivation

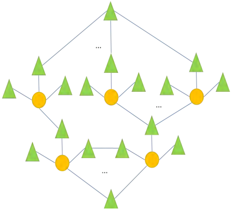
- The model 1 fits well real data in the long term for a specific combination of the random and preferential attachment rules, but not for the real time horizon.
- Limitations:
  - Distance is not included in the decision to enjoy certain services.
  - Rewiring of old links is not allowed in the model.
  - New links to old hotel are not allowed either.
  - Removal of lodging/services due to aging or closure is not included.

Model 2. Assumptions

- Assume two categories of nodes, lodgings (H) and services (S).
- A link between a lodging $i \in H$ and service $j \in S$ appears if a representative tourist of lodging i visits service j during his/her staying in the destination.
- Nodes (lodgings and services) are located in a topological space which is represented by a planar network where distance is defined by the length of geodesic paths.



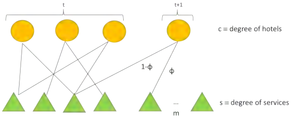
Representation of the lodging-services network in Maspalomas, Spain. Node's degree is represented by the ball size.



Spatial location of hotels (green triangles) and services (gold circles)

Model 2. Growing rules

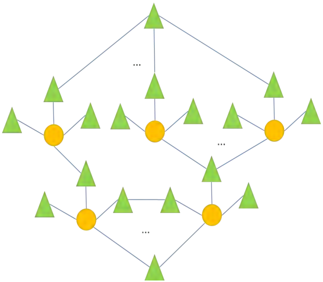
- At any time $t > t_0$, one new lodging and m new services are created in the destination.
- A representative tourist of a new hotel i visit c services, $\phi \in [0, 1]$ of them at random and $1 - \phi$ by preferential attachment, following the rule $\Pi_{i \to j} \sim s_j d_{ij}^{-\alpha}$, where $\alpha > 0$, s_j is the service j's degree and d_{ij} is the length of a geodesic path form i to j.



Representation of the growing rule of the supply network of a tourist destination

Model 2. Simulations and comparison with real data

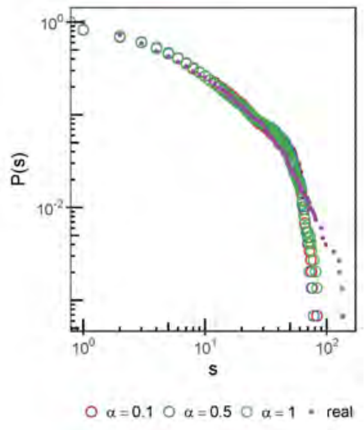
- Initially, we assume the existence of two extreme services in the network.
- The initial lodgings are linked to the extreme services.
- The hotels and the closest new services appear at every time step.



Spatial representation of hotels and services.

Model 2. Simulations and comparison with real data

- The degree distribution of services agree in the long term, but not for empirical size.



Comparison among the services degree distribution (s): Empirical (red points), simulated (circles).

Conclusions

- We have built an evolving bipartite network model that represents the supply network formed by the tourist visits from hotels to services, having into account the spatial restrictions of service location.
- The model fits well the real data in the long term for a specific selection of parameters, but not in the real time.
- Limitations and extensions:
  - Other economic characteristics from the supply-side (promotion, quality of the service) or demand-side (visitor's preferences) are not included.
  - Aging and links rewiring are not included either.

Thank you!