THE DEVELOPMENT OF SOCIAL NETWORK ANALYSIS

A STUDY IN THE SOCIOLOGY OF SCIENCE

Linton C. Freeman

ΣP
Empirical Press
Vancouver, BC Canada

Contents

Chapter 2

Prehistory: The Origins of Social Network Ideas and Practices

In the first chapter I defined social network analysis as an approach to social research that displays four features: a structural intuition, systematic relational data, graphic images and mathematical or computational models. In this chapter I will consider each of those four features in turn and try to specify its first use in social research. My aim here is to describe the earliest research I have been able to find that embodied each feature.

Much of the research described here will include work that is often cited—work done by people who are usually credited with having paved the way for social network analysis. Other research I will describe, however, was produced by people who are relatively unknown in the network field, people who are not usually recognized as having introduced an idea or practice, but who did innovate in ways that provide precursors of current network research practices. These are people whose contributions have not been—but perhaps should be—acknowledged.

2.1 Early Structural Intuitions

In a previous report (Freeman, 1989, p. 18), I reproduced a descent list from the Book of Genesis. My aim there was to suggest that human beings have, since the earliest days, recognized the importance of the ties that link social actors. But that early recognition was implicit.

The earliest explicit statement of a structural perspective on social life that I have found was proposed by Isidore Auguste Marie François Xavier Comte (usually known simply as Auguste Comte). Comte is pictured in Figure 2.1. Although he is usually not mentioned in reviews of the history of social network analysis, I suspect that he had a large, albeit indirect, influence on the development of the field.

Comte was born in 1798 in Montpellier in the south of France. He was the eldest son in a middle class family. He entered the *lycée*—the equivalent of high school—at the age of nine, and studied mathematics with Daniel Encontre. Comte was always an outstanding student, but from his earliest days in school he also established himself as a something of a rebel.

Figure 2.1. Auguste Comte honored on a French postage stamp

Despite his rebellious tendencies, Comte scored fourth among 300 applicants and was admitted to France's top technical college, the *Ecole Polytechnique*, when he was sixteen (Boiteux, 1958). There he continued to study mathematics and he continued to excel. But this was a period of political unrest in France and as a result, the whole *ecole* was closed two years later in 1816. For a year Comte supported himself by tutoring. Then in 1817, at the age of nineteen, he met and went to work with Henri Saint-Simon.

Saint-Simon was then sixty years old. He was an established utopian thinker and is generally acknowledged as the founder of French socialism. According to Coser (1977, p. 15), Saint-Simon was "creative" and "fertile," but at the same time was a "disorderly and tumultuous man." Comte brought order and discipline into the partnership and the two worked and published together for seven years. During most of that period Comte lived in the VIth arrondissement in Paris. His

Figure 2.2. Plaque at Comte's Paris residence[3]

regard in France is demonstrated by the fact that his residence is still memorialized with the marble plaque shown in Figure 2.2.

Coser (pp. 1–17) described the collaboration between Saint-Simon and Comte:

> *A number of scholars have argued the question of who benefited the most from the close collaboration, Comte or Saint-Simon. There is no need to take sides in this somewhat byzantine quarrel. It suffices to say that Comte was influenced in a major way by his patron, even though his close contact with Saint-Simon may have brought to fruition ideas that had already germinated in Comte's mind. . . . The sketches and essays that Comte wrote during the years of close association with Saint-Simon, especially between 1819 and 1824, contain the nucleus of all his later major ideas.*

[3]In English, this plaque reads, "Auguste Comte, founder of positivism, lived here from 1818 until 1822. Here he conceived the sociological law of three states and formulated the system of classification of the sciences."

Comte's primary commitment was to the development of sociology as a science. In that respect his views are quite contemporary, although they were published first between 1830 and 1842 (Comte, 1830–42).[4] Comte coined the term *sociology* and he specified its goal as uncovering the laws of society. This goal, he argued, required both theory and systematic observation. On the importance of theory he said (Martineau, 1853/2000, v. II, p. 203) "No real observation of any kind of phenomena is possible, except in as far as it is first directed, and finally interpreted, by some theory." And about observation, he indicated that the study of social phenomena (Martineau, 1853/2000, v. II, p. 181) "…supposes first, that we have abandoned the region of metaphysical idealities, to assume the ground of observed realities by a systematic subordination of imagination to observation."

Comte argued also for the importance of systematic comparative research in sociology. He (Martineau, 1853/2000, v. II, p. 207) went so far as to suggest comparing the social structures of human beings with those of non-human animals, "It is a very irrational disdain which makes us object to all comparison between human society and the social state of the lower animals." This view is consistent with some of the latest work in social network analysis (Faust and Skvoretz, 2002).

In his definition of sociology, Comte specified the two main aspects of the field, *statics* and *dynamics*. Statics, he said (Martineau, 1853/2000, v. II, pp. 192, 204 and 207), is focused on the investigation of the "laws of social interconnection" or (v. II, p. 183) "the laws of action and reaction of the different parts of the social system." At the most fundamental level, according to Comte, these parts are nuclear families. He argued (Martineau, 1853/2000, v. II, p. 234) that since

> *…every system must be composed of elements of the same nature with itself, the scientific spirit forbids us to regard society as composed of individuals. The true*

[4]In the present work I have drawn heavily on a new edition of the translation of Comte's *Cours de philosophie positive* by Harriet Martineau (1853/2000). Comte liked Martineau's translation of his work so much that he substituted Martineau's two-volume version for his own six-volume version in his list of books that he believed should survive forever (Standley, 1981, p. 160).

> *social limit is certainly the family—reduced, if necessary, to the elementary couple which forms its basis.*

Comte (Martineau, 1853/2000, v. II, p. 187) went on to show how the parts of the social system are interconnected: "Families become tribes and tribes become nations." His view, then, is strikingly contemporary. He defined society using the kinds of structural terms that are found today in social network analysis. Whether he is given credit or not, Comte was the first scholar I could find that proposed a way of looking at society in terms of the interconnections among social actors.

Most other prominent nineteenth and early twentieth century sociologists embraced Comte's structural perspective. A common theme involved describing differences in the patterning of social connections in traditional versus modern societies. In his study of ancient law, for example, Sir Henry Maine (1861/1931) did just that. He suggested that in small traditional, family-oriented societies, individual ties were governed by universal rights and obligations; these arrangements he called *status*. In contrast, he saw most ties in large modern societies as grounded on negotiated agreements; in his terms, they are based on *contract*.

Ferdinand Tönnies (1855/1936) made a similar distinction when he used the word *gemeinschaft* to characterize the traditional social form that involved personal and direct social ties that linked individuals who shared values and beliefs. He contrasted that with modern forms based on *gesellschaft*, where social links are formal, impersonal and instrumental.

Emile Durkheim (1893/1964) described traditional societies in which *solidarité mechanique* linked similar individuals with repressive regulations. This he distinguished from modern societies in which a division of labor led individuals to form cooperative links based on *solidarité organique*.

Sir Herbert Spencer (1897) in England and Charles Horton Cooley (1909/1962) in America both described traditional small-scale societies in which individuals were linked by intimate, *primary* relations. And they both contrasted those with modern, large-scale societies where individuals are often linked by impersonal, *secondary* relations.

These early sociologists all tried to specify the different kinds of social ties that link individuals in different forms of social collectivities. Thus, since they were all concerned with social linkages, they all shared a structural perspective.

An entirely different structural perspective was developed at about the same time by Gustave LeBon (1897/1995). LeBon examined the phenomenon of crowd behavior. He suggested that when individuals become members of crowds they lose their individual identities. As members of crowds, people imitate those around them, he said, and ideas and behaviors diffuse from person to person by a process of contagion. Thus, since LeBon focused on the flow of information among individuals, his concern was also structural.

Figure 2.3. Georg Simmel

Perhaps the most explicitly structural perspective adopted by any of the late nineteenth and early twentieth century social thinkers was displayed in the work of Georg Simmel (pictured in Figure 2.3). Simmel (1908/1971, p. 23) said, "Society exists where a number of individuals enter into interaction." He (pp. 24–25) went on to specify this idea:

> *A collection of human beings does not become a society because each of them has an objectively determined or subjectively impelling life-content. It becomes a society only when the vitality of these contents attains a form of reciprocal influence; only when one individual has an effect, immediate or mediate upon another, is mere spatial aggregation or temporal succession transformed into society. If, therefore, there is to be a science whose subject matter is society and nothing else, it must exclusively investigate these interactions, these kinds and forms of sociation.*

In these statements Simmel expressed the core belief that underlies modern social network analysis. For Simmel, sociology was no more and no less than the study of the patterning of interaction. And Simmel's student, Leopold von Wiese (Wiese and Mueller, 1931/1941, p. 30), went even further and talked in contemporary terms about a "system of relations" and a "network of lines between men." The views expressed by Simmel and his students, then, were—and remain—explicit statements of the social network perspective.

Overall, then, there were a good many early writers—particularly in the field of sociology—who laid out the intuitive groundwork for network analysis. Early contributions, however, were not limited to intuitive ideas. In the next section I will show some early work that involved the systematic collection of structural data.

2.2 Systematic Empirical Data

At the beginning of the nineteenth century, long before Comte defined a structural approach to sociology, systematic data on social structure were being collected by a Swiss naturalist, Pierre Huber. Huber was born in Geneva, Switzerland, in 1777. His father, François Huber, was an entomologist who was deeply involved in the study of honeybees. The elder Huber became blind in 1773 at the age of twenty-three. Nevertheless, he continued his research by drawing on the observational skills of his wife, Marie Lullin, and his servant, François Burnens. With their help as observers, Huber published a major work based on systematic observation of honeybees (Huber and Bonnet, 1792). That publication had enough impact that he was thereafter known as "Huber des abeilles (Huber of the bees)." Even now his contributions are honored with a walkway named "François Huber" off the rue de Saint-Victor in Geneva.[5]

Young Pierre was raised in this household—one that was almost totally focused on the systematic observation of insect behavior. It is no wonder then that he went on to a distinguished career in entomology. At twenty-four, he published a detailed

[5]There is even a recently published historical novel, *The Beekeeper's Pupil* by Sara George (2003), about the life of François Huber.

description of bumblebees (1802). He observed and reported every aspect of their anatomy and life cycle in painstaking detail. And included in that description was a thorough report of their dominance behavior with respect to one another. This, so far as I have been able to discover, provides the earliest example of a report of patterned social interaction based on systematic observation.

Pierre Huber went on to work with ants (1810) and achieved world fame as a naturalist. In fact, Charles Darwin (1859, Chapter VIII) described him "as a better observer even than his celebrated father." The younger Huber's work set the stage for the development of the ethological approach in biology. At the same time, it was a precursor of two developments that have become parts of contemporary social network analysis. First, studies of social patterning among nonhuman animals continues in social science. It was picked up first by Moreno's sociometric community in the 1940s (see the special issue of *Sociometry*, Volume VIII, Number 1, February 1945). Moreover, recent research in network analysis still includes observation-based studies of the patterning of social linkages among nonhumans. And second, network analysts still conduct systematic studies of dominance. Huber's work, then, provided a model for later research in both biology and social network analysis.

The earliest example of systematic data collection on humans came a half-century later. The lawyer-anthropologist Lewis Henry Morgan was born near Aurora in upstate New York in 1818. He attended Union College where he joined a secret society called the "Grand Union of the Iroquois." Members wanted to model the form of their organization after the Iroquois, but discovered that little was known about that form.

So several members of the society, including Morgan, went to nearby Iroquois reservations to learn what they could about Iroquois social practices. And soon Morgan found himself deeply involved in the fieldwork. He ended up publishing an ethnography of the Iroquois (1851).

At that point, Morgan left off his anthropological pursuits. He moved to Rochester and began the practice of law. But the 1856 meeting of the American Association for the Advancement of Science took place in Albany, and according to Tooker (1997), that meeting revived Morgan's interest in the field. Morgan presented a paper on the Iroquois descent system at the meeting in 1858. His conclusion was that the Iroquois system of naming kin and reckoning descent was dramatically different from that used by western Europeans.

Later that year, on a trip to Michigan, Morgan interviewed an Ojibwa woman and discovered that the Ojibwa had yet another distinct scheme for naming kin and reckoning descent. For Morgan, these differences provided enough motivation to determine his future career. He began to travel in order to interview representatives of other North American tribes. At the same time, he went to great effort to enlist the help of others in collecting data in other parts of the world. When he had amassed a mammoth collection of data, he quit his law practice and devoted his time to the task of bringing all the data together. The result was a huge volume, *Systems of Consanguinity & Affinity of the Human Family* (1871/1997).

In that book Morgan reported the terms used in describing lineages of peoples throughout the world. He presented data and proposed that kinship terminologies embodied various patterns and were correlated with forms of marriage and rules of descent.

A century later, John Atkinson Hobson (1894/1954) developed an approach to uncovering links among organizations. Hobson was born in 1858 in Derbyshire, England. He earned an M.A. degree at Lincoln College, Oxford. He taught literature in Faversham and Exeter until 1887 when he moved to London. There, he joined the Fabian Society and found a job as a reporter for the Manchester Guardian newspaper.

By the time he arrived in London, Hobson had already established a reputation for his liberal views on economics. He had published books on poverty and unemployment, and he had produced a major work on capitalism—one that was often cited by Lenin (Hobson 1894/1954). As part of his treatment of capitalism, Hobson (p. 271) presented systematic data on overlapping directorships among members of "the small inner ring of South African finance." His table is reproduced as Figure 2.4.

	De Beers.	Premier.	Rand Mines.	Goldfields.	Chartered Company.
Beit, A.	1	—	1	—	1
Jameson, L. S. ...	1	—	—	—	1
Maguire, R. ...	—	—	—	1	1
Michell, Sir L. L.	1	—	—	—	1
Neumann, S. ...	—	1	1	—	—
Wernher	1	—	1	—	—

Figure 2.4. Hobson's two mode data

It shows the five major South African companies and the six board members that linked them.[6]

Thus, Hobson innovated by presenting two-mode (board member by company) data that reveal links of individuals to companies. At the same time, these data display links among the companies in terms of their shared board members as well as links among the individuals who are co-members of the same boards. Hobson, then, was the first investigator to collect systematic data on corporate interlocks. Both of these innovations are still major parts of contemporary social network analysis (Levine, 1972; Stokman, Ziegler and Scott, 1985).

The next several contributions all came from developmental and educational psychologists who worked in the 1920s. During those years, large long-term grants were made to the child-welfare institutes at the Universities of Iowa and Minnesota, Yale, Columbia Teachers College and the University of California, Berkeley (Renshaw, 1981). One result of this funding was a huge increase in the amount of research focused on children, and a good deal of that new research was centered on the study of children's interpersonal relationships.

These works are for the most part unrecognized in the field of social network research, but they did succeed in innovating a

[6]Fennema and Schijf (1978/79) credited Otto Jeidels (1905) with "the first extensive and systematic study on corporate interlocks." They did mention that Hobson's research was earlier, but they described it as "very partial." But since Jeidels opened his report with a quotation from Hobson, the direct line of influence is clear.

number of important network ideas and practices. In a study of homophily among school children, for example, John C. Almack (1922) developed a way of using interviews to collect network data. He asked children in a class to name those they would like to invite to a party. Then he examined the data in terms of the similarities between choosers and those chosen. Thus, Almack anticipated by more than ten years the data collection procedure usually credited to Moreno. Questions of this sort continue to provide a standard method for collecting data in network analysis. And, by examining the similarities between the choosers and the chosen, Almack anticipated the very keen interest of contemporary network analysts in homophily as a basis for social choice.

Beth Wellman[7] (1926) collected network data by recording systematic observations of who played with whom among preschool children during periods of free play. She pioneered, then, in extending Huber's ethological approach to the study of human interaction. In network analysis involving human subjects, data generated by questioning actors are common, but observational data of social links are still gathered.

Helen Bott[8] (1928) went even further, refining Beth Wellman's approach in several ways. First, she used ethnographic methods to uncover the various forms of interaction that occurred regularly among preschool children. These methods enabled her to limit systematic observations to the specific kinds of interaction relevant in the context of her research. Second, she was the first to employ a focal-child method for collecting detailed observations of who displayed each particular form of interaction with whom. Thus, she was able to avoid many of the biases inherent in other methods of observing interaction (Altmann, 1973). And third—eighteen years before Forsyth and Katz (1946) suggested using matrices to record interaction patterns—Bott recorded her data in matrix form. Although Bott apparently had little direct influence on the development of social network re-

[7]Beth Wellman was not a relative of Barry Wellman, who is treated in later chapters.

[8]Helen Bott was the mother of Elizabeth Bott, a major figure in social network analysis. Elizabeth was one of the experimental subjects in Helen Bott's research. Interestingly, Elizabeth Bott denies that her mother's work had any influence on her involvement with the social network perspective.

search, and is seldom if ever cited by network analysts, she certainly anticipated many modern sophisticated methods of data collection and presentation.[9]

In 1933 Elizabeth Hagman brought these two approaches to data collection—observation and interview—together. She explicitly raised a data-related issue that is still a central concern among contemporary network analysts. She observed which children played with which others during a period of free play. Then, at the end of the school term, she interviewed each child. They were asked to name who their playmates had been at the beginning of the term, in the middle of the term and at the end of the term. She then compared the observed data with the reports and examined the discrepancy between the two. that the discrepancy she found defined a research problem that remains a key issue to a great many recent investigators in the social network field (Bernard, Killworth, Kronenfeld and Sailer, 1985; Freeman, Romney and Freeman, 1987).

2.3 Graphic Imagery

Graphic images have had an important place in structural studies from the earliest days (Freeman, 2000b). The earliest examples were all focused on kinship. Christine Klapisch-Zuber (2000) showed that tree-based images were drawn as early as the ninth century. Those early images, designed to display the general patterning of kinship, showed pictorially the proximity between the occupants of any two kin categories. Klapisch-Zuber (2000, p. 37) went on to report that by the thirteenth century, drawings of trees were commonly used to depict the lineages of particular families. So the application of imagery in examining kinship began very early.

The data on kinship collected by Lewis Henry Morgan were described above. However, Morgan not only collected a huge amount of data on kinship terminology, he drew diagrams to try to specify the positions of equivalent relatives. His drawing

[9]Bott's contributions are recognized by developmental psychologists (Renshaw, 1971), but until her work was rediscovered by Freeman and Wellman (1995), the only citation from any social network analyst to Helen Bott that I could locate was made by Eliot D. Chapple (1940).

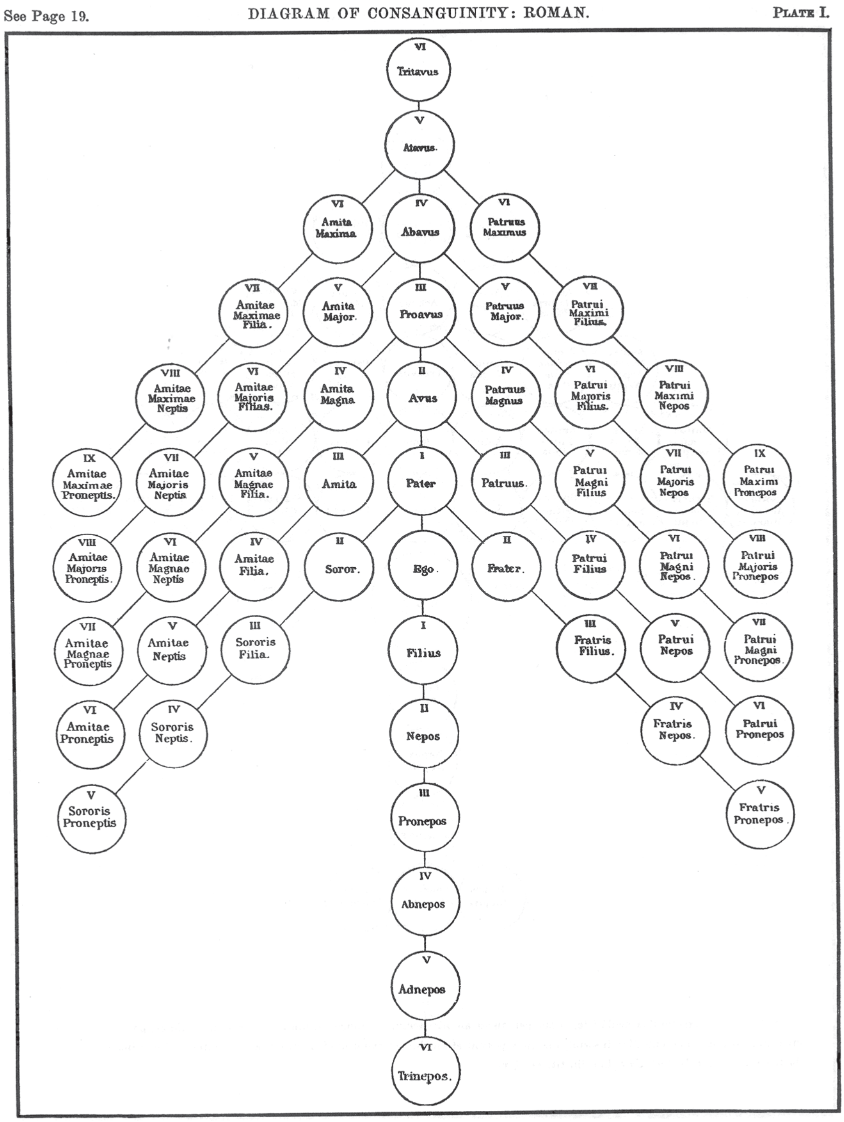

Figure 2.5. Morgan's descent system of ancient Rome

of the system of descent in ancient Rome is shown in Figure 2.5. In it, positions are drawn as circles and descent is shown by lines. Clearly, this image is an ancestor of contemporary pictures of social networks.

Near the end of the nineteenth century, Alexander Macfarlane (1883) developed a formal model of the British kinship system. Macfarlane was born in Ontario, Canada, in 1851. He received a DSc degree from the University of Edinburgh, and stayed on as a Lecturer in Mathematics there. In 1878 he became a Fellow of the Royal Society of Edinburgh, but left Scotland in 1893 to become a Professor of Physics at Texas University (now the University of Texas).

Macfarlane was an algebraist, and his work on kinship was primarily algebraic, but a part of that work involved constructing visual representations of various degrees of kinship. In the illustration shown in Figure 2.6, Macfarlane specified all the two-step marriage relationships that are prohibited in English law. From left to right, the picture shows that the law prohibits a male (designated by a +) from producing offspring with his granddaughter, his sister or his grandmother. Put another way, it prohibits a woman (○) from producing a child with her grandfather, her brother or her grandson. In the figure a short line

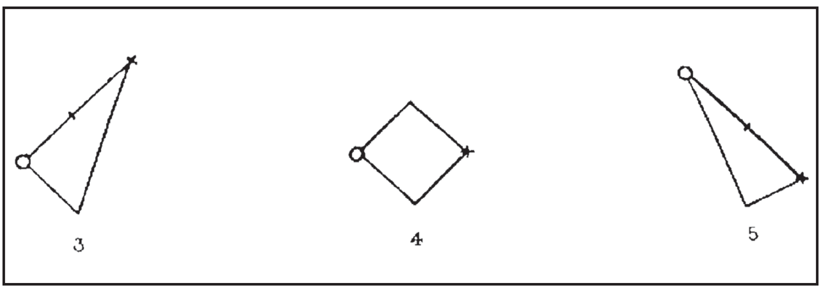

Figure 2.6. Macfarlane's images of two-step marriage prohibitions

crossing a longer one simply designates a person of either sex in another generation.

Both of these early attempts produced images of graph-like structures. They both used points to represent social actors and lines to represent linkages among them. In both cases, however, the images as they were drawn were not quite graphs. The whole meaning in a graph is in its structure—which point shares an edge with which other point. So when they are drawn, graphs place points in locations that are arbitrary; their relative positions—above, below, left or right—have no meaning. But both Morgan and Macfarlane drew images that were oriented top to bottom. Both placed genealogical antecedents higher on the page and descendants lower. Nonetheless, their images anticipated the pictures produced by network analysts using graph theory.

Hobson's (1894/1954) data on interlocking directorates was described in Section 2.2 above. But Hobson not only collected a new kind of data, he went on to introduce a way to display two-mode network data. He drew on the data shown in Figure 2.4 along with data from another table in his book and produced a picture of overlaps, a hypergraph, to show his readers how a small number of large corporations—notably the De Beers Group and Rand Mines—could use interlocking directorates to control many other firms. That hypergraph is shown in Figure 2.7. As far as I can discover, it is the earliest example of an image of

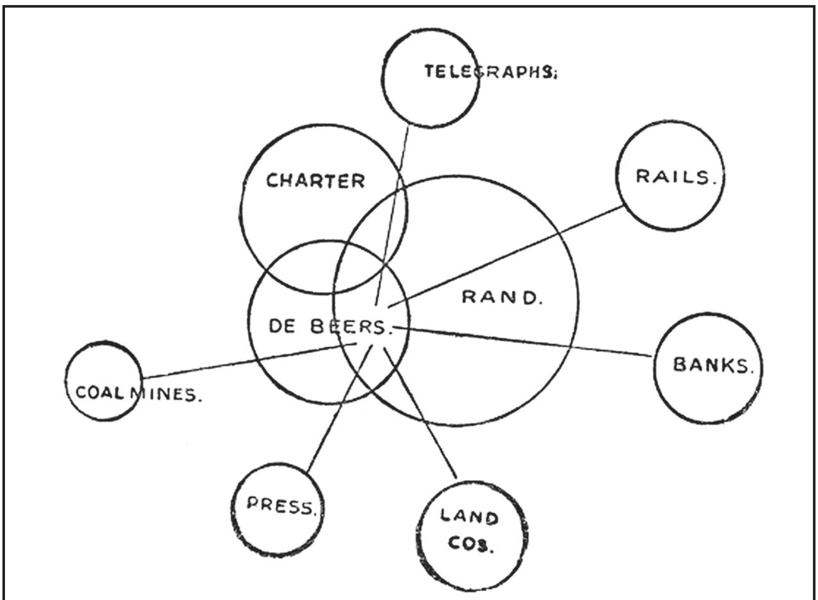

Figure 2.7. Hobson's hypergraph

overlaps used to display social patterning. Images of this sort are still widely used in social network analysis.

2.4 Mathematical and Computational Models

Unlike many other approaches to social research, network analysis has consistently drawn on various branches of mathematics both to clarify its concepts and to spell out their consequences in precise terms. The earliest application of mathematics to a problem of social structure that I have been able to uncover was algebraic.

Figure 2.8. Sir Francis Galton, on the right, 1909, with Karl Pearson[10]

Macfarlane was mentioned above in the discussion of early graphic tools. Those graphics, however, were secondary to his main contribution. Macfarlane (1883) worked out an explicit way of characterizing the English system of kinship by concatenating kin terms. He began with a simple symbolic system:

> *There are two fundamental relationships of the highest generality, namely,* child *and* parent, *the one relationship being the reciprocal of the other. These can be combined so as to express any of the complex relationships; thus grandchild is expressed as child of child…let* c *be used to denote child,* p *to denote parent, and let "of" be expressed by juxtaposition, then grandchild will be expressed by* cc.

[10]Pearson (1914, 1924, 1930), the famous statistician, wrote a three-volume biography of Galton.

Macfarlane made his system more complex and more interesting by adding a notation for gender, *m* and *f*. Then he went on to designate the incest prohibitions specified under English law using his system of symbols and drawings. An example was shown in Figure 2.6.

At that time, one of the leading figures in the British scientific community was Sir Francis Galton, pictured in Figure 2.8. Galton worked as an African explorer, a meteorologist, a geographer, an anthropometrist, a statistician and a geneticist. Like many of his nineteenth century colleagues, he was something of a racist. His belief that almost all human characteristics are inherited led him to develop the controversial notion of eugenics. As a consequence, he is less respected in the current world than he might otherwise have been.

Galton was born in 1822 to a well-to-do Quaker family in Birmingham, England. He was a gifted child and, following his mother's desire, he began the study of medicine as an apprentice at the age of sixteen. He attended medical school at Kings College and Trinity College, Cambridge.

Galton transferred his medical studies to London, but when his father died in 1844, he dropped out of school and used his inheritance to become a gentleman-explorer. He explored in Southwestern Africa, and his successes promoted his recognition by the members of the scientific community. In 1856, when he was thirty-four, he was elected a Fellow of the British Royal Academy.

The most profound influence in his life occurred in 1859 when his cousin, Charles Darwin, published the *Origin of the Species*. As Galton later recalled:

> *The publication in 1859 of the* Origin of Species *by Charles Darwin made a marked epoch in my own mental development, as it did in that of human thought generally. Its effect was to demolish a multitude of dogmatic barriers by a single stroke, and to arouse a spirit of rebellion against all ancient authorities whose positive and unauthenticated statements were contradicted by modern science.*

From then on, Galton's major pursuit involved the study of inheritance, which led indirectly to his contributions to the development of tools for social network analysis. It led, for example, to a concern with kinship. So when Macfarlane presented his paper at a meeting of the Royal Anthropological Institute in 1883, Galton was present. And the minutes of the meeting show that Galton was the major discussant of Macfarlane's paper. The minutes paraphrased Galton's remarks:

> *Dr. Macfarlane had attacked the problem of relationship with thoroughness, ability, and success, and that he had done a very acceptable work for all who concerned themselves with genealogies of the complicated descriptions referred to by Dr. Macfarlane. The diagrammatic form seemed to himself the most distinctive and self-explanatory.*

So Galton was involved in Macfarlane's presentation, but soon thereafter he made his own contribution. He participated in developing an early probability-based model in the social network area. Galton had been working on the problem of hereditary genius, and had noticed that families that were prominent at one time seemed to decline—even disappear—at later times. He had enough mathematical training to suspect that such disappearances were the result of some probabilistic process, but he was not a good enough mathematician to spell out the process.

Consequently Galton enlisted the help of the Reverend Henry William Watson, who was both a minister and a statistician. Galton and Watson (1875) took a stochastic approach to the study of networks. Galton put the problem as follows:

> *Let $\boldsymbol{p}_0, \boldsymbol{p}_1, \boldsymbol{p}_2, \ldots$ be the respective probabilities that a man has 0, 1, 2,…sons, let each son have the same probability for sons of his own and so on. What is the probability that the male line is extinct after **r** generations?*

Drawing on probability theory, Watson built a stochastic model of the disappearance of family names—a network process.

His model was essentially computational; his solution involved assigning values to their basic parameters and deriving a numerical solution based on those parameters.

Watson began with the notion of a population of family names. He proposed a set of parameters involving the probabilities of a given man producing 0, 1, 2,…, q male offspring. From these properties, Watson calculated the expected proportions of each surname in each succeeding generation. Of course, any surname holder who produced no male offspring would contribute to the reduction of representatives of his name in succeeding generations. So Watson was able to show that, simply by a random process of reproduction, "We have a continual extinction of surnames going on." His conclusion was that any family name would ultimately disappear with a probability of 1.

Among network analysts, Galton and Watson are traditionally credited with having been the first to solve the problem of the disappearance of family names (Mullins and Mullins, 1973, p. 257). But, as Kendall (1974) documented, that credit is misplaced. There were two important limitations of the Galton and Watson work. First it was incomplete, and second it was actually preceded by an earlier work that had produced a more complete answer.

In 1845, a French demographer and mathematician, Irénée Jules Bienaymé, had already examined the same problem (Heyde and Seneta, 1977, p. vii). Bienaymé's work was motivated by reports by Doubleday (1842) and de Châteauneuf (1845) both of whom had observed that the surnames of noble families tend to disappear over time. Bienaymé developed a model designed to explain that observation. He said (translation by Heyde and Seneta, 1977, pp. 117–118):

> *If…the mean of the number of male children who replace the number of males of the preceding generation were less than unity, it would be easily realized that families are dying out due to the disappearance of the members of which they are composed. However, the analysis shows further that when this mean is equal to unity families tend to disappear, although less rapidly…The analysis also shows clearly that if the mean*

> *ratio is greater than unity, the probability of the extinction of families with the passing of time no longer reduces to certainty. It only approaches a finite limit, which is fairly simple to calculate and which has the singular characteristic of being given by one of the roots of the equation (in which the number of generations is made infinite) which is not relevant to the question when the mean ratio is less than unity.*

This last point illustrates the incompleteness of the later Galton and Watson work. Watson dealt only with the case in which the mean number of male children was equal to one. His conclusion that names would always disappear is true only under that condition or when the number is less than one.

It is interesting to speculate about why the work of Bienaymé was lost and why Galton and Watson have traditionally been given credit for solving the problem. The answer, I think, lies in the relative prominence of Bienaymé and Galton. Bienaymé spent most of his career working as a French civil servant. He did publish mathematical papers, but relatively infrequently; there were thirty-four in all. Moreover, it was not until late in his career that he received any academic recognition at all: in 1848 he was awarded a temporary appointment as Professor of the Calculus of Probabilities at the Sorbonne, and in 1852 he was elected to the *Académie des Sciences de Paris* (Heyde and Seneta, 1977, p. 7).

In contrast, Galton came from a prominent family and was a major figure in the British Academy from the time he was in his early thirties. Moreover, he produced, "…considerably more than a hundred and seventy publications…" (Galton, 1908, Chapter XX). It is clear, then, that Galton's relative prominence put him in a much better position to be recognized and remembered. Some justice is evident, however, in the fact that in current works by probabilists the problem of the disappearance of lineages is called the "Bienaymé-Galton-Watson process."

The materials introduced in this chapter are all early examples of important innovations in each of the four features that characterize contemporary network analysis. Various early writers (1) produced structural intuitions, (2) collected systematic who-to-whom data, (3) produced graphic images of structural forms and (4) developed both mathematical and computational models.

Some, like Morgan (1871), Macfarlane (1883) and Hobson (1894), produced work that embodied two of the features. Morgan collected huge amounts of systematic data on kinship and displayed his results in graphic images. Macfarlane developed an algebraic model of kinship and he too used graphic images to display its properties. Hobson collected systematic data on corporate interlocks, then drew hypergraphs to reveal their observed interlock patterns. By employing two of the four tools that define modern social network analysis, these nineteenth century investigators began to approach current practice.

From a network perspective, this chapter has raised some interesting issues. First, of all the people discussed here, only one, Georg Simmel, is generally recognized as having been influential in forming a foundation for social network analysis. Maine, Tönnies, Durkheim, Spencer, Cooley and von Wiese have been mentioned in the network literature, but rarely. There has been at least one citation to Hobson and one to Galton and Watson. This suggests that both works may have had some impact on the development of the field. But, as far as I can discover, even though their work could be expected to have influenced the development of the field, all the others I have discussed in this chapter are simply unknown in the social network community.

In any case, in the early 1930s a broad research effort called sociometry was introduced. It was the first work that included all four of the defining features of social network analysis. In the next chapter I will review its development.