MODELING AND ANALYZING THE UNITED STATES COURTS OF LAST RESORT'S LEGAL CITATION SYSTEM AS A COMPLEX SYSTEM

by

Louai Magbol Mohammed

A dissertation submitted to the faculty of
The University of North Carolina at Charlotte
in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in
Computer Science

Charlotte
2021

Approved by:

Dr. Mirsad Hadzikadic

Dr. Jason Windett

Dr. Zbigniew Ras

Dr. Mohamed Shehab

ABSTRACT

LOUAI MAGBOL MOHAMMED. Modeling and Analyzing the United States Courts of Last Resort's Legal Citation System as a Complex System. (Under the direction of DR. MIRSAD HADZIKADIC)

Courts of last resort in the United States are becoming increasingly important in American politics as the number of cases, influential decisions, and controversial issues continue to rise in the states. In discussions of federalism in the United States, these critical institutions are often overlooked as a complex system, due to substantial data limitations on the behavior and outcomes of these courts. I situate state courts of last resort as a complex adaptive system in the broader U.S. framework. I then seek to redress the data shortcomings by introducing a comprehensive database on state courts of last resort from 1953-2010. Using advanced data-capture techniques, I evaluate my parsers to capture the ever-changing structures of the source documents. This database will be the largest in scope and case detail to date. Moreover, it should further our understanding of judicial decision making and assist the prediction of the impact of institutional change on the system.

In addition, I modeled and analyzed the system as a complex adaptive system. Since the system has network characteristics, I used the approach of network science to model the system based on the citation behavior. Moreover, I created an automated dictionary-based classification model to extract and classify the citation treatments for the court cases.

Using state-of-the-art algorithms in network science and natural language processing, I was able to analyze the system and test the performance of the algorithms based on the system characteristics.

ACKNOWLEDGMENTS

At the very outset, all praises to ALLAH, the source of wisdom and knowledge, for His blessing and giving me the strength, knowledge, ability, and opportunity to complete this work successfully.

I extend my sincere gratitude to my esteemed supervisor, Prof. Mirsad Hadzikadic, for his invaluable supervision, support, and tutelage during the course of my Ph.D. degree.

I also extend my sincere gratitude to Dr. Jason Windett for his treasured support. His immense knowledge and plentiful experience were truly influential in shaping my research and critiquing my results.

In addition, I want to thank Dr. Zbigniew Ras and Dr. Mohamed Shehab for their thoughtful and constructive suggestions and feedback.

Finally, I would like to offer my special thanks to the Institute of Public Administration in Saudi Arabia for the scholarship opportunity to undertake my studies in the United States.

DEDICATION

The culmination of this Ph.D. journey would not have been achievable without the support of many wonderful people. Although it may seem like one person’s effort, pursuing this path succeeds with the effort and support of remarkable individuals. Thus, it is with great pleasure that I take this opportunity to sincerely thank those who have helped me during this journey, and I dedicate this work to them.

Words cannot express how grateful I am to my parents, Magbol and Saima. Your unwavering support and belief in me through my entire life have helped me get to this day. Your prayers for me have sustained me.

I am deeply grateful to my wife, Amani, for your sacrifices and support throughout my lifelong educational pursuits, and especially over the last five years of this process. With all of my responsibilities and stress during this time, I would not have been able to accomplish this goal without your love, encouragement, assuming my family responsibilities, and your willingness to give me the time to devote to this endeavor.

This work is dedicated especially to my daughters: Nerceyan, Taleen, Noran, and Tasneem. You are my inspiration to achieve greatness. You have made me stronger, better, and more fulfilled than I could have ever imagined. Having you by my side made these years sweet and memorable.

My sincere gratitude extends to my siblings: Afnan, Amani, Wail, Tahani, and Suhayb. Without your tremendous encouragement during the past few years, it would not have been possible for me to complete my studies.

In addition, I would like to thank all the members and alumni of the Complex Systems Institute (CSI) at UNC Charlotte: Drs. Malak, Maryam, Gizem, Riyi, Robert, and Elizabeth. It is your kind help and support that have made my study and life in the U.S. a wonderful time.

Pursuing my education abroad has introduced me to the brightest friends: Khaldoon, Marran, Abdul Majeed, Sultan(s), Ali(s), Abdul Rahman(s), Hamad, Husain, Nawaf, Awadh, Faisal, and Naser. We shared our time through happiness, sadness, joy, and stress together. The rivalry between us has pushed me to succeed. I wish you all the best in your careers and future lives.

Despite our intermittent communication overseas, my appreciation also goes out to my relatives, family-in-law, friends, and colleagues. Your thoughts, courage, strength, kindness, and love inspire me in every step of my life and motivate me to overcome any obstacle.

Finally, I want to devote this work to the memory of my uncle, Ahmed, who I lost in 2019. His shared memories and experience about life in the U.S. with me encouraged me to pursue my education abroad. I still remember the moment when he promised to attend my graduation ceremony. Unfortunately, that didn’t work out quite the way it was planned. Accomplishing this goal while missing out on time with him makes this a bittersweet occasion.

Table of Contents

List of Figures

List of Tables

LIST OF ABBREVIATIONS

SCOLR	The state courts of last resort
SQL	Structured Query Language
NLP	Natural Language Processing
ML	Machine Learning

Chapter 1: Introduction

The complexity of the various systems in our life led scientists from different disciplines to collaborate to solve the challenges in those systems. The study of complex systems has become a popular research area in this decade. According to Stephen Hawking, in his “millennium” interview on January 23, 2000 (San Jose Mercury News), “I think the next century will be the century of complexity.” In general, the study of complex systems aims to extract hidden information and associations between objects in real-world systems and to predict the future of those systems.

However, the study of complex systems in some organizations suffers from a lack of analyzable data to understand their complexity. Therefore, these organizations cannot take advantage of automated tools to analyze their system behaviors, support their decision making, and predict the future changes in their business. Moreover, scholars and business-related domain experts still know relatively little about these organizations and their interaction with others.

The state courts of last resort (SCOLR) are examples of organizations with this significant issue. They have unstructured and noisy textual documents that are hard for machines to read and use.

The legal system in SCOLR is an interesting real-world system. Despite the importance of these courts and their decisions, there are many interactions between the system elements that are interdependent and unpredictable, which make it a complex system. The legal definition of a court of last resort, typically called a supreme court or a state supreme court, is as follows:

> It is the highest court in its jurisdiction. It decides the most important issues of constitutional and statutory law and is intended to provide legal clarity and consistency for the lower appellate and trial courts. Because it is the court of last resort, a supreme court's

> decisions also produce finality. In addition, a supreme court oversees the administration of the jurisdiction's judicial system. (Legal Dictionary, n.d., par.1)

The SCOLR have made numerous important decisions that have attracted both praise and hostility from federal courts, national politicians, and voters. Therefore, the components of these court systems and their behavior have become of interest to the nation, legal experts, politicians, and voters.

In the remaining part of this chapter, I provide background and context of the research's application, the SCOLR. Then, I discuss the importance of citation in the legal system, and I describe citation treatment and its effect on the legal system. I conclude the chapter by introducing the research problem statement, the hypotheses and validation approaches, and the research methodology.

Chapter 2 provides background and related work on the United States legal structure as a complex adaptive system. Chapter 3 discusses the data collection methods and results. Chapter 4 introduces the network modeling and analysis for the citation network in the supreme courts. Chapter 5 explains the methods of classifying the citation treatment in state supreme courts' cases. Chapter 6 concludes with a discussion, conclusions, and recommendations for future work.

Background and Context of The Application (The Courts of Last Resort)

Judicial independence is one of the most basic and controversial features of the American legal system. As Alexander Hamilton explained, "The complete independence of the courts of justice is peculiarly essential in a limited Constitution" (Hamilton, 1788). Basic precepts of fairness, impartiality, and liberty require that courts resist external political pressures and, instead, base their decisions on neutral principles of law and precedent (Wechsler, 1959). Yet, judicial independence also raises serious normative difficulties. As Alexander Bickel argued, "coherent, stable—and morally supportable—government is possible only on the basis of consent, and… the secret of consent is the sense of common venture fostered by institutions that reflect and represent us and that we can call to account" (Bickel, 1962, pg. 20). Accordingly, policymaking by unaccountable judges may be morally unsupportable.

Despite its normative importance, the empirical understanding of judicial independence remains hotly contested. Hamilton argued that "the permanent tenure of judicial offices" would contribute to "independent spirit in the judges"; but he also predicted that judges might bend to public pressure: "It would require an uncommon portion of fortitude in the judges" to resist "…the major voice of the community" (Hamilton, 1788). Numerous empirical studies validate Hamilton's prediction. Despite the insulating effect of life tenure, judges tend to make decisions in line with popular preferences and rarely stray from the dominant political coalition (Dahl, 1956; Friedman, 2009; McCloskey, 2010; McGuire & Stimson, 2004). A wide range of political, social, and institutional factors may drive the congruence between the judges' decisions and their political environments, but the mechanisms driving this association remain unclear.

The SCOLR in the American states have become increasingly salient in contemporary political and policy debates. These courts have made numerous important decisions that have attracted praise and hostility from federal courts, national politicians, and voters. Consequently, the membership, selection mechanisms, and organization of state supreme courts[1] have become topics of increasing interest and controversy across the nation, as politicians, legal experts, and voters debate the proper role and design of these influential institutions.

[1] Names of courts of last resort vary across the states. I use "state supreme courts" generally to identify state courts of last resort. Our database only contains cases appealed to the final appellate court within a state. In some states, these are called Supreme Courts, while in others they are called Courts of Appeal, Supreme Judicial Courts, or Courts of Criminal Appeals.

Yet, despite the increasing importance of these questions, scholars still know relatively little about state courts, especially as they interact with other institutions and influences. Arguments regarding particular judicial candidates, judicial policies, judicial reforms, and the proper role of courts frequently rely on untested empirical assumptions. This debate would undoubtedly benefit from a deeper understanding of the complex relationships between state political institutions, judicial decision making, selection mechanisms, organization, impact, attitudes, and identity within the broader social and political context. The study of the interaction of these state supreme courts with legislative and executive branches of government at the local, state, and federal levels is limited to a handful of case studies typically focusing on a single court case or ruling. This dearth of attention to state courts is largely hampered by limitations in data availability. Despite the importance of these institutions, scholars lack the types of data for state courts that we take for granted in the study of federal-level institutions.

In this dissertation, I seek to redress this deficiency by designing a framework to overcome the shortage of data and system analysis in the study of the judiciary cases of SCOLR. I am particularly focusing on the judiciary cases of SCOLR as a real-world complex system, but my framework can be adapted to other problems with similar data issues.

The Importance of Citation in the Legal System

In the legal system of the United States, judges make decisions regarding a case based on the evidence and decisions made in prior cases with similar issues—called precedent. The judge's decision is called an opinion. The judge in the current case uses the opinions in prior cases through the citation system: the judge in the current case identifies the opinions in prior cases, together with an explanation regarding the similarity between the current case and the cited cases, and the judge's decision in the current case is based on the opinions in the cited cases.

According to Cross, Spriggs, Johnson, and Wahlbeck (2010), the practice of citing prior cases is not an optional practice for judges but a rule in U.S. supreme courts. A judge's opinion cites prior cases as precedents. In fact, "justices place their holding in the existing body of the law by demonstrating that prior decisions directed their opinion" (Cross, Spriggs, Johnson, & Wahlbeck, 2010). According to (Macey, 1989), the precedents serving as citations "may be viewed as the principal asset of a judicial system," and the higher their quality, "the better the judicial system may be said to be." Another study done by (Segal & Spaeth, 1996) added that an "appeal to precedent is the primary justification justices provide for the decisions they reach." In addition, the authority of precedent "is generally thought to be one of the most important institutional characteristics of judicial decision making." (Young, 2002). Cross, Spriggs, Johnson, and Wahlbeck (2010) also reported that "a recent study, for example, shows that variation in the authority of precedent influences the way in which the Court chooses to legally treat those cases; even after controlling for the ideological position of the Court and other factors related to the citation of precedent", and their research found that "the Court is more likely to follow a precedent if it has greater legal authority".

Citation Treatment and Its Effect on The Legal System

Citation behavior plays an important role in the permanence and revision of the law. Judges do not always agree with the outcomes of previous cases related to their cases. A judge may adopt the position of the law from other cases by citing those cases, thus, showing his/her acceptance and agreement with the interpretation, or he/she may reject that interpretation and come to a different one.

The citation treatment affects the extension and the history of the law regarding the case. For example, if a certain case receives negative treatment (such as being overruled) in subsequent cases, it is an indication that the law might no longer be applicable. Therefore, it is important for the judges and the lawyers to keep track of the history of laws and any changes in particular laws. Citation treatment assists judges and everyone who works with them (such as attorneys) to study cases carefully.

Only a few studies have been done on the automatic recognition of case treatments. Lexis Advance has a shepardization feature that allows the tracking of the history of a case. However, Lexis Advance shepardizes only U.S. Supreme Court cases, not state supreme court cases. Moreover, this service from Lexis Advance is behind a paywall. Similar work has been done to classify citation treatment for Australian court cases (Galgani, Compton, & Hoffmann, 2015) but not for U.S. state supreme court cases. History Assistance was also introduced in Jackson, Al-Kofahi, Tyrrell, and Vachher (2003). However, this classifier was not successful when phrases such as “declined to follow” and “superseded by statute” were used. Conrad and Dabney (2001) used a model-based filtering application and evaluated their performance through having five professional editors annotate their data. Their project took approximately four years to complete. The dataset used in their analysis included approximately 31,000 cases. Conrad and Dabney (2001) achieved a recall of approximately 80% for negative treatment and a precision of approximately 50%. However, they only considered indirect history; our research focuses on direct history.

> There are two chief kinds of history. Direct history involves cases in the same appellate chain as the current case. Thus, the instant case may be part of the direct history of an earlier case via an appeal from an earlier decision. Indirect history involves cases in other

appellate chains, which usually appear because they are cited by the judge or counsel as logical precedents, with which the judge will often agree (positive history), but sometimes will disagree (negative history) (Conrad and Dabney, 2001).

To overcome this shortage of research on the citation treatment, there is a need for an automated tool to extract and classify the citation treatment of state supreme court cases.

Research Problem Statement

These are the goals of this dissertation: review the efforts to collect data and analyze the system of SCOLR; create a novel dataset for the system of SCOLR; and model and analyze the system of SCOLR as a real-world complex system.

These are the four steps planned to meet the goals: 1) Overcome the shortage of data of the system of SCOLR by constructing an automated tool to read and parse the cases of the system to produce a dataset. 2) Design an automated tool to extract case citations and classify the citation treatments. 3) Model the system as a complex system and apply the approach of network science to analyze it. 4) Compare the analysis results with the results of other approaches.

Hypotheses and Validation

As described in the previous sections, this research is at the intersection of complex systems, legal science, and political science, applied to the SCOLR system. These are the hypotheses of my research:

> ***Hypothesis 1:*** Automated methods of data collection are better than manual methods of data collection, as measured by the coverage of collected data and the time required to collect it.

Hypothesis 2: A case cites another case based on the similarity of the legal issues of both cases. Therefore, the cases network has clusters that represent each case's main legal issues.

Hypothesis 3: The emergence of the Internet and the availability of technology to the court increased the number of citations in the courts' opinions because judges could easily search across geographic regions.

Hypothesis 4: Adding the citation treatments to the citation network will affect the results of a network analysis.

In order to evaluate the research outcomes based on these hypotheses, the following reliability and performance approaches will be used. I conduct checks for reliability and validity by comparing my computer-generated dataset produced in the first study (described in Chapter 3) with existing datasets as well as our own hand-coding of state supreme court dockets. The algorithms and methods of network science and natural language processing (discussed in Chapters 4 and 5) are evaluated based on their performance.

Research Methodology

In this dissertation, I am investigating the potential of using methods of complex systems and data science to close the gap in the literature of the SCOLR system. I started my research by collecting source data and creating an automated tool to build a novel dataset for the state supreme court cases during the period 1953–2010. I verified the validity of my dataset by modeling and analyzing the system as a complex system. Since the system has network characteristics, I used the approach of network science to model the system based on the citation behavior. Then, I created an automated dictionary-based classifier model to extract and classify the citation treatments for the court cases, and I added the results to improve my analysis. Since the cases' opinions that include citations and treatments are in the form of text, I used concepts of natural language processing (NLP) to create the model.

Chapter 2: Background and Related Work on the United States Legal Structure as a Complex Adaptive System

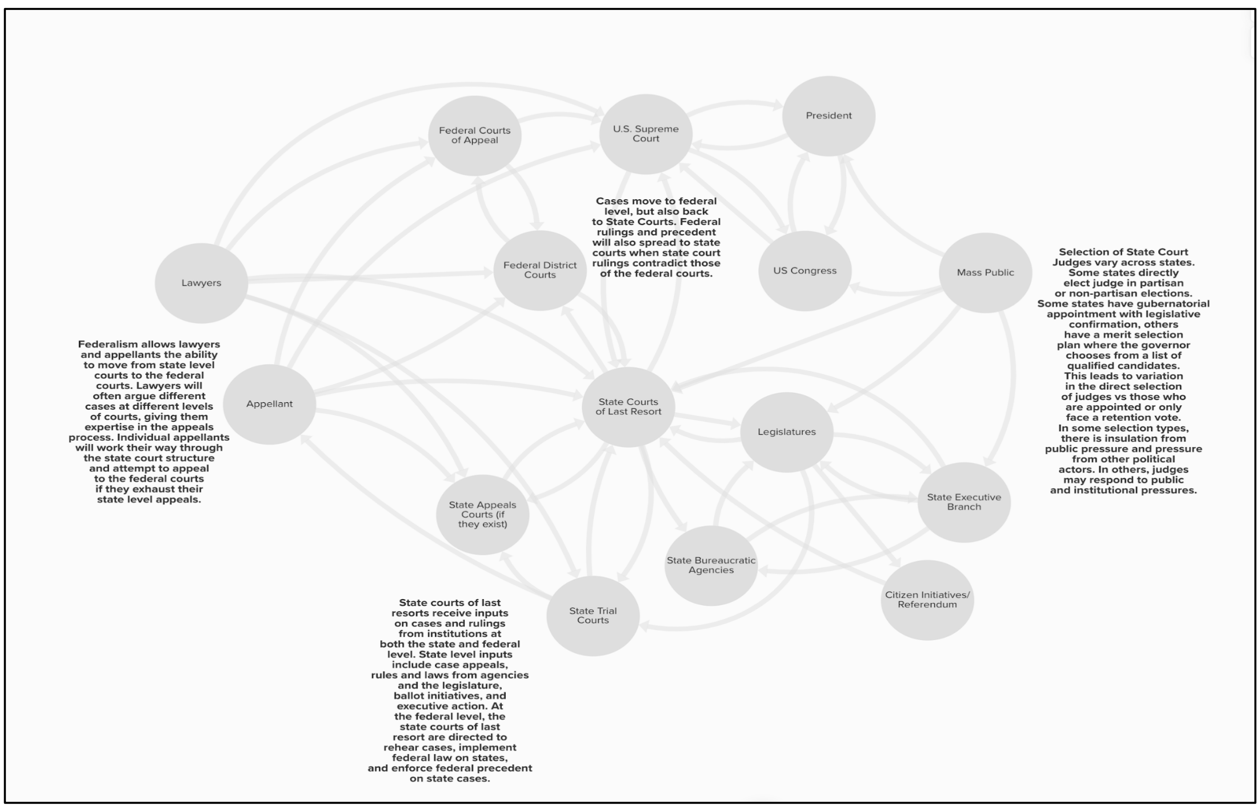
The bifurcation of the United States legal structure into overlapping and evolving units at the state and federal level is a cornerstone of federalism in the country. This court structure was intended to allow states to remain autonomous when deciding cases based on state law, while federal courts would be focused on federal law and conflicts between the states. An unintended consequence at the time of the nation's founding was a seemingly weak judiciary at the federal level, with little in terms of formal powers. As the country grew, however, new issues, jurisdiction, and types of legal challenges did so as well. The judicial authority at the state level was equally, if not more powerless, until the expansion of state capacity and influence following World War II, when state courts began a substantial increase in their activity.

As state courts have become increasingly salient in local, state, and even national political outcomes, scholars should adopt a complex systems approach to understanding the interactions of courts and other political units within and across states. As Ruhl, Katz, and Bommarito (2017) argue:

> legal systems exhibit what complexity scientists identify as hallmark elements of CAS. The diverse institutions (e.g. legislatures, agencies, and courts); norms (e.g. due process, equality, and fairness); actors (e.g. legislators, bureaucrats, and judges); and instruments (e.g. regulations, injunctions, and taxes); are interconnected through stochastic process (e.g. trials, negotiations, and rulemakings) with feedback mechanisms (e.g. appeals to higher courts and judicial review of legislation. (1377)

Figure 1 is a small example of the complexity involving actors and institutions interacting with SCOLR. We have classified four major areas of ***external*** actors and institutions that directly influence the inputs, outputs, and behavior of the state courts. First, federal institutions can overturn individual case rulings if a state case is deemed to violate federal law or if legislation, executive orders, or other rules are made where federal supremacy is established. In some instances, Congress will write legislation in response to a state court case; for example, the 1996 Defense of Marriage Act was a Congressional response to the Hawaii Supreme Courts' ruling allowing same-sex marriage. Most of the influence on state courts comes from the federal judiciary. In fact, Hall (2014) argues that vertical cases, those that emerge from lower level trial courts, are more likely to be overturned in the Federal Supreme Court due to the institutional control and implementation power the federal courts have at local and state level courts. With this in mind, judges at the state level will be mindful of higher rates of their decisions being overturned by higher level courts. This type of action is seen as the power of the federal judiciary challenging the institutional stability of the state court.

Figure 1: Sample Complex System of state and federal legal and political institutions



The mass public influences court decision making through direct selection, retention, and replacement of state judges. Numerous studies argue that elections prompt state judges to follow popular preferences. Most of this literature focuses on the manner in which judges reach the bench and retain their jobs. The American states vary widely in judicial selection mechanisms. Seven states elect justices through partisan elections, 15 hold nonpartisan elections, and 16 appoint members initially, then hold retention elections to remain on the court. Finally, 12 states use gubernatorial or legislative appointment, with many of these states requiring reappointment after varying term lengths or a mandatory retirement (American Judicature Society, 2013).[2] These selection differences influence the choices judges make (P. Brace, Hall, & Langer, 2001); (P. Brace & Hall, 1990, 1993). Elected trial court judges ((Gordon, 2007); see also (Huber & Gordon, 2004)) and judges selected through appointment are less influenced by the public preferences than other SCOLR justices (Johnson, 2017).

Other studies emphasize electoral competition; i.e., state judges may respond to public opinion only when they rationally anticipate future competition. These studies test the effects of competition in different ways. For example, judges tend to follow popular preferences when they are near the end of their terms (P. Brace & Hall, 1995; Caldarone, Canes-Wrone, & Clark, 2009) and when general partisan competition in the state is high (P. R. Brace & Hall, 1997; M. G. Hall, 1995). These findings suggest that judges respond to public opinion because they anticipate future competition; however, they do not capture rational expectations at the individual level. State-wide

---

[2] Reappointment mechanisms vary across these 12 states. Massachusetts and New Hampshire allow judges to serve until they turn 70 and then must retire. Connecticut, Delaware, Maine, New Jersey, New York, and Rhode Island have gubernatorial reappointment processes that include a confirmation vote by one or both legislative chambers. Vermont calls for a vote of the general assembly to reappoint a judge. Hawaii requires reappointment by a commission. Finally, Virginia and South Carolina hold legislative elections for both initial appointments and retention.

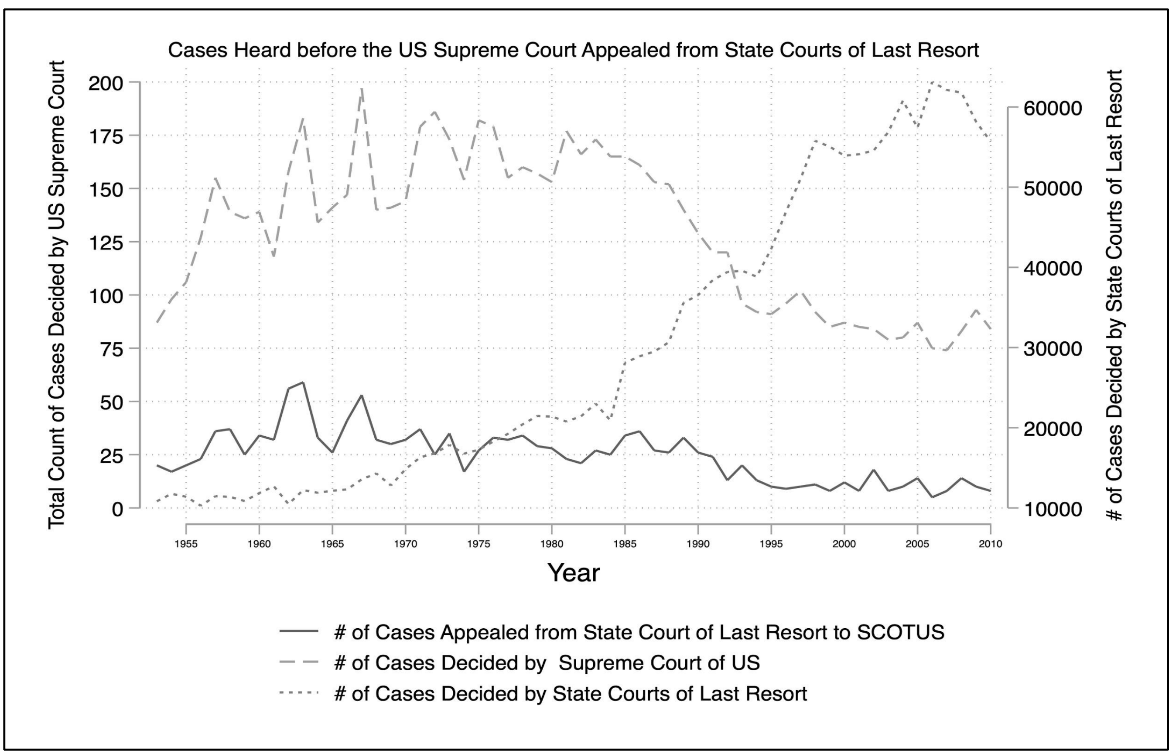
competition and the proximity of a judge's reelection contest are only indirectly related to that judge's rational expectations regarding future competition.

This rich, extensive literature offers valuable insights into policy responsiveness; yet, this groundbreaking work has been hampered by a lack of available data. For example, some studies examine judicial behavior over only a few years (e.g., Brace & Boyea, 2008; Cann & Wilhelm, 2011). Because these studies essentially employ cross-sectional rather than longitudinal data, they are unable to test whether judges actually change their voting behavior in response to shifts in public opinion over time. Other studies include only a small number of states (P. R. Brace & Hall, 1997; M. G. Hall, 1987, 1995) or only states with certain selection mechanisms (Caldarone, Canes-Wrone, & Clark, 2009). Finally, many studies focus exclusively on rare case types---namely abortion (P. Brace, Hall, & Langer, 2001) and capital punishment (P. Brace & Boyea, 2008; P. Brace & Hall, 1995; P. R. Brace & Hall, 1997; Canes-Wrone, Clark, & Kelly, 2014). As an example, the analysis of capital punishment decisions by Brace and Boyea (2008) uses only 889 of the 15,000 cases in the State Supreme Court Data Project (SSCDP). More recently, Cann and Wilhelm (2011) used the SSCDP to examine judicial responsiveness across a wide range of issue areas. They find that only judges facing contestable elections respond to the public and then only in “highly visible”' cases, which are just 1.3% of the cases in their data.

The mass public and federal institutions are only two of the external pressures that may exert influence over SCOLR. A more direct relationship often comes from other institutions within the state. Much like inter-branch interactions at the federal level, state institutions can be conflictual or retaliate against the actions of the other institution. For example, state legislatures may pass court-curbing legislation aimed at removing powers from a state court due to ideological differences (Blackley, 2019). Likewise, SCOLR may invalidate or rule state laws, administrative rules/orders, and executive actions as unconstitutional. Even governors can retaliate against state courts in budget requests and in extreme cases, remove or fail to re-appoint judges to office.

An even more crucial element of inputs in this system are cases from trial courts and intermediate appellate courts within the states. In 2017, state courts reported nearly 83.5 million cases (National Court Statistics), down 22% from 10 years prior. Of these cases, 207,321 were cases appealed to state intermediate appellate or SCOLR. This is not an insignificant number of cases when placed in the context of the Supreme Court of the United States. As Figure 2 shows, the number of cases decided by SCOLR has steadily increased from the 1950s, while the number of cases decided by the U.S. Supreme Court has steadily declined. Meanwhile, roughly 20% of all cases heard by the U.S. Supreme Court were previously argued in SCOLR.

Figure 2: Caseload Information on U.S. Supreme Court and State Courts of Last Resort Appeals



Understanding the relationship between these external pressures, institutional structures, and judicial decision making has important consequences for the nature of representative democracy, legal theory, and judicial politics. By exploiting the institutional and political variation in state supreme courts, we will allow scholars the ability to directly test the role of different individual actors, selection effects, public pressure, and interbranch relations in shaping elite policy decisions on courts. The lack of data currently available to scholars is also causing potential normative implications when it comes to questions of accessibility to courts, trends in areas of increased attention by law enforcement, as well as general inefficiencies in court output.

Chapter 3: Automated Data Extraction to Evaluate Courts of Last Resort in the American States

As I mentioned before, scholars suffer from the lack of available data from the state supreme courts. As a result, unlike for other judicial institutions, there are not many studies on the state supreme courts. On the one hand, this shortcoming has led some scholars to rely on untested assumptions in their studies and debate these courts and their behaviors. On the other hand, some scholars chose to focus on data for only a few years (e.g., Brace & Boyea, 2008; Cann & Wilhelm, 2011). Other studies include only a small number of states (P. R. Brace & Hall, 1997; M. G. Hall, 1987, 1992, 1995) or only states with certain selection mechanisms (Caldarone et al., 2009). Moreover, other studies focus exclusively on rare case types – for instance, abortion (P. Brace et al., 2001) and capital punishment (P. Brace & Boyea, 2008; P. Brace & Hall, 1995; P. R. Brace & Hall, 1997; Canes-Wrone et al., 2014).

In this chapter, with my research team, we seek to redress this deficiency by presenting a database of state supreme court cases and outcomes that we will make publicly available to the scholarly community. We use automated textual analysis through the Python programming language and the Structured Query Language (SQL) to quickly and reliably collect data on opinions and decisions for all 52 SCOLR.[3] Our computer program extracts critical pieces of information from text files of state supreme court opinions and converts the information into quantitative data. Our data-collection method produces reliable measures of SCOLR decision making and facilitates a wide range of empirical analyses that are currently impossible using

[3] Oklahoma and Texas have separate courts for criminal appeals.

existing data[4]. Researchers will be able to explore various influences on judicial decision making and the impact of state supreme court rulings through time-series cross-sectional analysis. Our data will also inform important practical debates about judicial selection, judicial organization, interbranch relations, federalism, and how the increasing diversity in the U.S. is reflected in the courts.

The main result of this study is a novel dataset of all cases of the SCOLR from 1953 to 2010. The dataset contains valuable data regarding the cases and judges of SCOLR. It contains more than 1,867,500 cases and more than 2,100 judges who worked on those cases in 52 state supreme courts. It has 32 variables that characterize the cases, the judges, and the related entities including courts, states, and regions. To the best of our knowledge, this is the first dataset that covers this period of time. This study was published in 2020 in the Journal on Policy and Complex Systems (JPCS) (https://doi.org/10.18278/JPCS.6.2.5.).

This chapter proceeds as follows. First, we describe the data-collection methods. Next, we introduce the dataset preprocessing. Then, we discuss the study results and conduct reliability and validity checks by comparing my computer-generated data with existing datasets as well as our own hand-coding of state supreme court dockets.

[4] For a full review of these data collection efforts, see (M. E. K. Hall & Windett, 2013)

Data-collection methods

Legal source documents

The source document collection is one of the critical phases of our project. Data is only useful if it is high quality, systematically collected, and free of bias. Bad data is at best inconsequential, but flawed source data could cost time in cleaning errors and ensuring the accuracy of information. Even though we have 52 state courts in the United States and millions of cases that were challenged, we ensured that we use a quality data source in terms of accuracy, completeness, relevance, validity, timeliness, and consistency. One technique to collect all the case details is to directly get information from the 52 state courts. However, these courts each format their documents differently. To compile a comprehensive tool, we would need to write a state-year specific program to extract our data. This is inefficient, and it increases the likelihood of non-random error; it also relies on documents from the states that are often not machine readable.

To overcome these issues and maintain the metrics mentioned above, we compile our source documents from LexisNexis Academic. This tool is a knowledge center that maintains the deep archive of federal and state court cases dating back to 1789. We have sourced our data from this central repository and organized them in a hierarchical, alphabetical, and chronological order. The level of hierarchy is as follows: 1. State 2. Year 3. Cases heard in the court. To easily traverse in the dockets and locate a case, we can make use of the unique identifiers Lexis assigns to each individual case. Finally, we have the outcomes of the SCOLR in 50 states, for 58 years and over 1.7 million appeals. Of this figure, 584,253 were full written opinions released by the court. The remaining appeals are orders, denials of a varied nature (certiorari, dismissal, habeas corpus, etc.), legal suspensions and disbarments, and other obscure cases that sometimes originate in the states.

Judges' Biographical Information

A major hurdle of this project is ensuring the accuracy of the judges involved in each individual case. For the United States Supreme Court, it is easy to identify which judges serve on which cases, as there are only nine judges and a relatively small number of cases. Given the national level office, each judge who has served on the U.S. Supreme Court can be identified with exact dates of service. SCOLR are not as simple, nor is the data readily available. To begin, we need to identify the population of judges that appear in these judicial decisions. We then cross-validate each name with state or regional reporters, state Bluebook biographical information, information obtained directly from the court, or information from the state archives or law library.

To build the knowledge base of judges, we extract all information that appears in a Lexis section titled "Judges." The "Judges" section is not consistent within the state, or across states. We identify all of the surnames, aliases, and combinations of first and last names when available, as well as the first and last date the name appears in each state. After confirming with source documents from the states and other resources, we discovered 2,101 unique judges serving between 1953 and 2010. We then combined these confirmed rosters with the results of our scraper. We matched all instances of last names, combinations of names, and even misspellings, typos, and combinations of multiple names. This informed us of 5,699 different aliases used by sitting judges. We may now match these names with the text of names in different fields as outlined below.

In Figure 3, we show three cases and how Alabama listed the judges that appear in each case. As evident from this figure, some cases will report very limited information on which judges appear in each case, while other cases offer a very detailed description of not only the judges on the case, but also how they decide each case. This type of variation makes background data collection on the judges a priority, to ensure the highest level of accuracy with our data processing.

Figure 3: Sample Judge section formatting for three cases in the same state, released on the same day

1953January19
262 Ala. 197; 77 So. 2d 903; 1953 Ala. LEXIS 18
Foster, Justice. All the Justices concur.

1953January19
258 Ala. 317; 62 So. 2d 918; 1953 Ala. LEXIS 234
FOSTER, Justice. LIVINGSTON, C. J., and SIMPSON and GOODWYN, JJ., concur.

1953January19
258 Ala. 303; 63 So. 2d 796; 1953 Ala. LEXIS 232
LAWSON, Justice. LIVINGSTON, C. J., and BROWN and GOODWYN, JJ., concur. FOSTER, J., concurs in the reversal of the judgment of the trial court. SIMPSON and STAKELY, JJ., adhere to the views expressed in the opinion prepared on original submission and therefore dissent.

Courts, States, and Regional Reporter information

All the information of the courts, states, and regional reporter volumes was collected from *The Bluebook: A Uniform System of Citation*, guide version 20. This information includes the names, abbreviations, volume numbers, and dates. This information was collected to help in processing and understanding the dynamic of the cases information.

Dataset preprocessing

The raw format of our source documents is the written description of a case, the actors involved, the outcome, and the written opinion(s). These documents are extremely useful for legal researchers, judges, law scholars, and lawyers looking for information on a single case. However, social scientists, data scientists, and those interested in studying complex systems would need a dataset that has separated critical fields of these data for comparison, statistical analysis, and other empirical evaluations. We discuss below the pertinent information extracted for each case.

Parsing the dockets

To create this database, all the cases in the collected dockets have been parsed and preprocessed to ensure high quality results. Two types of files have been produced after the parsing and preprocessing were complete. The first type of files is the individual Majority, Dissenting, and Concurring opinions for each case, if these opinions were written. The second type of files is comma-separated values (CSV) files where we stored other variables of pertinent information from the cases.

The Majority, Dissent, and Concurring Opinions. These opinions are the official ruling decided by the court, along with any explanation of a disagreeing position. The Majority opinion is the position held by half or more of the court. These written opinions are always reported if the case has any sections labeled as “Opinion”, “Dissent”, and “Concur”. All the written opinions are extracted and stored in separate text files. We produced up to three text files for each case. Each file has only one of the case’s opinions. These texts are given a unique name including the year, state abbreviation, the unique Lexis identifier, and the initial of the text type (M, D, or C). Using the Lexis identifier allows us to connect the text files with the CSV spreadsheets. Other identifiers are available, as discussed below, but they are not always unique to a single case.

After creating the text files of the majority, dissent, and concur texts of each case, we created 52 CSV files, one for each court. The CSV file of each court has eighteen variables describing the cases of the court. All of these variables were extracted from the text. We describe and discuss those eighteen variables in the following paragraphs.

The case number. This number indicates the internal case identifier used within each state. The standard manner for citing cases as advised by The Bluebook is to use the regional reporter.

The case identifiers. The citations of the case are the first variables to be extracted from the case's texts. These citations are the case identifiers in the reporters. In the documents collected from LexisNexis, each case should have between one and three citations - State Reporter citation, Regional Reporter citation, and Lexis citation. The Lexis citation is produced for all appealed cases if the state releases information on those cases. Prior to the widespread use of computers for storing large amounts of data, courts would only publish information pertinent to appeals that were accompanied by a written appeal. The other cases that were dismissed, issued orders, decided without an opinion, or sent to another court were not reported on. As computing capacity increased, all of these different cases would become more readily available. States would send their published opinions to be printed in the state and regional reporters, but other case types could be archived online. While the *Lexis citation* is unique for each case, the *Regional Reporter* and the *State Reporter* may not be unique variables. Multiple cases may share the same *reporter* if they were reported on the same page and volume of the correspondent reporter. In addition, some cases do not have *Regional Reporter* or *State Reporter* if they have not been published in the correspondent reporter. Table 1 shows the definitions, structures and examples of all these case identifiers.

We found three issues after extracting the reporter citations of the cases. The first issue is that some of the citations have typing errors including, but not limited to, extra spaces between the elements of the reporter, wrong capitalization, or missing letters. The second issue is that there is no specific order for which citations come first. The third issue is that there is no specific punctuation used as a separator between the reporters. To overcome these issues and extract the reporters, we used the regular expressions (regex) that can identify all the case reporters even if they have some typing errors. After that, we reconstructed those reporters in the correct formats.

Table 1: The Case's Identifiers

Case reporter	Definition	Structure	Example
State Reporter	The case identifier for the published cases in an official state reporter by the state	Three main elements: The number of the volume of the reporter The state abbreviation The number of the first page of the case in the volume	414 Ill. 120
Regional Reporter	The case identifier for the reported cases in the regional reporter	Three main elements: The number of the volume of the reporter The region abbreviation The number of the first page of the case in the volume	110 N.E. 2d 256
Lexis citation	The case identifier in LexisNexis reporter	Four main elements: The year of the case published date The court abbreviation The word 'LEXIS' The case number	1953 Ill. LEXIS 257

The court name. One of the main variables to be extracted from the case text is the court where the case was decided. However, the case text often does not include the name of the court due to typographical errors. Consequently, we used the court abbreviation in the *Lexis Citation* of the case. As you can see in Table 2, each court has a unique court abbreviation. Therefore, we used the court abbreviation to link the cases with their courts. For example, to get the court name for the case with Lexis number (1953 Ala. LEXIS 268), we extracted the court abbreviation (Ala.). Based on Table 2, this abbreviation (Ala.) belongs to The Supreme Court of Alabama.

Table 2: Reporter, Lexis, and State Information and Coverage

Reporter	Abbreviation and Volume	States Covered in the Reporter	Lexis Court Abbreviations
Atlantic Reporter	A., A.2d, A.3d	Connecticut, Delaware, Maine, Maryland, New Hampshire, New Jersey, Pennsylvania, & Rhode Island, Vermont,	Conn., Del., Md., Me., N.H., N.J., Pa., R.I., VT
North Eastern Reporter	N.E., N.E.2d	Illinois, Indiana, Massachusetts, New York & Ohio	Ill., Ind., Mass., N.Y., Ohio
North Western Reporter	N.W., N.W.2d	Iowa, Michigan, Minnesota, Nebraska, North Dakota, South Dakota & Wisconsin	Iowa Sup., Mich., Minn., N.D., Neb., S.D., Wisc.
Pacific Reporter	P., P.2d, P.3d	Alaska, Arizona, California, Colorado, Hawaii, Idaho, Kansas, Montana, Nevada, New Mexico, Oklahoma, Oregon, Utah, Washington & Wyoming	Alas., Ariz., Cal., Colo., Haw., Ida., Kan., Mont., N.M., Nev., Okla., Okla. Crom. App., Ore., Utah, Wash., Wyo.
South Eastern Reporter	S.E., S.E.2d	Georgia, North Carolina, South Carolina, Virginia & West Virginia	Ga., N.C., S.C., Va., W. Va.,
South Western Reporter	S.W., S.W.2d, S.W.3d	Arkansas, Kentucky, Missouri, Tennessee & Texas	Ark., Ky., Mo., Tenn., Tex., Tex. Crim. App.
Southern Reporter	So., So.2d, So.3d	Alabama, Florida, Louisiana & Mississippi	Ala., Fla., La., Miss.

The case date. Like most of the other variables, we could not find a standard format for the date of each case. There are eight different formats, and some cases do not report a date. Therefore, our tool has eight different regexes to capture and extract the case date if there is one. If the case does not have a date, the tool creates the date as the beginning of its published year based on the year in its Lexis reporter. Some cases have more than one date - the submission date and the decided date. The submission date is the date when the case was submitted to the court. The decided date refers to the case's decision and publishing date. To distinguish between those dates, the decided date usually is followed by one of four words - Decided, Released, Filed, or

Delivered. For those cases with a couple of dates, our tool extracts the decided date, which refers to the case's decision and publishing date.

Majority, dissenting and concurring opinion writers. These writers are the judges who write the opinion(s) of the case. The Majority writer is the person, persons, or court (per curium) who wrote the rationale behind the court's ruling in a case. In the source data, there are labeled sections to report the majority opinion, dissent, and concur writers, if the case has any. Those sections are Opinion by, Dissent by, and Concur by, consecutively. Even with the labeled sections, extracting and processing the judges' names from the corresponding sections faced three challenges: judges' names were associated with stop words; the names were separated by different punctuation in each case; and the names were shared by more than one judge. In the following paragraphs, we explain these challenges and how we overcame them.

The first challenge on extracting the judges' names was that the judges' names were sometimes associated with descriptive words and phrases that need to be cleaned first. We considered those words and phrases as stop words. To find and clean those stop words and keep the names only, we used dozens of regex commands with textual processing techniques.

The second challenge was that in most cases, the judges' names were separated using a semicolon or colon, but in some other cases, a comma was used to separate the names. Using the comma as a separator was a problem, since a comma was also used in other cases to separate the first name and last name of the judges. To overcome this challenge, we replaced all the colon separators by semicolons. After that, we check if the judge's name has more than one comma and no semicolon, then replace all the commas by semicolons. At the end, we split the names by the semicolons.

The third and most difficult hurdle was to map the extracted names from the opinion writers' sections to the judges themselves. Most of the judges have multiple aliases that were used in the source documents. What makes it worse is that we encounter more than one judge working on the same court and reported with the same alias--i.e. George Smith (1937-1955), Griffin Smith (1949-1987) William Smith (1958) and Lavenski Smith (1999-2000) all served on the Arkansas Supreme Court, with multiple instances of overlapping time. To overcome this challenge, we utilized the list of all the judges' names, their aliases, and their service times in the court as outlined above. We added the first initial or first and middle initial to distinguish between individuals with a common first initial. We also added "Sr" and "Jr" to surnames of judges who had parents serve on the court. We then matched names and aliases between the opinion-writer names and judge-section names. This added additional aliases to our master roster file to include even more iterations of judges' names. Most common in the opinion sections are to use a judge's first and last name, while the judge section (where we compiled the initial list of judges) would typically include a last name and first initial.

After dealing with all the challenges and issues of the judges' names, we generate a unique id, called a judge code, for each judge in the dataset. The generated judge codes consist of two parts: the Lexis court abbreviations (as shown in Table 2) of the court where the judge serves; and a unique number representing the order of the judges in the court based on their service start date. These judge codes were aimed to be used to represent the judge's names, to overcome the challenge of shared names.

Table 3: Additional Text Variables Extracted

Variable(s)	Section Head in Source Document	Definition
case_party_1_appellant	No label. It is in the first line of the text until "v."	Appellant or Appellee for each case. There is not distinction or ordering of which case appears first or second. We extract this information from the title and create an indicator variable for appellant status.
Case_party_2_appellant	No label. It is written after "v." in the first line of the text.	Same as above for the 2nd party in the case
procedural_posture	PROCEDURAL POSTURE:	Ground of appeal to SCOLR
overview	OVERVIEW:	A paragraph summary of the case history and facts
outcome	OUTCOME:	The Court's ruling on the appeal
counsel_1	COUNSEL: (the first line)	Name and law firm/agency for attorney for case_party_1
counsel_2	COUNSEL: (the second line)	Name and law firm/agency for attorney for case_party_2

The Disposition and the Disposition Description of the case. One of the most critical components of the court's ruling is how they treat lower level decisions. Most of the cases in our dataset are appeals from lower level trial courts or appellate courts. The disposition of the case can take numerous different positions depending on where the case is in the appeal process, the type of case, and the parties involved. These are examples of dispositions: overturning or upholding lower level decisions; the result of an attorney disciplinary hearing; validation of ballot initiatives; ordering of a new trial; new sentencing; and granting or denying petitions. Overall, we have 50 types of dispositions. Each case has a section for the disposition. As there is no uniform requirement or even a recommendation for disposition reporting, we began our process by simply scraping the disposition section and removing words that are not associated with the ruling. Overall, 235 unique phrases were classified into specific disposition categories.

Other variables have been extracted from the text report of the cases, since they were reported without any kind of preprocessing. Those variables are listed with their definitions in Table 3.

Results

The data collection and processing methods led to the creation of a novel dataset of the State Supreme Courts' Cases for the years 1953 to 2010. To the best of our knowledge, this is the first dataset that covers this period of time. The dataset contains valuable data about more than 1,867,500 cases with more than 2100 judges who have worked on those cases in 52 courts. The Supreme Court of California has the largest number of cases (258,175 cases), while the Supreme Court of Alaska has the smallest (7,512 cases). The average number of cases among all courts is 35,913 cases. Figure 4 shows a map of each court's case load geographically by groups of 10 in terms of caseload, while Figure 5 shows the total cases for each state.

Figure 4: Map of States with Caseloads

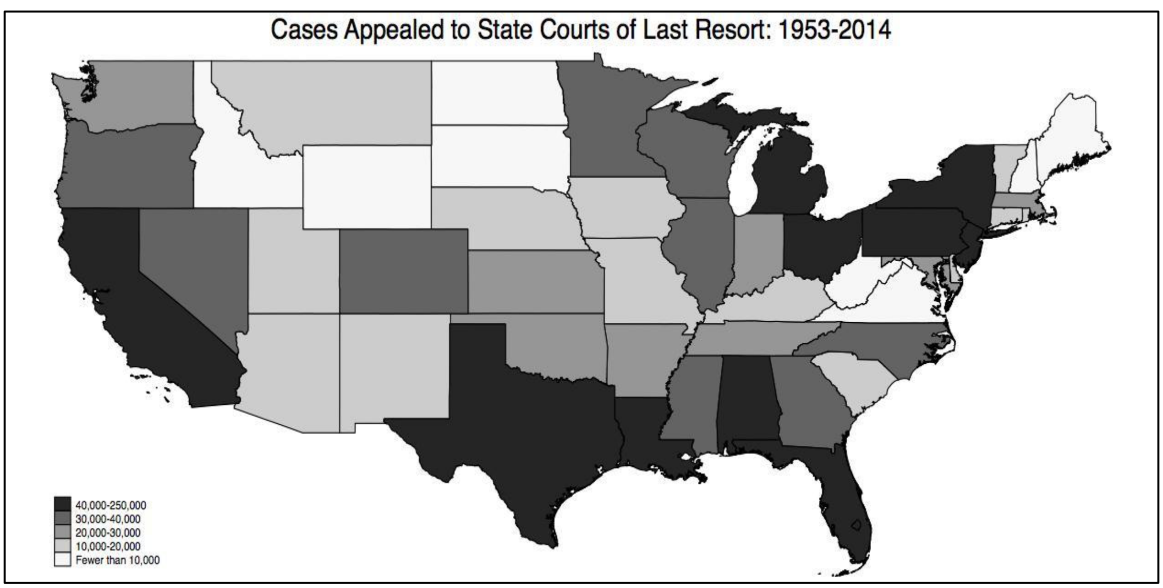


Figure 5: State by State case breakdown

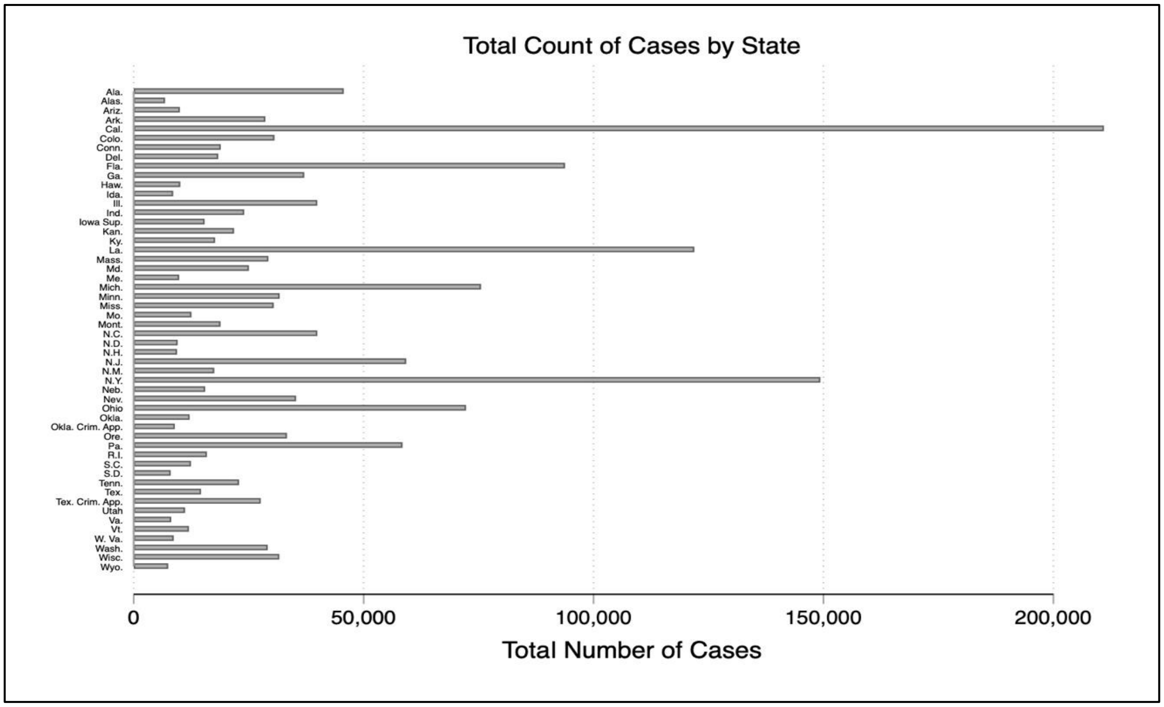
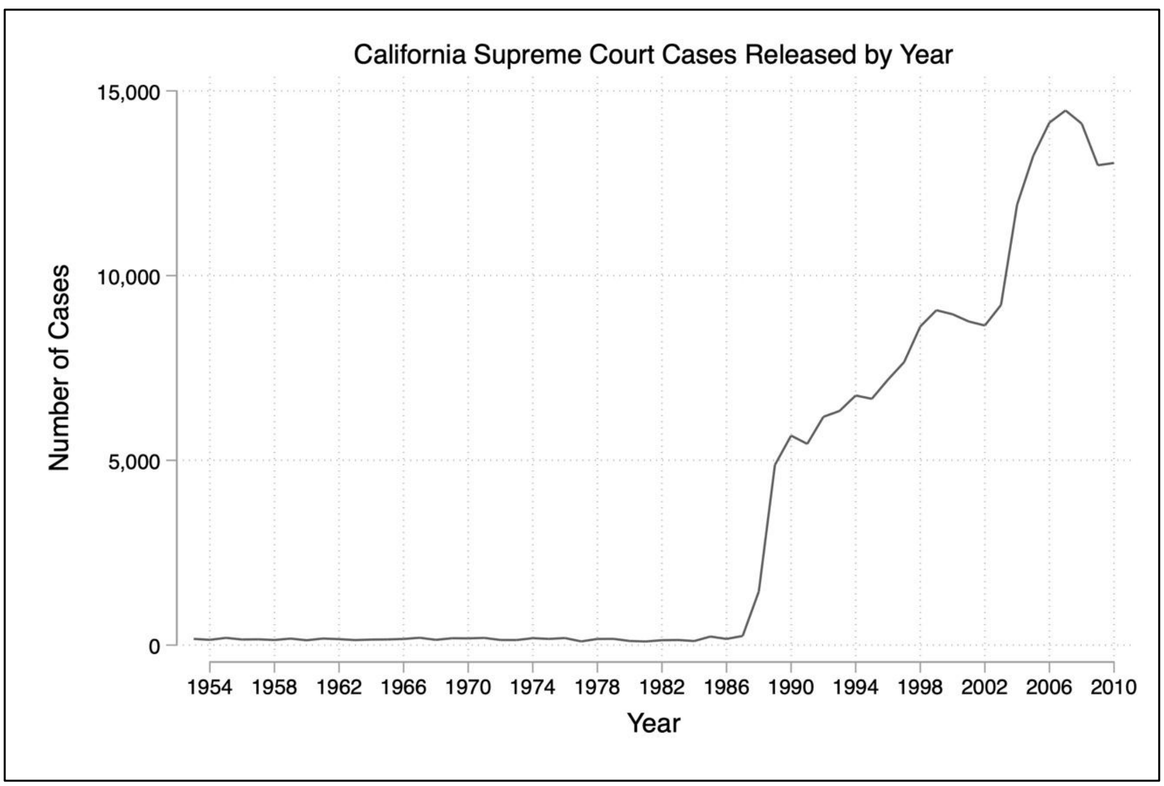


Figure 6: Increase in California Caseload by Year



Interestingly, we see changes over time in the number of cases released by state. In Figure 6, we show the annual number of cases released by the California Supreme Court. As evident in this graphic, in 1987, the court's caseload dramatically increased. This is not a function of cases increasing in a single year, but for one reason or another, California began to have the capacity to release all orders, certiorari denials, petition decisions, etc. This is a large amount of the work a court does, but these types of cases have often been overlooked in the collection of court data.

Beyond showing simple descriptive statistics of caseloads, Figure 8Figure 9 show examples of what happens to a state's dissent rate when there is an institutional change forced upon the court from the outside: a constitutional change, a ballot initiative, or legislation that changes the institution's rules. First, we show the impact of Tennessee implementing a merit selection plan in 1970, removing Supreme Court judges from retention elections beginning in 1974, and finally re-instituting retention elections in 1994. As evident in Figure 8, there was a slow but pronounced increase in cases with dissents immediately upon the removal of the retention requirement in the 1970s, and there was a small shift downward upon the 1994 modified Tennessee Plan's adoption.

Unlike Tennessee, when Florida instituted judicial reform in the 1970s, there was little change in judges' behavior. When the state moved from partisan elections to non-partisan elections in 1972, and to merit selection in 1977, the dissent rate did not see meaningful change. Most of the behavior on this court in terms of dissension occurred during the 1960s, prior to any reforms.

Not all reforms undertaken by the states were related to judicial selection. In decades prior to the 1970s, numerous midwestern states had "commissioners" who would write opinions to be voted on by the SCOLR, taking much of the opinion writing out of the hands of the judges. In Missouri, when the state removed commissioners from writing opinions in 1973, we see a dramatic

25 percent increase in cases with a dissent. It would appear as if requiring judges to write their own opinions resulted in increased disagreement within the state.

Figure 8: Tennessee's Multiple Selection Mechanism Changes

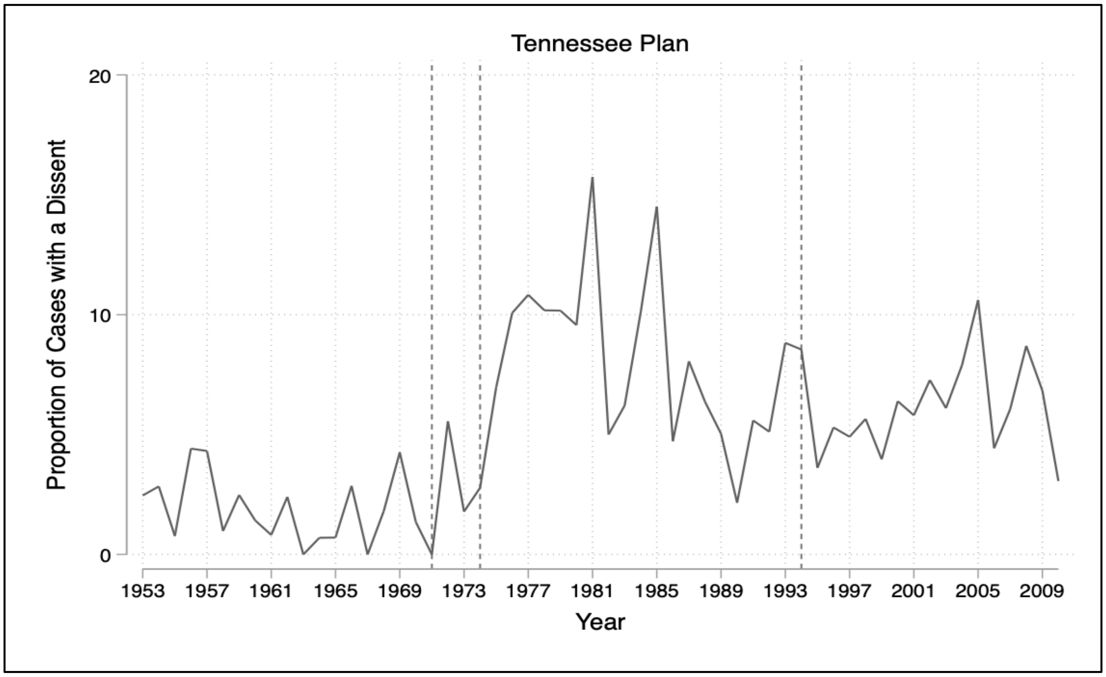


Figure 7: Florida's Multiple Selection Mechanism Change

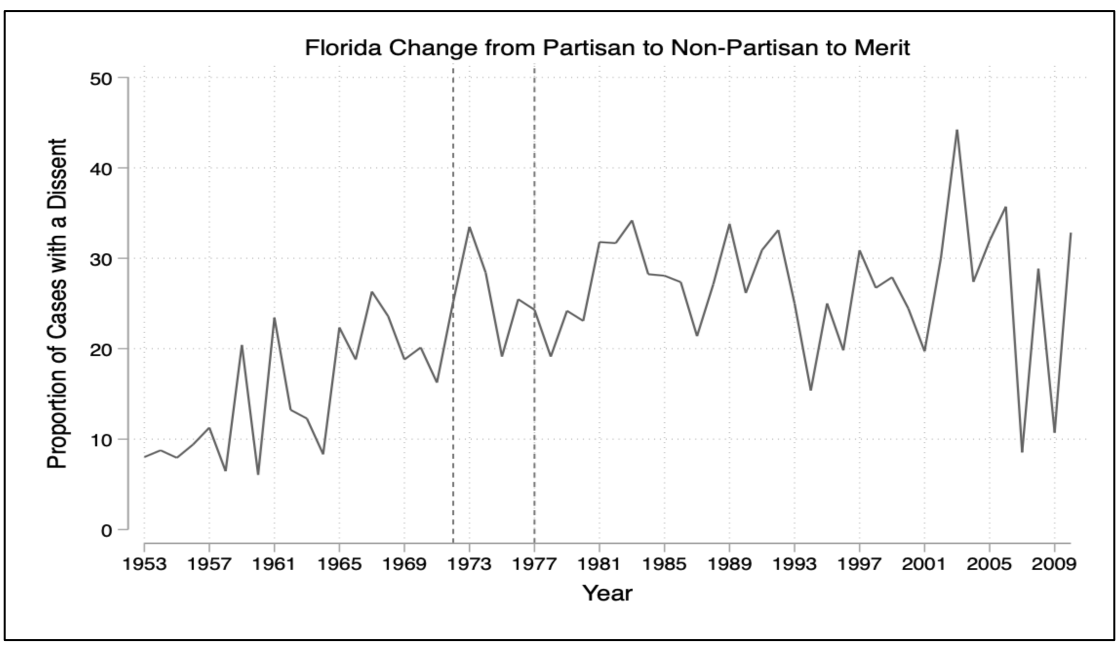
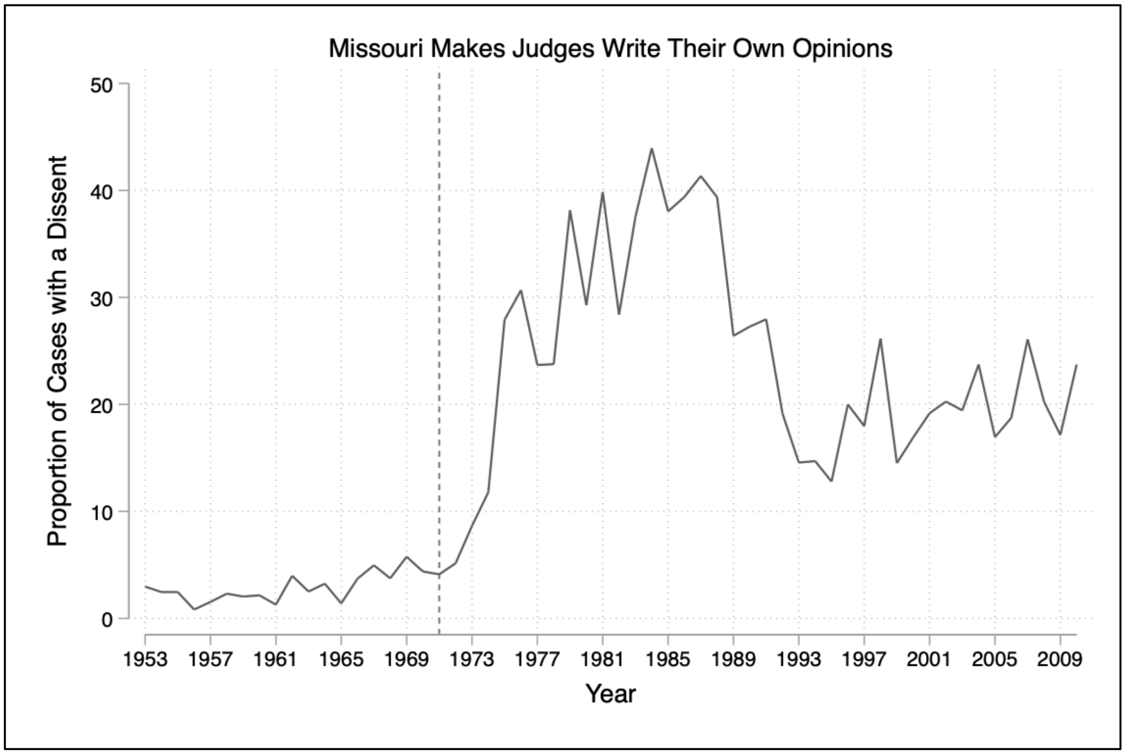
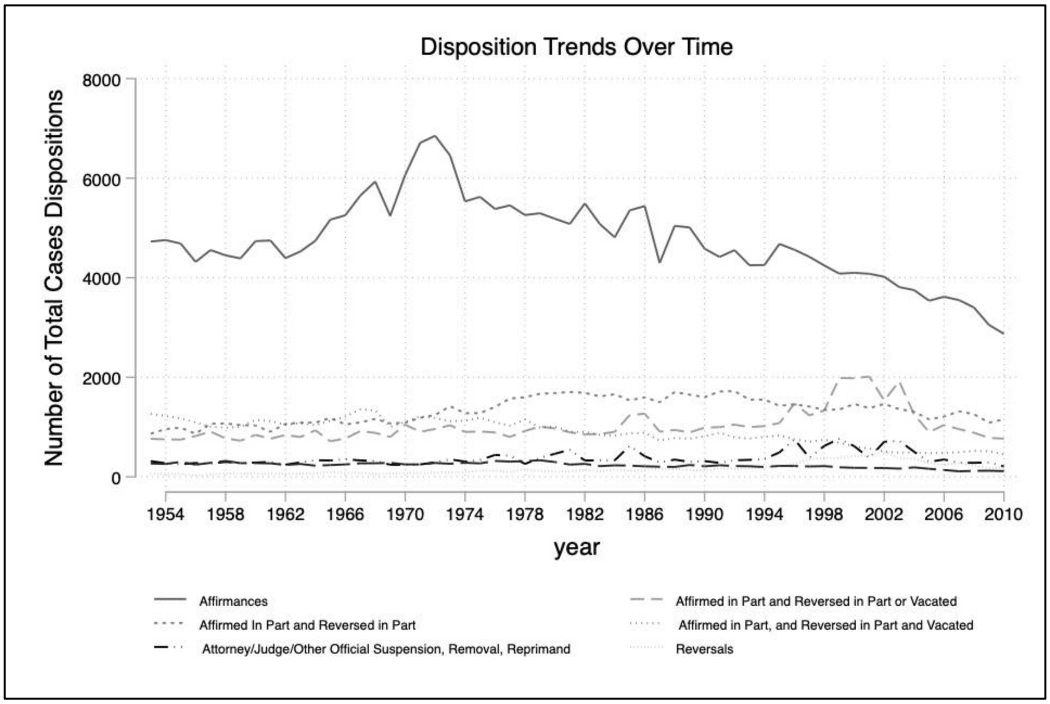


Figure 9: Missouri's Supreme Court Removes Commissioners from Opinion Writing



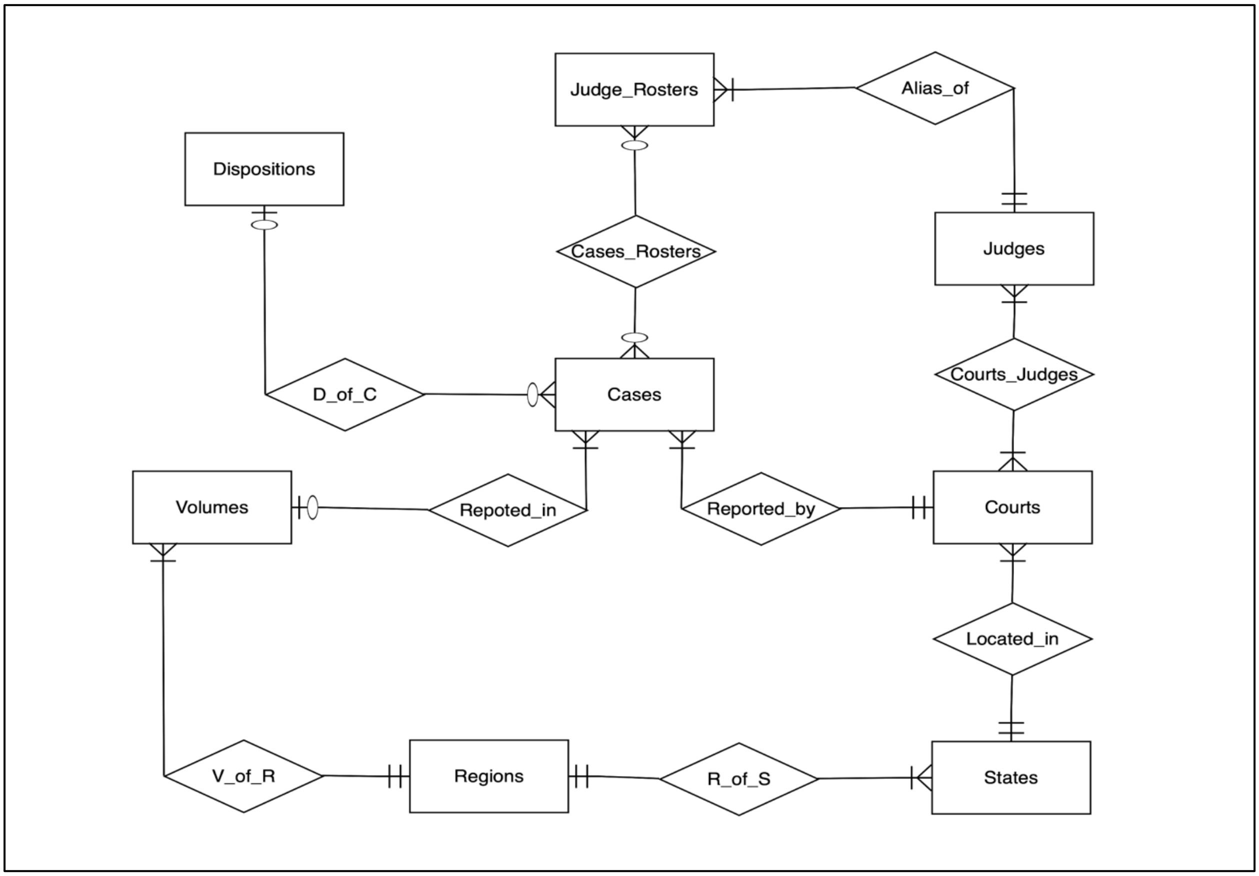
Lastly, we show how the SCOLR have been treating lower court decisions. Our collection of case dispositions allows us to closely examine these trends over time. Figure 10 plots the six most frequent disposition types. The first thing that jumps out in this graphic is the frequency of cases where the courts simply affirm the lower court's ruling. This number hits an all-time high in 1972 with nearly 7,000 affirmances. This number declined to just under 3,000 cases in 2010. Interestingly, reversals remain relatively stable over this period of time. Most of the changes in dispositions involve SCOLR remanding a case (or at least a portion of a case) back to a lower court. This is essentially allowing appellees an additional step in the appeals process with a lower court re-considering its ruling, often with instructions from the higher court.

Figure 10: Dispositions Over Time



The database has been reorganized and structured as a relational database to make it a searchable dataset and more usable for research related to SCOLR. Figure 11 shows the entity-relational diagram (ERD) of the database design. The Appendix goes into greater detail regarding the structure, entities, and relationships between all of our variables. The text of the opinions will become available in a searchable format.

Figure 11: The entity-relational diagram (ERD) of the database Design



Summary

In this study, I overcame the data shortage by collecting data and creating a novel dataset for all the cases of the SCOLR from 1953 to 2010. To the best of my knowledge, this dataset is the largest in scope and case details to date and should further our understanding of judicial decision-making. The row data was collected from LexisNexis Academic, the SCOLR's websites, and the Bluebook.

Multiple Python and SQL scripts have been created to build a tool that was able to parse and clean the collected data to produce a novel searchable and analyzable dataset. The dataset has been stored in a relational database that was created and designed based on the data relationships. Figure 11 shows the entity-relational diagram (ERD) of the database design.

The dataset contains valuable data regarding the cases and judges of the SCOLR. It contains more than 1,867,500 cases and more than 2,100 judges who worked on those cases in 52 state supreme courts. It contains 32 variables that characterize the cases, the judges, and the related entities including courts, states, and regions.

I conclude this study by presenting some statistical analyses of the data, to show some examples of using this database and how it will enrich the future of SCOLR-related research.

Chapter 4: The Network Modeling and Analysis for the Citation Network in the State Courts of Last Resorts

Network science is a promising approach to the modeling and study of complex systems. It is an interdisciplinary research area where scientists from different disciplines have a language that enables their interaction. Each scientist brings different goals and challenges, and they can share common issues across different systems (Barabási, 2016).

To use network science to study a real-world system, we need to identify some of the system's network properties; this can help us understand the characteristics of the dynamic changes in the system. Several properties that characterize a system's network have been studied by scientists, including (but not limited to) degree distribution, small-world property, scale-free property, network growth, network community, and network robustness. These properties are used to compare the network of the real system with a random network to show that the real system has different mechanisms from the random network. Once we have proved that the system network is not random, we use the network properties to understand its behavior and to predict some of the future changes in the system.

Creating the system networks will allow us to obtain a better understanding of the system and its components' behaviors. An understanding of the system networks is the first step to understanding the system and its complexity (Barabási, 2016). An analysis of these networks can lead to exploring information and finding answers to some research questions that have remained unanswered for decades.

As discussed earlier in Chapters 1 and 2, the U.S. legal system is a precedent system, and citations to prior cases are essential to reach decisions in new cases. Consequently, citation plays an important role in the system behavior and outcomes. There are many internal and external factors that have mutual impact with the citations acts and quality. Therefore, to analyze the legal system of the State Courts of Last Resort (SCOLR) and be able to predict the system's future, scholars need to identify those factors, observe the changes in citation behavior over time, identify the similarity between cases' decisions, and discover the influences between the system's internal components such as cases, judges, and courts.

The approach of network science is one of the best approaches that can be used to satisfy scholars' needs in analyzing the SCOLR system. Based on the citation behavior, the system can be modeled and analyzed as a network. In this network model, the cases will be the network's nodes, and the citations between the cases are the links that connect the nodes.

In this chapter, I have modeled the SCOLR system as a network based on its citation behavior, and then I used the approach of network science to analyze the system's complexity. I created a network of all SCOLR cases from 1953 to 2010. Also, to model the system and analyze it from different perspectives, I projected four other networks based on the cases network. These networks were projected based on four cases' features: the name of the opinion-writer judge, the state where the case was decided, the reporter region where the case's court belongs, and the published date of the case.

The results of the analysis prove the capability of network science to analyze the SCOLR system and introduce simple tools to obtain interesting results. For example, the algorithms of network community were able to identify the main legal issue of the cases, which is one of the most difficult results to reach by other approaches.

In the following sections, I introduce the research motivations and methodologies, and I describe the network construction procedure. Then, I show the analysis and results for the created networks, followed by a summary and conclusions.

Motivations

These are the motivations for this study: 1) Model the system as a network, to understand its complexity. 2) Analyze the system as a complex system. 3) Create a citation dataset for scholars to use in future research.

Objectives

These are the objectives of this study: 1) Understand the citation methods in the legal system. 2) Develop an automated tool to create the study's dataset. 3) Design a network model and construct the system network. 4) Analyze the system using the approach of network science. 5) Compare the results of the analysis with other research studies that used different approaches.

Citations in SCOLR Cases

In the legal system of the United States, judges make decisions regarding a case based on the evidence and decisions made in prior cases with similar issues—called precedent. The judge's decision is called an opinion. The judge in the current case uses the opinions in prior cases through the citation system. Citing other cases means that the judge in the current case identifies the opinions in prior cases, together with an explanation regarding the similarity between the current case and the cited cases, and the decision in the current case is based on the opinions in the cited cases. For more details, see the discussion in the section "The Importance of Citation in the Legal System" in Chapter 1 of this dissertation.

According to the Bluebook recommendations, cases should cite other cases by using their regional reporters[5]. Similar to the citations in academic articles, in the discussion of the current case's opinion, the judge needs to add the regional reporter of other cases and show how these cases directed the decision of the current case. Figure 12 shows an example of a SCOLR case that cited multiple cases in its opinion text using the regional reporters of those cases.

To build a citation network between the cases, I need to find all the cases that have cited other cases and record those citations. After I have all the records of citations, I can build the citation network. The network's nodes represent the cases, and the network's edges represent the citations between the cases. If case A cited case B in the opinion of A, I added a direct edge from A to B. For example, based on the opinion of the case (606 P.2d 310) that is shown in Figure 12, we will have a network with four nodes and three direct edges, as shown in Figure 13. By adding all the cases as nodes to the network with direct edges between them based on their citations, I was able to model the SCOLR system as a network based on the citation behavior.

[5] For more details about the cases' identifiers including regional, state, and Lexis reporters, please refer to section "The case identifiers" in Chapter 3.

Figure 12: Example of citations in legal cases.

This is a part of the opinion text of the case with regional reporter (606 P.2d 310) and Lexis reporter (1980 Wyo. LEXIS 227). This part shows multiple citations to other legal cases. Three citations have been highlighted for clarification.

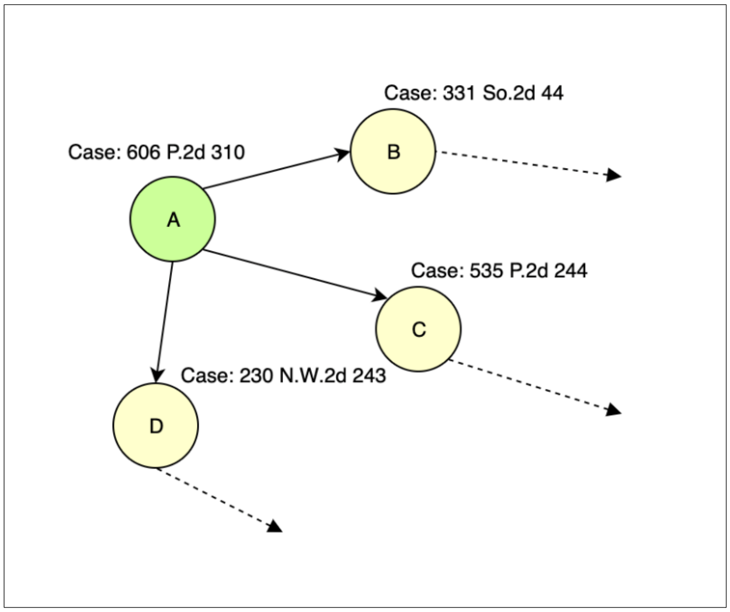
"Standing to sue doctrine. [HN7] 'Standing to sue' means that party has sufficient stake in an otherwise justiciable controversy to obtain judicial resolution of that controversy. Sierra Club v. Morton, 405 U.S. 727, 92 S. Ct. 1361, 1364, 31 L. Ed. 2d 636. Standing is a concept utilized to determine if a party is sufficiently affected so as to insure that a justiciable controversy is presented to the court. The requirement of 'standing' is satisfied if it can be said that the plaintiff has a legally protectible and tangible interest at stake in the litigation. Guidry v. Roberts, La. App., 331 So.2d 44, 50. Standing is a jurisdictional issue which concerns power of federal courts to hear and decide [**10] cases and does not concern ultimate merits of substantive claims involved in the action. [*317] Weiner v. Bank of King of Prussia, D.C.Pa., 358 F. Supp. 684, 695.

[HN8] "Standing is a requirement that the plaintiffs have been injured or been threatened with injury by governmental action complained of, and focuses on the question of whether the litigant is the proper party to fight the lawsuit, not whether the issue itself is justiciable. Carolina Environmental Study Group, Inc. v. U. S. Atomic Energy Comm., D.C.N.C., 431 F. Supp. 203, 218. Essence of standing is that no person is entitled to assail the constitutionality of an ordinance or statute except as he himself is adversely affected by it. Sandoval v. Ryan, Colo.App., 535 P.2d 244, 247."

[HN9] Standing is a concept used to determine whether a party is sufficiently affected to insure that a justiciable controversy is presented to the court. 67A C.J.S. Parties § 12, p. 662. It is a necessary and useful tool to be used by courts in ferreting out those cases which ask the courts to render advisory opinions or decide an artificial or academic controversy without there being a palpable injury to be remedied. However, [**11] it is not a rigid or dogmatic rule but one that must be applied with some view to realities as well as practicalities. Standing should not be construed narrowly or restrictively. Wisconsin's Environmental Decade, Inc. v. Public Service Commission of Wisconsin, 1975, 69 Wis.2d 1, 230 N.W.2d 243, 249. Further, as said in Residents of Beverly Glen, Inc. v. City of Los Angeles, 1973, 34 Cal.App.3d 117, 109 Cal.Rptr. 724, 727:

"In recent years there has been a marked accommodation of formerly strict procedural requirements of standing to sue [citation] and even of capacity to sue [citation] where matters relating to the 'social and economic realities of the present-day organization of society' [citation] are concerned. Accordingly, we have seen a retreat from * * * * formalism and rigidity * * * *."

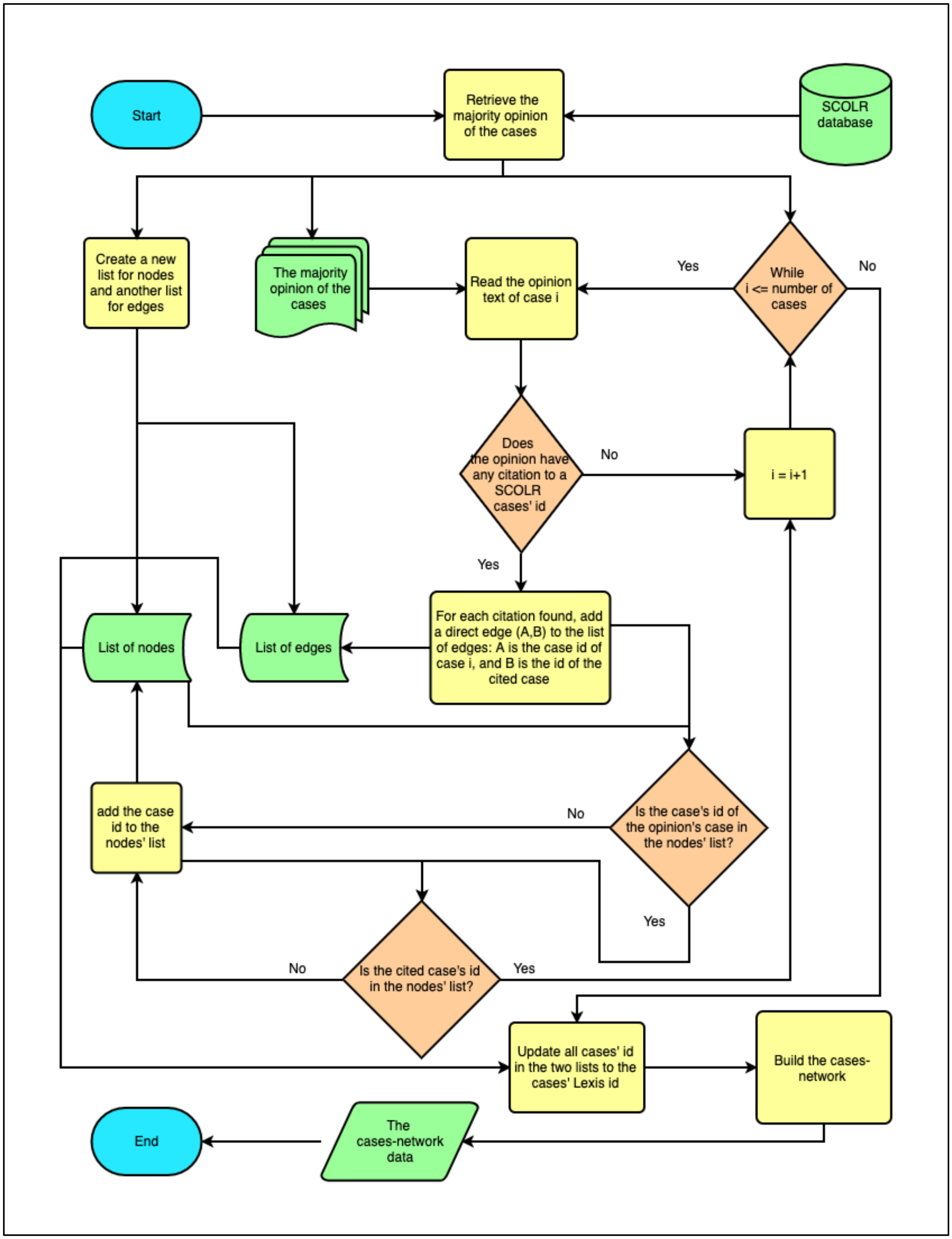
Figure 13: A sample citation network based on the citations that have been highlighted in Figure 1.



Network Construction Tool

As mentioned earlier, the system can be modeled as a network based on the citations among cases. Cases should cite other cases by using their regional reporters. However, I found many citations that used the state reporter or the Lexis reporter to cite other cases. Thus, I developed a tool to read the majority opinion text of all cases and extract the citations of all the cases' identifiers (IDs): regional reporter, state reporter, and Lexis reporter. Once the tool stores all the records of citations, the tool converts the regional reporters and state reporters that have been cited to the corresponding Lexis reporter, to have a unified ID for all cases. Figure 14 shows the flowchart of the tool. The tool was developed with Python code and SQL scripts that use Regex and other NLP methods.

Figure 14: A flowchart of the citation network construction tool.



System's Networks and Analysis

Using the network construction tool, I built the main system network model between the SCOLR cases based on the citation behavior between them. Also, as mentioned earlier, I projected four other networks based on the cases network. These networks were projected by grouping the nodes and summing the links between them based on four cases' features: the name of the judge writing the opinion, the state where the case was decided, the reporter region where the case's court belongs, and the published date of the case. Table 4 shows the essential characteristics of these networks.

In this section, I describe all five networks and their main characteristics, as well as the goal in building these networks and how their analysis will lead to understanding the system behavior.

Table 4: The system networks and their characteristics

The Network	Number of Nodes	Number of Edges	Network's Type
Cases-network	798,023	3,687,433	Unweighted & Directed
Judges-network	1,920	225,277	Weighted & Directed
States-network	50	2,500	Weighted & Directed
Regions-network	7	49	Weighted & Directed
Years-network	58	1,711	Weighted & Directed

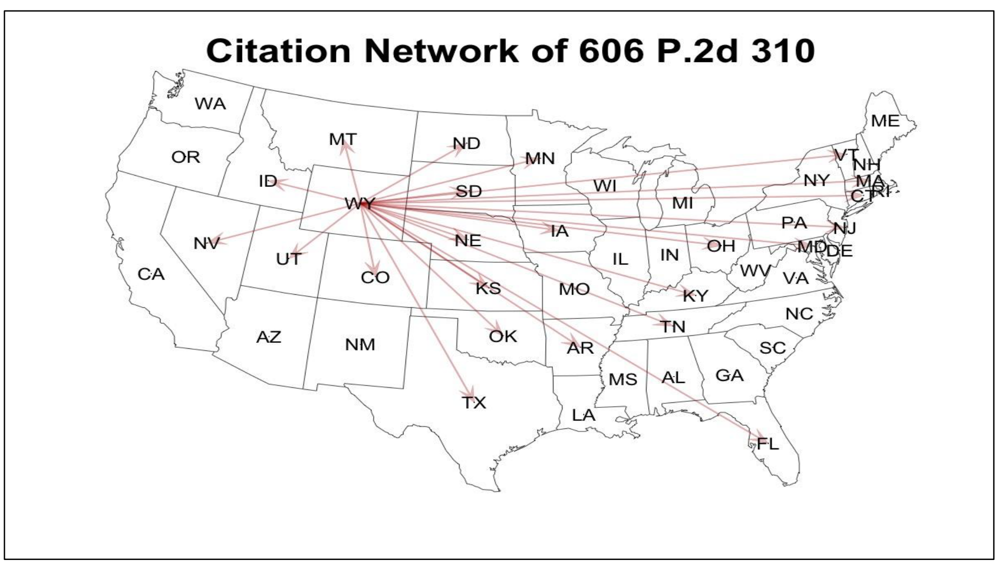
Cases-network

This network is the foundation network that models the system based on the citation behavior between the cases. The network's nodes represent the cases, and the network's edges represent the citations between the cases. If case A cited case B in the majority opinion of A, I added a direct edge from A to B. For example, Figure 15 shows the citation network of the case that has the regional reporter (606 P.2d 310) and the Lexis reporter (1980 Wyo. LEXIS 227).

The cases network is an unweighted and directed network. It has N = 798,023 nodes and L = 3,687,433 edges. Also, since each case cites one or more previously decided cases, it is an acyclic network that forms a directed tree.

Special kinds of networks like the directed tree cannot be analyzed by all the network science tools. Studies have not yet focused on such a network type. However, other networks can be projected from the tree network based on some of the system components. Therefore, in this section, I will analyze the cases network using some of the network's properties that can be applied to a tree network to get meaningful results.

Figure 15: The citation network of the case (606 P.2d 310)



The nodes' degrees. In network science, the degree k of a node is the total number of edges that connect the node to its neighbor nodes, regardless of whether the edges are directed edges or undirected edges. Also, in a directed network, each node has in-degree k_{in} and out-degree k_{out} , where k_{in} is the total number of directed edges pointed to that node from its neighbors, and k_{out} is the total number of directed edges from the node to its neighbors.

As mentioned earlier, the cases network is a directed network. The average degree $\langle k \rangle$ of a cases-network's node is (4.621). A node's degree k ranges from 1 to 5928. A node's out-degree k_{out} ranges from 0 to 873. That means that some of the cases have not cited any other cases, while other cases made up to 873 citations in their opinions. On the other hand, a node's in-degree k_{in} ranges from 0 to 5906. That means that some cases have not received any citations, while other cases received up to 5906 citations.

The degree distributions. The degree distribution is one of the network properties that is used to understand the network behavior. The degree distribution shows the number of nodes in the network with degree k for each k in the network. In directed networks, since each node has in-degree k_{in} and out-degree k_{out}, there are two degree distributions: the in-degree distribution and the out-degree distribution. Figure 16 shows the distributions of the cases-network.

Based on the in-degree and out-degree distributions, the degree distributions of the cases-network have the power-law distribution. The power-law distribution is the distribution of scale-free networks and has been found in most real-world networks. The cases-network's distributions show that there are a small number of nodes that have a high degree (called hubs) and many low-degree nodes.

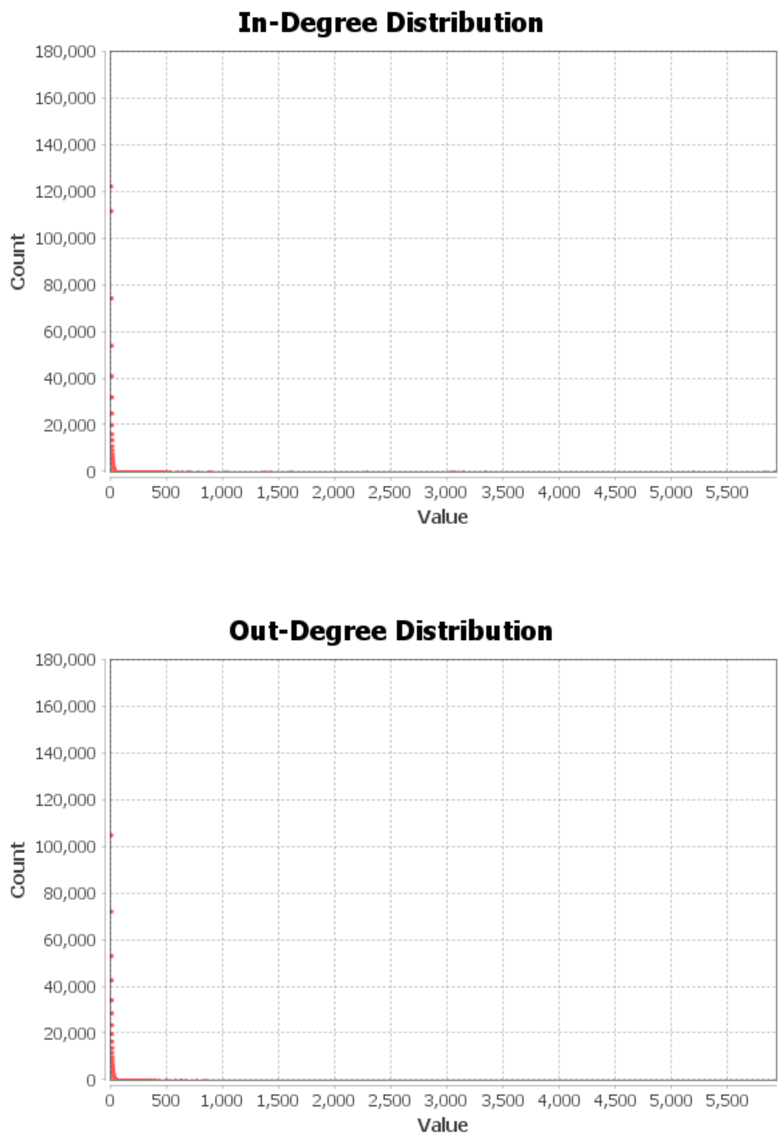


Figure 16: The in-degree and the out-degree distributions of the cases-network

Random network. A random network is a network that we can build to simulate a real network with the same number of nodes and edges.

There are two definitions of a random network (Barabási, 2016):

- G(N, L) Model: N labeled nodes are connected with L randomly placed links. Erdős and Rényi used this definition in their string of papers on random networks (Erdős & Rényi, 1959, 1960, 1961a, 1961b, 1963, 1966a, 1966b, 1968)

- G(N, p) Model: Each pair of N labeled nodes is connected with probability p, a model introduced by Gilbert (1959).

Network scientists start their analysis by testing the real network's properties and comparing them to the random-network properties, to determine whether the real network was randomly constructed. The dynamic changes in random networks can be easily predicted. This will be helpful if the real network has the characteristics of a random network.

Based on the random-network model, the cases-network is in the supercritical regime because its average degree $<k> = 4.621$ is greater than 1 and less than $ln\ N$ (ln 798,023 = 13.59). In other words, this network is not a random network, and it is expected to have a giant component, which is in agreement with the observation.

Moreover, as mentioned before, the degree distribution of the cases-network is the power-law distribution, which is not a distribution that can be found in a random network. Also, random networks don't have hubs like those found in the cases-network. Therefore, the cases-network doesn't have random-network properties.

Network Communities. The community in a network is a group of nodes that have internal links between them more than the external links that connect them with other nodes in the network. Different methods can be applied to the network to discover its communities. Modularity optimization is considered the best method for a large network, based on the accuracy of its results and the speed of its algorithms compared to other approaches (Barabási, 2016). Therefore. the approach of modularity optimization introduced by Newman (2004) has been applied to the cases-network using the algorithm of Blondel, Guillaume, Lambiotte, and Lefebvre (2008).

In the cases-network, 2,868 communities have been detected. As shown in Figure 17, most of these communities are very small communities that have less than 700 nodes, while only 24 communities absorb 55% of the nodes.

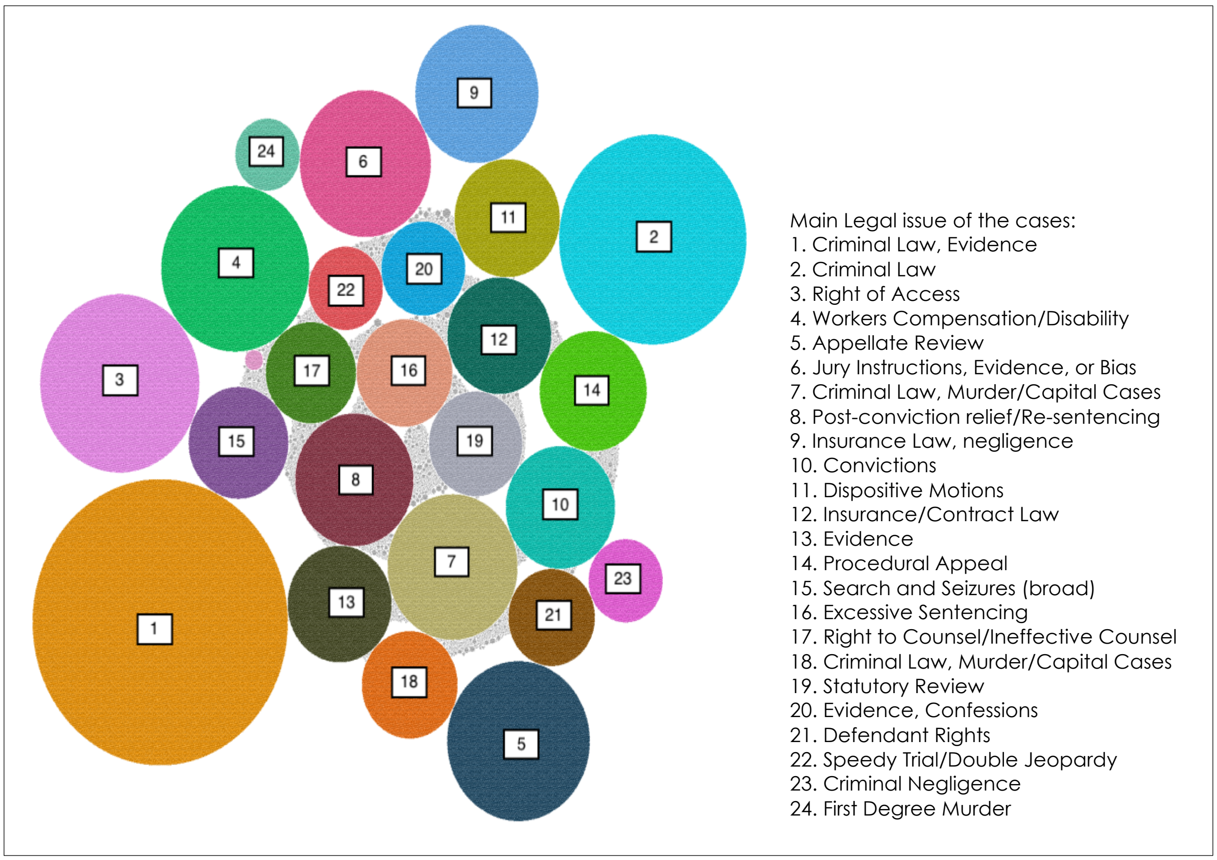
The second hypothesis of this research is that each network community represents the shared legal issues of the community's cases:

Hypothesis 2: A case cites another case based on the similarity of the legal issues of both cases. Therefore, the network of the case has clusters that represent the cases' main legal issues.

In order to test this hypothesis, we need to identify the legal issues of the cases. So, I created a random dataset sample for the cases of the 24 distinguished communities. This dataset sample has been labeled by a domain expert. The labeled dataset shows a correlation between the cases' communities and the cases' legal issues. In other words, each of the cases' communities is representing a cluster of cases that share the same legal issue. Figure 17 shows the detected communities and the legal issues they represent.

Ravi (2019) has used my database and applied the topic-modeling approach to define the legal issue of each case based on the case's details. I compared his research results with the detected network communities. I found that there is no correlation between the network communities and the topics of the cases. However, based on the hand-coding labeled dataset, the network communities gave better classification than the topic modeling for the cases' legal issues extraction. That is because the community-detection procedure encompasses more cases than the approach of topic modeling. The community-detection algorithms rely on the network structure and that led to include hundreds of thousands of more cases. On the other hand, the approach of topic modeling needs a great level of detail in the text to produce accurate results; this level of detail is missing in most SCOLR cases' records.

Figure 17: The network communities in the cases-network.
The nodes are clustered and colored based on their communities. Note: the edges between the nodes have been removed from the graph for clarification.



Judges-network

This network has been projected from the cases-network based on the majority-writer feature of the cases. The majority writer is the person, persons, or court (per curium) who wrote the rationale behind the court's ruling in a case. The majority writer is known for most of the cases in my dataset, but the majority writer is missing for some of them. Therefore, I excluded all the cases without the majority writer in the construction process of the judges-network.

The judges-network is the richest network among the networks tested, in regard to the information and interactions between the nodes. It reflects the judges' citation behaviors along with the effect of some organizational rules on selecting the judges in general and selecting the majority writer specifically.

Building the judge-network was done by merging the cases based on the majority writers' similarities between cases. All the cases with the same majority writer became one node. The node ID is the majority writer's ID. After creating all the network nodes, direct edges were added between the nodes based on the citations between the cases. A direct edge (A, B) is added if judge A wrote a case X that cited case Y that has been written by judge B. The weight of edge (A, B) is the total number of citations between the cases that have been written by judge A to all the cases that have been written by judge B.

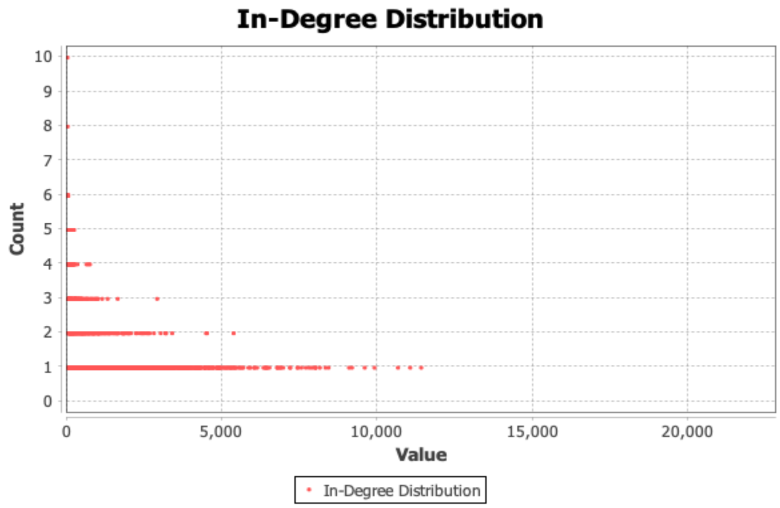
The resultant network is the judges-network, which has N = 1,920 nodes and L = 225,277 edges. It is a weighted and directed network with loops and self-loops. A loop is a group of nodes and edges that make a circle by themselves. A self-loop is an edge from the node to itself.

The nodes' degrees. Since the judges-network is a weighted network, it has two types of degrees. The first one is the node degree, which shows the number of links between the node and its neighbors. The second type is the node weighted degree, which shows the number of interactions between the node and its neighbors. So in the judges-network, the weighted degree shows the number of citations between the judge's cases and other judges' cases. In addition, because it is a directed network, each node has in-degree (representing the incoming links to the node) and out-degree (representing the outgoing links to other nodes).

The average degree $<k>$ of the judges-network's nodes is (117.33). The nodes' degree k ranges from 1 to 1,283. The nodes' out-degree k_{out} ranges between 0 to 915. That means that some of the judges have not cited any other judges, while others cited up to 915 judges in their opinions. The nodes' in-degree k_{in} ranges from 0 to 735. That means that some judges have not received any citations while other judges received citations to their cases from up to 735 judges.

The average weighted degree of the judges-network's nodes is 1,279. The nodes' weighted degree k ranges from 1 to 22,801. The nodes' weighted out-degree k_{out} ranges between 0 to 11,774. That means that some judges have not cited any other cases, while others cited up to 11,774 cases in their opinions. The nodes' weighted in-degree k_{in} ranges from 0 to 11,382. That means that some judges have not received any citations, while other judges received up to 11,382 citations to their cases.

The degree distributions. The distributions of the weighted in-degree and weighted out-degree of the judges-network are shown in Figure 18. These distributions satisfy the power-law distribution; this is the distribution of scale-free networks, and it has been found in most real-world networks. The distributions show that there are a small number of nodes that have high degrees (hubs) and many nodes that have low degrees.



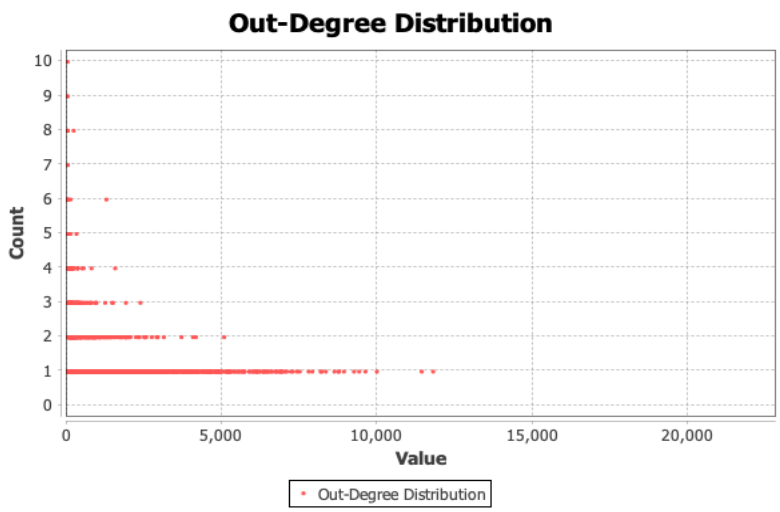


Figure 18: The weighted in-degree and the weighted out-degree distributions of the judges network

Centrality and influence. The centrality measurement is a property that shows the influence of the nodes on the network behavior. The type of influence and how to measure it depends on the type of network centrality that is used. Different types of centralities can be measured in a network: degree centrality, closeness centrality, betweenness centrality, and page-rank centrality. Each type gives different insights to understand the interaction between the network components. Also, the type of centrality can be selected based on the information needed or the network type (Das, Samanta, and Pal, 2018). Each type will be described below.

Degree Centrality (DC). DC is the simplest and easiest centrality measurement to compute. It was introduced by Shaw (1954), defined by NIEMINEN (1974), and mathematically developed by Freeman (1978). In DC measurement, each node is ranked based on the weighted degree of the node. The highest centrality (highest influence) node is the node with the highest weighted degree. In a directed network, the in-degree or the out-degree is mostly used for DC; it depends on the network's behavior that needs to be analyzed. This type of centrality can be clear if we think about it in the social network, where a person with a high number of followers in the network has more influence on the network compared to a person with a few followers. So, in the judges-network, the weighted in-degree is more suitable to use. Table 5 shows the top ten influencer judges based on DC score.

Closeness Centrality (CC). CC was introduced by Bavelas (1948) and defined by Sabidussi (1966), then generalized to weighted networks by Newman (2001). The CC score for each node is the total of all lengths of shortest paths between the node and all other nodes in the network. So, it considers the node that is the closest to all other nodes in the network to be the most influential in the network. CC is considered one of the best centrality algorithms to discover the influence in a citation network. Table 6

shows the top ten influencer judges based on CC score.

Table 5: The top ten judges based on the degree centrality score.

The weighted degree is the number of citations. For example, the weighted indegree is the number of citations the judge received.

Judge ID	Name	State	Regional Reporter	Weighted Indegree	Weighted Outdegree	Weighted Degree
Cal.14	Mosk	CA	P.	11382	7397	18779
Del.15	R. Holland	DE	A.	11027	11774	22801
Miss.30	J. Robertson	MS	So.	10633	7885	18518
Okla. Crim. App.5	Bussey	OK	P.	9876	8742	18618
Pa.17	Roberts	PA	A.	9564	7001	16565
Ind.15	Hunter	IN	N.E.	9154	6437	15591
N.D.17	Vandewalle	ND	N.W.	9061	11407	20468
Ind.21	Pivarnik	IN	N.E.	8399	8598	16997
Alas.5	J. Rabinowitz	AK	P.	8304	7480	15784
Miss.29	Prather	MS	So.	8120	7194	15314

Table 6: The top ten judges based on the closeness centrality score

Judge ID	Name	State	Regional Reporter	closeness centrality
Mich.7	Dethmers	MI	N.W.	1.00
W. Va.20	Miller	WV	S.E.	0.63
Vt.23	Dooley	VT	A.	0.61
W. Va.27	Davis	WV	S.E.	0.60
W. Va.21	Mchugh	WV	S.E.	0.59
Conn.44	Katz	CT	A.	0.58
Alas.12	Matthews, Jr.	AK	P.	0.59
Md.20	Eldridge	MD	A.	0.59
Okla.24	Kauger	OK	P.	0.58
Tenn.21	Drowota	TN	S.W.	0.58

Betweenness Centrality (BC). BC was introduced by Shaw (1954) and Freeman (1977, 1978). In BC, each node A is given a betweenness value that is equal to the number of shortest paths that pass through node A to connect any pair of nodes in the network. In other words, the betweenness of node A is the number of times node A appears in the set of shortest paths of the network's nodes. So, the betweenness of a node shows how important this node is to the connectivity of the network and the information traveling within the network (Brandes, 2001). Table 7 shows the top ten influencer judges based on BC score.

Page-Rank Centrality (PRC). PRC is the first algorithm used by Google in their search engine to rank web pages. It was introduced in 1999 by the Google founders Lawrence Page and Sergey Brin with their co-authors Rajeev Motwani and Terry Winograd. In PRC, each node is given a score that represents the degree of the node and the degree of its neighbors. PRC gives nodes a high centrality value by considering the importance of their connections to their neighbors. PRC also takes into account the direction and the weight of the links between the nodes. Therefore, PRC is more informative when used on weighted directed networks. Table 8 shows the top ten influencer judges based on PRC score.

Table 7: The top ten judges based on the betweenness centrality score

Judge ID	Name	State	Regional Reporter	Betweenness Centrality
Cal.14	Mosk	CA	P.	104751
Alas.5	Rabinowitz	AK	P.	67201
Kan.11	Schroeder	KS	P.	66244
Alas.12	Matthews, Jr.	AK	P.	57601
Va.10	Carrico	VA	S.E.	52255
Wisc.21	Abrahamson	WI	N.W.	48157
Ark.5	G. R. Smith	AR	S.W.	46729
Utah.5	Crockett	UT	P.	44387
Neb.9	L. Boslaugh	NE	N.W.	42675
Mass.18	H. Wilkins	MA	N.E.	41471

Table 8: The top ten judges based on the page-rank centrality score

Judge ID	Name	State	Regional Reporter	PageRank
Pa.5	Bell	PA	A.	0.006451
Wisc.7	Currie	WI	N.W.	0.006344
Cal.5	Traynor	CA	P.	0.006242
Pa.4	C. Jones	PA	A.	0.005307
Tex. Crim. App.2	Morrison	TX	S.W.	0.005220
Md.5	Hammond	MD	A.	0.004857
Tex. Crim. App.3	Woodley	TX	S.W.	0.004381
N.C.7	H. Parker	NC	S.E.	0.004148
Ark.5	G. R. Smith	AR	S.W.	0.004036
Ga.1	Duckworth	GA	S.E.	0.003922

Degree centrality, closeness centrality, betweenness centrality, and page-rank centrality are great centrality measurements to discover the influence of the judges on their network based on the citations from different views. Table 5Table 6

Betweenness Centrality (BC). BC was introduced by Shaw (1954) and Freeman (1977, 1978). In BC, each node A is given a betweenness value that is equal to the number of shortest paths that pass through node A to connect any pair of nodes in the network. In other words, the betweenness of node A is the number of times node A appears in the set of shortest paths of the network's nodes. So, the betweenness of a node shows how important this node is to the connectivity of the network and the information traveling within the network (Brandes, 2001). Table 7 shows the top ten influencer judges based on BC score.

Page-Rank Centrality (PRC). PRC is the first algorithm used by Google in their search engine to rank web pages. It was introduced in 1999 by the Google founders Lawrence Page and Sergey Brin with their co-authors Rajeev Motwani and Terry Winograd. In PRC, each node is given a score that represents the degree of the node and the degree of its neighbors. PRC gives nodes a high centrality value by considering the importance of their connections to their neighbors. PRC also takes into account the direction and the weight of the links between the nodes. Therefore, PRC is more informative when used on weighted directed networks. Table 8 shows the top ten influencer judges based on PRC score.

Table 7Table 8 show the ten most influential judges on the network based on the four mentioned centrality measurements. You can see that each centrality measurement has different results from the others, with very few similarities. That is normal because each one is calculated based on different aspects. The best centrality that can be used in this case depends on the application and the information needed from the analysis. For example, DC can be used to find the most-cited judges based on the weighted in-degree. Another example is that PRC can be used to find the judges with high influence on the network's hubs because they have been cited by hubs.

One observation that I found in the results of all the applied centralities is that the top 10 judges of each centrality are from different states and even different regions. Another observation is in the PRC result: All of the 10 most influential judges are senior judges, unlike other centralities' results where the judges served in different periods. You can refer to the database to see the judges' dates of service.

Hubs and self-citations. One of the most interesting findings in this analysis is that I found that the network's hubs, the judges who have high degree, tend to cite themselves more than other judges. This could be a sign of their self-confidence or other reasons that need to be discovered in the future. Table 9 shows the top ten hub judges and the number of citations they made to themselves.

Table 9: The top ten hubs in the judges-network and the number of citations the hub judges made to themselves

Judge ID	Name	self-citations	Total citations
Del.15	Randy J. Holland	2573	7397
N.D.17	Vandewalle	2138	11774
Cal.14	Stanley Mosk	1410	7885
Okla. Crim. App.5	Bussey	3805	8742
Miss.30	J. Robertson	1835	7001
Ind.21	Pivarnik	3043	6437

Judge ID	Name	self-citations	Total citations
Pa.17	Roberts	2512	11407
Ind.19	Givan	2011	8598
Mass.20	Abrams	1080	7480
Conn.29	Peters	1563	7194

Degree Correlation. The degree correlation is a network property that shows the likelihood that nodes connect to nodes with a similar degree. In other words, the degree correlation shows if small-degree nodes tend to link with hubs or to any-degree nodes, and if hubs tend to link with hubs or any-degree nodes. So, the degree correlation will show the correlation between a node's degree and the degree of the node's neighbors.

The degree correlation of a network can be assortative, dis-assortative, or neutral. In assortative correlation, the nodes tend to link to the nodes with a similar degree, while dis-assortative correlation is the opposite. In the neutral correlation, the nodes are connected with any nodes, regardless of their degree (Barabási, 2016).

In a directed network like the judges-network, each node has two degrees: in-degree and out-degree. So, to discover the degree correlation, we need to define four degree-correlations based on the in-degree and out-degree of each node and the in-degree and the out-degree of the neighbors of the nodes with the same degree. Consequently, the four correlations are in-in, in-out, out-in, and out-out.

Two methods have been used in this analysis to find the degree correlations: K-nearest Neighbors and Pearson Degree Correlation.

K-Nearest Neighbors (K_{nn}). The correlation is computed between each degree in the network and the average degree of the neighbors of each node of a given degree. Then, plot the results as a linear graph to visually observe the correlations.

$K_{nn}(K)$ = the average degree of the neighbors of all degree-k nodes.

As shown in Figure 19 Figure 20, it is hard to visually determine the correlation types. However, the four correlations (in-in, in-out, out-in, and out-out) show neutral correlations for small degrees and assortativity for large degrees. That means that small-degree nodes are links to both small-degree nodes and large-degree nodes, while the large-degree nodes (the hubs) tend to link with similar-degree nodes. This result supports the previously mentioned observation about the hubs and self-citations.

Pearson Degree Correlation (PDC). This is a statistical method that computes the correlation between all the degrees of the nodes and the degrees of their neighbors. I used this method to prove my observation on K_{nn} results, since it was hard to visually decide the degree correlation of the network.

The formula for the Pearson degree correlation is

$$r = \frac{n\left(\sum xy\right) - \left(\sum x\right)\left(\sum y\right)}{\sqrt{\left[n\sum x^2 - \left(\sum x\right)^2\right]\left[n\sum y^2 - \left(\sum y\right)^2\right]}}$$

where r = the correlation score, n = number of nodes, x = the degree of the node, and y = the degree of the node's neighbor

The value of r can be any number from 1 to -1; positive values mean assortative correlation, negative values mean dis-assortative correlation, and zero means neutral correlation.

By applying the formula to the four correlations, these are the results:

r (in-in) = 0.038

r (in-out) = - 0.011

r (out-in) = 0.054

r (out-out) = 0.099

The results show that the judges network indicates a lack of in-out correlations, but it shows the presence of assortativity for the other three correlations (in-in, out-in, out-out). However, the correlation scores r of the four correlations (in-in, in-out, out-in and out-out) are very close to zero.

Figure 19: Degree correlation out-in and out-out of the judges-network using K-Nearest Neighbors (Knn).

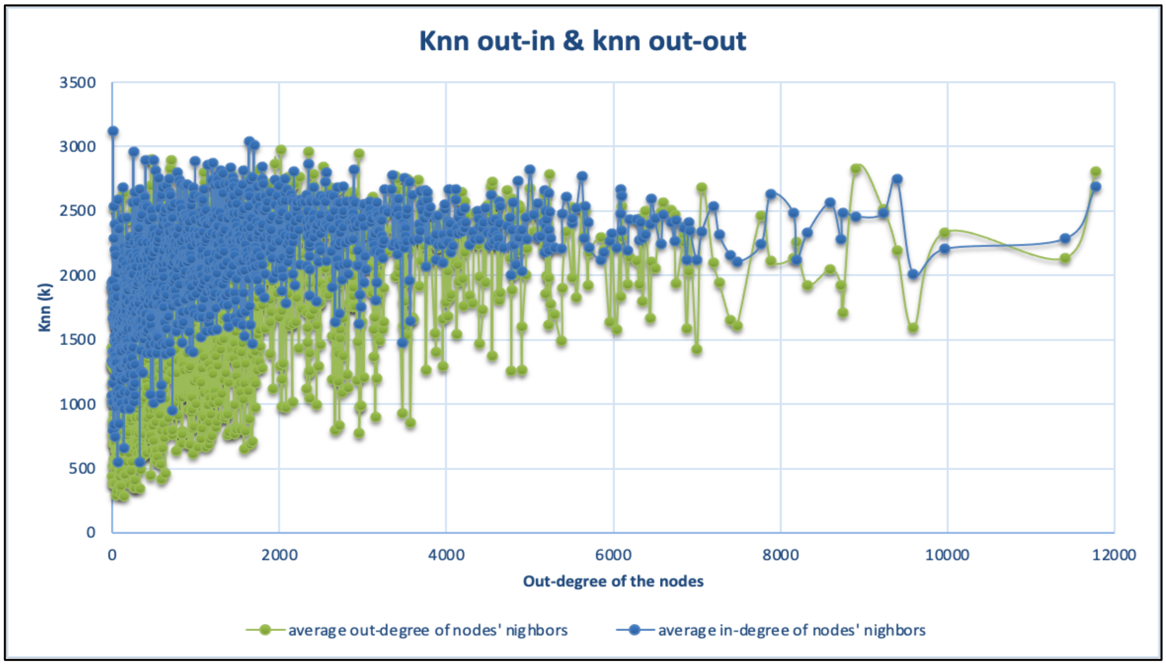
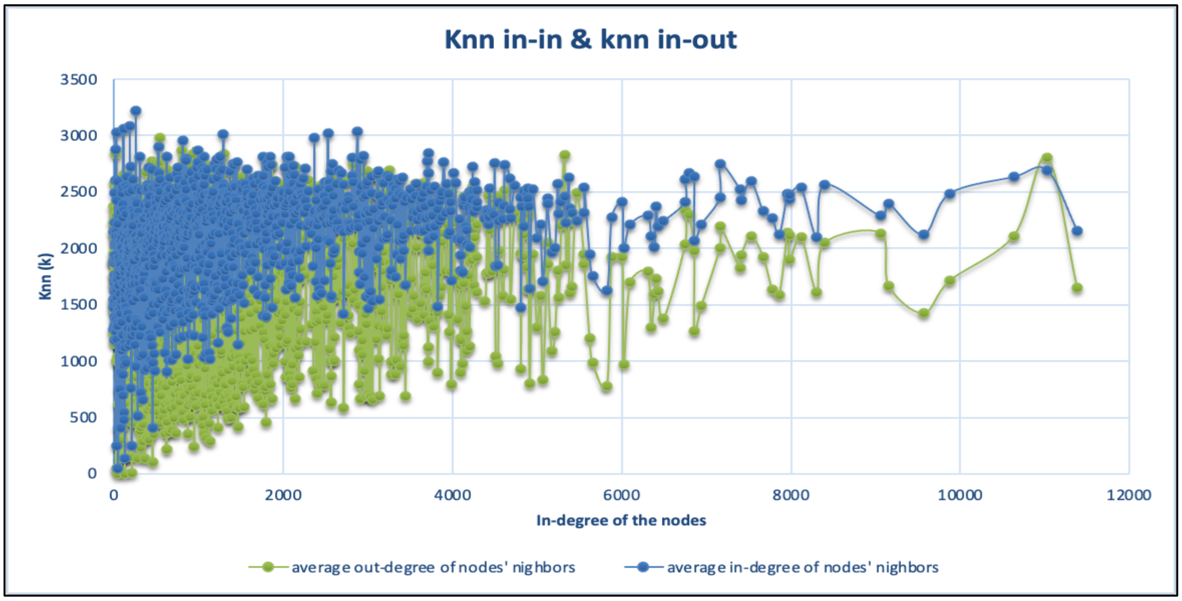


Figure 20: Degree correlation in-in and in-out of the judges-network using K-Nearest Neighbors (Knn).



Considering the results of degree correlation and the previous section of hubs and self-citations, I conclude that the degree correlation of the judges-network is assortative for hubs but neutral in general. In other words, the judges in SCOLR cite other cases regardless of the popularity of the opinion writers of the cited cases.

Network Robustness. Network robustness is a property of the network that indicates how many nodes we need to remove from the network to break it down. Networks with a big giant component, such as the judges-network, are usually very robust and very hard to break down (Barabási, 2016). By applying the Molloy-Reed criteria to quantify scale-free network robustness, we get the following (Molloy & Reed, 1995):

The network has a giant component if $k > 1$. (Therefore, a giant component exists in the cases-network, since $k = 117.33$.) The critical breakdown threshold for a real network is:

$f_c = 1 - \frac{1}{\frac{<k^2>}{<k>} - 1}$, where <k> is the average degree and < k^2> is the average degree of the second moment.

So, this is the critical breakdown threshold for the judges-network:

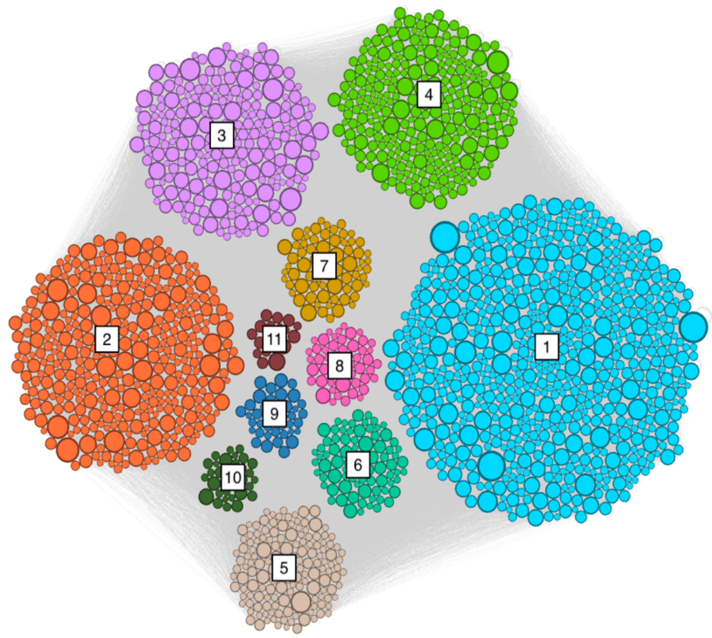
$f_c = 1 - \frac{1}{\frac{45,570}{117.33} - 1}$ = 0.997,

This fraction represents the percentage of nodes that needs to be removed to break the network apart; thus, we need to remove 99.7% of the nodes to fragment the network. The breakdown threshold of a random network is $(f_c^{ER}) = 1 - \frac{1}{<k>}$ = 0.991. If $f_c > f_c^{ER}$, the network is robust. Thus, the judges-network is robust.

Network Communities. By applying the modularity-optimization approach introduced by Newman (2004) using the algorithm of Blondel, Guillaume, Lambiotte, and Lefebvre (2008) to the judges-network, 11 communities were detected.

By merging the judges-network communities and the cases-network communities, I found that each of the judges-network communities is representing a cluster of judges that are expert in the same legal issue. In other words, using the network community's extraction algorithms on the judges-network, I was able to classify the judges into 11 classes. Each class is representing a legal issue. Figure 21 shows the detected communities and the legal issues they represent.

Figure 21: The network communities in the judges network.
Each communities' nodes are clustered and colored with the same color.

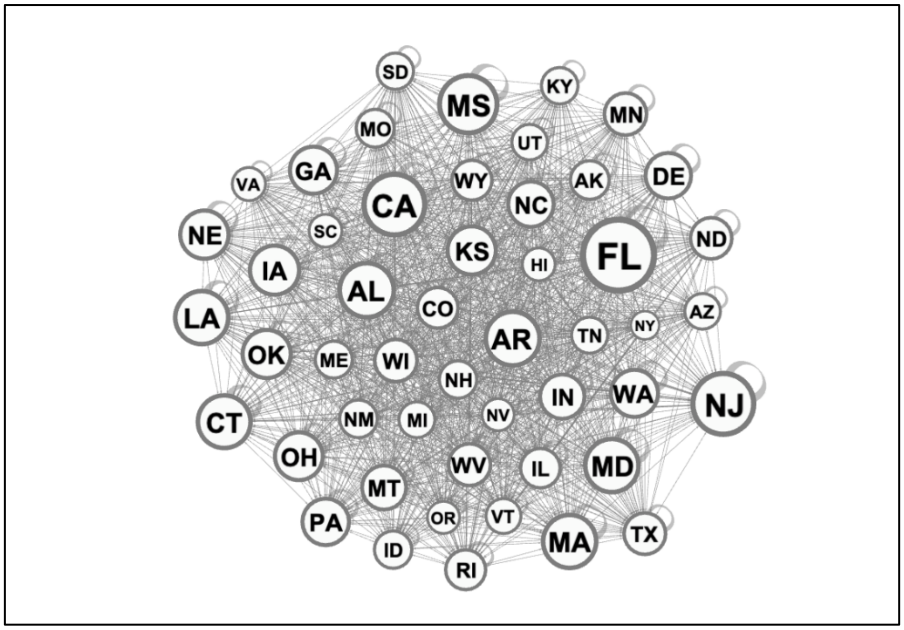


Main Legal issue of the cases:

1. Criminal Law, Murder/Capital Cases
2. Insurance/Contract Law
3. Convictions
4. Post-conviction relief/Re-sentencing
5. Criminal Law, Murder/Capital Cases
6. Appellate Review
7. Insurance Law, negligence
8. Jury Instructions, Evidence, or Bias
9. Criminal Law
10. Evidence
11. Criminal Law, Evidence

States-network

Figure 22: The states-network



This network has been projected from the cases-network based on the state where the court that decided the case is located. The states-network has been projected to examine the influences between the states. Therefore, the network has N= 50 nodes that represent the 50 states in the USA. Each state has one court except TX and OK, which each have two courts (one for criminal cases, and one for civil cases). So, I merged the cases of the two courts in TX and OK to be consistent with other states. The edges between the nodes represent the citations between the cases of each state to other states. Also, the nodes in the states-network have a self-loop that represents citations between the cases of the same state. Based on the citations between the cases, the states-network has L = 2500 edges. Since $L = N^2$, this network is fully connected: each node has a direct edge to each other node in the network, including itself. In network science and graph theory, a fully connected network is called a clique. Moreover, the states-network is a weighted network: the edges' weight represents the number of citations between the cases of each state at the two ends of the edge, based on the direction of the edge.

States and cases statistical analysis. Table 10 shows the number of cases that each state has reported, the number of cases that made citations (along with the percentage of these cases out of all reported cases), and the number of cases that received citations (along with the percentage of these cases out of all reported cases). In this table, we can see that only 29% of the cases have cited other cases, while 39% of the cases have received citations. This indicates that most of the cases in the SCOLR have been orders, dismissed, or not heard. Also, the number of cases that received citations is more than the number of cases that made citations, since I have not counted the citations to the cases before 1953. So, there are some cases that have cited cases decided before 1953 that also received citations from other cases. These cases made citations that have not been counted but received citations that have been counted.

In Table 10, we can see that California has the maximum number of cases and the maximum number of cases that made citations, but Florida has the maximum number of cases that received citations. On the other side, Alaska reported the minimum number of cases, New York has the minimum number of cases that cited other cases, and Hawaii has the minimum number of cases that have been cited by other cases. The average number of reported cases among all states is 34,242; the average number of cases that cited other cases is 9,998, and the average of the cases that have been cited is 13,274.

Moreover, we can see in Table 10 that West Virginia has the highest percentage, 93%, of cases that made citations out of its reported cases, while only 1% of New York's cases cited other cases. The average percentage of cases that made citations among all states is 48%. Nebraska has the highest percentage for the cases that have been cited out of all reported cases: 90%. The lowest percentage is 4%, which is the percentage of cases that have been cited out of all cases reported by California. The average percentage of cases that received citations among all states is 56%.

Table 10: Total number of cases of each state, number and percentage of cases that made citations out of all cases, and number and percentage of cases that received citations out of all cases.

State	Number of Cases	Cases Made Citations		Cases Received Citations	
		n	%	n	%
AK	6,647	6,035	91%	5,534	83%
AL	44,674	17,652	40%	19,205	43%
AR	28,223	20,677	73%	17,340	61%
AZ	9,990	6,282	63%	6,252	63%
CA	210,936	26,550	13%	8,167	4%
CO	30,394	9,945	33%	10,344	34%
CT	18,703	8,736	47%	14,340	77%
DE	18,365	10,244	56%	12,806	70%
FL	91,461	25,355	28%	48,627	53%
GA	36,604	16,777	46%	16,507	45%
HI	9,830	4,265	43%	3,713	38%
IA	15,349	12,244	80%	11,994	78%
ID	8,365	6,688	80%	6,476	77%
IL	39,826	6,869	17%	12,785	32%
IN	24,269	9,625	40%	14,469	60%
KS	21,648	10,183	47%	14,599	67%
KY	17,272	10,108	59%	10,255	59%
LA	120,561	17,530	15%	30,454	25%
MA	29,244	12,669	43%	17,683	60%
MD	26,592	9,597	36%	18,806	71%
ME	9,926	8,806	89%	7,945	80%
MI	73,372	3,884	5%	15,415	21%
MN	31,591	11,479	36%	11,366	36%
MO	12,192	8,966	74%	8,598	71%
MS	31,080	13,803	44%	16,209	52%
MT	18,627	12,894	69%	12,127	65%
NC	39,916	9,248	23%	14,962	37%
ND	9,514	8,232	87%	8,427	89%

State	Number of Cases	Cases Made Citations		Cases Received Citations	
		n	%	n	%
NE	15,182	13,729	90%	13,623	90%
NH	9,352	8,323	89%	7,734	83%
NJ	58,852	6,326	11%	38,413	65%
NM	17,432	5,285	30%	10,691	61%
NV	35,149	5,905	17%	6,639	19%
NY	148,840	1,302	1%	9,359	6%
OH	70,399	10,668	15%	26,159	37%
OK	20,716	15,398	74%	15,272	74%
OR	33,123	2,664	8%	12,016	36%
PA	57,376	11,367	20%	17,613	31%
RI	15,869	9,368	59%	9,727	61%
SC	12,310	7,695	63%	7,788	63%
SD	7,669	5,881	77%	5,736	75%
TN	22,851	7,615	33%	6,028	26%
TX	44,532	12,920	29%	13,491	30%
UT	11,027	6,997	63%	7,503	68%
VA	8,208	6,661	81%	6,509	79%
VT	11,874	6,618	56%	7,169	60%
WA	28,944	6,961	24%	20,641	71%
WI	31,503	8,714	28%	12,828	41%
WV	8,315	7,723	93%	7,002	84%
WY	7,396	6,412	87%	6,367	86%
Total	1,712,090	499,875	29%	663,713	39%
Maximum	210,936	26,550	93%	48,627	90%
Minimum	6,647	1,302	1%	3,713	4%
Average	34,242	9,998	48%	13,274	56%

States and Citations Statistical Analysis. Table 11 gives some details about the states and their citation behaviors. In this table, we can see that the total number of citation interactions in the network is 7,197,816 interactions. However, these interactions include each citation twice: the first time as an in-citation, which is the incoming citation to the state from any state; the second time as an out-citation, which is the out-going citation from the state to any state. So, the total number of citations is half of the interactions, which is 3,598,908 citations.

Florida has the maximum number of citations in general and the maximum number of in-citations and out-citations. In contrast, New York has the minimum number of total citations, out-citations, and in-citations. For New York, 70% of its total citations were in-citations that have been received from other states. That is the maximum percentage of in-citations out of all citations. On the opposite side, West Virginia has the maximum percentage of out-citations: 54% of its total citations.

Out-State and Self-Citations. Table 11 shows that 90% of the total citations were self-citations and only 10% were cross-state citations. The average number of self-citations was 63,147 citations, and the average number of out-state citations was 7,238. That indicates that states tend to cite themselves more than they cite other states. Florida has the maximum number of self-citations, and West Virginia has the maximum number of out-state citations. On the other hand, New York has the minimum number of self-citations and out-state citations. Georgia has the maximum percentage of self-citations and the minimum percentage of out-state citations, the opposite of West Virginia.

Table 11: Total number of citations of each state, number and percentage of citations received and sent by each state, and number and percentage of self-citations and citations to other states.

State	No. Citations	Incoming Citations		Outgoing Citations		Self-Citations		Out-State Citations	
		n	%	n	%	n	%	n	%
AK	113,551	53,932	47%	59,619	53%	48,709	82%	10,910	18%
AL	212,289	105,384	50%	106,905	50%	99,549	93%	7,356	7%
AR	219,322	110,013	50%	109,309	50%	104,557	96%	4,752	4%
AZ	84,265	43,015	51%	41,250	49%	34,522	84%	6,728	16%
CA	264,079	138,036	52%	126,043	48%	120,002	95%	6,041	5%
CO	125,812	62,805	50%	63,007	50%	53,576	85%	9,431	15%
CT	223,909	107,077	48%	116,832	52%	101,131	87%	15,701	13%
DE	164,687	81,944	50%	82,743	50%	77,908	94%	4,835	6%

State	No. Citations	Incoming Citations		Outgoing Citations		Self-Citations		Out-State Citations	
		n	%	n	%	n	%	n	%
FL	274,485	144,443	53%	130,042	47%	125,936	97%	4,106	3%
GA	164,774	84,208	51%	80,566	49%	79,288	98%	1,278	2%
HI	66,634	31,531	47%	35,103	53%	28,174	80%	6,929	20%
IA	194,326	94,727	49%	99,599	51%	85,730	86%	13,869	14%
ID	101,082	49,137	49%	51,945	51%	43,839	84%	8,106	16%
IL	109,320	58,068	53%	51,252	47%	45,966	90%	5,286	10%
IN	152,257	76,788	50%	75,469	50%	71,992	95%	3,477	5%
KS	178,108	89,365	50%	88,743	50%	80,764	91%	7,979	9%
KY	84,759	43,229	51%	41,530	49%	38,273	92%	3,257	8%
LA	204,198	103,093	50%	101,105	50%	96,940	96%	4,165	4%
MA	216,375	110,341	51%	106,034	49%	97,698	92%	8,336	8%
MD	225,304	108,709	48%	116,595	52%	99,792	86%	16,803	14%
ME	99,188	48,621	49%	50,567	51%	43,776	87%	6,791	13%
MI	70,353	38,948	55%	31,405	45%	30,431	97%	974	3%
MN	130,847	67,476	52%	63,371	48%	56,377	89%	6,994	11%
MO	98,442	49,129	50%	49,313	50%	44,015	89%	5,298	11%
MS	255,002	126,495	50%	128,507	50%	121,759	95%	6,748	5%
MT	148,654	73,947	50%	74,707	50%	68,803	92%	5,904	8%
NC	161,155	81,529	51%	79,626	49%	73,811	93%	5,815	7%
ND	135,156	65,043	48%	70,113	52%	60,933	87%	9,180	13%
NE	177,845	87,376	49%	90,469	51%	80,349	89%	10,120	11%
NH	104,364	49,915	48%	54,449	52%	44,723	82%	9,726	18%
NJ	267,162	134,928	51%	132,234	49%	120,381	91%	11,853	9%
NM	102,883	51,097	50%	51,786	50%	43,488	84%	8,298	16%
NV	61,321	29,313	48%	32,008	52%	25,013	78%	6,995	22%
NY	21,480	15,091	70%	6,389	30%	5,415	85%	974	15%
OH	152,816	77,306	51%	75,510	49%	70,118	93%	5,392	7%
OK	180,627	88,977	49%	91,650	51%	82,169	90%	9,481	10%
OR	46,967	29,614	63%	17,353	37%	16,234	94%	1,119	6%
PA	158,293	81,636	52%	76,657	48%	72,202	94%	4,455	6%
RI	122,179	59,685	49%	62,494	51%	55,670	89%	6,824	11%
SC	65,098	32,459	50%	32,639	50%	28,907	89%	3,732	11%
SD	88,321	40,904	46%	47,417	54%	37,536	79%	9,881	21%
TN	92,658	45,176	49%	47,482	51%	41,353	87%	6,129	13%
TX	120,975	59,639	49%	61,336	51%	54,774	89%	6,562	11%
UT	89,768	42,498	47%	47,270	53%	37,541	79%	9,729	21%
VA	66,651	33,853	51%	32,798	49%	29,515	90%	3,283	10%

State	No. Citations	Incoming Citations		Outgoing Citations		Self-Citations		Out-State Citations	
		n	%	n	%	n	%	n	%
VT	84,347	40,093	48%	44,254	52%	36,859	83%	7,395	17%
WA	166,133	84,947	51%	81,186	49%	72,778	90%	8,408	10%
WI	129,306	68,403	53%	60,903	47%	55,770	92%	5,133	8%
WV	133,822	60,047	45%	73,775	55%	56,522	77%	17,253	23%
WY	127,131	59,250	47%	67,881	53%	55,770	82%	12,111	18%
Total	7,038,480	3,519,240	50%	3,519,240	50%	3,157,338	90%	361,902	10%
Maximum	274,485	144,443	70%	132,234	55%	125,936	98%	17,253	23%
Minimum	21,480	15,091	45%	6,389	30%	5,415	77%	974	2%
average	140,770	70,385	50%	70,385	50%	63,147	89%	7,238	11%

The nodes' degrees. Recall that the states-network is a weighted, directed, and fully connected network. Since the network is fully connected, the average degree $<k>$ of the network's nodes is 100, since each node has an incoming and outgoing edge to all 50 network nodes, including itself. All nodes have the same degree: $k = 100$. Also, all nodes have the same out-degree: k_{out} = in-degree $k_{in} = 50$.

The average weighted degree of the states-network's nodes is 73,748.66. The nodes' weighted degree k ranges from 22,128 to 345,898. The nodes' weighted out-degree k_{out} ranges from 6,554 to 165,319. That means the cases of one or more states have cited 6,554 other cases, while the cases of other states cited up to 165,319 cases in their opinions. On the other hand, the nodes' weighted in-degree k_{in} ranges from 15,574 to 180,579. That means that each state receives a minimum of 15,574 and a maximum of 180,579 citations to their cases. Table 12 shows all the weighted degrees of all the nodes in the states-network.

Table 12: The network analysis results of the states-network

State	Regional Reporter	Weighted Indegree	Weighted Outdegree	Weighted Degree	PageRank	Modularity Class
AK	P.	55672	61507	117179	0.014869	1
AL	So.	115118	116752	231870	0.017554	3
AR	S.W.	113641	112839	226480	0.015971	2
AZ	P.	43930	42171	86101	0.022493	1
CA	P.	140271	127861	268132	0.044185	1
CO	P.	63313	63477	126790	0.023229	1
CT	A.	109520	119487	229007	0.017043	0
DE	A.	82201	83039	165240	0.012497	0
FL	So.	180579	165319	345898	0.045922	3
GA	S.E.	87777	83849	171626	0.015458	3
HI	P.	32970	36604	69574	0.010847	1
IA	N.W.	95724	100681	196405	0.024045	2
ID	P.	50582	53390	103972	0.016717	1
IL	N.E.	58906	51864	110770	0.031918	3
IN	N.E.	77448	76120	153568	0.01447	3
KS	P.	92188	91578	183766	0.022416	1
KY	S.W.	47174	45320	92494	0.015462	3
LA	So.	110855	108738	219593	0.017758	3
MA	N.E.	110984	106616	217600	0.030961	0
MD	A.	111660	119924	231584	0.025218	3
ME	A.	49082	51052	100134	0.013458	0
MI	N.W.	44741	36803	81544	0.025596	3
MN	N.W.	69009	64782	133791	0.027694	2
MO	S.W.	52584	52793	105377	0.016104	3
MS	So.	127785	129886	257671	0.014186	3
MT	P.	76387	77137	153524	0.013826	1
NC	S.E.	81756	79848	161604	0.020504	3
ND	N.W.	65723	70865	136588	0.011386	2
NE	N.W.	92492	95820	188312	0.018422	2
NH	A.	50861	55469	106330	0.015174	0
NJ	A.	141196	138331	279527	0.039165	3
NM	P.	52057	52619	104676	0.020598	1
NV	P.	30461	33236	63697	0.012846	1
NY	N.E.	15574	6554	22128	0.026241	3
OH	N.E.	90758	88822	179580	0.022368	3
OK	P.	92086	94921	187007	0.018687	1

State	Regional Reporter	Weighted Indegree	Weighted Outdegree	Weighted Degree	PageRank	Modularity Class
OR	P.	30531	17845	48376	0.030746	3
PA	A.	86210	80888	167098	0.027037	3
RI	A.	60173	63014	123187	0.011801	0
SC	S.E.	33216	33367	66583	0.011506	3
SD	N.W.	46317	53363	99680	0.010487	2
TN	S.W.	45845	48216	94061	0.01242	3
TX	S.W.	66466	68824	135290	0.015599	3
UT	P.	43244	48119	91363	0.014866	1
VA	S.E.	34535	33417	67952	0.012861	3
VT	A.	41300	45699	86999	0.010434	0
WA	P.	88135	84346	172481	0.030536	3
WI	N.W.	69868	62233	132101	0.033915	3
WV	S.E.	67838	82412	150250	0.011869	3
WY	P.	60690	69616	130306	0.010633	1

Centrality and influence. Two types of centralities can be applied to a clique network to get meaningful results: degree centrality (DC) and page-rank centrality (PRC). Other types like closeness centrality and betweenness centrality are meaningless, since all the nodes are connected to each other and have the same CC and BC scores.

Degree centrality (DC). In the states-network, the weighted in-degree is more suitable to use for DC. That's because the state that received more citations is the state that made most of the judgments that judges cited and used to make their opinions; hence, it has more influence on other states.

Based on the weighted in-degree of all the nodes (Table 12), we can see that Florida is the most influential state in the network, while New York is the least influential state. Table 13 shows the top ten influencer states based on the degree centrality score.

Table 13: The top ten states based on the degree centrality score

State	regional reporter	weighted indegree	weighted outdegree	Weighted Degree
FL	So.	180579	165319	345898
NJ	A.	141196	138331	279527
CA	P.	140271	127861	268132
MS	So.	127785	129886	257671
AL	So.	115118	116752	231870
AR	S.W.	113641	112839	226480
MD	A.	111660	119924	231584
MA	N.E.	110984	106616	217600
LA	So.	110855	108738	219593
CT	A.	109520	119487	229007

PageRank centrality (PRC). As mentioned earlier, PRC ranks nodes based on the importance of their neighbors. Table 14 shows the top 11 influencer states based on the PRC score. I added the eleventh influencer state because it is New York, which is the state with the lowest number of cases and citations (as mentioned earlier). Even with its low number of citations, New York ranked by PRC as the eleventh influencer state, since it has been in the four most-cited states by the top three network hubs that are considered important nodes in the network, as shown in Table 15. Moreover, the top three states in the DC and PRC results are the same states. That’s because these three states are the most cited in the network and they are in the top four cited states by each other, as shown in Table 15.

Table 14: The top 11 states based on the page-rank centrality score

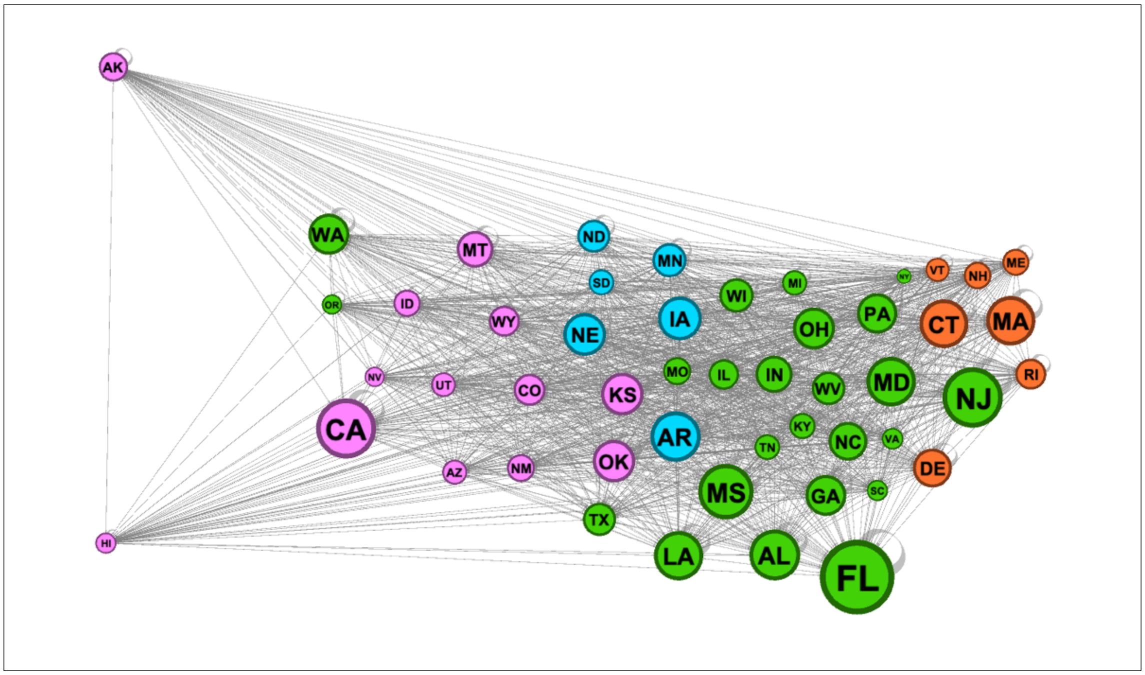
State	Regional Reporter	PageRank
FL	So.	0.045922
CA	P.	0.044185
NJ	A.	0.039165
WI	N.W.	0.033915
IL	N.E.	0.031918
MA	N.E.	0.030961
OR	P.	0.030746
WA	P.	0.030536
MN	N.W.	0.027694
PA	A.	0.027037
NY	N.E.	0.026241

Table 15: The top three hubs in states-network and the top four cited states by them with the number of citations

Source	Target	Citations
FL	FL	160698
FL	NJ	248
FL	CA	230
FL	NY	201
CA	CA	121590
CA	NJ	424
CA	FL	267
CA	NY	240
NJ	NJ	126064
NJ	FL	1211
NJ	NY	993
NJ	CA	699

Figure 23: The states-network.

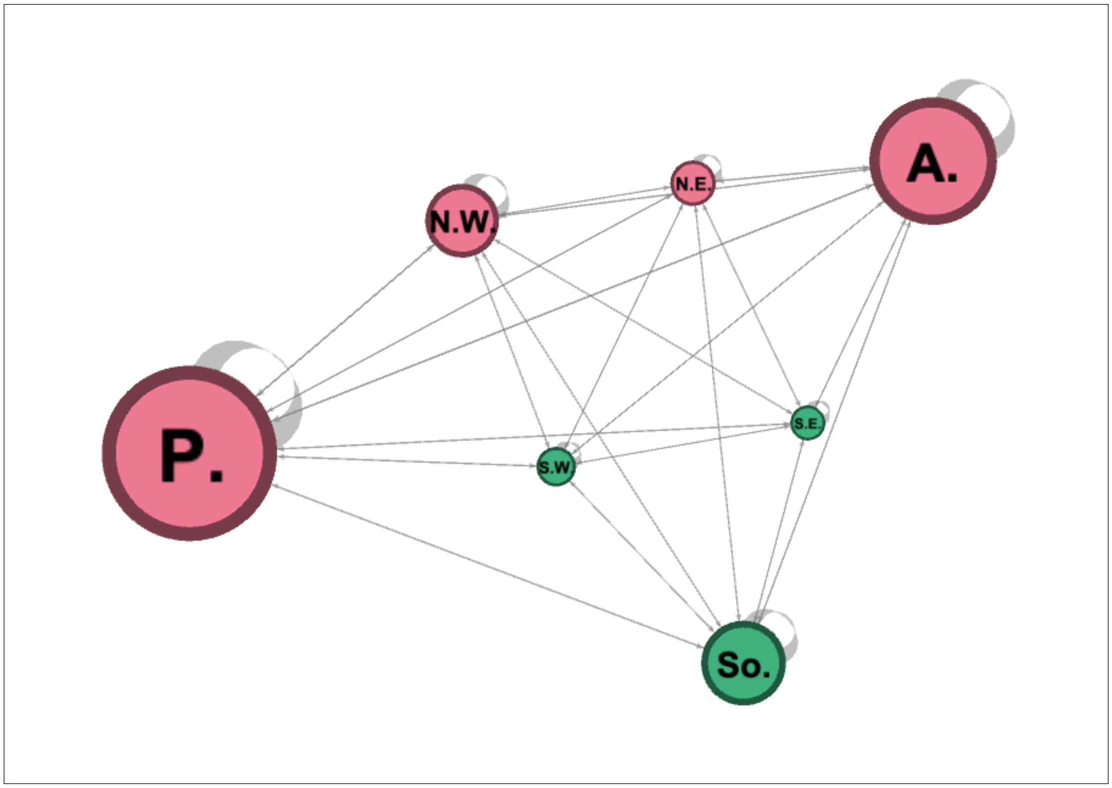
Nodes are positioned based on the state location in the USA map. The colors represent the communities and the size represent the degree centrality (DC)



Network Communities. By applying the modularity-optimization approach introduced by Newman (2004) using the algorithm of Blondel, Guillaume, Lambiotte, and Lefebvre (2008) to the states-network, four communities were detected. The Modularity Class column in Table 12 and the colors of the nodes in Figure 23, show the four communities and their nodes. In Figure 23, the nodes have been positioned based on their approximate locations on the USA map. Based on the colors and the location of the nodes, we can see that states tend to cite their neighbor states more than other states, except for Washington and Oregon. These two states interact with the states in the south and eastern regions more than with their neighbors in the western region. The four communities detected are not surprising, since neighboring states mostly have similar cultures, rules, and issues. Also, we can see that based on the nodes' size (which reflects the DC score of the node), each community has hubs and small-degree nodes like any real-world network. The cultural and political impact on the citation behavior can be observed clearly in the network communities if the network represents the cases by their regional reporter, as shown in Figure 24.

Regions-network

Figure 24: The regions-network.
Nodes are positioned based on the approximate regions' locations on the USA map.



This network has been projected from the cases-network based on the region of the cases' regional reporter where the court that decided the case is located. There are seven regions for the regional reporter, which divides the SCOLR into seven groups. Table 16 shows more details about the regions. In this network, I am looking to compare the in-region citations vs. the out-region citations. Also, I am looking to examine the influence of the regions on each other and how they are clustered based on the edges of the network.

The regions-network is a weighted direct network with a self-loop. It has N=7 nodes that represent the seven regions, and L= 49 direct edges represent the citations between the cases of the region to other regions' cases. The edges' weight represents the number of citations between the cases of each region. Each node's self-loop represents the citations between the cases of the same node's region. Also, the regions-network is a clique, like the states-network.

Table 16: The seven (7) regional reporters as part of the National Reporter System (Law, 2021)

Region name	Abbreviations	Description/Coverage
Atlantic Reporter	A. A.2d A.3d	Regional reporter containing cases from CT, DE, ME, MD, NH, NJ, PA, RI, VT; printed in three series.
North Eastern Reporter	N.E. N.E. 2d	Regional reporter containing cases from IL, IN, MA, NY, OH; printed in two series.
North Western Reporter	N.W. N.W. 2d	Regional reporter containing cases from IA, MI, MN, NE, ND, SD, WI; printed in two series.
Pacific Reporter	P. P.2d P.3d	Regional reporter containing cases from AK, AZ, CA, CO, HI, ID, KS, MT, NV, NM, OK, OR, UT, WA, WY; printed in three series.
South Eastern Reporter	S.E. S.E. 2d	Regional reporter containing cases from GA, NC, SC, VA, WV; printed in two series.
Southern Reporter	So. So. 2d	Regional reporter containing cases from AL, FL, LA, MS; printed in two series.
South Western Reporter	S.W. S.W. 2d S.W. 3d	Regional reporter containing cases from AR, KY, MO, TN, TX; printed in three series.

The nodes' degrees. As mentioned earlier, the regions-network is a weighted, directed, and fully connected network. The average degree $<k>$ of the network's nodes is (14), since each node has an incoming and outgoing edge to all 7 network nodes, including itself. All nodes have the same degree: $k = 14$. Also, all nodes have the same out-degree: k_{out} = in-degree k_{in} =7.

The average weighted degree of the regions-network's nodes is 526,776.14. The nodes' weighted degree k ranges from 618,015 to 1,906,944. The nodes' weighted out-degree k_{out} ranges from 312,893 to 954,427, and the nodes' weighted in-degree k_{in} ranges from 305,122 to 952,517.

Table 17: The results of the regions-network analysis. The nodes are sorted based on the DC score

Region	weighted indegree	weighted outdegree	Weighted Degree	modularity class
P.	952517	954427	1906944	0
A.	732203	756903	1489106	0
So.	534337	520695	1055032	1
N.W.	483874	484547	968421	0
N.E.	353670	329976	683646	0
S.W.	325710	327992	653702	1
S.E.	305122	312893	618015	1

Influence and Degree Centrality (DC). Based on the citation behavior of the SCOLR system and the regions-network properties, the weighted in-degree is the best measurement for centrality and influence in the regions network. As mentioned earlier, the weighted in-degree represents the number of citations the node received. Table 17 shows the analysis results of the regions-network. The table shows the nodes sorted based on the weighted in-degree values. We can see from the results that the node of the P. region has the highest weighted in-degree value, which means it is the most influential node on the network. If we look at Table 16 and Table 17, we can see that there is a positive correlation between the weighted in-degree values and the number of states that each region includes. However, the So. region, which has four states, has a higher number of in-citations than N.W. and N.E., which have more states. That is because So. includes Florida, which is the state with the highest weighted in-degree in the network.

Network Communities. By applying the modularity-optimization approach introduced by Newman (2004) using the algorithm of Blondel, Guillaume, Lambiotte, and Lefebvre (2008) to the regions-network, two communities were detected. In Figure 24, the nodes have been colored based on their communities and positioned based on their approximate locations on the U.S. map. From Figure 24, we can see that the first community colored with green has all the southern regions, while the other community includes all the other regions. By looking at the states that are covered by each region, I can conclude that the history of the regions in the civil war has an impact on the citation behavior between them. So., S.W., and S.E. include all the Confederate States of America and the border states in the civil war of the U.S., while the other regions include the remaining states.

Years-network

This network has been projected from the cases-network based on the "year of the decided date" property of the cases. This network will help to discover changes in the citation behavior over time. The main goal for analyzing the years-network is to test the third hypothesis of this dissertation.

> ***Hypothesis 3:*** The emergence of the internet and the technology availability in the court increased the number of citations in the courts' opinions because judges could easily search across geographic regions.

The years-network is a weighted direct network with a self-loop. It has N=58 nodes that represent the years from 1953 to 2010, and L= 1711 direct edges represent the citations between the cases of the year to other years' cases. The edges' weight represents the number of citations between the cases of each year. A node's self-loop represents the citations between the cases of the same year. This network forms a tree with no loops because the node can have an out-going edge to itself or nodes of older years. If we look at the years-network as an undirected network, the network will form a clique, since each node cited itself and all the older nodes, and each node has been cited by itself and all the younger nodes. So, every node is connected to every other node.

The nodes' degrees. The average degree $<k>$ of the years-network's node is 29.5. The nodes' degree k ranges from 1 to 59. The nodes' out-degree k_{out} and in-degree k_{in} range from 1 to 58.

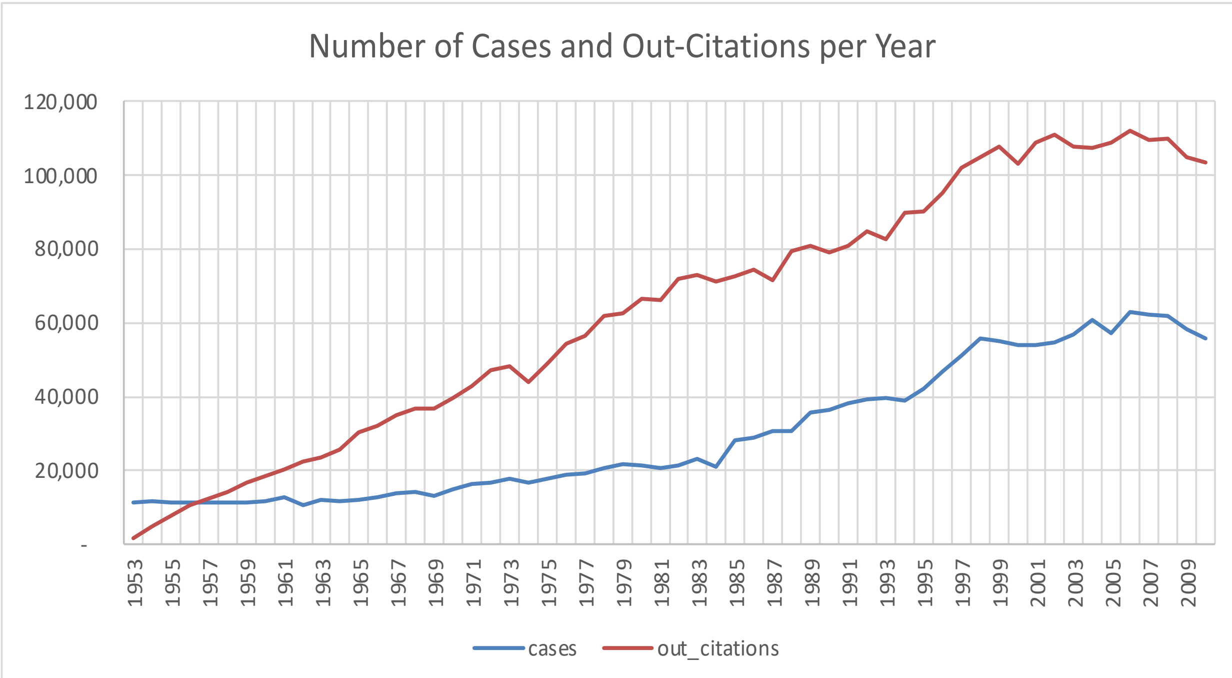
The average weighted degree of the years-network is 63,576.43. The nodes' weighted degree k ranges from 52,386 to 186,542. The nodes' weighted out-degree k_{out} ranges from 1,741 to 112,170. The nodes' weighted in-degree k_{in} ranges from 5,209 to 97,945. Table 18 shows all the degrees and the weighted degrees of all the nodes in the network.

Table 18: The years-network nodes' degrees and weighted degrees.

Year	In-degree	Out-degree	Degree	Weighted In-degree	Weighted Out-degree	Weighted Degree
1953	58	1	59	52689	1741	54430
1954	57	2	59	47508	4878	52386
1955	56	3	59	46044	7922	53966
1956	55	4	59	45355	10542	55897
1957	54	5	59	44473	12649	57122
1958	53	6	59	42620	14281	56901
1959	52	7	59	43765	16686	60451
1960	51	8	59	44557	18635	63192
1961	50	9	59	43951	20227	64178
1962	49	10	59	44478	22329	66807
1963	48	11	59	46939	23572	70511
1964	47	12	59	46975	25657	72632
1965	46	13	59	56956	30499	87455
1966	45	14	59	51402	32067	83469
1967	44	15	59	56063	35033	91096
1968	43	16	59	55780	36941	92721
1969	42	17	59	55102	36673	91775
1970	41	18	59	56547	39597	96144
1971	40	19	59	58760	42771	101531
1972	39	20	59	62628	47108	109736
1973	38	21	59	66868	48332	115200
1974	37	22	59	66195	44089	110284
1975	36	23	59	75962	49099	125061
1976	35	24	59	81347	54437	135784
1977	34	25	59	78982	56616	135598
1978	33	26	59	83173	61942	145115
1979	32	27	59	88131	62698	150829
1980	31	28	59	97945	66444	164389
1981	30	29	59	82610	66012	148622
1982	29	30	59	88275	71742	160017
1983	28	31	59	90171	72806	162977
1984	27	32	59	84548	71271	155819
1985	26	33	59	87304	72504	159808
1986	25	34	59	81502	74359	155861
1987	24	35	59	79416	71525	150941
1988	23	36	59	80507	79367	159874
1989	22	37	59	77580	81008	158588

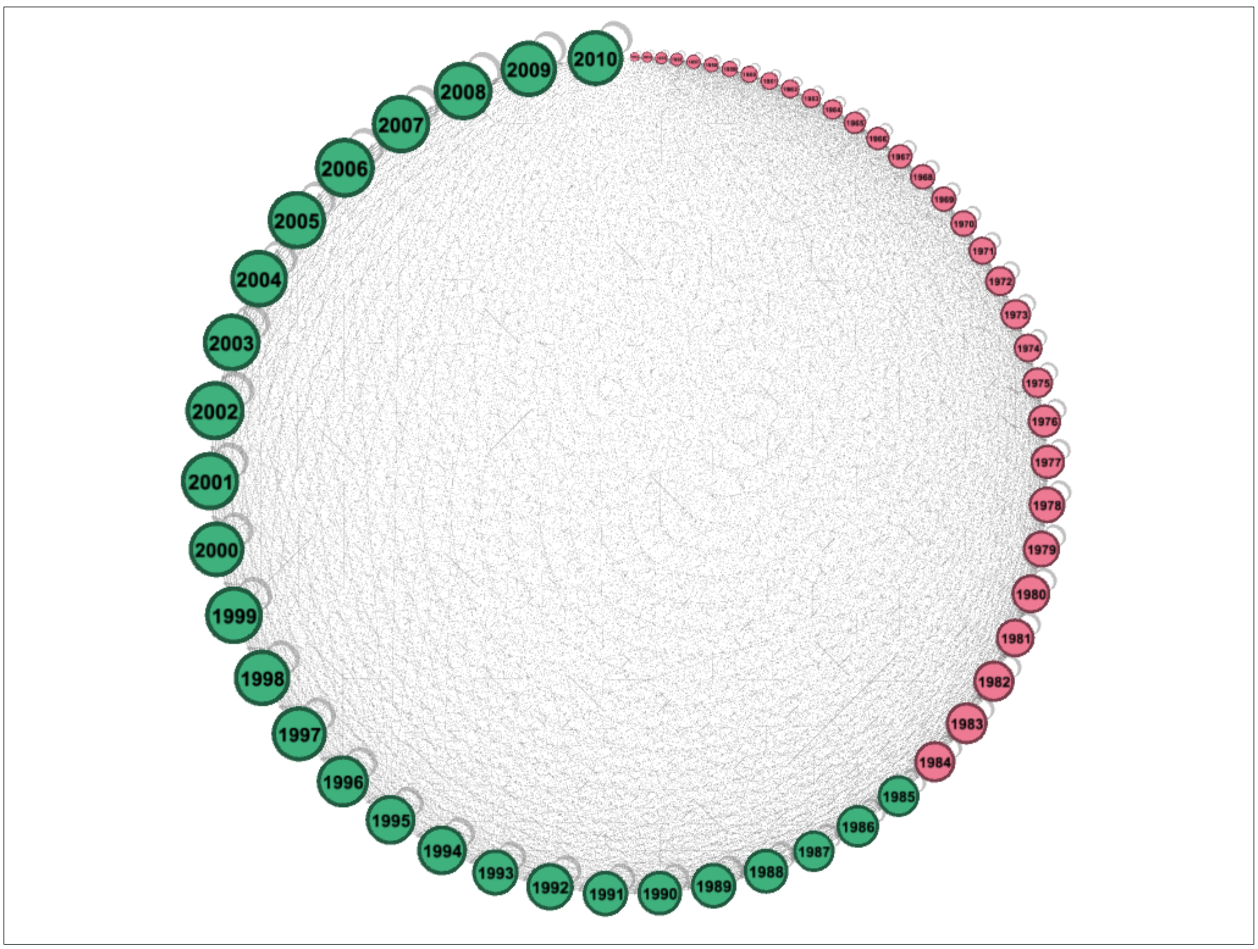
Year	In-degree	Out-degree	Degree	Weighted In-degree	Weighted Out-degree	Weighted Degree
1990	21	38	59	77136	78912	156048
1991	20	39	59	78951	80718	159669
1992	19	40	59	82044	84964	167008
1993	18	41	59	81759	82661	164420
1994	17	42	59	77081	89933	167014
1995	16	43	59	87566	90174	177740
1996	15	44	59	81352	95321	176673
1997	14	45	59	80733	101935	182668
1998	13	46	59	81556	104986	186542
1999	12	47	59	77202	107695	184897
2000	11	48	59	69960	103050	173010
2001	10	49	59	63716	108812	172528
2002	9	50	59	62018	110894	172912
2003	8	51	59	57948	107601	165549
2004	7	52	59	52080	107512	159592
2005	6	53	59	49079	108808	157887
2006	5	54	59	43007	112170	155177
2007	4	55	59	36764	109327	146091
2008	3	56	59	30212	109729	139941
2009	2	57	59	18952	104842	123794
2010	1	58	59	5209	103330	108539
Total	1711	1711	3422	3687433	3687433	7374866
Maximum	58	58	59	97945	112170	186542
Minimum	1	1	59	5209	1741	52386
Average	29.5	29.5	59	63576.43	63576.43	127152.86

Figure 25: Number of Cases and Out-Citations per Year



Years, cases, and citations correlation. Figure 25 and Figure 26 show clearly that there are positive correlations between the years, the number of cases that have been reported, and the number of citations that have been made. In my opinion, these correlations are results of incremental improvements in the SCOLR resources that include the increased adoption of technology and the availability of different resources to support the judges and the courts' activities.

Figure 26: The years-network.
Nodes are positioned based on the years' order. The colors represent the communities, and the sizes represent the number of out-citations of the nodes.



The network communities. By applying the modularity-optimization approach introduced by Newman (2004) using the algorithm of Blondel, Guillaume, Lambiotte, and Lefebvre (2008) to the years-network, two communities were detected. In Figure 26, the nodes have been colored based on their communities and ordered in a circle based on the years. From Figure 26, we can see clearly that the nodes have been divided into two communities, with the division occurring between the years 1984 and 1985. The first community has the years 1953-1984, and the second community has the years 1985-2010.

The years 1984 and 1985 are turning points in the history of technology. At this time, Apple introduced the Macintosh personal computer, Microsoft released the first Windows operating system, IBM released the PC Jr. and the PC/AT, Dell created PC's Limited and produced its first computer of its own design, Phillips introduced the CD-ROM, and IBM introduced its new 3480 cartridge tape system that could hold up to 200MB (*Computer History Museum*, n.d.). All of these inventions are related to storing data and accessing it in an easier way than using hardcopies.

Since the supreme courts are among the highest levels of government organizations, I assume that they adopted the technology from an early age. Using the technology and starting digital reporting affected the citation behavior. It led to creating more interaction and citations between the cases that have been decided in and after 1985. That could be the reason behind clustering the nodes in the years-network into two communities, as mentioned earlier.

Summary

In this chapter, I have confirmed that modeling the legal system of SCOLR as a complex system led to better understanding of the system's component behavior and interactions. I also proved the capability of the network-science approach to model and analyze the complexity of the SCOLR system to answer many unsolved questions about the system's behavior.

The SCOLR system has been modeled as a network based on the cases' citations. Then, four other networks between the cases variables were projected, to study the system from different aspects. This led to having five networks: a cases-network, a judges-network, a state-network, a regions-network, and a years-network.

The five networks were analyzed using a network-science approach. Different algorithms and methods have been applied to the networks to discover the network properties, including the degree distribution, the network centralities, robustness, correlation, and network communities. In the remaining part of this section, I will summarize the most important and interesting results based on these properties.

The degree distribution of the SCOLR system satisfies the power-law distribution similar to most real-world systems. The power-law distribution in the cases-network and the judges-network showed that most of the networks' nodes are small-degree nodes, and few nodes are high-degree nodes (hubs). Such a finding proves that the networks were not randomly formed.

Network centrality is another network property that has been discovered in the system networks. Network centrality shows the degree of influence of the network nodes. Different network centralities have been used for each network, based on the network characteristics and the information needed. For example, degree centrality in the judges-network ranks the judges based on the nodes' weighted in-degree, which is the number of citations they received. This centrality measurement helps to show the most influential judges in SCOLR who received the highest number of citations. One of the interesting results from studying the centralities is that Florida is the most influential state in the legal system, based on degree centrality and page-rank centrality. Additionally, the analysis of the judges-network and the states-network showed that the networks' hubs that are considered the most influential nodes tend to cite themselves more than others.

Moreover, the analysis presented a neutral-degree correlation, in general, in the judges-network. This means that judges tend to cite other cases regardless of the popularity of the opinion writers of the cited cases, which is considered as good practice in the legal system.

Finally, network communities were extracted in all the created networks. The community-detection algorithms of the modularity-optimization function resulted in the most interesting outcomes in this analysis. In the cases-network, 2,868 communities were detected; 24 of the communities were large communities that absorb more than 55% of the nodes, while the rest of the communities were very small. By consulting domain experts, I found that each of the 24 communities represents a group of cases that share the same legal issues. For example, the largest community has more than 100,000 nodes, and most of these cases are in the area of workers' compensation. The second-largest community has the cases in the area of real estate. There were 11 communities detected in the judges-network, as shown in Figure 21. The judges in each community are the more experienced judges in a legal-issue area. However, these findings have not been validated, since we couldn't find a labeled dataset of the cases for comparison. In the states-network and the regions-network, the detection of communities showed that there are some cultural and political influences on the network that led states and regions to interact with their neighbors more than with others.

Chapter 5: Classifying the Citation Treatment in the Cases of State Courts of Last Resorts

Citation behavior plays an important role in the permanence and revision of the law. Judges do not always agree with the outcomes of previous cases related to their cases. A judge may adopt the position of the law from other cases by citing those cases, thereby showing his/her acceptance and agreement with the interpretation, or he/she may reject that interpretation and come to a different one.

The citation treatment affects the extension and the history of the law regarding the case. For example, if a certain case receives negative treatment (such as being overruled) in subsequent cases, it is an indication that the law might not be applicable anymore. Therefore, judges and attorneys need to keep track of the history of laws and how particular laws may have changed. Citation treatment assists judges and everyone who works with them (such as attorneys) at carefully studying cases.

In addition, considering the citation treatments in network creation and analysis will assist us in discovering valuable, interesting information. To advance our understanding of judicial citation behavior, we should add the citation treatment to the system network analysis.

Classifying citation treatments is a subjective task. Moreover, only a few works have been done in automatic recognition of case treatments. Unfortunately, as discussed in Chapter 1, those works were done on some judicial institutions, but not SCOLR. Therefore, there is a need to model and develop an automated tool to extract and classify the citation treatments of the cases of SCOLR.

In this study, we built a model to create an automated dictionary-based classifier that classifies citation treatments as positive or negative, based on agreement or disagreement between the cases. The classifier was designed using NLP and machine-learning (ML) concepts, including (but not limited to) text-preprocessing techniques, sentiment analysis, and classification algorithms.

In general, not only in the legal system, most citations are positive or discussion citations, since they have been added to support the argument and outcome of the case. However, there are some negative citations brought up to show different opinions and new interpretations of the law. Therefore, our study scope is to identify the negative citation treatments in the majority opinions.

In the following sections, we introduce the study motivations, methodologies, and results. Then we conclude with our summary, conclusions, and next steps.

This study was conducted by Ali Al-Madan and myself, under the supervision of Dr. Samira Sheikh and Dr. Jason Windent.

Motivations

These were the motivations for this study: 1) Create a labeled dataset of the citations. 2) Design and develop an automated model to extract and classify the citations. 3) Add the citation treatment to the citation network, and update the analysis accordingly.

Steps and Methodologies

We conducted this study using the following methodologies:

1) Extract the citations from the majority opinion with the citation paragraph. To perform this task, we used the citation-extraction tool that was introduced in Chapter 4. However, we updated the tool to extract the citations and the citation paragraphs.
2) Create a sample of the citations and their paragraphs, to be annotated by experts.

3) Apply some of the well-known text-classifier algorithms to classify the citations. We selected five machine-learning algorithms to perform this task. The algorithms were selected based on their complexity and accuracy.
4) Build an automated dictionary-based classifier based on NLP concepts, to extract and classify the citations of SCOLR cases.

Dataset

In this study, we used the dataset introduced in Chapter 3, which consists of SCOLR cases from all 50 states in the U.S. We sampled the data by focusing on cross-state citations for all the cases during the period between 1990 to 2010 that had all the required data for the analysis (approximately 132,153 citations).

After creating the study dataset, graduate students from the department of political science at the University of North Carolina at Charlotte (UNCC) helped to annotate the data and label the citation treatments. We were provided with 454 citations and their treatments (183 negative treatments and 271 positive treatments).

To prepare the dataset for the classification algorithms, we divided the dataset into training (77%) and testing (23%) datasets. Before training our models, we performed preprocessing steps that are widely used in natural language processing. These steps include converting the citation paragraph to lowercase, removing numbers, punctuation, non-English words, and stop-words, and applying the stemming function. Stemming in NLP is the process of reducing the words to their stem by removing the suffixes and prefixes and returning the stem of the word. Stemming is a useful step in dealing with textual data analysis because it changes the different inflectional forms of a word to a single stem of the word and gives all of them the same treatment.

We have studied the effects of each of these preprocessing steps individually on the accuracy of our models. We found that removing the stop words negatively affected the accuracy of the Treatment-Classifier model, but it improved the accuracy of the other models. Therefore, we removed the stop words for all models except the Treatment-Classifier model. We found that stop words such as “but” and “no” are important in assigning treatments to citations and assisted the Treatment-Classifier model to identify some of the negative treatments.

Classification Algorithms and Results

To classify the citations, we identified and selected seven well-known algorithms in text classification. The selection was based on the complexity and accuracy of the algorithms in text classification. These algorithms are categorized as dictionary-based and supervised ML algorithms. Our focus in this study was on binary classification for the positive and negative classes. These are the algorithms used for classification:

- Dictionary-based classifications:
  - Bluebook classification
  - Sentiment Analysis
- Supervised machine learning classifications:
  - Multinomial Naive Bayes
  - Logistic Regression
  - SGD
  - Linear SVC
  - NuSVC

As you will see in the following discussion, none of the algorithms performed well on the classification of the citation treatment. Consequently, we designed and built a new dictionary-based classifier that we named *Treatment-Classifier*. In this section, we describe the classifying models that have been used or built in this study, and we discuss their results in detail.

Bluebook Classification

The Bluebook, a guide that is used for the standards of legal text and citations, has recommended ten keywords to be used for cases' citations by the judges in their opinions. However, judges do not follow the Bluebook recommendation on citations, which makes it difficult to detect citation treatments.

The Bluebook classification is a dictionary-based approach. In this classification task, we built a dictionary of keywords using the ten treatment words that have been recommended by the Bluebook for use in citations. Five of the treatment words are used to positively cite other cases (such as "accord") and three of the treatment words are used to negatively cite other cases (such as "contra"). Table 19 shows the keywords that are suggested for judges to use in their opinions.

Table 19: Keywords recommended by the Bluebook for citation treatments

Keywords for positive citations	Keywords for comparison/discussion citations	Keywords for negative citations
See	Compare	Contra
See e.g		But see
See also		But cf
Accord		
Cf		

A binary dictionary-based classifier was built to classify the citations based on the keywords. The classifier simply searches for negative keywords in a citation paragraph. If it finds at least one negative keyword in the text, it classifies the citation as negative; otherwise, it classifies it as positive, since that is the default in citations.

Table 20: The accuracy of results for the Bluebook classifier

	Precision	Recall	F1-Score	Support
Negative	0%	0%	0%	183
Positive	60%	100%	75%	271
Accuracy			60%	454
Macro Avg	30%	50%	37%	454
Weighted Avg	36%	60%	45%	454

The Bluebook classifier did not perform well to detect negative treatments. The recall score for our main class (negative class) was zero. Therefore, none of the Bluebook treatment words were found in the text, and none of the negatively-treated citations were assigned negative treatment by the classifier. This is an indication that judges do not follow the Bluebook recommendations.

Table 20 shows the accuracy results for the Bluebook classifier.

Sentiment Analysis

Sentiment analysis is contextual mining to detect and classify opinions expressed and emotion in texts. It can be a multi-class classification or just a binary classification (positive or negative). In the context of citations, sentiment analysis can be used to capture the attitude of the judges toward the interpretation and application of the law on the cited cases.

We applied sentiment analysis as a binary classification. We used the library *TextBlob* with the function *sentiment* to get the polarity.

Table 21: The accuracy of results for the sentiment analysis classifier

	Precision	Recall	F1-Score	Support
Negative	42%	40%	41%	183
Positive	61%	62%	62%	271
Accuracy			54%	454
Macro Avg	51%	51%	51%	454
Weighted Avg	53%	54%	53%	454

Sentiment analysis did not give good results. We believe that the legal language has its unique terms and meanings, and therefore using a general sentiment analysis tool such as TextBlob could not identify negative sentiments correctly.

Table 21 shows the accuracy of the classifier. From the table, we can see that only 40% of the negative citations were captured by the algorithms, and only 42% of the citations classified as negative were negative citations.

Multinomial Naive Bayes

The Multinomial Naive Bayes algorithm is one of the widely used supervised ML algorithms in text-classification problems. It is fast and simple to use (Géron, 2017). It works based on the concept of the bag of words. The bag of words is an NLP approach in which the algorithm creates a set of words or vectors and their frequency of appearing in the predefined classes in the training data. Then it can be used to classify new texts.

Table 22: The accuracy of results for the Multinomial Naive Bayes classifier

	Precision	Recall	F1-Score	Support
Negative	22%	25%	24%	32
Positive	65%	61%	63%	72
Accuracy			50%	104
Macro Avg	43%	43%	43%	104
Weighted Avg	52%	50%	51%	104

We applied the Multinomial Naïve Bayes algorithm on the citations paragraphs and built word vectors for both training and testing data. As in all our machine-learning classifications, we removed stop words. We also considered unigrams, bigrams, and trigrams.

Table 22 shows the accuracy of the model. The recall and the precision of the negative class were very low: 25% of the negative citations were classified as negative citations, while 78% of the citations classified as negative were positive citations.

Logistic Regression

Logistic regression, or logit regression, is a supervised machine-learning algorithm that uses probability to predict the classification. For example, it assigns a negative treatment to a citation when the probability of the citation belonging to the negative class is more than 50% (Géron, 2017). Logistic regression is used for binary classification, which makes it suitable for our task. The results for the classifier are shown in Table 23. We can see from the results that the recall for the negative class is 25%, which means that the algorithm failed to capture 75% of the negative citations. Moreover, the precision of the negative class is 28%. That means that only 28% of the citations classified as negative are correct, and 72% are incorrectly classified.

Table 23: The accuracy of the results for the Logistic Regression classifier

	Precision	Recall	F1-Score	Support
Negative	28%	25%	26%	32
Positive	68%	71%	69%	72
Accuracy			57%	104
Macro Avg	48%	48%	48%	104
Weighted Avg	56%	57%	56%	104

Linear Support Vector Classification (Linear SVC)

Support vector classification is one of the most popular machine-learning algorithms for binary classification. It simply converts the feature of the training data to vectors in multidimensional space, then adds a line called a hyperplane to separate the vectors into two classes. SVC can be a linear classifier or a nonlinear classifier. However, in NLP problems, using SVC as a linear classifier gives better results.

Linear SVC is a faster implementation of SVC. It is a linear support vector classifier that scales better than SVC with different numbers of samples. Linear SVC gives great classifications for many NLP problems. Unfortunately, it didn’t do a great job at classifying the citations. It was able to classify 31% of the negative treatments correctly. The results for the classifier are shown in Table 24.

NuSVC

According to the documentation of Scikit Learn, the Nu support vector classifier is similar to linear SVC, but it provides an option to specify the number of vectors. As shown in Table 25, the NuSVC field made better classifications than the previously applied algorithms.

Table 24: The accuracy of the results for the Linear SVC classifier

	Precision	Recall	F1-Score	Support
Negative	29%	31%	30%	32
Positive	68%	65%	67%	72
Accuracy			55%	104
Macro Avg	48%	48%	48%	104
Weighted Avg	56%	55%	55%	104

Table 25: The accuracy of the results for the NuSVC classifier

	Precision	Recall	F1-Score	Support
Negative	29%	28%	29%	32
Positive	68%	69%	69%	72
Accuracy			57%	104
Macro Avg	49%	49%	49%	104
Weighted Avg	56%	57%	57%	104

Stochastic Gradient Descent (SGD) Classifier

According to the documentation of Scikit Learn, the SGD Classifier is a linear classifier optimized by stochastic gradient descent (SGD). SGD optimization is used to increase the efficiency of the classifier. When the loss function is set to “log”, it acts like logistic regression. However, when the loss function is set to “hinge”, it acts like a support vector machine (SVC). In our case, we aim to improve the results of the SVC. So, we used loss function “hinge” to have an SVC classifier that has been optimized with SGD.

SGD optimization slightly improved the results of the SVC algorithm, as shown in

Table 26. It was able to capture 50% of the negative treatments, but 67% of negative classifications were misclassified.

Figure 27: the 1st-voting classifier structure

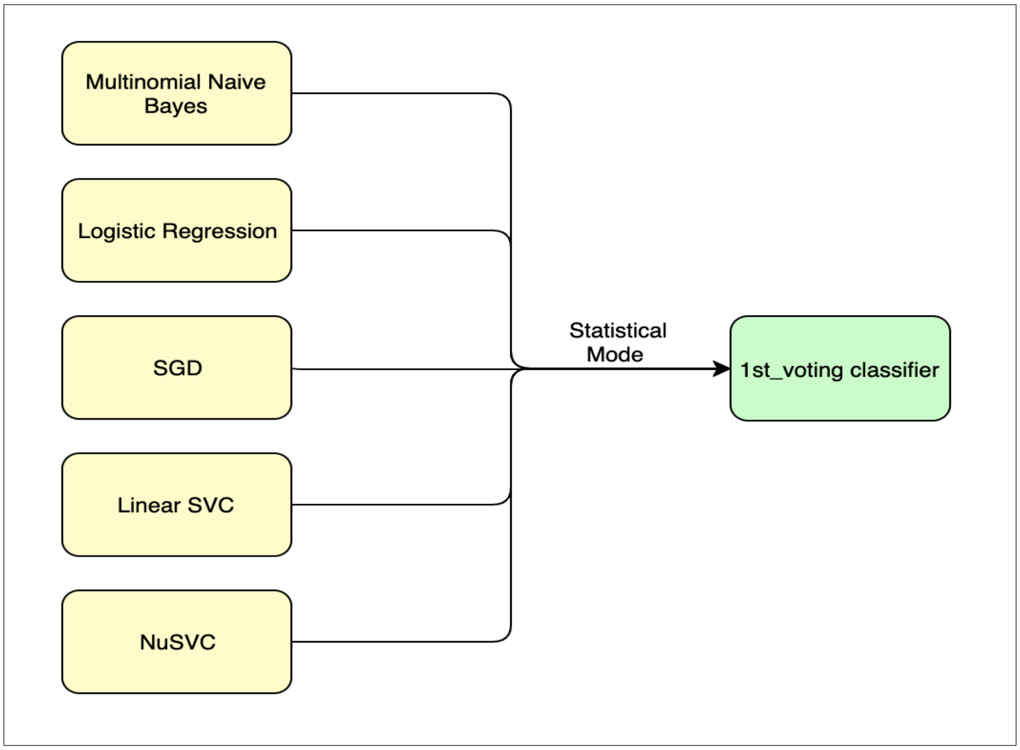


Table 26: The accuracy of the results for the SGD classifier

	Precision	Recall	F1-Score	Support
Negative	33%	50%	40%	32
Positive	71%	56%	63%	72
Accuracy			54%	104
Macro Avg	52%	53%	51%	104
Weighted Avg	60%	54%	56%	104

1st-voting Classifier

As we saw, none of the applied machine-learning algorithms performed well on the task of classifying citation treatments. Therefore, we built a voting classifier, which we called "1st-voting classifier", that aims to improve on the outcomes of the ML algorithms. The classifier was built by using the statistical mode function over the results of the five machine-learning classifiers mentioned before. The mode function returns the most frequent value in the input list. In our case, for each citation, the mode function will return the most frequent class among the results of the five ML algorithms. In other words, if a citation, for example, was treated positively by three machine-learning classifiers and negatively by two machine-learning algorithms, the 1st-voting classifier will assign a positive treatment to the citation. Figure 27 shows the structure of the 1st-voting classifier.

Unfortunately, the 1st-voting classifier couldn't improve the results of the combined algorithms. This indicates that most of the applied ML algorithms have similar results, but the results don't match the label of the citation. The accuracy of the 1st-voting classifier is shown in Table 27.

Table 27: The accuracy of the results for the 1st-voting classifier

	Precision	Recall	F1-Score	Support
Negative	28%	28%	28%	32
Positive	68%	68%	68%	72
Accuracy			56%	104
Macro Avg	48%	48%	48%	104
Weighted Avg	56%	56%	56%	104

Treatment-Classifier

Since none of the previous classifiers performed well at classifying citations, we built a novel dictionary-based classifier. In collaboration with legal domain experts, we built a dictionary of 300 keywords to distinguish negative citations from positive citations. Examples of words/phrases include "declined" and "rejected by the majority of the court". Similar to the Bluebook dictionary, in this classification task, the classifier gives the treatment to each citation based on the keywords included in the citation paragraph. For each negative keyword found in the text, it gave the citation a negative treatment (encoded as -1). Similarly, it gave a positive treatment (encoded as 1) for each positive keyword found. Then, it used the statistical mode function on the given treatments to decide the final treatment of the citation. Table 28 shows the accuracy of the classifier.

Using our keywords dictionary gave better results than all the previously applied classifiers. The recall of our model, the Treatment-Classifier, of the negative treatment was 80%, but the precision was only 42%. In other words, the Treatment-Classifier was able to correctly classify 80% of negative citation treatments. At the same time, only 42% of the negatively classified citations were negative, and 58% were incorrectly classified. On the other hand, the treatment classifier was not able to classify 75% of the positive citations correctly. The recall of the positive class was 25%. Therefore, the Treatment-Classifier needed some improvements to give better classification for the citation treatment.

Table 28: The accuracy of the results for the Treatment-Classifier

	Precision	Recall	F1-Score	Support
Negative	42%	80%	55%	183
Positive	65%	25%	36%	271
Accuracy	47%			454
Macro Avg	54%	53%	46%	454
Weighted Avg	56%	47%	44%	454

2nd-Voting Classifier

The classification accuracy of the Treatment-Classifier on the negative class was the highest among all the applied algorithms in this study. However, it was the worst application on the positive class. On the other hand, the classification accuracy of sentiment analysis and the 1st-voting classifier on the positive class were much better than their accuracy on the negative class. To come up with a better classifier, we built a new voting classifier, which we called "2nd-voting classifier", to classify citations based on the results of these three classifiers: Treatment-Classifier, sentiment analysis, and 1st-voting classifier.

The 2nd-voting classifier has a similar design as the 1st-voting classifier. It uses the statistical mode function over the classification results of the three mentioned classifiers on each citation. Figure 28 shows the design of the 2nd-voting classifier.

Unfortunately, the 2nd-voting classifier gave lower classification accuracy on the negative class than the Treatment-Classifier. Recall that the negative treatment is the scope of this study. Consequently, the 2nd-voting classifier was unsuccessful to improve the study outcome.

Table 29 shows the accuracy of the classifier.

Figure 28: Design of the 2nd-voting classifier.

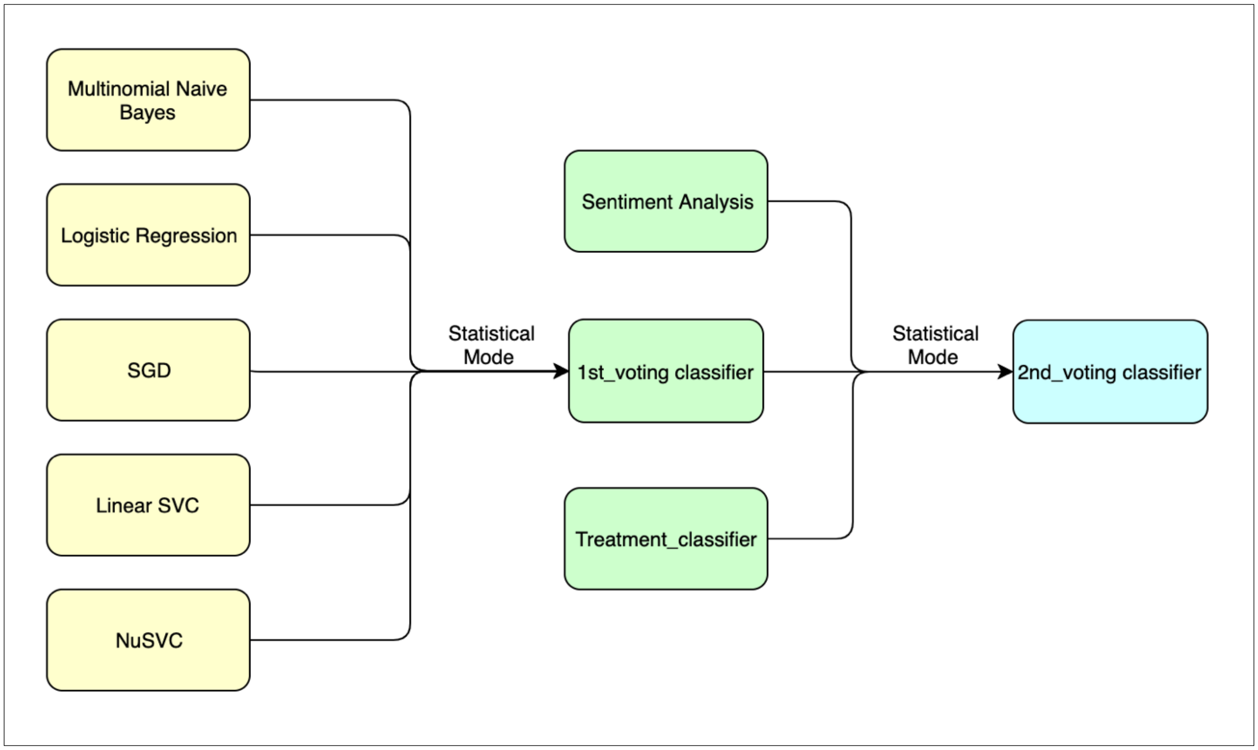


Table 29: The accuracy of the results for the 2nd-voting classifier

	Precision	Recall	F1-Score	Support
Negative	33%	47%	38%	32
Positive	71%	57%	63%	72
Accuracy			54%	104
Macro Avg	52%	52%	51%	104
Weighted Avg	59%	54%	56%	104

Summary

The main purpose of this study was to classify the citation treatments and more specifically to identify the negative treatments in SCOLR cases from a large number of unstructured text files. In this study, we updated our citation-extraction tool that was designed in Chapter Four to extract the citations and the citation paragraphs from the opinions of the cases. We performed a few classification tasks that we categorize as dictionary-based and based on machine learning. Moreover, we followed important preprocessing steps such as removing punctuation, digits, and stop words. Even though our Bluebook approach was not able to identify negative treatment (0% negative class recall), our Treatment-Classifier approach performed the best (80% negative class recall).

We applied five machine-learning algorithms (Naive Bayes, logistic regression, Linear-SVC, NuSVC, and SGD-Classifier) and sentiment analysis to classify the citations. However, these algorithms failed to classify the negative treatments well. Therefore, we built a new dictionary-based model, the Treatment-Classifier, to classify the citations. In our Treatment-Classifier model, we used a dictionary of keywords that was built in collaboration with our annotators. Since these classifiers did not perform as expected, we built two voting classifiers to improve the accuracy of the machine-learning classifiers. The highest recall we could get for the negative class was 47%.

Figure 29Figure 30 show the accuracy of results for all of our classifiers in terms of precision, recall, and F1 scores. From the results, we can see that building Treatment-Classifier (the dictionary-based model) was the best approach in this study to classify the negative treatments, which was the goal of the study.

Figure 29: Accuracy scores of the negative class (negative citations) for all the classifiers

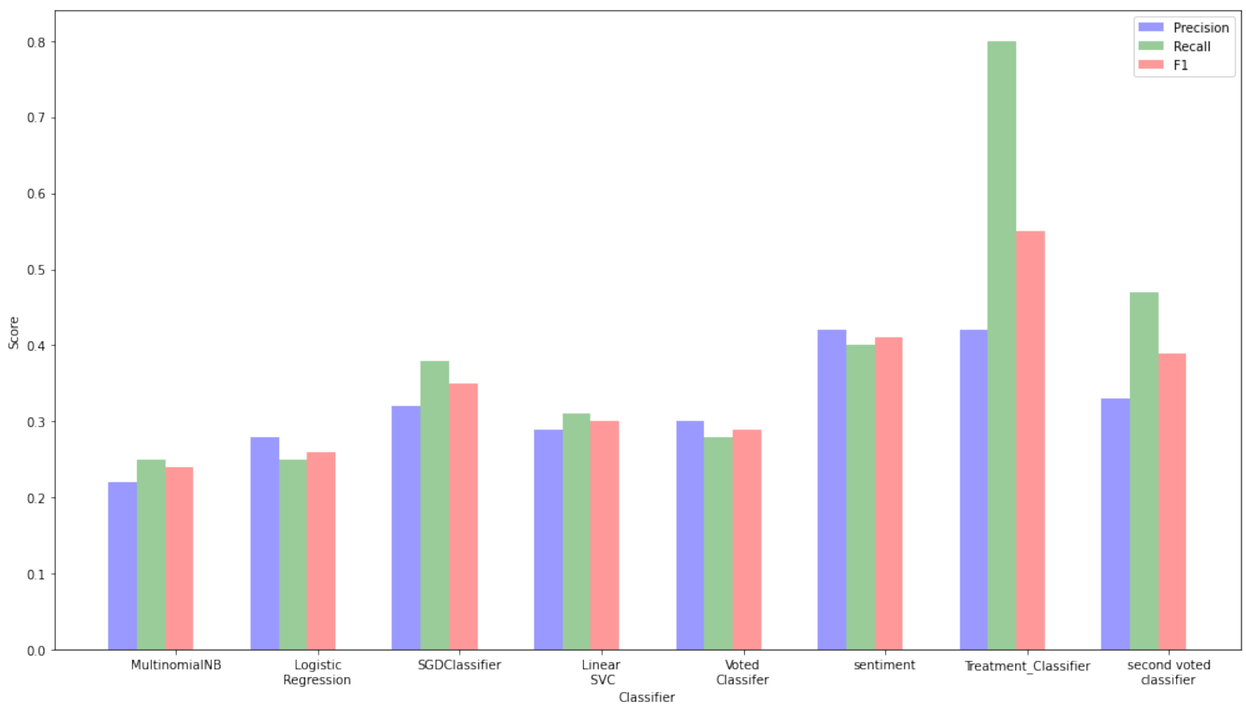
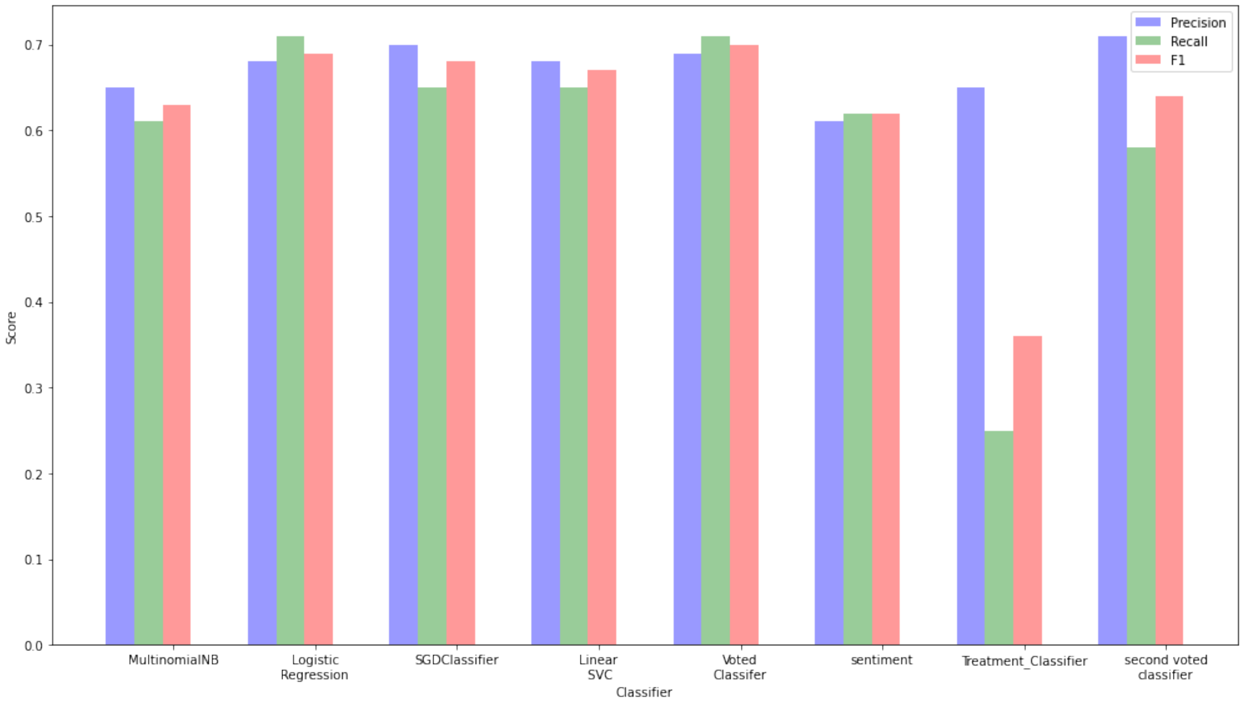


Figure 30: Accuracy scores of the positive class (positive citations) for all the classifiers



Chapter 6: Conclusion and Future Research Directions

The conventional IT infrastructures in many organizations do not have the capability for dynamic analysis. Yet, they contain vital insights for the organization-domain experts, social scientists, data scientists, and those interested in studying complex systems. SCOLR are examples of organizations with this significant issue. I reviewed numerous research studies on the United States legal systems. However, I found few studies regarding the complexity and behaviors of the SCOLR system.

Research on SCOLR suffers from two main limitations: a shortage of clean and analyzable data, and a lack of computational methods for utilizing the information and generating useful insights. To address these issues, the main goal for this dissertation is to design a framework to understand the dynamics of a problem with unstructured and noisy textual documents, which inter-cite each other. I am particularly focusing on the judiciary cases of SCOLR as one of the real-world complex systems, but the framework can be adapted to other problems with similar data issues.

This dissertation consists of three main studies: creating a novel dataset for the cases of the SCOLR, modeling and analyzing the citation network as a complex system, and designing an automated tool to classify the citation treatments.

In the first study, I have situated SCOLR in the broader context of complex adaptive systems. As the saliency of state court cases continues to increase, and the interactions of institutions become increasingly politicized, the importance of examining courts in a broader complex systems framework will allow for a more holistic analysis of federalism and institutions competing for power. To understand these interactions, I have constructed a novel database consisting of pertinent case facts, as well as a large amount legal text data in the forms of written majority, dissenting, and concurring opinions.

In the second study, I modeled the system and analyzed it as a complex system using the approach of Network Science. Selecting network science to model and analyze the SCOLR system was a successful decision. It helped to analyze the complexity of the system and understand the system behavior. Moreover, it produced interesting findings that are difficult to discover using other approaches. For example, the community-detection algorithms can be used to discover the issues of the cases. This process usually is an expensive process in terms of time and effort. Also, if we used community detection with the network-centrality measurements, we would be able to identify the most influential cases or judges in a specific legal area.

Creating the SCOLR system network based on the citation behavior (one of the most important behaviors in the system) is the first step to more analysis of the system as a complex system. However, there is a need for more research and analysis to better understand the system and predict its future.

Considering the citation treatments in network creation and analysis will assist us in discovering valuable, interesting information. To advance our understanding of judicial citation behavior, I proposed to add the citation treatment to the system network analysis.

Therefore, in the third study, I built a classification model to classify the citation treatments. Dealing with legal text was neither easy nor straightforward, especially when it involves court language and citations. Therefore, the accuracies of the sentiment analysis and ML classifications were not high enough to classify the citations treatments. (The highest I achieved was 50%.) However, I was able to build a dictionary of words/phrases that helped me to develop a new classifier model, the Treatment-Classifier, to identify the negative citation treatments. The model produced great results and better than what a human can produce. The recall for the negative class is 80% and the precision is 42%.

To improve the model accuracy, I need to collaborate with legal-domain experts to add more keywords to the dictionary and create a label dataset for the training and testing process of the classifier. Therefore, I decided to improve the accuracy of the model in future research before I add it to the network analysis.

This dissertation is a foundation for many future research directions. These directions include (but are not limited to) the following:

- Utilize my parsing tool to collect data on other judicial institutions. Federal court and state trial court data are essentially small sample of cases with little in terms of connections to other data. As I further develop my tool, I will be able to quickly and accurately construct large datasets with similar variables and case information to make comparisons across states and to the federal level possible. This effort is just scratching the surface on a plethora of other attributes I can construct from this data.
- Analyze the citation network of SCOLR as a dynamic network to explore the changes in the system behaviors over time.
- Use the approach of network science to build and analyze a network of collaboration between judges based on their votes on the cases.
- Improve the accuracy of the Treatment-Classifier model. This can be done by collaborating with legal domain experts to select the best citation-paragraph size, update the treatment dictionary, and increase the labeled-citations dataset to be used in the training and testing processes.
- Create a legal text lexicon to be used for legal text mining.

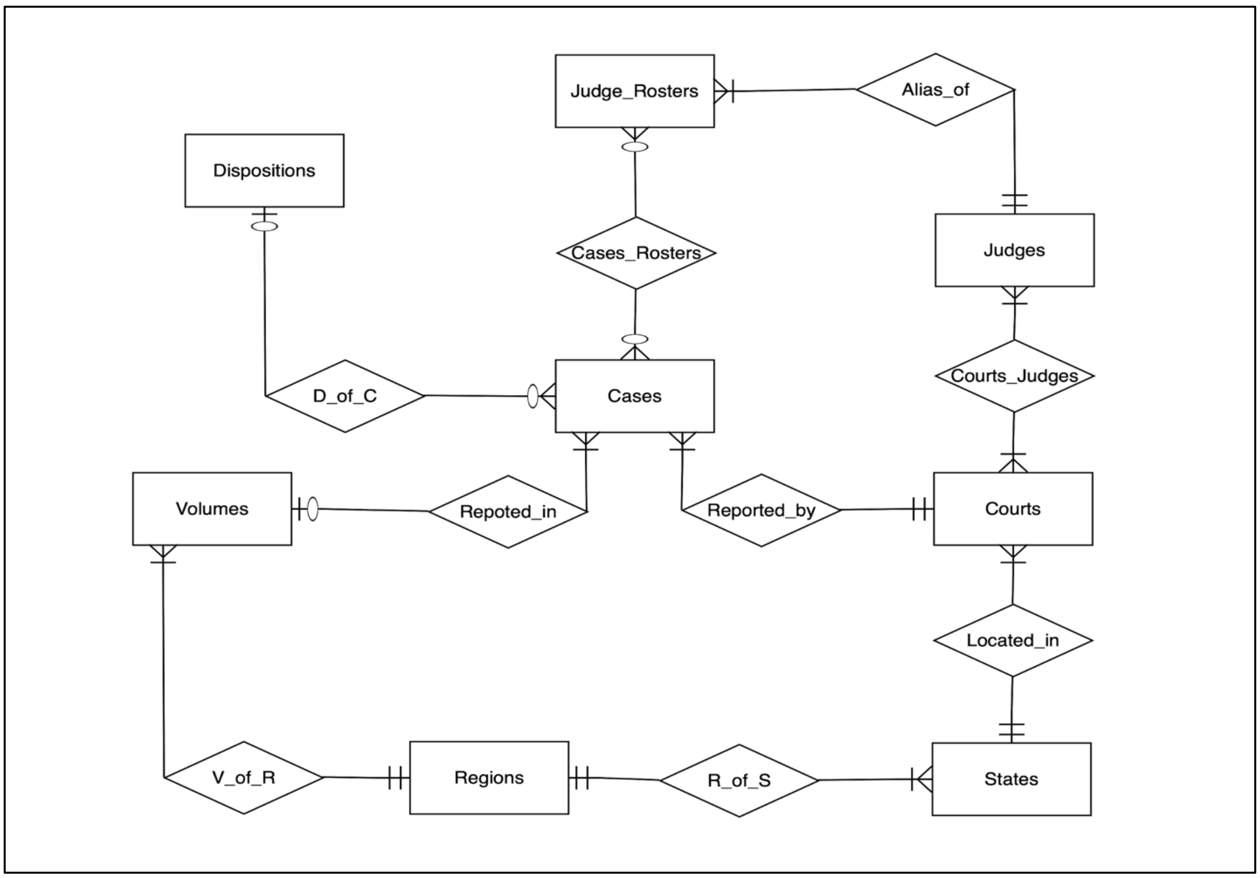
- Add the citation treatments to the network analysis.
- Create a citation prediction model based on the judges' citation behavior. It can be used to predict the cases that the judges will use and cite to make the decision on a new case.

Appendix

The Database Description

The dataset has been reorganized and structured as a relational database to make it a searchable dataset and more usable for the research community. Figure 31 shows the entity-relational diagram (ERD) of the database design. In this Appendix, we show the database design and the description of each table in the database.

Figure 31: ERD of the database Design



The Tables

1. **Cases table**. This table is the main table in the dataset. It has all the information about the cases of all the State Supreme Courts. This table has about 1,724,264 records and 14 columns. Each record presents the data of a case and each column is a variable to describe the cases.

Column name	Data type	Description
case_lexis_reporter	Text	Primary key of the table. The case id in the lexis reporter.
case_regional_reporter	Text	The case id in the regional reporter.
case_state_reporter	Text	The case id in the state reporter.
case_date	Date	The decision and publishing date of the case.
case_party_1	Text	Appellant or Appellee for each case.
case_party_2	Text	Appellant or Appellee for each case.
procedural_posture	Text	Ground of appeal to the court of last resort
overview	Text	A paragraph summary of the case history and facts
outcome	Text	The Court's ruling on the appeal
counsel_1	Text	Name and law firm/agency for attorney for case_party_1
counsel_2	Text	Name and law firm/agency for attorney for case_party_2
lexis_court_abbreviation	Text	Foreign key from the table "Courts" to connect the case to the court data of the court where the case published.
disposition_value	Number	Foreign key from the table "Dispositions" to connect the case with the disposition data of its disposition.

Column name	Data type	Description
rr_volume_abbreviation	Text	Foreign key from the table "Regional_Reporters_Volumes" to show the regional reporter volume number in which the case has been reported. The number and abbreviation of each volume.

2. **Regional_Reporters tables**. The courts of last resort have been divided into 7 groups based on their geographic region of the United States. Each region has its own reporter that includes all the published cases of the region's courts. This table has the data of the seven regions. As publishing decisions online became more readily available, cases in more recent decades may be published or unpublished decisions, with unpublished decisions holding no precedential authority in future cases.

Column name	Data type	Description
rr_abbreviation	Text	Primary key of the table. The abbreviation of the region name.
rr_name	Text	The name of the region.

3. **Regional_Reporters_Volumes tables.** This table has the data of each volume that has been published in each region.

Column name	Data type	Description
rr_volume_abbreviation	Text	Part of the table composite primary key. The number and abbreviation of each volume.

Column name	Data type	Description
rr_abbreviation	Text	Part of the table composite primary key. Foreign key from the table "Regional_Reporters" to connect the volume with the region. The abbreviation of the region.
rr_volume _start_date	Date	The start date of the volume.
rr_volume_end_date	Date	The end date of the volume.

4. **States tables.** This table has the states data. Each state has one court except the states of Colorado and Texas that each have two courts (one for criminal cases, and one for other types of cases). Each state is described in this table with three columns.

Column name	Data type	Description
postal_state_abbreviation	Text	Primary key of the table. The abbreviation of the State.
state_name	Text	The name of the state
fips_code	Text	The Federal Information Processing Standard state code
rr_abbreviation	Text	Foreign key from the table "Regional_Reporters" to connect the state with the region. The abbreviation of the region

5. **Courts tables.** This table has the courts data.

Column name	Data type	Description
lexis_court_abbreviation	Text	Primary key of the table. The abbreviation of the Courts based on Lexis reporter
court_name	Text	The court name

Column name	Data type	Description
court_type	Text	The court types. It will be helpful when different courts types and level add to the database. In this current dataset all the courts type is States Supreme Courts.
postal_state_abbreviation	Text	Foreign key from the table "States" to connect the courts to the states. The abbreviation of the State.

6. **Judges tables.** This table has the Judges names with their IDs.

Column name	Data type	Description
judge_code	Text	Primary key of the table. The judge id in the system.
judge_name	Text	The judge name

7. **Courts_Judges table**: since the relationship between the Judges and the Courts tables is many-to-many, this table serves as an associative table to connect those tables. This table shows the court that each judge worked in and also the judge's service start date and end date. Recall that some judges worked in more than one session at the same court. That is why the relationship is many-to-many.

Column name	Data type	Description
judge_code	Number	Primary key The generated judge code
lexis_court_abbreviation	Text	Part of the table composite primary key. The abbreviations of the courts. Foreign key from the table "Court" to connect the court with the judges who worked in that court.

Column name	Data type	Description
judge_code	Text	Part of the table composite primary key. The judge code. Foreign key from the table "Judge" to connect the judge with the court data and show his/her work sessions.
judge_start_date	Date	The date of the first day the judge served in the court.
judge_end_date	date	The date of the last day the judge served in the court.

8. **Judge_Rosters tables.** This table has the judge's ID and all of his/her aliases used in the dataset. Recall the fact that the judges have been reported by different aliases in the case reports.

Column name	Data type	Description
judge_alias	Text	Part of the table composite primary key. The alias used in the cases report for the judges.
judge_code	Text	Part of the table composite primary key. The code of the judge that the alias belongs to.

9. **Cases Rosters table**: This is an associative table for the many-to-many relationship between the cases and Judge_Rosters table. It shows the majority, dissent, and concur writers (judges) of the case. The relationship is many-to-many because the case may have one or more writers and the judge may vote and write one or more cases. The relationship is between the cases and the Judge_Rosters table, not to the Judge table, to connect the case with the judge's alias that was used in the case report.

Column name	Data type	Description
case_lexis_reporter	Text	Part of the table composite primary key. The *case_lexis_reporter* Foreign key from the table "cases" to connect the cases with the judges who voted on the case and their decisions on that case.
judge_alias judge_code	Text Text	Part of the table composite primary key. The judge alias and code. Foreign key from the table "Judge_rosters" to connect the cases with the judges who voted on the case and their decisions on that case.
decision_type	Text	The judge's decision on the case which could be opinion writer (O) or dissent writer (D) or concur writer (C).

10. **Dispositions tables.** This table has all the disposition values that have been created in the processing stage and their descriptions.

Column name	Data type	Description
disposition_value	Number	Primary key of the table. The code number of the disposition.
disposition_description	Text	The disposition description.

References

1984 | Timeline of Computer History | Computer History Museum. (n.d.). Retrieved October 26, 2021, from https://www.computerhistory.org/timeline/1984/

Barabási, A.-L. (2016). *Network Science*. Cambridge university press.

Bavelas, A. (1948). A MATHEMATICAL MODEL FOR GROUP STRUCTURES. *Human Organization*, *7*(3), 16–30.

Bickel, A. M. (1962). The Least Dangerous Branch: The Supreme Court at the Bar of Politics. *Bobbs-Merrill*. https://doi.org/10.2307/1952586

Blondel, V. D., Guillaume, J.-L., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, *2008*(10), P10008. https://doi.org/10.1088/1742-5468/2008/10/P10008

Brace, P., & Boyea, B. D. (2008). State public opinion, the death penalty, and the practice of electing judges. *American Journal of Political Science*, *52*(2), 360–372. https://doi.org/10.1111/j.1540-5907.2008.00317.x

Brace, P., & Hall, M. G. (1990). Neo-Institutionalism and Dissent in State Supreme Courts. *The Journal of Politics*, *52*(1), 54–70. https://doi.org/10.2307/2131419

Brace, P., & Hall, M. G. (1993). Integrated Models of Judicial Dissent. *The Journal of Politics*, *55*(4), 914–935. https://doi.org/10.2307/2131942

Brace, P., & Hall, M. G. (1995). Studying Courts Comparatively: The View from the American States. *Political Research Quarterly*, 5–29.

Brace, P., Hall, M. gann, & Langer, L. (2001). Placing State Supreme Courts in State Politics. *State Politics. State Politics & Policy Quarterly*, *1*(1), 81–108.

Brace, P. R., & Hall, M. G. (1997). The Interplay of Preferences, Case Facts, Context, and

Structure in the Politics of Judicial Choice. *The Journal of Politics*, *59*(4), 1206–1232.

Brandes, U. (2001). A Faster Algorithm for Betweenness Centrality. *Journal of Mathematical Sociology*, *25*(2), 163–177.

Caldarone, R. P., Canes-Wrone, B., & Clark, T. S. (2009). Partisan Labels and Democratic Accountability: An Analysis of State Supreme Court Abortion Decisions. *The Journal of Politics*, *71*(2), 560–573. https://doi.org/10.1017/S002238160909046X

Canes-Wrone, B., Clark, T. S., & Kelly, J. P. (2014). Judicial selection and death penalty Decisions. *American Political Science Review*, *108*(1), 23–39. https://doi.org/10.1017/S0003055413000622

Cann, D. M., & Wilhelm, T. (2011). Case visibility and the electoral connection in state supreme courts. *American Politics Research*, *39*(3), 557–581. https://doi.org/10.1177/1532673X10385778

Conrad, J. G., & Dabney, D. P. (2001). Automatic recognition of distinguishing negative indirect history language in judicial opinions. *Proceedings of the Tenth International Conference on Information and Knowledge Management*, 287–294. https://doi.org/10.1145/502585.502634

Cross, F., Spriggs, J., Johnson, T., & Wahlbeck, P. (2010). Citations in the US Supreme Court: An Empirical Study of Their Use and Significance. *U. Illinois L. R.*, *2010*(2), 489–576.

Dahl, R. A. (1956). *A Preface to Democratic Theory*. University of Chicago Press.

Das, K., Samanta, S., & Pal, M. (2018). Study on centrality measures in social networks: a survey. *Social Network Analysis and Mining*, *8*(1), 1–11. https://doi.org/10.1007/s13278-018-0493-2

Erdős, P., & Rényi, A. (1959). On random graphs I. *Publications Mathematicae (Debrecen)*, *6*, 290–297.

Erdős, P., & Rényi, A. (1960). ON THE EVOLUTION OF RANDOM GRAPHS. *Institute of*

Mathematics Hungarian Academy of Sciences, *5*, 17–61.

Erdős, P., & Rényi, A. (1961a). On the evolution of random graphs ii. *Bull. Inst. Internat. Statist.*, *38*, 343–347.

Erdős, P., & Rényi, A. (1961b). ON THE STRENGTH OF CONNECTEDNESS OF A RANDOM GRAPH. *Acta Math. Acad. Sci. Hungary*, *12*, 261–267.

Erdős, P., & Rényi, A. (1963). ASYMMETRIC GRAPHS. *Acta Mathematica Acad. Sci. Hungarica*, *14*, 195–315.

Erdős, P., & Rényi, A. (1966a). On random matrices. *Publ. Math. Inst. Hung. Acad. Sci.*, *8*, 455–461.

Erdős, P., & Rényi, A. (1966b). ON THE EXISTENCE OF A FACTOR OF DEGREE ONE OF A CONNECTED RANDOM GRAPH. *Acta Math. Acad. Sci. Hungary*, *17*, 359–368.

Erdős, P., & Rényi, A. (1968). On random matrices II. *Studia Scientiarum Mathematicarum Hungarica*, *13*, 459–464.

Freeman, L. C. (1977). A Set of Measures of Centrality Based on Betweenness. *Sociometry*, *40*(1), 35. https://doi.org/10.2307/3033543

Freeman, L. C. (1978). Centrality in Social Networks Conceptual Clarification. *Social Networks*, *1*(3), 215–239. https://doi.org/10.1016/0378-8733(78)90021-7

Friedman, B. (2009). *The Will of the People: How Public Opinion has Influenced the Supreme Court and Shaped the Meaning of the Constitution*. Farrar, Straus and Giroux.

Galgani, F., Compton, P., & Hoffmann, A. (2015). LEXA: Building knowledge bases for automatic legal citation classification. *Expert Systems with Applications*, *42*(17–18), 6391–6407. https://doi.org/10.1016/j.eswa.2015.04.022

Géron, A. (2017). *Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow:*

Concepts, tools, and techniques to build intelligent systems. O'Reilly Media.

Gilbert., E. N. (1959). Random graphs. *The Annals of Mathematical Statistics*, *30*, 1141–1144.

Gordon, S. C. (2007). The Effect of Electoral Competitiveness on Incumbent Behavior. *Quarterly Journal of Political Science*, *2*(2), 107–138. https://doi.org/10.1561/100.00006035

Hall, M. E. K. (2014). The Semiconstrained Court: Public Opinion, the Separation of Powers, and the U.S. Supreme Court's Fear of Nonimplementation. *American Journal of Political Science*, *58*(2), 352–366. https://doi.org/10.1111/ajps.12069

Hall, M. G. (1987). An Examination of Voting Behavior in the Louisiana Supreme Court. *Judicature*, *71*, 40–47.

Hall, M. G. (1992). Electoral Politics and Strategic Voting in State Supreme Courts. *The Journal of Politics*, *54*(2), 427–446. https://doi.org/10.2307/2132033

Hall, M. G. (1995). Justices as representatives: Elections and Judicial Politics in the American States. *American Politics Research*, *23*(4), 485–503. https://doi.org/10.1177/1532673X9502300407

Hamilton, A. (1788). *Federalist 78*.

Huber, G. A., & Gordon, S. C. (2004). Accountability and Coercion: Is Justice Blind when It Runs for Office? *American Journal of Political Science*. https://doi.org/10.1111/j.0092-5853.2004.00068.x

Jackson, P., Al-Kofahi, K., Tyrrell, A., & Vachher, A. (2003). Information extraction from case law and retrieval of prior cases. *Artificial Intelligence*, *150*(1–2), 239–290. https://doi.org/10.1016/S0004-3702(03)00106-1

Johnson, G. (2017). Legislative "Allies" and Judicial Oversight of Executive Power. *Justice System Journal*, *38*(2), 116–134. https://doi.org/10.1080/0098261X.2016.1243022

Legal Dictionary. (n.d.). *legal definition of Court of last resort*. Retrieved September 21, 2020, from https://legal-dictionary.thefreedictionary.com/Court+of+last+resort

Macey, J. (1989). Internal and External Costs and Benefits of Stare Decisis, The. *Chi.-Kent L. Rev.*, *65*(c), 1–15. http://heinonlinebackup.com/hol-cgi-bin/get_pdf.cgi?handle=hein.journals/chknt65&section=11

McCloskey, R. G. (2010). *The American Supreme Court* (5th ed.). Duke University Press.

McGuire, K. T., & Stimson, J. A. (2004). The least dangerous branch revisited: New evidence on Supreme Court responsiveness to public preferences. In *Journal of Politics* (Vol. 66, Issue 4, pp. 1018–1035). https://doi.org/10.1111/j.1468-2508.2004.00288.x

Molloy, M., & Reed, B. (1995). A Critical Point for Random Graphs-with a Given Degree Sequence. *Random Structures and Algorithms*, *6*.

Newman, M. E. J. (2001). Scientific collaboration networks. I. Network construction and fundamental results. *Physical Review E*, *64*(1). https://doi.org/10.1103/PhysRevE.64.016131

Newman, M. E. J. (2004). Fast algorithm for detecting community structure in networks. *Physical Review E*, *69*(6). https://doi.org/10.1103/PhysRevE.69.066133

NIEMINEN, J. (1974). On the centrality in a graph. *Scandinavian Journal of Psychology*, *15*(1), 332–336. https://doi.org/10.1111/J.1467-9450.1974.TB00598.X

Ravi, K. (2019). *Preparation of United States' State Court Cases Dataset and Deriving an Efficient Embedding for Cases*. The University of North Carolina at Charlotte.

Ruhl, J. B., Katz, D. M., & Bommarito, M. J. (2017). Harnessing legal complexity. In *Science* (Vol. 355, Issue 6332, pp. 1377–1378). https://doi.org/10.1126/science.aag3013

Sabidussi, G. (1966). THE CENTRALITY INDEX OF A GRAPH. *Psychometrika*, *31*(4), 581–603.

Segal, J. A. ., & Spaeth, H. J. . (1996). The Influence of Stare Decisis on the Votes of United States Supreme Court Justices Source : American Journal of Political Science , Vol . 40 , No . 4 , (Nov ., 1996), pp . 971-1003 Published by : Mi. *American Journal of Political Science*, *40*(4), 971–1003.

Shaw, M. E. (1954). Group Structure and the Behavior of Individuals in Small Groups. *The Journal of Psychology*, *38*(1), 139–149. https://doi.org/10.1080/00223980.1954.9712925

Wechsler, H. (1959). *Toward Neutral Principles of Constitutional Law Author (s): Herbert Wechsler Published by : The Harvard Law Review Association Stable URL : http://www.jstor.org/stable/1337945 HARVARD LAW REVIEW* |. *73*(1), 1–35.

Young, E. A. (2002). JUDICIAL ACTIVISM AND CONSERVATIVE POLITICS. *UNIVERSITY OF COLORADO LAW REVIEW*, *73*(4).