The Icelandic sagas are thousand-year-old stories of brave deeds and lives and loves, with a cast of thousands – but how true to life are they? Pádraig Mac Carron and Ralph Kenna analyse the interactions between the characters and find that social networks of the Viking era were very similar to those of today.

The *Íslendinga sögur* – or sagas of Icelanders – are a collection of prose texts purporting to describe events in the period following the settling of Iceland in the late ninth to the early eleventh centuries. They tell of families and feuds, of warfare and the deeds of warriors, of their loves, lives, betrothals and separations, of their betrayals and vengeances. Despite their antiquity they are still a living part of Icelandic culture: it is said that if you take a taxi in Iceland today, the driver will be able to quote you long sections from the sagas. They are widely available in translation and still provide very good reading.

The antiquity of the texts and their unique narrative style make them an important element of world literature. As with other ancient narratives, their historicity has long been the subject of scholarly debate – some consider them to contain information on Viking life and society, while others object that such tales are entirely fictional, with no basis in reality and no historical value. It is believed that the *Íslendinga sögur* were committed to writing in the thirteenth and fourteenth centuries, but their authors are unknown or uncertain. The stories describe daily struggles and conflicts in plausible detail and the narratives and characters are interwoven, with overlapping plots in different texts involving thousands of characters and their interactions. This huge network of interactions makes them an ideal study for statistical analysis. Studies of literature traditionally focus on individuals and events. Our new statistical approach looks instead at the collections of interactions between characters. It gives a way to make quantitative comparisons within the sagas, and between the sagas and other literary genres. We can try to answer questions about how much they are based on fact or on fiction, and to uncover what quantitative information may have been hidden for centuries within the pages of these ancient manuscripts.

The statistical approach follows our previous study of mythological and epic literature, in which we compared the Anglo-Saxon epic *Beowulf*, Homer’s *Iliad* and an Irish epic, the *Táin Bó Cúalnge*. Most historians consider the first two to be partly based on real events (though clearly embellished); the third is

---

1. December 2013

---
works have been identified and classified on the comparisons; we wanted statistics to inform about the remote past.

There is a huge amount of literary scholarship important part of our shared cultural heritage. Did fictional ones? We also wanted to compare most resemble real-life social networks, or might give some sort of indication. Would it between the huge cast of characters in the stories invented or perhaps based on a real society? Icelandic sagas, to investigate whether they are based epics.

Much more similar to those of the historically-adjusted, the remainder of the network was their too-good-to-be-true interactions were exaggerated them. When the few Viking helmets which survive have horns.) (Beards, on the other hand, probably were worn).

generally thought fictional. We looked at the social networks which underpin each of them and compared them to each other, as well as to real, imaginary and random networks. We found that the apparent artificiality of the Táin network was mainly associated with a few leading characters who are too well connected to be realistic (in the same way that comic-book superheroes all know each other and fight crime together), suggesting that perhaps the tales’ composers had exaggerated them. When their too-good-to-be-true interactions were adjusted, the remainder of the network was much more similar to those of the historically-based epics.

Could we do something similar with the Icelandic sagas, to investigate whether they are invented or perhaps based on a real society? We hoped the network of associations between the huge cast of characters in the stories might give some sort of indication. Would it most resemble real-life social networks, or fictional ones? We also wanted to compare the sagas to each other and to other social networks, whether historically accurate or not, to give a unique statistical insight into an important part of our shared cultural heritage. There is a huge amount of literary scholarship on the comparisons; we wanted statistics to contribute to that too.

In recent years, different types of networks have been identified and classified according to their properties. Studies have shown that social networks, in particular, tend to have distinguishing features. For example, they are small-world and scale-free, they tend to be assortatively mixed and exhibit community structures. In the next section we explain these terms before reporting on how we applied them to medieval literature in a quest for information about the remote past.

Network theory

A network is a collection of nodes connected by links. In a social network, the nodes represent people – or characters in the case of literature – and the links are relationships or interactions between them. The most fundamental quantity in a network is the degree, which is the number of links associated with a given node. For example, on Facebook your degree is simply the number of “friends” you have. In Figure 1, character A has degree three.

The probability that nodes in a network have specific degrees is given by the degree distribution. In regular networks, – a square grid, for example – every (interior) node has the same number of links. In other networks most nodes might have a small number of links, but a few highly-connected nodes might have many – just as, in real life, most tweeters tend to mingle with other popular people. Networks for which the degree distribution follows a power law are called scale-free, implying that only a few nodes tend to form links with a large number of other nodes.

One statistical measure of the extent to which a network is connected is its mean path length. This is simply the average number of steps needed to link pairs of individuals in the network. For example, in Figure 1 the shortest path (or geodesic) linking G and D has only two steps (namely G → H → D). F and D, on the other hand, need at least four steps to link them. The average length of all of the geodesics in the entire network is its mean path length. This notion was famously demonstrated in sociology through the concept of six degrees of separation: despite the world’s population of over 7 billion people, it is claimed that everyone is, on average, only about six steps away from everyone else. It is also famous as the Hollywood version known as the “Kevin Bacon number”: nearly every movie actor, it is said, is no more than a few steps from having worked in a film with Kevin Bacon. Even mathematicians are not immune: their version is the Erdős number.

The clustering coefficient measures the extent to which a network is cliqued. In real social networks, if an individual knows two others, there tends to be a high probability that the two are also acquainted – classmates in school are an example. In Figure 1 node A has 3 neighbours (B, G and H). Of the three potential relationship triads (ABG, ABH and AGH), only one (AGH) is realised. So the clustering coefficient for node A is 1/3. The clustering coefficient for the entire network is determined simply by averaging over all nodes. A network is called small-world if it has a small path length and a high clustering coefficient compared to a random graph of the same size and average degree. In small worlds, everyone is close to everyone else.

In modern society, people tend to form friendships with people who are similar to themselves: popular people (e.g., celebrities) tend to mingle with other popular people. Here, similarity of nodes is determined by comparing their degrees. Networks which tend to have links between nodes of similar degrees are called assortative. The opposite
feature is termed disassortativity (see Figure 2). Assortativity is an important property which helps distinguish social networks from other networks, most of which are disassortative. Assortativity can be quantified by statisticians as the Pearson correlation coefficient of the degrees between all pairs of linked nodes. All that non-statisticians need to know is that it comes out as a number between −1 and +1. A positive correlation indicates assortativity and a negative value means the network is disassortative.

Social networks often have community structures, and one way to measure this is through the modularity, $Q$. This is determined by a community-detection algorithm which seeks to split up densely connected clusters (see Figure 3). If the structure comprises a single community, $Q$ is close to 0. If it is evenly partitioned into a large number of sparsely interconnected communities, $Q$ approaches 1.

There have been many analyses of real-world social networks in recent years – examples include Hollywood actors, company directors and scientific co-authors. Such social networks are small-world and usually have power-law degree distributions distinguishing them from random and regular networks. They are also assortative (“birds of a feather flock together”) and tend to have a high modularity. While each of these characteristics is not unique to social networks, they are all commonly found in them and are therefore characteristic of them. The three epics of our first analysis, Beowulf, the Iliad, and the Táin, exhibit some or all of these properties to varying degrees.

The question we wish to ask is: how do these properties measure up in the Icelandic sagas? Do the ancient tales perhaps hold information about Viking society? Do they look like real social networks? How do the different characters meet in conflict. Characters can be connected by both friendly and hostile links (for example, if one changes sides).

We gathered data for 18 narratives, five of which contain over 100 characters each. These are Njál’s saga, Laxdæla saga, Vatnsdæla saga, Egils saga Skallagrímssonar and Gísla saga Súrssonar. We examined these five individually in order to compare different sagas to each other. We also studied the sagas collectively – a network of 1549 individual characters – to gain insight into the structure of the overall saga society.

A notable feature of the Íslendinga sögur is their objective narrative style, their consistency, and their presentation as chronicles, with the main characters in one text appearing as minor ones in another. The sagas were initially widely accepted as giving more or less accurate accounts of early Icelandic society. In recent times, however, historians have been more critical. Some modern historians dismiss the sagas as pure fiction. But even if the events are fictional, it is conceivable that they may play out against a backdrop which includes real history. This is why we chose to apply a statistical approach using the tools of network theory outlined in the previous section. We wanted to bring to the fore the societal backdrop instead of events and individuals – to focus on the set

---

**Figure 3.** The network on the left illustrates modularity – the community-detection algorithm can break it down to two communities (easily identifiable by eye here). Its modularity, $Q$, is approximately 0.5. The random network on the right has no community sub-structure and the algorithm is unable to break it down into clusters. It has a modularity of 0.
of relationships between the characters depicted in the texts.

*Njáls saga* has the most surviving manuscripts – around 60, if you include fragments – and is widely regarded as the greatest of the prose literature of the Viking era. It deals with feuds and how minor incidents could escalate into major bloodshed. More than 500 named characters appear in its narrative. The saga of the people of *Laxdæla* has the second highest number of preserved medieval manuscripts and also contains the second largest network, and it is believed that the author of *Njáls saga* may have used this as a source. It focuses on a love triangle which ultimately leads to enmity and death. The saga of the people of *Vatnsdal* is a family chronicle which tells how the grandson of a Norwegian chieftain arrived in Iceland and tracks his family until the arrival of Christianity in the late tenth century. *Egils saga* also begins in Norway where, after a dispute with the king, a poet and his family leave to settle in Iceland. *Gísla saga* is an outlaw narrative in which the eponymous character is “on the run” for 13 years before finally being killed.

Thus there are different types of sagas – ranging from outlaw sagas, mostly centred on one character, to family sagas that cover generations. In the next section, we report upon the network analysis of the *Íslendinga sögur* in an attempt to gather statistical information on interrelationships between characters and to compare saga societies to each other and to other social networks.

**Saga statistics**

The basic properties are summarised in Table 1. Each major network has a similar mean degree of about 5 – each character interacts with about five others on average. In each case the mean path length is comparable to that of a random network with the same number of nodes and the same average degree, but the saga networks are far more clustered than their random counterparts – therefore the saga society is small-world. Each saga network has a giant connected component containing over 97% of the nodes. The very few characters who are disconnected from the main component typically have only one or two links. For each network, no more than 10% of closed triads contain odd numbers of hostile links. This propensity to disfavour odd numbers of hostile links is related to the notion of *structural balance* – in the Viking era, the enemy of an enemy is a friend. This is also illustrated in Figure 1, for example, where green represents friendly links and red represents hostility: the enmity between characters G and H makes it difficult for E to remain on good terms with both of them. In our example, E has chosen to ally with G.

About 20% of the characters in each of the five major sagas are female. However, at 8%, the proportion of female–female interactions in *Laxdæla saga* is about twice that of the others. Indeed, it has been speculated that *Laxdæla saga* may have had a female author.

Are the saga networks scale-free? If so, as we discussed earlier, their degree distribution would follow a power law, with an exponent

![Figure 1: The network underlying Egils saga. Vertices coloured red represent characters who appear in the first part of the tale, set in Norway, while those in blue appear later, during Egil’s time in Iceland.](image)
\( \gamma \), so that \( p(k) \sim k^{-\gamma} \). For the various sagas, we found they are indeed well described by power laws, with \( \gamma \) falling between 2 and 3, as is usually found for real social networks. However, \textit{Laxdæla saga} in particular is better fitted by an exponential distribution, \( p(k) \sim \exp(-\text{const.} \, k) \).

The social networks of the sagas are close to those of real life and of society today

We found that some of the family sagas are assortative and only the outlaw saga is strongly disassortative – Gisla the outlaw interacts with many minor characters. This reflects the fact that the story is centred on a single protagonist’s exploits instead of on a larger society. \textit{Egils saga} has assortativity close to 0, in particular when we just consider the friendly network. This is a saga about a poet set in two different time periods, initially with his family in Norway and later in Iceland where it follows Egil’s life. The corresponding network, depicted in Figure 4, does not have a single cohesive social structure, which may account for its disassortativity. We see that, although the five sagas have many common features, the statistical analysis is capable of picking up differences in detail between them.

To test how strongly connected the sagas were, we merged the five large networks (Figure 5) and then applied a community-detection algorithm to see if we could break the merged network back down into five components again. This process failed to separate \textit{Njáls saga} and \textit{Laxdæla saga}. The overlap between these two tales offers support to the theory that one saga may have been used as a source for the other.

We also merged the entire corpus of all 18 sagas into one huge network. The entire society is also small-world, structurally balanced, assortative, and the giant component contains 98.6% of the 1547 unique characters. The average path length is 5.5 – remarkably close to the six degrees of separation of modern society discussed above. Rather than a simple power law, the degree distribution for the entire saga society is better described by a power law with exponential cut-off. This truncation is due to the fact that no single protagonist appears as a major player in multiple sagas – a high-degree character in one only turns up as a minor character in another, and thus the associated degree does not grow as the network size increases.

We can also use several iterations of the community-detection algorithm to try to break the amalgamated network back down into components, evaluating the modularity at each stage. We found the modularity – the measure of the number of sparsely-interconnected communities – reaches a plateau at over 0.7 with \( n = 9 \) communities – indicating that it is not possible to split the amalgamation back down to the 18 separate sagas. This again
indicates the remarkable interconnectedness of the Íslendinga sögur.

It is also interesting to compare the Íslendinga sögur to the epics we analysed earlier, and to fiction. A striking difference between the Icelandic texts and the Iliad, for example, is that while the full saga network is assortative, that of the Iliad is not. The two networks are more similar when hostility is excluded from the Iliad (which is then assortative). In other words, the Íslendinga sögur have similar network properties to the Iliad only when hostility is removed from the latter.

Moreover, although it is acknowledged that Tolkien was influenced by Nordic literature, the sagas have a very different network structure than The Lord of the Rings and other works of fiction. Works of fiction that have so far been studied, including the superhero world of Marvel comics, typically have exponential degree distributions and are strongly disassortative.

Summary

The Íslendinga sögur have fascinated humanities scholars for generations; we think we have brought a unique statistical approach to the study of the texts. It is an example of how statistical techniques applied to the humanities can open new fields of inquiry.

By analysing not the characters themselves, but how they are interconnected, we could compare the social structures underlying the narratives and statistically characterise different types of sagas. This is part of a broader initiative under way at Coventry University comparing epic and mythological literature and identifying similarities and differences between them.

Obviously the comparisons and conclusions we make here are from a statistical, network-theoretic point of view; for a holistic view our work must be combined with that from other fields – archaeology, literature, and comparative mythology are a few. Nonetheless, we can conclude that whether the sagas are historically accurate or not, the properties of the social worlds they record are similar to those of real social networks. Although one cannot conclusively determine whether the saga societies are real, on the basis of network theory we can conclude that they are remarkably realistic.

Acknowledgements

This project is supported by the Leverhulme Trust and a Marie Curie IRSES grant within the 7th Framework Programme. We are grateful to Robin de Regt and Joseph Yose for help in gathering the data.

References


Pádraig Mac Carron is a PhD student at Coventry University’s Applied Mathematics Research Centre, of which Professor Ralph Kenna is Deputy Director.