Journal of Archaeological Science 72 (2016) 105–116

Contents lists available at ScienceDirect

Journal of Archaeological Science

journal homepage: http://www.elsevier.com/locate/jas

On graphical representations of similarity in geo-temporal frequency data

Daniel Weidele [a, *], Mereke van Garderen [a], Mark Golitko [b, c], Gary M. Feinman [c], Ulrik Brandes [a]

[a] *Department of Computer & Information Science, University of Konstanz, Germany*
[b] *Department of Anthropology, University of Notre Dame, United States*
[c] *Department of Anthropology, Field Museum of Natural History, United States*

ARTICLE INFO

Article history:
Received 27 November 2015
Received in revised form
27 May 2016
Accepted 30 May 2016
Available online 22 June 2016

Keywords:
Network visualization
Maya
Obsidian
MDS
Spatio-temporal data

ABSTRACT

Its focus on dependencies and patterns in relational data makes network science a promising addition to the analytic toolbox in archaeology. Despite its tradition in a number of other fields, however, the methodology of network science is only in development and its scope and proper usage are subject to debate. We argue that the historical linkage with graph theory and limitations in commonly available software form an obstacle to leveraging the full potential of network methods. This is illustrated via replication of a study of Maya obsidian (Golitko et al. Antiquity, 2012), in which it seemed necessary to discard detailed information in order to represent data in networks suitable for further processing. We propose means to avoid such information loss by using methods capable of handling valued rather than binarized data. The resulting representations corroborate previous conclusions but are more reliable and thus justify a more detailed interpretation of shifting supply routes as an underlying process contributing to the collapse of Maya urban centers. Some general conclusions for the use of network science in archaeology are offered.

© 2016 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

1. Introduction

1.1. Theoretical background

Network Science is the study of the collection, management, analysis, interpretation, and presentation of relational data (Brandes et al., 2013). It combines statistical, combinatorial, algorithmic, and graphical methods to address research questions amenable to a network perspective. As for any science, a precise understanding of the potentials and interrelations as well as limitations of network science methods is vital in order to apply them appropriately and obtain meaningful results.

Network approaches are becoming increasingly commonplace. A range of examples demonstrate that also in archaeology new insight can be obtained. A network perspective was used to analyze the use of raw materials and knapping techniques in the pre-colonial Caribbean (Mol, 2014), to understand the collapse of inland Maya urban centers (Golitko et al., 2012; Golitko and Feinman, 2015), to study the transformation of social networks in the late pre-Hispanic US Southwest (Mills et al., 2013, 2015), to explore the co-occurrence and trade routes of Roman table wares (Brughmans, 2010; Brughmans and Poblome, 2012), to study information diffusion through Roman space (Graham, 2006), to model maritime interaction in the Aegean Bronze Age (Knappett et al., 2008), and to identify social and cultural boundaries in Papua New Guinea (Terrell, 2010), to name but a few examples.

However, the methodology of network science is only in development and proper usage standards are the subject of debate. Brughmans (2013) identifies two critical issues regarding the current status in this domain: (1) a lack of awareness and understanding of the broad range of formal network methods within the archaeological discipline has led to a limited methodological scope; (2) the application of network methods in archaeology has been driven mostly by possibility, rather than by specific archaeological research questions. As a result of these two issues, network science applications in archaeology have been dominated by a few popular

---

* Corresponding author.

E-mail addresses: daniel.weidele@uni-konstanz.de (D. Weidele), mereke.van.garderen@uni-konstanz.de (M. van Garderen), mgolitko@nd.edu (M. Golitko), gfeinman@fieldmuseum.org (G.M. Feinman), ulrik.brandes@uni-konstanz.de (U. Brandes).

http://dx.doi.org/10.1016/j.jas.2016.05.013
0305-4403/© 2016 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

methods.

One such popular method is binarization, replacing valued data with zeroes and ones. This converts a weighted network, in which each pair of nodes is connected with a link of some value, into a binary network, in which links can only be present (1) or absent (0). This technique, though very useful in principle, should be applied only with care and double-checking of conclusions, as was illustrated by Peeples and Roberts Jr. (2013) using a number of case studies. Due to the strong link of network science with graph theory, networks are often represented as binary and methods designed to handle valued data are less commonly used in current network science applications. Since, however, binarization incurs information loss, it should be avoided where possible.

1.2. Our contribution

We consider a chain of operations which obtained a prominent place among network methods used in archaeology. In this approach (see for example Mills et al., 2013; Golitko et al., 2012; Golitko and Feinman, 2015), a network is built from similarities between site assemblages. The network is then binarized using some threshold value. Unless sites are shown at the geographic locations, a layout of the graph is computed, typically using a spring-embedder algorithm. While this often serves to visually communicate results, exploration of the network diagram can also lead to new conclusions for the authors themselves.

In this paper, we consider the steps during this process at which information loss occurs. We demonstrate that binarization, which may sometimes appear necessary to be able to apply the intended methods, can actually be avoided. To do so we suggest methods able to handle valued data at each step of the analytic pipeline. We also note that common spring-embedder algorithms do not result in layouts that can be interpreted reliably. With the nature of archaeological research questions in mind, we introduce a variant method for visualizing and analyzing geo-temporal frequency data that gives a more accurate representation of the raw data. We illustrate that this new method can lead to slightly different results by reanalyzing data of Golitko et al. (2012) on Maya obsidian. We stress that this case study replication is only an example to illustrate the techniques we introduce. Due to the omnipresence of geospatial frequency data in the archaeological discipline, the method is in fact widely applicable.

The present contribution should not be understood as a competing analysis of particular archaeological hypotheses. Instead, our contribution is methodical: we point out a strategy to obtain more reliable visual representations and use the archaeological case study on Maya obsidian as a concrete example.

1.3. Data and case study

We identify a class of data that regularly constitutes the basis for archaeological studies. We refer to this class as geo-temporal frequencies, which can be defined given.

- a set of geographic locations L,
- a set of discrete time points T,
- a set of classes of artifacts C

as a three-dimensional tensor $X \in \mathbb{N}^{L \times T \times C}$, so $X_{l,t,c}$ represents the number of, for instance, pottery sherds of ware $c \in C$ found at site $l \in L$ dated to time $t \in T$.

As a case study we consider the work of Golitko et al. (2012) on Maya trade relations in eastern Mesoamerica between 250 CE and 1520 CE. In this study, network methods are applied to archaeological data on material culture, which in turn is used as a proxy for trade. We evaluate the methods used and suggest a number of improvements and extensions. We replicate the case study together with an application of the suggested method which leads to a more precise visualization of the data that allows some new observations.

The data set consists of obsidian assemblages from 121 archaeological sites. Obsidian is considered an ideal material to use for the reconstruction of trade relations since the original source of an obsidian artifact can be chemically determined with high confidence. The three main sources of obsidian in the eastern Mesoamerican Maya area are San Martin Jilotepeque (SMJ), El Chayal (ELC), and Ixtepeque (IXT), all currently located in Guatemala. For ease of viewing and analysis, all Mexican obsidian sources have been compiled into one category (MEX), and all non-major sources in Honduras and Guatemala have been grouped into one category (OTHER).

Fig. 1 shows a map of the study area on which the sites and sources are indicated. The node area corresponds linearly to the absolute number of sourced obsidian objects found at this site, which makes clear how large the differences really are. For ease of viewing, we will use a logarithmic scaling in the remainder of this paper, which makes the differences in node sizes a lot smaller as compared to this figure. Sites are colored according to their geographical zone after Adams and Culbert (1977). We will use the same encoding throughout this paper.

The assemblages have been dated to four time intervals: the Classic period (~250 CE/300–800), the Terminal Classic period (~800–1050 CE), the Early Postclassic period (~1050–1300 CE), and the Late Postclassic period (~1300–1520 CE). Fig. 2 shows the geographical distribution of obsidian from the different sources throughout the four periods as small multiples: a matrix with a column for each period and a row for each obisidian source. Sites [sources] are represented by dots [triangles] in their geographical locations. The node sizes correspond to the logarithmically scaled absolute number of sourced obsidian objects found at this site for a given source and period. The color intensities represent the proportion of the obsidian found at this site for this period that came from this source. A small, black node in the Classic-ELC cell means that for this site, (almost) all of the material found for the Classic period came from source ELC, but that there were not many pieces in total. A large, medium grey node in the Terminal Classic-IXT cell means that for this site only about half of the objects found for the Terminal Classic period came from source IXT, but that this was still quite a large number of objects.

1.4. Preliminaries

In the following we describe how to build a network out of the data described above. Following Brandes et al. (2013) we represent a network variable from geo-temporal frequency data as a mapping $x : \mathscr{D} \rightarrow \mathscr{W}$ of dyads from a finite domain $\mathscr{D} \subseteq \mathscr{N} \times \mathscr{A}$ comprised of ordered pairs of nodes $\mathscr{N}$ and affiliations $\mathscr{A}$ to values in a range $\mathscr{W}$.

Of the possible combinations with $\mathscr{N}, \mathscr{A} \in \{L, T, C\}$ we focus on site-site interaction domains $\mathscr{D}^{LL}$ where $\mathscr{N} = \mathscr{A} = L$. These provide a natural way of directly preserving the geographical context, and are presumably therefore frequently subject of study in archaeological research. We consequently define the network mapping x^{LL} on the interaction domain $L \times L$ as

$$x^{LL} : L \times L \rightarrow \mathscr{W}. \tag{1}$$

This means that we look at all possible combinations of two sites (the nodes in our network), and assign a weight to the link between each of these pairs.

Like Golitko et al. (2012), we rely on the assumption that the

Download English Version:

https://daneshyari.com/en/article/7441275

Download Persian Version:

https://daneshyari.com/article/7441275

Daneshyari.com