DOCTOR OF PHILOSOPHY

A network scientific approach to the quantitative analysis of epic texts

Yose, Joseph

Award date:
2017

Awarding institution:
Coventry University

A Network Scientific Approach To The Quantitative Analysis of Epic Texts

Joseph Yose

Applied Mathematics Research Centre, Coventry University

Director of studies: Prof Ralph Kenna

Supervisors: Dr Thierry Platini and Dr Pádraig Mac Carron

A thesis submitted for the award of

Doctor of Philosophy

2017

Acknowledgements

I would like to thank the following people who have inspired, supported, guided and encouraged me during the work of this thesis.

I am most grateful for my director of studies Prof Ralph Kenna whose patience, teaching and guidance went far beyond the call of duty. I have been so privileged to have someone who cared so much about my work and my progression. I am so thankful for all the opportunities that were made available to me, I met good friends, and I had fruitful research collaborations.

I am indebted to my supervisor Pádraig MacCarron for his unwavering support in and outside academia. Completion would have been difficult without his support. I appreciate how he got me into coding.

A very special thank you to Thierry Platini who became a sounding board for many of my ideas and for helping me with the mathematics behind this work.

And of course a big thank you, to Robin de Regt and Petro Sarkanych, for their enormous help with proof reading and other ad hoc tasks.

I would like to say a heartfelt thank you to my wife Kumbulani for her encouragement and support especially during the writing-up stage. “I did not know you are such a powerful motivational speaker”. Thanks also to my mum gogo Yose and my two lovely daughters, Kay and Kaylee for making it possible for me to complete what I started. You have not only been helpful but done so in a very much needed humorous way.

And finally but most importantly, thanks (Dyakuyu) to my friends and fellow PhD students at AMRC and ICMP, Anthony, Erin, Gerry, Hamid, John, Lee, Marjana, Natalia, Oksana, Ravi, Richard and Sanjay.

Contents

List of Figures

Chapter 1

Introduction

Network science has evolved out of various well-established disciplines including mathematics, statistics, physics, computer science and sociology. It enables the statistical treatment of certain types of systems comprising large numbers of interconnected elements. The collective behaviour that can be exhibited by such a system can manifest emergent features; i.e properties of the whole can be greater than the sum of those of its individual parts. Network science has been applied to a wide range of areas that include, biology, epidemiology, transportation, telecommunication, finance, scientometrics and more (see, e.g., Refs. [1–10]). Its application to different fields is still developing and it is still finding new applications in areas where it amalgamates more than one branch of knowledge, bringing together diverse traditional academic disciplines. Besides delivering new quantitative insights when applied to old problems, network science approaches inspire new questions and opens new avenues of research.

Graph theory is the mathematical platform behind much of network science and can be traced back to Leonhard Euler's mathematical investigations of the Königsberg-bridge problem in the 18^{th} century. Königsberg was a capital city of Prussia on the Preger river up to 1701. The city had seven bridges that connected two islands and the question was whether one could walk around the city in such a way that transverses each of the bridges precisely once. Euler presented the arrangement of the bridges as a graph and provided a mathematical proof that no such path exists [11, 12]. The graph corresponding to the bridge was an early representation of a network [13, 14].

From the 1970s, the empirical study of social networks has also played an important role in social science. Indeed, some of the mathematical and statistical tools were first developed in the field of sociology. In such applications, social-network analysis was employed to analyse the propagation of information such as news, ideas, inventions and culture as well as to understand the spread of diseases and opinions. Social networks have also been studied in economic systems, for example to investigate exchange relationships in trading. The spread of political ideology, the movement of traffic, and the study of scientific collaborations have also been targeting network analysis from the sociological point of view.

It was in 1998 that Watts and Strogatz produced their now famous paper [13] containing a model that could explain the clustering observed in many real-world networks while still manifesting the short average path lengths of the Erdős-Rényi random graphs. This was an important step mathematically and one that drew the attention of many from the statistical physics community. Striking similarities were noticed between the structure of an electrical grid, social networks of Hollywood actors and the wiring of the neural system of a nematode worm. Since then, the study of networks has accelerated and it now forms a large component of complexity science and applied mathematics, overlapping with many other disciplines.

Very recently a new sub-discipline has emerged from the applications of network science to study the relationships between characters in works of literature. The study of so-called character networks has gained in popularity in recent years. Previous analyses have focused on the universal properties and distinguishing features of character networks, particularly in areas such as mythology [8, 15, 16]. There it facilitates new quantitative ways to compare mythological narratives within and between cultures. This thesis follows on from these pioneering investigations and examines two important works of the Gaelic tradition of epic narratives. These are *Cogadh Gaedhel re Gallaibh* which is an account of the Viking age in Ireland and the *Poems of Ossian*, a controversial but highly influential Scottish narrative.

In this thesis, we mainly focus on the character networks embedded in these two epic narratives and compare them to others of the broad genre. In character networks, the basic elements in question are individuals (characters) depicted in the narrative and these are represented by nodes or vertices of the network. Interactions and relationships between these characters are represented by edges or links. Network science offers the mathematical and statistical resources to capture a plethora of quantitative features associated with the societies depicted in the texts. By comparing the statistics from various networks to each other, one can gain insight into similarities and differences between such societies. Of course, the application of networks to epic narratives does not supplant current methods of analysing such texts in the humanities. Rather our approach provides an additional tool — a quantitative methodology — that can extract new features and inspire new angles for investigation.

In the following chapter, we start by giving a background to *Cogadh Gaedhel re Gallaibh* and describe the characters and events in the Viking age in Ireland from the late 8^{th} century to 1014 AD. The text presents conflict of the time as a war between the Irish and the Vikings (i.e. an international conflict, a view supported by traditionalists). However, revisionists oppose this view and instead consider hostilities as part of a civil war. Our aim is to apply network theory to the collection of interactions contained in *Cogadh Gaedhel re Gallaibh*. Our primary aim is to gain some insight into the mixing patterns of interactions in the text to determine what these have to say about the traditionalist versus revisionist debate.

The second part of Chapter 2 begins by contextualising the origins of the Ossianic controversy, outlining the cultural circumstances in which the work appeared. It also documents how it was received both positively and negatively. The poems published in 1760, reached international acclaim and invited comparisons with major works of Homer. However, the work was plagued by authenticity issues which ignited a persistent debate that is still existing in humanities today. Our primary aim here is to compare the networks of *Ossian* with those coming from other works which are claimed to have been influential in its composition. We hope to identify which of these is closest to *Ossian* and if we can explain similarities detected.

In Chapter 3, we introduce various metrics from network theory that we use in this thesis. Historically, new methods and approaches in complexity science and applied mathematics are inspired through necessity — applications require the appropriate tools. Here too, we find the need to invent new measures. We introduce a renormalised version of the so-called *categorical assortativity* which facilitates meaningful comparison between some network systems in a manner which was not previously accessible. Our main aims are addressed in the subsequent chapters. We use the quantities that we discussed in Chapter 3 to provide empirical results that capture the statistics of networks found in Cogadh Gaedhel re Gallaibh and *Ossian*. These results are presented in chapters 4 and 5 respectively.

Chapter 6 carries on from our analysis of *Ossian* in Chapter 5 and addresses further questions about its origin. Here we employ sampling methods known as sparsification in network theory, to investigate if there is a process that generates an Ossianic type network from a likely source. We summarise and give our conclusions in the last Chapter.

The investigation and analysis of *Ossian* in Chapter 5 is published in *Advances in Complex Systems* as Ref. [9]. An online visualisation of *Ossian*'s character network is also available on the github page `https://yosej.github.io/Ossian_Full_Network`. The work on Cogadh Gaedhel re Gallaibh contained in Chapter 4 has been submitted to the journal *Royal Society Open Science* and on the arXiv in Ref. [17]. The visualisations of the corresponding networks will likewise be placed on github. The material in Chapter 6 is currently being prepared for publication.

Chapter 2

Epic Texts

In this thesis two epic narratives associated with the Gaelic tradition are investigated from a perspective of network science. The adjective Gaelic pertains to the language of the Gaels, i.e., Irish, Scottish Gaelic and Manx (from the Isle of Man). The first is *Cogadh Gaedhel re Gallaibh* (hereafter sometimes referred to simply as "the *Cogadh*"), one of the fifteen most important and extensive Irish chronicles [18]. The title, in Irish, translates to "The war of the Gaedhil with the Gaill" or "War of the Irish with the Foreigners". It has been described as a "romantic tale" which tells of events in Ireland between the late 8$^{\text{th}}$ and early 11$^{\text{th}}$ centuries. It differs from most of the remaining Irish chronicles mainly in that it is not "a simple record of events" but a dramatic account "in which heroes shine and villains play their sinister parts" [19]. The earliest version of the text itself is dated between the 11$^{\text{th}}$ and 13$^{\text{th}}$ centuries. The text recounts how Brian Boru, an Irish king, led a coalition of forces to triumph against Viking invaders. The general outlines of the story have entered popular tradition in Ireland in a manner reflective of this (i.e., as a clear-cut war between the Irish and Vikings). However, modern scholars view the chronicle as largely propagandistic. A debate has arisen in the humanities as to whether the events it describes really amount to an international conflict between Irish and Vikings or whether they describe a domestic dispute, internal to Ireland, primarily between two Irish factions. Our aim is to contribute to that debate by using the modern science of network theory to analyse Viking-age interactions recorded in the text.

The second set of texts that we examine are from Scotland and come under the title *The Poems of Ossian* [20–23]. These are purported to have been composed in Scottish-Gaelic by a third-century bard named Ossian and translated into English by James Macpherson in about 1760. The poems received international attention and were compared to Homer's *Illiad* and *Odyssey*. However, they also caused great controversy amid claims they were essentially misappropriated from Irish tradition. An epic is defined as a long narrative detailing heroic deeds and describes events of a nation or culture. Our aim is to compare the network structures contained within *The Poems of Ossian* to other epic works to which they have been matched in the past. In particular, we wish to determine whether they more closely resemble those from the Irish tradition or of the Classics.

In this chapter we provide some background to the network studies contained in subsequent chapters. We begin with *Cogadh Gaedhel re Gallaibh* in Section 2.1. We first provide some context in the form of a short discussion in Sub-section 2.1.1 of the main political divisions of Ireland in the Viking age. In Sub-section 2.1.2 we then pose the main question that we wish to address and discuss why network theory may deliver a meaningful answer. In Sub-section 2.1.4 more details of the context are given in the form of a summary of the main events of the Viking Age in Ireland and set down in *Cogadh Gaedhel re Gallaibh*. The authenticity and deficiencies of *Cogadh Gaedhel re Gallaibh* as a source are discussed in Sub-section 2.1.5. After giving all of these contextual details and the humanities background to our study, we restate the main question that we wish to address in the context of *Cogadh Gaedhel re Gallaibh* in Sub-section 2.1.6.

We turn our attention to the *Poems of Ossian* in Section 2.2. Following a very brief summary of the controversy surrounding the texts and of recent revisionist re-interpretations, we again declare our main outcome for the impatient reader. In the Sub-section 2.2.1, we outline the cultural circumstances in which *Ossian* appeared and in Sub-section 2.2.2 we discuss how it received both praise and condemnation. We discuss the enormous cultural influence of *Ossian* its legacy including the controversy surrounding the work. Despite the vast amount of humanities literature on the topic, the *Poems of Ossian* continue to appeal as an interesting area of academic research, recently revitalised by revisionist and counter-revisionist arguments. This is the context in which we pose our main question in Sub-section 2.2.3 and explain why we expect network theory to provide an answer.

2.1 Cogad Gáedel re Gallaib

The year 2014 marked the millennial anniversary of the famous Battle of Clontarf, an event of great importance in the history of Ireland. In popular tradition and collective memory, the battle is viewed as marking the end of Viking power after two centuries in the country. After Clontarf, Viking ambitions in Ireland were essentially reduced as they submitted to the overlordship of Irish kings. The recent anniversary inspired academics to revisit the period, and multiple articles appeared in academic journals, booklets, books, online commentaries and various media engagements (e.g., Refs. [24–34]). Quite naturally, these approached the subject matter using long-standing, traditional tools of the humanities, extending and building upon earlier investigations which go back centuries (e.g., Refs. [19, 35–61]). There has been an ongoing debate within the humanities community regarding the precise nature of the relationships and interactions described in *Cogadh Gaedhel re Gallaibh*. The debate, which has lasted at least seven decades, is between what we may call "traditionalist" and "revisionist" interpretations of the exchanges between the conflicting parties [19, 28–30, 34, 43]. (In fact, there is an argument that the discussion can be traced back as far as 250 years [28, 62, 63].)

In this thesis, we present an alternative investigation into *Cogadh Gaedhel re Gallaibh* using a novel approach based within complexity science. Although limitations of *Cogadh Gaedhel re Gallaibh* are multi-fold and well documented, it remains an extensive and useful source of information for the Viking Age in Ireland. The text recounts the exploits of multitudes of characters, conflicts, alliances, relationships and interactions between them.[1] The statistical tools appropriate to tackle networks like these have recently been developed [9, 15, 16]. They require large casts of characters to deliver reasonable statistical power. Here our aim is to apply them and to further develop them to deliver a new type of investigation into the Viking age in Ireland as presented in *Cogadh Gaedhel re Gallaibh*.

Nowadays, the set of events, interactions and relationships associated with the Viking Age in Ireland and, indeed, with the Battle of Clontarf, are frequently believed to exist in the public memory in an excessively simplified and even incorrect manner. The popular picture of the public at large is one of an "international" conflict between native Irish and "foreign" Viking. After centuries of conflict in the country, victory for the former ended the invaders ambitions in the country. We are told that this picture is at variance with the truth, which is more complex and nuanced [19, 43]. Instead of a conflict whose nature is international, it is claimed that the main issue at stake at the climactic Battle of Clontarf was actually a domestic dispute or civil war. It was all about an internal, Irish struggle for sovereignty. Specifically, the new orthodoxy views the conflict as primarily concerning the determination of Leinster (the eastern province of Ireland) to remain independent of the high king [19, 43].[2] Some such interpretations, wherein

[1]There are a number of theories for the origins and meaning of the word "Viking". In this thesis we use it to refer to the medieval raiders and invaders from Scandinavia who attacked Ireland, amongst other countries, by sea, between the late 8th and 11th centuries [61]. Many of these subsequently settled in the country. There are stricter definitions of the term "Viking" which may involve the notion of "piratical". When used in this sense, the term encompasses people beyond Scandinavia. Therefore, one may argue that not all Vikings were Scandinavian and, indeed, that not all Scandinavians were Vikings [61]. But here we use the term in the looser sense — along the lines in popular, modern usage. This is also keeping with much of the literature, e.g., Refs. [45, 48, 51, 52, 54, 56, 58, 59]). The Vikings we refer to are the *Gaill* (singular *Gall*) referred to in the title of the narrative. There are also Hiberno-Norse (Hibernia is the Classical Latin name for Ireland) or Norse-Gaels, also called Gall-Ghaedheil or "foreigner Gaels". These emerged as a mixture of Gaelic (Irish) and Norse culture through intermarriage. They include, for example, the kingdom of Dublin.

[2]The term "national" may be viewed by some as anachronistic and misleading here [30]. It is used in the sense of a large group of people with common characteristics such as language, traditions, customs and ethnicity [51] rather than in a governmental sense [64].

Figure 2.1: Left: The kingdoms of Ireland in 1014 (courtesy of Wesley Johnston, `http://www.wesleyjohnston.com`). Right: The first page of the oldest existing copy of *Cogadh Gaedhel re Gallaibh* (from the twelfth century Book of Leinster).

the Vikings play a secondary role, tend to downplay the significance of Clontarf [34] and have been partly ascribed to revisionist fashions [28, 54]. The aim of the investigations described herein is to determine what the character networks contained in *Cogadh Gaedhel re Gallaibh* have to say on the matter.

2.1.1 Background

Figure 2.1 contains a map outlining the political structure of Ireland in 1014. According to legend, Ireland was divided into two parts in ancient times. The northern part was called *Leath Cuinn* (meaning Conn's Half) and included what became the medieval provinces of Connacht, Ulster and Meath. The southern half was called *Leath Moga* (Mugh's half) and included the provinces of Munster and Leinster. Because there were five provinces altogether, these were called *cóiceda* or "fifths". The medieval provinces of Ulster, Connacht and Munster approximately correspond to the modern Irish provinces of the same names. A combination of the ancient provinces of Leinster and Meath approximate what is the modern Irish province of Leinster. In modern times, *Cúige* remains the Irish word for "province", even though they number only four instead of five. In the Viking Age, the main contest for the overlordship of Ireland (the so-called "high kingship") was between two branches of Uí Néill dynasty, both based in the northern half of the country. These were the so-called northern and southern branches of the Uí Néill, The dominance of the Uí Néill was finally ended by Brian Boru, as we shall see. In Figure 2.1 the red regions are Cork, Dublin, Limerick, Waterford and

Wexford. These were all developed from Vikings base camps in the ninth century to more permanent settlements and formed the foundations of the modern towns and cities of those names in Ireland.

Figure 2.1 also contains an image from the famous Book of Leinster (one of a number of famous Irish annals which have survived). It is the opening page of the copy of *Cogadh Gaedhel re Gallaibh* recorded in that volume (which is also the oldest existing version). The narrative style used in the *Cogadh* is "inflated and bombastic" [35] and modern scholars consider it to be a brilliant piece of propaganda, whose main purpose was to praise Brian Boru and his dynasty — the *Dalcassians* or *Dál gCais*. It is viewed as being of particular importance in asserting the claim of the *Dál gCais* to the high kingship of Ireland [35, 45, 48]. This is achieved through the usage of long, detailed tracts declaring the virtues of Brian and his army while portraying the Vikings as piratical and brutal. However such qualitative, rhetorical features are largely irrelevant for quantitative character-network analysis. Instead, the approach used here draws only from the most basic information, namely the presence or absence of interactions between characters which may be positive or negative in a sense to be described below. The depictions of the characters themselves — whether they be virtuous or piratical, for example — do not enter the analysis. Still, even such basic information is imperfect; e.g., the versions of the *Cogadh* which have come down to us contain imagined interactions in the form of interpolations (discussed below). If, despite this, the text contains networks which are reasonably or approximately reliable in the aggregate, they deliver essential information on the society at the time.

2.1.2 The networks contained in the text

A network visualisation of the full set of interactions between the characters of *Cogadh Gaedhel re Gallaibh* is depicted in Figure 2.2. Here and throughout individual network nodes represent characters identified in the text. Relationships or interactions between these characters are represented by edges or curved line segments connecting them. The nodes are coloured so that Irish characters are represented in green. Green edges represent interactions between pairs of Irish characters. The counterpart set of Viking nodes and their interlinks are in blue. Brown edges represent interactions between Irish and Viking nodes. Any remaining nodes and edges are in grey. Our aim is to analyse the network represented in Figure 2.2 and sub-networks associated with it in order to gain quantitative insight into the complexity and conflicts of the Viking Age in Ireland as described in *Cogadh Gaedhel re Gallaibh*.

The primary question we wish to ask in relation to this text is the following: Are the networks underlying the *Cogadh*

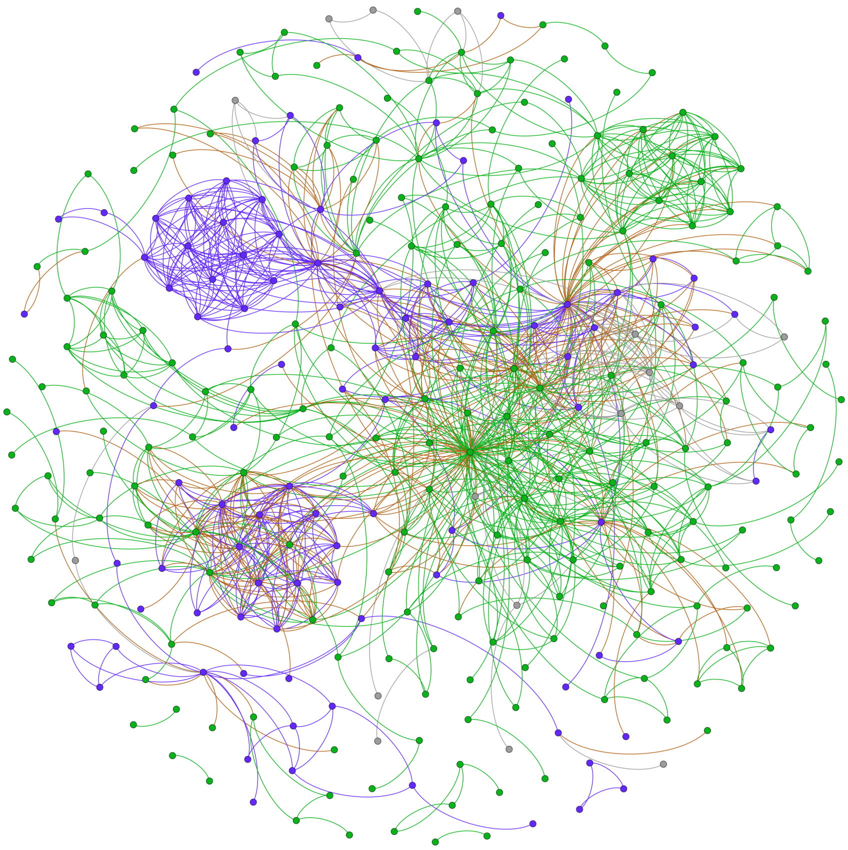

Figure 2.2: The entire *Cogadh* network of interacting characters. Characters identified as Irish are represented by green nodes and those identified as Vikings are in blue. Other characters are in grey. Edges between pairs of Irish nodes are also coloured green while those between Viking pairs are blue. Edges linking Irish to Viking nodes are brown and the remaining edges are grey.

(a) consistent with the traditional depiction of a contest which is clear-cut international or

(b) supportive of the revisionist notion of a power-struggle which is mostly domestic or

(c) suggestive of something between both pictures?

In other words, our aim is to determine whether the conflictual edges in Figure 2.2 are mainly blue and green or brown. There is a related question which we are also able to address, namely whether a set of positive interactions mostly connects nodes of the same or different cultural types. A simple count of edges is not adequate as such an approach would not take into account different numbers of nodes of different type. Instead, a proper quantitative approach necessitates the networks-science concepts of *assortativity* and *disassortativity*. The former is the tendency for nodes which have similar attributes to connect to each other. The opposite tendency is disassortativity, whereby links tend to be between nodes of different types. The type of attribute we are interested in here is cultural identity; we wish to gauge whether nodes linked by different types of edges represent Irish or Viking characters (the Gaedhil or the Gaill). We use the term "cultural assortativity"[3] to denote associated measures which we denote generically by ρ. One such quantity will be used as the primary determinant to distinguish between the alternatives (a), (b) and (c) listed. A network with a positive value of ρ is said to be culturally assortative. A negative value of the metric signals disassortativity. A value close to zero indicates the absence of any such correlations (neither assortative nor disassortative).

The attributes we wish to investigate in the cultural context are positive (i.e., social) and negative (i.e., hostile) interactions between characters of the text. Our main aim is to investigate the configurations of hostile links, as these are the ones which are relevant for the above-mentioned debate. One could argue that in an extreme situation (a limit, in a sense, of the overly simplified, traditional picture), a scenario of purely "international conflict" (a) might be expected to contain networks where hostility is "clear-cut" between nodes of different cultural identities.. This ("clear-cut") is the term employed by Ryan in a paper which challenged the traditional picture Ref. [19]. The value of the cultural assortativity metric ρ for such a network would be expected to be strongly negative. The revisionist picture of domestic or intranational conflict (b), on the other hand, would have a positive value of ρ because hostility is primarily between sides of the same cultural identity tags.[4] A moderate value of cultural assortativity, lying between the two extremes (i.e. a negative value of ρ which is moderate in magnitude) would then point to a more complex picture (c).

2.1.3 The outcome of the study of *Cogadh Gaedhel re Gallaibh*

The unique strength of our analysis is that, unlike previous qualitative investigations which tended to polarise the issue, we can deliver quantifying statements. We will show that the spectrum (see Figure 2.3) of potential values of cultural assortativity for networks of the conflictual type contained in *Cogadh Gaedhel re Gallaibh* ranges from $\rho = -0.88$ to $\rho = 1$. The traditional "clear-cut" picture corresponds to the left-hand limit. The modern, revisionist

[3]The term is motivated by a discussion in Ref. [51] of "the strong sense of identity, achievement, and cultural cohesion that had been created by the Irish learned classes." Ó Corráin states "The island was united culturally and linguistically" and "Self-consciously, the literati saw the Irish as a people or *natio*, to be compared with the Goths, the Franks, or the people of classical antiquity. As far as the genealogists were concerned, the Vikings were outsiders, and were called *Gaill* 'Foreigners' to the end. Irish reaction to the Vikings is to be understood in terms of these cultural traits." For further discussions of Hiberno-Scandinavian relations, see Ref. [52].

[4]We loosely use the term "intranational conflict" to describe Irish-on-Irish and Viking-on-Viking hostility although there are two sets of Vikings — Danes and Norse which frequently fight with each other. We deal with that issue in Section 4.7

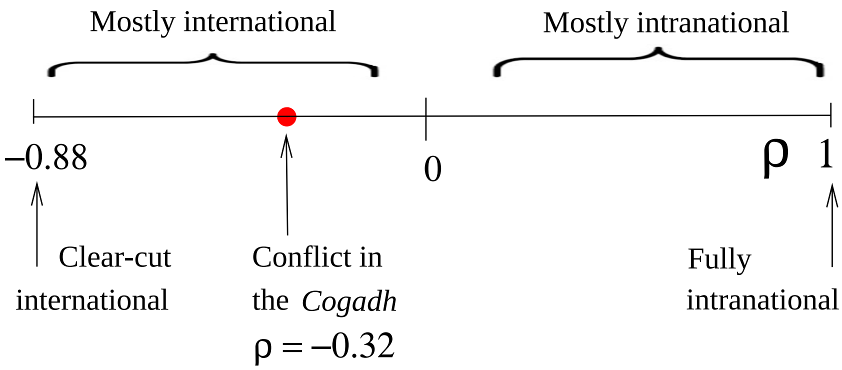

Figure 2.3: Graphical representation of the main conclusion of the analysis of *Cogadh Gaedhel re Gallaibh*. The spectrum of values of cultural assortativity for networks of the conflictual-*Cogadh* type ranges from $\rho = -0.88$ to $\rho = 1$. Negative values of ρ correspond to various degrees of the traditional picture of international hostilities with $\rho = -0.88$ representing a clear-cut Irish-versus-Viking conflict. Positive values correlate with the revisionist picture of mostly intranational conflict. The analysis presented in this paper shows that the *Cogadh* hostile network delivers a value −0.32 which, although not clear-cut, lies on the traditional side of the spectrum.

picture is associated with the right-hand side. We determine a value for the actual *Cogadh* network of $\rho = -0.32$. Its negative sign signals that the depiction of conflict in the narrative is mainly international. Its moderateness is indicative of significant intranational conflict too. Therefore the account is not as clear cut as either the traditional or revisionist pictures suggest, but is located between their two extremes, on the traditional side.

The principal aims of our investigation into *Cogadh Gaedhel re Gallaibh* in later chapters of this thesis are (i) to present visualisations for the positive and negative networks, (ii) to develop the notion of cultural assortativity to estimate where a network of interactions is positioned on the spectrum from the international to the intranational and (iii) to apply that tool to the networks recorded in *Cogadh Gaedhel re Gallaibh*. Another objective is to (iv) quantify various statistical properties of the *Cogadh* networks. These will allow us to compare with other epic narratives. However, before presenting these details, we first provide more contextualisation of the problem with a summary of *Cogadh Gaedhel re Gallaibh* and previous studies of it (all of which belong to the humanities).

2.1.4 Context and text: The war of the Gaedhil with the Gaill

Cogadh Gaedhel re Gallaibh comes down to us in three manuscripts. The oldest is in the twelfth century *Book of Leinster* (so named because of the associations of its content with that province of Ireland), which contains a fragment of the tale. The second (also incomplete) is the *Dublin Manuscript* (so called because the copy is kept in Trinity College Dublin), dated to the 14th century. The third, and only complete text is the *Brusssels Manuscript.* This was saved by the famous Franciscan friar Mícheál Ó Cléirigh who in the 17th century was sent from Louvain in Belgium to Ireland to collect and preserve Ireland's ancient heritage. It is a copy of an earlier work, *Leabhar Chonn Chonnacht Ui Dhálaigh* (Book of Cuconnact O'Daly who died in 1139 [35]) which is now lost. The Brussels and Dublin manuscripts are close but not identical. Ní Mhaonaigh gives a detailed textual history of *Cogadh Gaedhel re Gallaibh* in Refs. [47, 48].

As a proxy for the originals, we use the nineteenth-century translation into English by James Henthorn Todd [35]. Todd's edition, which is 150 years old in 2017, is accompanied by an extensive introduction and by detailed explanatory footnotes. It was prepared for the series *Rerum Britannicarum Medii Ævi Scriptores* ("Chronicles and Memorials of Great Britain and Ireland during the Middle Ages") and is considered well balanced and thorough. It serves as a source for some scholars wishing to access the narrative today [56].

Todd considered *Cogadh Gaedhel re Gallaibh* as divisible into two parts. The first recounts the arrival and deeds of the Vikings in Ireland in a rough chronological fashion. The second part concerns Brian Boru and his Munster dynasty whose power-base was on the banks of the river Shannon. The lives and politics of his family are outlined along with numerous encounters with the Vikings, all leading to the events at Clontarf.

Brian Boru or Brian Bóruma[5] mac Cennétig was son of Cennétig mac Lorcáin, king of the Dál gCais in the northern part of the province of Munster (a map of Ireland during the Viking Age is provided in Subsection 2.1.1). On the death of his father, Brian's older brother Mathgamain assumed control. After capturing Cashel, the traditional capital of Munster, he claimed authority over the entire province. On his death, Mathgamain was replaced by Brian who, like his brother, proved to be an excellent military commander. After various battles at the provincial level, Brian and the Dál gCais consolidated rule of Munster, defeating their Irish and Norse challengers. Brian then turned his attention to the easterly province of Leinster and the westerly province of Connacht.

Máel Sechnaill mac Domnaill was king of Meath (Mide) and high king of Ireland. He belonged to the Clann Cholmáin sept of the Southern Uí Néill. Brian's interests in Leinster brought him into contest with Máel Sechnaill who led repeated forays there, as well as into Munster. However, in 997, Brian and Máel Sechnaill agreed a truce, whereby the former would rule over the (approximate) southern half of Ireland, while Máel Sechnaill kept the (approximate)

[5]The epithet "Bóruma" in Old Irish ("Borumha" in Todd [35]) is an alternative to "Boru" and may signify "of the cattle tributes" [38].

northern half as well as the high kingship. By these means, Brian came to control Munster and Leinster as well as the Hiberno-Norse cities within, while Máel Sechnaill held the provinces of Connacht, Meath and Ulster.

In 998, Brian and Máel Sechnaill worked together against the Dublin Norse. The Vikings had established a settlement in Dublin in 838 and during the following century they developed a kingdom comprising large areas surrounding the town and controlling parts of the Irish Sea. Viking Dublin was politically linked to the Isle of Man and the Hebrides, as well as to Viking settlements in Britain and Scandinavia. Indeed, in 980, an unsuccessful Dublin campaign against Máel Sechnaill mac Domnaill had support from the Isle of Man and the Hebrides.

Dublin was joined by Leinster under a new king, Máel Morda mac Murchada, in opposing Brian and Máel Sechnaill. Leinster traditionally rejected the rule of both Munster and Meath and the Hiberno-Norse city of Dublin was ruled by Máel Morda's cousin, Sigtrygg Silkbeard. The two sides met at Glenmama in late December 999. The Irish annals agree that the combined forces of Munster and Meath decisively defeated those of Leinster and Dublin. Sigtrygg, however, was allowed to retain his rule over Dublin. Indeed, one of Brian's daughters married him while Brian himself married Sigtrygg's mother Gormflaith (Máel Morda's sister). Gormflaith was also the former wife of Máel Sechnaill.

The river Shannon presented a barrier to Meath receiving support from his ally Cathal mac Conchobar mac Taidg, king of Connacht, when Máel Sechnaill came under attack by Brian in the year 1000. By 1002, Máel Sechnaill had surrendered the high kingship to Brian at Athlone. The next target for Brian was Ulster. It took ten years, a combination of forces and coordinated use of sea and land attacks, and support from the Church in Armagh for the Northern Uí Néill and regional kings of Ulster to submit to Brian. By 1011, Brian had achieved his aim of bringing all the regional rulers of Ireland under his control.

In 1012, Máel Mórda mac Murchada of Leinster rose in rebellion. Allied with Flaithbertach Ua Néill, regional king of Ailech in Ulster, he again attacked Meath. Máel Sechnaill sought Brian's help and the following year Brian and his son led a combined force from Munster and Connacht into Leinster, reaching Dublin in September. Out of supplies near the end of the year, they abandoned their siege of the walled city, with an intention to return.

Thus was the background to the famous Battle of Clontarf. In 1014, Máel Morda's cousin, Sigtrygg, journeyed to Orkney and the Isle of Man seeking Viking support. These Norsemen came under Sigurd Hlodvirsson (Earl of Orkney, known as Sigurd the Stout) and Brodir, apparently of the Isle of Man. Brian's forces came from Munster and southern Connacht possibly supported, at least initially, by Máel Sechnaill mac Domnaill's Meathmen (the precise role of Meath in the Battle itself is a matter of some contention [19, 28, 54]). The Battle of

Clontarf is believed to have taken place on Good Friday, 23 April 1014 [35] (see, however, Ref. [30,65]). After a day's fighting, the battle ended with the routing of the Viking and Leinster armies. Their retreat was cut off by the high tide. Many of the nobles died. Brodir killed Brian, having found the old man in his tent. Brodir in turn was killed by Wolf the Quarrelsome, a relative of Brian Boru. Sigurd the Stout of Orkney was also killed, as was the Leinster king Máel Morda mac Murchada. Sigtrygg Silkbeard survived and remained king of Dublin, and the king of Meath, Máel Sechnaill mac Domnaill, resumed his high kingship of Ireland, at least in name, supported by Flaithbertach Ua Néill.

2.1.5 Authenticity and deficiencies of *Cogadh Gaedhel re Gallaibh*

It is nowadays widely accepted that one of the main aims of *Cogadh Gaedhel re Gallaibh* was to praise Brian Boru. Its purpose was to document the achievements of the Dál gCais and establish their right to the high kingship of Ireland. Although it is a valuable resource for studies of the Viking Age in Ireland, it is considered a flawed one. The question of the extent of its reliability has been the topic of a very long-standing debate [19, 28, 35, 43, 48]. Besides some clear interpolation, many of these flaws appear in the descriptive detail of the narrative. Ours, however, is a statistical analysis and, as such, is rather concerned with the totality of the interactions between characters rather than rhetorical levels of detail. As with any statistical analysis, what it delivers is a summary which captures aggregate characteristics, largely insensitive to individual elements. In this sense, one may hope that it delivers useful statistical information on the Viking Age in Ireland.

Estimates for the date of *Cogadh Gaedhel re Gallaibh* are various. Todd stated its author "was a contemporary and strong partizan of King Brian Borumha" [35]. Following consideration of the possibility of a more recent date, he concluded that the author was "either himself an eye-witness of the Battle of Clontarf, or else he compiled his narrative from the testimony of eye-witnesses" [35]. Flower also considered the chronicle "almost contemporary" [42]. Goedheer gives a date as late as 1160 [40] but Ryan argues that *Cogadh Gaedhel re Gallaibh* "might have been composed about 1130 or earlier" [41] In Ref. [43], Ó Corráin described the *Cogadh* as "concocted ... some two-and-a-half centuries after the date of the events it purports to narrate". Elsewhere in Ref. [43] Ó Corráin refers to it as "written in the twelfth century". He also describes the hypothesised text known as *Brian's saga* (discussed below) as written about 1100 in response to *Cogadh Gaedhel re Gallaibh*, a suggestion that implies a date before 1100 for the creation of the latter [43]. More recent scholarship by Ní Mhaonaigh gives the likely composition date of *Cogadh Gaedhel re Gallaibh* as between the years 1103 and 1113 [47]. (She dates the common source for the Dublin/Brussels recension as the 1120s or 1130s [45,47].

Casey, who also reviews dating estimates, sets this at about one hundred years after the events described [59].) Duffy believes it may be "based on contemporary annals and, no doubt, local memory" [28]. He suggests that *Cogadh Gaedhel re Gallaibh* gives "a vivid picture of what happened at Clontarf as related perhaps to the writer of the Cogadh by a veteran" and gives the possibility that it "was written by someone who may well have lived through these last year's of Brian's life". This bringing us back to Todd's original estimate [35].

The interpretation of *Cogadh Gaedhel re Gallaibh* as propagandistic is linked to the question of the date of its composition because "Heroic stature presupposes nurturing by time" [45]. Thus its propagandistic nature "implied that it could no longer be considered contemporary with any of the events it describes" [45]. The greater the distance between the events of Clontarf and the setting down of *Cogadh Gaedhel re Gallaibh*, the more room there is for a distorted view to take hold. This is the reason why a good estimate date for the composition of the *Cogadh* is important in the present context. "In the course of the eleventh century, ... the view seems to have gained universal acceptance that the Battle of Clontarf was par excellence the great decisive struggle of Irish history. Brian in the retrospect was everywhere acclaimed as a national hero" [19]. The claim is that time distorted reality; "The Norse were a substantial section of the opposing force, and in the mellow haze of popular imagination the battle tended to be transformed into a clear-cut issue, Irish versus Norse, with the former victorious. Even in the Northern countries the battle passed rapidly from history into saga" [19]. The above estimates for the interval between Clontarf and composition of *Cogadh Gaedhel re Gallaibh* range between contemporary and two hundred and fifty years. Our approach cannot deliver an independent estimate for the date of composition and the above estimates should be kept in mind. While the above considerations suggest that the *Cogadh* may distort in favour of an overly international picture of conflict, on the other hand it should also be kept in mind that, in places, it identifies Leinster as the principal enemies of Brian [28,30]. We return to these issues in Section 4.8.

In his Introduction to *Cogadh Gaedhel re Gallaibh*, Todd acknowledges the defects of the work and expresses regret that it is "so full of the feelings of clanship, and of the consequent partisanship of the time, disfigured also by considerable interpolations, and by a bombastic style in the worst taste ...". In chronicle literature, an interpolation of the type mentioned by Todd is a later addition not written by the original author. As scribes copied ancient material by hand, extraneous material frequently came to be inserted for a variety of reasons [66]. These may have been for bona fide intentions, perhaps as explanations; for fraudulent purposes; or they may simply have crept in through errors and inaccuracies arising from manual copying or, indeed, as attempts "to enhance the appeal of the narrative" [45]. One way to detect interpolation is through comparing different manuscripts.

For example, there occurs in the Dublin version a passage describing the actions of Fergal Ua Ruairc (also written O'Ruairc or O'Rourke) of Bréifne and associate chieftains. (For the location of Bréifne, see Figure 2.1.) The Brussels manuscript, however, "omits everything connected with Fergal and his presence in the battle" [35]. As stated by Todd, "the whole story bears internal evidence of fabrication, for Fergal O'Ruairc was slain A.D. 966 ..., and our author had already set him down amongst Brian's enemies". Ryan [19], Duffy [28] and others also identify Ua Ruairc as an interpolation and Ní Mhaonaigh gives a detailed account of Bréifne bias in *Cogadh Gaedhel re Gallaibh* [45]. She states "one of the main aims of the interpolator was to portray Fergal Ua Ruairc and his followers in as favourable a light as possible, sometimes regardless of the effect this had on his text". The point is that pro-Ua Ruairc reviser of the narrative may have deemed it politically expedient to alter the record of relations between the Uí Ruairc and the Dál gCais by demonstrating assistance given by the former to Brian at Clontarf. Ní Mhaonaigh estimates the period when the Uí Ruairc were likely to have gained maximum advantage from such an association to have been the mid- to late 1140s, over a hundred years after Clontarf [45].

The possibility of interpolation applies not only to Ua Ruairc and allies. Ryan claims that "Many of the names mentioned are names only, for nothing is known of the persons who bear them. Some of the levies in important positions were certainly absent. In a word, no effort is made to distinguish between the genuine and the spurious, to criticise suspect sources, and to reconcile contradictions" [19]. Ó Corráin states that the author of *Cogadh Gaedhel re Gallaibh* "drew his material from the extant annals but he telescoped events, omitted references to other Viking leaders and concocted a super-Viking, Turgesius, whose wholesale raiding and, particularly, whose attack on Armagh was intended to demonstrate the inefficiency of the Uí Néill as defenders of the church and of the country in contrast of the achievements of the great Brian" [43]. (Turgesius is elsewhere referred to as "exaggerated" rather than "concocted" [61].) Duffy, on the other hand argues that whatever about the detail of *Cogadh Gaedhel re Gallaibh* "and its slightly cavalier approach to chronology", the gist of the account "seems sound" [28]. Duffy also discusses difficulties in using the annals to check the historicity of *Cogadh Gaedhel re Gallaibh*. By his reckoning, although some of the names of individuals drafted in from beyond Ireland are indeed suspicious, "up to half of them appear to be real and their presence at Clontarf is historically credible, if not corroborated by some other source" [28].

In Ref. [48], Ní Mhaonaigh shows that genuine annals underlie *Cogadh Gaedhel re Gallaibh* and that the compiler of *Cogadh Gaedhel re Gallaibh* "remained fairly true to his exemplar".

“Provided, therefore, that we keep the redactor’s political purpose firmly in view, we may tentatively add the annalistic material preserved in *Cogadh Gaedhel re Gallaibh* to our list of sources for information on the history of Ireland in the Viking Age” [48].

Todd himself also reports what he considers to be “curious incidental evidence” for reliability of at least some of the *Cogadh* account in that it “was compiled from contemporary materials” [35]. “It is stated in the account given of the Battle of Clontarf, that the full tide in Dublin Bay on the day of the battle (23rd April, 1014), coincided with sunrise” [35]. In a piece of “mathematical detective-work” [28] that precedes our own by 150 years, Todd’s colleague established that the full tide that morning occurred at 5:30 am and indeed coincided with sunrise. For Todd, this “proves that our author, if not himself an eye-witness, must have derived his information from those who were” [35]. We have already seen the importance of the time of the evening tide; calculated to have been at 5:55 pm, consistent with the account in *Cogadh Gaedhel re Gallaibh*; it prevented the escape of the Viking forces and considerably aided Brian’s victory.

This is certainly amongst the most striking evidence in support of the account of *Cogadh Gaedhel re Gallaibh*. Duffy provides multiple other instances where the *Cogadh* may be reliable [28]. Certainly bombastic statements that are not backed up by the annals have to be treated warily. Notwithstanding this, he considers the narrative as having “some credibility”, although unreliable” in its precise detail [28]. There is “usually a germ of truth” in its narrative and he often finds reason to give *Cogadh Gaedhel re Gallaibh* the “benefit of the doubt”. For example, although there is a tendency to assume that the author of the *Cogadh* “dreamt up much” of the detail, such as of fortifications built by Brian, Duffy argues that since these structures would still have been standing at the time of writing, the author could not get away with much invention; “the Cogadh’s audience comprised individuals in a position to contradict it if it were inaccurate” [28]. (For criticism of Duffy’s counter-revisionist views, see e.g., Ref. [30].)

Along with the Irish annals, one may compare the contents of *Cogadh Gaedhel re Gallaibh* to those of texts from other countries. There are similarities, for example to some of the content of *Njáls Saga*, which takes place in Iceland between 960 and 1020. *Njáls Saga* is believed to have been composed between 1270 and 1290 [67]. Einar Olafur Sveinsson postulated that some of the content of *Njáls Saga* may have been drawn from a now lost *Brjáns saga* (Brian’s saga) [68]. Following Sophus Bugge [37], Ó Corráin [51] suggests that *Brjáns saga* was written in about 1100 as a response to *Cogadh Gaedhel re Gallaibh*, in order to affirm the loyalty of the Hiberno-Norse Dubliners to descendants of Brian. There are also accounts of the Battle of Clontarf in Wales, France and Germany. The consistency between *Cogadh Gaedhel re Gallaibh* and these accounts is discussed in Ref. [28].

To summarise, there is a vast amount of humanities scholarship concerning *Cogadh Gaedhel re Gallaibh*. Although some dispute its reliability, others consider its version of events mainly credible and largely consistent with other sources and evidence. As stated by Duffy, "even though it is exaggerated and biased", *Cogadh Gaedhel re Gallaibh* can be useful "if we use it judiciously" and "make allowance for its propagandist tendency". The composer surely did not think in terms of network science but, in recording a cast of hundreds connected with well over a thousand links between them, he nevertheless imprinted networks in the narrative. Thus we may expect that the bulks of the networks contained in *Cogadh Gaedhel re Gallaibh* might not be too far away from the reality of the networks of the Viking Age in Ireland. Objections listed above are largely irrelevant to our approach as static networks are immune to "bombastic" descriptions, "telescoping" of events and "cavalier" attitudes to chronology. We will see that the aggregate approach is even resistant to isolated cases of interpolation. It is with this perspective that we interrogate the narrative with a networks-science methodology. To recap, our primary aim is to determine whether the character networks in *Cogadh Gaedhel re Gallaibh* are implicative of an "international contest" or "local quarrel" [29].

2.1.6 International contest or local quarrel?

O'Connor [62] in the 18th century, with Ryan [19] and Ó Corráin [43], in the 20th, are considered early debunkers of the traditional myth of Clontarf [28, 47]. O'Connor describes the conflict as a "civil war" in which "the whole province of Leinster revolted, and called the Normans from all quarters to its assistance" [62]. Ryan's main claim in Ref. [19] is that "In the series of events that led to Clontarf it was not …the Norse but the Leinstermen, who played the predominant part". His thesis is that the conflict is not a "clear-cut" one between Irish and Viking. Firstly, Brian's army was not a national one, but one of Munstermen supported by two small Connacht states. Secondly, the opposition "was not an army of Norse, but an army composed of Leinster and Norse troops, in which the former were certainly the predominant element and may have constituted two-thirds of the whole" [19]. The battle, then, was not a contest for the sovereignty of Ireland — it was not a clear-cut issue of Irish versus Norse. Instead, the issue at hand was "the determination of the Leinstermen to maintain their independence against the High-King".

It was in the course of the eleventh century, Ryan argues, that the picture of a decisive struggle of Irish history gained "universal acceptance" in the popular imagination. This came about because of the parts played by forces from the Isle of Man and the Orkney Islands together with the partisan nature of *Cogadh Gaedhel re Gallaibh*. It was only in this retrospect that Brian was acclaimed as a national hero.

Ó Corráin's view is similar [43]: "The battle of Clontarf was not a struggle between the Irish and the Norse for the sovereignty of Ireland. It was part of the internal struggle for sovereignty and was essentially the revolt of the Leinstermen against the dominance of Brian, a revolt in which their Norse allies played an important but secondary role".

Duffy points out that this revisionist interpretation is not supported by the other ancient annals. E.g., the Annals of Inisfallen gives a short but reliable account "reflective of contemporary reaction to what occurred" [28]. It is stated that "the Foreigners of Dublin gave battle to Brian" and Leinstermen are also slain. According to Duffy, "Whereas some modern historians see the Leinstermen as Brian's primary enemy at Clontarf, the annalist was in no doubt that the enemy was the Norse of Dublin. In fact he has the same black-and-white picture of the opposing sides that we tend to think of as later legend ...". "The entry in the Annals of Ulster also echoes the Annals of Inisfallen in emphasising the primacy of the Norse as Brian's adversaries". Duffy states that the Annals of Ulster suggest "it was fundamentally a contest between the Irish and Norse (although the latter too had Irish allies)".

Duffy provides multiple items of evidence in support of his view that "Brian's principle opponents were the Hiberno-Norse allied to Leinster; and [the Battle of Clontarf] was notable in particular for the great numbers of overseas Norse forces present, and for the huge losses they incurred by fighting and drowning". "Implicitly, for the Cogadh's author, two centuries of Irish opposition to Viking invasion, spearheaded by Brian's dynasty, reached a climax at Clontarf. That picture was imprinted too, with remarkable correspondences, on the minds of those thirteenth-century Icelandic writers. Those who did battle with Brian came from the Norse world seeking a kingdom for themselves in Ireland".

Thus, the debate about Clontarf has spanned the centuries and frames our present investigation. Here we broaden the question to conflictual and social relationships in the Viking Age in Ireland, and how they are reported in *Cogadh Gaedhel re Gallaibh*.

2.2 The Ossianic Poems

The second main aim of our investigations concerns the *Poems of Ossian*. More than 250 years ago, the Scottish writer, poet and politician James Macpherson published what he declared was ancient Scottish Gaelic poetry, translated into English. Macpherson claimed that the poetry dated from the "most remote antiquity" [20–23]. Although, as we shall see, the poems caused enormous controversy, they are recognised as some of "the most important and influential works ever to have emerged" [69] from the islands of Britain or Ireland. In its time, *Ossian* was widely significant, especially in regard to the development of romantic nationalism in

Europe. The controversy associated with the work (which is today regarded as one of the most famous literary controversies of all time) concerned their authenticity and its consequences rebound in Ossianic studies today [70–77]. In the 1990s the entire controversy was again brought under academic scrutiny thanks to an enormous amount of revisionist scholarship. Since then, the general topic has undergone a renaissance of sorts [69, 70, 78–92]. In this thesis a fresh investigation into Macpherson's famous work is presented, based upon the networks approach [15]. The aim of this part of the work is to compare and contrast the social-network structures of *Ossian* with other works which have been identified by humanities scholars as playing influential roles in its composition. Specifically, these are texts from Irish mythology and Homer's epics. Our main aim is to determine if the society described in Macpherson's poems bears any resemblance to those of either corpus. We will see that the Ossianic networks are quite dissimilar to those of Homer but that there is a remarkable structural similarity to the society underlying an Irish text: *Acallam na Senórach* (*Colloquy of the Ancients*) from the Fenian Cycle in Irish mythology. The significance of this finding is that it suggests a structural closeness between Macpherson's works and the Irish texts which he explicitly rejected and a distance to the classics with which he sought to draw parallels [20–22].

2.2.1 Contexts for *Ossian*

In the mid 1700s, the society of the Scottish Highlands was very different to that of the rest of Britain and indeed, Europe, both culturally and structurally as well as in terms of customs and language [82]. In 1746 the Battle of Culloden, the final showdown of the Jacobite Rising, took place there. The Jacobites were a political movement whose ambition was to restore the Catholic king James VII of Scotland, who was also James II of England and James II of Ireland, to the throne. (The name "Jacobite" comes from "Jacobus", the Renaissance Latin for "James".) James belonged to the House of Stuart, nine monarchs of which had ruled Scotland from 1371 until 1603. The Jacobites rebelled against the British government on various occasions between 1688 and 1746. The Battle of Culloden saw the defeat of Charles Edward Stuart by British forces and after their defeat the Highlands were absorbed into Great Britain. This was accompanied by efforts to suppress native Scottish Gaelic culture and the old Jacobite clans.

This was the background to which the first fragments of a Scottish epic narrative appeared. In 1760, James Macpherson published *Fragments of Ancient Poetry, Collected in the Highlands of Scotland, and Translated from the Galic or Erse Language* [93] in Edinburgh. The claim was that the *Fragments* were authored by a blind poet named Ossian, who lived in the third century AD. The poetry told of battles, struggles and unhappy love and its rhythmic narrative style created an atmosphere which captured the attention of a wide section of the public. This is how the romantic image of the Scottish Highlands, which persists in many forms to the present day, was created [82, 90] as the images in the poems seemed to connect readers to a long-gone heroic age [88]. In his short publication, Macpherson promised to follow up on the work and hoped that one day a "work of considerable length, and which deserves to be styled an heroic poem, might be recovered and translated" [93].

Macpherson visited the Scottish Highlands and Islands in the early 1760s to look for evidence of ancient Gaelic epic poetry. In December 1761, his earlier promise was fulfilled and he published *Fingal: An Ancient Epic Poem in Six Books, Together with several Other Poems composed by Ossian the Son of Fingal* [20]. (The volume was published in London and dated 1762.) In 1763, Macpherson produced with another volume: *Temora: An Ancient Epic Poem in Eight Books, Together with several Other Poems composed by Ossian the Son of Fingal* [21]. Two years later, both works were re-published in *The Works of Ossian, the Son of Fingal*. This publication included "A Critical Dissertation on the Poems of Ossian, the Son of Fingal" by Hugh Blair, who was a Professor of Rhetoric at the University of Edinburgh [22]. In 1773 a new edition, simply called *The Poems of Ossian*, was published, which claimed to be "carefully corrected, and greatly improved" [23]. In 1996, Howard Gaskill produced a new edition, *The Poems of Ossian and Related Works*, which is based on the 1765 edition [69]. The network analysis presented here is based on Gaskill's modern work. There are differences between the 1765 and 1773 editions, but these are mostly in terms of how content is arranged (the order of the poems). There are also changes in style but neither of these changes affects character interactions in the text, which are the basis for the data and the analysis presented here. Throughout this thesis, *Ossian* (in italic font) refers to Macpherson's work and Ossian (in Roman font) refers to the lead character.

2.2.2 Reception of *Ossian*

An interpretation of Macpherson's work is that it was motivated (at least partly) by an intention to undo some of the damage caused to the Scottish Highlands after the Jacobite Rising [88, 89]. The supposed discovery of epic poetry was intended to increase a sense of national identity as people could look to an ancient and noble literary heritage for Scotland. Renowned scholars such as Adam Ferguson, David Hume and Hugh Blair were keen to support claims of *Ossian*'s authenticity and credibility. Scholarly support for the Ossianic poems was especially important to create Scottish unity and to counter the fragmentation of its culture caused by Highland Clearances (the eviction during the 18th century of large numbers of tenants in the Scottish Highlands) [94]. Certainly, *Ossian* was pivotal in popularising Highland literary traditions amongst audiences beyond the north of Scotland [90].

The impact of *Ossian* beyond Scotland was enormous. It directly influenced literary figures such as Blake, Byron, Coleridge, Goethe, Scott, and Wordsworth. Brahms, Mendelssohn, and Schubert were also influenced by the work. Ossianic themes featured in paintings such as by Nicolai Abildgaard, Franois Gérard, Anne-Louis Girodet de Roussy-Trioson, Jean-August-Dominique Ingres, Angelika Kauffmann, Johann Krafft, and Philip Runge. Thomas Jefferson (principal author of the United States' Declaration of Independence and its third President) believed that *Ossian* was "the greatest poet that has ever existed" [70, 95, 96]. Napoleon kept a copy of *Ossian* in the portable library that accompanied him on his military campaigns. The town of Selma, Alabama, which became famous during the Civil Rights Movement of the 1960s, derives its name to "The Songs of Selma" in *Ossian*.

The poetry instigated a revival of interest in national literature, folklore, mythology, and poetry around Europe [81]. Accounts of the growth of romantic nationalism usually acknowledge *Ossian* as important. In particular, the work was perceived to "vindicate a fallen nation" [94] and it appealed to other countries which were subjected to conquest. It inspired these to look to their own ancient epics as reminders of a proud and illustrious past [81, 94]. The romantic imagery used in *Ossian* was extended and adapted for literary nationalism emerging in other countries [91] including the Celtic nations. In 1764, *Some Specimens of the Poetry of the Ancient Welsh Bards* by Evans [97] showed that Wales could also boast of an interesting native corpus of poetry [98]. In England, Thomas Percy published *Reliques of Ancient English Poetry* in 1775 [99]. In 1789, Charlotte Brooke produced a volume called *Reliques of Irish Poetry* [100]. The reader is referred to Bär and Gaskill's Refs. [81, 89] for a wider discussion on the reception of *Ossian* and the development of the Ossianic tradition in various European contexts.

By the 1770s, supporters of Macpherson's *Ossian* already believed it to be comparable to Homer's epics and they saw a particular similarity with the *Iliad* [92, 101]. In Ref. [101] it is suggested that Macpherson was perhaps assisted by Hugh Blair in providing comparative passages from Homeric epics in the footnotes to *Fingal*. The purpose of such footnotes was to try to draw parallels between episodes in Homer and Ossian. A large part of Blair's "Critical Dissertation" was also devoted to establishing *Ossian*'s epic character [22]. The comparison of Ossian to Homer has been an important and central element of analysis of the poetry and, in this context, the phrase "Homer of the North," emerged. The phrase has been attributed to Madame de Staël in Ref. [101]. (See also Refs. [70, 81, 90, 92] for further developments of this association.)

Macpherson's work also came in for heavy criticism, however. Some came from Samuel Johnson, the English scholar famous for compiling *A Dictionary of the English Language*. Johnson's expertise and interests were in the classical and neo-classical world centered on the civilizations of ancient Greece and Rome. Macpherson's poems formed a contrast to this world because they focussed instead on *native* literature. This focus was far more relevant for the Romantic mode of literature than was Johnson's experience. Johnson famously stated that there is a "strong temptation to deceit" in the Ossianic poems [71] and, in 1773, he went to the western islands of Scotland to inquire about their supposed origins. He argued that there were no ancient manuscripts in the local Scottish language and he did not believe that texts of *Ossian*'s length could be delivered intact from generation to generation and through word of mouth. However, his statement that Gaelic was "the rude speech of barbarous people" [71] only served to show his bias against that culture and to demonstrate widespread English hostility towards Scotland at the time [94].

During the days of Empire, British intellectuals and administrators tended to see themselves as inheritors of the classical civilisations. They saw imperialist ideologies as in line with the classical history of ancient Greece and Rome. Homer was viewed as representative of this culture and the epic narrative form was seen as reflecting its splendour. This is to be contrasted with the view by some of Scottish culture as inferior and even uncivilised. This picture allowed and justified the conquest of Scotland and its colonisation. However, the apparent discovery of an ancient epic tradition in the Highlands threatened to undermine these attitudes. As a consequence, besides representing a clash between emerging Romanticism and Classicism, the disagreements around *Ossian* inflamed tensions between national identities within Great Britain [84]. In the end, over subsequent decades, Macpherson's *Ossian* did play a central role in the rise of Romanticism as Classicism declined.

Whatever about the reaction to *Ossian* across Europe, the reaction in Ireland was one of outrage. Here, however, the offence caused was entirely different as were the objections. Irish antiquarians immediately recognised the poems as variants of their own mythology and of the tales of the Fenian Cycle in Irish mythology, in particular. They identified poorly disguised characters, place-names, and events from well-known Irish epics and accused Macpherson of forgery. Cuchullin – "the General or Chief of the Irish tribes" [93] in Macpherson was easily identified as Cúchulainn, the hero of an entirely different epoch in Irish mythology. The antiquarian and Gaelic scholar Charles O'Connor identified the character Ossian – "an illiterate Bard of an illiterate Age" for Macpherson [62] – as Oisín, the warrior-poet of the Fenian Cycle [62]. Fingal, Ossian's father, who was a third-century Scottish king according to Macpherson, was identified as a corruption Finn MacCumhail, the famous leader of a heroic Irish band, the Fianna Éireann. Sylvester O'Halloran characterised the Irish response in 1763 by denouncing what he saw as Macpherson's attempt to misappropriate Ireland's Gaelic heroes for Scotland [72] (see also Refs. [73–76]).

Matters were worsened by Macpherson's reversal of the ancient relationship between Ireland and Scotland [73, 88]. O'Halloran directly denounced *Ossian* as an "imposture" and "Caledonian plagery" [72]. ("Caledonia" is Roman name for the parts of Scotland beyond their control. It is nowadays used as a romantic or poetic name for Scotland, similar to the way "Hibernia" is used for Ireland.) O'Connor, claimed that Macpherson lacked "decency, in the illiberal abuse of all ancient and modern writers, who endeavoured to throw lights upon the ancient state of Ireland, and North-Britain" [62]. A more extensive analysis of Irish reactions to *Ossian* is given in Ref. [102].

2.2.3 "Homer of the North" or an "imposture"?

The question of whether Ossian is a "Homer of the North" or an "imposture" reflects the two sides of the early days of the debate. The majority of scholarship in the humanities has moved towards other aspects of *Ossian*, but still the questions of originality and inspiration are relevant. In this thesis, we wish to add to the complex ideas surrounding the Ossianic controversy by using the modern concept of social network analysis. Our main aim is to investigate how *Ossian* compares to the literary traditions from which it is claimed that it emerged. As stated, two comparisons were made throughout: to the Homeric epics on the one hand and to Irish mythology on the other, especially to the Fenian Cycle. While previous studies in the humanities focused on the identities of individual characters, such as Ossian, Fingal and Cuchullin, discussed above, the networks approach allows for a complementary angle, namely the set of interactions between the characters. This is our main motivation, to compare the social network structure of *Ossian* to those of Greek and Irish narratives in Chapter 5. However, we first provide further context for the comparative texts.

The Homeric epics, the *Iliad* and the *Odyssey*, are widely considered to mark the beginning of Western literature. The former is dated to the eighth century BC, during the final year of the Trojan War. It recounts a quarrel between Agamemnon, king of Mycenae and leader of the Greeks, and Achilles, their greatest hero. The *Odyssey*, in part a sequel to the *Iliad*, describes the journey home of the Greek hero Odysseus to his wife Penelope after the fall of Troy.

The Fenian Cycle of Irish mythology mostly takes place in Leinster and Munster, the eastern and southern provinces of Ireland. The tales focus on the adventures of Fionn mac Cumhail and the Fianna and there are also strong links with the Irish of Scotland. The most important source is *Acallam na Senórach* (*Colloquy of the Ancients* or *Tales of the Elders*). In the stories, Caílte mac Rónáin and Oisín are the last of the Fianna, and having survived into Christian times. They tell tales of Fionn and his warriors to the recently arrived Saint Patrick, explaining how the stories came to be preserved in written form [103].

As stated earlier, we use Gaskill's version of *Ossian* for our network analyses [69]. For the *Iliad* and *Odyssey*, we use Refs. [104] and [105], respectively. We also use a relatively recent translation of *Acallam na Senórach* [103]. Part II (*The Fianna*) of *Lady Gregory's Complete Irish Mythology* [106] is a more recent rendition of parts of the Fenian Cycle. Since it is clear that this derives from *Acallam na Senórach*, it provides a benchmark for our analysis which should be able to pick up similarities between it and the original. While other versions or translations of any of these texts may deliver differences in respect of minor characters and relationships, the large numbers of characters and links between them in each case leads one to expect that the network properties be robust with respect to small variations.

Chapter 3

Network Theory

A *network* is defined as a collection of nodes connected by edges. The nodes are sometimes called vertices and they are interconnected by the edges, which are also called links. In mathematics, the word "network" is almost synonymous with "graph" and these terms will be used interchangeably from here on. (Technically a network is more general than a graph because it can entail multiple types of links which can be time dependent. However we do not deal with such objects here.) Regular lattices and Erdős-Rényi random graphs are examples of networks. Complex networks that are encountered in many situations are neither regular nor random. Graphs are an abstract representation of real networks in which interactions between elements lead to the emergence of non-trivial topological features. Examples include social or transport networks in which nodes may represent people or stations and edges represent social interaction or journeys. In this manuscript we consider graphs generated from epic-type narratives, each node representing a character, while each edge is the representation of interactions between characters.

A graph is defined by a pair $G = (V, E)$ where V is a finite set of vertices and where E is the set of edges between the vertices. The total number of nodes $|V|$ in a network is denoted by N and the total number of edges $|E|$ is denoted by M. Nodes v_i and v_j are said to be adjacent if an edge $e_{i,j} = (v_i, v_j)$ exists between them, that is $e \in E$. A graph can be directed or undirected, the former being defined when an orientation is associated with each edge. In an undirected graph, edges connect unordered pairs of nodes, that is $(v_i, v_j) \in E \Leftrightarrow (v_j, v_i) \in E$. The networks that we envisage here are undirected. In addition, our edges are anti-reflexive i.e $(v_i, v_i) \notin E$ for all v in V. This means that an edge cannot link a node to itself. It is sometimes useful to assign a weight to each edge, representing relative strength between nodes e.g., frequency of interaction. A graph containing weighted edges is called a weighted network.

Statistical quantities have been used to characterise the structure of complex networks, and in this chapter we discuss some of these network measures in detail. In the next section, we present the topology of a network in the form of an adjacency matrix and some network properties that includes but not limited to degree, degree distribution, assortativity, clustering coefficient, path length, closeness, eigenvalue and betweenness centralities.

3.1 Adjacency Matrix, Mean Degree and Density

The information on a network structure can be stored in the form of an $N \times N$ matrix called the *adjacency matrix* often denoted by A. The *adjacency matrix* with elements $A_{i,j}$ is defined by

$$A_{i,j} = \begin{cases} 1 & \text{if } (v_i, v_j) \in \text{E} \\ 0 & \text{otherwise.} \end{cases}$$

The adjacency matrix is symmetric ($A_{i,j} = A_{j,i}$) with zeros along the main diagonal, since we are only considering undirected networks with no self-loops.

In graph theory, local and global network properties help to identify the most important nodes and to understand the structure of a network. The most basic localised structural property of a network is the *degree* denoted by k_i for node i. It measures the number of edges incident to node i and is given by

$$k_i = \sum_{j=1}^{N} A_{i,j}. \tag{3.1.1}$$

Its maximum value $k_{\max}$ identifies the most connected node $v_{\max}$ and the various values of k_i can be used to rank nodes. The *mean degree* of the network is defined by

$$\langle k \rangle = \frac{1}{N} \sum_{i=1}^{N} k_i = \frac{2M}{N}, \tag{3.1.2}$$

where the factor 2 in Eq.(3.1.2) accounts for the fact that there are two nodes at the ends of each edge.

In addition to the network's mean degree, we are also interested in quantifying the extent to which a graph is close to being a complete graph. A complete graph is when every pair of vertices is connected by an edge. The *network density* D is the ratio of edges that are contained in G to the possible maximum number of such edges. It is normalised to give a value of 1 for a completely connected graph and it is defined as

$$D = \frac{2M}{N(N-1)} = \frac{\langle k \rangle}{N-1}, \tag{3.1.3}$$

where $0 \leq D \leq 1$.

3.2 Degree Distribution

Besides the mean, variance and various moments, one is also interested in the spread of degrees between the nodes of the network. This is its *degree distribution*. If N_k is the number of nodes with degree k, the probability that any given node is of this type is

$$p_k = \frac{N_k}{N}. \tag{3.2.4}$$

Therefore p_k gives the probability for a randomly chosen node to have k links. The convention is to display the complementary cumulative degree distribution $P(k)$ rather than $p(k)$ itself. This is because the latter can be rather noisy, especially in the tails of the distribution. The preferred quantity $P(k)$ is the frequency of nodes whose degree is greater than or equal to the given value k.

It is also common to talk of the degree distribution when referring to an idealised probability distribution from which a given graph G is generated. Both quantities, the degree frequency and degree distribution, share the same properties defined below.

The distribution is normalised and satisfies

$$\sum_{k=0}^{\infty} p_k = 1. \tag{3.2.5}$$

The associated generating function defined by

$$g(z) = \sum_{k=0}^{\infty} z^k p_k \tag{3.2.6}$$

and is such that $g(1) = 1$. The first derivative,

$$g'(z) = \sum_{k=0}^{\infty} z^{k-1} k p_k, \tag{3.2.7}$$

generates the *first moment* or mean of the distribution

$$g'(1) = \sum_{k=0}^{\infty} k p_k \equiv \langle k \rangle. \tag{3.2.8}$$

Similarly, for the *second moment* of the degree distribution we have

$$\langle k^2 \rangle = \sum_{k=0}^{\infty} k^2 p_k = g''(1) + g'(1). \tag{3.2.9}$$

This is useful in constructing the variance about the mean defined by

$$\langle (k - \langle k \rangle)^2 \rangle = \langle k^2 \rangle - \langle k \rangle^2 = g''(1) - g'(1)(g'(1) - 1). \tag{3.2.10}$$

In general, the n^{th} moment is defined as

$$\langle k^n \rangle = \sum_{k=0}^{\infty} k^n p_k = \left[\left(z\frac{d}{dz} \right)^n g(z) \right]_{z=1}. \tag{3.2.11}$$

The higher moments can be used to quantify other features of the distribution such as its skewness (third moment), kurtosis (fourth moment) etc. *Skewness* is a measure of the shape asymmetry of a probability distribution and can be positive or negative. *Kurtosis* is another shape metric and quantifies the "tailedness" of a distribution. Degree distributions which have kurtosis less than that of the Gaussian distribution are called *platykurtic* and generate fewer (and less extreme) outliers (they have thin tails). *Leptokurtic* distributions, on the other hand, have tails that approach zero slower than a Gaussian (they have fat tails).

3.2.1 Degree Distribution of Complex Networks

The degree distributions of complex networks tend to be right-skewed, meaning there is a small probability for high degree nodes and higher probability for low degree nodes [107]. Complex networks with leptokurtic (fat-tailed) degree distributions are sometimes described by a *power-law*,

$$p(k) = \frac{C}{k^\gamma}, \tag{3.2.12}$$

where γ and C are constants. The shape of such a distribution is unaffected by increasing the degree k by a factor of c. That is

$$p(ck) = C(ck)^{-\gamma} = Cc^{-\gamma} p(k) \propto p(k).$$

Hence, networks whose node degrees follow a power-law distribution, are sometimes referred to as scale-free networks meaning they are scale invariant.
There are however restrictions to the possible values of γ. In Eq.(3.2.12) the distribution diverges at $k = 0$, which is impossible for a finite system therefore, there must be some minimum degree $k = k_{\min}$ above which the power-law behaviour is observed.
Real-world systems have a finite number of nodes N, and a finite number of edges M. As a consequence, all moments $\langle k^n \rangle$ of the distribution $p(k)$ are finite. This is in contradiction with predictions from an exact power-law distribution as we shall now see. First, summing over all k in Eq.(3.2.12), leads to $C = 1/\zeta(\gamma)$, where ζ is the Riemann zeta function which diverges for $0 < \gamma \leq 1$. Using Eq.(3.2.11), we see that

$$\langle k^n \rangle = \frac{\zeta(\gamma - n)}{\zeta(\gamma)}. \tag{3.2.13}$$

Hence, for a given γ it is always possible to find a value of n for which $\langle k^n \rangle$ diverges (if $0 < \gamma - n < 1$). It follows that the degree distribution can only be best fitted by a power-law function between some lower bound $k_{\min}$ and some upper bound $k_{\max}$. Often networks have a cut-off beyond which a simple power-law no longer describes their degree distributions [108, 109]. Such a cut-off sets a scale and, if it occurs at $k = \kappa$, then the distribution can be described by a *truncated power-law* given by

$$p_k \sim k^{-\gamma} e^{-k/\kappa}. \tag{3.2.14}$$

It follows that the cumulative degree distribution has the form

$$P(k) \sim k^{1-\gamma} e^{-k/\kappa}. \tag{3.2.15}$$

For example, a truncated power-law distribution was observed in networks generated from human mobility patterns. Gonzalez *et. al* modelled travel patterns for 100 000 mobile phone users by capturing each user's position between consecutive calls [110]. The distribution of displacements measured in kilometres was found to follow a truncated power-law with an exponent of 1.75 ± 0.15 and a cut-off value of 400 km.

Another distribution generally found in complex networks is the *exponential distribution*, given by

$$P(k) \sim e^{-k/\kappa}, \tag{3.2.16}$$

with parameter $\kappa > 0$ setting the scale of the distribution [107]. Other exponential functions often met in the literature include the *normal distribution* also called the Gaussian, of the form

$$P(k) \sim \exp\left[-\frac{(k-\mu)^2}{2\sigma^2}\right], \tag{3.2.17}$$

where μ and σ^2 are the mean and variance of the distribution, respectively. Additionally the *log-normal* is more skewed than the normal distribution and takes the form

$$P(k) \sim \frac{1}{k}\exp\left[-\frac{(\ln k-\mu)^2}{2\sigma^2}\right]. \tag{3.2.18}$$

The *stretched exponential* is given by

$$P(k) \sim e^{-(k/\kappa)^\beta}. \tag{3.2.19}$$

Setting $\beta = 1$ recovers the exponential distribution function Eq.3.2.16 and $\beta = 2$ gives the Gaussian distribution. The stretched-exponential is also the complementary cumulative function of the *Weibull* [111]. The Weibull is described by

$$P(k) =\sim \left(\frac{k}{\kappa}\right)^{\beta-1} e^{-(k/\kappa)^\beta}. \tag{3.2.20}$$

3.2.2 Parameter Estimators

We can fit the data set to each of the functional forms given in Sub-section 3.2.1. Using the method of maximum likelihood estimators (MLE), we then determine best estimates of the parameters for each probability distribution. The likelihood $\mathcal{L}$ is the probability that independent and identically distributed data of N observations were drawn from a distribution $p(k)$ with parameter θ,

$$\mathcal{L}(\theta|k) = \prod_{i=1}^{N} p_\theta(k_i). \tag{3.2.21}$$

The estimated parameters are those that maximise the probability of obtaining the given distribution. The logarithm of the likelihood is generally used instead of the likelihood itself because they share the same maximising values of the parameters.

The *Akaike Information Criterion* (AIC) is then used to determine the most likely model that best describes the data [112]. If a model estimates q parameters for a data set, the AIC is given by

$$\mathrm{AIC} = \frac{2qn}{n-q-1} - 2\ln\mathcal{L}, \tag{3.2.22}$$

where n is the sample size.

To compare between various models, one relies on the relative probability that a given model minimises the estimated information loss given by $\exp\left[(\mathrm{AIC}_{\min} - \mathrm{AIC})/2\right]$ where $\mathrm{AIC}_{\min}$ is the minimum AIC value for a given set of data.

3.2.2.1 Maximum Likelihood Estimators

We can estimate the probability distribution parameters using definitions given in Sub-section 3.2.2 above.

Power-law distributions

Taking Δ to represent the unknown distribution prior to $k_{\min}$, Eq. (3.2.12) can be normalised to obtain

$$p_k = \frac{1-\Delta}{\zeta(\gamma)} k^{-\gamma}. \tag{3.2.23}$$

Taking the likelihood on Eq.(3.2.23) yields

$$\ln\mathcal{L} = N\ln(1-\Delta) - N\ln(\zeta) - \gamma\sum_{i=1}^{N}\ln k_i. \tag{3.2.24}$$

The approximate value of γ can be obtained numerically however, [113] gives the following approximation

$$\hat{\gamma} = 1 - N\left[\sum_{i=1}^{N}\ln\frac{k_i}{k_{\min}-\frac{1}{2}}\right]^{-1}. \tag{3.2.25}$$

Truncated Power-law Distribution

For the truncated power-law, we can normalise Eq.(3.2.14) to get

$$p(k) = (1 - \Delta)\frac{\mathrm{e}^{\mathrm{k_{min}}/\kappa}}{Z(k_{\mathrm{min}})}k^{-\gamma}\mathrm{e}^{-\mathrm{k}/\kappa}, \tag{3.2.26}$$

where $Z(k) \equiv \sum_{m=0}^{\infty}(k+m)^{-\gamma}\mathrm{e}^{-\mathrm{m}/\kappa}$. The log likelihood becomes

$$\ln\mathcal{L} = N\ln(1-\Delta) + \frac{Nk_{\mathrm{min}}}{\kappa} - N\ln Z(k_{\mathrm{min}}) - \sum_{i=1}^{N}\left(\gamma\ln k_i\right) + \frac{k_i}{\kappa}, \tag{3.2.27}$$

which can be maximised numerically to obtain the value of $\hat{\gamma}$.

Stretched Exponential Distribution

From Eq.(3.2.19) it follows that the cumulative degree distribution of the stretched exponential can be expressed as

$$P(k) \sim \left(\frac{k}{\kappa}\right)^{\beta-1}\mathrm{e}^{-(\mathrm{k}/\kappa)^{\beta}}. \tag{3.2.28}$$

The log likelihood using Eq.(3.2.19) is therefore given by

$$\ln\mathcal{L} = N\ln(1-\Delta) - N\ln\left[\sum_{k_{min}}^{\infty}\mathrm{e}^{-\left(\frac{\mathrm{k_{min}}}{\kappa}\right)^{\beta}}\right] - \sum_{\mathrm{i}=1}^{\mathrm{N}}\left(\frac{\mathrm{k_i}}{\kappa}\right)^{\beta}. \tag{3.2.29}$$

Normal Distribution

The log likelihood of the *normal* or *Gaussian* distribution in Eq.(3.2.17) is given by

$$\ln\mathcal{L} = N\ln(1-\Delta) - N\ln\left[\sum_{k_{min}}^{\infty}\mathrm{e}^{-\frac{\left(\mathrm{k_{min}}-\mu\right)^2}{2\sigma^2}}\right] - \sum_{\mathrm{i}=1}^{\mathrm{N}}\frac{(\mathrm{k_i}-\mu)^2}{2\sigma^2}. \tag{3.2.30}$$

Additionally, the log likelihood for the *log-normal* is expressed by

$$\ln\mathcal{L} = N\ln(1-\Delta) - N\ln\left[\sum_{k_{min}}^{\infty}\frac{1}{k_{min}}\mathrm{e}^{-\frac{\left(\ln\mathrm{k_{min}}-\mu\right)^2}{2\sigma^2}}\right] - \mathrm{N}\sum_{\mathrm{i}=1}^{\mathrm{N}}\ln\mathrm{k_i} - \sum_{\mathrm{i}=1}^{\mathrm{N}}\frac{(\ln\mathrm{k_i}-\mu)^2}{2\sigma^2}. \tag{3.2.31}$$

Weibull Distribution

Applying the log likelihood on Eq.(3.2.20) for the Weibull distribution yields

$$\ln\mathcal{L} = N\ln(1-\Delta) - N\ln\left[\sum_{k_{min}}^{\infty}\left(\frac{k_{min}}{\kappa}\right)^{\beta-1}\mathrm{e}^{-\left(\frac{\mathrm{k_{min}}}{\kappa}\right)^{\beta}}\right] - \mathrm{N}(\beta-1)\ln\kappa + (\beta-1)\sum_{\mathrm{i}=1}^{\mathrm{N}}\ln\mathrm{k_i} - \sum_{\mathrm{i}=1}\left(\frac{\mathrm{k_i}}{\kappa}\right)^{\beta}. \tag{3.2.32}$$

Estimates for the parameters of the model distributions in Sub-section 3.2.1 are obtained by numerically maximising the log-likelihood.

3.3 Assortativity

In sociology, *homophily* is the tendency for people to associate with similar people. In networks with this type of feature, nodes tend to connect with other nodes that are similar in some way. In network theory, this tendency is measured using *assortativity*. There are two types of assortativity measures: *scalar* and *categorical*. Scalar characteristics, can be measured and ordered — e.g., age, weight, height or popularity. Categorical features, by contrast, cannot be ordered or ranked and they lack an associated metric. Examples include gender, race, or nationality.

Degree assortativity r is an example of *scalar assortativity*. It quantifies the tendency of nodes whose degrees have similar values to associate with each other. In determining r, it is important to account for the possibility that nodes may have similar but not identical values; e.g., high degree nodes may tend to mix with other high degree nodes without them having to have precisely the same k-values.

In the categorical case, one is interested in whether two nodes have the same attributes or not. The notion of degrees of similarity is absent here. In categorically assortative networks, there is a tendency for nodes belonging to the same category to link to each other. Our aim is to quantify statements such as this. Because we have two different types of assortativity, we require two different formulae to quantify them.

Scalar assortativity can be measured using Pearson's correlation coefficient. This is the covariance of two variables divided by the product of their standard deviations. This normalisation guarantees that the scalar assortativity takes values which lie within the range $[-1, 1]$. Networks with positive values of degree assortativity (i.e., for which $r > 0$) are deemed to be *degree assortative*. If, on the other hand, r is negative, the network is considered to be *degree disassortative*. Since the scalar assortativity is limited by the same theoretical bounds (i.e., -1 and 1) for all networks, comparisons are straightforward. It is meaningful to state that one network is more (or less) scalar assortative than another.

There is a tendency for networks to evolve towards their maximum-entropy (most disordered) state unless there is some sort of "force" which prevents this from happening [114]. The interactions between people form an example of such a "force" in social circumstances. The homophilic tendency of people to gravitate (i.e., to form links with) to others who are like themselves is an example of such a "force" [107, 115]. The maximum-entropy configurations encountered in networks without such a force are usually disassortative because disassortative states are more plentiful than assortative ones [116]. For this reason, networks which have no organising pressures (e.g., many non-social networks) are usually disassortative by degree. Many social networks, on the other hand, are either assortative or else they are uncorrelated in this sense (i.e., they are neither assortative or disassortative). It has been found that this property is also contained in many character networks for narratives of the epic genres, especially if one restricts interactions to so-called positive (i.e., friendly) ones [8, 9, 15, 16]. This may be considered to reflect the presence of a non-trivial social or "narrative force" — driving the networks away from their maximum-entropy, disassortative states.

Categorical assortativity differs from scalar assortativity in a number of ways. Firstly it does not carry the notion of *relative similarity*: two vertices either possess the same feature or they do not. The feature under question may be thought of as a colour. We cannot say that green and red, for example are closer to each other than, say orange and purple. Secondly, the standard definition of categorical assortativity does not usually extend over the full range $[-1, 1]$. This is another way in which it differs from the Pearson-correlation-coefficient measure of scalar assortativity. Denoting catagorical assortativity by ρ, it turns out that it ranges from a theoretical minimum value ρ_{min} to a theoretical maximum of 1. Here ρ_{min} is a value which itself lies somewhere between -1 and 0. A network is categorically assortative if ρ is bigger than 0 and it is categorically disassortative if $\rho < 0$ but bigger than ρ_{min}. Therefore, while it is easy to compare the ρ-values for two different assortative networks, one has to be quite careful when looking at disassortative ones. Two different types of disassortative network may have two different values of ρ_{min}, so it becomes difficult to compare them. A low value of ρ for one network may not be so low for another. In other words, the assortativity metric ρ is not particularly good when comparing two networks which are disassortative in some categorical feature.

In the next section, we discuss categorical assortativity in some more detail. After giving the standard measure for it, we introduce a new quantity $\bar{\rho}$ which measures *disassortativity* directly, instead of considering it merely as the negative of assortativity. The metric $\bar{\rho}$ is developed in such a way as to take the value 0 for uncorrelated networks and the value -1 for fully disassortative networks. In a sense, therefore, it is more suitable than ρ for comparing two categorically disassortative networks. This will become important when we come to investigate conflict in certain character networks. The disassortativity measure $\bar{\rho}$ is as suitable for measuring scalar disassortativity as ρ is for measuring assortativity. As we shall see, there is a simple relationship between $\bar{\rho}$ which will enable us to invent a measure that would be good for both circumstances.

3.4 Categorical Assortativity

Let N and M denote the numbers of nodes and edges, respectively. The nodes are endowed with a categorical feature, which, for concreteness we can think of as colour. In constructing the assortativity coefficient, the aim is to measure the difference between the fraction of edges that exist between nodes of the same attribute and the fraction of such edges we would expect if the nodes were connected at random regardless of the nodes' attributes (i.e., if the linking process were "colour blind"). Here, we only consider undirected networks, although one could extend to include directed networks too.

The fraction of edges between nodes of the same type (colour) is

$$\frac{1}{M}\sum_{(i,j)}\delta_{c_i,c_j} = \frac{1}{2M}\sum_{i,j}A_{i,j}\delta_{c_i,c_j}, \tag{3.4.33}$$

where the sum on the left hand side is over all edges of the network while that on the right hand side is over each pair of nodes i, j and where $A_{i,j}$ is an element of the adjacency or connection matrix for the graph in question. This is one if there is an edge between nodes i and j and zero otherwise. Here, δ_{c_i,c_j} is the Kronecker delta function which is 1 if $c_i = c_j$ and 0 if $c_i \neq c_j$. If we have an undirected network, every edge i, j is counted twice, and hence the factor $1/2$ on the right-hand side.

If nodes were linked at random (i.e., in a colour-blind manner), the probability that i and j are connected is independent of their category but dependent on their degrees k_i and k_j. The number of ways available to connect them is k_ik_j and the total number of such pairs is $(2M)^2$ so that the adjacency matrix is $k_ik_j/(2M)^2$. Therefore the fraction of edges between nodes of the same colour if linking is performed in a colour-blind manner is

$$\frac{1}{2M}\sum_{i,j}\frac{k_ik_j}{2M}\delta_{c_i,c_j}. \tag{3.4.34}$$

The *modularity* is defined as the difference between the actual proportion of similar-type edges from Eq.(3.4.33) and the expected frequency of colour-blind edges from Eq.(3.4.34), and is

$$Q = \frac{1}{2M}\sum_{i,j=1}^{N}\left(A_{i,j} - \frac{k_ik_j}{2M}\right)\delta_{c_i,c_j}. \tag{3.4.35}$$

Because of the Kronecker delta function in Eq.(3.4.35), the summation is effectively over edges connecting nodes of the same colour. The next step to replace the summation over nodes by one over colours.

The total degree of the network is $\sum_{i=1}^{N} k_i = 2M$ (twice the number of edges because each edge is double-counted). It is convenient to introduce $2M$ *chevrons* or oriented markers as a visual representation of these degrees. Each chevron is associated with the end of an edge and each is endowed with a colour, namely that of the nearest node. Let c and c' denote categorical variables which take their values from the set of node colours. Let $E_{cc'}$ denote the number of chevrons in the network pointing from nodes of colour c to nodes of colour c'. We note that $E_{cc'} = E_{c'c}$ if the network is undirected. If the two chevrons associated with an edge are of the same colour, we say that the entire edge is of that colour. Thus E_{cc} edges are of colour c. If the two chevrons at the ends of an edge are of different colour, we say that edge is colourless. We define the chevron density as

$$e_{cc'} = \frac{E_{cc'}}{2M}. \tag{3.4.36}$$

The quantities E_{cc} count the colour-c edges so that

$$E_{cc} = \sum_{i,j} A_{i,j}\delta_{c_i c}\delta_{c_j c}. \tag{3.4.37}$$

Summing over the colours,

$$\sum_c E_{cc} = \sum_{i,j} A_{i,j}\delta_{c_i,c_j}. \tag{3.4.38}$$

This is the first term in Eq.(3.4.35).
To express the second term in Eq.(3.4.35) as a sum over colours, we first define the total number and density of chevrons of colour c (the number of colour-c chevrons is simply the total degree of nodes of that colour):

$$A_c = \sum_{c'} E_{cc'} = \sum_i k_i\delta_{c_i c}, \qquad a_c = \frac{A_c}{2M} = \sum_{c'} e_{cc'} = \frac{1}{2M}\sum_i k_i\delta_{c_i c}. \tag{3.4.39}$$

We have the sum rule

$$\sum_c a_c = \sum_{cc'} e_{cc'} = 1. \tag{3.4.40}$$

To express the second term in Eq.(3.4.35) as a sum over colours, we write

$$-\sum_{i,j} \frac{k_i k_j}{(2M)^2}\delta_{c_i,c_j} = -\sum_c \left(\sum_i \frac{k_i}{2M}\delta_{cc_j}\right)\left(\sum_j \frac{k_j}{2M}\delta_{cc_j}\right) = -\sum_c a_c^2,$$

having used Eq.(3.4.39). We finally are in a position to express the modularity of Eq.(3.4.35) as a sum over colours [117]:

$$Q = \sum_c \left(e_{cc} - a_c^2\right). \tag{3.4.41}$$

The categorical assortativity ρ is obtained by normalising the modularity so that its maximum value is 1 (as is the case for the scalar assortativity). If the network is fully assortative, *all* edges

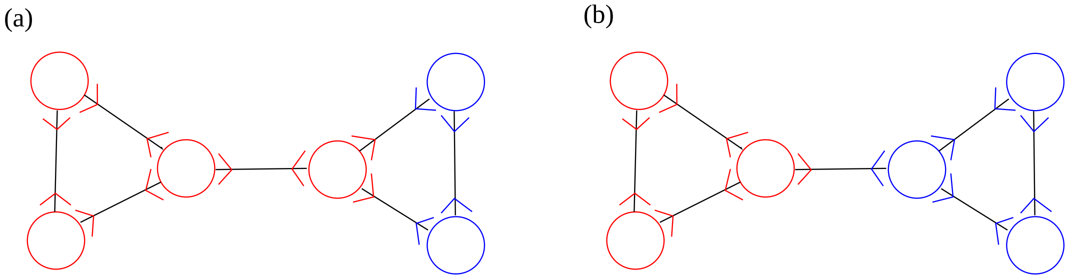

Figure 3.1: Two networks, (a) and (b) each with 6 nodes and 7 edges. The colours of the chevrons indicate the type of the node associated with that edge and these are oriented towards the partner node. The assortativity values are $\rho = 0.3$ and $\rho = 0.714$, respectively.

are coloured; they all connect nodes of the same colour. Therefore the normalising factor for Q is given by Eq.(3.4.41) with $\sum_c e_{cc} = 1$ set to 1. This suggests the definition

$$\rho = \frac{\sum_c \left(e_{cc} - a_c^2\right)}{1 - \sum_c a_c^2}. \tag{3.4.42}$$

The minimum value that this quantity can attain is achieved when all edges are colourless ($e_{cc} = 0$ for all c). It is

$$\rho_{\min} = \frac{-\sum_c a_c^2}{1 - \sum_c a_c^2}. \tag{3.4.43}$$

We will shortly see that $\rho_{\min}$ can be -1 if there are only two categories and if the network is fully disassortative. Otherwise it will be in the range $-1 \leq \rho_{\min} < 0$ [118].

An example is given in Figure 3.1(a) which has $N = 6$ nodes, $M = 7$ edges, and $2M = 14$ chevrons. (Both examples in Figure 3.1, but without the chevrons, are borrowed from Ref. [117].) With $c \in \{r, b\}$ (r and b standing for red and blue, respectively), we have

$$E_{rr} = 8, E_{rb} = E_{br} = E_{bb} = 2, \quad \text{or} \quad e_{rr} = \frac{4}{7}, e_{rb} = e_{br} = e_{bb} = \frac{1}{7}. \tag{3.4.44}$$

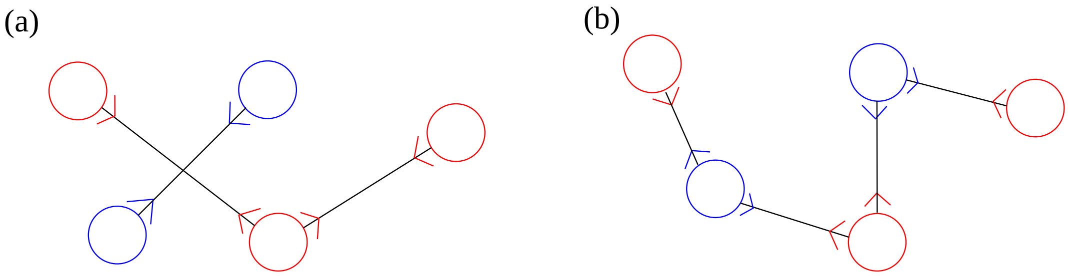

Figure 3.2: (a) and (b) are fully assortative and fully disassortative two-colour networks. The values of their scalar assortativities are $\rho = 1$ and $\rho = -1$, respectively.

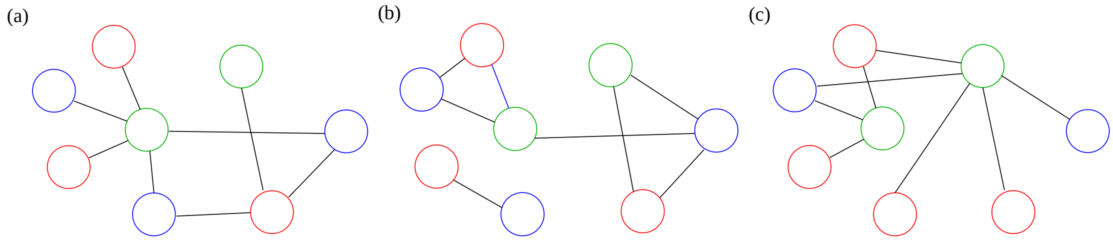

Figure 3.3: Three fully disassortative networks, each with three colours. Network (a) and (c) have the same degree distributions but their assortativities ρ differ. Network (a) and (b) have different degree distributions but the totals of their degrees for different colours are distributed similarly so they have the same assortativity value.

So the densities of red and blue edges are, respectively, $e_{rr} = 4/7$ and $e_{bb} = 1/7$. Also,

$$A_r = 10, \quad A_b = 4 \quad \text{and} \quad a_r = \frac{5}{7}, \quad a_b = \frac{2}{7}.$$

Eq.(3.4.42) then gives $\rho = 3/10 = 0.3$ and the network is assortative. A similar calculation for Figure 3.1(b) gives $\rho = 5/7 \approx 0.714$. Since the theoretical bounds on positive assortativity for both networks in Figure 3.1 are the same, i.e., between 0 and 1, we can say that the network in Figure 3.1(b) is more assortative than that in Figure 3.1(a).

An example of a fully assortative network is given in Figure 3.2(a). This has $M = 3$ edges and the total numbers of red and blue chevrons are $A_r = 4$ and $A_b = 2$, respectively. Their proportions are therefore $a_r = 2/3$ and $a_b = 1/3$. Moreover, $E_{rr} = 4$, $E_{bb} = 2$ and $E_{rb} = E_{br} = 0$, so that $e_{rr} = 2/3$, $e_{bb} = 1/3$, and $e_{rb} = e_{br} = 0$. The fact that $e_{rr} + e_{bb} = 1$ means that all links are red or blue and ensures that $\rho = 1$ from Eq.(3.4.42).

An example of a fully disassortative network is given in Figure 3.2(b). Here we have $M = 4$, $A_r = A_b = 4$ so that $A_r = A_b = 1/2$. There are no coloured edges so that $E_{rr} = E_{bb} = 0$. The undirectedness of the network ensures symmetry between the two colours, giving $E_{rb} = E_{br} = 4$. Normalised, these give $e_{rr} = e_{bb} = 0$, $e_{rb} = e_{br} = 1/2$. Eq.(3.4.42) then gives $\rho = -1$. It is a general result that a fully disassortative, undirected network with two colours has $\rho_{\min} = -1$. However, this minimum value for ρ is misleading; it is not generally -1 if more colours are involved. While the absence of assortativity means that $\sum_c e_{cc} = 0$ for any number of colours, the undirectedness that assures the symmetry between the colours only happens when there are two of them. This property, together with $a_r + a_b = 1$ from Eq.(3.4.40) means that $a_r = a_b = 1/2$. Eq.(3.4.42) then trivially gives $\rho = -1$.

The three networks in Figure 3.3 illustrate that the result $\rho = -1$ does not carry over to fully disassortative networks with more than two colours. Each network in the figure is fully

disassortative because no edge links two nodes of the same colour: $E_{rr} = E_{bb} = E_{gg} = 0$ (here the green nodes are labelled g and we omit the chevrons from the illustrations from now on). Each network has $M = 8$. Networks (a) and (c) have the same degree distributions given by $16p(k) = 4, 2, 1, 0, 1$ for $k = 1, 2, 3, 4$ and $k = 5$, respectively. Here, $p(k)$ is the probability (relative frequency) of nodes with degree k. Network (b) has $16p(k) = 2, 4, 2, 0, 0$ and network 3 has $16p(k) = 4, 2, 1, 0, 1$. Network (a) has $A_r = 5$, $A_b = 5$, $A_g = 6$; network (b) has $A_r = 5$, $A_b = 6$, $A_g = 5$; and network (c) $A_r = 4$, $A_b = 4$, $A_g = 8$.
Although networks (a) and (b) have different degree distributions, because the totals of their degrees for different colours have the same distributions, they have the same assortativity values, namely $\rho = -86/170 = -0.506$. Although network (c) has the same degree distribution as network (a), because the totals of their degrees for different colours differ, they have the same assortativity values; that of network (c) is $\rho = -3/5 = -0.6$. Moreover, although each of the three networks is fully disassortative, none of their ρ values is -1.
Fully disassortative, undirected networks with only two categories have $\rho_{\min} = -1$. However, the minimum value for ρ is not generally -1 if more categories are involved. While the absence of assortativity means that $\sum_c e_{cc} = 0$ for any number of categories, the lack of directedness that assures the symmetry between the categories only happens when there are two of them. This property, together with Eq.(3.4.40) trivially gives $\rho = -1$. More generally, $\rho_{\min}$ lies between -1 and 1.
The reason why $\rho_{\min} \neq -1$ for a fully disassortative network with three or more colours is that such a network more closely resembles a random network than a perfectly assortative one does [118]. When many colours are available, random colour-blind mixing will tend to connect nodes of different colours. It is for this reason that the value $\rho = 0$ for a random network is closer to the value of a fully disassortative network than it is to that of a perfectly assortative network. For this reason it is more difficult to compare disassortative networks using ρ than it is to compare assortative ones.
In developing Eq.(3.4.42) for the modularity, one seeks to compare to the expected density of edges between nodes of the same category in the circumstance that the network is assembled without regard to category. This is an appropriate approach when it is that *assortativity* that we seek to measure. However, as we have seen, this leads to a bad normalisation for disassortative networks. To develop a measure that is nicely normalised in the latter case, one may seek to directly measure *disassortativity* instead. To this end, we focus on edges between node of *different* categories. We start by introducing

$$\bar{\rho} = -\frac{\sum'_{c,c'} (e_{cc'} - a_c a_{c'})}{1 - \sum'_{c,c'} a_c a_{c'}}. \tag{3.4.45}$$

Here the prime on the summation signals that it runs over values of c and c' that are not equal ($c = c'$ is omitted). Also, the leading minus sign is introduced in order to make sure that disassortative networks have negative values of $\bar{\rho}$. This is to ensure that we keep the same sign conventions as in the ρ-cases.
Eq.(3.4.40) gives

$$\sum\nolimits'_{c,c'} e_{cc'} = 1 - \sum_c e_{cc}.$$

This allows us to write

$$\bar{\rho} = \rho \left(\frac{1}{\sum_c a_c^2} - 1 \right). \tag{3.4.46}$$

From Eq.(3.4.43), this may simply be written

$$\bar{\rho} = -\frac{\rho}{\rho_{\min}}, \tag{3.4.47}$$

so that $\bar{\rho}$ is just the standard expression for the assortativity, but normalised by its minimum value (which is, of course, negative).
We can also introduce a renormalised version of the categorical assortativity that is more suitable for all circumstances:

$$\hat{\rho} = \begin{cases} \rho & \text{if } \rho > 0, \\ -\frac{\rho}{\rho_{\min}} & \text{if } \rho < 0. \end{cases} \tag{3.4.48}$$

This quantity is 1 in the case of a fully assortative networks; it is −1 if the network is fully disassortative; and is 0 in the case of a colour-blind network. Therefore it has all the features that we desire, including a normalisation which facilitates easy and meaningful comparisons of categorical assortativities between networks.

3.5 Clustering Coefficient

Another network topological measure is *clustering coefficient*. This measures the concentration of edges in a vertex neighbourhood. The maximum potential number of edges between node i and its nearest neighbours is $k_i(k_i - 1)/2$. If n_i is the number of links that actually exist between node i and its neighbours, we define the clustering coefficient of node i as

$$C_i = \frac{2n_i}{k_i(k_i - 1)}. \tag{3.5.49}$$

The *average clustering coefficient* C for the entire network is the mean of these quantities taken over all nodes of the network and is bound between 0 and 1:

$$C = \frac{1}{N} \sum_{i=1}^{N} C_i. \tag{3.5.50}$$

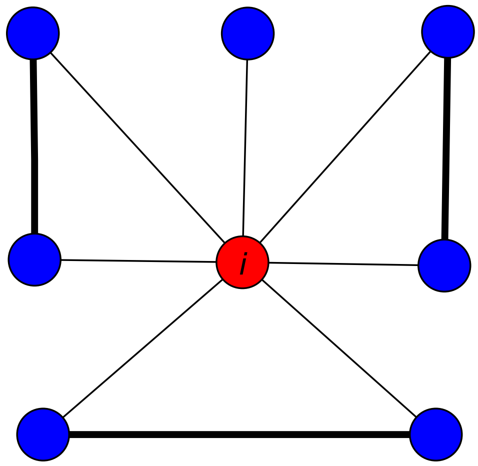

Figure 3.4: A schematic graph showing central node i with seven of its neighbours.

This measure of clustering coefficient was proposed by Watts and Strogatz in [13] and we use Figure 3.4 to illustrate its application. In the figure, the central node i has degree $k_i = 7$. Only three edges connect the neighbours of node i. Therefore, $C_i = 3/21 = 1/7$. The clustering coefficient of the entire network is

$$C = 0.77$$

It can be noticed however, from Eq.(3.5.49) that clustering coefficient is biased towards low k_i values because of the denominator $k_i(k_i - 1)$. An alternative measure of clustering coefficient, sometimes called *transitivity* C_T minimises this bias. It is given by Newman et al.(2002) and can be quantified as follows. If node a connects with b and node b connects with c, then abc makes a path of two edges. If a also connects with c then a triangle forms. In this case, the path between a, b and c forms a closed triad representing a transitive relationship. Transitivity is defined by

$$C_T = \frac{3N_\Delta}{N_t},$$

where N_Δ is the number of triangles and N_t is the number of triads of paths length two. The factor of three accounts for the symmetry of the triangle. Figure 3.4, gives a clustering transitivity of 0.33.

3.6 Path length

A path in a network is any sequence of consecutive edges that connects two nodes. The path length, also known as distance, is then defined by the number of edges transversed along the path. The shortest path between two nodes is called a *geodesic*. The geodesic length is the number of edges between two nodes such that no other shortest path exists. If $d_{i,j}$ is the shortest distance between node i and j, the average shortest path length in a network is defined as

$$\ell = \frac{1}{N(N-1)} \sum_{i=1}^{N} \sum_{j=1}^{N} d_{i,j},$$

where N is the number of nodes in the network. In networks, ℓ tends to be short (typically the average path length is comparable to the logarithm of N) [119] . The longest geodesic path ℓ_{max} is the network's diameter.

It is possible for a network to be fragmented into a number of disconnected components. In this case, the component with the largest size is called the *giant component*. By convention, the geodesic path length between two nodes is said to be infinite if the nodes belong to different components of the network [119]. In circumstances where the network is fragmented, one calculates the average shortest path length over paths that run between nodes of the same component (i.e., one simply omits the infinite paths).

3.7 Closeness

Closeness centrality measures the proximity of a node to other nodes in the network. It can be used to identify the most central node in a network as that with the least mean distance to all other nodes in the network. The closeness centrality of node i is calculated as

$$C_c(i) = \frac{1}{N-1} \sum_{j(\neq i)} \frac{1}{d_{i,j}}. \tag{3.7.51}$$

The mean closeness centrality for the network is

$$\langle C_c \rangle = \frac{1}{N} \sum_{i=1}^{N} C_c(i).$$

The idea of closeness centrality has found uses in various applications. For example, in transport networks, closeness was applied to rank airports in Ref. [120]. Erjia Yan and Ying Ding, in their study of co-authorship networks, refer to closeness as a useful proxy for scientific impact [121]. In addition, Ref. [122] used this metric to measure how Computer Science and Mathematics are progressively drawing closer to each other through interdisciplinary research.

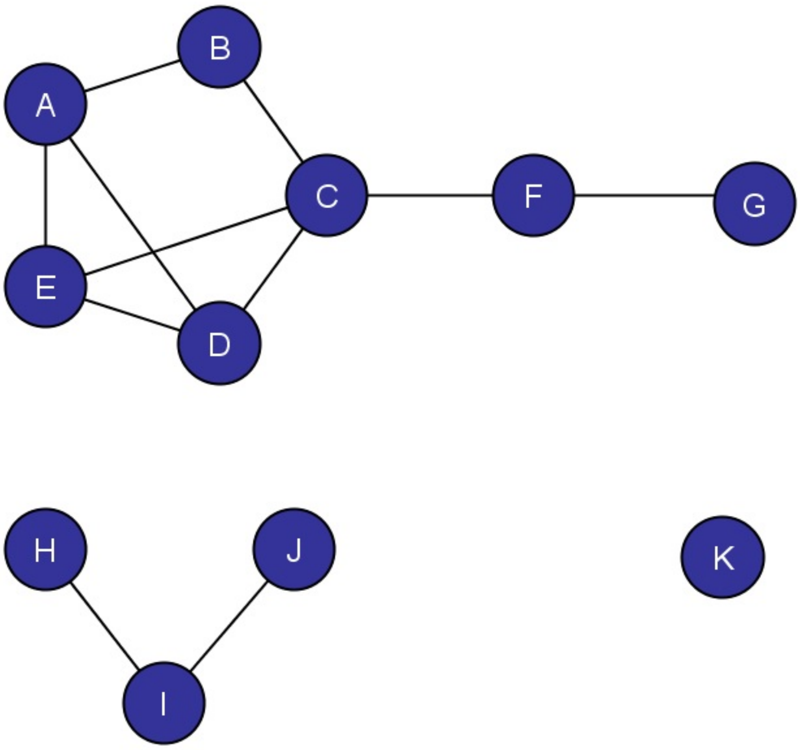

Figure 3.5: The network is used in Table 3.2 to illustrate how Closeness centrality is calculated in a network.

3.8 Betweenness centrality

Another measure of the relative importance of nodes is the *betweenness centrality* b_l. This measures the degree to which a node acts as a broker between pairs of other nodes in the network. Freeman [125] is typically accredited for this concept of betweenness, although Freeman himself [126] acknowledges Anthonisse [127] for the original idea. It is computed as the fraction of shortest paths between pairs of nodes passing through a given node [119]. It is normalised to give a bound $0 < b_l < 1$. To define it, we first write the number of geodesics (shortest paths) between nodes i and j as $\sigma(i, j)$. We denote the number of these which pass through node l as $\sigma_l(i, j)$. The betweenness centrality of vertex l is then defined as

$$b_l = \frac{2}{(N-1)(N-2)} \sum_{i \neq j} \frac{\sigma_l(i, j)}{\sigma(i, j)}. \tag{3.8.52}$$

If $b_l = 1$, all geodesics pass through node l. Nodes associated with high betweenness have a considerable influence in the network as they control the flow of information between other

Nodes												Closeness	
	A	B	C	D	E	F	G	H	I	J	K	$\sum_{j(\neq i)} \frac{1}{d_{i,j}}$	Normalised
A	...	1	0.5	1	1	0.33	**0.25**	0	0	0	0	4.08	0.41
B	1	...	1	0.5	0.5	0.5	0.33	0	0	0	0	3.83	0.38
C	0.5	1	...	1	1	1	0.5	0	0	0	0	5.00	0.50
D	1	0.5	1	...	1	0.5	0.33	0	0	0	0	4.33	0.43
E	1	0.5	1	1	...	0.33	0.25	0	0	0	0	4.08	0.41
F	0.33	0.5	1	0.5	0.33	...	1	0	0	0	0	3.67	0.37
G	0.25	0.33	0.5	0.33	0.25	1	...	0	0	0	0	2.67	0.27
H	0	0	0	0	0	0	0	...	1	0.5	0	1.50	0.15
I	0	0	0	0	0	0	0	1	...	1	0	2.00	0.20
J	0	0	0	0	0	0	0	0.5	1	...	0	1.50	0.15
K	0	0	0	0	0	0	0	0	0	0	...	0.00	0.00

Table 3.1: Closeness centralities for the network shown in Figure 3.5. For example, the geodesic path between nodes A and G is four and the entry in boldface is the inverse of this distance. Where no path exist between node i and j, for example A and J, the quantity $1/d_{i,j}$ is replaced by zero [123, 124]. The normalized values are bound between 0 and 1. Node C is the most central node in this example.

nodes. The removal of these nodes disrupts communication between other nodes as they act as a bridge for the largest number of paths between pairs of nodes in a network [119].

The concept of betweenness is additionally useful in defining one of the most common and novel measures of edge centrality, proposed by Girvan and Newman in [128]. In this case, i, j and l in Eq.(3.8.52) can be used to represent edges rather than nodes. The quantity b_l becomes *edge betweenness centrality* instead. Edge betweenness finds a wide range of application in network analysis including help with detecting communities [129].

3.9 Eigenvector centrality

Eigenvector centrality characterises node importance based on centralities of its neighbours. The measure recognises that not all interactions are equal. Nodes are deemed influential according to how they are linked to other important nodes. Eigenvector centrality is a variant of the "pagerank" score used to rank websites [119]. For a graph G, if we let x_i denote the average centralities of node i's neighbours, then

$$x_i = \frac{1}{\lambda} \sum_{j=1}^{n} A_{i,j} x_j, \tag{3.9.53}$$

where A is the adjacency matrix and λ is a constant. If $x = [x_1, x_2, x_3, ...]$ is the vector of

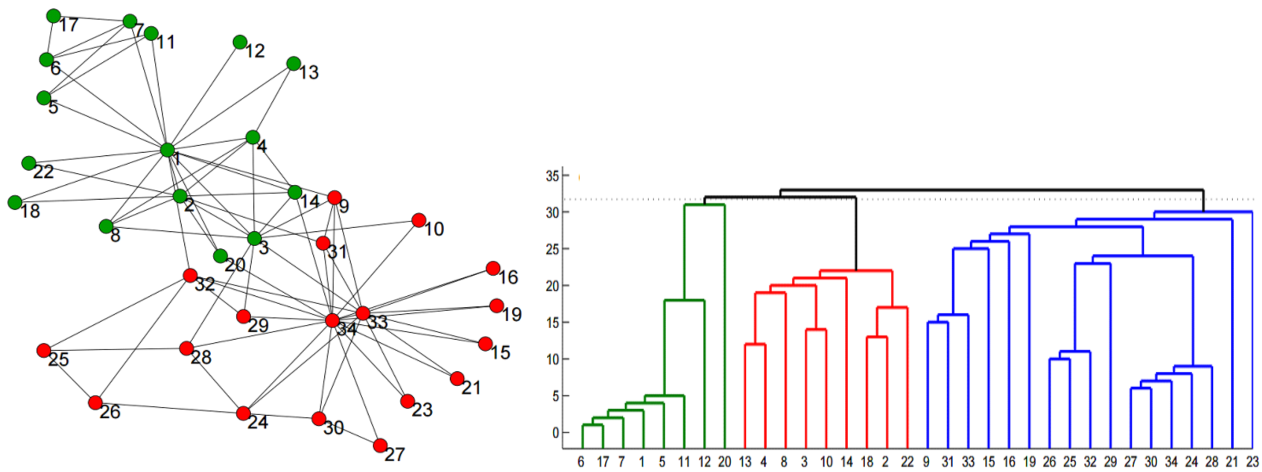

Figure 3.6: Community structure of the Zachary karate club network displaying three communities detected using the Girvan-Newman algorithm with modularity Q = 0.39 [130]. Panel to the left shows a network coloured into two different communities from earlier steps of the algorithm.

centralities then Eq.(3.9.53) leads to $\lambda x = A \cdot x$, where x is the corresponding eigenvector of the adjacency matrix A with its largest eigenvalue λ.

3.10 Community Structure

Another property of many real-world networks is the presence of a community structure. Communities are tightly knit groups i.e, nodes with multiple interconnections. Links between groups are measurably less dense [128]. An example from social networks could be groups of people with similar interests, beliefs, education, background or behaviours. Such a network structure where a heterogeneous large structure is built from homogeneous clusters is often refereed to as modular or a hierarchical community structure. This structure is sometimes represented using a dendrogram that exhibits how smaller groups are nested within larger groups which in turn make-up the entire network.

The problem of community detection is closely related to graph partitioning in computer science [131]. In the Girvan-Newman algorithm, the edge betweenness for the network is calculated and an edge with highest betweenness removed, edge betweenness is then recalculated and again removing the edge with highest betweenness [128, 129]. The process is repeated until an optimal number of communities is achieved. If we gradually remove edges with the highest betweenness, we obtain a hierarchical map, called a dendrogram as shown by an example in

Figure 3.6. The leafs are the individual nodes that represents the whole graph.

Girvan and Newman [128] introduced a modularity quantity Q as a qualitative measure that helps to evaluate the optimal number of communities (partitions) given by the algorithm. This is given by Eq.(3.4.41)

$$Q = \sum_c \left(e_{cc} - a_c^2\right), \tag{3.10.54}$$

where instead of colour, c represents a set of communities.

The karate club studied in [132] is a good example of a network that exhibit a community structure. This particular network is quite famous because has been used to test different community detection algorithms [128, 130, 133, 134].

The anthropologist Zachary conducted observations over two years relating to the friendships between 34 members of a karate club. A disagreement between the club's administrator and a teacher prompted a split which led to the club members taking sides between the two factions. The network depicted in Figure 3.6 contains 78 edges between 34 nodes.

3.11 Structural Balance

We have earlier touched upon the notion of positivity to describe friendly connections in social networks. The converse concept is negativity; links which represent hostile or conflictual relationships between pairs of nodes are deemed negative. Obviously the concepts of positive and negative edges relate to social networks. In a full social network, there is often a tendency to suppress odd numbers of negative links in a closed triad. This is related to the notion of *structural balance* [135]. This is the notion that "the enemy of my enemy is my friend"; hostility between two individuals in a social network (or characters of a text) is suppressed if they have a common foe. One way to quantify this is by introducing the quantity Δ, defined as the fraction of closed triads that contain an odd number of positive links. Clearly Δ is only relevant to the complete network rather than to its positive or negative subsets, formed out of only positive or only negative relations, respectively. (In the former $\Delta = 1$ by definition, while in the latter it is zero.)

3.12 Random networks

A random network is a graph containing N nodes where each pair is connected with some probability p [136]. There are two regularly considered models of random graphs, the Gilbert model [137] and the Erdős-Rényi model [138]. The Gilbert model is labelled by $G(N, p)$. In it,

M edges appear independently with probability $0 < p < 1$. In the closely related Erdős-Rényi model, $G(N, M)$, N nodes are connected by randomly placing M links between them. If we consider a random network where each edge is connected with probability p and a given node with degree k which have a maximum of $N-1$ connections. The probability of having degree k in the network will then be proportional to $p^k(1-p)^{(N-1-k)}$. The degree distribution of a random network therefore follows a binomial distribution of the form

$$p_k = \binom{N-1}{k} p^k \, (1 - p)^{N-1-k}. \tag{3.12.55}$$

The corresponding generating function is defined by

$$g(z) = [1 + p(z-1)]^{N-1}. \tag{3.12.56}$$

The first and second moments are given by

$$\langle k \rangle = p(N-1) \tag{3.12.57}$$
$$\langle k^2 \rangle = p(N-1)[p(N-2)+1]. \tag{3.12.58}$$

In the limit of large N, the following approximation can be used for Eq.(3.12.55):

$$P(k) = \frac{\langle k \rangle^k}{k!} e^{-\langle k \rangle}. \tag{3.12.59}$$

From Eq.(3.12.59), it shows that the degree distribution for a random network can be approximated by a Poisson distribution for large N with a small connecting probability, p.
Random graphs are characterised by a small average path-length ℓ_{rand} that behaves like the logarithm of the network size N. They are also associated with low clustering coefficient C_{rand}. The average path length of a random network ℓ_{rand} is suggested as

$$\ell_{\text{rand}} \propto \frac{\log N}{\log \langle k \rangle}. \tag{3.12.60}$$

This quantity in Eq.(3.12.60) was a proposal of Albert and Barabási [109] but later adjusted by to include the Euler-Mascheroni constant $a \approx 0.5772$ [139] and given as follows

$$\ell_{\text{rand}} \approx \frac{\log N - a}{\log pN} + \frac{1}{2}. \tag{3.12.61}$$

We recall that clustering coefficient is the average probability that there is a connection between two neighbours of a node. The clustering coefficient of a random network is given by

$$C_{\text{rand}} = \frac{\langle k \rangle}{N-1} \equiv \text{p}. \tag{3.12.62}$$

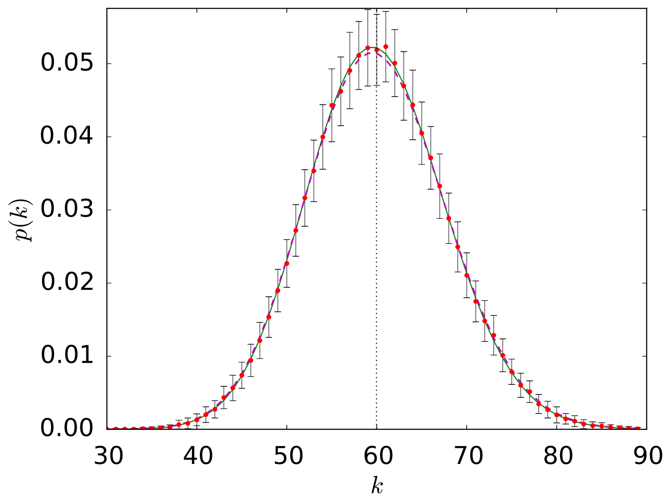

Figure 3.7: Degree distribution averaged from 100 random graphs, each with 2000 nodes connected with probability $p = 0.03$. The continuous green line shows the binomial distribution and the Poisson distribution is represented by a dotted line in red. The point shows the average fraction of nodes with degree k and error bars represent the standard deviation around that point.

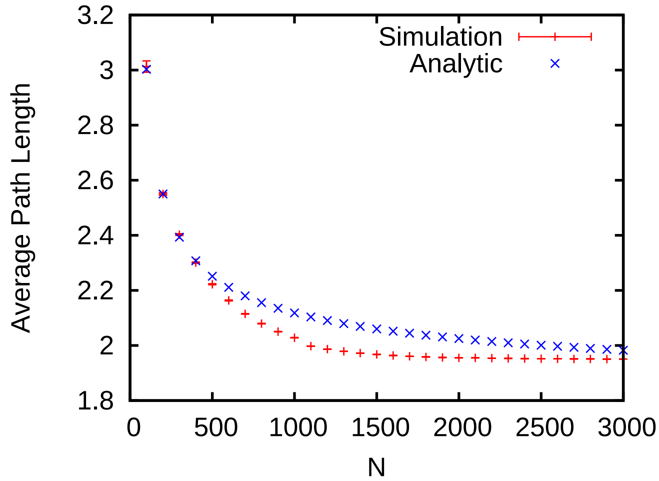

Figure 3.8: A plot of system size vs average path-length obtained from Erdős-Rényi random graph with probability $p = 0.05$. Even though the analytical results obtained from Eq.(3.12.61) do not follow the simulation results exactly, they follow similar behaviour. The simulation results are averaged over 100 realisations.

Both random and real-world networks inherently have a small average path length. However, there exist many differences between random and real-world systems. There is no correlation between the degrees of adjacent nodes in random networks as nodes link at random unlike in real-world networks where they are usually correlated. Another difference is found in the degree distribution. As discussed earlier, the degree distribution of real networks tend to be right skewed with a small number of nodes displaying hubs and a high number of low-degree nodes. On the other hand, random networks display a poisson degree distribution. Moreover, nodes in real networks tend to group into communities, a structure that is absent in random networks [119].

3.13 Small world networks

Small-worldness is related to the concept of "six degrees of separation". It is a phenomenon that a short path exists between any two strangers picked from anywhere in the world. Stanley Milgram conducted a study (the now famous "small-world experiment") to examine the average path-length between people in the United States. The research findings suggested that any two strangers are separated by not more than six steps between them. This notion became popularly known as the "six degrees of separation" [140]. Although the notion is generally accredited to Milgram, the concept can be traced back to 1929 to an experiment conducted by Frigyes Karinthy. Karinthy, a Hungarian poet and journalist [141], suggested that the world's population lives close enough to connect any two individuals by not more than five steps between them. He was able to establish a relationship with a very distant stranger in an American factory through not more than five individuals between them. He contacted a factory worker at Ford Motor Company, who knew his foreman, who knew Mr. Ford, who was acquainted with an American publisher who had Hungarian links – William Randolph Hearst, who in turn knew Mr. Árpád Pásztor a friend of Frigyes [142].

A small-world network is a graph in which most nodes can be reached from every other node by a small number of steps [13, 143]. Conventionally, random network models such as the Erdős-Rényi have been used to discern the structural properties of small-world networks [8,109, 139,144]. There are two main metrics that characterise the behaviour of small-world networks, $C \gg C_{rand}$ and $\ell \approx \ell_{rand}$. That is, a network G with N nodes and M edges is a small-world if its average path length is similar yet having significantly higher average clustering coefficient than that of a random network of the same size.

(a) (b)

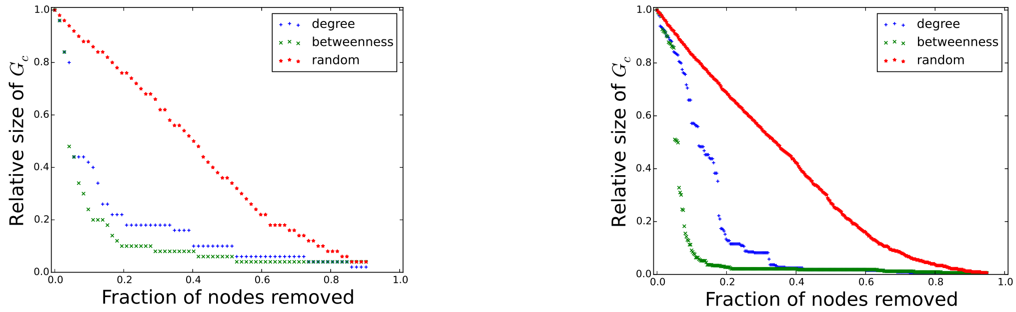

Figure 3.9: Robustness of the Beowulf network panel (a) and Laxdæla's saga (b).

3.14 Robustness

Robustness is measured by a network's ability to maintain it's basic functions amid network attacks. This implies the presence of parallel functions that can replace missing functionalities from removed nodes or edges. This concept is used in many fields. In economics, the concept is applied to study risks and stability in the financial markets [145]. In biology, the principal can be used to give an insight into mutations and diseases [146]. It is often useful to identify influential nodes in character networks. Nodes of high degree are likely to have more influence on the network than those of low degree. Likewise, a node with high betweenness centrality controls the connectivity of other nodes in the network. To test robustness, the size of the giant component is measured as nodes are removed [8]. A network is fragile when its giant component fragments quickly [147]. Node removal can be targeted: nodes are removed according to their degree or betweenness centrality. In the later case, betweenness centrality has to be recalculated every time a node is removed. Node removal can also be done randomly. Scale-free networks tend to be robust against the random removal of nodes [147]. This seems to be the same behaviour observed in mythological networks looked at by MacCarron [8].

3.15 Gathering the Data

As with previous studies [9, 15, 16, 148], we consider epic texts in Chapter 2 as playing out on character network comprising N nodes and M edges which link the nodes through relationship or interactions. We distinguish between two types of edge: positive and negative. Positive edges are established when any two characters are related, communicate directly with each another, or speak about one another or are present together when it is clear that they know each other. So positive edges basically represent friendly, social or neutral relationships. Negative links, on the other hand, represent hostility between nodes. They are formed when two characters meet in conflict or when animosity is explicitly declared by one character against another and it is clear they know each other.

It is possible that two characters are linked by both positive and negative edges as relationships between characters may change over time. However, because *Cogadh Gáedhel re Gallaibh* is temporally relatively non-extended, covering a period of only two centuries, its connectivity properties can be reasonably accommodated through a static analysis. Indeed, previous studies, to which we wish to compare, used the static approach [9, 15, 16]. Nonetheless, it should be noted that the study of dynamical properties of networks constitutes an active, broad and developing area of research and such an approach would be of interest in the future [148].

In the following analysis, we focus primarily on the *topology* of the networks underlying texts in Chapter 2. For this purpose, we consider undirected, unweighted networks. This means that (i) the features which connect the various nodes are not oriented and (ii) the statistics we report upon do not take into account varying levels of intensity of interactions between nodes. To account for (i), one would have to introduce a level of detail which is finer that just positivity or negativity. However, what one gains in refining this detail, one loses in statistical power. To account for (ii), one may place higher weight on more intense interactions. However, besides using the number of interactions between characters in the narrative, there is no established standard mode of weighting edges in character networks. Moreover, we are primarily interested in the presence or absence of conflict, not on the details of varying intensity of such hostility. Therefore we defer consideration of directed, weighted and temporal networks for future studies and restrict the current study to network topology and related matters.

Finally, we remark that our approach to constructing the networks follows the methodology of Refs. [9,15,16] in that nodes and links are identified by carefully reading the texts with multiple passes through all of the material by multiple readers. In our experience, such an approach is required to minimise errors and omissions as well as to reduce levels of subjectivity. As an example, *Cogadh Gáedhel re Gallaibh* is an extremely dense text and meticulous care is required to interpret extremely subtle tracts containing large amounts of explicit and implicit information. It is currently beyond technological capabilities to extract such information automatically owing to the inherent complexity of such texts (see, e.g., Ref. [149]). Establishing the technology for such an approach is another active area of research.

Together the set of nodes and edges form a *network*. Analysis of such networks delivers outputs in the form of statistics which can be considered as metrics. We can also study the distributions of edges amongst the nodes of the various networks as well as their connectivity information.

By comparing these we can gauge similarities and dissimilarities between the social structures underlying the narratives.
Our analyses begin by comparing some standard network topological observables underlying each text. This is followed by the analyses of their degree distributions. We finally look at the spectral distance from the adjacency matrices of the graphs.

3.16 Remarks

In this chapter we have presented a number of generic network statistics that are commonly used in the literature to characterise the structure of networks. We have also highlighted the shortcomings of the current metric used to quantify categorical assortativity, particularly in cases that involve disassortativity. In a complete disassortative network, one would expect the extremely minimum of this value (-1). Instead, the current measure gives a value below zero but not its expected minimum (-1). We, therefore introduce a new measure of disassortativity, an adjusted renormalised version of assortativity found in [118]. The new measure caters for a completely disassortative network and for more than two categorical classifications. Chapters 4 and 5 will show how these network topological measures can be used to quantify character networks.

Chapter 4

Network Analysis of *Cogadh Gaedhel re Gallaibh*

In this chapter, we present our character-network analysis for *Cogadh Gaedhel re Gallaibh*. A network visualisation of the full set of interactions recorded in the text is shown in Figure 2.2 of Sub-section 2.1.1. We begin here with Section 4.1 where the summary statistics are given. In Section 4.2 we consider networks formed from Irish and Viking characters separately. The degree distributions are examined in Section 4.3 and the question of network robustness and the importance of individual characters is addressed in Section 4.4. We compare the *Cogadh* networks to those from other epic traditions in Section 4.5. In Section 4.6 we look at a particular instance of interpolation in *Cogadh Gaedhel re Gallaibh*. Our main results are addressed in Section 4.7. It is here that we examine the question of what the data embedded in the *Cogadh* have to say regarding the long-standing debate about the nature of conflict in the Viking Age in Ireland. We conclude with some remarks in Section 4.8.
The results of this chapter have been published on the arXiv as Ref. [17]. The work has also been submitted in to the journal *Royal Society Open Science*.

4.1 The entire network and its positive and negative sub-networks

We identified 326 individual characters in the main part of Todd's translation of *Cogadh Gaedhel re Gallaibh*. The introduction, footnotes, appendices and index were also used to aid the identification of characters and links between them but individuals mentioned only in these paratexts do not form part of the *Cogadh* network. A small number of characters appear in the main text which are omitted in Todd's index. Of the 326 characters, eleven are isolated in the sense that they do not interact in the narrative. We consider these as not forming

Table 4.1: Full-cast networks comprise Irish, Viking and other nodes together with interactions between them. Unsigned networks comprise positive and negative edges as well as the nodes they connect. Thus, for example, the positive, full-cast network comprises all nodes but only positive links. The unsigned, Irish network comprises only Irish nodes but both positive and negative links between them. The *entire* network comprises all interacting nodes and all links.

		Edges	
		Positive	Negative
Nodes	Irish		
	Viking		
	Other		

Full cast

Unsigned

part of the *Cogadh* network and they are omitted from our analysis. Thus we are left with $N = 315$ individual interacting characters in the entire network. These nodes are interconnected by $M = 1190$ edges. We also identified 34 groups of unnamed characters. If these were to be considered as nodes, they would bring an additional 187 edges. However because these are neither individuals or named, we omit them too. (Note that for completeness, the networks with these extra nodes and links were analysed too and it was found that they deliver only very small changes to the statistics presented here.)

We refer to the assemblage of 315 nodes and 1190 edges as the entire network. We also consider the positive and negative sub-networks, which are formed only of positive or negative edges, respectively. Examination of these allows us to gain additional insight into the social and conflictual statistics contained in the narrative.

It is known from sociology that societies exhibit what is called *homophily*. As previously discussed, the tendency of individuals to associate with others who are similar to themselves in some way (homophily) is know as generally known as assortativity [107, 115, 150]. In previous studies of epic literature [8, 9, 15, 16], degree assortativity was a focus of study. It was shown that some positive sub-networks exhibit degree assortativity, or are uncorrelated. The opposite feature — degree disassortativity — was found to be characteristic of negative sub-networks. This means that, compared to full social networks, positive sub-networks give, in some sense, a "cleaner" picture of the non-conflictual societies underlying such narratives. This makes it valuable to study positive sub-networks in isolation [144]. A new feature of this thesis is the additional focus on the negative sub-network to statistically measure levels of hostility.

From here on, we use the term *unsigned* to refer to networks which contain both positive and

Table 4.2: Statistics for the entire network and the various sub-networks. Columns 1 and 2 indicate whether the sub-network is unsigned, positive or negative with full cast of characters (Irish, Viking and other) or only the Irish or Vikings are taken into account. N represents the number of nodes and M is the number of edges; $\langle k \rangle$ is the average degree and $k_{\max}$ its maximum; ℓ is the mean path length and C is the average clustering coefficient. Random counterparts are indicated with the subscripts "rand". The relative frequency of triads that contain an odd number of positive links is represented by Δ; the proportion of nodes belonging to the giant component is represented by G_c and the degree assortativity is denoted by r.

		N	M	$\langle k \rangle$	$k_{\max}$	ℓ	ℓ_{rand}	C	C_{rand}	Δ	G_c	r
Unsigned	Full cast	315	1190	7.6	105	3.6	3.1	0.58	0.02	0.93	0.95	-0.09(2)
	Irish	193	530	5.5	63	3.7	3.3	0.53	0.03	0.93	0.88	-0.08(3)
	Vikings	91	313	6.9	26	4.4	2.6	0.67	0.08	1.00	0.90	0.31(7)
		N^+	M^+	$\langle k \rangle^+$	$k^+_{\max}$	ℓ^+	ℓ^+_{rand}	C^+	C^+_{rand}		G^+_c	r^+
Positive	Full cast	287	957	6.7	53	4.1	3.2	0.59	0.02		0.87	0.00(4)
	Irish	186	475	5.1	47	3.9	3.4	0.53	0.03		0.85	-0.02(4)
	Vikings	88	301	6.8	26	3.0	2.5	0.68	0.08		0.73	0.34(7)
		N^-	M^-	$\langle k \rangle^-$	$k^-_{\max}$	ℓ^-	ℓ^-_{rand}	C^-	C^-_{rand}		G^-_c	r^-
Negative	Full cast	180	264	2.9	63	3.7	4.8	0.06	0.02		0.89	-0.25(3)
	Irish	62	72	2.3	25	2.6	4.7	0.06	0.04		0.69	-0.26(6)
	Vikings	18	16	1.8	4	1.5	4.5	0.00	0.10		0.33	-0.08(18)

negative edges. Networks made up of only positive (or only negative) edges are themselves termed *positive* (or *negative*, respectively). The term *full-cast* is used to refer to networks containing the full cast of characters, i.e., with Irish, Viking and other characters. Networks containing only Irish characters are themselves referred to as *Irish*. An analogous statement holds for Viking networks. This terminology used is summarised in Table 4.1. Statistics for the entire network and various sub-networks are collected in Table 4.2.

The mean number of edges per node for the entire network is determined to be $\langle k \rangle = 2M/N \approx 7.6$. This varies between 1 for the least connected characters and $k_{\max}$ for the most connected one. For the entire network, the most connected character is Brian himself. He has $k_{\max} = 105$ edges, and is linked to 33% of the other characters in the narrative. Of course, besides Brian's degree, we are also interested in the connectedness of the other characters in the narrative and we discuss the distribution of k_i-values (the so-called *degree distribution*) in Section 4.3. We also rank the first few characters according to their individual degrees, and according to other measures of importance, in Section 4.4.

Henceforth, we use the superscripts "+" and "−" to identify statistics associated with the positive and negative networks, respectively. We refrain from using a superscript for the unsigned networks. *Cogadh Gaedhel re Gallaibh* has $N^+ = 287$ interacting characters in its positive sub-network, and they are interconnected by $M^+ = 957$ edges. This corresponds to a mean degree of $\langle k \rangle^+ \approx 6.7$. Again we have omitted isolated nodes from the positive and negative sub-networks. (We do this because the negative sub-network can have a large number of zero-nodes and may end up with average degree less than 1, which is not sensible for comparisons to other networks.) The counterpart numbers for the negative network are $N^- = 180$, $M^- = 264$ and $\langle k \rangle^- \approx 2.9$. We note that the total number of positive and negative links $M^+ + M^- = 957 + 264 = 1221$ is greater than the number $M = 1090$ which we previously identified for the entire network. This is because some relationships involve both positive and negative aspects. Brian has the highest degrees in both positive and negative sub-graphs, with the former measured at $k^+_{\max} = 53$ and the latter at $k^-_{\max} = 63$.

The mean path length for the full network is measured to be $\ell \approx 3.6$. This is shorter than the *six degrees of separation* that is supposed to characterise modern society [140]. The counterpart statistics for the positive and negative sub-networks are $\ell^+ \approx 4.1$ and $\ell^- \approx 3.7$, respectively.

The clustering coefficient of a given individual is the proportion of acquaintances which are themselves mutually linked and the average value over all vertices of the network is the mean clustering coefficient. We found that both the full network and the positive sub-network have significant amounts of clustering with $C \approx 0.58$ and $C^+ \approx 0.59$, respectively. By contrast, the negative sub-network is unclustered with $C^- \approx 0.06$.

It is usual to compare the structure of a given network to that of a random one. To do this in a sensible manner, the random network is chosen to have the same size and mean degree as the one we wish to compare to. As mentioned in Section 3.13, the term small world applies to the network if its mean path length ℓ is similar to that of the random graph ℓ_{rand} and if the clustering coefficient of the network C is much larger than that of the same random graph C_{rand} [13, 151]. We find that the path lengths of the associated random networks are of the same order of magnitude for both the unsigned and positive networks but the clustering is far smaller (see Table 4.2). The unsigned and positive *Cogadh* networks are therefore small worlds but the negative network does not exhibit this feature.

Triads with odd numbers of positive edges are considered structurally balanced and the overall network is loosely also called structurally balanced if a high proportion of them have this feature. We find this to be the case for the entire network underlying *Cogadh Gaedhel re Gallaibh* (which has 3041 triads) with $\Delta \approx 93\%$.

As mentioned in Section 3.6, a network may be fragmented into a number of disconnected components. The largest is called the *giant component*. The giant component of the unsigned *Cogadh* network contains 95% of its nodes and the equivalent proportions for the positive and negative networks are 87% and 89%, respectively (Table 4.2).

In common with other character networks, we find that the negative full-cast network is disassortative by degree, with $r = -0.25(3)$. The error here is estimated using the method given in Refs. [107, 115]. Error estimates for other network statistics are small (see discussion in the final paragraph of Section 4.6) and we refrain from reporting them here. The reason why we display errors for assortativity values is that they provide useful information when comparing systems which are, or nearly are, uncorrelated (r close to zero). The negative value of r for the negative sub-network means that high-degree characters are hubs and their negative links preferentially attach to low-degree ones. This is a generic feature of heroic tales where a small number of individuals encounter multitudes of lesser characters and defeat them in battle. The positive full-cast network, by contrast, is uncorrelated within errors, with $r = -0.00(4)$. This means it is neither assortative nor disassortative. These features are quite typical of social networks and of character networks with positive interactions [15, 144].

4.2 The Irish and Viking sub-networks

Beside the networks comprising the full cast of characters, we can also analyse the networks containing only Irish or only Viking nodes. The statistics are also listed in Table 4.2. As usual, isolated (degree-zero) nodes are removed. For example, there are 202 Irish nodes in total (see Table 4.8), but nine of these are disconnected from other Irish nodes. These are omitted in calculating the statistics for the unsigned Irish network in Table 4.2. Apart from the value of N, reinstating the nine does not alter the statistics listed within the precision of Table 4.2.

We observe the following average properties of the various sub-networks. In the Irish, and Viking networks (as, indeed, in the full-cast cases), the mean degrees are maximum for the unsigned networks and minimum for the negative ones. For the Irish sub-network, the mean path length is largest for the positive sub-network, just as it is for the full-cast of characters. The clustering coefficients of the unsigned and positive networks are similar in size and far smaller for the negative sub-networks. The unsigned networks and their positive sub-networks are considered small worlds but the negative sub-networks are not. This is quite a typical feature of epic literature [8, 9, 15, 16]. The unsigned Viking network is more structurally balanced than the Irish counterpart. Structural balance for the Irish network, with 830 triads, is 93% whereas all of the 881 Viking triads contain odd numbers of positive links.

4.3 Degree distributions

Degree distributions deliver further important insights into the nature of complex networks and facilitate comparisons between them. Here a brief report on the degree distributions of the cast of *Cogadh Gaedhel re Gallaibh* is presented and how they compare to other epic narratives.
The simplest possibility is to seek to describe a degree distribution by a power law using the form $p(k) \sim k^{-\gamma}$ [108, 152]. Actually, this is more common in the tail of the distribution, i.e., some minimum value degree $k_{\min} > 0$ [153]. Following Ref. [9] we use $k_{\min} = 2$ to capture as much information as possible.
As described in Section 3.2 power-law degree distributions often exhibit a cut-off at some value $k = \kappa$. Section 3.2 provides details of this and other distributions including the log-normal and Weibull categories. We find that the positive network underlying *Cogadh Gaedhel re Gallaibh* is of the truncated power-law type. The negative network of *Cogadh Gaedhel re Gallaibh* is best fitted by a log-normal distribution whilst the entire network is best described by a Weibull distribution. Notwithstanding these fits, truncated power laws and Weibull distributions cannot be ruled out as an alternative for any of the networks considered here.
In Table 4.3 we present the relative probabilities with the most likely values highlighted in boldface. In Table 4.4 we present the maximum-likelihood estimates. Statistics in boldface are those which come from the most likely distributions. These most likely distributions (along with other candidate functions) are plotted for the unsigned, positive and negative networks in Figure 4.1.

Table 4.3: Relative minimum-information-loss probabilities for the degree-distribution functions of the unsigned, positive and negative networks. The values of distributions which are most likely are highlighted in boldface.

	Power law	Truncated power law	Log normal	Exponential	Weibull
Unsigned	~ 0	0.26	0.22	~ 0	**0.52**
Positive	~ 0	**0.62**	~ 0	0.04	0.34
Negative	0.04	0.27	**0.44**	~ 0	0.25

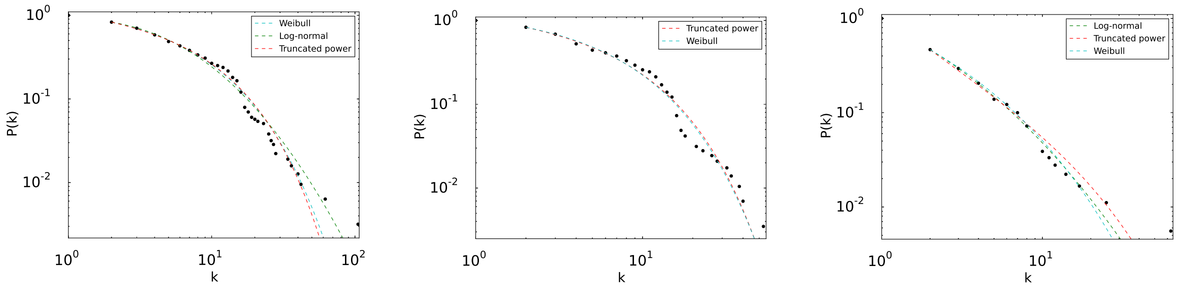

Figure 4.1: The complementary cumulative degree distributions for the unsigned network of *Cogadh Gaedhel re Gallaibh* (left), its positive sub-network (middle) and negative sub-networks (right).

Table 4.4: Maximum-likelihood estimates for the parameters associated with the various probability distributions, as fitted to the data for the unsigned, positive and negative, full-cast networks. Values which correspond to the most likely distributions are highlighted in boldface.

	Power law	Truncated power law		Log normal		Exponential	Weibull	
	γ	γ	κ	μ	σ	κ	β	κ
Unsigned	1.7(1)	0.5(2)	11.4(1)	1.6(1)	1.0(1)	7.4(1)	**0.8(1)**	**5.3(1)**
Positive	1.8(1)	**0.5(2)**	**9.4(1)**	1.5(1)	1.0(1)	6.3(1)	0.8(1)	5.0(1)
Negative	2.2(2)	1.8(1)	30.9(1)	**0.0(1)**	**1.3(1)**	3.6(1)	0.5(1)	0.5(2)

4.4 Network robustness and importance of individual characters

The leading characters of *Cogadh Gaedhel re Gallaibh* are listed in Table 4.5, ranked according to four different measures: degree; betweenness; closeness and eigenvector centrality. It can be noticed that Brian is the most central node based on all four measures. Brian, a high king of Ireland, is portrayed in *Cogadh*as a heroic king who ended Viking invasion of Ireland. Sitriuc is also recognised as an important character in the unsigned network following Brian. Sitriuc, the main Viking leader, is associated with a number of raids into Irish territories and played a prominent role at the battle of Clontarf in 1014.

An interesting question is: how reliant is the integrity of the giant component on the most important characters? This is related to the question of robustness and it can be investigated by examining the effects of systematic removal of nodes or edges. Another possibility is to remove elements of the network in a random manner. In the systematic approach, we can remove the most important nodes one after the other and keep tabs on how the giant component reduces

Table 4.5: The most important characters of *Cogadh Gaedhel re Gallaibh* ranked according to various criteria: degree, betweenness centrality, closeness, and eigenvector centrality.

	Rank	Degree	Betweenness	Closeness	Eigenvector
Unsigned	1	Brian (105)	Brian (0.42)	Brian (0.44)	Brian (0.53)
	2	Sitriuc (62)	Sitriuc (0.21)	Sitriuc (0.41)	Maelmordha (0.28)
	3	Maelmordha (42)	Ottir (0.16)	Ottir (0.39)	Malachy II (0.22)
	4	Ottir (40)	Aedh Finnliath (0.13)	Gormflaith (0.38)	Sitruic (0.21)
	5	Malachy II (36)	Ossill (0.11)	Maelmordha (0.38)	Gormflaith (0.21)
Positive	1	Brian (53)	Brian (0.28)	Sitriuc (0.34)	Brian (0.48)
	2	Sitriuc (40)	Sitriuc (0.17)	Brian (0.34)	Murchadh (0.30)
	3	Maelmordha (38)	Malachy II (0.11)	Gormflaith (0.34)	Maelmordha (0.26)
	4	Gormflaith (34)	Ottir (0.10)	Maelmordha (0.32)	Malachy II (0.26)
	5	Ottir (32)	Gormflaith (0.10)	Malachy II (0.32)	Conaing (0.23)
Negative	1	Brian (63)	Brian (0.63)	Brian (0.44)	Brian (0.66)
	2	Sitriuc (25)	Ottir (0.23)	Malachy II (0.35)	Maelmordha (0.23)
	3	Mathgamhain (17)	Sitriuc (0.23)	Sitriuc (0.34)	Brodir (0.22)
	4	Cathal (14)	Aedh Finnliath (0.16)	Ottir (0.33)	Malachy II (0.17)
	5	Olaf Cuaran (12)	Olaf Cuaran (0.12)	Ivar (0.32)	Ivar (0.17)

in size. We can then compare this to the results of the random approach, in which removal of nodes is a random process.

There are many ways in which one can decide which are the most important or influential nodes. One possibility is to consider that nodes with highest degree are most important and to remove them first. Another is to consider nodes with the highest betweenness centralities [154]. Other measures of importance include nodes' closeness and eigenvector centralities as defined in Section 3.7.

We present the results of the study of robustness for the networks underlying *Cogadh Gaedhel re Gallaibh* in Figure 4.2. As discussed in Section 3.14, we test robustness by measuring the size of the giant component as nodes are removed. Nodes are removed either randomly or in a targeted manner (nodes with highest degree or betweenness centralities). When nodes are removed by betweenness centrality, betweenness is recalculated after each removal.

The relative sizes of the giant component of the unsigned network as nodes are removed randomly are depicted in the main left panel by the red data points. The blue and green data points correspond to removal by highest betweenness and by degree, respectively. A similar behaviour is observed for the positive network, which is shown in the insert. The equivalent information for the negative sub-network is contained in the next main panel. Random removal of nodes has only a relatively gradual effect on the giant-component size in each of the three networks. Removal by betweenness or by degree has rather more devastating consequences.

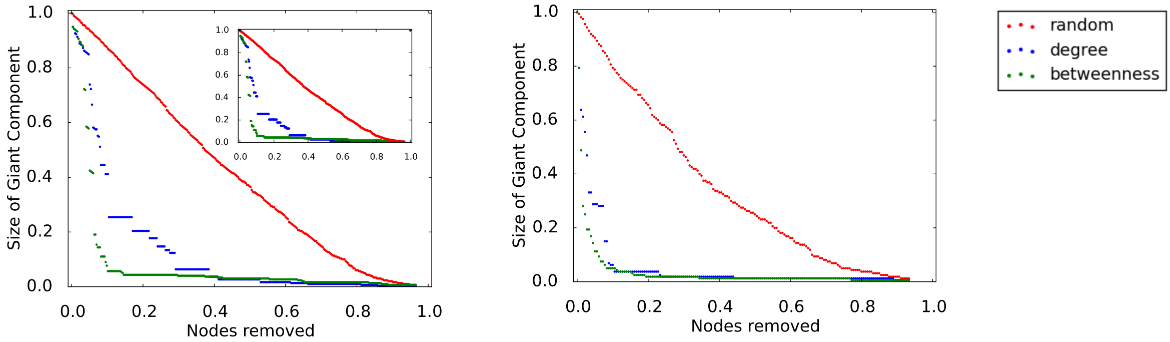

Figure 4.2: The relative sizes of the giant components as depending on the percentage of nodes removed. The left panel gives the size of the giant component for the unsigned network. That for the positive network, which has a very similar type of decay, is given as an insert. The rightmost panel displays the decay of the giant component of the negative network as nodes are removed. The red data points track the effects of random removal of nodes and the blue and green data correspond to removal by highest degree and betweenness, respectively.

Removal by betweenness is particularly effective for the destruction of the full and positive networks. For the negative network, removal by betweenness and degree are about equally effective. More details of the effects of node-removal on the relative sizes of the giant components are given in Table 4.6.

4.5 Comparisons to epic networks

It is interesting to compare the network structures in *Cogadh Gaedhel re Gallaibh* to those from epic narratives previously studied in Refs. [15, 16]. We decided to compare to the Irish tale *Táin Bó Cuailnge* (*The Cattle Raid of Cooley*) because a recension (version) of this story is contained in the *Book of Leinster* (which also, as we have seen, contains the oldest extant fragmentary recension of *Cogadh Gaedhel re Gallaibh*). Actually, it has been claimed that the style of *Cogadh Gaedhel re Gallaibh* is not unlike that used in the Book of Leinster's version of the *Táin* [19, 39–41, 47]. We also decided to compare the *Cogadh* to the Icelandic *Njáls Saga*. This decision was motivated by the fact that the latter also contains a record of the Battle of Clontarf and its relation with *Cogadh Gaedhel re Gallaibh* is described above. Finally we decided to also compare to the *Iliad*. This decision was motivated by Refs. [36, 40, 47, 60] which claim that as the account given in *Cogadh Gaedhel re Gallaibh* is partly a pseudo-history borrowed from the story of Troy.
The statistics collected in Refs. [15, 16] are listed in Table 4.7. Again the corresponding data

Table 4.6: The effects of removing the most important characters or of randomly removing characters. Here, the entries in the table give the relative size of the giant component after removal of the top 10% of characters systematically and randomly; the top five characters; and after removal of the most important character, namely Brian Boru.

	Remove 10% by degree	Remove 10% by betweenness	Remove 10% randomly	Remove top 5 by degree	Remove top 5 by betweenness	Remove Brian Boru
Unsigned	43%	6%	92%	90%	91%	92%
Positive	47%	7%	83%	85%	85%	86%
Negative	6%	5%	81%	69%	58%	85%

omit zero-degree nodes. There are only six such nodes in the *Iliad* and even fewer for the Táin and *Njáls Saga*. The positive *Cogadh* network contains 91% of all *Cogadh* nodes. The corresponding amounts for *Táin Bó Cuailnge*, *Njáls Saga* and the *Iliad* are 96%, 98% and 92%, respectively. In contrast, the negative *Cogadh* network contains 57% of all *Cogadh* nodes. The counterpart proportions for *Táin Bó Cuailnge*, *Njáls Saga* and the *Iliad* are 32%, 46% and 25%, respectively. These statistics suggest that hostility plays a greater overall role in *Cogadh Gaedhel re Gallaibh* than it does in the other narratives.

Next we turn to the unsigned networks. We observe that even though *Cogadh Gaedhel re Gallaibh* is closer in size to *Táin Bó Cuailnge*, in terms of node numbers, the network statistics are different. The *Cogadh* network is also quite dissimilar to that of *Njáls Saga*. The data suggest that the overall network of *Cogadh Gaedhel re Gallaibh* is more alike that of the *Iliad* to either that of the Irish or Icelandic works. The mean and maximal degrees and the mean path length are very close ($\langle k \rangle \approx 7.6$ for *Cogadh Gaedhel re Gallaibh* and 7.7 for the *Iliad*; $k_{\text{max}} = 105$ and 106; $\ell \approx 3.6$ and 3.5 respectively). The degree assortativities are also similar ($r = -0.09$ for the *Cogadh* and $r = -0.08$ for the *Iliad*).

4.6 Interpolations

The most famous interpolation in the narrative is that associated with Fergal Ua Ruairc [19,45]. It is suggested by Ní Mhaonaigh that "the period in which the Uí Ruairc stood to gain most from associating themselves with the Uí Briain seems to have been the mid- to late 1140s" [45]. This implies such a date for the interpolation. (Note that the name "Ua Ruairc" refers to an individual, Fergal in this instance. The "Uí Ruairc" refer to the associated clan or dynasty.)

Table 4.7: Here we list the statistics for the unsigned, positive and negative networks of *Táin Bó Cuailnge*, *Njáls Saga* and the *Iliad*.

		N	M	$\langle k \rangle$	$k_{\max}$	ℓ	ℓ_{rand}	C	C_{rand}	Δ	G_c	r
Unsigned	*Táin*	422	1266	6.0	168	2.8	3.5	0.73	0.01	0.92	0.99	-0.35(2)
	Njáls Saga	575	1612	5.6	83	5.1	3.9	0.42	0.01	0.90	1.00	0.01(2)
	Iliad	694	2684	7.7	106	3.5	3.4	0.44	0.01	0.98	0.99	-0.08(2)
		N^+	M^+	$\langle k \rangle^+$	$k^+_{\max}$	ℓ^+	ℓ^+_{rand}	C^+	C^+_{rand}		G^+_c	r^+
Positive	*Táin*	405	1118	5.5	164	3.0	3.7	0.74	0.01		0.93	-0.33(2)
	Njáls Saga	564	1388	4.9	77	5.4	4.1	0.39	0.01		0.96	0.07(2)
	Iliad	640	2329	7.3	85	3.8	3.5	0.44	0.01		0.86	0.09(2)
		N^-	M^-	$\langle k \rangle^-$	$k^-_{\max}$	ℓ^-	ℓ^-_{rand}	C^-	C^-_{rand}		G^-_c	r^-
Negative	*Táin*	134	148	2.2	82	2.9	6.0	0.03	0.02		0.91	-0.45(4)
	Njáls Saga	145	224	3.1	22	4.9	4.4	0.0	0.02		0.82	-0.30(4)
	Iliad	321	355	2.2	40	4.5	7.0	0.0	0.01		0.90	-0.44(4)

We saw earlier that 93% of triads in the unsigned network are structurally balanced. The same proportion of triads in the Irish network are also structurally balanced. The triad formed by the enmity between Ua Ruairc's and Máel Sechnaill, the alliance between Máel Sechnaill and Brian, and the interpolated support of Ua Ruairc for Brian has two positive and one negative edges, meaning that it is structurally *imbalanced*. A possible explanation for the interpolation strategy is given in Ref. [45]. Since only a small proportion of triads in *Cogadh Gaedhel re Gallaibh* are imbalanced in this manner, this makes the Ua Ruairc episode stand out as unusual. To investigate further, Ua Ruairc and his three associates (Domhnall mac Raghallach, Gilla-na-naomh, and Mac an Trin) were omitted from the networks to test the effects on the statistics. Obviously the number of edges decreases and M reduces by 44 to 1146 in the entire network. But beyond this, the effects of this removal are tiny. The degree assortativies, for example, do not change within the error estimates for each of the unsigned, positive and negative networks.

Having established that the most famous interpolation has only a very minor effect on the overall networks, we do not attempt to remove other interpolated material from our analysis. Besides, any attempt to do this would be partial or incomplete as there can be no certainty all interpolations have been identified and removed. Moreover, since this is a study of the networks in *Cogadh Gaedhel re Gallaibh* as represented by Todd in Ref. [35], we considered it right to present it in its entirety. However, for completeness, we simulated some effects of interpolation by randomly removing up to 15% of nodes or edges. Repeating the process 1000 times and then taking averages delivers no sizeable difference to the statistics of Table 4.2. This, again, indicates their robustness (see Section 4.4 for a network-robustness analysis). For example, removal of 15% of the nodes changes the assortativity from $r = -0.09$ to $r = -0.08$ (an imperceptible amount within errors). Removal of 15% of the edges leaves r unchanged within the level of precision identified. While interpolation appears not to affect the overall (global) network statistics, a more systematic and targeted (local) quantitative study of the effects of interpolation might be interesting for future study.

4.7 Interactions between cultures in *Cogadh Gaedhel re Gallaibh*

As discussed in Chapter 2, traditional "memory" of the events surrounding the Battle of Clontarf is of an international conflict between two clear, distinct sides: Irish vs Viking [19]. This is rejected by some revisionist historians who instead argue that the conflict is primarily between two Irish sides (each with Viking support) [19, 43, 62]. A viewpoint of a clear-cut contest could lead one to expect a network wherein most of the negative (hostile) interactions are between Irish and Viking nodes representing the primacy of hostility being between these two groups. On the other hand, a network supporting the revisionist view might be expected to deliver something rather different: the majority of negative edges would mainly link pairs of Irish nodes. In the following, we have also to monitor Viking versus Viking conflict because, as explained by Todd, there are "two distinct nations of the Gaill" [35]. "They are distinguished as white or fair-haired, and black or dark-haired foreigners, the Danes being the dark, and the Norwegians, including, perhaps, Swedes, the white race. The two nations are represented as hostile to each other, and battles between them not infrequently took place. But it is to be regretted that our author does not always very clearly distinguish between them in his descriptions of their devastations in Ireland. We cannot even be sure that the name Dane is not sometimes given to the Norwegians." We present visualisations of the positive and negative networks in Figure 4.3 and Figure 4.4. The same colour coding we used as for Figure 2.2 (i.e., green edges stand for Irish-on-Irish interactions, blue for Viking-Viking, brown for Irish-Viking and grey for interactions involving unassigned nodes). In Table 4.8, we record the proportions of Irish, Viking and other nodes in each of the unsigned networks and in its positive and negative sub-networks. (Some of the

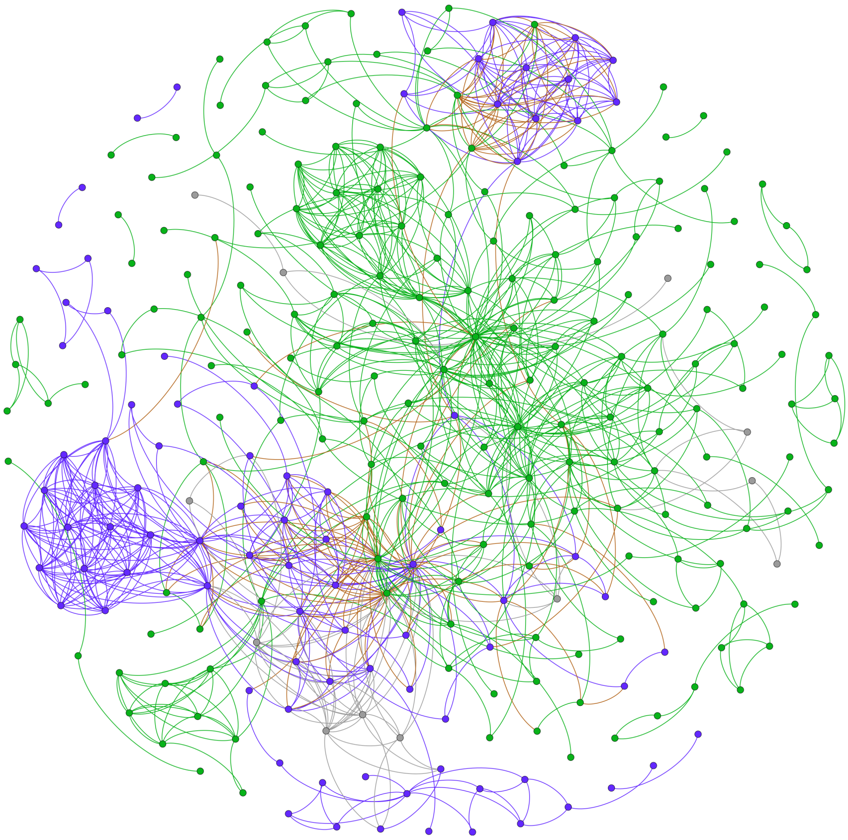

Figure 4.3: The complexity of the positive network underlying *Cogadh*. About 50% of the positive interactions connect pairs of Irish nodes (green); 31% are between Viking pairs (blue) and only 12% of positive interactions are between Irish and Viking (brown). The cultural assortativity for this network is measured to be $\rho^+ = 0.65$.

entries in the second and third rows of Table 4.8 are different to the corresponding entries in the third column of Table 4.2. This is because isolated nodes are not removed from sub-networks in Table 4.8. The reason for this is that Table 4.8 concerns cultural profiles of unsigned, positive and negative networks, rather than the Irish and Viking sub-networks of Table 4.2. Of course, numbers of edges match across both tables because, by definition, these do not involve isolated nodes. The proportions of Irish nodes in each of the three graphs lie between 61% and 65 %, so are nearly constant. The proportion of Viking nodes is also quite stable, lying between 31% and 34%.

In the same table, the proportions of interactions which link Irish to Irish nodes are listed, as are those connecting Viking-to-Viking and Irish-to-Viking pairs. Half of the edges in the positive network of Figure 4.3 connect pairs of Irish nodes; 31% link pairs of Viking nodes; 12% of positive interactions connect mixed Irish-Viking pairs. Twenty-seven percent of edges

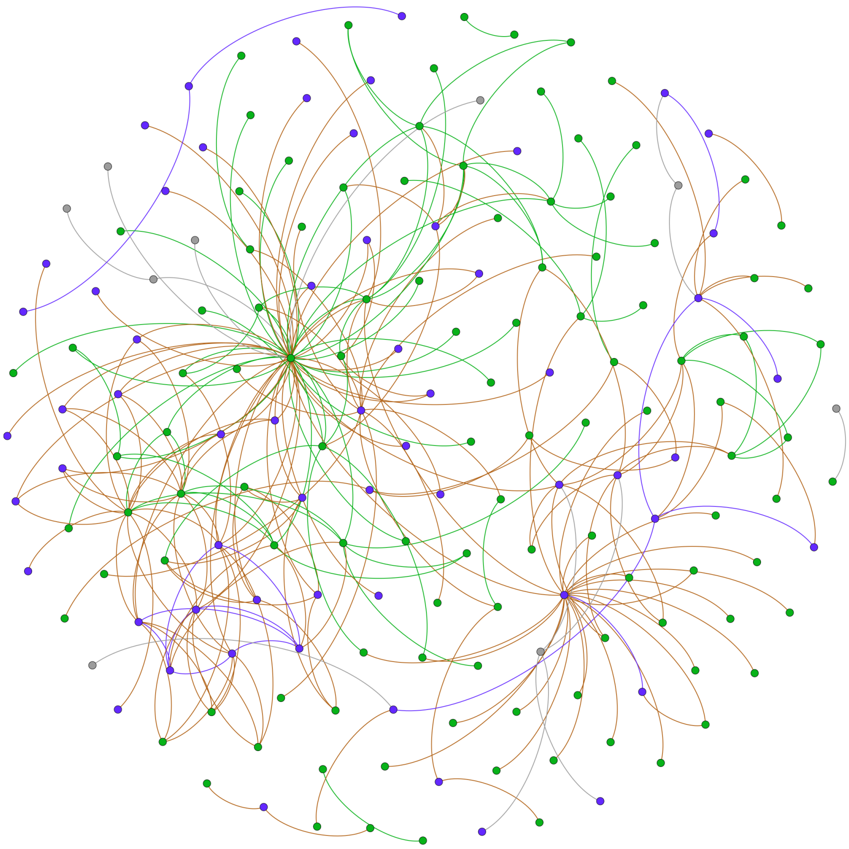

Figure 4.4: The complexity of the negative network underlying *Cogadh*. again, the same colour coding is used as in Figure 2.2 and Figure 4.3. Twenty-seven percent of the conflictual interactions connect pairs of Irish nodes (green); 6% are between Viking pairs (blue). Sixty-two percent of the negative links are between mixed Irish-Viking pairs (brown). The cultural assortativity for this network was found to be $\rho^{-} = -0.32$.

in the negative network of Figure 4.4 connect Irish to Irish nodes; 6% connect pairs of Viking nodes; and over 62% of negative interactions connect mixed Irish-Viking pairs. This means that the positive (i.e., essentially social) network of Figure 4.3 is dominated by interactions between characters of the same cultural identities (what we term intranational interactions). The negative (conflictual) network of Figure 4.4 is dominated by Irish-Viking (international) interactions. This indicates that the largest proportion of *Cogadh* conflict is international, but with significant levels of intranational hostilities too (especially Irish versus Irish). From Table 4.8, we observe that the number of international edges in the negative network is over two times the number of Irish-Irish negative edges, which, in turn is over four times the number of Viking-Viking negative edges.

To properly evaluate the levels of mixing between Irish and Viking, one should account for the fact that they have different numbers of nodes in the networks. In fact, there are about twice as many Irish nodes as there are Viking. To account for this, we introduce what we call the *cultural assortativity* of the various networks, and we represent this generically by ρ. The precise definition of ρ is given in Section 3.4. It is a metric which runs between $\rho_{\min}$ and 1 where $\rho_{\min}$ is itself a non-trivial, negative value, which lies between -1 and 0 if there are more than two categories under consideration [107, 115]. Therefore, although the largest possible value of ρ is one, its smallest value can be network-dependent. This is because, when there are more than two categories present, disassortativity connects dissimilar nodes — like randomness tends to. Assortativity, however, connects only like nodes and is quite different to randomness. We have to keep this asymmetry in mind when we interpret the cultural assortativity for the negative networks with three categories of node (Irish, Viking and unassigned). It turns out that $\rho_{\min} = -1$ only when there are two categories.

A purely culturally assortative network is signalled by a value $\rho = 1$. For example, if this were so for our positive network it would imply that the only positive interactions are *within* rather than between cultural groups (it would mean that friendly interactions are intranational). The value $\rho = -\rho_{\min}$, on the other hand, implies that the network is completely culturally

Table 4.8: Cultural profiles of the cast of *Cogadh Gaedhel re Gallaibh* and their interactions. The second, third and fourth rows list the numbers (and percentages) of nodes which are identified as Irish, Viking and others (i.e., those not identified as Irish or Viking) in the entire, unsigned network as well as in the positive and negative sub-networks. The next row gives the total number of nodes in each network (these values are N, N^+ and N^- for the full-cast networks, respectively). The sixth and seventh rows give the numbers (and proportions) of edges which connect pairs of like nodes. The eighth row gives the numbers (with proportions) of edges which connect Irish and Viking nodes. The final row gives the total number of edges in each case as (M, M^+ and M^- for the full-cast networks). The remaining edges involve other (not assigned as Irish or Viking) nodes.

	Entire network	Positive network	Negative network
Irish nodes	202 (64 %)	187 (65 %)	110 (61 %)
Viking nodes	97 (31 %)	88 (31 %)	61 (34 %)
Other nodes	16 (5 %)	12 (4 %)	9 (5 %)
Total # nodes	315 (100 %)	287 (100 %)	180 (100 %)
Irish-Irish edges	530 (45 %)	475 (50 %)	72 (27 %)
Viking-Viking edges	313 (26 %)	301 (31 %)	16 (6 %)
Irish-Viking edges	272 (23 %)	119 (12 %)	163 (62 %)
Total # edges	1190 (100 %)	957 (100 %)	264 (100 %)

disassortative. If this were the case for our positive network it would signal that the only positive interactions are *between* rather than within cultural groups (positive interactions would be international). A zero-value of ρ would indicate that the cultural assortativity is the same as should be expected for random mixing between nodes, oblivious of their Irish or Viking character. It turns out that $\rho^+ = 0.65(3)$ for the full-cast positive network. Restricting our attention to Irish and Viking nodes only by removing the other nodes, gives $\rho^+ = 0.72(3)$. These statistics are recorded in Table 4.9 and they are supportive of the picture that most (but not all) positive interactions are intranational.

We now focus our attention on the negative networks as these connect with the debate in the humanities discussed in Chapter 2. A "clear-cut" version of the "international-conflict" picture would correspond to the value $\rho^- \approx \rho^-_{\text{min}}$ (where ρ^-_{min} is the minimum possible value of ρ^-, i.e., -1 when unassigned nodes are excluded). A limiting negative value would indicate a purely Irish-versus-Viking conflict. At the other end of the spectrum would be a narrative in which all conflict is intranational. In this case one should expect $\rho^- \approx 1$. The revisionist picture of a primarily (but not entirely) intranational conflict might be expected to correspond to a positive value of ρ^-. Between these two extremes, we a more even distribution of negative edges would correspond to conflict between nodes being "blind" to their cultural identities. Such a completely colour-blind narrative would deliver $\rho^- \approx 0$ for the negative network.

Our data delivers t $\rho^- = -0.32(6)$ if all three kinds of node (Irish, Viking and other) are included in the negative network. One should compare this statistic to the theoretical minimum $\rho^-_{\text{min}} = -0.88(4)$. If unassigned nodes are omitted, (so that $\rho^-_{\text{min}} = -1$), one finds $\rho^- = -0.37(6)$. Therefore the measured values for cultural assortativity on the negative (conflictual) networks are negative. As discussed, this suggests that revisionist picture of a primarily intranational conflict is not supported by data contained in *Cogadh Gaedhel re Gallaibh*. However, the conflict is not clear-cut international either. Instead it is a narrative in which the highest proportion of conflict is between Irish and Viking but with large amounts of green-on-green and blue-on-blue conflict too. On the spectrum ranging from international to intranational conflict, which represents various degrees of the traditional to the revisionist views, the negative *Cogadh* networks are clearly on the traditional side albeit at a moderate rather than a limiting value (Figure 2.3). This is the main conclusion of our analysis of *Cogadh Gaedhel re Gallaibh*.

The analysis of assortativity so far investigates the extent to which conflict or positivity reigns within or between the two cultural sets. However, it may be argued that the revisionist concern is only with the Irish side. The revisionist view is that the hostility is mainly within the Irish community — it is not that it is both within the Irish cast and within the Viking set. The humanities literature tells us that there was a great degree of such conflict too; e.g., Ryan states

Table 4.9: Cultural assortativities. The left column identifies whether all nodes (Irish, Viking and other) are included in the calculation of ρ or if the unassigned (other) nodes are omitted. In the former case, $\rho_{\min}$ is calculated by Eq.(3.4.43) and in the latter case, it is -1. The next column identifies whether all remaining links are included or whether the Viking-on-Viking edges are omitted.

Nodes	Edges	Positive Network (ρ^+)	Negative Network (ρ^-)
All nodes included	Include all edges	0.65(3)	-0.32(6)
	Omit Viking-on-Viking edges		-0.45(5)
	$\rho_{\min}$	-0.62(3)	-0.88(4)
Other nodes omitted	Include all remaining edges	0.72(3)	-0.37(6)
	Omit Viking-on-Viking edges only		-0.53(4)
	$\rho_{\min}$	-1	-1

"The Norse were traditionally unscrupulous in preying upon one another" [19]. Therefore, Viking-on-Viking conflicts could contaminate our measurements. Our intention is to determine whether the Irish are mostly fighting with other Irish or if they are mostly in conflict with Vikings. In this sense, whether Vikings were also fighting amongst themselves is irrelevant.

To investigate this further detail, and to take these aspects into account, we remove Viking-on-Viking links from the negative sub-network. Redetermining the cultural assortativity, we find that $\rho^- = -0.45(5)$ [$\rho^- = -0.53(4)$ if the unassigned nodes are omitted]. This is indeed larger in magnitude than the previous measure (because the assortative Viking-on-Viking edges having been removed). But it is still not a clear-cut Irish-versus-Viking picture. In other words the value of ρ^- is still not close to $\rho^-_{\min} = -0.88(4)$ (or -1 in the case where unassigned nodes are removed). Therefore our conclusions are unchanged. The statistics are listed in Table 4.9.

In Section 3.4, to overcome the difficulties presented by network-dependent $\rho_{\min}$-values, we introduce a renormalised categorical assortativity metric which ranges from -1 in the case of fully disassortative networks through zero for uncorrelated networks to 1 for fully assortative networks. In Table 4.10 we also present an alternative to Table 4.9, using these renormalized values.

Table 4.10: The renormalised cultural assortativity values $\hat{\rho}$ from Eq.(3.4.45) presented here is an alternative to Table 4.9. Completely disassortative, uncorrelated, and assortative networks have $\hat{\rho} = -1$, $\hat{\rho} = 0$ and $\hat{\rho} = 1$, respectively. The introduction of a renormalised version of the categorical assortativity ($\hat{\rho}$) is suitable for all circumstances:

Nodes	Edges	Positive Network ($\hat{\rho}^{+}$)	Negative Network ($\hat{\rho}^{-}$)
All nodes included	Include all edges	0.65(3)	-0.32(6)
	Omit Viking-on-Viking edges		-0.43(5)
	$\rho_{\min}$	-1	-1
Other nodes omitted	Include all remaining edges	0.72(3)	-0.33(6)
	Omit Viking-on-Viking edges only		-0.44(5)
	$\rho_{\min}$	-1	-1

4.8 Remarks

The popular tradition associated with the Viking Age in Ireland and the events leading to the Battle of Clontarf in 1014 is that Brian's principal opponents were Vikings. In 1938, Ryan stated that contemporary annalists "portray Clontarf in terms consistent with the later popular tradition" [19]. Ryan's paper [19] contains what has been described as an "assault" on that traditional interpretation [29]. Instead of a "clear-cut" Irish-versus-Norse conflict, the revisionist argument is that it was a struggle primarily between two Irish forces. Munster and some allies on the one side were in conflict with Leinster and allies on the other, each having some Viking support. The fundamental issue was not, therefore, Irish versus Norse, but the intention of Leinster to be independent of the high king [19].

With the 1000th anniversary of the Battle of Clontarf, Seán Duffy attacked what he calls "the new orthodoxy" [29] and launched a counter-revisionist defence of the traditional picture [28]. Notwithstanding the bombastic and partisan tone of *Cogadh Gaedhel re Gallaibh*, Duffy provides a great deal of evidence that it is a valuable source if used correctly. His own use of that and other texts leads him to conclude that "The Battle of Clontarf was an international contest" [29]. This counter-revisionist stance has itself come in for criticism [30]. Thus the anniversary reinvigorated a lively set of discussions and healthy debate amongst experts and the wider public. This debate, together with the 150th anniversary of Todd's famous translation

[35], form the backdrop to which the above results are presented.

It is widely agreed that *Cogadh Gaedhel re Gallaibh* is a skillfully written propagandistic text, containing bias and exaggerating virtues and vices of many of its characters [35, 45, 48]. It has been used by both sides of the debate to support their arguments. For example, Duffy describes it as a "long narrative of Irish conflict with the Vikings" [28]. Etchingham says that "even Cogadh actually identifies the Leinstermen as principal rebels" [30]. From the side which fought against Brian at Clontarf, the *Cogadh* gives the majority of the slain (3100 out of 5600) as Irish [19, 35], amounts which could perhaps be viewed as supporting the picture of a mostly domestic conflict. At least these head-counts indicate that *Cogadh Gaedhel re Gallaibh* does not pretend that Viking slain exceed the numbers of Leinstermen in order to "internationalise" the story. This suggests that, even though *Cogadh Gaedhel re Gallaibh* exaggerated qualities, it might not have exaggerated quantities. Ryan himself believes that the account of the actual battle of Clontarf in the *Cogadh* is "incomparably the most reliable".

Here we have gone beyond a simple head-count of the slain and carried out a character-network analysis of *Cogadh Gaedhel re Gallaibh*. Because this is independent of the tone of the account (its bombastic and partisan nature) and its shortcomings ("telescoping" of events and "cavalier" attitude to chronology), we believe this is a judicious use of the text. To contribute to the debate about the nature of the Viking Age in Ireland as set down in the *Cogadh*, we developed a new measure, ρ, which we termed cultural assortativity and which can take proportions of Irish and Viking nodes into account. As we have seen, a literal interpretation of "the popular tradition of Clontarf as wholly an Irish-Norse" conflict [19] suggests a strongly negative value of cultural assortativity for the negative (conflictual) network. The revisionist picture of a "civil war" [62] or an "internal struggle" [43], on the other hand, with Leinster as the "predominant element" [19] or "principal rebels" [30], suggests a positive value of cultural assortativity for the negative network. The main outcome of the investigation reported upon in this Chapter is the measured values of the associated metric and we find negative values, supportive of the traditional picture. But magnitudes are moderate, suggesting that, at least in network terms, *Cogadh Gaedhel re Gallaibh* does not describe a "clear-cut" Irish versus Norse conflict. Complementary information is gained from the positive network. The cultural assortativities measured on the positive networks suggest that positive (social) interactions are mainly, but not entirely, between nodes of the same cultural identity. To summarise, *Cogadh Gaedhel re Gallaibh* describes the Viking Age in Ireland as predominantly an Irish-Norse conflict but not wholly so.

Obviously all of the conclusions drawn here refer to the content of *Cogadh Gaedhel re Gallaibh*. A statistical analysis can only be as good as the data it draws from and, all the claims to the contrary notwithstanding [19, 28], if one considers *Cogadh Gaedhel re Gallaibh*, in the main, to be unreliable, invented or concocted then little can be drawn from it about reality. Even if the author of the *Cogadh* tried to record as much as possible about encounters and relationships between personages of the time, we have no way of knowing if some types of interactions are under-reported. For example, the Irish author may not have been able to record interactions between Vikings with an accuracy comparable to that with which he recorded interactions involving Irish players. Even in these cases, texts such as the *Cogadh* still contain important information on how writers tried to, or were able to, portray the society and the perceptions, expectations and concepts of the time they was committed to writing. Since the author of the *Cogadh* could not have anticipated a complexity-scientific analysis nearly 1000 years later, the networks approach delivers unique insights in that it extracts a perhaps unintended message from that time. One may be more optimistic than this and hope that a reasonable proportion of characters and their interactions reflect the reality of the age. Indeed, we have seen that omitting Viking-Viking interactions does not alter the broad conclusions of our study so, even if the author did under-report them, our conclusions are unchanged. If one considers the main thrust of the *Cogadh* account to be sound, the network approach delivers new, quantitative knowledge of the Viking Age in Ireland.

A second question which we addressed in this Chapter is how the *Cogadh* narrative compares to others of the epic genre [9, 15, 16]. We measured a number of other statistics characterising the *Cogadh* network and found it has properties typical of the genre, resembling the *Iliad* more than *Njáls Saga* or *Táin Bó Cuailnge*. The similarity to the *Iliad* is interesting because a link has been suggested before by scholars from humanities, using traditional methods [36, 40, 47, 60]. It would be fascinating to continue such comparative investigations at a more detailed level in future studies.

In summary, we conclude from this Chapter that the character networks contained in *Cogadh Gaedhel re Gallaibh* support neither clear-cut traditionalist nor clean-cut revisionist depictions of the Viking Age in Ireland. Instead they suggest a moderate traditionalist-type picture of conflict which is mostly between Irish and Viking characters, but with significant amounts of hostilities between both sides as well.

Chapter 5

Network Analysis of the *Poems of Ossian*

In this Chapter, we compare the network structures in the *Poems of Ossian* to those of four other texts. These are *Acallam na Senórach* (referred to as *Acallam*, henceforth); Lady Gregory's text; the *Iliad*; and the *Odyssey*. We begin in Section 5.1 by comparing some quite standard topological measures of the networks underlying each text. These statistics, together with the degree distribution studied in Section 5.2, encapsulate a variety of characteristics of networks. By comparing them between different networks, we obtain quantitative indications of similarities and differences between them. In previous studies, the literary usage of such an approach was comparison and classification of epic narratives and sagas [15, 16]. Our objective here is to measure these statistics for the social network underlying *Ossian* and those from texts to which Macpherson's has been compared. Our aim is to use these to formulate quantitative comparisons. We also probe them in more detail and analyse their degree distributions and look at the spectral distances from the adjacency matrices of these graphs.

After establishing that the Ossianic networks are more similar to networks appearing in Irish mythology in Section 5.2, we redo the entire analysis restricted to the giant components of the various networks in Section 5.3. This is to establish that, not only is *Ossian* similar to *Acallam*, but Lady Gregory's text is too. This is what one would expect if the latter is derived from the former and, in a sense, this benchmarks our approach. We take this opportunity to list the relative importance of individual characters in Section 5.4 in the hope that such information may be useful to humanities researchers in the future. We end the Chapter with a short discussion of the significance of its results.

The results of this chapter have been published in *Advances in Complex Systems* as Ref. [9].

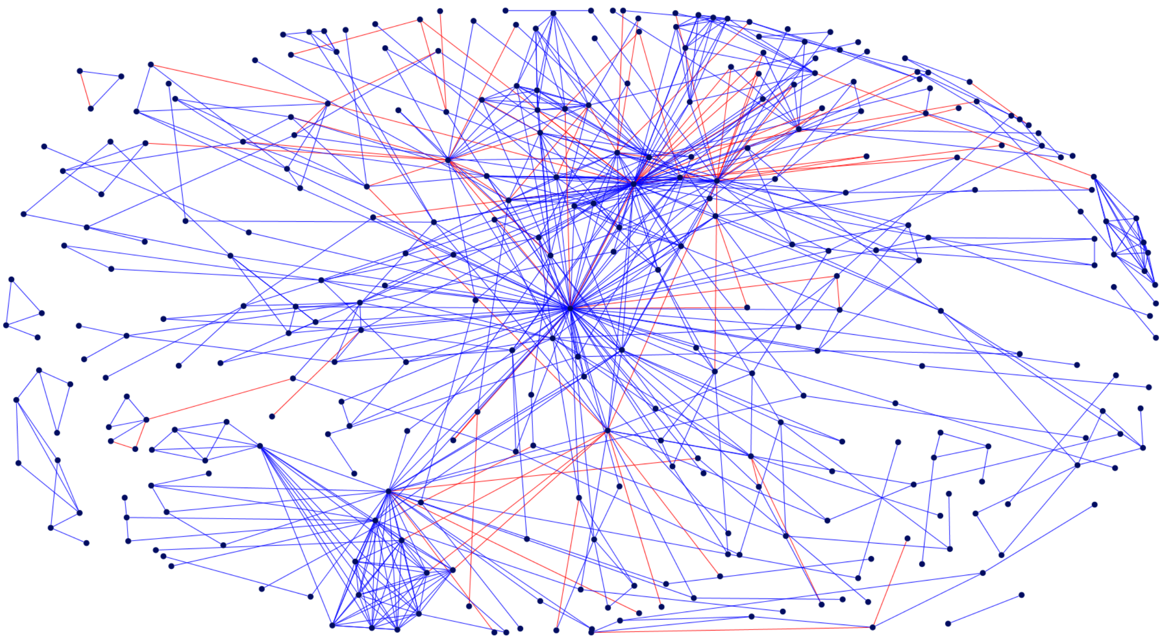

Figure 5.1: The entire network of 748 edges between 325 characters from MacPherson's *Ossian*. Positive (friendly) and negative (hostile) edges are in blue and red, respectively.

5.1 Network statistics

To construct the Ossianic network we use Gaskill's modern scholarly edition, *The Poems of Ossian and Related Works*. This has five mains sections: *Fragments of Ancient Poetry* (1760); *Preface to* Fingal (1761/62); *The Works of Ossian* (1765): Fingal; *The Works of Ossian* (1765): Temora; and *Preface to* the Poems of Ossian (1773). We use the material from 1765 only (*Fingal* and *Temora*) because these comprise the entire Ossianic corpus. (The poems in *Fragments* are reproduced and expanded in *Fingal*, with occasional modifications to the spelling of characters' names.) The full network is displayed in Figure 5.1.

In Table 5.1, we list some of the network statistics for *Ossian* and the texts to which we wish to compare. The statistics generally indicate that, in each case, the full networks have quite similar properties to the positive ones from the same narrative and are quite different to their negative counterparts. This can be interpreted as reflecting that, although conflict is an important element of each narrative, the overall network properties are dominated by positive social interactions. Indeed, although *Ossian* is generally seen as a work replete with melancholy, its 666 positive edges between 309 nodes and 82 negative edges between 87 nodes would imply that the former dominate the Ossianic system. Strong degrees disassortativity values suggest that the negative networks are held together by hubs of strongly connected characters who are in conflict with multitudes of lesser significant individuals. Short average path lengths compared to equivalent sized random networks together with the absence of clustering mean they are not small-world. The negative networks are therefore not reflective of genuine societies. Moreover, similarities between their relatively trivial topologies do not allow the negative networks to deliver information that may be used to categorise or differentiate between the different narratives.

Like many complex networks that have previously been studied, the full and positive societies listed in Table 5.1 are structurally balanced small worlds. They have non-trivial topologies and large proportions of their nodes belong to their giant components [107, 109]. They are structurally balanced and the full networks exhibit high values of Δ. These statistics therefore capture what appear to be universal properties of mythological networks (see Refs. [15, 16]) and invite comparison. We see, for example, that the average degrees of *Ossian* are more similar to those of *Acallam* and Lady Gregory's text than they are to those of the Homeric epics. This might be interpreted as suggesting that networks in Macpherson's text may bear closer structural similarities to the Irish than the Greek. However, although the path lengths of *Ossian* are close to those of *Acallam*, they are also quite close to those of the *Iliad* and rather different to those of Lady Gregory's text. These likewise have quite different values of

Table 5.1: Properties of the full, positive and negative networks for *Ossian* and the texts to which we compare it. Here, "*Acallam*" means *Acallam na Senórach* and "*Gregory*" represents Lady Gregory's text.

	Narrative	N	M	$\langle k \rangle$	ℓ	ℓ_{rand}	C	C_{rand}	C_T	r	Δ	G_c
Full	*Ossian*	325	748	4.60	3.62	3.91	0.49	0.01	0.27	-0.08	0.95	88.62%
	Acallam	732	1606	4.39	3.79	4.57	0.37	0.01	0.19	-0.10	0.97	76.91%
	Gregory	355	913	5.14	3.10	3.73	0.44	0.01	0.16	-0.18	0.97	94.65%
	Iliad	694	2684	7.74	3.49	3.42	0.44	0.01	0.45	-0.08	0.98	99.42%
	Odyssey	301	1019	6.77	3.29	3.18	0.45	0.02	0.38	-0.08	0.97	98.34%
Positive	*Ossian*	309	666	4.31	3.65	4.03	0.42	0.01	0.31	-0.06		82.20%
	Acallam	722	1513	4.19	3.72	4.69	0.38	0.01	0.20	-0.09		71.33%
	Gregory	337	833	4.94	3.23	3.78	0.45	0.01	0.18	-0.17		91.99%
	Iliad	640	2329	7.28	3.80	3.47	0.44	0.01	0.58	0.02		85.94%
	Odyssey	299	989	6.62	3.42	3.21	0.45	0.02	0.40	-0.08		97.32%
Negative	*Ossian*	87	82	1.89	5.30	6.61	0.00	0.02	0.00	-0.31		70.11%
	Acallam	86	93	2.16	2.32	5.53	0.00	0.00	0.00	-0.30		24.42%
	Gregory	95	80	1.68	4.75	8.13	0.00	0.01	0.00	-0.30		45.26%
	Iliad	321	355	2.21	4.46	7.00	0.00	0.00	0.00	-0.45		90.34%
	Odyssey	41	30	1.46	1.88	8.74	0.00	0.04	0.00	-0.18		26.83%

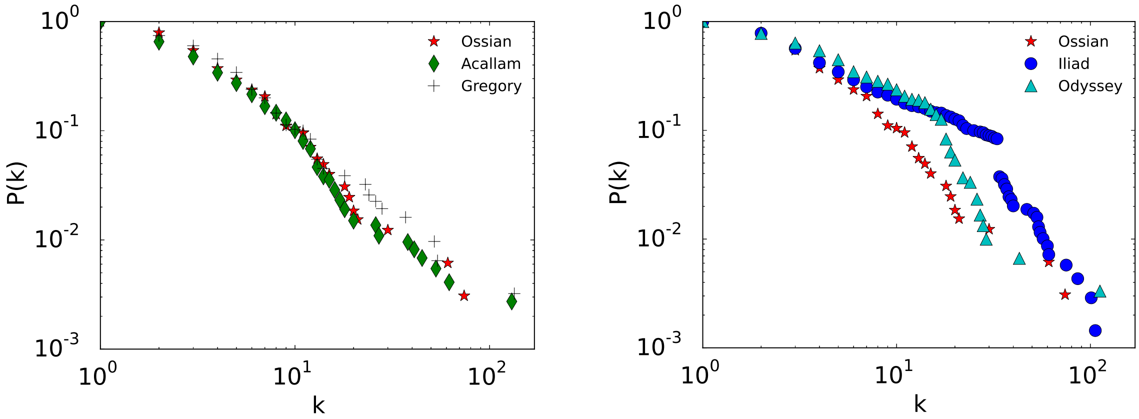

Figure 5.2: The complementary cumulative degree distributions of the full networks suggest that the society depicted in *Ossian* more closely resembles those of the Irish *Acallam na Senórach* and Lady Gregory's text (denoted here by "*Gregory*") than the *Iliad* or *Odyssey*.

assortativity. The various clustering coefficients fail to distinguish similarities too. Therefore, although the network statistics are informative in a wide sense, they do not constitute precise enough a tool to signal whether the Ossianic networks are more similar to those of the Irish or the Classics. To investigate further we turn next to the degree distributions.

5.2 The Degree Distributions

As mentioned before, the degree distribution $p(k)$ of a network governs the probabilities for nodes to have certain numbers of edges. In the next sub-section, we examine and compare the functional forms of our various networks.

5.2.1 Functional Forms

The functional form of the decay gives us important information about the structure of the connectivity of the network. Rather than displaying the probability distribution functions themselves, the convention is to present the complementary cumulative degree distributions $P(k)$ because it reduces the noise in the tail [155]; These cumulative distributions for the networks are plotted in Figure 5.2. In the left (right) panel, the degree distribution for the full *Ossian* network (all components and including both positive and negative edges) is compared to those of the Irish (Greek) texts. (See Section 5.3 for equivalent plots restricted to the giant components of the positive networks.) The distribution for *Ossian* more closely resembles those

from the Irish corpus. To analyse further, we fit different standard functions to these various distributions. (In Sub-section 5.2.2, we also apply the Kolmogorov-Smirnov test to investigate the similarities between degree distributions in a parametric-independent manner.)
Here we fit data for each narrative to each of the functional forms described in Chapter 3. Again we use maximum likelihood estimators to determine the parameters for each probability distribution as described in Sub-section 3.2.2.1. To determine which of these models are most likely, we use the Akaike information criterion (AIC) Sub-section 3.2.2. The results are listed for each network in Table 5.2. We also present the outcomes of the maximum-likelihood fits to the various distributions in Table 5.3.
The results indicate that, of the various candidate models that we have examined for the degree distributions, the networks underlying *Ossian*, *Acallam na Senórach*, and Lady Gregory's text are most likely to be log-normal with parameters μ between 0.5 and 1.2 and σ about 1 or a little above. However, in the case of *Ossian*, one cannot rule out a Weibull distribution. A truncated

Table 5.2: Relative minimum-information-loss probabilities for the degree distribution functions of the full networks underlying to the considered narratives. The probabilities corresponding to the most likely distribution for a given narrative is highlighted in boldface.

	Power law	Truncated power law	Log normal	Exponential	Weibull
Ossian	~ 0	0.11	**0.54**	~ 0	0.35
Acallam	~ 0	~ 0	**0.97**	~ 0	0.03
Gregory	~ 0	~ 0	~ **1**	~ 0	~ 0
Iliad	~ 0	~ **1**	~ 0	~ 0	~ 0
Odyssey	~ 0	0.32	0.16	~ 0	**0.52**

Table 5.3: Maximum-likelihood estimates for the various parameters associated with the probability distributions fitted to the data for the full character networks.

	Power law	Truncated power law		Log normal		Exponential	Weibull	
	γ	γ	κ	μ	σ	κ	β	κ
Ossian	2.1(7)	1.4(1)	15.7(6)	**0.5(7)**	**1.2(0)**	4.0(1)	0.5(2)	0.6(6)
Acallam	2.0(5)	1.4(1)	18.9(4)	**0.8(5)**	**1.1(8)**	4.7(1)	0.5(1)	0.8(4)
Gregory	1.9(6)	1.2(1)	17.3(6)	**1.2(6)**	**1.0(3)**	5.2(1)	0.5(2)	1.3(4)
Iliad	1.8(4)	**1.4(0)**	**39.6(3)**	0.0(8)	1.8(0)	8.1(1)	0.4(1)	0.5(6)
Odyssey	1.8(6)	0.7(1)	12.2(12)	1.5(7)	1.0(2)	6.9(2)	**0.7(4)**	**4.3(2)**

power law with $\gamma \approx 1.4$ and $\kappa \approx 39.6$ is the most probable distribution for the *Iliad* network and the Weibull distribution with $\beta \approx 0.7$ and $\kappa \approx 4.3$ is most likely for the *Odyssey*. However, a truncated power-law or a log-normal distribution cannot be dismissed for the *Odyssey*. In Section 5.3 we give the relative probabilities and maximum-likelihood parameter estimates for the giant components of the positive networks too.

5.2.2 Kolmogorov-Smirnov Comparisons

The AIC cannot deliver information about the quality of a particular model in an absolute sense. It can only deliver information on relative quality. To offer an alternative, we next investigate similarities between the degree distributions associated with different narratives in a non-parametric manner. An advantage of the Kolmogorov-Smirnov test is that it does not require knowledge of the distributional forms. Although it was developed for continuous distributions, it can be used in the discrete case provided one is interested in comparing two samples to each other rather than comparing one sample to a probability distribution function [156–158]. This is the situation that exists in the present instance. If $f(k)$ and $g(k)$ are the two empirical distribution functions, the test statistic to be used is $D = \sup|f(k) - g(k)|$. The appropriate null hypothesis is that the samples are drawn from the same underlying distribution. Assuming this, the p value gives the probability that the two distributions are as different as those observed. If the p value is large (i.e., if $p > 0.05$), we cannot reject the null hypothesis and if it is small (i.e., if $p < 0.05$), we may conclude that the two samples are likely to have come from different populations with different distributions. In the current context, we interpret large p values as indicating similarities between the degree distributions of the networks underlying two narratives. We describe such a circumstance as a "match". Small p values suggest dissimilarities. The results for the full and the positive networks (considering all components in each case) are presented in Table 5.4. Unlike in the previous chapter, here we are not interested in hostile networks as any such similarities merely reflect their trivial topologies rather than interesting societal structure. Indeed we checked, and the negative networks from every text analysed matches every other, indicating another universal, non-discriminating feature of such narratives.

From the table, it is clear that whether we use the full network or its positive sub-network, there are strong matches between the social structures of *Ossian* and *Acallam na Senórach*. The test do not detect a match between *Ossian* and either of the Homeric texts. In Section 5.3 we apply similar tests to the giant components of the various networks. These detect matches between *Ossian* and both of the Irish texts in addition to the match between *Acallam na Senórach* and

Lady Gregory's version. Again, however, they detect no match between *Ossian* and the Homeric networks. (See Table 5.9 of Section 5.3.)

5.2.3 Spectral Distances

Besides comparing the structural properties of the societies underlying the various narratives through network statistics and degree distributions, it is interesting to compare two networks directly. A network's adjacency matrix contains all of its connectivity information and one way to compare these is to use the Ipsen-Mikhailov (IM) distance [159]. Although originally defined for dynamical biological networks, the approach is quite robust and can be considered wherever a quantitative comparison between networks is needed [160].

The adjacency matrix A of a network is constructed so that element A_{ij} is 1 if there is an edge from node i to node j and 0 otherwise. In our case, all diagonal elements are zero because there are no edges from a node to itself (loops) in the simple graphs considered here. The matrix A is also symmetric because our networks are undirected. We construct the network's so-called Laplacian matrix L by defining its elements as $L_{ij} = \delta_{ij}k_i - A_{ij}$ where δ_{ij} is the Kronecker delta function which is 1 if $i = j$ and 0 otherwise. The diagonal of the Laplacian is the *degree matrix*. This Laplacian matrix has $N - 1$ eigenvalues λ_i [119]. The spectral function $\rho(\omega)$ is a sum of

Table 5.4: Results of the Kolmogorov-Smirnov tests, in terms of p values, applied to the full and positive networks. When they differ, data for the latter are given in square brackets. The large values of p (highlighted in boldface) in the cell corresponding to the *Ossian* column and *Acallam* row imply similarities in the corresponding degree distributions. The small values in remaining cells suggest the absence of similarities between the corresponding character networks.

	Ossian				
Full	**0.79 [0.95]**	*Acallam*			
[Positive]	$< 10^{-2}$ [0.04]	0.03 [0.02]	*Gregory*		
Networks	$< 10^{-5}$	$< 10^{-5}$	$< 10^{-2}$	*Iliad*	
	$< 10^{-3}$	$< 10^{-3}$	$< 10^{-3}$	$\leq 10^{-2}$	*Odyssey*

Table 5.5: Spectral distances from the networks in *Ossian* to those of the other texts we examined. Lower values indicate a greater degree of similarity between the networks.

Distance from *Ossian*	*Acallam*	*Gregory*	*Iliad*	*Odyssey*
Full network	0.03	0.02	0.07	0.09
Positive sub-graph	0.03	0.03	0.08	0.10

Cauchy-Lorentz distributions, defined as

$$\rho(\omega) = K \sum_{i=1}^{N-1} \frac{\gamma_w}{(\omega - \sqrt{\lambda_i})^2 + \gamma_w^2}, \tag{5.2.1}$$

in which K is a normalisation constant given by $\int_0^\infty \rho(\omega)d\omega = 1$. Here, γ_w is a scale parameter which represents the width of each distribution. Its value will be made explicit once the distance is defined.

Each network has a Laplacian matrix and an associated spectral function. It follows that a measure of the distance between two any spectral functions can be interpreted as the distance between the networks themselves. Labelling the networks by 1 and 2, the IM-distance is defined by

$$\epsilon = \sqrt{\int_0^\infty [\rho_1(\omega) - \rho_2(\omega)]^2 d\omega}. \tag{5.2.2}$$

This quantity vanishes when the two networks are identical $[\rho_1(\omega) = \rho_2(\omega)]$. The parameter γ_w acts as a scale and one sets it by considering a complete graph (of the same size as Network 1, for example) and an empty graph (of the same size as Network 2). Choosing γ_w so that $\epsilon = 1$ for these networks sets the IM-distance to be 1 when two networks are maximally different. Keeping this particular value of γ_w then allows us to evaluate the distance for the two networks of interest.

As the networks considered here are all relatively sparse, low values of ϵ are expected. Still, they can serve to indicate which networks are most similar and dissimilar to each other. The results are listed in Table 5.5 as the IM-distances from *Ossian* to each of the other texts. They suggest that the Ossianic networks are two or three times further from the Homeric networks than they are to those in *Acallam na Senórach* or Lady Gregory's text. The results for the giant components are given in Section 5.3.

Therefore all three approaches that we have used (namely based on parametric fits, Kolmogorov-Smirnov tests and IM distances) deliver the same result. That result is that the social network structure of *Ossian* bears a closer resemblance to those of the Irish corpus than to the Homeric narratives. This is interesting and significant because Macpherson and his allies strove to parallel the Classical texts and explicitly tried to distance themselves from the Irish [101].

5.3 The Giant Components

In Table 5.4 of this chapter, we saw evidence for matches between the character networks of *Ossian* and *Acallam na Senórach*. However, the test did not match the full networks *Acallam na*

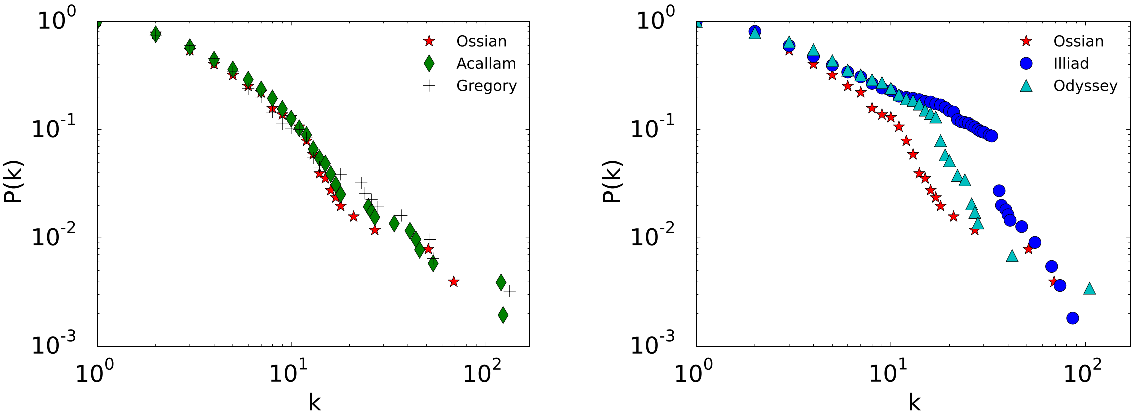

Figure 5.3: As with Figure 5.2, the degree distributions for the giant components of the positive sub-networks depicted in *Ossian* more closely resembles those of *Acallam na Senórach* and Lady Gregory's text (denoted here by "*Gregory*") than the *Iliad* or *Odyssey*.

Senórach and Lady Gregory's text. It is reasonable to expect that these should match because the former is acknowledged as a source for the latter. If our approach is valid, it should be able to detect this. For these reasons we repeat part of the analysis for the giant components of the positive sub-networks. We begin with Table 5.6 where we list the network statistics.

In Section 5.2 we also established that the degree distribution of the social network underlying Macpherson's *Ossian* is similar to those of *Acallam na Senórach* and dissimilar to the *Iliad* and *Odyssey*. In Figure 5.3 we display the equivalent plot for the giant components of the positive sub-networks. Table 5.7 lists the corresponding probabilities for the various degree distribution functions. The corresponding estimates of the various parameter are given in Table 5.8. In Table 5.9 the results of the Kolmogorov-Smirnov tests are displayed.

The match between *Ossian* and *Acallam na Senórach* that was identified in Table 5.4 is repeated

Table 5.6: Properties of the giant components of the positive networks for *Ossian* and the four comparative texts.

Narrative	N	M	$\langle k \rangle$	ℓ	C	C_T	r_k
Ossian	254	596	4.69	3.62	0.45	0.30	-0.10
Acallam	515	1364	5.30	3.71	0.41	0.20	-0.13
Gregory	310	808	5.21	3.24	0.46	0.18	-0.18
Iliad	550	2275	8.27	3.78	0.48	0.57	0.05
Odyssey	291	987	6.78	3.38	0.46	0.39	-0.08

here but this time there is also a match to Lady Gregory's text. The expected match between Lady Gregory's text and *Acallam na Senórach* is also evident. There is still no sign of similarity between *Ossian* and either of the Homeric texts. Comparing the results of Table 5.9 with those of Table 5.4, then, it would seem that the giant components of the positive sub-networks deliver even stronger results than those of the full networks.

Table 5.7: Relative minimum-information-loss probabilities for the degree distribution functions of the giant components of the positive networks. The probabilities for the most likely distribution for a given narrative are highlighted in boldface.

	Power law	Truncated power law	Log normal	Exponential	Weibull
Ossian	~ 0	0.17	**0.45**	~ 0	0.38
Acallam	~ 0	~ 0	0.98	~ 0	0.02
Gregory	~ 0	~ 0	**0.99**	~ 0	0.01
Iliad	~ 0	~ **1**	~ 0	~ 0	~ 0
Odyssey	~ 0	0.33	0.17	~ 0	**0.50**

Table 5.8: Maximum likelihood estimates for the parameters associated with the probability distributions fitted to the data coming from the giant components of the positive sub-networks.

	Power law	Truncated power law		Log normal		Exponential	Weibull	
	γ	γ	κ	μ	σ	κ	β	κ
Ossian	2.1(8)	1.2(1)	12.4(8)	**0.8(8)**	**1.1(11)**	4.3(1)	0.5(3)	1.2(5)
Acallam	1.9(5)	1.2(1)	15.0(6)	**1.1(5)**	**1.0(3)**	5.1(1)	0.5(2)	1.5(3)
Gregory	1.9(6)	1.2(1)	15.7(7)	**1.2(6)**	**1.0(3)**	5.1(1)	0.5(3)	1.4(4)
Iliad	1.8(4)	**1.3(0)**	**40.0(4)**	0.3(8)	1.7(0)	8.5(1)	0.4(1)	0.7(5)
Odyssey	1.8(6)	0.6(1)	11.6(13)	1.5(7)	1.0(2)	6.9(2)	**0.7(5)**	**4.5(2)**

Table 5.9: Results of the Kolmogorov-Smirnov tests for the giant components of the full and positive sub-networks. Large values of p indicate similarities in degree distributions and are highlighted in boldface.

Giant	*Ossian*				
Components	**0.30 [0.72]**	*Acallam*			
of Full	**0.07 [0.20]**	**0.75 [0.64]**	*Gregory*		
[Positive]	$< 10^{-2}$	$< 10^{-5}$	$< 10^{-2}$	*Iliad*	
Networks	$< 10^{-2}$	$< 10^{-2}$	$< 10^{-3}$	$\leq 10^{-2}$	*Odyssey*

This conclusion is reinforced by the spectral distances. These are listed for the giant components of the positive networks in Table 5.10. These results can be compared to those of the full networks listed in Table 5.5. The distance from the networks of *Ossian* to the *Acallam na Senórach*, which was 0.03 in Table 5.5 is reduced to 0.01 and 0.02 here.
Therefore we conclude again that there are strong resemblances between the character networks of *Ossian* and those of the Fenian cycle of Irish mythology. Additionally we can confirm that Macpherson's networks are unlike those of Homer.

Table 5.10: Spectral distances from the giant component of the positive sub-networks of *Ossian* to those in the other texts. Lower values indicate greater similarity between networks.

Distance from *Ossian*	*Acallam*	*Gregory*	*Iliad*	*Odyssey*
Giant Component	0.01	0.02	0.07	0.08
Positive GC	0.02	0.03	0.09	0.08

Table 5.11: The most important characters of the Ossianic texts ranked according to their betweenness centrality, closeness, eigenvector centrality, and degree. We employ Gaskill's superscript notation to identify characters that share their names with others in the text.

	Rank	Betweenness	Closeness	Eigenvector
Full	1	Fingal (0.38)	Fingal (0.43)	Fingal (0.53)
	2	Ossian (0.26)	Ossian (0.40)	Ossian (0.48)
	3	Cuchullin (0.12)	Swaran (0.36)	Gaul^{1} (0.35)
	4	Swaran (0.08)	Carril (0.35)	Fillan (0.32)
	5	Gaul^{1} (0.06)	Gaul^{1} (0.35)	Fergus^{1} (0.17)
Positive	1	Fingal (0.38)	Fingal (0.40)	Fingal (0.44)
	2	Ossian (0.26)	Ossian (0.37)	Ossian (0.38)
	3	Cuchullin (0.12)	Carril (0.33)	Gaul^{1} (0.21)
	4	Swaran (0.08)	Swaran (0.32)	Fillan (0.19)
	5	Gaul^{1} (0.06)	Cuchullin (0.32)	Fergus^{1} (0.18)
Negative	1	Cairbar^{2} (0.34)	Cairbar^{2} (0.20)	Gaul^{1} (0.49)
	2	Swaran (0.27)	Ardan^{2} (0.20)	Ossian (0.48)
	3	Ardan^{2} (0.24)	Swaran (0.19)	Cremor^{1} (0.24)
	4	Gaul^{1} (0.24)	Fingal (0.18)	Cathmin^{1} (0.24)
	5	Fingal (0.17)	Gaul^{1} (0.18)	Leth^{1} (0.24)

Rank	Degree (k)
1	Fingal (74)
2	Ossian (61)
3	Cuchullin (30)
3	Gaul^{1} (30)
5	Cairbar^{2} (21)
1	Fingal (68)
2	Ossian (51)
3	Cuchullin (27)
4	Gaul^{1} (20)
5	Connal^{5} (18)
1	Cairbar^{2} (10)
1	Gaul^{1} (10)
1	Ossian (10)
4	Swaran (7)
5	Fingal (6)

5.4 Importance of Individual Characters in *Ossian*

The leading characters of *Ossian* are listed in Table 5.11. Fingal and Ossian are the two most ranked characters according to all four network measures for the full and positive sub-network. The network statistics listed in Table 5.1 are averaged over the entire network. We consider that literary scholars may also find statistics corresponding to individual characters in the Ossianic narratives of interest. Therefore we rank top characters according to various criteria in Table 5.11. In the index of names that accompanies Gaskill's text, superscripts are employed to identify characters that share a common name with other characters in the narrative. We employ this same notation in our identification of characters in Table 5.11.

5.5 Remarks

James Macpherson's *Ossian* had a profound and wide-reaching influence on Western culture. It acted as a catalyst for Romanticism in art and literature and in drawing attention to national folklore, mythology, and poetry throughout Europe and, indeed, the wider world. However, it has also been called a "curious blend of misreading and wilful misrepresentation" [161] and "one of the greatest literary hoaxes of all time" [83]. Ossianic scholarship has experienced a major revival in recent decades. In part this is due to revisionist interpretations surrounding the production of the texts and its context in the history of cultural colonisation. As a result of this, James Macpherson has enjoyed a degree of rehabilitation. Additionally Ossianic scholarship has seen a growing focus on research questions that are more subtle than the original and divisive questions of authenticity and fraud [92].

Macpherson employed footnotes, prefaces, and introductory dissertations to claim the value of the *Ossian* poems and their supposed relevance for the early history of Scotland and Ireland. He uses such paratexts to argue for the consideration of *Ossian* alongside established examples of ancient epic poetry. With the help of Hugh Blair, he deliberately sought to place *Ossian* within an established classical context. His aim in doing so was to legitimise and lend authority to the new Scottish epic. The paratexts were used to repeatedly compare the poems to those of Greek and Latin sources. Especially important to them were the comparisons to Homer. Moreover they tried to distance the poems from Irish sources. At the time, Irish sources were seen as having less cultural worth than the classical ones. Here, it has been shown using network analysis that, despite the efforts of Macpherson and his allies, *Ossian* has structural similarities to Irish epic tradition and not the Greek.

By determining the statistics which describe the structural connectedness of the character networks present in *Ossian* and those of the Irish and Greek texts, meaningful quantitative comparisons have been made. Whether full or positive networks are used, or their giant components, there is a clear match between *Ossian* and *Acallam na Senórach* (as well as Lady Gregory's tales of the Fianna in the case of the giant components). On the other hand, there is little or no detectable structural resemblance between *Ossian* and the works of Homer, whose networks are also unlike those of the Irish texts. Almost two hundred and fifty years ago, Charles O'Connor and Sylvester O'Halloran identified characters in the Ossianic network as appropriated from Irish tradition. These characters are nodes in the social-network interpretation that we have given here. Here it has been shown that the configuration of edges of the network also resembles that of Irish mythology, and not that of Homer. In this sense this chapter work complements and completes that of O'Connor and O'Halloran.

Chapter 6

“Ossianisation” of the Fenian Cycle

Here we follow on from the analysis in Chapter 5 in which it was established that the Ossianic networks are similar to those of *Acallam na Senórach* and (to a lesser extent) Lady Gregory’s text and are different to those of the Homeric epics. In Ref. [161], Thomson investigates the parallels between the narratives in *Ossian* and those from the Irish corpus which further our understanding of Macpherson’s relationship to oral and written Gaelic culture both in Ireland and in Scotland. However, Thompson’s qualitative approach cannot resolve the debate about the authenticity or otherwise of *Ossian*. It also cannot determine Macpherson’s intentions in constructing the work. Similarly, the structural similarity of the character networks present in *Ossian* to those of the Irish epics do not prove that these were Macpherson’s sources; correlations do not imply causation.

Certainly we cannot suppose that Macpherson had network theory in mind when he went about his work. Nonetheless, we can ask the question: how *would* one construct one network from another if one did have the *Acallam na Senórach* networks to hand. In this chapter we attempt to partially answer such a question. The groundwork prepared here will hopefully lay a route for further investigations of the issue and further publications on the theme.

The analysis in the previous chapter demonstrated that the KS-test failed to detect a match between the degree distributions of the entire networks of *Acallam* and Lady Gregory’s text. However, restricting the analysis to the giant components of the positive networks, we found that there is such a match. This was the benchmark for our analysis - for the method to be viable, it should detect a match between *Acallam* and Lady Gregory’s text. Therefore we restrict the further analysis in this chapter to the giant components of the positive networks. Indeed, one may surmise that both Macpherson and Lady Gregory may have, for the most part, been interested in characters taking part in the main action of the various texts — i.e., those in the giant components. So, the question we wish to ask is: starting from the giant component of the positive network of *Acallam na Senórach*, how would one go about constructing a network like that of *Ossian*? We refer to the former network as the *parent* and it has $N = 515$ characters. The counterpart *Ossian* network has 254 characters. We ask, how could one decimate the parent network to create *daughter* networks with approximately 254 characters and which (a) resemble the parent and (b) resemble the corresponding network in *Ossian*. In network theory the type of process corresponding to part (a) is called "sparsification". Here, we have the additional requirement (b) that the daughter network should match Macpherson's *Ossian*. We refer to the process that satisfies (a) and (b) as *Ossianification*.

Our approach is inspired by how we think Macpherson might have gone about his work, had he some knowledge of networks science. Having heard the stories of the Fianna, he may have randomly selected or remembered some characters from *Acallam* and used them and their interactions for the basis of the Ossianic networks. In doing so, he might have recalled to mind especially the stronger characters or indeed those centered around Oisín. We simulate such approaches using various methods, testing whether the daughter networks have properties similar to those of the parent as well as the Ossianic network. We are especially looking for similarities in the degree distributions, assortativity and other network properties like path length and clustering coefficient, which relate to small-worldness. We will show that the sparsification processes have varying degrees of success. We also consider the inverse process whereby a parent network is grown from the daughter. In this case we start with *Ossian*, grow it to the size of *Acallam* and try to match the Irish network.

In fact we start in Section 6.1 with this network-growth approach. We look at *uniform attachment* and *preferential attachment* [162] and find that results from both methods successfully generate matches. This suggests that some form of "inverse growth" method may be created to generate *Ossian* from *Acallam*. In Section 6.2 we set about looking for such an approach by decimating *Acallam* to scale it down to the size of *Ossian*. This is done using various sparcification techniques and we will find that some are indeed capable of delivering daughter networks which resemble that of *Ossian*.

6.1 Network growth

If we let $N = 515$ denote the size of the giant component of the positive *Acallam* network. $n = 254$ to that of *Ossian*. We also take $M = 1364$ and $m = 596$ for their respective numbers of edges. Taking *Ossian* to be the daughter graph g and Acallam to be the parent G, we apply growth processes so that the daughter network has to grow to about twice its original size. Here we use models of uniform and preferential attachment to try to achieve matches between parent and daughter.

6.1.1 Uniform attachment

Before we examine the process empirically, we look at some of the analytic considerations already extant in the literature [163]. Let us start with a connected network with N initial vertices. Additional vertices are randomly introduced to the network one at a time. Nodes are indexed in order of their time of birth, so that node i is born at time $t = i$.

We let $k_i(t)$ be the degree of node i at time t. At time $t = 0$, the connected network contains N nodes and at every time $t = N + j$, a new node is born. A newly born node i forms η undirected links with already-existing nodes at the time of birth $t = i$ so that $k_i(i) = \eta$. Because the new node distributes its η new links randomly over the t existing nodes at time t, the change in the expected degree of node i at time $t > i$ is

$$\frac{dk_i(t)}{dt} = \frac{\eta}{t}. \tag{6.1.1}$$

which leads to

$$k_i(t) = \eta \ln i + \text{constant}.$$

The constant must be such that $k_i(i) = \eta$ as previously stated. Therefore

$$k_i(t) = \eta + \eta \log\left(\frac{t}{i}\right). \tag{6.1.2}$$

If a newly born node, introduced to the network at time τ, has degree k at time t, it follows that the fraction of nodes having a degree less or equal to k (at time t) is

$$\frac{t - \tau}{t}. \tag{6.1.3}$$

Let us consider the node which is such that at time t its degree is $k = d$. We write $\tau_d(t)$ the time of its birth. It follows that, at time t, the fraction of nodes of degree less than d is given by

$$F_t(\leq d) = 1 - \frac{\tau_d(t)}{t}. \tag{6.1.4}$$

Using Eq.(6.1.2) we have

$$k_\tau(t) = \eta + \eta \log\left(\frac{t}{\tau_d(t)}\right), \tag{6.1.5}$$

leading to

$$\tau_d(t) = t \exp\left(\frac{\eta - d}{\eta}\right). \tag{6.1.6}$$

It follows that the cumulative distribution is an exponential distribution [163]:

$$F_t(\leq d) = 1 - \exp\left(-\frac{d - \eta}{\eta}\right). \tag{6.1.7}$$

Of course, in the previous chapter, we found no support for exponential distributions for any of our networks. However, in the above calculations we did not take into account the existing degree distribution functions of the daughter networks. One might expect that if these are large relative to the parent, a uniform attachment scheme may not generate sufficiently large numbers of new nodes and links so as to overwhelm the degree distribution of the daughter. However, if the parent is far greater in size than the daughter, one might expect uniform attachment to eventually result in a network whose degree distribution is exponential. In any case, while uniform attachment may be able to grow the daughter to have the same number of nodes as the parent, it does not achieve a similar number of edges. To achieve this, we need a second step. Therefore our empirical growing process is implemented in two parts:

1. 261 new nodes are added to the *Ossian* network containing $n = 254$ to achieve a network size of $N = 515$. Each newly born node arrives with one edge and is randomly linked to existing nodes thereby increasing the edges of the network to 857 edges.

2. The second step involves the addition of extra edges at random until a target of $M = 1364$ is achieved.

The above process is repeated 100 times before taking averages. The resultant networks are compared to those of *Acallam* for similarities in the degree distributions. Using the KS-test, results demonstrate that 67% of the networks grown in this way have similar degree distributions to *Acallam*. The results are displayed in Table 6.1. The average degree matches by construction and we find that the path length of the parent also matches the target. However, neither the clustering coefficient nor the degree assortativity are close to those of *Acallam*. KS tests show that uniform attachment changes the degree distribution so that only 1% of the newly grown networks match the original. Sixty-seven percent of them match the degree distribution of the mother, however. Encouraged by this result we explore an alternative approach, namely preferential attachment.

6.1.2 Preferential attachment

This idea of preferential attachment has been around since the beginning of the 20^{th} century. However it has been known under different names until the end of the 20^{th} century. Barabási referred to it as *preferential attachment* as it is known in the scientific corpus today [164, 165]. Derek de Solla Price called it the *cummulative advantage* in his work that describes the network of scientific publications [166,167]. An American sociologist Robert Merton in his contribution to science called it the *Mathew effect* in an attempt to address a phenomenon related to the reward system in science. His endeavour was to quell the imbalance between eminent scientists and comparatively unknown researchers in the way they receive credit for their contribution to science. The inspiration for his theory came from the gospel of St. Matthew which says "For everyone who has will be given more, and he will have an abundance. But the one who does not have, even what he has will be taken away from him" [168] (the rich gets richer phenomenon). Preferential attachment provides an understanding of the universal mechanism responsible for the emergence of the scale-free nature found in the degree distributions of networks [119, 164, 165]. Assuming that a network starts with N nodes. A new node, pairs with a pre-existing node to form an edge but unlike random attachment, the pre-existing node i attracts a node with probability p_i proportional to its degree. As an example, node i of degree k_i is k_i times more likely to be linked to a new node than a node of degree one. The probability p_i that an existing node i connects to a new node is k_i relative to the overall degree of all existing nodes. Thus,

$$p_i = \frac{k_i}{\sum_j k_j} \equiv \frac{k_i}{2m}.$$

Hence new nodes tend to favour highly existing connected nodes. This technique generates networks in which all nodes are connected in a single component [169]. The probability that node i is connected to $k \geq 1$ other nodes is

$$p(k) = \frac{4}{k(k+1)(k+2)} \approx k^{-3}. \tag{6.1.8}$$

Therefore, the degree distribution is a power-law with an exponent of −3 [164]. The preferential attachment is often generalised by considering the addition of c edges for every new node. This however does not affect the power-low exponent.

Our implementation is again a variant on this and follows the following growth process:

1. Firstly, we add 261 nodes, one at a time, with preference to higher degree nodes. This gives a resultant network with same number of nodes as those found in *Acallam*. However this process only creates 261 extra edges resulting in 857 edges.

2. The second and final step involves the addition of 507 edges at random to make up for the 1364 remaining edges.

Eighty-two percent of the 100 grown networks, pass the KS-test in comparison with the *Ossian*'s network. Compared with the *Acallam*, the results given are even more favourable with a KS pass of 95%. N and M match by default and we also have a match in the average degree. However, again the clustering coefficient and the degree assortativity of the grown network do not match the target. Table 6.1 lists the results of the network growth techniques with preferential attachment achieving more successful results than random attachment.

6.2 Network Sampling (Sparsification)

Sparsification is a process that reduces the size of a network whilst preserving structural and statistical properties. It is sometimes used for improved visualisation and data reduction to manageable sizes [170–174]. For instance, analysing the world-Wide-Web may require intensive computation, which would take a prohibitive amount of time to finish on the full network. In such cases, it becomes viable to work with the smaller network (subset of nodes and edges). However, the difficulty is in finding a sampling technique that maintains the properties of the main graph.

Network sampling is not a new problem, although its application here is. Sparsification is a non-trivial problem [175], and remains a challenging topic of research interest in network science. Importantly, the way sub-networks are constructed is likely to have an impact on its properties (degree distribution, average degree, clustering coefficient, etc). Research efforts in this area focus on understanding how the properties of the selected sub-network (daughter) are dependent on its original network (mother) and of the selection process. This problem has been identified as one of the main issues in network modelling [176]. In general the daughter and mother network's degree distribution are expected to be different. For some important type of networks, it has been shown that random sampling gives rise to networks with the same type of degree distribution but characterised by different parameters. A study of random sampling applied on to scale free networks was presented in [177], clearly showing that the degree distribution of a random sub-network would typically differ from its mother network. Random sampling as well as sampling strategies dependent on nodes' degree was presented in [178], again showing that for the majority of networks, daughter and mother networks properties are different. The authors of [179] have presented a study of the assortativity and clustering coefficient. There have been discussions on various sampling methods searching for one that

Table 6.1: Properties of the giant components of the positive networks for *Ossian*, *Acallam* and the two grown networks from uniform and preferential attachment. The columns "KS-test" shows the percentage pass out of 100 realisations.

Narrative	N	M	$\langle k \rangle$	ℓ	C	r	*Ossian* KS-test	*Acallam* KS-test
Ossian	254	596	4.69	3.62	0.45	-0.10		
Acallam	515	1364	5.30	3.71	0.41	-0.13		
Uniform	515	1364	5.30	3.87	0.10	0.07	1%	67%
Preferential	515	1364	5.30	3.58	0.10	0.00	82%	95%

preserves the properties of the mother network [170, 171]. A number of sampling methods were explored by Leskovec [172] who also was interested in addressing questions on what constitutes sampling success. Using the Kolmogorov-Smirnoff test (KS), the author concludes that the so-called random-walk and forest fire methods are amongst the two best, having also looked at random-node selection, random-edge selection and sampling by exploration where a node is selected at random before exploring it's neighbourhood. However the authors acknowledge that there is no one perfect method for all applications but recommend hybrid methods that combine random node selection and some sampling exploration. To evaluate sampling success, a comparison of some network properties is used. These properties includes the network diameter, average clustering coefficient, size of the giant component and the degree distribution. The authors recognise no conclusive list of network properties that can be used to evaluate sampling success. Some network properties of sampled networks behave differently depending on a particular application, a result also confirmed in [179]. As an example, disconnected components are expected where edge sparsification techniques are used. Mark Granovetter was interested in sampling methods that estimated network density and average degree for large populations [180]. The sampling method employed involved taking random sub-networks of size n from a population of size N. On the other hand, Sokolov *et. al* [181] in their work on sampling scale-free networks have shown that the sample mean degree is much smaller compared to that of the population. This result was confirmed in [182] in which they observed that the sample mean displays systematic deviations for $n \ll N$.
Figure 6.1 illustrates the task at hand; our aim is to find sparsification methods whose sub-network's degree distribution and network statistics approximately estimate observables from its mother network. The objective is to exploit various sparsification techniques by sampling from the original graph (G) to create a sub-network (sub-graph g). Successful or effective sampling methods are those that preserve the spectrum of the original network. The next sections look at various sampling methods.

6.2.1 Random Node selection

This method involves randomly selecting the desired number of nodes from the mother G. Already existing edges are considered only between selected nodes. The simplest approach uses a uniform random selection method whereby the chance of a node's inclusion in the daughter graph g is equiprobable. Another possibility is to attach a weight to all nodes in G based on some topological characteristics e.g the degree k_i and to randomly select nodes according to this weight. This sampling technique gives an inclination towards highly connected nodes. Random node selection has been claimed to preserve some properties for daughter network down to a

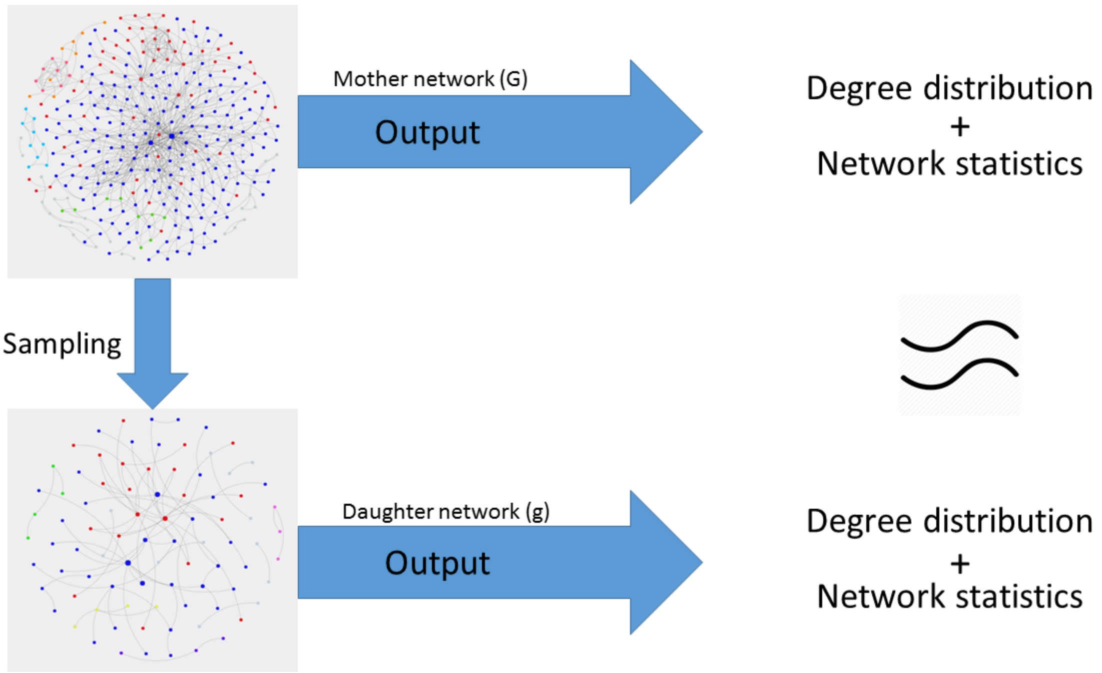

Figure 6.1: Network Sampling: Representative diagram of the task of sparsification in which one seeks to generate a daughter network from a parent while preserving network properties. The daughter should also satisfy a second condition which we call Ossianification: that is match network properties of *Ossian* .

size of 30% of the mother network [172]. However [181] found that results generally differ, with sampled networks producing a smaller average degree relative to that of the mother network.
Let us start by considering a graph G of size N, for which we write N_k the number of nodes of degree k. In addition, we write M the total number of edges and remind the reader that k_j denotes the degree of a particular node j. It follows that the degree distribution is given by $p(k) = N_k/N$, so that the average degree of the mother network G is

$$\langle k \rangle_G = \sum_k p(k)k = \frac{2M}{N}. \tag{6.2.9}$$

A daughter network g is constructed by randomly and independently selecting n nodes in G (the mother network) and including all mutual edges between the n nodes. The construction process of g is assumed to be Markovian: meaning that, at the step $n + 1$, the probability to add a given node j to the sub-graph is only a function of the sub-network g at the step n. We wish to consider two different situations:

1. A fully random selection process, where all nodes are equiprobable. This implies that the probability to add a particular node j to a sub-graph is n/N.

2. A preferential selection process where nodes are chosen according to their degree.

The probability to select a given node j is then proportional to k_j^β, where β is a parameter we are free to adjust. Note that by choosing $\beta = 0$, the probability to select a node is independent of its degree, reducing therefore to the fully random process defined in 1). Taking the limit $\beta \to \infty$, gives a systematic selection process, where nodes of maximum degree are selected.
One should mention that the probability to add a given node j does not depend on the existence of links connecting nodes together. Therefore, it is possible that such process generates disconnected sub-graphs and isolated nodes.

6.2.1.1 Unweighted node selection: $\beta = 0$

When considering a sub-graph g we define as κ_j the apparent degree of a node j. It is defined by the total number of links connecting j to all other nodes in the sub-graph g: $\kappa_j(g) = \sum_{i \in g} A_{j,i}$. Its average defines the average degree of the sub-network. It is given by the ratio

$$\langle \kappa \rangle_g = \frac{2\langle M \rangle_g}{n}, \tag{6.2.10}$$

where $\langle M \rangle_g$ denotes the average number of edges in g. This calculation remains quite simple as $p(i \in g)$: the probability to find the node i in the sub-network g, is constant, given by $p(i \in g) = n/N$. Summing over all edges M, and writing e_1 and e_2 the nodes at its extremities, we have

$$\langle M \rangle_g = \sum_e p(e_1 \in g) p(e_2 \in g) = M \left(\frac{n}{N} \right)^2. \tag{6.2.11}$$

In agreement with the results presented in [177, 178], it follows that the average degree is

$$\langle \kappa \rangle_g = \frac{n}{N} \langle k \rangle_G. \tag{6.2.12}$$

A simple argument can help to understand the previous equation. Allow yourself to select a node at random in the parent network G. On average, this node has a degree $\langle k \rangle_G$. Assuming this node is also part of the daughter network g. Its apparent degree $\langle \kappa \rangle_g$ will be smaller than its actual degree, as only a fraction f of its neighbours will be in the daughter network too. It follows that $\langle \kappa \rangle_g = f \times \langle k \rangle_G$, with $f = n/N$.
Using the data presented in table 5.6 where $n = 254$ (the number of nodes in the *Ossian*), leads to the prediction $\langle \kappa \rangle_g = 2.61$. This is quite far from the average degree measured in *Ossian*, namely $\langle k \rangle = 4.69$. Therefore we deem this approach to be unsuccessful.

6.2.1.2 Weighted node selection: $\beta > 0$

When looking to evaluate the average degree of the daughter network, we follow the previous calculation keeping in mind that this time, the probability for a node i to be in g is proportional to k_i^β. Using the equality

$$p(i \in g) = \frac{n}{N}\frac{k_i^\beta}{\langle k^\beta \rangle}, \tag{6.2.13}$$

and summing over all edges e, and writing e_1 and e_2 the nodes at its extremities, we have

$$\langle E \rangle_g = \sum_e p(e_1 \in g) p(e_2 \in g) = \left(\frac{n}{N}\right)^2 \frac{1}{\langle k^\beta \rangle^2} \sum_e k_{e_1}^\beta k_{e_2}^\beta. \tag{6.2.14}$$

To proceed further, we need to evaluate the sum over all edges in the previous equation. Let us approximate the adjacency matrix by

$$A_{i,j} \simeq \frac{k_i k_j}{N\langle k \rangle}, \tag{6.2.15}$$

which holds for uncorrelated networks. The previous approximation has a simple interpretation if one choose to think of $A_{i,j}$ of the probability to find a link between nodes i and j. It is then obvious that this probability is proportional to both, the number of links connected to i and j. Equation 6.2.14 becomes

$$\langle E \rangle_g \simeq \frac{n^2}{2N}\frac{\langle k^{\beta+1} \rangle^2}{\langle k \rangle \langle k^\beta \rangle^2}, \tag{6.2.16}$$

and leads to

$$\langle \kappa \rangle \simeq \frac{n}{N}\frac{\langle k^{\beta+1} \rangle^2}{\langle k \rangle \langle k^\beta \rangle^2}. \tag{6.2.17}$$

In the case of $\beta = 1$, this leads to $\langle k \rangle_g \approx (n/N)(\langle k^2 \rangle_G^2 / \langle k \rangle_G^3)$. Using $(\langle k^2 \rangle_G = 113.51$ we get $\langle k \rangle_g \approx 42.68$. The latter value is ten times larger than the average measured in the *Ossian*. This is due to nodes of a higher degree having a greater chance of being selected than nodes of a lower degree.

6.2.1.3 Taking fragmentation into account

Our strategy so far has been to start with the giant component of the positive *Acallam* network and to sparsify it by randomly selecting nodes to deliver a daughter of the same size as the giant component of the positive *Ossian* network. We have not attempted to take into account the fact that this process of sparsification generates isolated nodes, and indeed, small fragments, in the daughter networks. Therefore, we are comparing daughters which are fragmented to an unfragmented giant component of *Ossian*. Thus we are not comparing like with like and this is a potential reason why our efforts have failed so far.

Next we modify the above process so that, rather than identifying the full daughter networks with *Ossian* (which has size $N = 254$), we identify only their giant components. We have to permit the full daughter networks have varying sizes, dependent on the stochastic realisation at hand, in excess of $N = 254$, in order to deliver giant components comparable to $N = 254$. The results are listed in Table 6.2. We see that the network statistics broadly match the results found in both *Ossian* and *Acallam* (i.e. $\langle k \rangle$, ℓ, C and r).

This is an encouraging step and indicates that the process of fragmentation is in important one that has to be taken into account in any ossianification of *Acallam*. However, it turns out that this approach fails to match the degree distributions of the daughter networks to *Ossian*. This can be seen visually in Panel (a) of Figure 6.2 where the degree distributions are plotted in the case where $\beta = 0$. We can confirm with KS-tests that the process fails in this case and in the higher-β cases.

6.2.2 The random walk approach

Given a mother network $G = (V, E)$, with $N = |V|$ and $M = |E|$ denoting the total number of nodes and edges respectively, a random walk is a stochastic process that starts from a given node, and then randomly selects one of its neighbours to visit and select. This technique starts from a randomly selected seed node. At every step, a neighbour of node i is either uniformly selected or selected in a weighted manner. The process is repeated for t steps until the daughter network g has the desired number of nodes (n) or edges M_g.

Consider a daughter network g constructed from a random walker starting at node i and taking t random steps. Let $p_i(t)$ denote the probability that a walker is at node i at time t. If we

Table 6.2: Statistics for k^β sparsification. The daughter networks have been sampled to match the size (G_c) of *Ossian*. Although statistics match for $\beta = 0$, non of the sparsification techniques so far provide a degree distribution similar to the mother network

Narrative	N	M	$\langle k \rangle$	ℓ	C	r_k
Ossian	254	596	4.69	3.62	0.45	-0.10
Acallam	515	1364	5.30	3.71	0.41	-0.13
$\beta = 0$	253(6)	530(18)	4.17(7)	3.90(2)	0.40(1)	-0.14(1)
$\beta = 1$	255(1)	938(3)	7.37(3)	2.98(2)	0.63(2)	-0.15(1)
$\beta = \infty$	254(1)	995(1)	7.87(2)	2.93(2)	0.70(2)	-0.15(1)

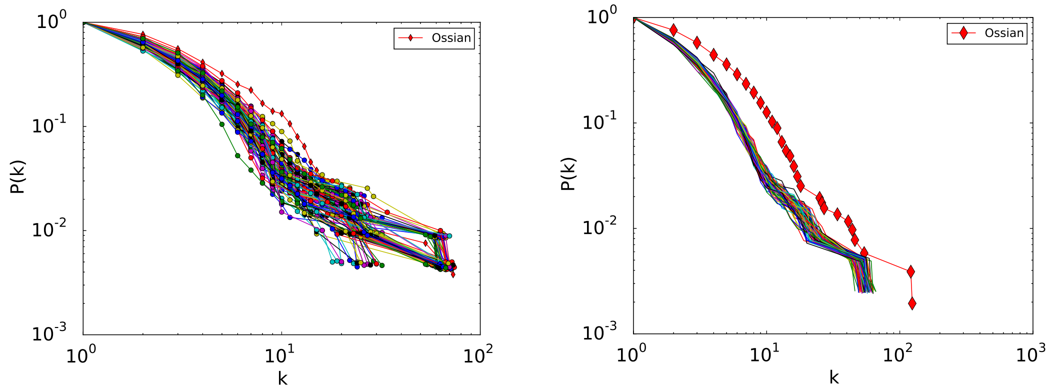

Figure 6.2: Daughter networks generated by (a) Random node sparsification and (b) unweighted edge sparsification. The degree distribution of *Ossian* is marked by a red diamond. The shaded circles (node sparsification) and lines (edge sparsification) represent the generated daughter networks. Neither type is Ossianic.

consider that a walker is at node j at time $t-1$ then the probability of traversing to any one of k_j neighbours of j is $1/k_j$. Hence, for an undirected network, the probability that a walker is at node i is given by

$$p_i(t) = \sum_j \frac{A_{i,j}}{k_j} p_j(t-1). \tag{6.2.18}$$

In the limit $t \to \infty$, this becomes $t = \infty$ [119] so that

$$p_i^*(\infty) = \sum_j \frac{A_{i,j}}{k_j} p_j^*(\infty). \tag{6.2.19}$$

The probability that a random walker will be found at node i in the limit of long time is proportional to k_i, the degree of node i, that is

$$p_i^* = \frac{k_i}{\sum_j k_j} = \frac{k_i}{2M}. \tag{6.2.20}$$

Besides a low clustering coefficient of $C = 0.22$ compared to $C = 0.45$ observed in *Ossian*, the remaining statistics are comparable. Additionally, 15% of samples obtained in this way pass the KS-test. Table 6.3 provides averaged statistics from 100 realisations.

6.2.3 The snowball approach

We also consider a more deterministic approach, called *snowball sampling* [180]. This is sometimes called chain-referral and is a non-probabilistic approach where we begin from a seed

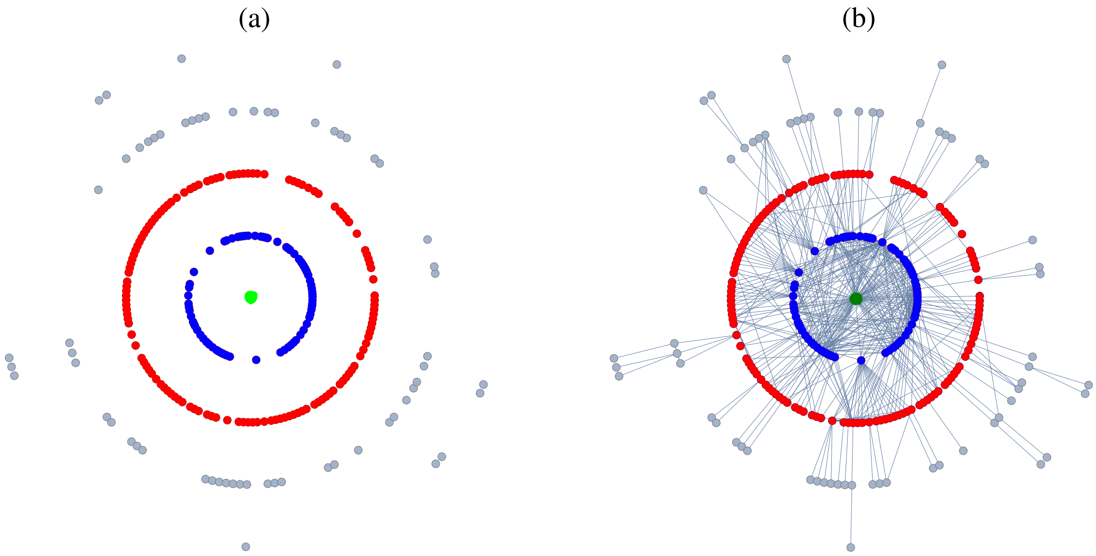

Figure 6.3: Snowball sampling of radius two: Panel (a) shows the seed node in the middle (green) and its neighbours (first layer in blue). The second layer (path length of radius two) is represented by nodes in red and the greyed out nodes are from additional layers excluded in this sample. Panel (b) includes edges between the sampled nodes (blue and red) and all nodes in grey are discarded.

node and recruit additional subjects from among their neighbours. This method is sometimes used to find a hidden sub-population [119, 183]. The initial step involves choosing a single node and all other nodes directly linked to it [179]. The second step includes neighbours of subjects recruited from the previous step. The process is repeated until the desired sample size n is reached. The steps otherwise called layers is akin to the idea of a radius in ego centric networks. To control the number of required nodes in the sample graph g, nodes can be randomly selected from the last n^{th} layer. Only interactions between the sampled nodes form the edges of the sample network g. Ebbes *et. al* in [171] considers an analogous method where a seed node, also called an ego, is randomly selected. All neighbours of the seed node are included in the sample including edges between the sampled nodes. Figure 6.3 provides an illustration of snowball sampling.

Here we acknowledge Ossian as the most important character. We then shine a spotlight around him by making him the seed node to construct an ego-centred daughter network. We include in our sample, Ossian's nearest and second-nearest neighbours. This gives 309 characters in total. The final step involves randomly removing extra nodes and edges to make sure that our resultant daughter network is of the same size N and M as that of *Ossian*. Before doing this, we note that all nearest neighbours pass through Ossín. This imposes a constraint on the network and its diameter, which is restricted to a maximum of four. To remove this constraint we allow any node (including that representing Ossian himself) to be a candidate for random removal. The randomisation process is repeated 100 times before taking averages. The statistics and degree distributions obtained using this method delivers an Ossianic type network which also resembles the mother network *Acallam*. The results of the KS-test gives p-values > 0.05, 76% of the time in support of similar distributions. A match here means the daughter network resembles both the degree distribution of the target network *Ossian* and the mother network *Acallam*. Errors from the random walk method were too small to report.

6.2.4 Other Sparsification Methods

We also consider other sparsification methods found in literature but none of them deliver a sample that matches both *Ossian* and *Acallam*. These methods are briefly discussed below.

The approaches that we have looked at so far are based upon node sparsification. We can also consider analogous edge sparsification techniques. Similarly, a sample network g can be conceived by randomly selecting the desired number of edges from the mother network G, a process called *random edge sampling*. Therefore, the daughter network will contain a total of n nodes attached at the end of each randomly selected edge. However the sampled network formed in this way is likely to be sparsely connected with a large network diameter as a result. It is suggested that edge sparsification fails to preserve some network properties of the original graph including the average clustering coefficient which is underestimated as a result of disconnected components that originates from this technique [184] . The degree distributions generated from 100 realisations of the random edge sparsification are shown in panel (b) of Figure 6.2.

The 515 characters in the giant component of *Acallam*'s positive network are connected by almost 1364 edges. *Ossian*, on the other hand, has 596 edges between 254 characters. Selecting 596 edges at random from *Acallam* delivers approximately 400 vertices for the daughter

Table 6.3: Statistics for random walk and snowball sampling. Both methods deliver a match in statistics. However, we obtain better results from a more deterministic process (snowball) especially on the degree distributions. About 76% of snowball samples match both *Ossian* and *Acallam* compared to a random walk technique with a success rate of 15% .

Narrative	N	M	$\langle k \rangle$	ℓ	C	r_k	Match
Ossian	254	596	4.69	3.62	0.45	-0.10	
Acallam	515	1364	5.30	3.71	0.41	-0.13	
Random walk	254	596	4.69	3.35	0.22	-0.13	15 %
Snowball	254(2)	598(1)	4.70(3)	3.20(3)	0.37(1)	-0.22(1)	76 %

network, on average. We generated 100 realisations of such daughter networks. The KS test delivers a p-value of less than 0.0001 suggesting the sparsified network is not ossianic. Panel (b) of Figure 6.2 illustrates this by plotting the 100 realisations alongside *Ossian*'s degree distribution.

As with the weighted case in random node selection, a weight can also be attached to an edge based on the frequency of interaction between two nodes. One way to do this is to partition *Acallam* into 11 sections according to the chapters in [103] and to assign a numerical weight to each link based on how many chapters two characters interact in. If the same edge appears in two chapters, for example, it is twice as likely to be included in the sample than an edge appearing in only one chapter of the narrative. However, results of the KS test give p-values of less than 0.0001 so that this is not a good ossianification process

Another method, the *forest-fire* approach, was initially proposed by Leskovec *et. al* in [175] and applied in [172] and [185]. The concept is comparable to the behaviour of fire spreading in a woodland. Contrary to a random walk that begins from a single seed node, the forest-fire may start from a multiple nodes. It is likewise possible for multiple walkers to transverse multiple edges at the same time step. The idea involves a forward burning probability p and the process is implemented as follows:

1. The process starts with randomly selecting a seed node i. The process can start from either one seed node or a multiple of nodes.

2. Node i will arbitrarily pick x of its unvisited neighbours and burn them. Where $k_i < x$, then all k_i neighbours of node i will be burnt. Nodes and edges burnt by the forest fire are returned as the sampled graph g.

3. The process is repeated until g achieves the required number of n nodes. Where the forest-fire dies before the required sample size then additional runs of forest-fire is possible.

Samples obtained through this technique preserve the original network's degree distribution and diameter. A match in these statistics is observed for up to 15% reduction of the original network [172].

6.3 Remarks

In this section, we reported upon various attempts to "Ossianify" the *Acallam* networks. By "ossianification" we mean to decimate a parent network in such a way that the daughter both resembles the parent and resembles the target network — *Ossian* in this instance. This goes a step beyond what has been studied previously in the literature, namely sparsification. It has been pointed out in literature that there is no single sparcification method that performs better than others across all network statistics. So too there may not be a unique best method when it comes to ossianification. But the idea is not so much to closely match the *Ossian* — rather it is to find out if it is relatively easy to do so.

Here sampling success is evaluated by comparing between mother and daughter networks. Network characteristics compared include, average path-length, clustering coefficient, assortativity and degree distribution. The Kolmogorov-Smirnoff test is used to evaluate similarities in degree distributions.

Although not all methods work, we have demonstrated that it is possible to Ossianify *Acallam na Senórach* stochastically. Moreover, the methods that seem to work best are those that place some emphasis on node importance in particular sparsification by Oisín's ego network appears to work best. They also work best when the network integrity is maintained - i.e., when one focusses on characters who are part of the main action. We suggest that if Macpherson wished to employ network science to Ossianify *Acallam*, he would have had a number of options at his disposal. He only needed to remember pockets of Acallam na Senórach and some important characters around Oisín. It would be interesting, in the future, to couple this approach with methods from psychology although, but that is beyond the remit of the current investigation.

Chapter 7

Conclusions

We have performed extensive statistical analyses of character networks from two epic narratives belonging to the Irish and Scottish Gaelic traditions. The first of these is concerned with the Viking Age in Ireland up to the battle of Clontarf in 1014 AD. We contextualised the study through a discussion of the basis from which the events were chronicled, and explained a long-standing debate between the views of traditionalists and revisionists. Whilst the traditional memory of the events has been described as a "clear-cut" war between the Irish and Vikings, the revisionist view is of a domestic conflict in which the bulk of hostilities are internal to the Irish players.

In this thesis, we have gone beyond traditional humanities approaches and basic descriptive statistics and studied the character networks of *Cogadh Gaedhel re Gallaibh*. We consider our approach to be a prudent use of the chronicle since it is completely independent of the tone of the account. Indeed, the bombastic style and partisan nature of the *Cogadh* has long been foundations for its criticism (although it has not prevented it being used to support both sides of the debate). Our primary aim was to place the charater networks on a spectrum between the traditionalist and revisionist views, and to do this we applied the notion of categorical assortativity. This evaluates the extent to which interactions mainly connect nodes from either category: Irish or Viking. A strong negative categorical assortativity for the negative network would suggest a "clear-cut" war between the Irish and Vikings and a positive value would support the revisionist's interpretation.

Instead our results show a moderate negative value in categorical assortativity. We interpret this as supportive of the traditionalist view. However the magnitude of the value is not large and therefore does not back a "clear-cut" conflict between the Irish and Vikings. On the other hand, results from the positive network show a positive value in categorical assortativity, an indication that social interactions are mostly between characters of the same attribute but not wholly so.

The observed results are also robust, e.g. our conclusions are not influenced by disregarding Viking-Viking interactions.

Additionally, we compared the network of *Cogadh Gaedhel re Gallaibh* to those of other epic narratives [15, 16]. We found that the network properties of the *Cogadh* are closer to those from *Iliad* than *Njal's Saga* and *Táin Bó Cuailnge*. The closeness to the *Iliad* is fascinating considering some claims from humanities scholars that *Cogadh* was partly a rhetoric "pseudo-history" borrowed from the account of Troy [36,40,47,60]. It would be fascinating to perform more extensive comparative investigations in future studies.

The second part of our analysis focused on the character networks of *The Poems of Ossian*. The poems, purported to have been translations from Scottish-Gaelic sources into English, drew worldwide attention at the time of publication, and their scholarly study has enjoyed a revival in recent decades. James Macpherson's *Ossian* had great influence on the introduction of romanticism in literature. The works achieved international success and were paralleled to Homer's *Iliad* and *Odyssey*. However, from the outset, a great amount of controversy concerned the authenticity of the works and provoked a divided view in humanities. Some scholars branded them "the most successful literary fraud in history" implying that they were essentially a distortion of tales from Irish mythology [84]. Here we sought to put *Ossian* on a spectrum between the Greek and the Irish to try and make a quantitative contribution to this debate from a new perspective using character networks.

Our results show that the character network of *Ossian* resembles those from Irish mythology and a derivative text by Lady Gregory but is different to the *Iliad* or *Odyssey*. Our outcomes may be considered as completing investigations by the likes of Sylvester O'Halloran 250 years ago [63,72]. O'Halloran's identification of some characters in *Ossian* as appropriated from Irish mythology was essentially a study of the nodes of the networks. Ours is a study of its links.

These two studies have formed the basis for two papers. One of these has been published as Ref. [9] and the second has been submitted as detailed in Ref. [17]. A third publication [10] emerged during the course of these investigations but is not reported upon in this thesis because it refers to a different genre of literature.

One additional question that is addressed in this thesis (and which will form the basis of a fourth publication) is that of "ossianification". The thesis explores the possibilities of creating an *Ossian*-type network from that of *Acallam*. To demonstrate this is possible, we began by growing the network of *Ossian* to the size of *Acallam* to see if it would match the Irish network. Having verified this, we suggested that some form of an inverse process may be employed to create an *Ossian* type network from *Acallam*. Our results show that there are various successful sparsification methods capable of delivering an Ossianic type network that resembles both

Acallam and the target network *Ossian*. The methods that seem to work best are those that place an emphasis on node importance, especially for characters involved in the main action of the tales and in particular nodes that appear in Oisín's ego network. The results demonstrate that if Macpherson could have had network science to hand, and if he desired to use some of the methods employed here, he would have had a number of options available to him. He only needed to recall the main characters and interactions of the *Acallam* network, particularly those around Oisín, and build from there. However, these correlations do not decisively prove the sources of James Macpherson's *Ossian*. (Certainly, we cannot suppose that Macpherson had network theory in mind when he went about his work.) But they do show that creation of an *Ossian*-type network from an *Acallam*-type network is a relatively easy task.

This work can be extended in a number of ways. Like Refs. [15,16], the present analysis is based on static networks. These freeze the narrative progress and capture the plot "all at one glance in a visual display of its character network" [186]. Static networks are particularly appropriate for the texts that were investigated here which paid little regard to chronology. Nonetheless, dynamical properties are also of interest and ought to be explored in the future [148]. It would be interesting to see if temporal networks can help to re-establish some of the chronology. Directed and weighted networks also offer obvious routes for wider study. Moreover, the use of machine learning techniques can be extended to character networks to include predictive analysis.

Furthermore, inspired by the Ua Ruairc example in *Cogadh Gaedhel re Gallaibh*, it would also be interesting to investigate if the structural imbalance in some network triads could be developed to give a way to spot other potential interpolations in this and other texts. In the case of *Ossian*, it would be interesting to advance the work through interdisciplinary collaborations with experts from psychology to explore connections with the theories of memory and of the mind. A plethora of other quantitative approaches and suggestions are contained in the compendium [187]. Substantially more work is needed to extend this approach to compare with the Irish annals, genealogies, and other contemporary sources. We feel this is desirable in the long-term, and we hope that this work will help generate enthusiasm for such a project.

Having evolved out of applied mathematics, statistics and complexity science, the character-network approach is sometimes criticised by humanities scholars for bringing little new, merely re-discovering or verifying knowledge already established from conventional studies in that discipline. The standard reply to this type of criticism is that such agreement is exactly what one should hope to obtain using an approach which is new, provided that approach is valid. One has to be patient and allow the sub-discipline to evolve and mature. The precise placement of *Cogadh Gaedhel re Gallaibh* along a spectrum from the international to the intranational using quantitative techniques (namely categorical assortativity) in this thesis is a new development in this regard. The demonstration of the possibility to “ossianify” Irish narratives by focusing on the main characters and their interlinks also exemplifies new developments in addition to showing how the sub-discipline can open up new questions. In these ways, our work represents a step forward in the evolution of the sub-field. It goes beyond previous works in that it generates new quantitative elements to unfinished debates in the humanities and opens up new avenues for interdisciplinary research.

Bibliography

[1] Sharan, R., Suthram, S., Kelley, R.M., Kuhn, T., McCuine, S., Uetz, P., Sittler, T., Karp, R.M. and Ideker, T. (2005). Conserved patterns of protein interaction in multiple species. *Proceedings of the National Academy of Sciences of the United States of America*, **102**,1974-1979.

[2] Von Ferber, C., Holovatch, T., Holovatch, Yu. and Palchykov, V. (2009). Public transport networks: empirical analysis and modelling. *Eur. Phys. J. B*, **68**,261-275.

[3] Von Ferber, C., Berche, B., Holovatch, T. and Holovatch, Yu. (2012). A Tale of Two Cities. Vulnerabilities of the London and Paris Transit Networks, *J. Transp. Secur.*, **5**, 199216.

[4] Newman, M.E. (2001a). Scientific collaboration networks. I. Network construction and fundamental results. *Physical review E*, **64**,016131.

[5] Newman, M.E.J. (2001b). Scientific collaboration networks. II. Shortest paths, weighted networks, and centrality. *Physical review E*, **64**,016132.

[6] Newman, M.E.J. (2001c). The structure of scientific collaboration networks. *Proceedings of the National Academy of Sciences*, **98**,404–409.

[7] Agarwal, A., Corvalan, A., Jensen, J. and Rambow, O. (2012). Social Network Analysis of Alice in Wonderland. In *Workshop on Computational Linguistics for Literature*, 88-96.

[8] Mac Carron, P. (2014). *A Network Theoretic Approach to Comparative Mythology*, PhD thesis, Coventry University, UK.

[9] Yose, J., Kenna, R., Mac Carron, P., Platini, T. and Tonra, J. (2016). A Networks-Science Investigation into the Epic Poems of Ossian. *Advs. Complex Syst*, **19**, 1650008.

[10] Gramsch, R., MacCarron, M., MacCarron, P. and Yose, J. (2017). Medieval Historical, Hagiographical and Biographical Networks. In *Maths Meets Myths: Quantitative Approaches to Ancient Narratives*, 45-69. Springer International Publishing.

[11] Euler, L. (1741). Solutio problematis ad geometriam situs pertinentis. *Commentarii academiae scientiarum Petropolitanae*, **8**, 128-140.

[12] Alexanderson, G. (2006). About the cover: Euler and Königsbergs Bridges: A historical view. *Bulletin of the American mathematical society*, **43**,567-573.

[13] Watts, D.J. and Strogatz, S.H. (1998). Collective dynamics of 'small-world' networks, *Nature*, **393**, 440-442.

[14] Shields R. (2012). Cultural topology: The seven bridges of Königsberg, 1736. *Theory, Culture & Society*. **29**,43-57.

[15] Mac Carron, P. and Kenna, R. (2012). Universal properties of mythological networks. *EPL*, **99**, 28002.

[16] Mac Carron, P., Kenna, R. (2013). Network analysis of the Íslendinga sögur the Sagas of Icelanders. *Eur. Phys. J. B*, **86**, 407.

[17] Yose, J., Kenna, R., MacCarron, M. and MacCarron, P. (2017). Network Analysis of the Viking Age in Ireland as portrayed in Cogadh Gaedhel re Gallaibh. *arXiv preprint*:1707.07526 (under review *Royal Society Open Science)*, 1707.07526.

[18] Mc Carthy, D. (2017). Analysing and restoring the chronology of the Irish Annals. In *Maths Meets Myths: Quantitative Approaches to Ancient Narratives*, 177-194. Springer International Publishing.

[19] Ryan, J. (1938). The Battle of Clontarf. *Journal of the Royal Society of Antiquaries of Ireland*, **8**, 1-50.

[20] Macpherson, J. (1762). *Fingal, an Ancient Epic Poem, in Six Books; Together with Several Other Poems, Composed by Ossian, the Son of Fingal. Translated from the Galic Language* (T. Becket and P.A. De Hondt, London).

[21] Macpherson, J. (1763). *Temora, an Ancient Epic Poem, in Eight Books; Together with Several Other Poems, Composed by Ossian, the Son of Fingal. Translated from the Galic Language* (T. Becket and P.A. De Hondt, London).

[22] Macpherson, J. (1765). *The Works of Ossian, the Son of Fingal, in Two Volumes. Translated from the Galic Language by James Macpherson. The third Edition. To which is subjoined a critical dissertation on the poems of Ossian. By Hugh Blair, D. D.* (T. Becket and P.A. De Hondt, London).

[23] Macpherson, J. (1773). *The Poems of Ossian. Translated by James Macpherson, Esq., in Two Volumes. A new edition, carefully corrected, and greatly improved* (W. Strahan and T. Becket, London).

[24] Birkett, T. and Lee, C. (2014). *The Vikings in Munster*, Centre for the Study of the Viking Age, University of Nottingham.

[25] Casey, D. and Meehan, B. (2014). Brian Boru and the Book of Armagh. *History Ireland*, **22**, 28–29.

[26] Clarke, H.B. and Johnson, R. (eds) (2015). *Before and After the Battle of Clontarf: The Vikings in Ireland and Beyond*, Dublin: Four Courts Press.

[27] Duffy, S. (ed.) (2014). *Medieval Dublin XVI, Proceedings of Clontarf 1014–2014: national conference marking the millennium of the Battle of Clontarf, 11–12 April 2014*, Dublin: Gill and Macmillan Ltd.

[28] Duffy, S. (2014). *Brian Boru and the Battle of Clontarf*. Dublin: Gill and Macmillan Ltd.

[29] Duffy, S. (2014). Brian Boru at the Battle of Clontarf: a medieval version of 1916? *Irish Times*, 25 Jan 2014; http://www.irishtimes.com/culture/heritage/brian-boru-at-the-battle-of-clontarf-a-medieval-version-of-1916-1.1667155. Accessed 03 July 2017.

[30] Etchingham, C. (2015). His finest hour. *Irish Literary Supplement*, **34**, made available through The Free Library (22 March 2015); https://www.thefreelibrary.com/Hisfinesthour.-a0407227269. Accessed 03 July 2017.

[31] Llywelyn, M. (2014). *Brian Boru and the Battle for Ireland*. Dublin: The O'Brien Press.

[32] McGettigan, D. (2013). *The Battle of Clontarf, Good Friday, 1014*. Dublin: Four Courts Press.

[33] Royal Irish Academy. (2014). *Humanities: 1014 Battle of Clontarf — Royal Irish Academy Lunchtime Lecture Series Spring 2014*; https://soundcloud.com/the-royal-irish-academy/sets/royal-irish-academy-lunchtime-lecture-series-spring-2014-1014. Accessed on 03 July 2017.

[34] 1014: Brian Boru and the Battle of Clontarf - History Hub; http://historyhub.ie/1014-brian-boru-battle-of-clontarf Accessed 03 July 2017.

[35] Todd, J.H. (ed) (1867). *Cogadh Gaedhel re Gallaibh. The war of the Gaedhil with the Gaill, or, the invasions of Ireland by the Danes and other Norsemen*. London: Longmans, Green, Reader and Dyer.

[36] Bugge, A. (1905). *Caithréim Chellacháin Chaisil: the victorious career of Cellachan of Cashel or the wars between the Irishmen and the Norsemen in the middle of the 10th century*. Christiana (Oslo): Gundersen.

[37] Bugge, S. (1908). *Norsk sagafortælling og sagaskrivning i Irland*. Kristiania (Oslo): Grøndahl.

[38] Meyer, K. (1910). Brian Borumha. *Ériu*, **4**, 72–73.

[39] Thurneysen, R. (1921). *Die irische Helden- und Königsage bis zum siebzehnten Jahrhundert*. Halle: Max Niemeyer.

[40] Goedheer, A.J. (1938). *Irish and Norse Traditions about the Battle of Clontarf*. Haarlem: H.D. Tjeenk, Willink. and Zoon.

[41] Ryan, J. (1940). Review of "Irish and Norse traditions about the Battle of Clontarf" by Albertus Johannes Goedheer, *Irish Historical Studies*, **2**, 93–97.

[42] Flower, R. (1947). *The Irish Tradition*. Oxford: Oxford University Press.

[43] Ó Corráin, D. (1972). *Ireland Before the Normans*. Dublin: Gill and MacMillan.

[44] Ghiollamhaith, A.N. (1981). Dynastic warfare and historical writing in North Munster, 12761350. *Cambridge Medieval Celtic Studies*, **2**, 73–89.

[45] Ní Mhaonaigh, M. (1992). Bréifne bias in Cogad Gáedel Re Gallaib. *Ériu*, **43**, 135–158.

[46] Holm, P. (1994). Between apathy and antipathy: the Vikings in Irish and Scandinavian history. *Peritia*, **8**, 151–169.

[47] Ní Mhaonaigh, M. (1995). Cogad Gáedel re Gallaib: Some Dating Considerations. *Peritia*, **9**, 354–377.

[48] Ní Mhaonaigh, M. (1996). Cogad Gáedel re Gallaib and the Annals: a Comparison. *Ériu*, **47**, 101–126.

[49] Ní Mhaonaigh, M. (1997). Some Middle Irish declensional patterns in Cogad Gáedel re Gallaib. *Z. Celt. Philol.* **49**, 615–628.

[50] Ó Corráin, D. (1997). Ireland, Wales, and the Hebrides. In *The Oxford illustrated history of the Vikings*. 105–106. Oxford: Oxford University Press.

[51] Ó Corráin, D. (1998). Viking Ireland afterthoughts. In *Ireland and Scandinavia in the early Viking Age*. 421–452. Dublin: Four Courts Press.

[52] Ní Mhaonaigh M. (1998). Friend and foe: Vikings in ninth- and tenth-century Irish literature. In *Ireland and Scandinavia in the early Viking age* 381–402. Dublin: Four Courts Press.

[53] Ní Mhaonaigh, M. (2002). Tales of Three Gormlaiths in Medieval Irish Literature. *Ériu*, **52**, 1–24.

[54] Downham, C. (2005). The Battle of Clontarf in Irish history and legend. *History Ireland*, **13**, 19–53.

[55] Casey, D. (2010). Historical and literary representations of Brian Boru's burial in Armagh, 1014AD. *North Munster Antiquarian Journal*, **50**, 29–44.

[56] Dubois, T.A. (2011). Juxtaposing Cogadh Gáedel re Gallaib with Orkneyinga saga. *Oral Tradition*, **26**, 267–296.

[57] Ní Úrdail, M. (2011). *Cath Cluana Tarbh 'The battle of Clontarf'*. London: The Irish Texts Society/Cumann na Scríbheann nGaedhilge

[58] Ní Mhaonaigh, M. (2012). A neglected account of the battle of Clontarf. *Z. Celt. Philol.*, **59**, 143–167.

[59] Casey, D. (2013). A reconsideration of the authorship and transmission of Cogadh Gáedhel re Gallaibh. *P. Roy. Irish Acad. C*, **113**, 139–161.

[60] Ní Mhaonaigh, M. (2014). 'The metaphorical Hector': the literary portrayal of Murchad mac Bríain. In *Classical literature and learning in medieval Irish narrative*, Cambridge, UK: Boydell and Brewer.

[61] Kirwan, M. (2014). The Viking legacy in Ireland. In *The Vikings in Munster*. 4–6. Nottingham: Centre for the Study of the Viking Age, University of Nottingham.

[62] O'Connor, C. (1766). *Dissertations on the History of Ireland. To which is subjoined, a dissertation on the Irish colonies established in Britain. With some remarks on Mr. Mac Pherson's translation of Fingal and Temora*. (Second edition). Dublin: G. Faulkner.

[63] O'Halloran, S. (1778). *A General History of Ireland, from the Earliest Accounts to the close of the Twelfth Century, Collected from the Most Authentic Records* Vol.III. London: A. Hamilton.

[64] *Nation*, https://en.wikipedia.org/wiki/Nation. Accessed 03 July 2017.

[65] Harrison, S. (2014). The Battle of Clontarf: the archaeological evidence? Talk as part of the Royal Irish Academy Lunchtime Lecture Series, Spring 2014, available at https://soundcloud.com/the-royal-irish-academy/stephen-harrison. Accessed 03 July 2017.

[66] Mc Carthy, D.P. (2008). *The Irish annals — Their genesis, evolution and history*. Dublin: Four Courts Press.

[67] Bayerschmidt, C.F. and Hollander, L.M. (1998). *Njal's Saga*. Hertfordshire: Wordsworth Classics.

[68] Sveinsson, E.Ó. (ed). (1954). *Íslenzk fornrit 12: Brennu-Njáls saga* . Reykjavík: Hith íslenzka fornritafélag.

[69] Gaskill, H. (ed.) (1996). *The Poems of Ossian and Related Works* with an Introduction by Fiona Stafford. Edinburgh: Edinburgh University Press.

[70] Mulholland, J. (2009). James Macpherson's Ossian poems, oral traditions, and the invention of voice, *Oral Tradition*, **24**, 393-414.

[71] Johnson, S. (1775). *A Journey to the Western Islands of Scotland*. London: Strahan and Cadell.

[72] Ó Halloran, S. (1763). The poems of Ossine, the son of Fionne Mac Comhal, re-claimed, *Dublin Mag*. 2123; reprinted in *Ossian and Ossianism*, Vol. III, 87-89.

[73] Lyons, C. E. (2007). Sylvester Ó Halloran: Miso-Dolos, *Journal of the Galway Archaelogical and Historical Society*, **59**, 46-58.

[74] Warner, F. *Remarks on the History of Fingal*, 1762. https://ia700200.us.archive.org/14/items/remarksonhistory00warn/remarksonhistory00warn.pdf

[75] Webb, D. (1763). *Fingal Reclaimed*. (pamphlet).

[76] O'Brien, J. (1764). Mémoire au suject des poëmes de M. Macpherson, in *Journal des Scavans*.

[77] Mackenzie, H. (ed.) (1805). *Report of the Committee of the Highland Society of Scotland appointed to inquire into the Nature and Authenticity of the Poems of Ossian*. Edinburgh: Edinburgh University Press.

[78] Moore, D. (2000). Heroic incoherence in James Macpherson's The Poems of Ossian. *Eighteenth-Century Studies*, **34**, 43-59.

[79] Moore, D. (2004). *Ossian and Ossianism*. London: Taylor and Francis.

[80] Stafford, F. J. (1991). *Dangerous Success: Ossian, Wordsworth, and English Romantic Literature* in Ref. [87], 4972.

[81] Bär, G. and Gaskill, H. (2012). *Ossian and National Epic*. Frankfurt am Main: Peter Lang.

[82] Womack, P. (1989). *Improvement and Romance: Constructing the Myth of the Highlands*. London: Macmillan.

[83] Burnett, A. and Anderson, L. (2011). *Blind Ossian's Fingal: Fragments and controversy*. Edinburgh: Luath Press Ltd. A reprint of the first edition and abridgement of the follow-up with new material.

[84] Curley, T.M. (2009). *Samuel Johnson, the Ossian Fraud, and the Celtic Revival in Great Britain and Ireland*. Cambridge,UK: Cambridge University Press.

[85] Meek, D.E. (1991). *The Gaelic ballads of Scotland: creativity and adaptation*, in Ref. [87] 19-48.

[86] Fielding, P. (1996). *Writing and Orality: Nationality, culture, and nineteenth-century Scottish fiction*. Oxford: Oxford University Press.

[87] Gaskill, H. (ed.) (1991). *Ossian Revisited*. Edinburgh: University of Edinburgh Press.

[88] Stafford, F. (1996). Introduction: The Ossianic poems of James Macpherson, in *The Poems of Ossian and Related Works*. Edinburgh:Edinburgh University Press

[89] Gaskill, H. (ed.), (2004). *The Reception of Ossian in Europe*. London:Thoemmes Continuum.

[90] McNeil, K. (2007). *Scotland, Britain, Empire: Writing the Highlands, 1760-1860*. Columbus: State University Press.

[91] Kidd, C. (1993). *Subverting Scotlands Past: Scottish Whig Historians and the Creation of an Anglo-British Identity 16891830*. Cambridge: Cambridge University Press.

[92] Gaskill, H. (20072008). The Homer of the North, *Interfaces*, **27**, 1324.

[93] Macpherson, J. (1760). *Fragments of Ancient Poetry, Collected in the Highlands of Scotland, and Translated from the Galic or Erse Language*. Edinburgh: G. Hamilton and J. Balfour.

[94] Quint, D. (1993). *Epic and Empire: Politics and Generic Form from Virgil to Milton*. New Jersey: Princeton University Press.

[95] Golden, J.L. and Golden, A.L. (2002). *Thomas Jefferson and the Rhetoric of Virtue*. Rowman and Littlefield (Maryland).

[96] Johnson, A. L. (2016). Thomas Jeffersons Ossianic romance, *Stud. Eighteenth-Cent. Cult.*, **45** , 1935.

[97] Evans, E. (1764). *Some Specimens of the Poetry of the Ancient Welsh Bards. Translated into English, with Explanatory Notes on the Historical Passages, and a Short Account of Men and Places mentioned by the Bards, in order to give the Curious some Idea of the Tastes and Sentiments of our Ancestors, and their Manner of Writing*. London: R. and J. Dodsley.

[98] Skene, W. F. (1868). *The Four Ancient Books of Wales Containing the Cymric Poems Attributed to the Bards of the Sixth Century*. Edinburgh: Edmonston and Douglas.

[99] Percey, T. (1775). *Reliques of Ancient English Poetry: Consisting of Old Heroic Ballads, Songs, and Other Pieces of our Earlier Poets (Chiefly of the lyric kind.) Together with some few of later date* (London: J. Dodsley).

[100] Brooke, C. (1816). *Reliques of Irish Poetry: Consisting of Heroic Poems, Odes, Elegies, and Songs, Translated Into English Verse, With Notes Explanatory and Historical, and the Originals in the Irish Character, to which is Subjoined an Irish Tale* Dublin:J. Christie.

[101] Mitchell, S. (2012). Macpherson, Ossian, and Homers Iliad, in *Ossian and National Epic* (eds.) Frankfurt am Main: Peter Lang.

[102] O'Halloran, C. (1989). Irish Recreations of the Gaelic Past: the Challenge of Macpherson's Ossian, *Past and Present*, **124**, 69-95.

[103] Dooley, A. and Roe, H. (1999). *Tales of the Elders of Ireland: A New Translation of Acallam na Senórach*. Oxford: Oxford University Press.

[104] Homer. (2003). *The Iliad*. London: Penguin Classics. Originally translated by Rieu E.V., revised and updated by Jones, P. with Rieu, D.C.H.

[105] Homer. (1980). *The Odyssey*. Oxford: Oxford University Press. Translated by Shewring, W., with an Introduction by Kirk, G.S.

[106] Gregory, A. (2004). *Gods and Fighting Men: The Story of the Tuatha De Danaan and of the Fianna of Ireland, Arranged and put into English by Lady Augusta Gregory*. (J. Murray, London, 1904); reproduced in *Lady Gregorys Complete Irish Mythology* CA: Bounty Books, Vacaville.

[107] Newman, M. E. J. (2003). The structure and function of complex networks. *SIAM Rev.*, **45**, 167 256.

[108] Amaral, L.A.N., Scala, A., Barthélémy, N. and Stanley, H.E. (2000). Classes of small-world networks. *Proc. Nat. Acad. Sci.*, **97**, 11149–11152.

[109] Albert, R. and Barabási, A.L. (2002). Statistical mechanics of complex networks. *Reviews of Modern Physics*, **74**, 47-97.

[110] González, M.C., Hildago, C.A. and Barabási, A.L. (2008). Understanding individual human mobility patterns. *Nature*, **453**, 779-782.

[111] Picoli, Jr.S., Mendes, R.S. and Malacarne, L.C. (2003). q-exponential, Weibull, and q-Weibull distributions: an empirical analysis. *Physica A: Statistical Mechanics and its Applications*. **324**,678-688

[112] Akaike, H. (1974). A new look at the statistical model identification. *IEEE Trans. Autom. Control*, **19**, 716-723.

[113] Clauset, A., Shalizi, C.R. and Newman, M. E. J. (2009). Power-law distributions in empirical data. *SIAM Rev.*, **51**, 661703.

[114] Bianconi, G. (2007). The entropy of randomized network ensembles. *EPL*, **81**, 28005.

[115] Newman, M.E. (2002). Assortative mixing in networks. *Physical Review Letters*, **89**, 208701.

[116] Johnson, S., Torres, J.J., Marro, J. and Muñoz, M.A. (2010). The entropic origin of disassortativity in complex networks. *Phys. Rev. Lett.* , **104**, 108702.

[117] Clauset, A. (2013). *Network Analysis and Modeling.* Lecture 5. Retrieved from `http://tuvalu.santafe.edu/\~aaronc/courses/5352/fall2013/csci5352\_2013\_L5.pdf`.

[118] Newman, M. E. J. (2003). Mixing patterns in networks. *Phys. Rev. E*, **67**, 026126.

[119] Newman, M. E. J. (2009). *Networks: an introduction.* Oxford: Oxford University Press.

[120] Estrada, E. and Knight, P. (2015). *A First Course in Network Theory*. Oxford: Oxford University Press

[121] Yan, E. and Ding, Y. (2009). Applying centrality measures to impact analysis: A coauthorship network analysis. *J. Amer. Soc. Inf. Sci. Technol.*, **60**, 2107-2118.

[122] Ni, C., Sugimoto, C. and Jiang, J. (2011). Degree, Closeness, and Betweenness: Application of group centrality measurements to explore macro-disciplinary evolution diachronically. In *Proceedings of ISSI*, 1-13.

[123] Opsahl, T., Agneessens, F. and Skvoretz, J. (2010). Node centrality in weighted networks: Generalizing degree and shortest paths. *Social Networks*, **32**, 245-251.

[124] Freeman, L. C. (1978). Centrality in social networks: Conceptual clarification. *Social Networks*, **1**, 215-239.

[125] Freeman, L.C. (1977). A Set of Measures of Centrality based on Betweenness. *Sociometry*, **40**, 35-41

[126] Freeman, L. C. (2004). *The development of social Network Analysis*. Vancouver: Empirical Press.

[127] Anthonisse, J.M. (1971). The rush in a directed graph. *Stichting Mathematisch Centrum. Mathematische Besliskunde*, **9**, 1-10.

[128] Girvan, M. and Newman, M. E. J. (2002). Community structure in social and biological networks. *Proceedings of the National Academy of Sciences*, **99**, 7821-7826.

[129] Newman, M.E. and Girvan, M. (2004). Finding and evaluating community structure in networks. *Physical review E*, **69**, 026113.

[130] Danon, L., Díiaz-Guilera, A. and Arenas, A. (2006). The effect of size heterogeneity on community identification in complex networks. *Journal of Statistical Mechanics: Theory and Experiment*, **11**, 11010.

[131] Kernighan, B. and Lin, S. (1970). An ancient heuristic procedure to partition graphs. *Bell Syst. Tech. J.*, **49**, 291-307.

[132] Zachary, W. W. (1977). An information flow model for conflict and fission in small groups. *Journal of Anthropological Research*, **33**, 452-473

[133] Newman, M.E.J. (2006). Modularity and community structure in networks. *Proceedings of the National Academy of Sciences of the United States of America*, **103**, 8577-8582.

[134] Boccaletti, S., Latora, V., Moreno, Y., Chavez, M. and Hwang, D.-U. (2006). Complex networks: Structure and dynamics. *Physics Reports*, **424** 175-308.

[135] Heider, F. (1946). Attitudes and cognitive organization. *J. Psychol*, **21**, 107-112.

[136] Estrada, E. (2011). *The Structure of Complex Networks: Theory and Applications*. Oxford: Oxford University Press.

[137] Gilbert, E. N. (1959). Random graphs. *Annals of Mathematical Statistics*, **30**, 1141 1144.

[138] Erdős, P. and Rényi, A. (1959). On random graphs I. *Publ. Math. Debrecen*, **6**, 290297.

[139] Fronczak, A., Fronczak, P. and Holyst, J. A. (2004). Average path length in random networks. *Physical Review E*, **70**, 056110.

[140] Milgram, S. (1967). The small world problem. *Psychol. Today* , **2**, 60–67.

[141] *Frigyes Karinthy*, `https://en.wikipedia.org/wiki/Frigyes_Karinthy`. Accessed 16-02-17

[142] Karinthy, F. (1929). *Chain-Links*. Translated from Hungarian and annotated by Adam Makkai and Enikö Jankó. Everything is Different.

[143] Humphries, M.D. and Gurney, K. (2008). Network small-world-ness: a quantitative method for determining canonical network equivalence. *PLoS One*, **3**, e0002051.

[144] Newman, M.E.J. and Park, J. (2003). Why social networks are different from other types of networks. *Physical review E*, **68**, 036122.

[145] Haldane, A.G. and May, R.M. (2011). Systemic risk in banking ecosystems. *Nature*, **469**, 351.

[146] Motter, A.E., Gulbahce, N., Almaas, E. and Barabási, A.L. (2008). Predicting synthetic rescues in metabolic networks. *Molecular Systems Biology*, **4**, 168.

[147] Albert, R., Jeong, H. and Barabási, A.L. (2000). Error and attack tolerance of complex networks. *Nature*, **406**, 378382.

[148] Prado, S.D., Dahmen, S.R., Bazzan, A.L., Mac Carron, P. and Kenna, R. (2016). Temporal network analysis of literary texts. *Advs. Complex Syst.*, **19**, 1650005.

[149] Trovati, M. and Brady, J. (2014). Towards an automated approach to extract and compare fictional networks: An initial evaluation. *Proceedings - International Workshop on Database and Expert Systems Applications*, DEXA, 246-250.

[150] McPherson, J. M., Smith-Lovin, L. and Cook, J. M. (2001). Birds of a feather: Homophily in social networks. *Annual Review of Sociology*, **27**, 415444.

[151] Humphries, M.D. and Gurney, K. (2008). Network small-world-ness: A quantitative method for determining canonical network equivalence. *PLoS One*. **3**, e0002051.

[152] Clauset, A., Shalizi. C.R., Newman, M.E.J. (2009). Power-law distributions in empirical data. *SIAM Rev.*, **51**, 661–703.

[153] Jeong, H., Mason, S.P., Barabási, A.L. and Oltvai, Z.N. (2001). Lethality and centrality in protein networks. *Nature*, **411**, 41–42.

[154] Freeman, L.C. (1977). A set of measures of centrality based on betweenness. *Sociometry*, **40**, 35–41.

[155] Newman, M.E.J. (2005). Power laws, Pareto distributions and Zipfs law. *Contemporary physics*, **46**, 323351.

[156] Conover, W.J., (1972). A Kolmogorov goodness-of-fit test for discontinuous distributions. *Journal of the American Statistical Association*, **67**, 591-596.

[157] Conover, W.J. and Conover, W.J. (1980). *Practical nonparametric statistics*. New York: Wiley.

[158] Daniel, W.W. (1990). *Applied nonparametric statistics*. Boston: PWS-Kent.

[159] Ipsen, M. and Mikhailov, A.S. (2003). Evolutionary reconstruction of networks. *Physical Review E*, **66**, 046109.

[160] Jurman, G., Riccadonna, S., Visintainer, R. and Furlanello, C. (2011). Biological network comparison via Ipsen-Mikhailov distance. *arXiv*, 1109.0220.

[161] Thomson, D.S. (1952). *The Gaelic Sources of Macpherson's Ossian*. Aberdeen: Oliver and Boyd.

[162] Albert, R. and Barabási, A.L. (2002). Statistical mechanics of complex networks. *Rev. Mod. Phys.*, **74**, 47–94.

[163] Jackson, M.O. (2010). *Social and economic networks*. Princeton: Princeton university press.

[164] Barabási, A.L. and Albert, R. (1999). Emergence of scaling in random networks. *Science*, **286**, 509-512.

[165] Barabási, A.L., Albert, R. and Jeong, H. (1999). Mean-field theory for scale-free random networks. *Physica A: Statistical Mechanics and its Applications*, **272**, 173-187.

[166] Price, D.J.D.S. (1965). Networks of scientific papers. *Science*, **149**, 510-515.

[167] Price, D.D.S. (1976). A general theory of bibliometric and other cumulative advantage processes. *Journal of the Association for Information Science and Technology*, **27**, 292-306.

[168] Matthew 5:3-12, Holy Bible: King James Version.

[169] Callaway, D. S., Hopcroft, J. E., Kleinberg, J. M., Newman, M. E. and Strogatz, S. H. (2001). Are randomly grown graphs really random? *Physical review E*, **64**, 041902.

[170] Ahmed, N.K., Neville, J. and Kompella, R. (2014). Network sampling: From static to streaming graphs. *ACM Transactions on Knowledge Discovery from Data (TKDD)*, **8**, 7.

[171] P. Ebbes et al. Ebbes, P., Huang, Z., Rangaswamy, A., Thadakamalla, H.P. and Unit, O.R.G.B. (2008). Sampling large-scale social networks: Insights from simulated networks. In *18th Annual Workshop on Information Technologies and Systems, Paris, France*.

[172] Leskovec, J. and Faloutsos, C. (2006). Sampling from large graphs. In *Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining*, 631-636.

[173] Lindner, G., Staudt, C.L., Hamann, M., Meyerhenke, H. and Wagner, D. (2015). Structure-preserving sparsification of social networks. In *Proceedings of the 2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*, 448-454.

[174] Mathioudakis, M., Bonchi, F., Castillo, C., Gionis, A. and Ukkonen, A. (2011). Sparsification of influence networks. In *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining*, 529-537.

[175] Leskovec, J., Kleinberg, J. and Faloutsos, C. (2005). Graphs over time: densification laws, shrinking diameters and possible explanations. In *Proceedings of the eleventh ACM SIGKDD international conference on Knowledge discovery in data mining*, 177-187.

[176] Goldenberg, A., Zheng, A.X., Fienberg, S.E. and Airoldi, E.M. (2010). A survey of statistical network models. *Foundations and Trends in Machine Learning*, **2**, 129-233.

[177] Stumpf, M.P., Wiuf, C. and May, R.M. (2005). Subnets of scale-free networks are not scale-free: sampling properties of networks. *Proceedings of the National Academy of Sciences of the United States of America*, **102**, 4221-4224.

[178] Stumpf, M.P. and Wiuf, C. (2005). Sampling properties of random graphs: The degree distribution, *Phys. Rev. E* **72**, 036118

[179] Lee, S.H., Kim, P.J. and Jeong, H. (2006). Statistical properties of sampled networks, *Phys. Rev. E*, **73**, 016102

[180] Granovetter, M. (1976). Network Sampling: Some First Steps. *American Journal of Sociology*, **81**, 1287-1303

[181] Sokolov, I.M., Eliazar, I.I. (2010). Sampling from scale-free networks and the matchmaking paradox, *Phys. Rev. E*, **82**, 26107.

[182] Ahmed, N., Neville, J. and Kompella, R.R. (2011). Network sampling via edge-based node selection with graph induction.

[183] Frank, O. and Snijders, T. (1994). Estimating the size of hidden populations using snowball sampling. *Journal of Official Statistics*, **10**, 53.

[184] Ahmed, N.K., Berchmans, F., Neville, J. and Kompella, R. (2010). Time-based sampling of social network activity graphs. In *Proceedings of the eighth workshop on mining and learning with graphs*, 1-9.

[185] De Choudhury, M., Lin, Y.R., Sundaram, H., Candan, K.S., Xie, L. and Kelliher, A. (2010). How does the data sampling strategy impact the discovery of information diffusion in social media? *ICWSM*, **10**, 34-41.

[186] Rhody, L.M. (2011). *A Method to the Model: Responding to Franco Morettis Network Theory, Plot Analysis*, `https://magmods.wordpress.com/2011/08/22/a-method-to-the-model-responding-to-franco-moretti\%E2\%80\%99s-network-theory-plot-analysis/`, accessed on 08.02.2017.

[187] Kenna, R., MacCarron, M. and MacCarron, P. eds. (2017). *Maths Meets Myths: Quantitative Approaches to Ancient Narratives*. Springer International Publishing.