



2019-07-01

# Measuring Influence on Linear Dynamical Networks

Jaekob Chenina  
*Brigham Young University*

Follow this and additional works at: <https://scholarsarchive.byu.edu/etd>

---

## BYU ScholarsArchive Citation

Chenina, Jaekob, "Measuring Influence on Linear Dynamical Networks" (2019). *Theses and Dissertations*. 7523.  
<https://scholarsarchive.byu.edu/etd/7523>

This Thesis is brought to you for free and open access by BYU ScholarsArchive. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of BYU ScholarsArchive. For more information, please contact [scholarsarchive@byu.edu](mailto:scholarsarchive@byu.edu), [ellen\\_amatangelo@byu.edu](mailto:ellen_amatangelo@byu.edu).

Measuring Influence on Linear Dynamical Networks

Jaekob Chenina

A thesis submitted to the faculty of  
Brigham Young University  
in partial fulfillment of the requirements for the degree of  
Master of Science

Sean Warnick, Chair  
Jacob Crandall  
Casey Deccio

Department of Computer Science  
Brigham Young University

Copyright © 2019 Jaekob Chenina  
All Rights Reserved

## ABSTRACT

### Measuring Influence on Linear Dynamical Networks

Jaekob Chenina

Department of Computer Science, BYU

Master of Science

Influence has been studied across many different domains including sociology, statistics, marketing, network theory, psychology, social media, politics, and web search. In each of these domains, being able to measure and rank various degrees of influence has useful applications. For example, measuring influence in web search allows internet users to discover useful content more quickly. However, many of these algorithms measure influence across networks and graphs that are mathematically static. This project explores influence measurement within the context of linear time invariant (LTI) systems. While dynamical networks do have mathematical models for quantifying influence on a node-to-node basis, to the best of our knowledge, there are no proposed mathematical formulations that measure aggregate level influence across an entire dynamical network. The dynamics associated with each link, which can differ from one link to another, add additional complexity to the problem. Because of this complexity, many of the static-graph approaches used in web search do not achieve the desired outcome for dynamical networks. In this work we build upon concepts from PageRank and systems theory introduce two new methods for measuring influence within dynamical networks: 1) Dynamical Responsive Page Rank (DRPR) and 2) Aggregated Targeted Reachability (ATR). We then compare and analyze and compare results with these new methods.

Keywords: network reconstruction, targeted-reachability, page-rank, influence, authority, dynamical networks

## ACKNOWLEDGEMENTS

Thank you to all of those that believed in me. Thank you to all of those that coached, advised, taught and empowered me. Thank you to Sean for all your support. Thank you to my advisors for your involvement and advice. Thank you to all the BYU professors that educated me and shared their knowledge with me. Thank you to Galena and my kids, Alex, Leo and Ksenja for supporting me and inspiring me all along the way. Thank you.

## Table of Contents

<b>List of Figures</b>	<b>vi</b>
<b>List of Tables</b>	<b>viii</b>
<b>1 Influence</b>	<b>1</b>
1.1 The Beginning of Influence . . . . .	1
1.2 Influence in the Mathematical Sciences . . . . .	4
1.3 Influence in the Social Sciences . . . . .	5
1.4 Influence in Computer Science . . . . .	6
1.5 Other Areas of Influence . . . . .	7
1.6 Influence in Dynamical Networks . . . . .	8
1.7 Thesis Outline . . . . .	11
<b>2 Problem Statement</b>	<b>12</b>
2.1 The Complexity of Dynamical Networks . . . . .	12
2.2 Problem Illustration . . . . .	13
<b>3 New Approaches to Measuring Influence in Linear Dynamical Networks</b>	<b>17</b>
3.1 Page Rank . . . . .	17
3.2 Rank Sink Problem . . . . .	18
3.3 Responsiveness to Model Various Degrees of Relationships . . . . .	18
3.4 Dynamical Responsive PageRank (DRPR) . . . . .	19
3.5 Aggregated Targeted Reachability (ATR) . . . . .	22

<b>4</b>	<b>Technical Results</b>	<b>24</b>
4.1	Simple Canonical Networks . . . . .	24
4.2	Medium Sized Networks . . . . .	27
4.3	Larger Network . . . . .	30
4.4	DRPR and ATR Results for Problem Illustration Example Network . . . . .	30
<b>5</b>	<b>Conclusions and Future Work</b>	<b>36</b>
5.1	Conclusion . . . . .	36
5.2	Future Work . . . . .	37
	<b>References</b>	<b>39</b>

## List of Figures

1.1	The serpent influences Eve and then Eve influences Adam to eat the forbidden fruit . . . . .	1
1.2	Drug X and Drug Y do have an influence on health condition A . . . . .	3
1.3	Neither Drug X nor Drug Y have an influence on health condition A . . . . .	3
1.4	The relative magnitude of edge-to-edge influence, $\ Q_{ij}(s)\ _\infty$ , reconstructed by the B-VNR Algorithm on the Twitterdata [15]. . . . .	10
2.1	Problem Illustration . . . . .	14
2.2	Problem Illustration Results Using Standard PageRank . . . . .	14
2.3	$tf_1$ and $tf_2$ step responses compared (orange is $tf_1$ and blue is $tf_2$ ) . . . . .	15
4.1	Simple Canonical Example 1 . . . . .	24
4.2	Simple Canonical Example 2 . . . . .	25
4.3	Simple Canonical Example 3 . . . . .	25
4.4	Simple Canonical Example 4 . . . . .	25
4.5	Simple Canonical Example 5 . . . . .	25
4.6	Medium Example 6 . . . . .	27
4.7	Medium Example 7 . . . . .	28
4.8	Medium Example 8 . . . . .	28
4.9	ATR results for large canonical examples . . . . .	29
4.10	DRPR results for large canonical examples . . . . .	30
4.11	Normalized ATR results on large canonical Example 8 . . . . .	31
4.12	An influence diagram for a larger network . . . . .	32

4.13 Normalized ATR results on larger network. . . . .	33
4.14 DRPR on a larger network . . . . .	33
4.15 DRPR Results Computed on Problem Illustration Example from Section 2 . . . . .	34
4.16 Normalized ATR Results Computed on Problem Illustration Example from Section 2 . . . . .	35



## List of Tables

4.1	DRPR algorithm results with canonical examples . . . . .	26
4.2	Results for Targeted Reachability . . . . .	26
4.3	Normalized ATR Results . . . . .	26

## Chapter 1

### Influence

#### 1.1 The Beginning of Influence

Humans have an innate need to understand their environment and identify the factors that influence their ever changing circumstances. Judea Pearl, a renown Israeli-American computer scientist and philosopher, famously illustrates how this need to understand causal effects was documented at the beginning of one of the world’s oldest books - the Bible. After Adam ate the forbidden fruit in the garden of Eden and was awakened to his own nakedness, God asks him “Who told thee that thou wast naked? Hast thou eaten of the tree, whereof I commanded thee that thou shouldest not eat?” Adam responds “The woman whom thou gavest to be with me, she gave me of the tree, and I did eat.”



Figure 1.1: The serpent influences Eve and then Eve influences Adam to eat the forbidden fruit

Adam is essentially implying that Eve had an effect on Adam that causally influenced him to eat the fruit. God then follows up His inquiry with Eve “And the Lord God said unto the woman, What is this that thou hast done? And the woman said, The serpent beguiled me, and I did eat.” Yet again mankind is explaining the factors that causally influenced their actions and current circumstances.

If we take this Bible passage and abstract the dialogue into an influence diagram as described by [26], then we will get the diagram shown in Figure 1.1. Visual models are one way of understanding how distinct entities influence each other. The first step in understanding a causal relationship between two distinct entities is confirming whether there is any sort of influence taking place in the first place. Here is a list of questions that aim to answer such “is there a relationship” type of questions:

- Does smoking contribute to lung cancer?
- Does posting speed limit signs cause drivers to drive slower?
- Does teaching method X make students learn more?
- Is there a relationship between drug X and symptom Y?

As we apply the various information gathering methodologies regarding the question at hand (scientific method or day-to-day observation, etc.), we form an opinion about whether or not there is a relationship between the distinct entities. If we establish through controlled experimentation that drugs X and Y do have an influence on health condition A, then the influence diagram could look the diagram in Figure 1.2.

If, however, we discover that Drug X and Drug Y have no influence on health condition A, then our influence diagram will not have edges as in Figure 1.3. These simple influence diagrams are one way that mankind strives to understand the causal influences in their environment and in turn influence them directly. Throughout this thesis we will continue to use influence diagrams in combination with other tools like transfer functions, linear time invariant systems (LTI systems), PageRank computing, and other mathematical functions to enrich

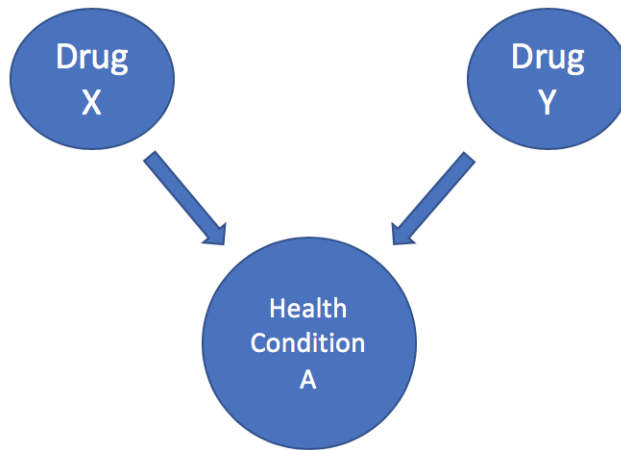


Figure 1.2: Drug X and Drug Y do have an influence on health condition A

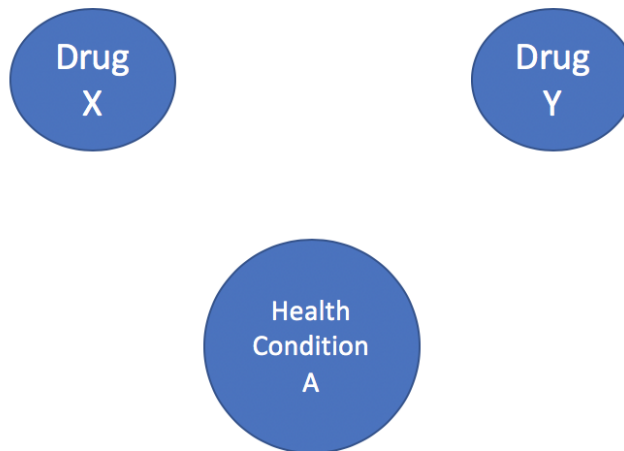


Figure 1.3: Neither Drug X nor Drug Y have an influence on health condition A

our understanding of these fundamentally important influence relationships. The influence diagrams that we have shown thus far are binary influence diagrams, meaning that either there is an edge or there is not an edge. Similarly, the world famous PageRank computing algorithms function with binary graphs (either there is a link or there is not a link). Throughout this

thesis we will introduce scenarios where the existing tools (including PageRank and binary influence networks) are insufficient to describe the rich nuance relationships that might exist within an LTI network (particularly a dynamical network). The goal of this research is to empower mankind with a more advanced tools for measuring influence and in turn deepen our understanding of the influencing factors that shape our world. Although there are many conjectures and theories regarding influence, it is a slippery concept that could be defined and measured in many different ways. However, among the various definitions across sociology, communications, marketing, and political science a common theme indicates that

- 1) There are distinct entities (example :  $A$  and  $B$ )
- 2) There is a change that is happening in at least one of the entities
- 3) The change in one entity is caused by the other entity.

In this research, we will focus on measuring aggregate level influence within LTI systems. Before we dive into this level of detail, we will provide an overview of the theory of influence across research disciplines.

## 1.2 Influence in the Mathematical Sciences

Visual diagrams are not the only way that people have tried to reason about the influences of the world. Much of the scientific method itself is designed to validate or invalidate whether or not certain influence relationships exist. This introduces the long-lasting dilemma of correlation versus causation. Without dwelling much on this fundamental challenge of parsing the two apart, it is worth discussing that there is a variety of tools for measuring patterns, which may or may not indicate causal influence.

As early as the 1800s, Legendre and Gauss used linear regressions to measure the influence of correlated variables and to estimate the planetary movement [43]. The chi-squared ( $x^2$ ) test was also introduced around the same time to test the “goodness of fit” in a hypotheses test. However, one weakness of the chi-squared test is that it is difficult to measure to what degree one variable changes as values of another variable change [24]. A

variation of the chi-squared test is the contingency coefficient, which automatically adjusts for different sample sizes; however, it is also limited in the possible ranges of values to show various strengths of correlation [48]. Then in the 1940's, Cramer introduced Cramer's V as another chi-based method of measuring causality across nominal variables, which made it easier to detect the strength of the correlation [18].

Later the concept of Granger causality was introduced. Granger causality is a statistical measure for estimating influence between two variables. It is based on linear regressions and stochastic processes. In practice, if including an  $X_1$  term reduces the variance  $E_1$  when forecasting  $X_2$ , then it is said that  $X_1$  granger-causes  $X_2$  [23]. A similar concept from information theory is mutual information, which measures how much more you know about a random variable given another random variable [30].

### 1.3 Influence in the Social Sciences

Another area of influence research focuses on practical aspects of how to influence the people and the world around us. Classic books such as Dale Carnegie's "How to Win Friends and Influence People" describe fundamental qualities and attributes that make a person more influential and likeable. This includes humility, accepting fault quickly, sincere appreciation of the efforts of others, genuine honesty, remembering other peoples' names, making others feel important, and smiling. Robert Cialdini's book "Influence, the Psychology of Persuasion" describes specific interpersonal and psychological strategies of influence such as social obligation through reciprocity, providing reasoning and explanations, social proof ("everyone else is doing this"), perceived authority, and scarcity ("your offer expires today!").

Moving beyond influence at the individual level, marketing research studies the various factors that influence consumers at scale [38]. This includes customer value [19], a desire to be recognized, company perception [20], product positioning [12], cultural factors [34], and the physical environment [41]. Business researchers have studied how to measure influence in organizational decision making [42] and how to identify and measure influential customers

[29]. In marketing research the notion of influence is especially important because if one can identify customers or potential brand-advocates that will have a network effect, then this information can be used to sell and market much more effectively [32, 33].

In politics there is also an active community of research seeking to understand the factors that influence public opinion. This field of research studies how individual politicians can advance their careers [28], influence through lobbying [13], economic power and influence [40], religious influence on the political system [21], influence through social media [8], the influence of education and universities [27], and the political influence of interpersonal relationships [36].

These areas of influence research deepened our understanding of the effects and influences that Adam and Eve described in the Bible. However, to what degree precisely did the Serpent really influence Eve to partake of the fruit? Fifty Percent? 10 percent? How do we quantify that influence and how do we quantify the degree to which smiling makes us more likeable and influential as Dale Carnegie observes? Experiments have been conducted in an attempt to answer some of these questions under arguably artificial controlled circumstances. However, with the birth of the internet a new unbiased data set became available. This data set has and is propelling quantifiable influence research to the next level.

#### **1.4 Influence in Computer Science**

With the adoption of the internet, a vast data set became available of websites, social media, blogs, discussion forums, and other publicly shared data sets where computer scientists have created a variety of use-case specific methods for computing influence. By far one of the most successful algorithms ever deployed for measuring influence in the world of computer science is indisputably the PageRank algorithm. Later we will have a dedicated section for the PageRank algorithm and build upon it to create a new and improved responsiveness-based method for computing influence and page rank. Before we do this, we will survey other non-PageRank contributions relevant to the topic of measuring influence.

With the rise of social media, many more quantitative models were created for measuring influence on social media [44, 47]. In the early days of social media, influence was often measured by evaluating graph structure [25]. However, it was noted that a node may have many social media connections, but those connections may not actually respond to a source node [5, 11, 14]. Hence others defined social media influence as the magnitude of a information propagation across a network originating from a source node [6, 39]. With Twitter the retweet metric can be helpful for measuring such propagation. However, people may be influenced by a tweet without explicitly re-tweeting or commenting on a post. Such influence is much more difficult to measure and the possible solutions that we outline in this thesis could potentially be used to measure such implicit influence.

## 1.5 Other Areas of Influence

Statistics, psychology, business, politics and computer science are a few areas with active communities that study influence-related topics that we have mentioned, but if we think of influence broadly, it truly touches nearly all disciplines. Even in academics we also use a variety of techniques for evaluating the influence of research itself including the h-index (used to measure the influence and importance of academic publications) [9, 10], and Lotka's Law and Bradford's Law ("bibliometrics" that help measure the impact and influence of researcher's and academic journals) [3]. There are even influential books that advise poets how to gain influence within the world of poetry [7]. Here we have just provided a sample of some of the influential highlights, however it is arguable that influence influences everything.

Now we will shift gears from discussing influence broadly and we will dive deeper into a more focused area of influence with gaps and opportunities for improvement that we will illustrate. Next we will discuss influence in the context of dynamical networks and transfer functions.



## 1.6 Influence in Dynamical Networks

Transfer functions are functions of complex variables that model a system's output for each possible input. Transfer functions can be determined from 1) experiments on a system, 2) inspection, or by 3) algebraic manipulations of differential equations that describe the system. The Transfer function of a system can be represented by the rational function

$$G(s) = \frac{b(s)}{a(s)}$$

where  $a(s)$  and  $b(s)$  are polynomials of a system of arbitrary order,  $n$  and  $m$ , although typically we consider systems that are proper, i.e. where  $n > m$ , given by

$$a(s) = s^n + a_1s^{n-1} + a_2s^{n-2} + \dots + a_{n-1}s + a_n$$

,

$$b(s) = b_0s^m + b_1s^{m-1} + \dots + b_{m-1}s + b_m$$

and where  $s$  is the Laplace variable. These system equations are useful because they can be computed from data and can model the overall system and subsystem behavior of interconnected networks. In previous work, we use a method called network reconstruction to obtain system models from raw data [16, 50, 50]. In this process, we use the dynamical structure function to relate inputs and outputs through the following equation:

$$Y(s) = Q(s)Y(s) + P(s)U(s), \tag{1.1}$$

where entries (often called *links* or *modules*) in  $P(s)$  define the direct causal mapping from individual inputs to individual outputs and entries in  $Q(s)$  define the direct causal mapping from individual outputs to other outputs (the diagonal entries of  $Q(s)$  are zero). As the state of one of the network nodes changes, it can have ripple effects on the other nodes in

the network. This downstream effects are essentially how one node influences another node. For certain types of reconstructed network models (such as the ones we evaluate in [15]) the direct inputs into the system are not observable and hence the  $U(s)$  component, which represents the inputs of the system goes away and we are left with

$$Y(s) = Q(s)Y(s). \tag{1.2}$$

The entries inside of  $Q$  are transfer functions and  $Y(s)$  are the measured output variables on each link. Once these values have been solved for, the magnitudes of the transfer functions can be visualized as in Figure 1.1. From this point however, it is difficult to identify which node in the network is overall the most influential given that this is a dynamical network. To simplify the calculations we can apply a norm to the  $Q$  matrix to simplify the calculations. For some of the examples going forward we will continue to use the  $Q$  matrix (the matrix of transfer functions that model a dynamic network) as the starting point and we will continue to apply a norm in order to simplify comparisons.

The dynamical structure function was introduced in 2008 [22], and further developed [1, 31, 45, 46, 49–51] to create input-output transfer functions from data that model how a network node influences another node as a function of input variables and time. In previous research of ours, we demonstrated how this framework can be adapted to model the influence between nodes in a social media network [15]. Figure 1.4 represents a network of social media nodes where the edges in between nodes represents the relative level of influence from node to node. After completing this work, we were left wondering which of the nodes in the network was the most influential node overall in the network. A naive solution to this problem might be to simply sum up the norm of the transfer function of each edge, however, this would not account for the impact of ripple effects across the network. Seeking to solve this problem was the genesis of the research in this thesis. As we thought deeply about this specific use case in social media, it became apparent that this is a generalizable problem that applies to a wide range of transfer-function based time-invariant systems networks. Much of the

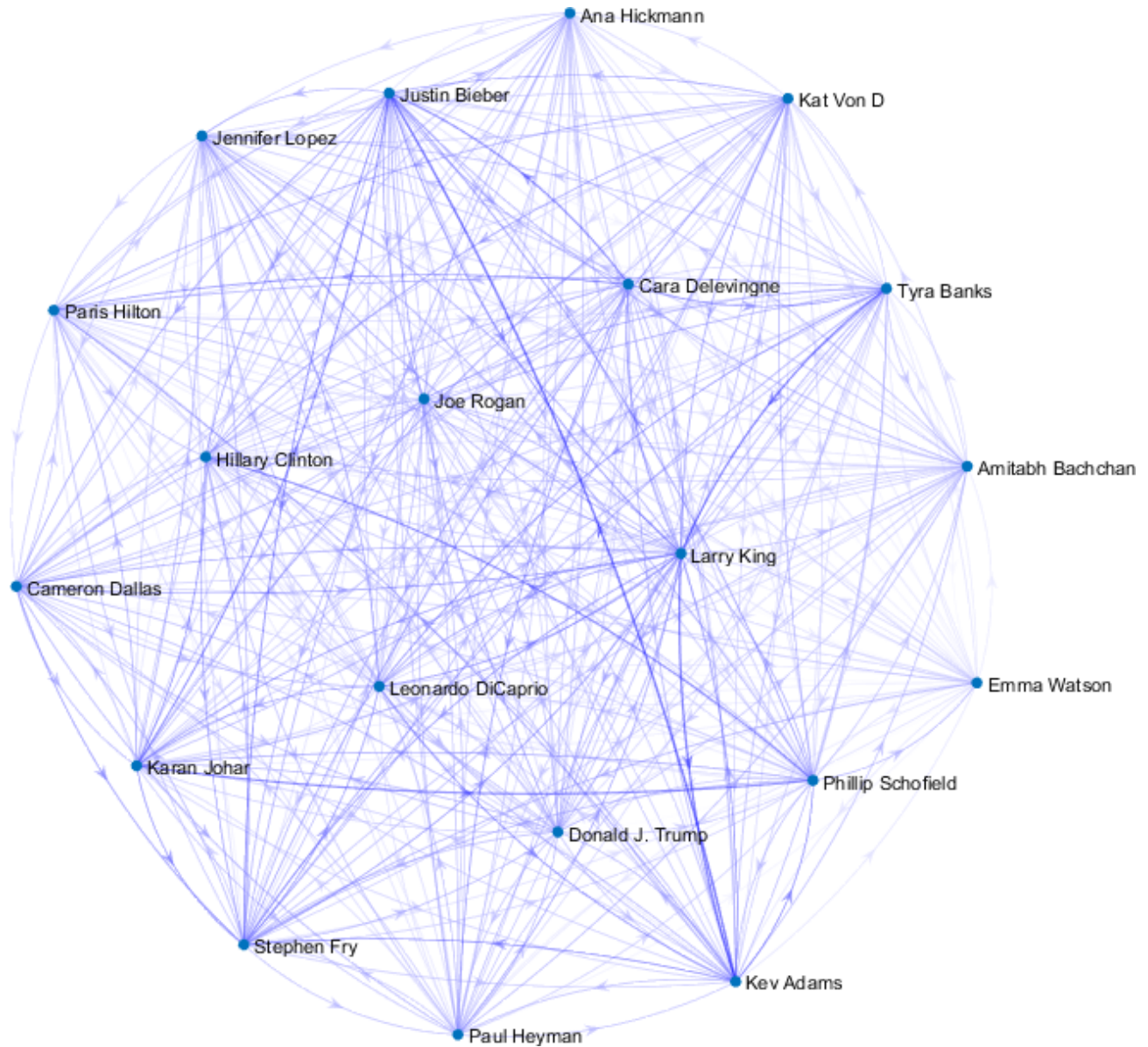


Figure 1.4: The relative magnitude of edge-to-edge influence,  $\|Q_{ij}(s)\|_{\infty}$ , reconstructed by the B-VNR Algorithm on the Twitterdata [15].

research related to the dynamical structure function was applied in areas of biology, finance, and operational research and solving this problem could help identify key influencer nodes across any modeled network in all of the respective domains mentioned.

## 1.7 Thesis Outline

When it comes to identifying “influencers” in a connected network, PageRank immediately comes to mind, but in the upcoming section on PageRank we will demonstrate some of the limitations with PageRank in solving the problem of identifying an influencer in a time-invariant system. In Chapter 3, we will then introduce two novel methods for computing influence in a dynamic network. In chapter 4, we will demonstrate experimental results and the final chapter will draw conclusions and discuss future work.

## Chapter 2

### Problem Statement

In this chapter, we discuss in greater detail the problem we are seeking to solve in this thesis. The focus of this thesis is to formulate and evaluate new methodologies for quantifying influence in dynamical networks.

#### 2.1 The Complexity of Dynamical Networks

Since Sergey Brin and Larry Page's first introduction of the random surfer model, PageRank computing has been primarily focused on static networks with binary links (i.e., there is or there isn't a link between any two nodes). Within dynamical networks, much like static networks, there is a need to measure influence within a complex body of information. In the case of a web search, the complexity is largely due to the sheer volume of individual graph nodes. Within dynamical network applications, the number of nodes may be smaller by comparison. However, the number of possible network states is infinite and instead of a binary value on an edge, it could be zero but it could also be a rational function such as

$$\frac{(s^2 + 5)}{(s + 2)}$$

or something even more complex such as

$$\frac{(s^2 + 6s + 9)}{(s^3 + 9s^2 + 25s + 21)}.$$

and where  $s \in \mathbb{C}$ . These rational functions can be summarized and simplified to a scalar value by applying a norm to the rational function (we will demonstrate using this method starting in Section 3). However, treating the existence of a link between two nodes as a binary value in order to apply the well-established PageRank algorithm has limitations that we will illustrate here.

## 2.2 Problem Illustration

This example illustrates that nodes may technically be influencing each other, but the rate, frequency and intensity of that influence may be very low. Hence PageRank with binary edges will not accurately represent the full dynamic level of influence that is taking place. In the image below (Figure 2.1), the dark edges are represented by transfer functions  $tf_1$  with the value

$$tf_1 = \frac{10}{s + 3}$$

which has a 2-norm of 4.0825. The lighter colored edges are represented by

$$tf_2 = \frac{1}{5s + 5}$$

which has a norm of 0.1414.

In the world of PageRank, a link implies that the receiving node has some sort of influence or authority that is represented by the incoming link. However, the actual intensity or severity of the link is not accounted for. As such, we will treat these edges as binary values and apply the PageRank algorithm as outlined in [37]. This yields the results displayed in Figure 2.2. However, comparing the step response of  $tf_1$  and  $tf_2$  we see that the other nodes in the network are much more “disturbed” by the effects of the  $tf_1$  edges coming out of  $Y_2$ . According to the Merriam-Webster dictionary, influence is defined as the power or capacity of causing an effect in indirect or intangible ways. Although there are other ways that influence could be defined, intuitively we can sense from the example that the PageRank algorithm

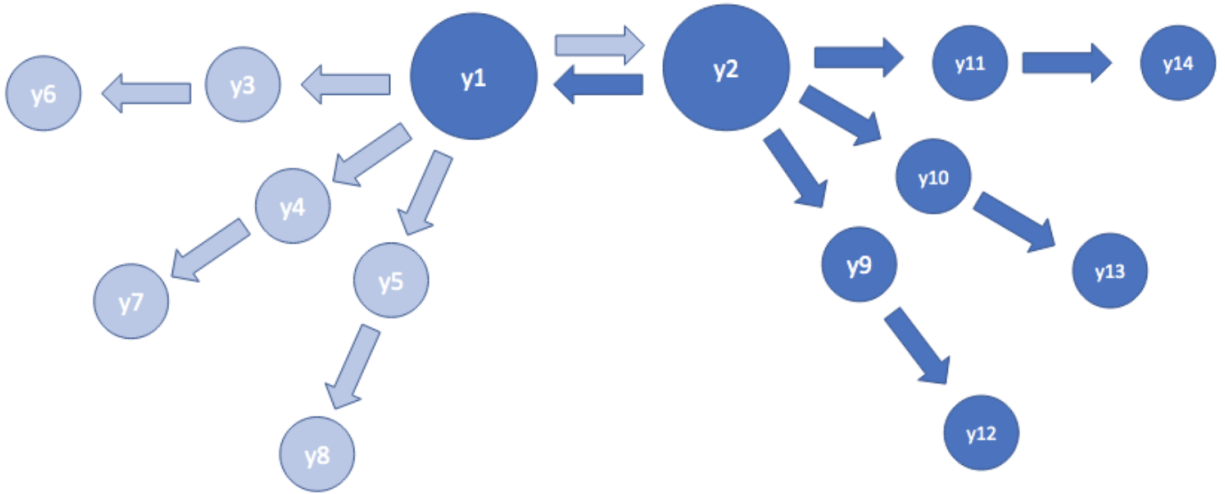


Figure 2.1: Problem Illustration

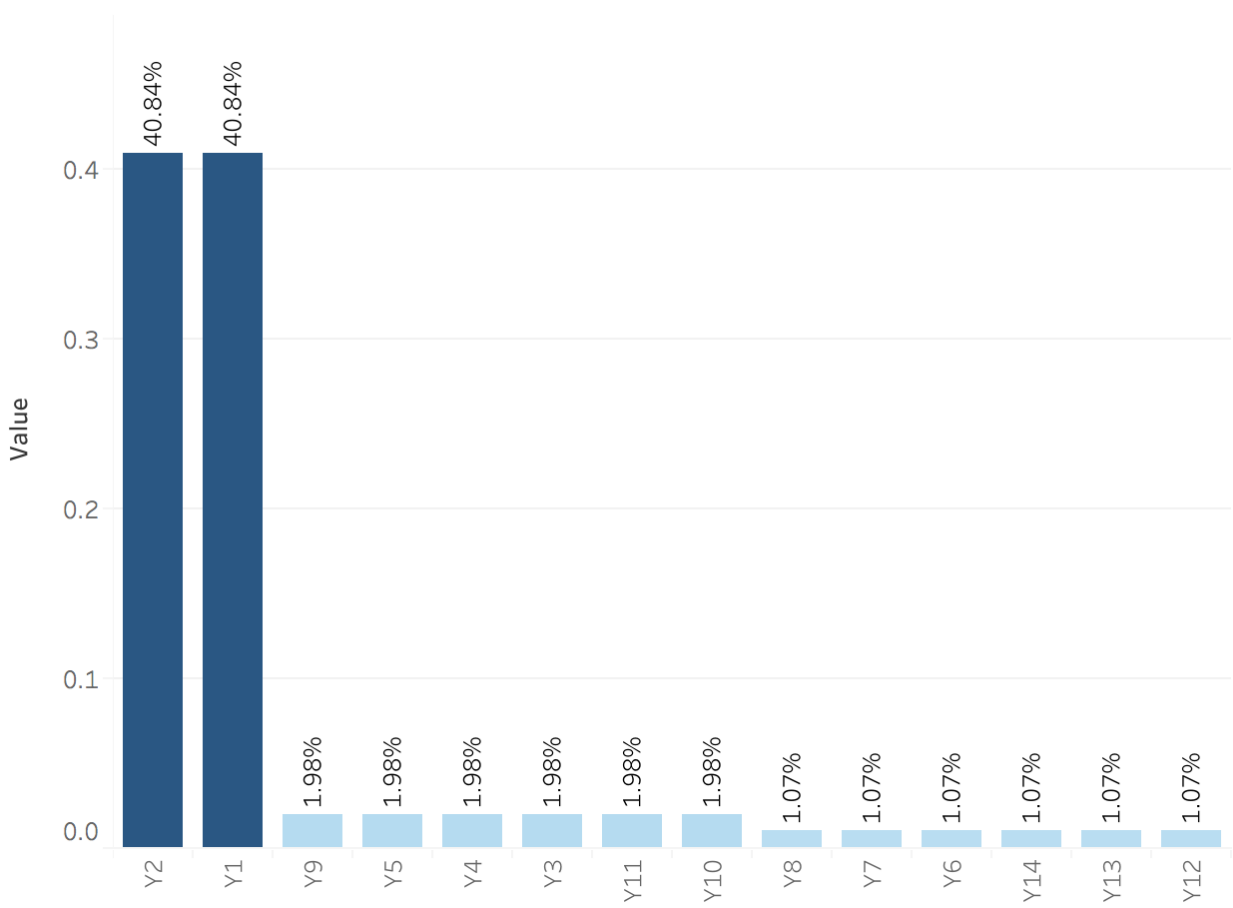


Figure 2.2: Problem Illustration Results Using Standard PageRank

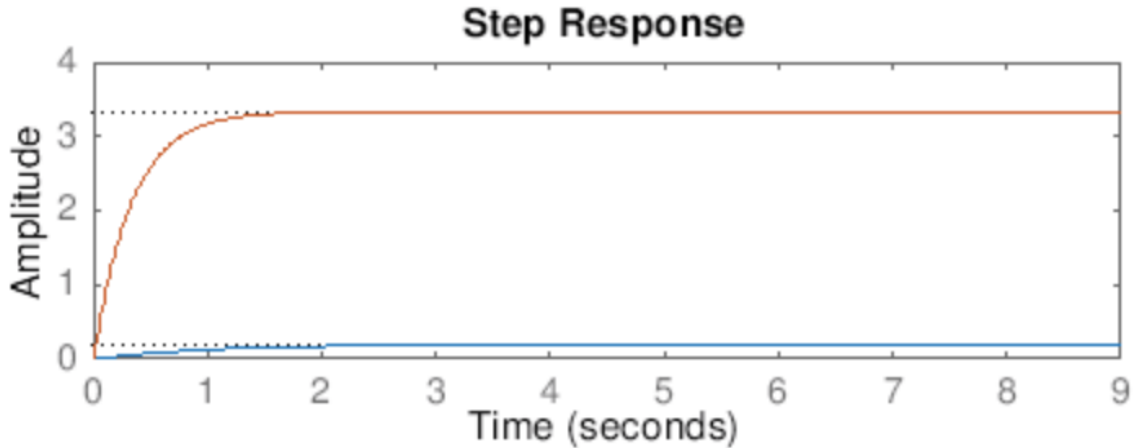


Figure 2.3:  $tf_1$  and  $tf_2$  step responses compared (orange is  $tf_1$  and blue is  $tf_2$ )

is not fully representing the effect or change that node  $Y_2$  is evoking across the network. After all,  $Y_1$  and  $Y_2$  are clearly not equivalent in their impact on the network with a value of 40.84%.

Social science research by Cialdini [17] suggests that influence evokes a change that causes other individuals to change their perspectives, opinions and overall emotional state [4]. Hence, influence is highly coupled with the concept of responsiveness and change. If somebody does not respond to a candidate “influencer” (i.e. their opinions, beliefs, and overall emotional state are not changing as a result of the influencer, then the candidate influencer is not truly influential). Using this social science conceptual model as a foundation for this paper, we will generally define influence as responsiveness and create a new responsiveness-based method for measuring influence in linear time invariant systems. This is evidenced by the step response (Figure 2.3) for  $tf_1$  and  $tf_2$ .

Being able to solve this problem with a responsiveness-based approach has a wide range of applications. In previous work of ours, we created dynamical networks that modeled the influence of social media accounts on one another on Twitter[16]. In the diagram from Figure 1.4 for example, there are 400 possible edges and each of the edges could have an infinite number of functions and values. The raw data itself is difficult to work with, but even



with the visualization it is nearly impossible to look at the diagram and easily understand which node in the network is the most influential, especially when you account for 2nd order and 3rd order ripple effects on the network. PageRank would not be able to clearly quantify which social media account in this network is the most influential since it is fully connected. Each of these edges using PageRank would be treated a binary edge and hence the influence would be evenly distributed across all nodes. However, if we define and apply a new responsiveness based approach to solving this problem, we would be able to discover which node in the network is the most influential. A social media marketer or a political strategy team could use this team to target and collaborate with specific influencers in the audience segment that is most relevant to their area of interest.

In the following section we will introduce two new methods for measuring influence that are designed for dynamical network applications. The first is Dynamical Responsive Page Rank (DRPR) and the second is Aggregate Targeted Reachability (ATR). Following this, we will review computed results using these methods and finally we will illustrate how these methods more intuitively model the network from the problem illustration detailed in this section.

## Chapter 3

### New Approaches to Measuring Influence in Linear Dynamical Networks

In this research, we introduce two new methods for measuring influence in Linear dynamical networks. We will begin with introducing an adaptation of PageRank that incorporates responsiveness for dynamical networks. After that, we will explain another model for measuring influence called aggregated targeted reachability (ATR).

#### 3.1 Page Rank

A simple version of PageRank introduced by Page and Brin [37] is defined as

$$\mathcal{PR}(u) = c \sum_{v \in B(u)} \frac{PR(v)}{N_v}$$

where  $u$  is a web page and,  $B(u)$  is the set of pages that point to  $u$ .  $PR(v)$  are the PageRank scores for page  $u$  and  $v$ .  $N_v$  is the number of outgoing links of page  $v$  and  $c$  is a factor that is used for normalizing the results across all nodes in the network that are being relatively compared in terms of importance and influence. For all the nodes that node  $p$  points to, the score is evenly divided and distributed among them in their respective calculations of  $PR(p)$ . When calculating the PageRank scores for a set of pages, the scores could be calculated iteratively starting with any of the candidates nodes. However, there are cases that there is a loop (which is common especially in a web search context where multiple pages on a single website point back to each other to form a loop). If these pages are linked by other web pages outside the loop, then they accumulate rank but never distribute scores. This scenario is called a rank sink problem [37].

### 3.2 Rank Sink Problem

Not all users on a website visit all existing links on a web page. After visiting a page, the user may not choose to visit any of the connected links and instead choose to go to an entirely new website in the relevant content domain. To solve this problem, a damping factor is introduced, which allows us to draw implicit connecting lines in the network to simulate the behavior of a web visitor that jumps from one page to another that is completely unrelated. These implicit lines don't only allow us to better model the user behavior, but it allows us to more holistically compare all the nodes in the network to one another, even though in reality they are not connected to each other. To dampen the affect of an edge in the network that in reality doesn't exist, a dampening factor is used. The damping factor  $d$  is often set to 0.85 and is incorporated as such:

$$\mathcal{PR}(u) = (1 - d) + d \sum_{v \in B(u)} \frac{PR(v)}{N_v}$$

Using this algorithm, the scores are iteratively computed for all nodes in target network until the calculations converge to a reasonable tolerance. For the use case of web search, this algorithm has been widely adopted for its usefulness. The search results are typically highly relevant and the calculation normally converges within  $\log(n)$  time. However, in the world of dynamical networks, the links are not binary, and the level to which one node influences another cannot be assumed to be equal. Before we build upon the PageRank algorithm to create an improved version to handle aggregated influence computation for linear time invariant systems, we will introduce the types of networks that we will use.

### 3.3 Responsiveness to Model Various Degrees of Relationships

In the domain of web search all relationships between one node and another are treated as equal (they are all given a binary value). This model has limitations when translated into scenarios where relationships are unequal. For example, an individual may be connected

with many people on a social media network. However, on a day to day basis this person interacts much more frequently with family members and close friends. For the use case of analyzing and computing influence in social dynamical situations (such as twitter) or even regular human-to-human interaction, a new system is needed to accurately model these various degrees of influence. For example, just because person  $A$  has one conversation with person  $B$  and another conversation of similar length with person  $C$  does not mean that each of those interactions influenced person  $A$  to the same degree. Based on social science research from [2], the level of responsiveness is highly reflective of the depth of influence that a person has.

### 3.4 Dynamical Responsive PageRank (DRPR)

Now we will demonstrate how we will incorporate responsiveness into the PageRank calculation. We will ground the discussion in a concrete use case from previous research in order to illustrate the usefulness of this method. In previous research we demonstrated how natural language data from a social media network could be computed to find causal relations between various accounts on Twitter using network reconstruction [16]. Let us assume we have already applied a network reconstruction algorithm in order to obtain a dynamical structure function with an equation  $Q$  and  $P$ . We will assume that there are no explicit inputs and as such  $P = 0$ .  $Q$  and  $P$  will have the form

$$Q = \begin{bmatrix} 0 & q_{12} & q_{13} & \dots & q_{1m} \\ q_{21} & 0 & q_{23} & \dots & q_{2m} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ q_{m1} & q_{m2} & q_{m3} & \dots & 0 \end{bmatrix}$$

$$P = \begin{bmatrix} 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 0 \end{bmatrix}$$

and where each value of  $Q$  is a transfer function that is a rational function of the Laplace variable  $s \in \mathbb{C}$ . From here there are several approaches we can use to incorporate PageRank. Since PageRank uses binary ranks, we could compute the binary adjacency state matrix.

$$P^{bin} = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & 1 \end{bmatrix} \quad \text{and} \quad Q^{bin} = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \dots & \dots & 0 & 1 \\ 1 & 0 & \dots & \dots & 0 \end{bmatrix}$$

where each value 1 indicates that there is non-zero transfer function and that there is some causal relationship. This is a binary matrix and PageRank is designed to work with binary links. The problem illustration example uses this binary approach. If, however, we want to create an algorithm that weights the influence transfer functions more “fairly” we will compute the norm for each individual index inside of the  $Q$  matrix. Thus we will compute  $\|q_{ab}\|$  for each entry inside of  $Q$  and obtain the  $Q_{Influence}$  matrix:

$$Q_{Influence} = \begin{bmatrix} 0 & \|q_{12}\| & \|q_{13}\| & \dots & \|q_{1m}\| \\ \|q_{21}\| & 0 & \|q_{23}\| & \dots & \|q_{2m}\| \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \|q_{m1}\| & \|q_{m2}\| & \|q_{m3}\| & \dots & 0 \end{bmatrix}$$

Where each individually normed transfer function represents an overall summary of the node-to-node influence, irrespective of time. If we then take the column sum for each index then we will have

$$Influence_q = \sum_{n=1}^m \|q_{mn}\|$$

for each individual node in the network. This *Influence* column sum can be computed for each column, corresponding to each node in the network.

$$INF_{network} = Influence_1, Influence_2, \dots, Influence_m$$

And these column sum measure for the total influence for each node in the network could be normalized. Then, when computing pageRank, this normalized influence weight can be applied to the page rank algorithm as such

$$DRPR(u) = (1 - d) + d \sum_{v \in B(u)} \frac{INF_v * PR(v)}{N_v}$$

such that the PageRank score is offset by the influence weighted-score that is computed. Thus the edges are not being treated equally and nodes with a higher score measured by their normalized transfer function influence will have a strong page rank score and vice versa. This method of computing the influence of a node is called Dynamical Responsive PageRank (DRPR) and is one of the novel contributions of this research. In an upcoming section we will share technical results using this method and compare it to alternative influence measurement techniques. In the next section we will introduce a 2nd candidate methodology for computing responsiveness-based influence. We call it Aggregated Targetted Reachability (ATR).

### 3.5 Aggregated Targeted Reachability (ATR)

Next we will introduce another algorithm for measuring influence in dynamical networks. We will leverage the same  $Q$  and  $P$  construct discussed earlier. Let us examine

$$Y = QY + PU$$

and take note that  $P$  is the identity matrix. Now we can rearrange the equation in order to obtain the net effect:

$$Y - QY = PU.$$

Rearranging we obtain

$$Y = (I - Q)^{-1}PU$$

and then finally obtain the matrix

$$G = (I - Q)^{-1}PU.$$

Where all  $U$ 's were just copies of  $Y$ 's and  $P$ 's in the identity. This essentially computes the net effect from and  $Y_i$  to  $Y_j$  which is the overall net effect that any node  $i$  has on any node  $j$ . This accounts for all the ripple effects of influence from node  $i$  to node  $j$ .

This effectively is another measurement of influence where we measure one node's overall ability to influence another node through the network. It is an influence metric, but not a metric that compares the source nodes influence overall throughout the whole network. However, this can be used for another use case where we measure one nodes "influence" over another node and can answer questions such as "which node has the greatest influence over node  $X$ ?"

We can then take the norm of each entry of the transfer function matrix,  $G$  in order to get a real number that summarizes this effect. This gives us a score that we can call

$$S_{ij} = \|G_{ij}\|$$

and sum up all of the normed scores for node  $i$  going to all other nodes in the network. Thus we obtain

$$ATR_i = \sum_{j=1}^m \|G_{ij}\|$$

which gives us a number of how node  $i$  impacts all the other nodes in the network. We can use this number as is, or we can normalize it so that it can more easily be compared side-by-side with DRPR (since this way the results for all nodes will also add up to 1).



## Chapter 4

### Technical Results

Now that we have introduced DRPR and ATR we will demonstrate technical results using these methods and discuss trade-offs.

#### 4.1 Simple Canonical Networks

We will define a handful of canonical examples for which we will execute DRPR and ATR. Let us define

$$TF_1 = \frac{1}{s+1}$$
$$TF_2 = \frac{1}{s^2 + 5s + 6}$$

where the 2-norms each compute as .7071 and 0.1291 respectively. Now, using these transfer functions, we will define the following canonical network examples and then compare results of DRPR and ATR.

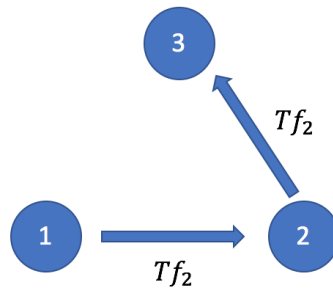


Figure 4.1: Simple Canonical Example 1

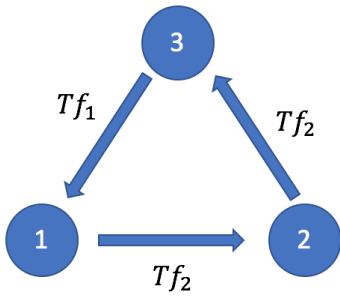


Figure 4.2: Simple Canonical Example 2

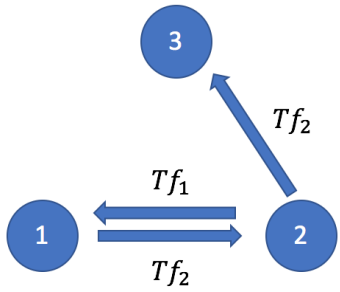


Figure 4.3: Simple Canonical Example 3

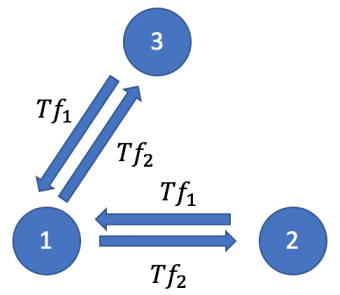


Figure 4.4: Simple Canonical Example 4

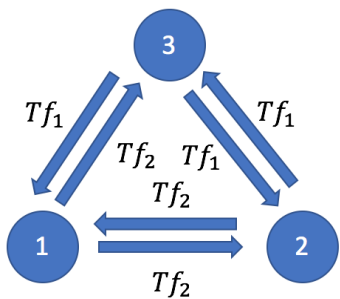


Figure 4.5: Simple Canonical Example 5

As a first step, we apply the 2-norm to each of the transfer functions. We used these as the edge weights of the DRPR algorithm. Here we present the results in the table below. Note that the relative influence of each node is a percentage of the total influence, so each row in the table sums to 1.

Example	Node 1	Node 2	Node 3
1	0.4489	0.3474	0.2036
2	0.4791	0.3253	0.1955
3	0.54936	0.3523	0.0982
4	0.57897	0.2105	0.2105
5	0.3489	0.4410	0.2099

Table 4.1: DRPR algorithm results with canonical examples

Next, we will evaluate the results of ATR for the same canonical examples in Table II.

Example	Node 1	Node 2	Node 3
1	1, 0, 0	0.71, 1, 0	0.5, 0.71, 1
2	1, 0.1, 0.13	.73, 1, .1	.53, .73, 1
3	1, 0, 0	.74, 1, .13	.55, .74, 1
4	1, 1, .14	.11, 1, .14	.11, 1, .14
5	1, .17, .27	.85, 1, .89	.89, .27, 1

Table 4.2: Results for Targeted Reachability

Summing the column values of Table 4.2, we obtain the ATR results found in Table 4.3.

Example	Node 1	Node 2	Node 3
1	.45	.35	.2
2	.44	.32	.24
3	.44	.34	.22
4	.46	.26	.26
5	.39	.30	.31

Table 4.3: Normalized ATR Results

Comparing ATR and DRPR, node 1 appears to generally be the most influential node across both methods, with the exception of Example 5. At a glance, it appears that node 1

should also be the most influential node for both methods because node 1 has the greatest sum of norms of the outgoing edges (a simplistic approximation of influence that seems to work for small examples). However, for the DRPR algorithm, the influence of node 1 appears to be diminished when the network is fully connected. It may be that this is due to the cyclic effects of PageRank and the timing of when the algorithm times out. Since, measuring influence is a subjective and slippery concept there is no ground truth against which we can compare these measurements, but as we introduce these new methods, we can intuitively argue which methods

### 4.2 Medium Sized Networks

The initial results that we just compared were some simple examples to get a good grasp of the general trends. Now we will do some more advanced examples. To make the networks more intuitively understandable at a glance, we will set the value of each transfer function to TF3 which has a value of  $q = 1/(s + 3)$  and a corresponding 2-norm of 0.4082. The purpose of these following examples is to analyze how the behavior of DRPR and normalized targeted reachability change as new edges are added to the network.

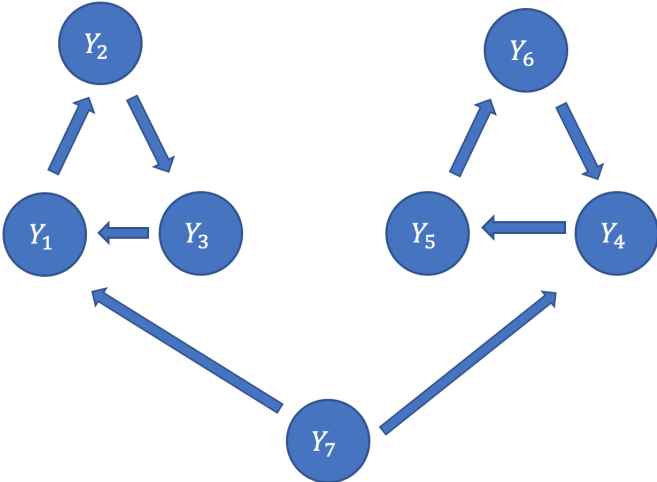


Figure 4.6: Medium Example 6

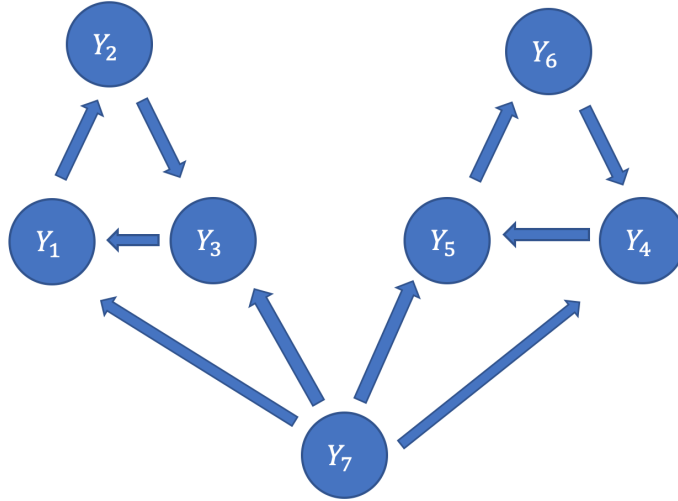


Figure 4.7: Medium Example 7

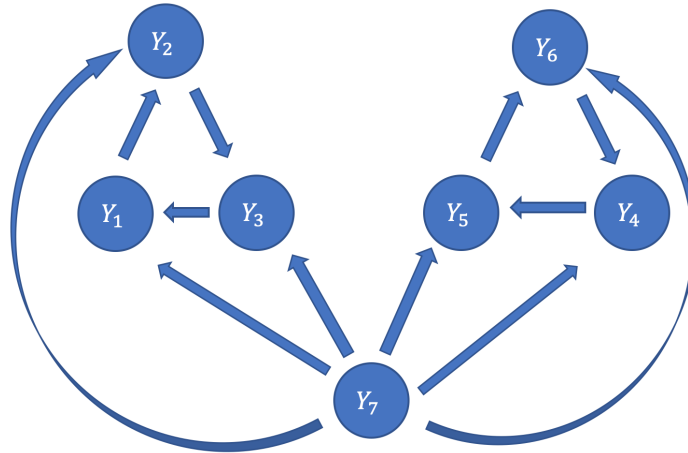


Figure 4.8: Medium Example 8

The only difference between canonical Examples 6 and 7 is that there are two additional edges that have been added from  $Y_7$  to  $Y_3$  and from  $Y_7$  to  $Y_5$ . Likewise, the biggest difference between canonical Example 7 and 8 is that there are two additional edges from  $Y_7$  to  $Y_2$  and from  $Y_7$  to  $Y_6$ .

Analyzing the results for ATR, it is clear that adding additional edges for  $Y_7$  increases its ability to influence other nodes and hence boosts the overall ATR score for  $Y_7$ . Adding the additional edges to node  $Y_7$  does not influence the ATR for the remaining six nodes in



Figure 4.9: ATR results for large canonical examples

the network. For ATR, unlike some versions of PageRank, having an incoming edge does not change the ATR for the node that is being pointed to. There are only increases in the ATR score for edges that come out of the source node. Incoming edges have no change on the score since the incoming edges do not immediately influence a destination node’s ability to influence other nodes.

However, for DRPR it is clear that adding the additional outgoing edges from a source node does influence the scores of all other nodes in the network. If we normalize the results from ATR then we get the graph in Figure 4.11. If you compare this to the DRPR results, the score of Y7 is very comparable across both algorithms. They both contain a score that is close to .30 and all other nodes have scores that is close to .10. However, for the network in Example 6 and 7, there appears to be more of a variety of dynamics across all other edges. We see that as we saturate a node with outgoing edges then the DRPR results begin to look more like the targeted ATR results in this scenario.

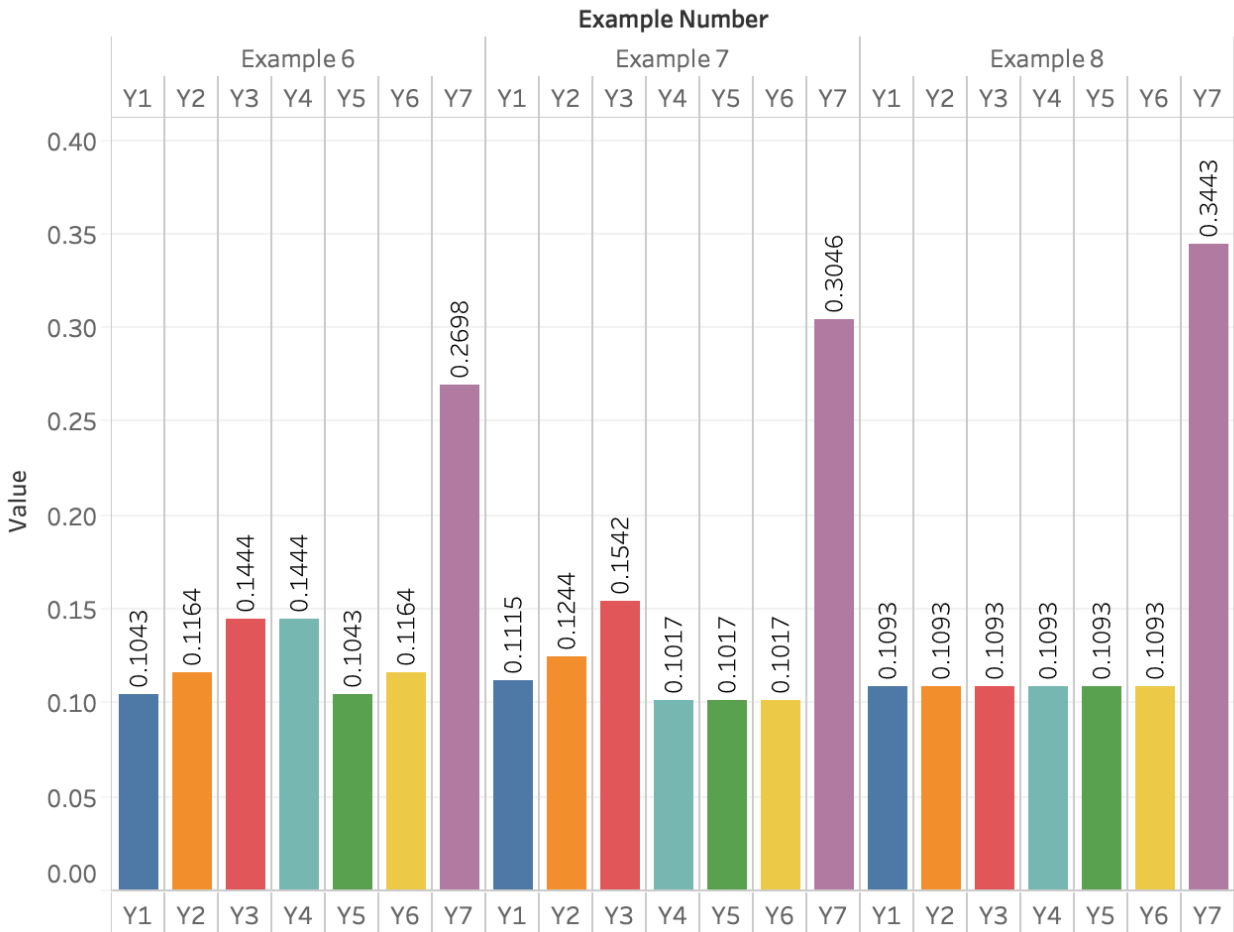


Figure 4.10: DRPR results for large canonical examples

### 4.3 Larger Network

On Figure 4.12, we see the graph of the larger network with the same default transfer functions where we ran DRPR and ATR. The results for DRPR are shown in Figure 4.13 and intuitively they seem to make sense given the structure of the network and given the fact that all transfer functions are the same. Likewise the results for ATR in Figure 4.14

### 4.4 DRPR and ATR Results for Problem Illustration Example Network

Finally, to illustrate the usefulness of these methods, we will compute results using ATR and DRPR for the problem illustration example network from Section 2. Computed results

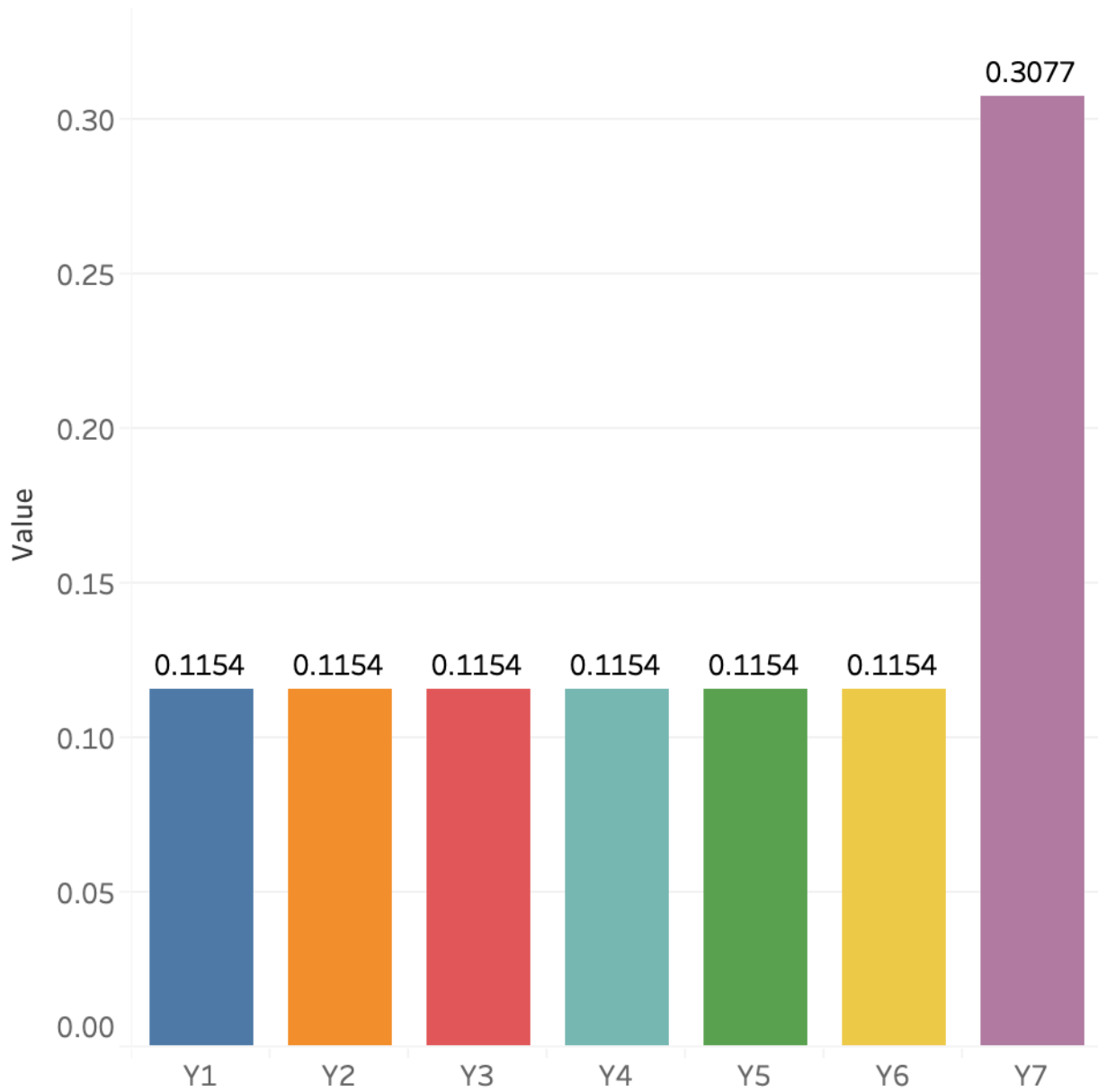


Figure 4.11: Normalized ATR results on large canonical Example 8

for DRPR are in Figure 4.15 and ATR results are shown in Figure 4.16. Analyzing and comparing the results, both ATR and DRPR both appear to model the influence behavior of the problem illustration example more intuitively than the regular binary PageRank results shown in Figure 2.2. Interestingly, it appears that normalized ATR appears to distribute the influence values more throughout the network, whereas with DRPR the influence value is more highly concentrated with Y2. Because of the iterative nature of PageRank computing,



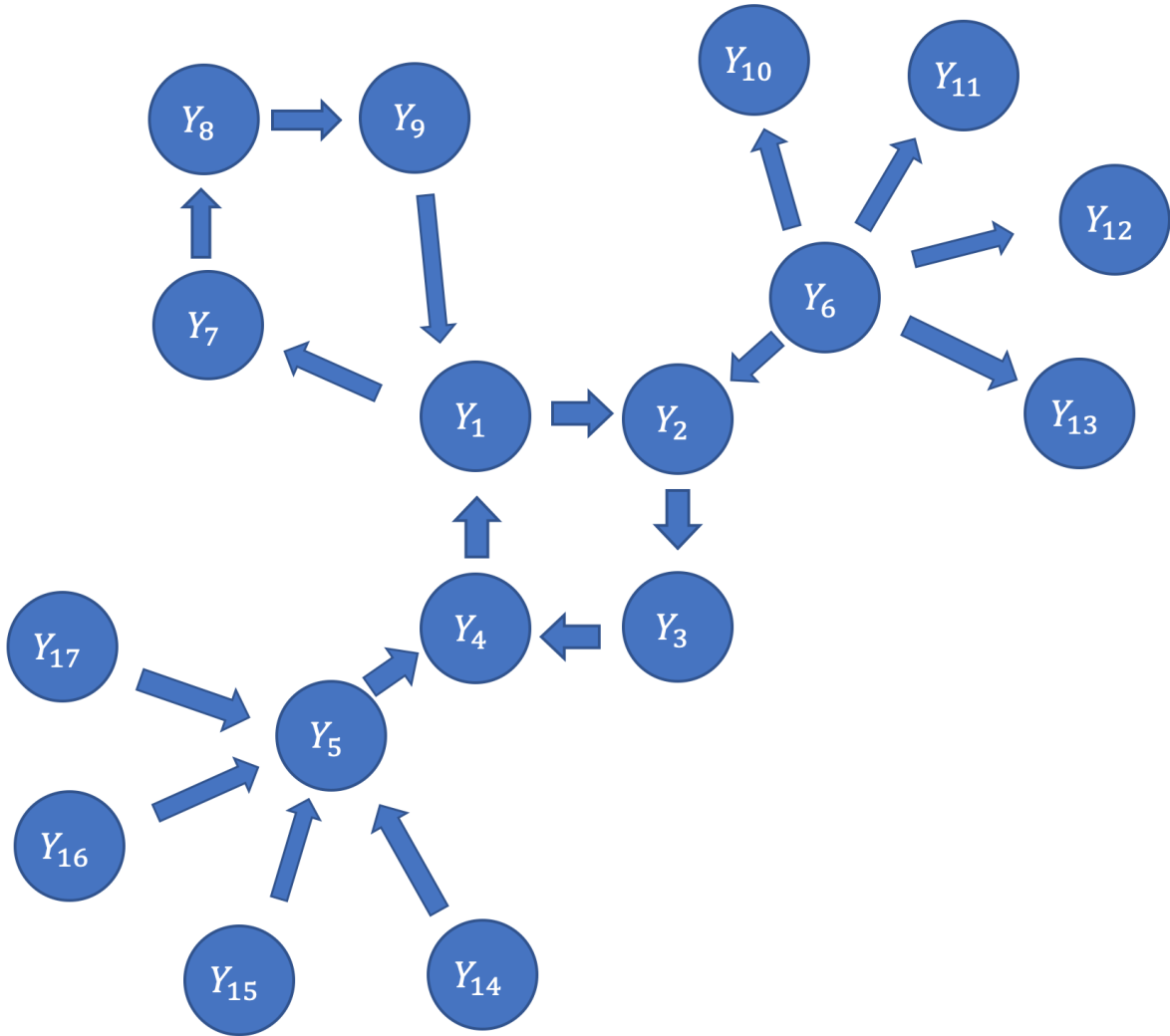


Figure 4.12: An influence diagram for a larger network

there is more opportunity for an influential node in the network to “claim” the influence value of other nodes in the network that have fewer outgoing edges. Likewise the results from the larger network example from Figure 4.12 appear to be more polarly distributed with DRPR whereas ATR appears to have a slightly more even distribution by comparison.

ATR Results on Larger Network

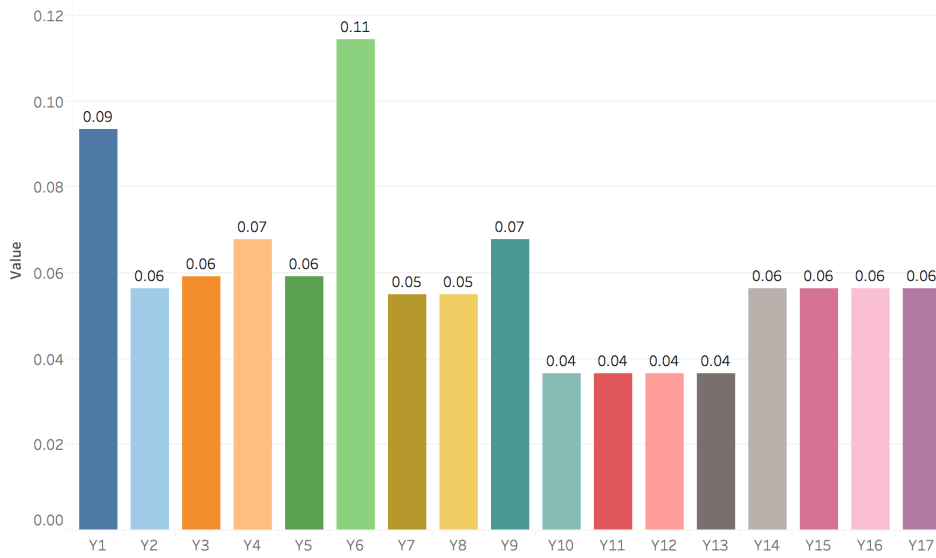


Figure 4.13: Normalized ATR results on larger network.

DRPR Results on Larger Network

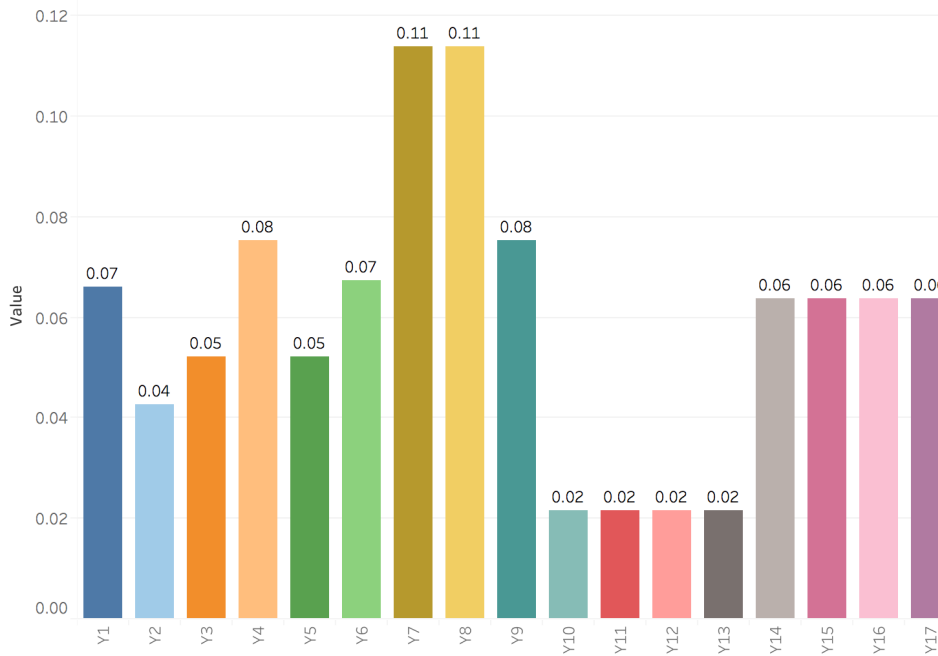


Figure 4.14: DRPR on a larger network

### DRPR Results for Problem Illustration

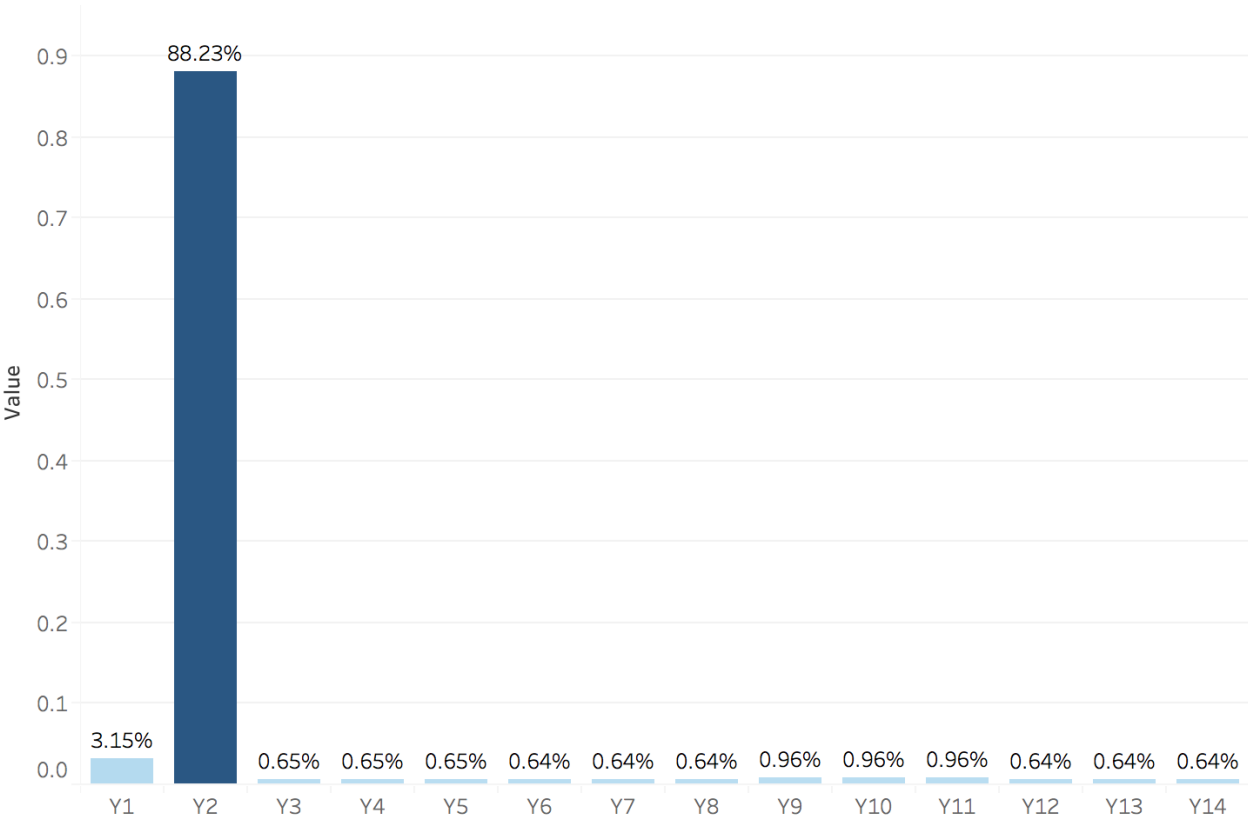


Figure 4.15: DRPR Results Computed on Problem Illustration Example from Section 2

### ATR Results for Problem Illustration

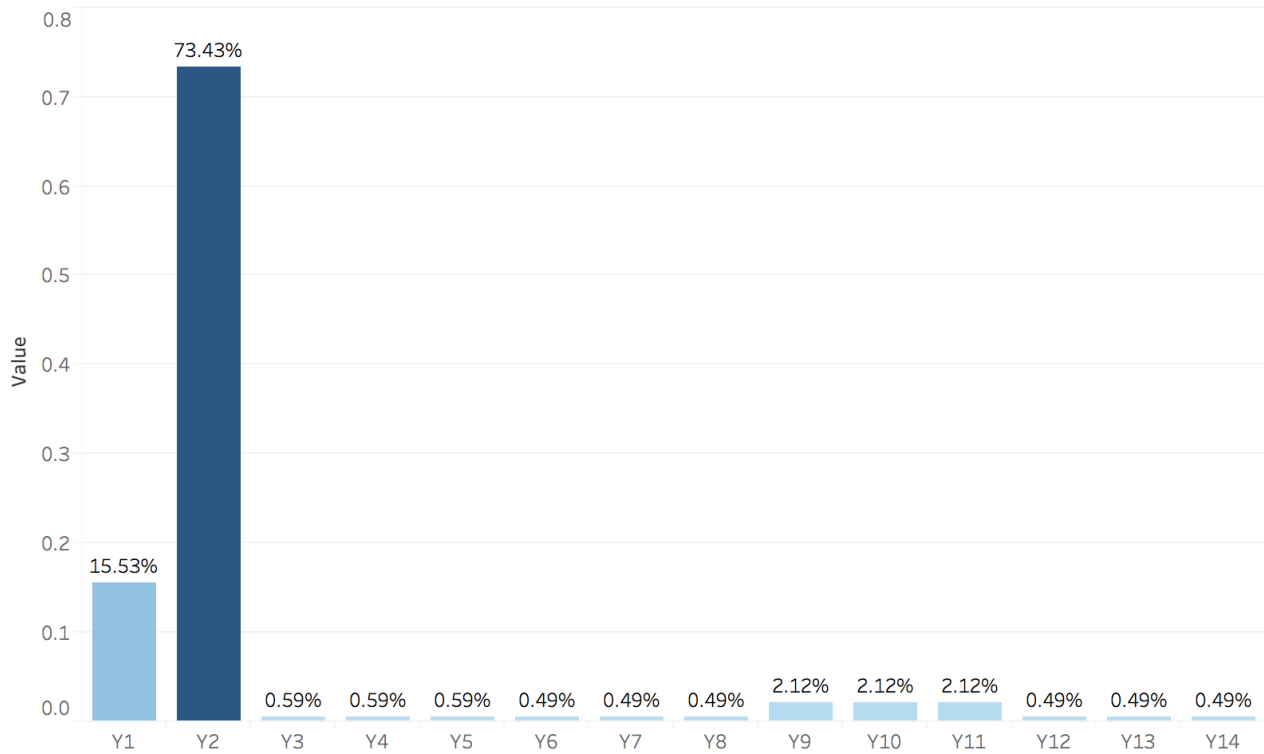


Figure 4.16: Normalized ATR Results Computed on Problem Illustration Example from Section 2

## Chapter 5

### Conclusions and Future Work

#### 5.1 Conclusion

In this work, we provide an overview of research across disciplines related to influence and how to measure influence. We show that for certain scenarios related to linear time invariant systems that existing approaches like PageRank do not fully represent different levels of influence that occurs across differing transfer function magnitudes. We introduce the DRPR algorithm which builds upon PageRank in order to incorporate a weighted link structure computation into the influence calculation. We also build upon systems theory concepts to introduce a second approach for modeling the influence inside of dynamical networks called targeted reachability. In the technical results, we show that these algorithms intuitively model the dynamical relations and more fully describe the various degrees of influence relationships in scenarios like the problem illustration.

Comparing the DRPR algorithm and ATR, it appears that for some scenarios the two approaches have very similar results (for example the network in Figure 4.8). However, for other networks such as the larger network (Figure 4.12) the algorithms appear to have quite different results. Overall, it appears that DRPR tends to accumulate more of the influence value from other nodes. Hence the results tend to be a bit more polarized with DRPR whereas with ATR they tend to be more evenly distributed.

It is difficult at this point to make a determination about which influence algorithm is “better” than the other. However, ATR appears to model the behavior of the larger network in Figure 4.12 more intuitively than does DRPR. It seems that DRPR doesn’t give  $Y_6$  that high

of a score because  $Y_{10}$ ,  $Y_{11}$ ,  $Y_{12}$ , and  $Y_{13}$  are sinks that don't have much influence. Instead it seems that DRPR allocates a much higher score to  $Y_7$  and  $Y_8$  because of their ability to influence other nodes in the network that are not sink nodes. Yet, it is arguable that  $Y_6$  should be given a higher score (as ATR allocates) because of  $Y_6$ 's ability to influence a larger quantity of nodes overall throughout the network.  $Y_7$  and  $Y_8$  have no ability to influence  $Y_{10}$  through  $Y_{13}$  and yet  $Y_6$  can influence  $Y_{10}$  through  $Y_{13}$  plus all the same nodes that  $Y_7$  and  $Y_8$  can influence. Hence it is our conclusion that ATR is the superior algorithm for the use-cases and examples that we have outlined in this research.

ATR also has an additional advantage over DRPR as well. Targeted Reachability (without aggregation) is very precise in its calculation of one single node's ability to influence another specific node. DRPR can only measure one nodes influence overall (not in respects to any single particular node). Comparing the normalized ATR results to that of DRPR on example 8, it is clear that there is a relationship between the two algorithms especially as you saturate individual nodes in the network and in certain network scenarios the output of DRPR and ATR appear to yield very similar results.

## 5.2 Future Work

The research of this thesis is the first of its kind that helps analyze aggregated influence across a linear time invariant network by introducing two new algorithms. Yet more work needs to be done to understand the mathematical relationship between these two algorithms. To date much of the influence modeling approaches happen in the world of static graphs. This research attempted to bridge the gap between the linear time invariant systems and the established static graph influence computation methods. There are likely other algorithms and other approaches to simplify the insight extraction process from working with dynamical networks. The analytical algorithms proposed in this work could also likely be extended or adapted to derive deeper insight other scenarios involving dynamical networks transfer

function modeling. Further work needs to be done to understand the advantages and the limitations of each of these algorithms.

## References

- [1] Julius Adebayo, Taylor Southwick, Vasu Chetty, Enoch Yeung, Ye Yuan, Jorge Goncalves, Julianne Grose, John Prince, Guy-Bart Stan, and Sean Warnick. Dynamical structure function identifiability conditions enabling signal structure reconstruction. In *2012 51st IEEE Conference on Decision and Control*, pages 4635–4641. IEEE, 2012.
- [2] Philippe Aghion and Jean Tirole. Formal and real authority in organizations. *Journal of Political Economy*, 105(1):1–29, 1997.
- [3] Ana Andrés. *Measuring academic research: How to undertake a bibliometric study*. Elsevier, UK, 2009.
- [4] Solomon E Asch. Studies of independence and conformity: I. A minority of one against a unanimous majority. *Psychological Monographs: General and applied*, 70(9):1, 1956.
- [5] Eytan Bakshy, Jake M Hofman, Winter A Mason, and Duncan J Watts. Everyone’s an influencer: quantifying influence on twitter. In *Proceedings of the fourth ACM International Conference on Web Search and Data Mining*, pages 65–74. ACM, 2011.
- [6] Eytan Bakshy, Itamar Rosenn, Cameron Marlow, and Lada Adamic. The role of social networks in information diffusion. In *Proceedings of the 21st International Conference on World Wide Web*, pages 519–528. ACM, 2012.
- [7] Harold Bloom et al. *The anxiety of influence: A theory of poetry*. Oxford University Press, USA, 1997.
- [8] Robert M Bond, Christopher J Fariss, Jason J Jones, Adam DI Kramer, Cameron Marlow, Jaime E Settle, and James H Fowler. A 61-million-person experiment in social influence and political mobilization. *Nature*, 489(7415):295, 2012.
- [9] Lutz Bornmann and Hans-Dieter Daniel. Does the h-index for ranking of scientists really work? *Scientometrics*, 65(3):391–392, 2005.
- [10] Lutz Bornmann and Hans-Dieter Daniel. What do we know about the h index? *Journal of the Association for Information Science and Technology*, 58(9):1381–1385, 2007.



- [11] Danah Boyd, Scott Golder, and Gilad Lotan. Tweet, tweet, retweet: Conversational aspects of retweeting on twitter. In *System Sciences (HICSS), 2010 43rd Hawaii International Conference on*, pages 1–10. IEEE, 2010.
- [12] Roger Brooksbank. The anatomy of marketing positioning strategy. *Marketing Intelligence & Planning*, 12(4):10–14, 1994.
- [13] Nauro F Campos and Francesco Giovannoni. Lobbying, corruption and political influence. *Public Choice*, 131(1-2):1–21, 2007.
- [14] Meeyoung Cha, Hamed Haddadi, Fabricio Benevenuto, and P Krishna Gummadi. Measuring user influence in twitter: The million follower fallacy. *International Conference on Web and Social Media*, 10(10-17):30, 2010.
- [15] V. Chetty, N. Woodbury, J. Brewer, K. Lee, and S. Warnick. Applying a passive network reconstruction technique to twitter data in order to identify trend setters. In *IEEE Conference on Control Technology and Applications*, Kohala Coast, 2017.
- [16] Vasu Chetty, Nathan Woodbury, Jacob Brewer, Kenneth Lee, and Sean Warnick. Applying a passive network reconstruction technique to twitter data in order to identify trend setters. In *2017 IEEE Conference on Control Technology and Applications*, pages 1344–1349. IEEE, 2017.
- [17] Robert B Cialdini. *Influence: Science and practice*, volume 4. Pearson Education, Boston, 2009.
- [18] Harald Cramér. A contribution to the theory of statistical estimation. *Scandinavian Actuarial Journal*, 1946(1):85–94, 1946.
- [19] J Joseph Cronin Jr, Michael K Brady, and G Tomas M Hult. Assessing the effects of quality, value, and customer satisfaction on consumer behavioral intentions in service environments. *Journal of Retailing*, 76(2):193–218, 2000.
- [20] Dwane Hal Dean. Consumer perception of corporate donations effects of company reputation for social responsibility and type of donation. *Journal of Advertising*, 32(4): 91–102, 2003.
- [21] Paul A Djupe and Christopher P Gilbert. *The political influence of churches*. Cambridge University Press, 2008.

- [22] Jorge Gonçalves and Sean Warnick. Necessary and sufficient conditions for dynamical structure reconstruction of LTI networks. *IEEE Transactions on Automatic Control*, 53(7):1670–1674, 2008.
- [23] Clive WJ Granger. Causality, cointegration, and control. *Journal of Economic Dynamics and Control*, 12(2-3):551–559, 1988.
- [24] Priscilla E Greenwood and Michael S Nikulin. *A guide to chi-squared testing*, volume 280. John Wiley & Sons, 1996.
- [25] Derek Hansen, Ben Shneiderman, and Marc A Smith. *Analyzing social media networks with NodeXL: Insights from a connected world*. Morgan Kaufmann, USA, 2010.
- [26] Ronald A Howard and James E Matheson. Influence diagrams. *Decision Analysis*, 2(3):127–143, 2005.
- [27] M Kent Jennings and Richard G Niemi. *Political character of adolescence: The influence of families and schools*. Princeton University Press, USA, 2015.
- [28] Timothy A Judge and Robert D Bretz Jr. Political influence behavior and career success. *Journal of Management*, 20(1):43–65, 1994.
- [29] Christine Kiss and Martin Bichler. Identification of influencers measuring influence in customer networks. *Decision Support Systems*, 46(1):233–253, 2008.
- [30] Alexander Kraskov, Harald Stögbauer, and Peter Grassberger. Estimating mutual information. *Physical Review E*, 69(6):066138, 2004.
- [31] H. Leung and S. Warnick. Necessary and sufficient conditions on state transformations that preserve the causal structure of LTI dynamical networks. In *IEEE Conference on Decision and Control*, Miami, FL, 2018.
- [32] Yung-Ming Li, Chia-Hao Lin, and Cheng-Yang Lai. Identifying influential reviewers for word-of-mouth marketing. *Electronic Commerce Research and Applications*, 9(4):294–304, 2010.
- [33] Yung-Ming Li, Cheng-Yang Lai, and Ching-Wen Chen. Discovering influencers for marketing in the blogosphere. *Information Sciences*, 181(23):5143–5157, 2011.
- [34] Naresh K Malhotra, James Agarwal, and Mark Peterson. Methodological issues in cross-cultural marketing research: A state-of-the-art review. *International Marketing Review*, 13(5):7–43, 1996.

- [35] Me. This is not a real paper. In *Proceedings of My Life (ML)*, 2006.
- [36] Diana C Mutz. *Impersonal influence: How perceptions of mass collectives affect political attitudes*. Cambridge University Press, 1998.
- [37] Lawrence Page, Sergey Brin, Rajeev Motwani, and Terry Winograd. The pagerank citation ranking: Bringing order to the web. Technical report, Stanford InfoLab, 1999.
- [38] J Paul Peter, Jerry C Olson, and Klaus G Grunert. *Consumer Behaviour and Marketing Strategy*. McGraw-Hill London, 1999.
- [39] Daniel M Romero, Wojciech Galuba, Sitaram Asur, and Bernardo A Huberman. Influence and passivity in social media. In *Proceedings of the 20th International Conference Companion on World Wide Web*, pages 113–114. ACM, 2011.
- [40] Lester M Salamon and John J Siegfried. Economic power and political influence: The impact of industry structure on public policy. *American Political Science Review*, 71(3): 1026–1043, 1977.
- [41] Bernd Schmitt. Experiential marketing. *Journal of Marketing Management*, 15(1-3): 53–67, 1999.
- [42] Alvin J Silk and Manohar U Kalwani. Measuring influence in organizational purchase decisions. *Journal of Marketing Research*, 19(2):165–181, 1982.
- [43] Stephen M Stigler. *The history of statistics: The measurement of uncertainty before 1900*. Harvard University Press, USA, 1986.
- [44] Michael Trusov, Anand V Bodapati, and Randolph E Bucklin. Determining influential users in internet social networks. *Journal of Marketing Research*, 47(4):643–658, 2010.
- [45] A. Vosughi, C. Johnson, S. Roy, S. Warnick, and M. Xue. Local control and estimation performance in dynamical networks: Structural and graph-theoretic results. In *Conference on Decision and Control*, Melbourne, Australia, 2017.
- [46] A. Vosughi, C. Johnson, M. Xue, S. Roy, and S. Warnick. Target control and source estimation metrics for dynamical networks. *Automatica*, 100:412–416, February 2019.
- [47] Jianshu Weng, Ee-Peng Lim, Jing Jiang, and Qi He. Twitterrank: finding topic-sensitive influential twitterers. In *Proceedings of the third ACM International Conference on Web Search and Data Mining*, pages 261–270. ACM, 2010.

- [48] Frank Yates. Contingency tables involving small numbers and the  $\chi^2$  test. *Supplement to the Journal of the Royal Statistical Society*, 1(2):217–235, 1934.
- [49] E. Yeung, J. Goncalves, H. Sandberg, and S. Warnick. Mathematical relationships between representations of structure in linear interconnected dynamical systems. In *American Control Conference*, San Francisco, 2011.
- [50] Ye Yuan, Guy-Bart Stan, Sean Warnick, and Jorge Goncalves. Robust dynamical network structure reconstruction. *Automatica*, 47(6):1230–1235, 2011.
- [51] Ye Yuan, Anurag Rai, Enoch Yeung, Guy-Bart Stan, Sean Warnick, and Jorge Goncalves. A minimal realization technique for the dynamical structure function of a class of lti systems. *IEEE Transactions on Control of Network Systems*, 4(2):301–311, 2015.