



UvA-DARE (Digital Academic Repository)

Structure from interaction events

de Nooy, W.

DOI

[10.1177/2053951715603732](https://doi.org/10.1177/2053951715603732)

Publication date

2015

Document Version

Final published version

Published in

Big Data & Society

[Link to publication](#)

Citation for published version (APA):

de Nooy, W. (2015). Structure from interaction events. *Big Data & Society*, 2(2), 1-4. <https://doi.org/10.1177/2053951715603732>

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

Structure from interaction events

Wouter de Nooy

Big Data & Society
 July–December 2015: 1–4
 © The Author(s) 2015
 Reprints and permissions:
sagepub.com/journalsPermissions.nav
 DOI: 10.1177/2053951715603732
bds.sagepub.com



Abstract

In this contribution to the colloquium, I argue why and how I lost interest in the overall structure of social networks even though Big Data techniques are increasingly simplifying the collection, organisation, and analysis of ever larger networks. The challenge that Big Data techniques pose to the social scientist, I think, is of a different nature. Big Data on social actors mainly record events, e.g. interactions between human beings that happen at a point in time. In contrast, social network analysts tend to think in terms of social relations that exist over a timespan. The challenge, then, is to rethink our conceptions and models of social relations and social structure. I conceptualize social structure and social relations as forces. I propose modelling these forces with regression models for longitudinal interaction data.

Keywords

Network analysis, regression models, relational events, agency, sociality

This contribution, as does most of my work, focuses on interaction data, a term I use as shorthand for data on actions by a social actor directed at someone or something else. This is the type of data that Big Data techniques may generate in large quantities and at low cost: time-stamped events, identifying a person or organization doing something towards another person, organization, or object, and possibly with a qualification of the action contents. How should we conceptualize and model social structure and social relations if we analyse this type of data?

Network analysis: Not the Big picture

Initially, I equated social structure with the overall network that I could piece together from interaction data. In my analyses of literary criticism, I coerced literary evaluations published within a time slice into a network and I compared the structure of this network to social characteristics of the authors and critics (de Nooy, 1991). The comparison was meant to show that field structure reflected overall social structure. In Bourdieu's words, I wanted to show the homology between the literary field and the social field (1993). Quite a few Big Data users still focus on overall network structure as the Big picture of social structure (e.g. in this journal, Sudhahar et al., 2015; Venturini et al., 2014).

I gradually lost interest in the overall network structure. It was evident that the selection of authors and critics to be included in the network mattered a lot to the resulting overall network structure – the perennial problem of network boundaries. In addition, overall structure was neither clear-cut nor stable. Different network analysis techniques and different time slices for aggregating evaluations yielded different results. More importantly, the overall network approach did not appeal to me theoretically because it did not address agency. I could speculate about why authors and critics evaluated each other positively or negatively but I could not theorize or model these decisions. Thus, social network analysis was separated from social theories about how people behave and from the multivariate statistical techniques commonly used for testing these theories. Two encounters made me abandon the analysis of overall network structure and embrace a

Amsterdam School of Communication Research (ASCoR), University of Amsterdam, Amsterdam, the Netherlands

Corresponding author:

Wouter de Nooy, Amsterdam School of Communication Research (ASCoR), University of Amsterdam, Nieuwe Achtergracht 166, Amsterdam 1001 NG, the Netherlands.
 Email: w.denooy@uva.nl



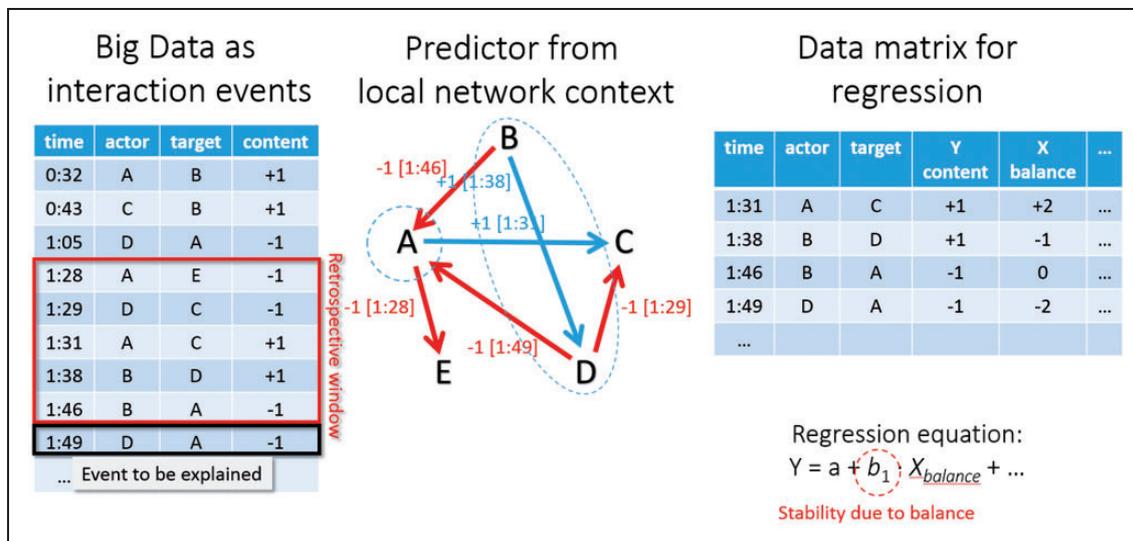


Figure 1. From time-stamped event data to a regression model with a predictor constructed from local network context.

local network approach that situates actors in the context of their direct network contacts.

The first encounter was with balance theory. Fritz Heider's concept of balance (1946, 1958) includes local network context in a behavioural hypothesis on affective relations. It states that people feel more comfortable if they agree rather than disagree with a person they like on a topic or a third person. Progressing from relations between two to three persons is a major sociological step (Simmel, 1902). It implies that a person's response to another person or object is neither based solely on his or her attributes, e.g. attitude, nor is it fully derived from (dis)similarities between the interacting persons. Instead, persons take into account the perceived or imagined response of another person and their affective relation with that person: What would my friend say about that? This action model makes sense to me. I can easily imagine that a human being is aware of and sensitive to interactions among people with whom she herself is interacting and that perceived interactions induce expectations of how one should interact.

Systematic application of the balance principle generates a remarkable overall network structure, namely polarization of the network into two (Cartwright and Harary, 1956) or more (Davis, 1967) antagonist groups. Relaxation of some aspects of the balance principle was shown to stratify groups into a hierarchy (Davis and Leinhardt, 1972). From a sociological point of view, the balance model addresses two dictionary meanings of *sociality*: (1) the tendency to form groups and (2) being friendly. The stratified overall network structure of layered antagonist groups also resonates with the central sociological concept of social stratification. Although obvious in retrospect, it took me quite

some time to realize that the overall network structure of interactions is merely a by-product of how social actors respond to their local network context. One should study the latter, not the former.

My second encounter was with statistical regression models for network data that apply a local network perspective. Stanley Wasserman and Philippa Pattison (1996) proposed to take pairs of network nodes as cases, the presence or absence of a line within the pair as the dependent variable, and construct predictor variables from local network structure. In a test of balance theory, for example, one would probably want to predict whether node B (Bob) evaluates node C (Christy) positively or negatively (dependent variable) from the number of balanced triples that would be created if Bob evaluates Christy positively in contrast to the number created if the evaluation would be negative. If one of the two counts outweighs the other, balance theory predicts that Bob will use the evaluation associated with that number. The linked video illustrates this (see also Figure 1).

Cross-sectional interaction data yields nasty observational dependencies, which are not easily solved in regression models. In addition, the causal order of events remains unclear: Did Bob's evaluation of Christy respond to Allison's evaluation of Christy or the other way around? Both the estimation and interpretation problems are alleviated if we work with longitudinal data; a realization that came to me when I studied the work by Tom Snijders (2001). In my opinion, regression models for longitudinal network data combine the best of two worlds, namely the world of relational analysis and the world of variable-based theorizing and testing.

Structure and social relations as forces

It has been noted that event data recorded with Big Data techniques require other theories and models than the more stable social ties studied by social network analysts (e.g. Burrows and Savage, 2014; Lazer et al., 2009). But what theories and models? Due to the encounters described above, I no longer see social structure as a stable overall pattern of social relations. Instead, I regard social structure as forces – I am under the spell of field theory – that stabilize interactions over time. As I see it now, social structure is the tendency of social actors to interact in predictable ways. To cope with their local interaction network, social actors must feel able to predict future interactions that involve them and they must interact in a way that is predictable to other actors. To be predictable, people must respond to previous interactions in systematic ways, which stabilizes their interactions: the iterational dimension of agency (Emirbayer and Mische, 1998).

Inspired by Jan Fuhse (2009), I distinguish between interactions as observable actions and social relations as relational expectations. In the present context, I focus on social relations as expectations about interactions with another social actor. Friendship, for example, may entail the expectation that both your actions towards your friend and your friend's actions towards you will be positive and more or less equally frequent. In addition, a friend may be expected to come to one's help when one is under attack. This is not to say that this set of expectations is always called friendship by the persons involved. It reflects a type of solidarity that may go under different names in different contexts, such as the family and the workplace, if it is named at all.

Expectations cannot be directly observed. They can be named and communicated – one person can declare himself to be another person's friend – but then the question arises: Which expectations does friendship entail for this person? Like probabilities, a researcher can infer relational expectations from patterns of observations, for example, with regression models that predict the occurrence or contents of interaction. Regression coefficients may indicate social relations because they capture the predictability of action; they represent stabilizing forces. I expect human beings to construct their relational expectations in a similar way from the interactions that they observe. Even if interaction partners explicitly discuss their expectations, they will check communicated expectations against the interactions that they observe.

If interaction regularities represent relational expectations, they are likely to vary for different partners. After all, we do not have the same social relation with all of our partners. For example, Allison may

have a higher probability of reciprocating an action by Bob than an action by Christy if her relation with Bob is stronger or different from her relation with Christy. Estimating different reciprocity coefficients for different directed pairs – Allison's relational expectations regarding Bob may differ from Bob's relational expectations towards Allison – requires many repeated observations for each pair. Such fine-grained longitudinal data can now be collected with Big Data techniques. In addition, the relative ease of continuous data collection over a longer period may even allow estimating effect changes over time, indicating that a social relation changes over time.

Big Data techniques also provide us with detailed information about interaction context. This information is indispensable for determining who are able to interact at a particular moment. Big Data more often than not tell us the exact whereabouts of persons in geographical space or on digital communication platforms, e.g. when a person is logged on to a web forum or opens an electronic message. With this information, we can distinguish persons who cannot interact from those who did not interact but could have interacted. Ignoring this distinction seriously compromises the interpretation of results in terms of agency and relations (Kitts, 2014). In addition, we can determine the importance of context to relational expectations. For example, if Bob meets Allison in a bar, he is more likely to reciprocate her actions than if he meets her in class. If an organization provides the context, the physical layout is usually geared towards permitting or stimulating particular types of interaction and interaction sequences. Staff who are present are instructed to interact in particular ways. In these ways organizations imbue people with relational expectations.

The strengths of Big Data techniques – capturing the context of interaction – highlight one of its current weaknesses, namely capturing the content and meaning of interaction. Spatial proximity tends to be measured rather than who interacts with whom (e.g. Madan et al., 2012). The contents of interaction, for example, a friendly versus unfriendly act, tend to be accessible only in the case of digital communication and even then it is difficult to (semi-)automatically tease out its probable meaning to the actors. Human coders may still be needed; additional qualitative data may have to be gathered.

Conclusion

Big Data techniques produce large quantities of event data, in particular data on social interactions as events. In the social sciences, event data is usually collected on a small scale, for example, in ethnomethodology, whereas large scale data collection tends to consist of

opinions and reported events in surveys. Surveyed opinions and reported events are typically regarded as relatively stable attributes reflecting a stable overall social structure. Event data problematizes the assumed stability at the individual and social level, suggesting that social actors need to work hard to create stability. People must interact in systematic ways to make their interactions somewhat predictable to their interaction partners, which helps the latter to respond and maintain a social relation. Providing large quantities of detailed event data, Big Data techniques may turn out to be a game changer for our theorizing of social networks and social structure.

In this essay, I argued that we should understand social relations as expectations based on previous interactions. Expectations help human beings to predict interactions within their local network context and respond accordingly. The expectations stabilize patterns of interactions and it is the stabilizing effect of these expectations that we should understand as social relations and social structure. In my opinion, the value of Big Data techniques resides in the unprecedented access to contextualized and longitudinal action at the micro level that allow us to model relational expectations. Big Data techniques offer us Big Detail, not Big Structure.

Acknowledgement

I am grateful to Jan Fuhse for our discussions and to this special issue's editors Robin Wagner-Pacifici, John Mohr, and Ron Breiger for their stimulating comments that have helped me to clarify and sharpen my argument.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

References

Bourdieu P (1993) *The Field of Cultural Production. Essays on Art and Literature*. Cambridge: Polity Press.

- Burrows R and Savage M (2014) After the crisis? Big Data and the methodological challenges of empirical sociology. *Big Data & Society* 1(1): 1–6.
- Cartwright D and Harary F (1956) Structural balance. A generalisation of Heider's theory. *Psychological Review* 63: 277–293.
- Davis JA (1967) Clustering and structural balance in graphs. *Human Relations* 20: 181–187.
- Davis JA and Leinhardt S (1972) The structure of positive interpersonal relations in small groups. In: Berger J (ed.) *Sociological Theories in Progress*. Boston: Houghton Mifflin, pp. 218–251.
- de Nooy W (1991) Social networks and classification in literature. *Poetics* 20(6): 507–537.
- Emirbayer M and Mische A (1998) What is agency? *American Journal of Sociology* 103(4): 962–1023.
- Fuhse JA (2009) The meaning structure of social networks. *Sociological Theory* 27(1): 51–73.
- Heider F (1958) *The Psychology of Interpersonal Relations*. New York, NY: John Wiley and Sons.
- Heider F (1946) Attitudes and cognitive organization. *Journal of Psychology* 21: 107–112.
- Kitts JA (2014) Beyond networks in structural theories of exchange. Promises from computational social science. *Advances in Group Processes* 31: 263–298.
- Lazer D, Pentland A, Adamic L, et al. (2009) Computational social science. *Science* 323(5915): 721–723.
- Madan A, Cebrian M, Moturu S, et al. (2012) Sensing the “health state” of a community. *IEEE Pervasive Computing* 11(4): 36–45.
- Simmel G (1902) The number of members as determining the sociological form of the group. I. *American Journal of Sociology* 8(1): 1–46.
- Snijders TAB (2001) The statistical evaluation of social network dynamics. *Sociological Methodology* 31: 361–395.
- Sudhakar S, Veltri GA and Cristianini N (2015) Automated analysis of the US presidential elections using Big Data and network analysis. *Big Data & Society* 2(1). DOI: 10.1177/2053951715572916.
- Venturini T, Baya Laffite N, Cointet J, et al. (2014) Three maps and three misunderstandings. A digital mapping of climate diplomacy. *Big Data & Society* 1(2): 1–19.
- Wasserman S and Pattison PE (1996) Logit models and logistic regressions for social networks. 1. An introduction to Markov graphs and p*. *Psychometrika* 61(3): 401–425.

This article is part of a special theme on *Colloquium: Assumptions of Sociality*. To see a full list of all articles in this special theme, please click here: <http://bds.sagepub.com/content/colloquium-assumptions-sociality>.