National College of Ireland

# Identifying Key Players in the 1916 Rising using Centrality Measures and K-Shell Decomposition

MSc Research Project
Data Analytics

James Nagle
Student ID: x20191014

School of Computing
National College of Ireland

Supervisor: Dr Mohammed Hasanuzzaman

**National College of Ireland**

**MSc Project Submission Sheet**

**School of Computing**

| | | | |
|---|---|---|---|
| **Student Name:** | James Nagle | | |
| **Student ID:** | X20191014 | | |
| **Programme:** | MSc in Data Analytics | **Year:** | 2022 |
| **Module:** | MSc Research Project | | |
| **Supervisor:** | Dr Mohammed Hasanuzzaman | | |
| **Submission Due Date:** | 1 February 2023 | | |
| **Project Title:** | Identifying Key Players in the 1916 Rising using Centrality Measures and K-Shell Decomposition | | |
| **Word Count:** | 6260 | **Page Count** | 17 |

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.
ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

**Signature:** James Nagle

**Date:** 29 January 2023

**PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST**

| | |
|---|---|
| Attach a completed copy of this sheet to each project (including multiple copies) | □ |
| **Attach a Moodle submission receipt of the online project submission,** to each project (including multiple copies). | □ |
| **You must ensure that you retain a HARD COPY of the project**, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer. | □ |

Assignments that are submitted to the Programme Coordinator Office must be placed into the assignment box located outside the office.

| **Office Use Only** | |
|---|---|
| Signature: | |
| Date: | |
| Penalty Applied (if applicable): | |

# Identifying Key Players in the 1916 Rising using Centrality Measures and K-Shell Decomposition

James Nagle
X20191014

**Abstract**

Covert network analysis is largely dependent on a limited number of datasets created after a major event or the breakup of an organisation has occurred, leading to possible bias. Due to the lack of reliable data and their subjective nature it is often difficult to accurately evaluate the results of analysis on these networks. This research proposes a new network, based on intelligence reports compiled in the lead up to the 1916 Rising, on which covert network analysis can be conducted and which also contributes to a growing body of quantitative research on the Irish Revolutionary period. Several centrality measures commonly used for identifying influential nodes are calculated for each individual in the network to determine key players. K-shell decomposition is used to determine the individuals who inhabit the core of the network and community detection is used to identify closely connected factions. These results are then compared to the historical consensus, established over a century of qualitative analysis, to determine the suitability of each method. The analysis largely conforms to the established historical consensus and highlights the importance of figures whose role in the Rising is being re-evaluated in light of recently released documentation, indicating that the network can play a valuable role in assisting historical research and in serving as a benchmark for covert network analysis.

# 1 Introduction

In 2015 the National Archives released a tranche of reports entitled "the movement of extremists in Ireland"[1], compiled 100 years beforehand by Superintendent Owen Brien of G Division of the Dublin Metropolitan Police (DMP). G Division consisted of a small number of plainclothes detectives who attended public meetings and observed political activists (McMahon, 2008); between May 1915 and April 1916 they reported on members of the Irish Republican Brotherhood (IRB), the Irish Volunteers, the Irish Citizen Army and others designated as "extremists". These observations were collated by Brien and submitted daily to Sir Matthew Nathan, the Under Secretary for Ireland, at Dublin Castle. The final report was submitted on Thursday the 20th of April 1916; 4 days later many of those under observation seized key buildings throughout Dublin and staged a rebellion against British rule in Ireland.

By treating the observations in the reports as edges in an edge list a social network is created on which network analysis techniques are applied. This paper addresses the contributions this network makes to the study of covert networks in general and to the application of data analytics methods to the study of Irish history, an area of digital humanities with great potential but which has seen limited investigation.

[1] https://www.nationalarchives.ie/article/chief-secretarys-office-crime-branch-dublin-metropolitan-police-dmp-movement-extremists-april-1916/

Covert networks are networks in which some, or all, of the members seek to maintain some degree of secrecy, concerning either their identity, their actions or their future plans (Crossley, Edwards, Harries, & Stevenson, 2012). There has been significant growth in the use of social network analysis to study covert networks since the early 2000's (Ficara, Curreri, Fiumara, De Meo, & Liotta, 2022) however, due to the nature of such organisations, the field is heavily dependent on criminal networks created from evidence presented after its members have been brought to trial (Ficara, et al., 2021) and on a limited number of empirically verifiable datasets (Crossley, Edwards, Harries, & Stevenson, 2012). The network presented in this paper consists of a greater number of edges and nodes than those analysed by (Ficara, et al., 2021) and contains members who were not involved in the planning of the 1916 Rising, making the identification of key players non-trivial. The results returned by key player detection methods can be compared to over a century of historical research into the participants of the 1916 Rising. The most famous individuals have multiple biographies each while renewed interest around the centenary of the Rising has led to new works on some of the less well-known figures. As such the network can serve to test and benchmark covert network analysis techniques against an empirically verifiable dataset.

Historians Tom Garvin and David Fitzpatrick are credited by (Hart, 2003) as the first to "begin the quantitative analysis" of the Irish Revolution (a time period that covers the formation of the Irish Volunteers in 1913 through to the end of the Irish Civil War in 1923), using statistical models to search for correlations between a variety of socio-economic factors and participation levels in revolutionary activity. Using witness statements released by the Bureau of Military History[2] in 2012, (Borgonovo, Ó'Drisceoil, Crowley, Murphy, & Hogan, 2017) (to which both Garvin and Fitzpatrick contributed) used quantitative methods to challenge or support several long-held beliefs, such as activity levels in specific counties or the age profile and economic background of those who took part.

While (Borgonovo, Ó'Drisceoil, Crowley, Murphy, & Hogan, 2017) makes extensive use of visualisation techniques often overlooked in Irish historical research, such as charts and maps, no attempt has been made to apply social network analysis to any of the new documentation becoming available; indeed the only previously existing application of social network analysis in the field of Irish history is to accounts of the Battle of Clontarf in 1014 (Yose, Kenna, MacCarron, & MacCarron, 2018). Thus, the results and visualisations presented in this paper serve as a major contribution to the application of data analytics principles to the study of Irish history. They can also facilitate a clearer understanding of the role played by less well-known figures in the planning of the 1916 Rising and will highlight the potential for further study in this area.

Section 2 of this paper consists of a literature review covering related work in the areas of Covert Network Analysis, Community Detection and Visualisation and Historical Analysis of the 1916 Rising. Section 3 covers the Methodology used and section 4 deals with the Design Specification. Section 5 deals with the final stage of the implementation and the most pertinent results are presented in Section 6. Conclusions and further suggestions are then offered.

[2] https://www.militaryarchives.ie/collections/online-collections/bureau-of-military-history-1913-1921

# 2 Related Work

## 2.1 Covert Network Analysis

Based on earlier works in the field of sociology, (Freeman, 1977) is credited with formulating the generalized degree, betweenness and closeness centrality measures used in the analysis of networks. These measures are applied to identify important nodes based on the structure of a network (Grando, Noble, & Lamb, 2016) and have been used extensively in the analysis of covert networks, notably (Krebs, 2002) which used them to rank participants in the 9/11 terrorist attacks based on different definitions of importance. The surge in terrorist attacks in the 2000's led to more sophisticated methods of analysing covert networks. (Memon, 2012) revisited the findings of (Krebs, 2002) following developments in applying centrality measures to weighted networks; previously weights were largely disregarded. Other centrality measures popular in social network analysis were also tested, predominantly eigenvector (Waniek, Michalak, Wooldridge, & Rahwan, 2022) and PageRank (Hu, Chu, Xu, Wu, & Lia, 2021) which sought to determine the popularity of a node. Attempts were made to combine or adjust centrality measures to identify specific leadership figures, as in (Berzinji, Kaati, & Rezine, 2012), but as the values returned by these measures are highly dependent on the structure of the network (Grando, Noble, & Lamb, 2016) it is unlikely that such methods will generalize well to other networks (Bright, Brewer, & Morselli, 2021).

Study has also been conducted on how covert networks structure themselves, particularly in light of the need to balance secrecy and operational efficiency (Crossley, Edwards, Harries, & Stevenson, 2012). Highly centralized "star graphs" form around leaders with high levels of control and are considered optimal for information flow over an "all-to-all graph" (Lindelauf, Borm, & Hamers, 2009). Individuals sitting at the centre of a star graph would be expected to score highly across multiple centrality measures but their position would not be obvious from these measures alone, justifying their use alongside a visual examination of the network's structure.

The most used covert networks for analysis come from criminal organisations, especially those involved in drug dealing, and are formed from surveillance data or evidence presented in court after the organisation is broken up (Ficara, Curreri, Fiumara, De Meo, & Liotta, 2022). Data collected in this way can often suffer from incompleteness (key figures missed), incorrectness (wrong name attributed to suspects) and inconsistency (merging data from different surveillance sources resulting in multiple entries for a single person or meeting) (Ficara, et al., 2021). The study of terrorist organisations in covert network analysis has been criticised by (Crossley, Edwards, Harries, & Stevenson, 2012) for being overly focused on contemporary Jihadist groups and examining only a small "action set" associated with a major event, such as the analysis of 9/11 carried out by (Krebs, 2002). By focusing on just a small group of actors, without reference to or knowledge of the wider network these actors are a part of, key players can be missed, and the structure of the wider network may be greatly different to that of the action set. (Crossley, Edwards, Harries, & Stevenson, 2012) also highlighted that often the results of convert network analysis are not empirically verified against known accounts identifying who the key players were. This makes it difficult to determine how suitable these methods of analysis are. Early covert network analysis tended to focus on unweighted networks where the strength of the connection between different actors was unknown due to a lack of information. Improvements in weighted implementations of centrality measures (Opsahl, Agneessens, & Skvoretz, 2010) led to a re-examination of covert networks using weights (Memon, 2012) though most analysis continued on unweighted networks.

Centrality measures were used extensively in network analysis without any thorough assessment of the value of their results. (Batool & Niazi, 2014) validated 4 of the 5 centrality measures which will be examined in this project by comparing their results to knowledge of several networks, such as instructors in the Zachary's Karate Club network. They found that betweenness centrality was impacted by the network topology and that degree and eigenvectory centrality were closely correlated.

## 2.2 Community Detection and Visualisation

Community detection algorithms are used in social network analysis to identify cliques of closely connected nodes (Moscato & Sperlì, 2021). K-shell decomposition can be used to find the nodes which constitute the "core" of the network by iteratively removing nodes with a low degree (Dumba & Zhang, 2018). The members of this core are found to influence or spread information to larger parts of the network (Miorandi & De Pellegrini, 2010). There have been no applications of k-shell decomposition to case studies of covert networks but it would be expected to find the most senior and longest serving members of a conspiracy residing in the core, regardless of results from centrality measures (Kitsak, et al., 2010). Advancements have been made on network visualisation tools such as Gephi which can colour or resize nodes and edges based on centrality measures or community membership (Majeed, Uzair, Qamar, & Farooq, 2020).

## 2.3 Historical Analysis of the 1916 Rising

A substantial body of work has been published on the 1916 Rising, both in isolation and in the context of the wider Irish revolution (Ó'Tuathaigh, 2017). Renewed interest around the centenary of the Rising led to the release of updated biographies on all of those who were executed (such as (O'Donnell, 2016), representative of the 16 Lives series) and specialized accounts of the event, such as (Townshend, 2005). Many historians also sought to fill the gap in historical research by authoring works for the first time on important but overlooked figures ((McGee, 2015) and (McGough, 2013) for example). As a result of its secretive nature, it has proven extremely difficult to provide more than a cursory account of the functioning of the Irish Republican Brotherhood, the organisation which planned and carried out the 1916 Rising. (Ó'Broin, 1976) and (McGee, 2005) represent the only attempts at a dedicated history of the IRB but due to a lack of information and the execution of its most central members in 1916, both deal mainly with events prior to 1900. Rough outlines of what actions the main planners took in the lead up to the Rising were set out by contemporaries such as (O'Donoghue, 1963). These outlines have been accepted by later academics and are in line with witness statements recently released by the Military Archives. Though the 1916 Rising has been a topic of study for over 100 years, no attempt prior to the publication of (Borgonovo, Ó'Drisceoil, Crowley, Murphy, & Hogan, 2017) was made to carry out a quantitative analysis of the event. This analysis mainly consists of maps showing the location and size of Irish Volunteer battalions prior to the Rising and the home location of those interned in its aftermath [for example (McCarthy & Wrynn, 2017)].

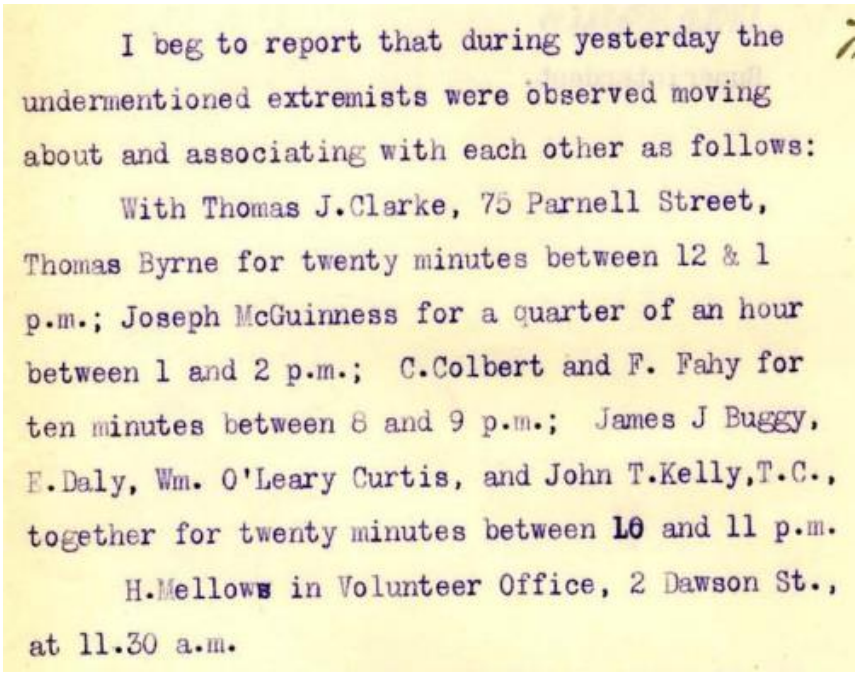
Quantitative statistics have been used since the 1970's, most notably in (Hart, 2003) which found that the level of violence a county witnessed during the Irish Revolution was positively correlated with emigration levels following the Great Hunger and violence during the Land War. Only recently have historians looked to use data to drive research and utilize methods from data science, mainly in the area of visualisation. No attempt has been made to

apply social network analysis to any of the documentation released over the last decade, such as intelligence reports and diaries maintained by prominent revolutionaries, nor has there been any examination as to the role data science methods can play in the study of Irish history.

# 3 Research Methodology

To complete this project a ground truth edge list was created from the intelligence reports available. This edge list was then used to generate a network on which centrality measures, k-shell decomposition and community detection algorithms were applied. The results were then compared to accounts of participants, contemporaries and historians to determine how useful they were.

The movement of extremists in Ireland intelligence reports consist of observations made by plainclothes detectives, compiled into typewritten reports and submitted to Dublin Castle over an 11-month period. An edge list is created from these observations consisting of a date, person a and person b who were observed together. Figure 1 shows a section of the report collected on the 30th of June 1915. Figure 2 shows how this section is represented as an edge list. A count of each pairing in the edge list can be used to represent the weight of the connection for that pair.



I beg to report that during yesterday the undermentioned extremists were observed moving about and associating with each other as follows:

With Thomas J.Clarke, 75 Parnell Street, Thomas Byrne for twenty minutes between 12 & 1 p.m.; Joseph McGuinness for a quarter of an hour between 1 and 2 p.m.; C.Colbert and F. Fahy for ten minutes between 8 and 9 p.m.; James J Buggy, E.Daly, Wm. O'Leary Curtis, and John T.Kelly,T.C., together for twenty minutes between 10 and 11 p.m.

H.Mellows in Volunteer Office, 2 Dawson St., at 11.30 a.m.

**Figure 1: A section from a typewritten report**

It is necessary to construct the edge list from a manual examination of the reports to limit the issues raised by (Ficara, et al., 2021). Several of the digitised files are in poor condition and the words appear faint; this would be difficult for Optical Character Recognition software to process and convert to text correctly. Individuals appear in the reports under different variants of their name, such as “Mr. Briscoe” and “J. M. Briscoe”. Domain knowledge is required to verify that these variants refer to the same individual. Misspellings are also common.

(Yose, Kenna, MacCarron, & MacCarron, 2018) also extracted the data used to build the network of the Battle of Clontarf by reading the historical account and manually creating an edge list. Multiple passes through the text were made by different researchers to remove subjectivity in deciding who should be connected.

The network used here will also suffer from incompleteness as many meetings of the IRB that were held in secret are not included in the reports and it can be surmised that many other public meetings are also not included. This will need to be taken into consideration when assessing the results of the centrality measures returned by network analysis. This is a problem faced by all covert networks built on surveillance data. The network can still be examined to determine the value of the data the DMP managed to collect and present an analysis of the interactions between major historical figures and groups in the lead up to the 1916 Rising. It is not expected that the missing meetings will have much, if any, impact on unweighted measures taken on the network as the members of the IRB met often at public events or at the business premises of Tom Clarke, though not all together. Weighted centrality measures may therefore be affected more.

```
30 June 1915,Tom Clarke,Thomas Byrne
30 June 1915,Tom Clarke,Joseph McGuinness
30 June 1915,Tom Clarke,Con Colbert
30 June 1915,Tom Clarke,Frank Fahy
30 June 1915,Con Colbert,Frank Fahy
30 June 1915,Tom Clarke,J.J. Buggy
30 June 1915,Tom Clarke,Edward Daly
30 June 1915,Tom Clarke,William O'Leary Curtis
30 June 1915,Tom Clarke,Seán T. O'Kelly
30 June 1915,J.J. Buggy,Edward Daly
30 June 1915,J.J. Buggy,William O'Leary Curtis
30 June 1915,J.J. Buggy,Seán T. O'Kelly
30 June 1915,Edward Daly,William O'Leary Curtis
30 June 1915,Edward Daly,Seán T. O'Kelly
30 June 1915,William O'Leary Curtis,Seán T. O'Kelly
```

**Figure 2: An edge list representation of the report in Figure 1**

The network will also be impacted by the heavy surveillance conducted on Tom Clarke and those meeting at his shop. Clarke was a senior revolutionary who had been imprisoned for 15 years in 1883 for his role in the Fenian Dynamite Campaign (McMahon, 2008) and his premises are the most surveyed in the reports, ahead of the offices of the Irish Volunteers. It can be expected that Clarke will be connected to many people more than others in the network, however his contacts that were not connected with the wider planning of the Rising will most likely be pushed to the outside periphery of the network as they will have little to no contact with other conspirators. A focus on known individuals is also common in other covert networks and the bias towards observing Clarke will also need to be considered when assessing the results.

No uniform definition of centrality exists and attempts to formulate what all centrality measures have in common have proven unsuccessful (Borgatti & Everett, 2006). (Landherr, Friedl, & Heidemann, 2010) building on the study of centrality measures in (Borgatti & Everett, 2006), examined four different concepts of centrality:

- Degree, in which the number of direct connections to other nodes represents a node's centrality, with those having many connections being considered more central than those with only a few.
- Closeness, where nodes which are closest to all other nodes by direct and indirect connections are considered more central than those who sit towards the outside of the network.
- Betweenness, which examines the centrality of a node's position within the network with reference to how many of the shortest paths between all other nodes it sits on. A node with high betweenness centrality would sit on many of these paths and would be involved in the flow of information around the network.
- Influence, where a node's centrality is determined by how well connected its neighbours are. Also known as popularity, a node with a few very popular neighbours would be more central than a node with many neighbours with low popularity.

Degree centrality, closeness centrality and betweenness centrality, initially outlined by (Freeman, 1977), are the most common centrality measures used in covert network analysis. To measure influence eigenvector centrality is most often used. These measures will be used in this project, along with PageRank centrality, the algorithm used to rank webpages by the Google search engine (Brin & Page, 1998), as it is becoming more widely used in network analysis.

Both k-shell decomposition and the modularity community detection algorithms will be used to examine the structure of the network. k-shell decomposition will partition the network into a hierarchy of substructures (Liu, Tang, Zhou, & Do, 2015) and modularity will be used to divide the network into communities where the nodes in each community are strongly connected but weakly connected to the nodes in other communities (Newman, 2006). K-shell decomposition will return a core containing influential nodes while modularity will detect relationships and structures within the network.

The results of the above measures can then be compared to the historical consensus surrounding the 1916 Rising. As one of the most important events in modern Irish history, leading to the foundation of the Irish state, the 1916 Rising has been studied for over a century. This body of knowledge can be used to empirically validate the results of applying centrality measures to the network. The lack of attempts to do so in covert network analysis before this has been criticised by (Crossley, Edwards, Harries, & Stevenson, 2012). Historical accounts were used by (Carvalho, 2022) to validate the rules of a board game created to simulate the late 13th century O'Brien civil war in Thomond, justifying the use of this method in studying historical events.

The central role played by the signatories of the Proclamation is well known, particularly Tom Clarke and Seán Mac Diarmada (Townshend, 2005) who served as Treasurer and Secretary of IRB respectively. As such, these figures would be expected to score highly across multiple centrality measures and appear in or close to the k-core. Their absence, or the presence of individuals known not to have been involved in the events leading to the 1916 Rising, will need to be explained by reference to the network structure, the nature of the data collected by the DMP or the suitability of the centrality measures used. The results are also

expected to highlight figures whose role is recognised, but often overlooked or downplayed, in the organisation of the 1916 Rising. The relationships and structures returned by modularity analysis will be compared to known divisions within the nationalist movement at the time. It is expected that the hierarchy of the Irish Volunteers, who were unaware that a rising was to be held and attempted to stop it, would form a cluster while another would form around senior members of the IRB's Military Council which planned the Rising in secrecy. Other clusters may also appear, representing the suffragette movement, nationalist politicians or other groups who did not feature often in the DMP Reports.

# 4 Design Specification

The edge list used in this project was created by reading each of the movement of extremists in Ireland reports and adding each pair of observed suspects to a text notepad. The python programming language was chosen to process this file to leverage the specialist libraries it contains, most notably pandas for manipulating data and networkx for building and analysing networks.

Functions were written in python that could filter the edge list between two date ranges, use this to build a network and calculate the chosen centrality measures. Functions were also written that would take an already existing network, remove edges below a given weight and then recalculate the centrality measures.

Weighted and unweighted measures were calculated for each node using all 11 months of data and a filtered version of the network where the largest connected component was retained after removing all edges with a weight less than 23. This was done to investigate if removing weak edges would return a simplified network containing only influential figures.

The number 23 was chosen from several trials and a visual analysis of the resultant networks. A weight of 23 would indicate that two individuals spoke or were seen together at least this many times in the 11-month period covered by the reports. While it is reasonable to consider that nodes who only have edges less than this value had little impact on the overall network, this method will be shown to be overly selective in comparison to k-shell decomposition and too reliant on user definition of what constitutes a weak edge.

# 5 Implementation

The project was implemented in IPython notebooks with additional functions written in a python script and imported as needed. Separate notebooks were written for the unweighted and weighted analysis of the networks and the application of k-shell decomposition. The Gephi software package was then used to visualise the networks generated in the previous steps and to run modularity analysis. Modularity analysis identifies nodes which are strongly connected to each other and assigns them to communities. This is used to identify underlying structures in the network.

The results from the centrality measures, k-shell decomposition and modularity analysis were then compared to several works on the Irish revolution, mainly (Borgonovo, Ó'Drisceoil, Crowley, Murphy, & Hogan, 2017), (Townshend, 2005), (Ó'Broin, 1976) and dedicated biographies of major figures, such as (O'Donnell, 2016). Other sources, such as the BMH files, were consulted for details on less well-known individuals.

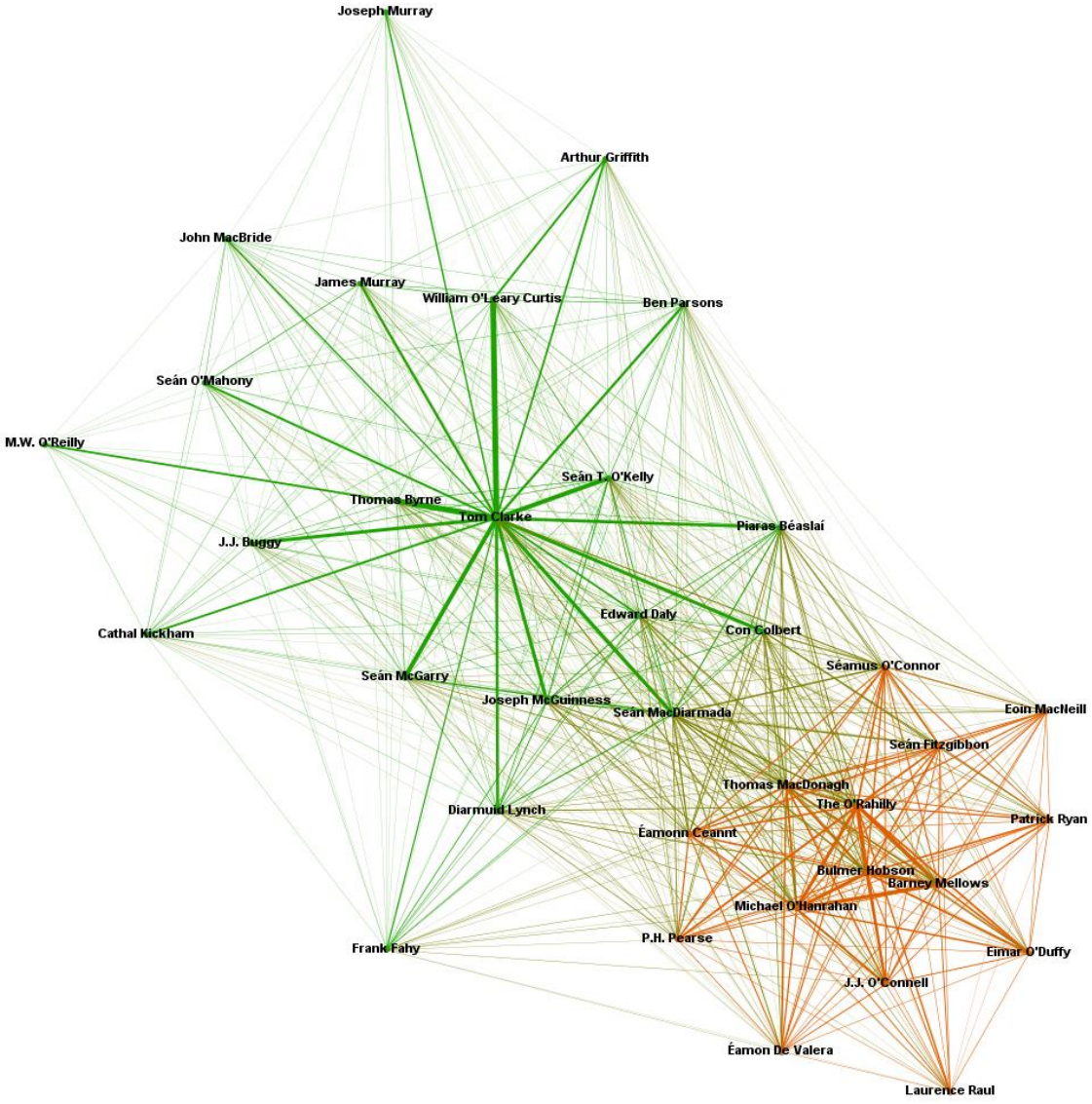



**Figure 3: Modularity analysis applied to the weight23 network, showing the IRB (green) and the Irish Volunteers (orange)**

# 6 Evaluation

Centrality measures were calculated for each node across different weighted and unweighted networks and a ranking table was then produced showing each node's rank across the different measures. Important figures such as the signatories of the Proclamation of the Irish Republic and dual members of the IRB and Irish Volunteers who held key positions during the 1916 Rising were found to rank highly across at least four of the measures on both the weighted and unweighted versions of the network. The evaluation of the results will focus on lesser-known figures who have ranked highly, well known figures who have not ranked highly and the suitability of the chosen methods for identifying key players.

For the full unweighted network there appears to be a correlation between degree, closeness and, to a lesser extent, eigenvector centrality as shown in Table 1. Five of the seven signatories of the Proclamation rank in the top 15 nodes in at least four measures. Of the

remaining signatories, Joseph Plunkett featured rarely in the reports due to trips to Germany and ill health while the political meetings of James Connolly were rarely observed by the DMP and he only joined the IRB late in the planning. The same measures applied on the weighted network show degree, eigenvector and PageRank to be somewhat correlated and generally return the same individuals, though many only rank highly in three measures.

**Table 1: Top 15 nodes ranked by Degree and their ranks in Closeness and Eigenvector Centrality, not taking weights into account.**

| Name | Degree Rank | Closeness Rank | Eigenvector Rank |
|---|---|---|---|
| Tom Clarke | 1 | 1 | 9 |
| Seán T. O'Kelly | 2 | 2 | 1 |
| The O'Rahilly | 3 | 3 | 2 |
| Piaras Béaslaí | 4 | 3 | 4 |
| Bulmer Hobson | 5 | 3 | 5 |
| Éamonn Ceannt | 6 | 6 | 3 |
| Thomas MacDonagh | 7 | 7 | 8 |
| Seán MacDiarmada | 8 | 7 | 6 |
| Joseph McGuinness | 9 | 9 | 7 |
| Edward Daly | 10 | 10 | 10 |
| Michael O'Hanrahan | 11 | 11 | 14 |
| P.H. Pearse | 12 | 13 | 13 |
| Séamus O'Connor | 12 | 13 | 12 |
| J.J. Walsh | 14 | 12 | 20 |
| Diarmuid Lynch | 15 | 16 | 11 |

In 1915 the IRB secured the election of Eoin MacNeill as President of the Gaelic League, but he would later be criticized for not attending meetings[3]. It could be argued that this played into the hands of the IRB and that they had earlier urged him to establish and lead the Irish Volunteers knowing that he would not attend its meetings, allowing them to place members into influential positions and guide the organisation in his absence. This is supported by the analysis conducted here; of the 12-person Central Executive elected in October 1915 only 3 rank lower than MacNeill across the measures used on the unweighted network. When weights are taken into account, he drops a further place. MacNeill is ranked 22nd to 26th on 4 centrality measures, just outside the top 10% of nodes on the unweighted network and performs similarly on the weighted network.

Of interest are the high ranks of Seán T. O'Kelly, Piaras Béaslaí and Diarmuid Lynch, all three of whom were members of the Irish Republican Brotherhood and the Gaelic League. It is known that O'Kelly and Lynch played an important role in organising the 1916 Rising and disseminating orders to IRB members in the Irish Volunteers though they are most often overlooked in favour of accounts which focus on those who were executed. Only Lynch has a dedicated biography and Béaslaí appears only a few times in (Townshend, 2005). All three rank slightly lower on measures applied to the weighted network.

On both the weighted and unweighted networks betweenness centrality ranks highly several individuals known to have had limited involvement in the wider republican movement at the time. Table 2 shows the unweighted network ordered by betweenness.

[3] https://www.oireachtas.ie/en/debates/debate/dail/1919-08-20/17/

**Table 2: Top 5 nodes ranked by Betweenness and their ranks in Degree and Eigenvector Centrality on the unweighted network.**

| Name | Degree Rank | Betweenness Rank | Eigenvector Rank |
|---|---|---|---|
| Tom Clarke | 1 | 1 | 1 |
| J.J. Walsh | 14 | 2 | 20 |
| James Murray | 52 | 3 | 62 |
| Piaras Béaslaí | 4 | 4 | 4 |
| Hanna Sheehy-Skeffington | 93 | 5 | 94 |
| Éamonn Ceannt | 6 | 6 | 3 |
| Seán T. O'Kelly | 2 | 7 | 1 |
| Bulmer Hobson | 5 | 8 | 5 |
| The O'Rahilly | 3 | 9 | 2 |
| Miss Burke Dowling | 204 | 10 | 222 |
| Seán MacDiarmada | 8 | 11 | 6 |
| Michael O'Hanrahan | 11 | 12 | 14 |
| Thomas MacDonagh | 7 | 13 | 8 |
| John MacBride | 37 | 14 | 50 |
| Joseph McGuinness | 9 | 15 | 7 |

A visual analysis of the network shows that Hanna Sheehy-Skeffington for example was the only node connecting a small suffragette cluster to the wider network which would increase her betweenness score. Similar situations account for the high betweenness centrality and low centrality on the remaining measures of the other figures. J.J. Walsh is a valuable exception. He ranks highly across four unweighted measures, which is difficult to reconcile with his relative lack of involvement in the IRB or the Irish Volunteers compared to others of similar rank. On the weighted network however, he performs poorly on all but betweenness. As an organiser with the GAA he has many links across the political spectrum but the weighted analysis shows these to be largely weak and mostly with marginal figures.

Weighted closeness centrality ranks highly some individuals on whom practically nothing is known, namely Ben Parsons and J.J. Buggy. Parsons only appears once and Buggy twice, both in passing, in the BMH witness statements. Little is known about William O'Leary Curtis or the role played by Thomas Byrne, who met almost daily with Tom Clarke throughout 1915. All were in close contact with Tom Clarke and were likely members of the IRB. These figures in particular are deserving of further study to determine the reasons for their high centrality.

Applying modularity to the full, weighted network found three large clusters and several smaller ones connected to groups like the Ancient Order of Hibernians and the suffragettes.

Of the two largest clusters one centred around Tom Clarke forming a "star graph" while the other formed a heavily connected "all-to-all graph" based around senior leaders of the Irish Volunteers, structures identified in (Lindelauf, Borm, & Hamers, 2009). This can be seen more clearly in the "weight23" network shown in Figure 3, where the largest connected component is retained after removing all edges with a weight of less than 23. Clarke controls the flow of information around a cluster which largely consists of IRB members and others, such as Arthur Griffith, who were kept in the dark as to the planning of the 1916 Rising. Those involved in the conspiracy, such as MacDiarmada and Pearse, appear at the edge of both clusters, controlling the flow of information in and out of the Irish Volunteers. Those in the Irish Volunteers cluster have much stronger ties to each other than in the IRB cluster where most nodes only have a single edge of weight 23 or higher.

As the "weight23" network only contains 36 nodes, recalculating unweighted centrality measures on this network returns identical results for many nodes as only a limited number of

edges exist. As such these results are of limited value. When recalculated using weighted centrality measures many of the same figures who ranked highly on the full network are returned. Tom Clarke ranks 1st in every measure and the four central figures who sit at the edge of the IRB and Irish Volunteer clusters (MacDiarmada, Colbert, MacDonagh and Ceannt) rank inside the top 10 across multiple measures. Due to the limited size of the network the betweenness score is the same for most of the nodes and is of limited use outside the top 8 ranks.

Applying k-shell decomposition to the full network of 229 nodes returns 40 cores with kmax (k40) containing 41 nodes and kmax-1 (k39) containing 30 nodes. The reports mention 11 individuals who would be executed in the aftermath of the 1916 Rising; 8 of these feature in the kmax core and the remaining 3 are in kmax-1. Of the 7 signatories, 6 appear in kmax and the 7th is in kmax-1. The kmax core contains a number of important IRB officers, such as Terence MacSwiney and Thomas Ashe, who do not appear in any of the analyses conducted so far. These individuals ranked low on centrality measures and featured rarely in the reports as they were based outside of Dublin. Ashe however commanded the Irish Volunteers in Ashbourne during Easter Week and MacSwiney would have commanded those in Cork had they risen.

The kmax core includes all of the Commandants and almost all of the Vice Commandants of the four Dublin Battalions and the Fingal Battalion which took part in the Rising. In total 30% of all nodes in the network are included in kmax and kmax-1, as shown in Figure 4, which are the two largest of the cores returned. Using domain knowledge, justification can be made for the inclusion of most of these figures and a graph generated from these two cores maintains the IRB and Irish Volunteer structures observed in the overall network.

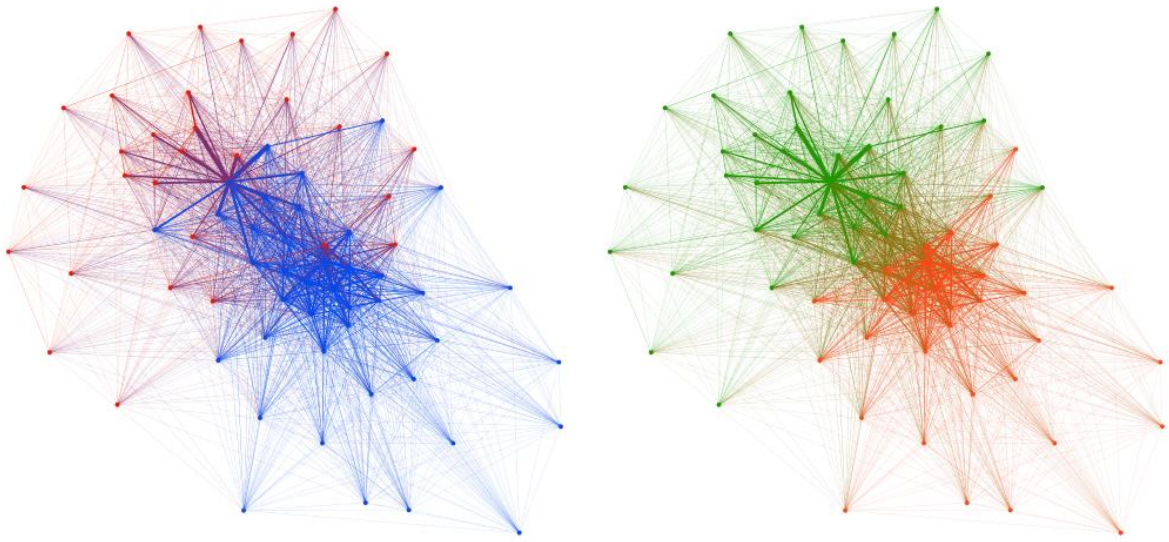


**Figure 4: K-Shell Decomposition applied to the network. The left shows kmax (blue) and kmax -1 (red), the right shows the results of applying modularity analysis.**

The analysis has not been greatly impacted by the absence of meetings, such as those of the Military Council, though the sporadic reporting on groups such as the suffragettes and the Ancient Order of Hibernians is responsible for the inconsistent betweenness scores. The importance of James Connolly and those close to him appears to have been negatively impacted by the lack of reports on meetings he was involved with. The heavy surveillance placed on Tom Clarke is somewhat mitigated by his departure to Limerick in January 1916 to recover from an injury though the quality of the reports submitted deteriorates from this time

on. A daily list of individuals who attended his shop in his absence is given but with no indication of times to determine who visited together. While this may lead to some edge weights being underrepresented it is unlikely to have caused the total absence of an edge which would materially impact the overall results.

Returning to the modularity analysis of the full network, shown in Figure 5, it can be seen that the third largest cluster (light blue) which does not appear in the weight23 network or the in the results of k-shell decomposition consists of members of the wider nationalist movement, individuals who were largely disconnected from the planning of the 1916 Rising or introduced into the conspiracy late. James Connolly, Constance Markievicz and the Irish Citizen Army appear here, as do Irish Parliamentary Party politicians such as Alfie Byrne. Reducing the size of the network and removing weak edges breaks down this cluster while assigning its most connected members to either of the two larger clusters, predominantly the IRB cluster due to its members attempts to involve themselves in and infiltrate various political and cultural movements. Small clusters representing the Gaelic Athletic Association (black), the Ancient Order of Hibernians (pink) and suffragette organisers (red) can be seen as well, along with two other clusters that contain only two members each. At all levels the graphs and measures generated are dominated by the hierarchy of the Irish Volunteers and a cadre of IRB members who had infiltrated the organisation with the intent of using it to stage a rebellion against British rule, keeping their intent secret from both the Volunteers and most of their own membership.

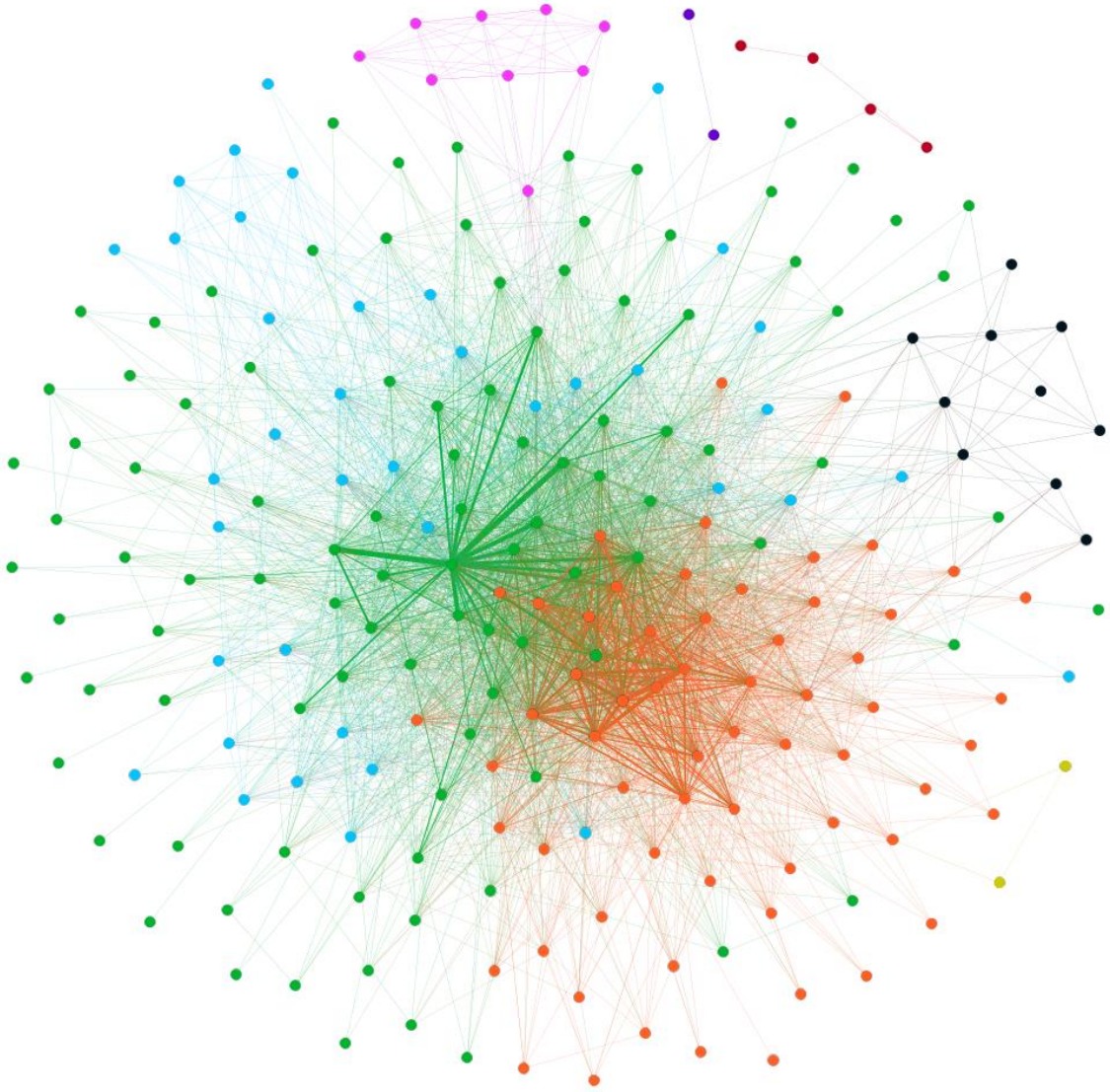


**Figure 5: Modularity applied to the full network. The three largest clusters represent the IRB (green), the Irish Volunteers (orange) and the wider nationalist movement (blue).**

# 7 Conclusion and Future Work

Centrality measures have been used extensively to identify key players in covert networks though few attempts have been made to validate their findings empirically. The results of applying centrality measures to a network made from intelligence reports compiled during the planning of the 1916 Rising show that betweenness centrality is heavily impacted by the structure of the network, to the point that its value is limited without a thorough visual inspection of the network or reference to other centrality measures. There was a correlation between the ranks awarded by degree and eigenvector centrality on both weighted and unweighted implementations and these were also correlated to closeness centrality on unweighted implementations. The results for the unweighted implementations are in line with the results of (Batool & Niazi, 2014). PageRank also shows correlation with eigenvector centrality, suggesting it is suitable for use in covert network analysis and might be more be of more value as it is similar to explain.

Individuals who ranked highly across multiple centrality measures were shown to have held important positions in the revolutionary movement prior to the 1916 Rising. The most important organisers and commanders involved were ranked consistently in the top 15 nodes across multiple measures, even in the absence of reports on meetings of the Military Council, the secretive body which planned the Rising. The measures showed Eoin MacNeill to be largely disconnected from the day-to-day operations of the Irish Volunteers and that Seán MacDiarmada was more central than P.H. Pearse. As the most public member of the IRB, Pearse's role is often overstated at MacDiarmada's expense, though the latter's importance is attested to in sources.

Removing weak edges and using k-shell decomposition returned highly important figures and allowed for the creation of simplified visualisations which showed the underlying structure of the IRB and Irish Volunteers. K-shell decomposition also returned many important figures who did not rank highly in the centrality measures used. This serves as a much more efficient way of reducing the size of the network than removing weak edges, which requires trials to arrive at a value below which edges will be removed and is susceptible to bias, in that a value may be selected that includes or excludes specific individuals in the resulting network. This project shows that k-shell decomposition can identify key players in covert networks and serves as a justification for its further study on other covert networks.

This project represents the first application of social network analysis to the events of the Irish Revolution and contributes to a small but growing body of work applying data science techniques to the study of Irish history. The results presented conform to over a century of historical analysis but have also highlighted overlooked figures who could serve as the basis for further qualitative research. The visualisations of the weight23 network and the two largest cores show the structure of the main organisations and the relative positioning of key players in a way that would take considerable effort to convey in writing. The results of this project encourage further research into the overlooked figures identified as being important and suggest that valuable findings and research topics can be identified by applying data science techniques to Irish historical documentation.

The findings of k-shell decomposition in particular show that the 1916 Rising network can serve as a valuable tool for benchmarking methods for covert network analysis and contributes to a small body of literature which validates the use of centrality measures to identify key players. It is suggested in future to conduct a temporal analysis of the network, examining how it changes over time with the goal of identifying when James Connolly and the Irish Citizen Army were brought fully into the conspiracy. The ability to identify organisational or structural change within a covert network would be of tremendous value in defense and security policy making. The network could also be expanded by including meetings not observed by the DMP but for which historical accounts exist, such as meetings of the Military Council, to build a more accurate understanding of one of the most important events in Irish history.

# References


Batool, K., & Niazi, M. A. (2014). Towards a methodology for validation of centrality measures in complex networks. *PLoS One, 9*(4).

Berzinji, A., Kaati, L., & Rezine, A. (2012). Detecting Key Players in Terrorist Networks. *2012 European Intelligence and Security Informatics Conference*, (pp. 297-302).

Borgatti, S., & Everett, M. G. (2006). A Graph-theoretic perspective on centrality. *Social Networks, 28*(4), 466-484.

Borgonovo, J., Ó'Drisceoil, D., Crowley, J., Murphy, M., & Hogan, N. (Eds.). (2017). *Atlas of the Irish Revolution.* Cork: Cork University Press.

Bright, D., Brewer, R., & Morselli, C. (2021). Using social network analysis to study crime: Navigating the challenges of criminal justice records. *Social Networks, 66*, 50-64.

Brin, S., & Page, L. (1998). The anatomy of a large-scale hypertextual Web search engine. *Computer Networks and ISDN Systems, 30*(1), 107-117.

Carvalho, V. M. (2022). War, hazards, and economic degradation in Thomond, 1276-1318: agent-based approaches to the Uí Bhriain civil war. [Doctoral dissertation, University of Sao Paulo].

Crossley, N., Edwards, G., Harries, E., & Stevenson, R. (2012). Covert social movement networks and the secrecy-efficiency trade off: The case of the UK suffragettes (1906–1914). *Social Networks, 34*(4), 634-644.

Dumba, B., & Zhang, Z. (2018). Uncovering the Nucleus of Social Networks. *WebSci '18: Proceedings of the 10th ACM Conference on Web Science*, (pp. 37-46).

Ficara, A., Cavallaro, L., Curreri, F., Fiumara, G., De Meo, P., Bagdasar, O., . . . Liotta, A. (2021). Criminal networks analysis in missing data scenarios through graph distances. *PloS one, 16*(8).

Ficara, A., Curreri, F., Fiumara, G., De Meo, P., & Liotta, A. (2022). Covert Network Construction, Disruption, and Resilience: A Survey. *Mathematics, 10*(16), 1-43.

Freeman, L. (1977). A Set of Measures of Centrality Based on Betweenness. *American Sociological Association, 40*(1), 35-41.

Grando, F., Noble, D., & Lamb, L. C. (2016). An Analysis of Centrality Measures for Complex and Social Networks. *2016 IEEE Global Communications Conference (GLOBECOM)*, (pp. 1-6).

Hart, P. (2003). *The I.R.A. at War 1916-1923.* Oxford: Oxford University Press.

Hu, J., Chu, C., Xu, L., Wu, P., & Lia, H. (2021). Critical Terrorist Organization in Terrorist Organization Alliance Network Based on Key Nodes Founding. *Frontiers in Physics, 9*.

Kitsak, M., Gallos, L. K., Havlin, S., Liljeros, F., Muchnik, L., Eugene Stanley, H., & Makse, H. A. (2010). Identification of influential spreaders in complex networks. *nature physics, 6*, 888-893.

Krebs, V. (2002). Mapping Networks of Terrorist Cells. *Connections, 24*(3), 43-52.

Landherr, A., Friedl, B., & Heidemann, J. (2010). A Critical Review of Centrality Measures in Social Networks. *Business & Information Systems Engineering, 2*, 371-385.

Lindelauf, R., Borm, P., & Hamers, H. (2009). The influence of secrecy on the communication structure of covert networks. *Social Neworks, 137*(2), 126-137.

Liu, Y., Tang, M., Zhou, T., & Do, Y. (2015). Improving the accuracy of the k-shell method by removing redundant links: From a perspective of spreading dynamics. *Scientific Reports, 5*.

Majeed, S., Uzair, M., Qamar, U., & Farooq, A. (2020). Social Network Analysis Visualization Tools: A Comparative Review. *2020 IEEE 23rd International Multitopic Conference (INMIC)*, (pp. 1-6).

McCarthy, M., & Wrynn, S. (2017). The Irish Volunteers In County Galway. In *Atlas of the Irish Revolution.* Cork: Cork University Press.

McGee, O. (2005). *The IRB: The Irish Republican Brotherhood from the Land League to Sinn Féin.* Dublin: Four Courts Press.

McGee, O. (2015). *Arthur Griffith.* Kildare: Merrion Press.

McGough, E. (2013). *Diarmuid Lynch: A Forgotten Irish Patriot.* Cork: Mercier Press.

McMahon, P. (2008). *British Spies & Irish Rebels.* Suffolk: The Boyle Press.

Memon, B. (2012). Identifying Important Nodes in Weighted Covert Networks Using Generalized Centrality Measures. *2012 European Intelligence and Security Informatics Conference*, (pp. 131-140).

Miorandi, D., & De Pellegrini, F. (2010). K-shell decomposition for dynamic complex networks. *8th International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks*, (pp. 488-496).

Moscato, V., & Sperlì, G. (2021). A survey about community detection over On-line Social and Heterogeneous Information Networks. *Knowledge-Based Systems, 224*.

Newman, M. E. (2006). Modularity and communty structure in networks. *PNAS, 103*(23), 8577-8582.

Ó'Broin. (1976). *Revolutionary Underground: The Story of the Irish Republican Brotherhood 1858-1924.* Dublin: Gill and Macmillan.

O'Donnell, R. (2016). *Patrick Pearse.* Dublin: The O'Brien Press ltd.

O'Donoghue, F. (1963). Plans for the 1916 Rising. *Edinburgh University Review, 3*(1), 3-21.

Opsahl, T., Agneessens, P., & Skvoretz, J. (2010). Node centrality in weighted networks: Generalizing degree and shortest paths. *Social Networks, 32*(3), 245-251.

Ó'Tuathaigh, G. (2017). The Historiography of the Irish Revolution. In *Atlas of the Irish Revolution.* Cork: Cork University Press.

Townshend, C. (2005). *Easter 1916: The Irish Rebellion.* London: The Penguin Group.

Waniek, M., Michalak, T., Wooldridge, M., & Rahwan, T. (2022). How Members of Covert Networks Conceal the Identities of Their Leaders. *ACM Transactions on Intelligent Systems and Technology, 13*(1), 1-29.

Yose, J., Kenna, R., MacCarron, M., & MacCarron, P. (2018). Network analysis of the Viking Age in Ireland as portrayed in Cogadh Gaedhel re Gallaibh. *Royal Society Open Science, 5*(1).