

A case study on the usage of social network analysis to complement participant observation on Twitter

Un estudio de caso sobre el uso de análisis de red social como complemento a observación participante en Twitter

Um estudo de caso sobre o uso de análise de rede social como complemento a observação participante no Twitter

Recebido em 27-08-2021
Modificado em 29-10-2021
Aceito para publicação 25-11-2021

 <https://doi.org/10.47456/simbitica.v8i4.37347>

 **Vlad Schüler-Costa**

Doutor em Social Anthropology pela University of Manchester, UK, com bolsa de doutorado pleno no exterior pela CAPES (Ministério da Educação, Brasil). Mestre em Antropologia Social pelo Museu Nacional/UFRJ, com bolsa pelo CNPq. E-mail: vlad.schulercosta@manchester.ac.uk

Abstract

Based on digital ethnography conducted between 2013 and 2015 on the social network service Twitter, this article develops a theoretical-methodological discussion on the usage of data mining and network visualisation as a “quantitative” complement to the “qualitative” approach of participant observation. Starting from a literature review, this article argues that ethnography, from its beginnings, has been open to “quantitative” research. This article also tackles the anthropological origins of terminology and methodology regarding social network analysis, thus recovering the anthropological interest on the subject. Finally, based on an empirical case study, this article shows a possible usage of such methods as a way to enrich ethnographic analysis.

Keywords: social anthropology; digital ethnography; network analysis; methodology.

Introduction

Through the last decade or so, we have seen an increasing interest in digital and computational methods within the humanities and social sciences. Indeed, burgeoning fields such as digital humanities, social data science, and computational social sciences show promising possibilities for the analysis of social and human data through digital methods. This article aims to contribute to these fields and conversations. Focusing on the disciplinary boundaries of social anthropology – and its main method, ethnography –, this article argues, based on a case study, that digital methods such as data mining and social network analysis can be fruitfully used as complimentary to “traditional” participant observation.

To that end, the article is composed of four main sections: I begin by covering the historical precedents of the usage of “quantitative” methods in social anthropology (which is nowadays thought of as a mainly qualitative discipline); I then show how social anthropologists – particularly those connected with J. Clyde Mitchell and the “Manchester School” – pioneered, from the 1950s through the 1980s, “social network” as a concept, as a method, and as a field; I then move on to the case study in question, first describing the community which preoccupied me and whose patterns of social interaction made me seek computational methods for its study; and finally I show how the usage of these methods helped me achieve a better understanding of this community than I would otherwise get with “just” qualitative methods.

Finally, in the conclusion, I reiterate my argument that neither approach – either “qualitative” or “quantitative” – is inherently superior, and that it is the reality of the field which will show which approach (or both) better suits the research project.

Ethnography and quantitative methods: a brief historical overview

Bronisław Malinowski is widely credited as the inventor of modern ethnography, anthropological fieldwork, and as coiner of the term “participant observation” (Vermeulen, 2008:149). He considered his new method to be innovative in that, unlike previous anthropologists, he was not content in merely writing about the “natives” from afar.¹ On the

¹ This is, however, contentious. It could be argued that budding forms of modern ethnographic fieldwork were being conducted – by historians, geographers, and folklorists – at the the early eighteenth century (Vermeulen, 2008). Pre-modern ethnography, it has been argued, dates as far back as the ancient Roman imperial conquer of Europe (Woolf, 2010).

contrary: it was the ethnographer's duty to be as completely immersed in local life as possible. As Adam Kuper writes:

[Malinowski] decided that the ethnographer 'must relinquish his comfortable position on the veranda', pitch a tent in the village, cultivate a garden, exchange gifts, listen in to conversations, flirt, argue, and generally hang about. Intimate personal histories, neighbourhood feuds, the tug of war between a person's emotional loyalties and his legal obligations, all this was accessible only to an observer who was immersed in the everyday life of the village. *The new ethnographer was a participant observer* (Kuper, 2014:18 [emphasis added]).

Even then, however, Malinowski did not completely disown the methods developed by his antecessors. It was not as if "the new ethnographer" had to abandon the questionnaire-based survey method, developed over the previous five decades by the Royal Anthropological Institute and published in its then famous handbook *Notes and queries in anthropology*.² What Malinowski argued for was a method which would add to, rather than make obsolete, earlier forms of anthropological data-gathering. In an oft-cited passage, he writes that "[i]n certain results of scientific work – especially that which has been called 'survey work' – we are given an excellent skeleton, so to speak, of the tribal constitution, but it lacks flesh and blood." (Malinowski, 2014[1922]:52-53). His proposed method – deep engagement through participant observation – would provide this much needed "flesh and blood", but still would work upon the "skeleton" provided by previous methods: "add something essential to the bare outline of tribal constitution, and to supplement it by all the details" (Malinowski, 2014[1922]:54).

This separation, and complementarity, between the "skeleton" provided by "survey work" and the "flesh and blood" provided by "participant observation" would remain central to much of anthropological research, particularly in Britain, throughout the twentieth century. Indeed, it was commonplace to start an anthropological account by what was sometimes called "tribal composition", "demography", or "social structure" – examples can be seen, among others, in the 1940 edited book *African Political Systems* (Fortes & Evans-Pritchard, 2016[1940]). Likewise, in a series of lectures given at Chicago in 1937, Radcliffe-Brown, proposing a "natural science of society", argues: "[t]he kind of mathematics which will be required ultimately for a full development of the science of society will not be metrical, but will be the hitherto comparatively neglected branch of mathematics, the calculus of relations" (Radcliffe-Brown, 1964[1937]:69).

² See Urry (1972) for a historical overview of the development and usage of the handbook between the years 1870 and 1920.

It would be, however, in the so-called “Manchester School”, spearheaded by Max Gluckman (one of Radcliffe-Brown’s disciples) at the department of Social Anthropology at Manchester University – and its associated Rhodes-Livingstone Institute (RLI) in then Northern Rhodesia, current Zambia –, that ideas of deeply integrating mathematical analysis with ethnographic description would take hold. Indeed, as Werbner tells us:

Such quantitative study accompanying the qualitative was a methodology Gluckman much admired, and he sought to have a demographer participate in the RLI research; his own genealogies and censuses for Barotseland were destroyed in a fire so that he never analysed them in relation to his other evidence on kinship and family relations. At his insistence, all of his students working in Africa took an obligatory highly detailed census in the field, often on punch cards after the model of the RLI census card which [J. Clyde] Mitchell designed (Werbner, 2020:103).

It would be hard to overstate Mitchell’s importance in establishing this interest in merging quantitative and qualitative analysis in social anthropology. His 1956 book *The Yao Village* (Mitchell, 1956), based on research conducted between 1946 and 1949, thoroughly mixes ethnographic description and statistical analysis, often in support of one another. Indeed, he was hired by Gluckman, in 1945, due to having “somehow hypnotized Gluckman with the magic of Chi-square values” (Werbner, 2020:100). His most well-known work, *The Kalela Dance* (Mitchell, 1959), likewise elegantly uses survey-based statistics to illustrate his ethnographic findings.

Indeed, in his chapter *On Quantification in Social Anthropology* (part of the 1967 edited book *The Craft of Social Anthropology*), Mitchell presents an overview of the usage of quantitative data in anthropology, but likewise argues that “quantification, however, while it is an important method of data collection, should be an aid to, and not the purpose of, fieldwork” (Mitchell, 1967:21). However, he also sings high praise on the potential of quantitative analysis – particularly the usage of statistical methods:

Statistical analysis is a powerful tool in the hands of the modern social scientist, and social anthropologists cannot afford to forgo the aid they can derive from using it. It should take – and is taking – its rightful place as one of the many techniques social anthropologists must use in their quest to document and understand man’s infinite social variety (Mitchell, 1967:45).

Mitchell’s interest in such methodology can be best seen in his 1980 edited book *Numerical Techniques in Social Anthropology* (Mitchell, 1980), itself based on a conference session held in 1973. In his introduction to the book, he maps the crescent interest in, and methodological development of, the usage of mathematical procedures in social anthropology. Here, unlike in *The Yao Village* or *The Kalela Dance*, the presence of mathematics is not (necessarily) part of the data-collection stage, but rather of the analysis of ethnographic data:

The contributions to this book merely illustrate how certain characteristics inherent in ethnographic data with which the authors were familiar may be thrown into relief through the usage of appropriate numerical procedures. [...] What is new is the way in which the various procedures have been applied to empirical anthropological problems. The contributions are directed, therefore, not so much to the specialist in mathematical or quantitative anthropology as to the general anthropologist who wishes to learn about possible ways of handling field work material (Mitchell, 1980:3-4).

Mitchell's concern, thus, was with making mathematical anthropology accessible to "non-mathematical" anthropologists, although one wonders, forty years later, how successful that was. Regardless, it is here that he suggests that, although the majority of highly complex mathematical analysis might be beyond most anthropologists' concerns, some of them are "relatively simple procedure[s] when a digital computer is available. Many standard data analysis packages, such as the *Statistical Package for the Social Sciences* (SPSS),³ have facilities for doing this" (Mitchell, 1980:5). This sentence anticipates much of what was to come regarding the presence of quantitative methods in social anthropology: the widespread usage of computing.

A full review of the usage of computing in social anthropology would be beyond the scope of this article,⁴ but suffice to say that through the following few decades anthropologists would increasingly merge the applications of quantitative methods and the usage of computing software – long would be gone the days in which the likes of Mitchell would do their Chi-square calculations in slide rules (Werbner, 2020:102).

Indeed, in his 1994 book, *Applications in Computing for Social Anthropologists*, Michael D. Fischer writes that "computers have the potential to make numerical methods more acceptable to anthropologists who currently avoid them and vastly to improve the research of those who do not" (Fischer, 1994:3), even recommending the short-lived (1989-1996) journal *Quantitative Anthropology* as a resource for quantitative computing in anthropology. Even earlier, in his introduction to the 1965 edited book *The Use of Computers in Anthropology*, Dell Hymes wrote "use of the computer in American anthropology is likely to increase as the next generation, having advantage of new approaches in the teaching of mathematical thinking, reaches professional careers" (Hymes, 1965:24). If computing literacy was not needed for the anthropological research of mathematics (Crump, 1992),⁵ mathematical literacy was, it seems, needed for the burgeoning field of computational

³ It is noteworthy that this was written at an age before personal computers, and SPSS 1.0 was meant for usage in mainframes. It was only in the mid-1980s that SPSS would first appear on, then, MS-DOS.

⁴ The best such review to date is, to my knowledge, Seaver (2014).

⁵ Or, likewise, for the *ethnomethodological* research of mathematics (Livingston, 1986).

anthropology. However – perhaps because of that requirement – the usage of computational methods in social anthropology remained (and arguably remains) marginal in the discipline: as Fischer himself promptly admits, most anthropologists would find computers most useful in their writing, using word processors instead of typewriters or, indeed, of handwriting (Fischer, 1994:16).

It would only be in the 2000s, with the emergence of digital anthropology (Horst & Miller, 2012; Sanjek & Tratner, 2015), that anthropological interest in computers would re-emerge – this time, however, not as quantitative method,⁶ but as qualitative fieldsite.

Social anthropology and the development of social network analysis

A similar pattern, of waxing and waning interest, can be seen on anthropological approaches to social network analysis – although its history is somewhat shorter.

In his presidential address to the Royal Anthropological Institute in 1940 (later republished as a chapter in *Structure and Function in Primitive Society*), Radcliffe-Brown develops his idea of a social science based on the “calculus of relations” by introducing into social science the concept of a (social) “network”:

[...] direct observation does reveal to us that these human beings are connected by a complex network of social relations. I use the term ‘social structure’ to denote this network of actually existing relations [...] A particular social relation between two persons (unless they be Adam and Eve in the Garden of Eden) exists only as part of a wide network of social relations, involving many other persons, and it is this network which I regard as the object of our investigations (Radcliffe-Brown, 1940b:2-3).

However, it was only after a bit more than a decade that the concept of “social network” caught on, particularly amongst scholars associated with (again) Max Gluckman at the department of Social Anthropology at Manchester University.⁷ Struggling with the burgeoning anthropological study of “complex” societies, Manchester School anthropologists found in this concept a helpful tool in analysing communities which were not “self-contained”, as was supposedly the case previously.

⁶ That being said, interest in quantitative methods in social anthropology never really disappeared, as can be attested by H. Russel Bernard’s continuously updated masterpiece (and classroom staple) *Research Methods in Anthropology: Qualitative and Quantitative Approaches* - initially published in 1988 but, as of writing, on its sixth edition (Bernard, 2017).

⁷ A historical overview of social network analysis as a wider discipline, within and without anthropology, can be seen at Scott (2000:7-37).

John A. Barnes, the first to engage with the concept in writing, tells us that in his study of an island parish in Western Norway, he had to account for the fact that the islanders – numbering 4,600 – did not know every other islander, but were together part of the same community (and, indeed, part of Norway). The image he paints to illustrate his analysis of such a community might seem commonsensical to today’s anthropologists, but this was not the case in the 1950s:

Each person is, at it were, in touch with a number of other people, some of whom are directly in touch with each other and some of whom are not. Similarly each person has a number of friends, and these friends have their own friends; some of any person’s friends know each other, some do not. I find it convenient to talk of a social field of this kind as a *network*. The image I have is of a set of points some of which are joined by lines. The points of the image are people, or sometimes groups, and the lines indicate which people interact with each other. We can of course think of the whole of social life as generating a network of this kind (Barnes, 1954:43, italics in original).

I want to reinforce here the fact that while this description might be exactly what nowadays we would see as a social network of any kind, with its nodes and links, it was nothing short of revolutionary when it was first introduced into the social sciences. Indeed, as Barnes shows us, it was a theory borne out of necessity: as anthropologists, the Manchester School researchers were first and foremost interested in face-to-face relationships and interactions, but it was not possible to paint a coherent and cohesive picture of such a community in the same way that, say, Malinowski had done with the Trobriand Islanders or Evans-Pritchard with the Nuer or the Azande.⁸ Indeed, another pioneer in anthropological network analysis – again associated (although more loosely) with the Manchester School –, Elizabeth Bott makes an explicit contrast between networks and what she calls “organised groups” in her own study of conjugal roles in London families:

The external social relationships of all families appeared to assume the form of a network rather than of an organized group. In an organized group, the component individuals make up a larger social whole with common aims, interdependent roles and a distinctive sub-culture. In network formation some but not all of the component individuals have social relationships with one another. They do not form an organized group and the component external units do not make up a larger social whole; they are not surrounded by a common boundary (Bott 1990[1955]:324).

Maybe nowhere else did this novel idea of “social network” take hold as in the Rhodes-Livingstone Institute. Faced with the rapid change in social structure brought upon by

⁸ I don’t mean, of course, that the Trobrianders, the Nuer, or the Azande *were* self-contained, coherent and cohesive communities – but that they were then *theorised* as being so.

urbanisation and industrialisation in Africa, Manchester School researchers associated with it found in it an explanatory power that they had struggled with beforehand. Writing in 1961, Alfred L. Epstein paints a vivid picture of Ndola, a rapidly developing town in the African Copperbelt: arising almost from nothing due to the mining industry, Ndola had a population of over 50,000 African inhabitants (thus, not counting European colonial settlers), most of whom were then-recent migrants from elsewhere in Africa. The town exhibited a then-unusual diverse environment, in which many ethnic groups, tribes, languages, dresses, and customs mixed seemingly at random – indeed, without social organisation into closer-knit neighbourhoods or ghettos traditionally thought of (by social scientists at the time) as the usual pattern of inhabitation and settlement of migrants into new spaces.⁹ Epstein writes in surprise:

There is a constant coming and going of people. Individuals move from one part of the town to another. [...] Yet despite the apparent confusion of the urban scene, it is equally patent that the Africans who live in Ndola do not compose a mere aggregation of individuals not a disorganized rabble. [...] The fact is that each individual African is involved in a network of social ties which ramify throughout the urban community and extend to other towns and to the tribal areas (Epstein 1969[1961]:79-80).

These anthropologists, faced with the change of their research sites from “small-scale” societies to “complex” ones – but still drawing from social anthropology’s traditional methodology of participant observation, face-to-face engagements, and an interest on what Malinowski called “the imponderabilia of actual life” (Malinowski, 2014[1922]:54) – started to conceptualise life in society – particularly in complex societies – as a social network.

It is important to say, however, that none of these authors – Barnes, Bott, Epstein – used “social network” as anything but a *concept*. That is to say that none of them presented the networks they analysed in graphical form – the closest we see is a kinship graph¹⁰ in Epstein’s paper (Epstein 1969[1961]:97). It would only be in the mid-1960s that we would start seeing the graphical representation of networks being used in published anthropological research. Of particular importance here is the 1969 book *Social Networks in Urban Situations*, edited by Clyde Mitchell (Mitchell, 1969). Based on a series of papers presented by the different contributors between 1965 and 1966 (but also containing some previously

⁹ The foremost study of this inhabitation pattern being Thomas & Znaniecki (1996 [1918]).

¹⁰ It is noteworthy that Barnes himself draws an approximation between his usage of the term “network” to Meyes Fortes’ idea of “web of kinship”. A full analysis of the influence of kinship studies (and genealogy graphs) upon network studies is, however, beyond the scope of this article.

published work), the book contains twenty network graphs,¹¹ none of which containing more than twenty nodes.

Still, this was groundbreaking for its time, particularly the book's introductory chapter, in which Mitchell tries to delineate a theory and methodology for social network analysis, borrowing heavily from mathematics (particularly set theory and graph theory) to define key terms and concepts which, as he tells his reader, were used haphazardly by previous authors. However, as Scott tells us, the book's groundbreaking aspect made it, perhaps, victim of its own success:

The arguments of Mitchell, Barnes and Bott were extremely influential in Britain, but their very success meant that social network analysis came to be identified with the specific ideas of the Manchester anthropologists. That is to say, network analysis was seen to be concerned specifically with informal, interpersonal relations of a 'communal' type, and the method was seen as specifically concerned with the investigation of ego-centric networks. As a result, the crucial breakthrough to the study of the global properties of social networks in all fields of social life was not made in Britain (Scott, 2000:33).

Thus, despite its success and influence, the "Manchester-style" social network analysis steadily started to lose steam, within and without anthropology. Attempts by anthropologists to broaden its analytical power did, however, happen. The 1973 book *Network Analysis: Studies in Human Interaction*, edited by Jeremy Boissevain and Clyde Mitchell (Boissevain e Mitchell, 1973) holds examples of this. In his contribution, Mitchell writes that "we will probably find it necessary to make more use of the notion of social network to explain social behaviour in large-scale societies" (Mitchell, 1973:34), and Boissevain shows that the biggest difficulty in network analysis was, then, collecting and analysing data:

[...] the researcher is faced with the problem of gathering the relevant data and processing it. This in my experience is prohibitively time consuming and likely to inundate the researcher with data to the point of immobility. I would suggest therefore that perhaps the most urgent problem confronting those who wish to make use of the network approach is to develop a method of gathering pertinent data rapidly (Boissevain, 1973:147).

Be that as it may, by the 1970s anthropologists were already criticising network analysis for, in fact, being too enthusiastic about networks for their own sake: Bruce Kapferer (1973) scathingly argues that network analysis was a method should be coupled with theory, instead of being seen as a theory of its own; similarly, Roger Sanjek argues that "One does not study networks; one uses network methods to answer anthropological questions" (Sanjek, 1974:589) such as "class and inter-class relations, about inter-ethnic relations, or about

¹¹ Half of which are merely hypothetical graphs used for illustrative purposes.

economic relations” (Sanjek, 1974:596); Boissevain himself would later write that “Far too much of the research now being done on networks lacks any clear formulation of the problems it seeks to resolve” and that “Network analysis is a research instrument which can help resolve certain social and theoretical problems. It must not become an esoteric end in itself” (Boissevain, 1979:393-394). Furthermore, network analysis was showing signs of facing the issue already hinted at by Mitchell in 1967: “Quantification has no magical property to confer accuracy on the data: if the basic observations are inaccurate or incomplete, statistics derived from them will assuredly also reflect these weaknesses” (Mitchell, 1967:26). Thus, without much fanfare, social anthropologists slowly but steadily fell out of love with network analysis, and other social scientists (such as sociologists, political scientists, economists, geographers, and psychologists) took the field further (Scott, 2000:8-27, 33-37).

Aside from relatively few exceptions (such as Barnes & Harary, 1983; Hage & Harary, 1983), it would be many decades before anthropologists would retake their interest in social networks: Clyde Mitchell himself, considered by some to be the founder of the discipline, claimed in 1990 that social network analysis was dead within anthropology (Werbner, 2020:143). He, thus, turned away from anthropology and towards this new discipline – involving himself heavily with the *International Network for Social Network Analysis*, the journal *Social Networks*, and the networks group at the British Sociological Association (Freeman, 2004; Werbner, 2020). As Werbner tells us: “At the end of Mitchell’s career [...] it was not anthropologists [...] but the young geographers, graduate students he supervised, who ‘kept me intellectually alive’, as he remarked” (Werbner, 2020:128).

Like with the computational anthropology from the previous section, it would be in the twenty-first century, with the emergence of digital anthropology, that social anthropologists would turn their interest towards social networks yet again – and, likewise, as qualitative fieldwork instead of quantitative method¹².

Digital ethnography on Twitter: a case study

And it is with digital anthropology that I start this section. More specifically, I will develop on my own research, in which I followed a community that used Twitter (among other platforms, such as Facebook, YouTube, Skype, and Discord) to communicate and

¹² A full review of digital anthropology’s interest on social networks is beyond the scope of this article.

socialise digitally. This community broadly calls itself “anitwitter”, and are Brazilian fans of Japanese media – what are usually called, in the West, “otakus” or “weebs”¹³.

My interest in this community pre-dated my online engagement with them – part of my previous research had been about so-called “eventos de anime” (anime conventions),¹⁴ and it was based on this previous interest that I was drawn into otaku online sociability, and towards anitwitter in particular.

Moving from “traditional”, face-to-face, ethnography into digital ethnography was a bit of a struggle,¹⁵ but soon enough I was able to follow and interact with many members of anitwitter – although, as I soon found out, a complete mapping of the community would be impossible. Be that as it may, between late 2013 and early 2015 I ended up following close to 150 community members.¹⁶ Among them, I became particularly interested in the 40 or so that contributed to blogs – as I learned, writing in blogs, recording podcasts, and producing YouTube videos was a way in which community members could acquire prestige and showcase their cultural capital (Bourdieu, 2004). Indeed, presenting your knowledge of Japanese media artworks (anime, manga, games, and so on) is a central feature of otaku sociability, both online and offline, and as such it became a central feature of my research.

I was soon drawn to the fact that there was, then, a striking division in anitwitter. A friendly rivalry existed between two “big players” in the blogging scene. One of these players was an older and established blog – which I refer to, here, as The Settled¹⁷ –, which already had constructed, over the previous four years, a loyal readership, and a number of respected contributors. The other one was an association of smaller blogs which, less than two years prior, joined together into a larger portal – which I refer to, here, as The Newcomers – and which, despite their smaller individual size, were, as a group, quite a force to be reckoned with.

The rivalry between the Settled and the Newcomers was very peculiar. On the one hand, it was considered that both sides were indeed competitors in what can be called the “political economy of attention” (Pedersen, Albris & Seaver, 2021) – that is, revenue generated by clicks and views of their published content, as well as affiliate links to

¹³ There is a significant difference between Japanese “otaku” and Western “otaku”, but that distinction is not relevant for this article. However, for an overview, see Galbraith *et al.* (2015), Galbraith (2019), and, particularly, Eng (2006). Overviews of otaku studies in Brazil can be seen in Lourenço (2006), and Schüler-Costa (2015).

¹⁴ See Schüler-Costa (2014).

¹⁵ I recount some of it in Schüler-Costa (2016).

¹⁶ The community itself had, easily, tens of thousands of members, but for my research I initially focused on those who held higher prestige.

¹⁷ Names mentioned in this article are pseudonyms.

merchandise, and so on. This attention to content was considered to be a limited resource: the more readers Settled had, the less people would pay attention to Newcomers, and vice-versa. On the other hand, however, this rivalry was also very much symbolic: it was about the prestige, visibility, and fame within the community that came with being associated with one of the big players. The possession of prestige could be sought for utilitarian reasons – it was often the case that prestigious community members knew the cool people, were invited to the fun parties, or were employed at the best jobs¹⁸ –, but it was also the case that prestige was pursued for its own sake.

However, as is very common elsewhere, this rivalry took the form of a joking relationship (Radcliffe-Brown 1940a), permeated by friendly teasing and banter between members of either side; as such, public (indeed, very public, as most of them were on online forums such as Twitter and Facebook) interactions between members of the Settled and the Newcomers would be sometimes in playful disrespect, often in sincere friendliness, and even, albeit rarely, in downright hostility. Regardless of the form, the content of these interactions would almost invariably be the same: which party, the Settled or the Newcomers, was on top on the community's hierarchy? Which side had more viewers, readers, listeners, which side had more visibility within the community, which side made more people join anitwitter, and which side appealed more to anitwitter veterans?¹⁹

However, as I kept my participant observation with this group going on – chatting with them on public (and some private) fora, observing their interactions between themselves, and just, in general, carrying out what has sometimes been called “deep hanging out” (Rosaldo in Clifford, 1996)²⁰ – I came to realise that despite the importance assigned by the community to these two big players, their duopoly itself was debatable. Not only there was a plethora of smaller blogs who were happy not to engage in this contest, but there were dissenting voices who argued against this particular brand of snobbery and hierarchy. Indeed, if the traditional way of acquiring prestige in the community was through art *criticism*, there were some people who were dedicating themselves to social *critique*. It was this interplay between criticism –

¹⁸ Usually in the anime/manga cultural industry, such as comic publishers or dubbing studios.

¹⁹ This dispute for prestige often took the form of a search for aesthetic distinction in artistic taste (Bourdieu, 2013). For an analysis on their production and reproduction of taste, see Schüler-Costa (2015).

²⁰ Clifford uses Renato Rosaldo's phrase while drawing a distinction between two ethnographic methods: “a practice of intensive dwelling (the 'tent in the village')” on one side, and “more a matter of repeated visiting, collaborative work” on the other (Clifford, 1996:5).

and the production and reproduction of social hierarchy resulting from it – and its critique that, then, became my main research focus²¹.

By then, however, my sole ethnographic method had been digital participant observation. I had interviewed some people, but was soon dissatisfied with the realisation that, going back to Boissevain, that was “prohibitively time consuming and likely to inundate the researcher with data to the point of immobility” (Boissevain, 1973:147). If I wanted to show (as I did) both that the Settled and the Newcomers dominated the community but also that there was a significant presence of “unaligned” community members, I had to find a way to gather that data and analyse it. And so I did.

Data mining and social network analysis on Twitter: methods and results

Before I tell you how I went on to do so, I need to preface this brief methodological discussion by informing my reader that I consider myself to be, now and then, at best a beginner in programming. I never had any real training, and at most dabbled with code now and then. I know real programmers and I’m definitely not one of them. However, I was (and am) interested in using computational methods in my research, and I had already played out with the idea of using a script to mine and store tweets from anitwitter to allow for asynchronous analysis, but soon gave up on the idea for methodological reasons.

So, when I found myself thinking about data mining and social network analysis, I roughly knew where to start. I had already learned how to navigate the Twitter API,²² and taught myself a bit of R, particularly `twitteR`,²³ following online tutorials. However, after doing some research (and thinking of making my life easier), I ended up following a tutorial which used NodeXL for data mining,²⁴ and Gephi²⁵ for data visualisation.

The first step, then, was to mine the relevant data. Based on my ethnographic data, I chose 43 Twitter profiles which would be the “central nodes”²⁶ of this network – some of which were the “official” profiles of the blogs themselves, and some of which were the “personal” profiles of their contributors. It is important to say, at this step, that this decision

²¹ A full exposure of this, however, goes beyond the scope of this article. See Schüler-Costa (2015) for this discussion.

²² The tool which gives access to tweets and users straight from Twitter itself.

²³ R is a programming language, and `twitteR` is a package which integrates R with the Twitter API.

²⁴ NodeXL was, then, a free Excel add-on, which allowed for the import of data from Twitter and Facebook into an Excel spreadsheet. In the six years since, the software has changed significantly and, as such, I cannot vouch for its current version.

²⁵ Gephi is an open-source network analysis and visualisation software.

²⁶ Which is to say, NodeXL would only gather data of who they follow, and who follows them.

was, inherently and unequivocally, subjective. My usage of a computational method does not, in any way, make it less so. To paraphrase Mitchell, computers have no magical property to confer objectivity onto the data. Be that as it may, based on these 43 profiles I was able to collect a network of 23573 nodes (profiles) and 49667 edges (follower-ships). This network was then exported from NodeXL to Gephi.

On Gephi, I was able to start working with data visualisation. To begin with, I ran two algorithms: Eigenvector Centrality (to determine how “well-connected” each node is) and Modularity (to identify “communities” within the network). Nodes were then re-sized based on centrality and coloured based on modularity. Centrality is quite simple to understand: in this case, is mostly a factor of how many links within the network an account has (the more links, the larger the node)²⁷. Modularity, however, is “is one of the outstanding issues in the study of networked systems” (Newman, 2006:8577), and there is not, thus far, an accepted way of applying such algorithms. Deciding whether an individual belongs to this or that community, as well as the number of communities within a network, are, in the end, inevitably arbitrary decisions (no matter whether the arbiter is the researcher or the algorithm).²⁸ Be that as it may, the Gephi modularity algorithm did (roughly) identify the groups I had identified through qualitative research. Below is a graph which showcases the different groups identified by the algorithm.

²⁷ Again, it is necessary to highlight that it is by number of links *within the network*. Otherwise accounts such as @barackobama, then with 59,486,793 followers, would overwhelm the graph – despite the fact that only one person within the 43 nodes of interest followed him.

²⁸ It is noteworthy that this introduces what could be called a “non-human agency” into the research process. A discussion on this, however, goes beyond the scope of this article.

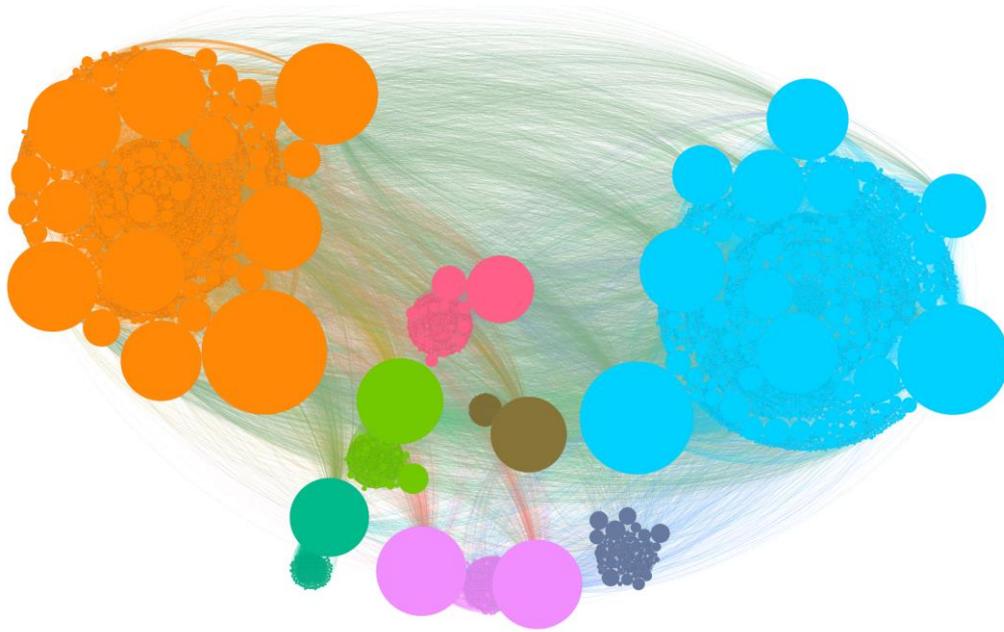


Figure 1: Communities identified via modularity algorithm

This image shows the Newcomers in orange and the Settled in blue. Interestingly, however, in purple we have the official Twitter accounts of *both* Settled and Newcomers. That is to say: although there are clear communities around the personal profiles of contributors, there is also a separate community surrounding the institutional profiles of both sides of the rivalry – people who possibly did not care for the rivalry at all. Also of interest, of course, are the five other communities in the figure – all of which surround blogs and/or contributors which did not belong to either side of the rivalry but whose presence in anitwitter was significant.

This graph, however, artificially separates the different communities identified by the modularity algorithm. Another graph, which I believe better represents anitwitter, can be seen below:

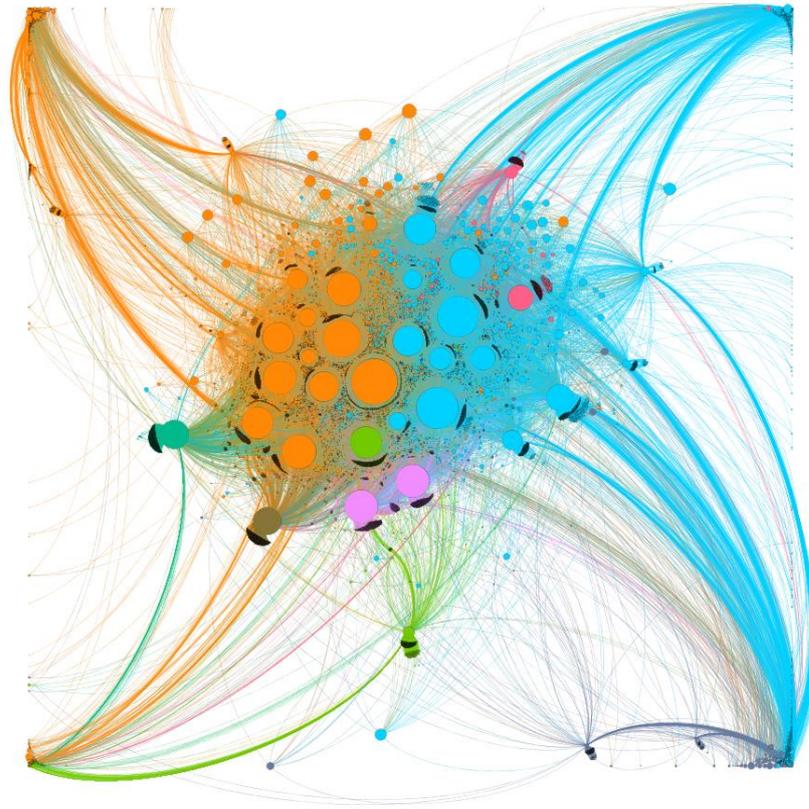


Figure 2: Anitwitter network graph

In this graph, we see a highly connected centre (where most of our nodes of interest are) and many edges at the graph's border (these are accounts which are the least connected within the network). Or, in other words, a community which is very *intra*-connected but also considerably *extra*-connected. The distinction between the two big players (the Newcomers in orange, the Settled in blue) is less accentuated, but we see they are still quite central (literally and metaphorically) to the network. Which helps illustrate the ethnographic argument I have previously shown: this is a community dominated by a rivalry between two sides but, also, with a significant presence of “unaligned” members.

Of particular interest to my argument here is how much the quantitative aligns with the qualitative data. My understanding of the anitwitter community had been, until then, supported by solid qualitative data gathered from participant observation and interviews. However, one of the reasons I decided to pursue social network analysis as a method was an

interest in seeing whether the quantitative method would paint a different picture from its qualitative counterparts – instead, they showed a surprising complementarity and synergy.

There was new information provided by the network analysis, of course. I was especially surprised to see the algorithm considering the two institutional profiles and its followers (in purple in both images) as a “community” of itself – which then led me to inquire whether, indeed, the rivalry between Settled and Newcomers was larger than the blogs themselves.²⁹

Furthermore, this complementarity between qualitative and quantitative methods illustrates yet another reason why a mixed-methods approach might be useful: it highlights the hard problem of community detection. It is widely recognised, both in the “social” and the “hard” sciences, that there is no easy way of identifying communities in a social network.³⁰ By using social network analysis to strengthen ethnographic insight, I was able to achieve a clearer picture of the communities I studied – in a way that neither of these methods alone would be able to achieve.

Conclusion

87

In this article, I explored how ethnography – the bread-and-butter of contemporary anthropology – can be enriched with the addition of computational methods. Using the case study of social network analysis and digital ethnography, I show the complementarity of qualitative and quantitative methods towards a better understanding of social reality.

It is important, however, for me to introduce a few caveats. First and foremost, this article is not an exercise on the field of network studies. My main interest here is to enrich anthropology by incorporating into it network analysis – network analysts would probably find my research uninteresting. A second caveat, following from this, is that I am not an expert on the computational methods I used here. Like Mitchell when he first started to develop upon the idea of social networks (Mitchell, 1974), I am merely a neophyte interested in having a clearer picture of social reality. A third caveat is that I do not propose that computational methods replace the qualitative research that social anthropologists do best. On the contrary: as I tried to show in this article, I consider my analysis to be meaningful because of my long-term qualitative engagement with the community studied. And a fourth caveat,

²⁹ However, due to financial and time constraints, this research question could not be pursued.

³⁰ See Fortunato (2010) for an extensive review of the question. Tantipathananandh *et al.* (2007) show that the problem is even harder for diachronic analysis of social networks.

following from this, is that I do not think all qualitative research can or will be enriched by quantitative and/or computational methods. For my current research I did one year of “traditional” participant observation among computer scientists and roboticists, and I did not use any quantitative or computational method, because I did not consider it would benefit my analysis and my research interests.

What I am arguing here, thus, is that we anthropologists (and other social scientists) think of our methods as tools in a toolbox. In most cases, good old participant observation might very well perfectly do the job we need. In some cases, it might be interesting to engage other methods depending on our research interests and demands from the field itself – the same way that we already sometimes engage in archival research, life histories, video documentary, and so on. Adding one or another tool to our toolbox is not too much of a problem – as long as we do not fall into the trap of, being armed with a hammer, see all research projects as nails. That is to say, we should use the methods that fit our research, rather than engage with methods for their own sake. Or, as Boissevain rather harshly puts it, regarding network analysis:

Far too much of the research now being done on networks lacks any clear formulation of the problems it seeks to resolve. Networks are compared with regard to density, size, and even composition, much in the way butterfly-collectors compare the colouring, wingspread, and number of spots of their favourite species. Trivial but extremely costly results based on samples of thousands are put forward with great solemnity by sociologists. Thus we learn that if you ask several hundred persons to name a few persons outside their household with whom they have close relationships, these turn out typically to be kin and friends. Other studies have discovered that affective relations change over time. Is this news? What is the social or theoretical significance of these ‘scientific’ discoveries? (Boissevain, 1979:393).

If what my reader takes from this article is the idea that “I should add network analysis (or indeed any computational approach) to my ethnography no matter what!”, I will have failed. Rather, I want my reader to take to heart what Diana Forsythe has argued regarding the skill of ethnography (Forsythe, 2001:148-149): the most important “research instrument” in the ethnographic fieldworker’s toolbox is the fieldworker themselves. It is through “calibrating” our data-gathering methods to the realities of fieldwork that we achieve a better understanding of social reality. Any one-method-fits-all approach to social research is, inevitably, doomed to fail.

In the preface for the recent edited book *Big Data in Computational Social Science and Humanities*, Shu-Heng Chen writes that “to get immersed in the deplorable conditions of workers under an industrial capitalist society, one can probably learn more from Charles

Dickens, say, in his *Great Expectations*, than from axiomatic or mathematical analysis alone” (Chen, 2018:vi).

I would go further and say (no disrespect to Dickens) that one can also learn a great deal in the original ethnography of worker life – Engels’ *The Condition of the Working Class in England*:

Such are the various working-people’s quarters of Manchester as I had occasion to observe them personally during twenty months. If we briefly formulate the result of our wanderings, we must admit that 350,000 working-people of Manchester and its environs live, almost all of them, in wretched, damp, filthy cottages, that the streets which surround them are usually in the most miserable and filthy condition, laid out without the slightest reference to ventilation, with reference solely to the profit secured by the contractor. In a word, we must confess that in the working-men’s dwellings of Manchester, no cleanliness, no convenience, and consequently no comfortable family life is possible; that in such dwellings only a physically degenerate race, robbed of all humanity, degraded, reduced morally and physically to bestiality, could feel comfortable and at home (Engels, 2009:75).

Quantitative and computational methods can tell us a lot about social reality. So can, clearly, qualitative methods. And, I argue, sometimes it takes a mixed-method approach to further elucidate social reality. As we anthropologists like to say, the field takes you where you need to go.

References

- BARNES, J. A. (1954), “Class and Committees in a Norwegian Island Parish”. *Human Relations*, v. 7, n. 1, pp. 39-58 [Consult. 25-11-2021]. Available in <https://doi.org/10.1177/001872675400700102>
- BARNES, J. A. & HARARY, Frank (1983), “Graph theory in network analysis”. *Social Networks*, v. 5, n. 2, pp. 235-244 [Consult. 25-11-2021]. Available in [https://doi.org/10.1016/0378-8733\(83\)90026-6](https://doi.org/10.1016/0378-8733(83)90026-6)
- BERNARD, H. Russel (2017), *Research Methods in Anthropology: Qualitative and Quantitative Approaches*. Blue Ridge Summit, Rowman & Littlefield Publishers.
- BOISSEVAIN, Jeremy (1973), “An exploration of two first-order zones”, in Boissevain, J. & Mitchell, J. C. (eds.), *Network Analysis: Studies in Human Interaction*. The Hague, Mouton & Co, pp. 125-148.
- BOISSEVAIN, Jeremy (1979), “Network analysis: a reappraisal”. *Current Anthropology*, v. 20, n. 2, pp. 392-394 [Consult. 25-11-2021]. Available in <http://www.jstor.org/stable/2741937>
- BOISSEVAIN, Jeremy & MITCHELL, J. Clyde (1973), *Network Analysis: Studies in Human Interaction*. The Hague, Mouton & Co.
- BOTT, Elizabeth (1990 [1955]), “Conjugal Roles and Social Networks”, in Trist, E., Murray, H., & Trist, B. (eds.). *The Social Engagement of Social Science, Volume 1*. Philadelphia, University of Pennsylvania Press, pp. 323-350.

- BOURDIEU, Pierre (2004), "The Forms of Capital", in Ball, S. J. (ed.), *The RoutledgeFalmer Reader in Sociology of Education*. London, Routledge, pp. 15-29.
- BOURDIEU, Pierre (2013), *Distinction: A Social Critique of the Judgement of Taste*. New York, Taylor & Francis.
- CHEN, Shu-Heng (2018), *Big Data in Computational Social Science and Humanities*. Springer International Publishing.
- CLIFFORD, James (1996), "Anthropology and/as Travel". *Etnofoor*, v. 9, n. 2, pp. 5-15 [Consult. 25-11-2021]. Available in <http://www.jstor.org/stable/25757889>
- CRUMP, Thomas (1992), *The Anthropology of Numbers*. Cambridge, Cambridge University Press.
- ENG, Lawrence (2006), *Otaku engagements: Subcultural appropriation of science and technology*. Thesis (Ph.D. in Science and Technology Studies). Rensselaer Polytechnic Institute, Ann Arbor. 239 p.
- ENGELS, Friedrich (2009 [1877]), *The Condition of the Working Class in England*. Oxford, Oxford University Press.
- EPSTEIN, Alfred L. (1969 [1961]), "The Network and Urban Social Organization", in Mitchell, J. C. (ed.). *Social Networks in Urban Situations: Analyses of personal relationships in Central African towns*. Manchester, Manchester University Press, pp. 77-116.
- FISCHER, Michael D. (1994), *Applications in computing for social anthropologists*. London, Routledge.
- FORSYTHE, Diana E. (2001), *Studying those who study us: An anthropologist in the world of artificial intelligence*. Stanford, Stanford University Press.
- FORTES, Meyer & EVANS-PRITCHARD, Edward E. (2016 [1940]). *African Political Systems*. London, Routledge.
- FORTUNATO, Santo (2010), Community detection in graphs. *Physics Reports*, v. 486, n. 3, pp. 75-174 [Consult. 25-11-2021]. Available in <https://doi.org/10.1016/j.physrep.2009.11.002>
- FREEMAN, Linton C. (2004), *The Development of Social Network Analysis: A Study in the Sociology of Science*. Vancouver, Empirical Press.
- GALBRAITH, Patrick W. (2019), *Otaku and the Struggle for Imagination in Japan*. Durham, Duke University Press.
- GALBRAITH, Patrick W., KAM, Thiam Huat, & KAMM, Björn-Ole (2015), *Debating Otaku in Contemporary Japan: Historical Perspectives and New Horizons*. London, Bloomsbury Publishing.
- HAGE, Per & HARARY, Frank (1983), *Structural Models in Anthropology*. Cambridge, Cambridge University Press.
- HORST, Heather A & MILLER, Daniel (2012), *Digital Anthropology*. London, Bloomsbury Academic.
- HYMES, Dell H. (ed.) (1965), *The Use of Computers in Anthropology*. Berlin, De Gruyter Mouton.

- KAPFERER, Bruce (1973), "Social networks and conjugal roles", in Boissevain, J.; Mitchell, J. C. (eds.), *Network Analysis: Studies in Human Interaction*. The Hague, Mouton & Co, pp. 83-110.
- KUPER, Adam (2014), "Foreword", in Malinowski, B. *Argonauts of the Western Pacific*. London, Routledge.
- LIVINGSTON, Eric (1986), *The ethnomethodological foundations of mathematics*. London, Routledge.
- LOURENÇO, André L. C. (2006), *Otakus: Construção e Representação de si entre Aficionados por Cultura Pop Nipônica*. Tese (Doutorado em Antropologia Social). PPGAS, Museu Nacional, UFRJ, Rio de Janeiro.
- MALINOWSKI, Bronislaw (2014 [1922]), *Argonauts of the Western Pacific*. London, Routledge.
- MITCHELL, J. Clyde (1956). *The Yao Village*. Manchester, Manchester University Press.
- MITCHELL, J. Clyde (1959), *The Kalela Dance: Aspects of Social Relationships among Urban Africans in Northern Rhodesia*. Manchester, Manchester University Press.
- MITCHELL, J. Clyde (1967), "On Quantification in Social Anthropology", in Epstein, A. L. (ed.), *The Craft of Social Anthropology*. London, Tavistock, pp. 17-45.
- MITCHELL, J. Clyde (1969), *Social Networks in Urban Situations: Analyses of Personal Relationships in Central African Towns*. Manchester, Manchester University Press.
- MITCHELL, J. Clyde (1973), "Networks, norms and institutions", in Boissevain, J.; Mitchell, J. C. (eds.), *Network Analysis: Studies in Human Interaction*. The Hague, Mouton & Co, pp. 15-35.
- MITCHELL, J. Clyde (1974), "Social Networks". *Annual Review of Anthropology*, v. 3, n. 1, pp. 279-299 [Consult. 25-11-2021]. Available in <https://doi.org/10.1146/annurev.an.03.100174.001431>
- MITCHELL, J. Clyde (1980), *Numerical techniques in social anthropology*. Philadelphia, Institute for the Study of Human Issues.
- NEWMAN, M. E. J. (2006), "Modularity and community structure in networks". *Proceedings of the National Academy of Sciences*, v. 103, n. 23, pp. 8577-8582 [Consult. 25-11-2021]. Available in <https://doi.org/10.1073/pnas.0601602103>
- PEDERSEN, Morten A., ALBRIS, Kristoffer, & SEAVER, Nick (2021), "The Political Economy of Attention". *Annual Review of Anthropology*, v. 50, n. 1, pp. 309-325 [Consult. 25-11-2021]. Available in <https://doi.org/10.1146/annurev-anthro-101819-110356>
- RADCLIFFE-BROWN, Alfred R. (1940a), "On joking relationships". *Africa*, v. 13, n. 3, pp. 195-210 [Consult. 25-11-2021]. Available in <https://www.jstor.org/stable/1156093>
- RADCLIFFE-BROWN, Alfred R. (1940b), "On Social Structure". *The Journal of the Royal Anthropological Institute of Great Britain and Ireland*, v. 70, n. 1, pp. 1-12 [Consult. 25-11-2021]. Available in <https://www.jstor.org/stable/2844197>
- RADCLIFFE-BROWN, Alfred R. (1964 [1937]), *A Natural Science of Society*. Glencoe, The Free Press.

- SANJEK, Roger (1974), "What is network analysis, and what is it good for?". *Reviews in Anthropology*, v. 1, n. 4, pp. 588-597 [Consult. 25-11-2021]. Available in <https://doi.org/10.1080/00988157.1974.9977130>
- SANJEK, Roger, & TRATNER, Susan W. (2015), *eFieldnotes: The Makings of Anthropology in the Digital World*. Philadelphia, University of Pennsylvania Press.
- SCHÜLER-COSTA, Vlad (2014), "O being e o becoming otaku: trajetórias e carreiras de fãs de animê e de mangá". *Simbiótica*, n. 6 [Consult. 25-11-2021]. Available in <https://periodicos.ufes.br/simbiotica/article/view/8078>
- SCHÜLER-COSTA, Vlad (2015), "*Mais do que fãs*": o universo da crítica de animê e mangá na Internet. Dissertação (Mestrado em Antropologia Social). PPGAS, Museu Nacional, UFRJ, Rio de Janeiro. 119p.
- SCHÜLER-COSTA, Vlad (2016), "On Conducting Ethnography in Twitter". *Medium*. [Consult. 25-11-2021]. Available in <https://medium.com/@vladschuler/ethnography-in-twitter-81b363493171>
- SCOTT, John (2000), *Social Network Analysis: A Handbook*. London, SAGE Publications.
- SEAVER, Nick (2014), "Computers and Sociocultural Anthropology". *Savage Minds*. [Consult. 25-11-2021]. Available in <https://savageminds.org/2014/05/19/computers-and-sociocultural-anthropology/>
- THOMAS, William I. & ZNANIECKI, Florian (1996 [1918]), *The Polish Peasant in Europe and America: A Classic Work in Immigration History*. Chicago, University of Illinois Press.
- TANTIPATHANANANDH, Chayant, BERGER-WOLF, Tanya, & KEMPE, David (2007), "A framework for community identification in dynamic social networks". *Proceedings of the 13th ACM SIGKDD international conference on Knowledge discovery and data mining*. San Jose, California, USA. Association for Computing Machinery. [Consult. 25-11-2021]. Available in <https://doi.org/10.1145/1281192.1281269>
- URRY, James (1972), "'Notes and Queries on Anthropology' and the Development of Field Methods in British Anthropology, 1870-1920". *Proceedings of the Royal Anthropological Institute of Great Britain and Ireland*, n. 1972, pp. 45-57 [Consult. 25-11-2021]. Available in <http://www.jstor.org/stable/3031732>
- VERMEULEN, Hendrik F. (2008), *Early history of ethnography and ethnology in the German enlightenment: anthropological discourse in Europe and Asia, 1710-1808*. Thesis (Ph.D.). Leiden University, Leiden.
- WERBNER, Richard (2020), *Anthropology after Gluckman: The Manchester School, colonial and postcolonial transformations*. Manchester, Manchester University Press.
- WOOLF, Greg (2010), *Tales of the Barbarians: Ethnography and Empire in the Roman West*. Oxford, Wiley.

Resumen

Basado en una etnografía digital conducida entre 2013 y 2015 en el servicio de red social Twitter, este artículo desenvuelve una discusión teórica-metodológica sobre el uso de la minería de datos y la visualización de redes como complemento “cuantitativo” al enfoque “cualitativo” de la observación participante. A partir de una revisión bibliográfica, sostiene que la etnografía, desde sus inicios, ha estado abierta a la investigación “cuantitativa”. Este artículo también aborda los orígenes antropológicos de la terminología y la metodología del análisis de redes sociales, recuperando así el interés antropológico por el tema. Por último, basándose en un estudio de caso empírico, demuestra un posible uso de dichos métodos como forma de enriquecer el análisis etnográfico.

Palabras-clave: antropología social; etnografía digital; análisis de red; metodología.

Resumo

Baseado em etnografia digital conduzida entre 2013 e 2015 no serviço de rede social Twitter, este artigo desenvolve uma discussão teórico-metodológica sobre o uso de mineração de dados e visualização de redes como complemento “quantitativo” à abordagem “qualitativa” da observação participante. Partindo de uma revisão bibliográfica, argumenta que a etnografia, desde seus primórdios, esteve aberta a pesquisa “quantitativa”. Este artigo também aborda as origens antropológicas da terminologia e da metodologia de análise de redes sociais, recuperando, portanto, o interesse antropológico no tema. Por fim, baseado em um estudo de caso empírico, demonstra uma possível utilização de tais métodos como forma de enriquecer a análise etnográfica.

Palavras-chave: antropologia social; etnografia digital; análise de redes; metodologia.
