1

Introduction

BALÁZS VEDRES AND MARCO SCOTTI

1.1 Introduction: applied network science

This volume provides an overview of network science applied to social policy problems. Network science is arguably the most dynamic and interdisciplinary field that has grown up to address problems of an increasingly interconnected world. Social problems transgress disciplinary boundaries, especially with the ever-increasing complexity of our globally interconnected world society. It is natural that with complex problems such as loss in ecological diversity, economic crisis, spread of epidemics, and the safety of our food supply, we turn to a field of research that focuses on explaining complex dynamics, and that is inherently interdisciplinary itself.

Networks have become part of our everyday experience as we routinely use online social network services, we hear reports on the operations of terrorist networks, and we speculate on the six degrees of separation to celebrities and presidents. Less manifestly, we rely on vast and complex infrastructural networks of electric power distribution, Internet data routing, or financial transfers. We only ponder the complexity of these systems when we are faced with avalanche-like dynamics in their collapse, as major blackouts, system stoppages, or financial meltdowns.

With networks on the collective mind, there is ample interest in tools to understand and manage complex network systems of social ties. At the time of writing this introduction (in October 2011), there were eighteen applications available on Facebook to visualize one's social network. The popularity of such software tools shows our fascination with the interesting new perspective that the graphic visualization of friendship ties provides. While many of these little apps are for the purposes of entertainment, high-power software platforms are being put in place to assist scientific communities, business work teams, or emergency response teams to efficiently map out not only their social network neighborhood, but also who possesses critical

Networks in Social Policy Problems, eds. Balázs Vedres and Marco Scotti. Published by Cambridge University Press © Cambridge University Press 2012.

pieces of information, tacit knowledge, or expertise in their vicinity (Contractor, 2009). Network analysis has become an essential tool in the counter-insurgency battlefield. As the enemy is increasingly organized into non-hierarchical networks, army operatives need to learn methods to visualize and measure the weak points of the insurgency network (Bunker, 2006). In the capture of Saddam Hussein and Osama Bin Laden, a crucial element was a social network analysis approach that helped to sort relevant ties and key couriers from the vast pool of intelligence data.[1]

Network science tools are also being deployed to understand and manage biological or engineered systems. With the increasing reliance on green energy, and the depletion of fossil fuels, electric energy is becoming ever more prominent, increasing the load on the power grid. Network science tools help in the design of power grids that are more resilient to failure and attack (Holmgren, 2006; Pagani and Aiello, 2011). With the steady decline in fisheries productivity, there is increasing attention on the fragility of marine food webs. The simulation tools in Ecopath and Ecosim have been developed to understand the impact of policy and economic changes (such as the size of fishing vessels or the size of fishing conglomerates) on marine ecosystems from a network perspective (Pauly *et al.*, 2000). Such tools can help avoid fishing to the point where catastrophic collapses happen in food web networks (Espinosa-Romero *et al.*, 2009). Network models are being adopted in cancer research as well. The prediction of cancer metastasis connections (Chen *et al.*, 2009), and the better understanding of the metabolic links of cancer cells (Jeong *et al.*, 2000) all contribute to a more targeted treatment of cancer.

Behind these tools there is the new science of networks, an interdisciplinary field that aims at explaining complex phenomena at larger scales, emerging from the simple principles of making network links locally. Witnessing a decade of explosive growth, academics from the natural and social sciences are engaged in basic research to unveil principles that govern network behavior in a wide range of empirical fields. The power of this field lies in its strong and diverse roots: in the social sciences, attention to the social web dates back to the works of Georg Simmel, from the last decade of the nineteenth century (Simmel, [1922] 1964). Drawing graphs of social ties (sociograms) became a key tool of psychology in the 1930s through the work of Jacob L. Moreno, who famously aspired to draw the social network graph of the entire New York City, and got as far as drawing a graph of 435 individuals (Moreno, 1934).[2] Attention to networks was revived using computing power in the sixties and seventies, to understand social structure as blockmodels

[1] On May 12, 2011 the BBC news featured Dan Pearson from a company named i2, who presented the capabilities of their social network reconnaissance software system in a video interview (http://www.bbc.co.uk/news/technology-13336281).

[2] A high quality digital version of the original sociogram of 435 nodes is available here: http://www.asgpp.org/docs/WSS/WSS.html

(White *et al.*, 1976). By the nineties social network analysts had developed rich methods to estimate individual network centrality and power, role structures, balance, and structural holes. In the natural sciences the roots of network thinking reach back to the beginnings of graph theory, to the work of Leonhard Euler in the eighteenth century. Research on ecological networks dates back to the pioneering work of the Italian zoologist Lorenzo Camerano (1880), who constructed the earliest known graph of feeding relations. The theory of random graphs of Paul Erdős and Alfred Rényi from 1959 was crucial in catalyzing thinking about complex systems. Within the physics of condensed matter, studies of spin glasses in the seventies drew attention to the complexity arising from the stochastic interactions of particles. In biology, increasing attention was paid to the complex network dynamics of metabolic networks and protein interactions.

While the roots of network science are diverse, there has been frequent borrowing amongst disciplinary, methodological, and thematic groups throughout its history. Network science entered a new phase of integration during the first decade after the turn of the millennium. Scholars and research centers emerged following research agendas that truly integrated disciplines. The field of network science in its recent phase of evolution maintains a network-like structure that continually bridges and bonds disciplinary groups, avoiding compartmentalization, yet consolidating a common language and methodological toolbox. For example, protein–protein interaction models in biology borrowed from sociological metrics of betweenness centrality (Hahn and Kern, 2005). Or, social movement studies adopted methods directly from the physics of phase transitions (Oliver and Myers, 2003). The key researcher who best exemplifies the new science of networks is Albert-László Barabási. His work on scale-free network structures, the dynamics of events unfolding in networks, and the possibilities of controlling networks attracted tens of thousands of citations. His readership spans disciplinary boundaries, and his center brings together researchers with diverse backgrounds.

A shared feature of all network research is the use of innovative, experimenting, highly visual methods. Methods in network science are highly modular. As a contrast, consider linear regression (or any regression-like method): such methods force the structure of the technique on the whole of the research project. Our metaphor is a ten-ton highly specialized machine tool. Once you buy it, you have to shape your workshop to suit the inputs of the machine, hire workers with specialized skills, and find specific buyers for the output. While there are fixes and modifications to regression methods, the basic logic stays the same. In contrast, network methods are built modularly from general purpose tools, and allow great flexibility. One can take an algorithm developed in anthropology to identify typical kinship arrangements (like, for example, "blood marriage" – the marriage of cousins) and take this pattern search idea to citation network research, where instead of parents and descendants, there are citing and cited works. The notion of a certain lineage pattern can be used to identify emerging disciplinary schools, rather than kinship clans. Then one can augment the kinship pattern search algorithm by a statistical evaluation of pattern frequencies.

With recent advances in the availability of computing power, ideas about complex network dynamics are being tested with large datasets. The enhanced potential of high-speed computation has opened up new vistas to take on previously unimaginable challenges. With a large number of nodes, qualitatively different dynamics can be observed, similarly to the notion of thermodynamic limit in physics (with the increase in the number of particles, one can switch from conceptualizing colliding particles to the notion of gas). Visualization has always been a key element in network science. An efficient map of network topography is a great tool to show group divisions, central players, and main directions of flow. At websites like www.visualcomplexity.com, vast collections of visualizations are available, that border on an art form. Again, the idea is that complexity can be grasped given efficient tools.

While the conceptual roots and the applications are diverse, the fundamental principles of network science are shared. What is distinctive about the conceptual foundations of network science? What is, and what is not a case of networked phenomena? While taking a network approach is productive for many problem areas, one should refrain from seeing networks everywhere. It seems that the concept of network is being evoked all too casually these days: there is much talk of social networking, the old boy network, or a network society. It seems that networks are everywhere, but network science uses this concept with a careful definition. What is a network in a network science sense? The mere possibility of identifying nodes and drawing lines between them is different from identifying a network mechanism that explains complex outcomes from simple building blocks. At a railway station, for example, one could identify travelers as nodes, and draw ties between any two of them if they are headed to the same destination, or, if they purchased their ticket from the same teller. These ties, however, would hardly figure as part of a theoretically sound explanation for the behavior of individuals, or for the functioning of the transit system. However, if you drew a network where destinations are nodes, and the number of travelers moving between them is the tie, you can start to identify mechanisms of systemic breakdown in the case of a major accident.

While nodes and ties are the fundamental building blocks to any network science approach, the power of network science lies in focusing on triads and beyond: the broader topographical context. A connectionist perspective (focusing on dyads, pairs of nodes) does not exhaust the possibilities that a network perspective can offer. Recognizing that connections matter is only the first step: recognizing that ties around dyads matter begins to unleash the power of network science. A tie to a political party might matter for a firm to have access to resources and timely information on policies. Thus far this is a question of connections – is a firm connected to a party? But one can also ask: once a firm is connected to a political party, how are its ties to other firms affected? This becomes a network question: a tie formed at one part of the network system influences the probability of ties elsewhere, and might help explain the emergence of a politically polarized economy (Stark and Vedres, in press).

While this all might seem rather abstract, network science offers a fresh, simple set of heuristics to tackle practical problems around us. When we speak of complex dynamics, we think of processes that are beyond the reach of common sense. For example, the cause of a massive wave of demonstrations with several thousand participants (such as those in the Middle East over the spring of 2011) might elude our common sense if the timing of the demonstration shows no correlation with the timing of grievances. Why now, why so many? A network science approach would start from the ties between activists, and the way they convince each other to participate. Assuming the simplest process of peer pressure ("I go if at least two of my friends go"), the resulting process shows complex dynamics akin to an avalanche: massive participation is sudden, following a long period of gestation (Kim and Bearman, 1997; Oliver and Myers, 2003).

While network science can appear complex, even esoteric at first, it is readily applicable to a wide range of social problems. Instead of providing complex answers to simple questions, network thinking offers new heuristics, based on fundamental and simple, yet startling regularities. Our common sense in the subjective assessments of probabilities predominantly rests on the assumption of unconnected units (cells, human individuals, businesses, states), and their autonomous decision-making (Tversky and Kahneman, 1974). Such heuristics lead us astray in many problem-solving situations where dynamics are influenced by connected agents. As we show in this volume, the range of such problems is truly wide, from team creativity to compensation of executives, to financial crises and ecological disasters. As a start let us consider three examples for heuristics based on the assumption of connections and interplay that one should adopt from network science.

The first example of such heuristics is the idea of a *small world*. We generally assume that a new epidemic outbreak in the heart of Africa or a new mobile phone virus spreading in Taiwan are too far and insulated from us. In reality, our social networks are much more connected than we would prefer. While many of our ties are within closely knit clusters of friends and colleagues, there are enough long distance ties in our network vicinity, so that there is a short path starting somewhere near you to any part of the world. The concept of a small world was coined from the experience of, as we discover, the shared acquaintances with a random person we happen to sit next to on a long flight – what a small world! Duncan Watts, the researcher who first described this phenomenon in mathematical terms, showed that social networks occupy a very narrow range between order and randomness, where local clustering is high (there are dense pockets of friends), while overall path lengths are low (there is close reachability globally). It is much more probable to catch a new strain of flu, or to have your smart-phone turned into paperweight by a virus infection than you would assume. This research insight is now an integral part of epidemic forecasting (Pastor-Satorras and Vespignani, 2001).

A second example for a network heuristic is that there are no typical participants: networks are *scale-free*. Social and technological networks are scale-free; there is not a particular peak in network embeddedness, such that at any scale, inequalities show a similar distribution. This means that there are huge inequalities in network systems, such as the World Wide Web. While naïve approaches would assume that "getting online" is the decisive step to transgress the digital divide, the distribution of incoming links – the main channels of attention – is extremely unequal (Barabási and Albert, 1999). Getting online is just the first step, and for most website startups the investment into hardware, programming skills, and graphic design brings no visibility whatsoever. Most websites have no links pointing to them at all, while there are very few that attain millions of incoming links. In the world of webmasters making links, but also in the realm of choosing new friends or making new business partnerships, the winner seems to take all: links are made to participants with many links already. This is true for our social networks as well. Instead of most people having an average number of friends, with a few isolated and a few well connected individuals, most people have a relatively low number of friends, while there are several well connected individuals, then there are enough extremely well connected individuals, and even more than a few that are unbelievably well connected. This means that success in a connected world takes persistence: the first links are difficult to attract, but then there are increasing returns to scale: having more friends makes it much easier to gather more friends still. This also means that the robustness of social and technical networks hinges on a few key hubs, which highly limits their tolerance to directed attack.

The third example of a heuristic from network science is about disproportionate sizes of causes and effects. Processes of peer pressure and imitation, the spread of innovations, or the collapse of energy delivery systems all unfold in a non-linear way. If something becomes fashionable, suddenly almost everyone seems to carry that item (a new gadget or a new piece of clothing). This contradicts our common sense assumption that the outcome of a process is linearly proportional to the amount of energy invested – that large events need big causes. We assume that a particularly popular brand of shoe must be a result of an expensive marketing campaign. Charles Tilly was among the first to point out how deeply this heuristic is rooted in our conceptions of social change (Tilly, 1984): historians assume that large revolutions are ignited by the gravest of grievances, for example. What network science found, instead, is that when social movements are near critical phase transition points, a small increase in activism can lead to massive protests and revolutions. In the case of engineered systems we are more accustomed to the notion that a small fault can lead to a major collapse, as in the case of the 2003 blackout of North America (Albert *et al.*, 2004). In the case of ecological systems, it is much less part of public consciousness that beside protecting charismatic species (that are easy to stand up for, such as whales or turtles) one also needs to pay attention to the keystone species (possibly a small crustacean). Keystone species show disproportionate effects on ecosystems in comparison with their relative abundance. Such species are at key positions in the ecological network, and losing them can set off major avalanches of ecological collapse (Ulanowicz, 2009).

1.2 The structure of this volume

This volume provides practical examples along these lines, arranged in three parts that correspond to the three network heuristics outlined above. The first part of the book focuses on the construction and structure of social networks among individuals, and how the structure of such networks influences information flows, and the emergence of new ideas. The heuristic of a small world connects the typical structure of social networks with cohesive groups and occasional long distance ties with the speed and dynamics of spreading information and epidemics. This part further elaborates on the impact of our social networks on the flow of ideas. Chapters in this part offer insights into how the decision to pass on ideas is conditioned by properties of network ties, and how network structures help in generating ideas in the first place.

Chapter 2 focuses on what predicts spreading information in a social network, using the small-world network approach. While counterintuitive at first, compared with mass media, social network ties can be more effective channels of information. Mass media can reach more people faster with a message but it is less likely that the message will be acted upon. This is of especially grave concern for public advisories and marketing campaigns. A message that spreads via a social network tie – passed on by an acquaintance – can be customized for the receiver, can be told in the right context and in the right moment and, most importantly, can be endowed by the trust that the receiver has in the sender. These information transmission ties, however, need to be placed in the network context of other ties. Depending on the network location where information starts spreading, the diffusion process might reach almost everyone fast, or it might peter out on the peripheries.

Networks are effective channels for information flow, but can they also contribute to the generation of new ideas? Chapter 3 focuses on teams in creative fields – science and Broadway musical production – to identify the significance of network connections in success. The prevalence of teamwork in creative production is clearly increasing: new ideas are increasingly products of connected individuals rather than solitary geniuses. Teams themselves can be more or less successful and this is a function of their position in a larger collaborative network. The most successful teams are composed of a mix of incumbents with trusted and tried collaborative ties, and newcomers bringing fresh ideas and access to as yet untapped areas of the network. Thus, again, the dual nature of our social networks leaves clear marks: we need cohesive groups to produce ideas, but these groups are most efficient if they make use of the long reaching ties, to be embedded in the broader flow of ideas.

Chapter 4 identifies a network structure that is closely associated with the generation of new ideas. The focus here is on the network structures that enable novel insights. Truly innovative ideas – in the first instance, a fresh conceptualization of the problem itself – are not free-floating outside social groups. Innovation is a generative recombination that requires familiar access to diverse resources. The chapter identifies a network structure where dense regions of ties overlap – a structural fold – that facilitates such recombination. There seems to be a tradeoff for social groups: maintain exclusivity for long-term stability with modest performance, or be open to overlaps for high performance and the risk of disintegration. The results of this chapter indicate though that business groups have devised ways to maintain stability while also reaping the innovation benefits of structural folding. Coherence of larger business groups – a business group proper – was achieved by a historical lineage of group breakup, merger, and reunification. Successful business groups seem to have an inner network structure with overlapping subgroups that are frequently re-shuffled to maintain a steady production of fresh ideas.

Chapter 5 focuses on the processes of composing a team in the context of nanotechnology, and potential entry points for policy-makers to intervene in team composition. While the preceding chapters aimed at uncovering the predictors of efficiently spreading and successfully inventing ideas, this chapter makes a further step to ponder the path by which the team is assembled. Successful teams possess both a favorable mix of attributes (along gender for example) and include stars with many ties at a broader network scale. A team brings together a mix of career histories and personal and social categories. From a policy-maker's perspective the question is: what are the factors that one can influence, encourage, or discriminate to help create better teams? A key balancing act here is not to stifle the power of self organization, while maintaining the tools that guide team formation. The key to successful intervention is knowledge about the emerging team network structure: the policy-maker can identify patterns and principles that none of the participants on the ground sees.

The second part broadens the conceptual horizon to organizations and institutions, and how network structures influence power and the abuse of power in an organizational context. In line with the heuristic of scale-free networks, the attention in this part is on inequalities of power and money. This part offers a shift of perspective: instead of the circuitry metaphor inherent in the small-world approach to the creation and spreading of ideas, this part focuses on power fields. Networks are far from being the domain of homogeneity and equity. To understand organizational structures of social movements, inequalities in CEO compensation, mechanisms of organizational capture, or the workings of political corruption, this part offers a look at power in networks.

Chapter 6 introduces the idea of organizational fields, focusing on the workings of social movements. The first question that arises when one tries to understand collaboration and power in a network, is where the boundaries of the network lie, and how these boundaries get defined. While organizations have natural boundaries, the perimeter of a collaborating network of organizations is much more blurred. The internal structure of social movement networks is strongly related to mechanisms of coordination. For example, denser, more horizontal network structures spend more energy on deliberation, while centralized network structures are more efficient in coordination; but they are also more fragile and dependent on central nodes.

Chapter 7 takes the tools of organizational network analysis to explain the size and structure of executive compensation. Compensation is supposed to provide incentives for executives, such that the executive (the agent) follows the priorities of shareholders (principals), in a complex landscape of decisions of which the principals have limited oversight. Principals are typically concerned with too little risk taking by agents, thus compensation packages include a large portion of a non-salary (equity) element. For the policy-maker the problem is the reverse: how to intervene and shape executive compensation such that excessive risk taking would be curbed. If financial firms are ranked according to CEO compensation in 2006, we observe that only one of the top five survived the financial crisis of 2008. To understand entry points for the policy-maker into shaping executive compensation, one needs to know the mechanisms by which companies benchmark their compensation packages. The strongest elements are the interaction of the peer network (the other companies that a given company chooses to follow) and the mobility network of CEOs. A CEO seduced from outside of the company usually enjoys a substantively higher compensation package, which then sets off cascades in the benchmarking network.

Chapter 8 introduces the notion of power and capture into such organizational fields. Powerful private actors endowed by resources and network centrality can capture public organizations, as in the case of the Commercial Court of Paris. Bankers managed to occupy key positions in this court, that originally aimed to exercise control over them; while finance represents only about five percent of the constituency of the court, almost a third of the judges came from the financial sector.

The explanation of this process of capture is in the network position of bankers: they occupied an increasingly central role in the informal advice network of judges. As the selection of new judges is by invitation from currently serving judges, the representation and centrality of bankers in this organization was increasing steadily.

As Chapter 9 shows, organizational networks can be misused for private gain. Corruption is essentially a parasitic use of organizational network resources; it is inherently a network phenomenon, where trust and covert resource movements are channeled through informal network ties. If we consider network patterns of corruption, organizational networks are integrated with covert personal networks, hidden principals, brokers and quasi-clients. Network embeddedness can be analyzed referring to four types of corruption schemes: bribery, extortion, embezzlement, and fraud. As examples from Hungary show, the smooth operation of these schemes relies on organizational power inequities, lack of transparency, exposure, and vulnerability.

The third part of the volume turns to dynamics at a larger scale, with cases of complex dynamics at the level of whole systems – biological, economic, ecological, and engineered. Chapter 10 brings lessons learned from biological crisis at the cellular level to systemic crises in the current financial meltdown. Interactions in yeast and human cells under stress become more modular, with a decrease in network ties across segments. The exceptions are stress proteins that maintain their diverse and far reaching portfolio of network ties. In the recovery of the cell after crisis, these proteins take the lead, and define the shape of the new network. The lesson for financial systems and social networks in general is that we can expect a decrease in connectedness (fewer transactions, an increased internal focus on organizations), and that those who can maintain a broader network during the crisis, will become key players as the system bounces back and reconstructs itself.

Chapter 11 takes a broad historical perspective on globally interconnected economic networks. Starting from early historical examples of global trade network systems (the silk road from the tenth to the twelfth century, Champagne fairs from the thirteenth to the fourteenth century, the silk road from the thirteenth to the fourteenth century, and East India trading in the seventeenth century), a mode of operation for global economic networks based on a code of transparency is outlined. The dynamics of the global system becomes markedly different when a new code of operation emerges based on guardian ethics. The global system with European dominance that emerged from the mid fifteenth century – based on financial capital backed by military power – produced hyper-inequalities, and a dynamics with deeper crises. Policy lessons are gathered from the alternative logics of earlier global systems.

Chapter 12 introduces network science tools to understand the functioning of ecological systems. As the stress from human activities on natural habitats is increasing