The Adoption of Electronic Medical Records by U.S. Hospitals: An Exploration of Network Methods and Models

A Dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at George Mason University

by

Yinyue Hu
Master of Arts
Georgetown University, 2010
Bachelor of Arts
University of Shanghai for Science and Technology, 2008

Director: Laurie Schintler, Associate Professor
School of Policy, Government and International Affairs

Spring Semester 2016
George Mason University
Fairfax, VA

DEDICATION

This dissertation is dedicated to my grandfather, YIN Guangqi, who had always inspired me to never stop learning.

ACKNOWLEDGEMENTS

First and foremost, I would like to express my deep gratitude and appreciation to my committee members, without whose persistent help this dissertation would not have been possible. I owe a debt of gratitude to my committee chair Professor Laurie Schintler for her guidance and extreme patience from the very early stage of my doctoral study through the completion of this dissertation. I would like to thank my committee member Professor Robert Axtell, who introduced me to agent based modeling and complexity theories, for his friendly guidance and thought-provoking suggestions. My earnest thank you goes to my committee member and supervisor for nearly six years, Professor Arnauld Nicogossian, for his mentorship throughout my doctoral study. I am beyond grateful to have had the opportunity to work with him. His continued direction and support is what makes me survive the doctoral study. He has taught me what a true researcher should be in innumerable ways.

Also, I would like to thank my external reader Professor Azhar Rafiq from Virginia Commonwealth University for offering his help to review my dissertation. I would also like to recognize Professors Naoru Koizumi, Phil Auerswald and Ray Sommer for their input during my years of study at George Mason University. I would like to thank my friends and colleagues at George Mason University for their feedback, support and friendship. Also, a thank you to Professor Charles Doarn from University of Cincinnati for all the constructive discussions throughout the years.

This dissertation would not be possible without the Dissertation Completion Grant from the Office of the Provost at GMU during my last semester of the study. In a similar vein, I would like to thank the School of Policy, Government and International Affairs for the financial support granted at various stages of this dissertation. I would like to acknowledge the HIMSS Foundation for granting me the access to the research data.

I would like to thank my parents for continuously encouraging me in all my pursuits. They always put education at the first priority and have faith in me. Finally, I would like to thank my husband, Kai He, who supported me in every possible way to see the completion of this work. Thank you for being such a great life and intellectual partner.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

Electronic Medical Records..EMR
Agent Based Modeling .... ABM
Health Information Technology...HIT
Healthcare Information and Management Systems Society .... HIMSS
American Recovery and Reinvestment Act... ARRA
Health Information Technology for Economic and Clinical Health.......................HITECH
Clinical Data Repository.. CDR
Nursing Documentation ....DOC
Electronic Medication Administration Records..EMAR
Clinical Decision Support..CDS
Computerized Physician Order Entry .... CPOE
Full-time Equivalent .... FTE
American Hospital Association ....AHA
Distance-based Spatial Clustering of Application with Noise.............................. DBSCAN

ABSTRACT

THE ADOPTION OF ELECTRONIC MEDICAL RECORDS BY U.S. HOSPITALS: AN EXPLORATION OF NETWORK METHODS AND MODELS

Yinyue Hu, Ph.D.

George Mason University, 2016

Dissertation Director: Dr. Laurie Schintler

Why do people adopt innovations at different rate? Studies from the diffusion of innovations and network models suggest that network structure and properties provide good explanations of the influence mechanism about how attitudes and behaviors change. This study investigates the adoption of Electronic Medical Records (EMR) among U.S. hospitals with network methods and models. Three research questions are addressed. First, what is the structure of the networks among hospitals? Second, how does the presence of network contribute to the diffusion of EMR? And third, what network-based policies can accelerate the adoption? The study employs network analysis tools, event history models and agent based modeling to present the structure of the network, its role in the diffusion process and to test different policy scenarios.

CHAPTER 1 INTRODUCTION

Health Information Technologies (HIT) have been considered of great importance to transform health care industry by health providers, consumers, and policy makers (MF Smith 2004; Hillestad et al. 2005). Many believe that HIT has the potential to enable safer, more effective and efficient delivery of health care with better quality (Agency for Healthcare Research and Quality 2006; Wu et al. 2006; McCullough 2008; Bodenheimer and Grumbach 2003). Among the varieties of health IT applications, Electronic Medical Records (EMR) system that “integrates electronically originated and maintained patient-level clinical health information derived from multiple sources into one point of access” and “replaces the paper medical record as the primary source of patient information” (American Hospital Association 2007, 19) has been one of the most promising components(Jha et al. 2009; Kazley and Ozcan 2007). The hope comes not only from the many wonders that information technologies have brought to the twenty-first century, but also from the fundamental role that information management plays in health care delivery (Chassin and Galvin 1998). It has been shown that the convergence of technological development, government policy and economy determines that HIT, and EMR in particular, is the wave of the future in the United States (Berner, Detmer, and Simborg 2005).

However, the implementation of EMR in the U.S. has not been without challenges

(Heisey-Grove et al. 2014). Although an overall trend of adopting the EMR has been observed (Heisey-Grove et al. 2014), it has also been shown that hospitals have been taking different trajectories towards EMR implementation at various rates (Angst et al. 2010). Explanation to the mechanisms behind the differences can inform us about the means to encourage the adoption and increase the welfare from the new technologies. The studies in the diffusion of innovation and network models suggest that the adoption of an innovation has to do with the relational influence from an individual's social network (Valente 1995; Jackson 2010). Network methods and models have thus been well employed to examine the heterogeneity of individuals through their links and positions in the network.

1.1 ELECTRONIC MEDICAL RECORDS

It is noted that there is no universally accepted definition of EMR (Jha et al. 2006). Consensus is that EMR considers the means to store, organize, and retrieve information about patient with information technologies for better delivery of health care. It is, in general, a data repository sharable within a health system and across physicians, insurance companies and other stakeholders (Angst et al. 2010). The notion of EMR does not indicate any single technology but rather the implementation of various technologies at different levels of hospital facilities. This may include Clinical Data Repository (CDR), Computerized Practitioner/Physician Order Entry (CPOE), Clinical Decision Support System (CDSS), and standardized clinical information transactions. Hospitals may take different trajectories to implement the technologies, but in general they proceed from adopting simple data retrieving technologies at separate ancillaries into fully integrating data sharing and analyzing systems at all hospital entities. The most commonly adopted measure of EMR adoption is the EMR Adoption Model developed by HIMSS Analytics (2015), as listed below in Table 1-1. HIMSS surveys U.S. hospitals each year regarding the implementation status of nearly 100 information technologies and assesses their state of EMR adoption. From Figure 1-1, one can tell that hospitals have been migrating towards higher degrees of EMR penetration. The percentage of hospitals at Stages 0, 1 and 2 has been decreasing. The number of Stage 4 hospitals reached its peak at 2009 but declines afterwards. In contrast, the number of hospitals at Stages 5 to 7 has been increasing dramatically over the past eight years. By end of year 2015, over sixty percent of U.S. hospitals have cumulative EMR capabilities of stage 5 or above.

Table 1-1 HIMSS EMR Adoption Model

Stage	**Cumulative Capabilities**
Stage 7	Complete EMR; CCD transactions to share data; Data warehousing; Data continuity with ED, ambulatory, OP
Stage 6	Physician documentation (structured templates), full CDSS (variance & compliance), full R-PACS
Stage 5	Closed loop medication administration
Stage 4	CPOE, Clinical Decision Support (clinical protocols)
Stage 3	Nursing/clinical documentation (flow sheets), CDSS (error checking), PACS available outside Radiology
Stage 2	CDR, Controlled Medical Vocabulary, CDS, may have Document Imaging; HIE capable
Stage 1	Ancillaries – Lab, Rad, Pharmacy – All Installed
Stage 0	All Three Ancillaries Not Installed

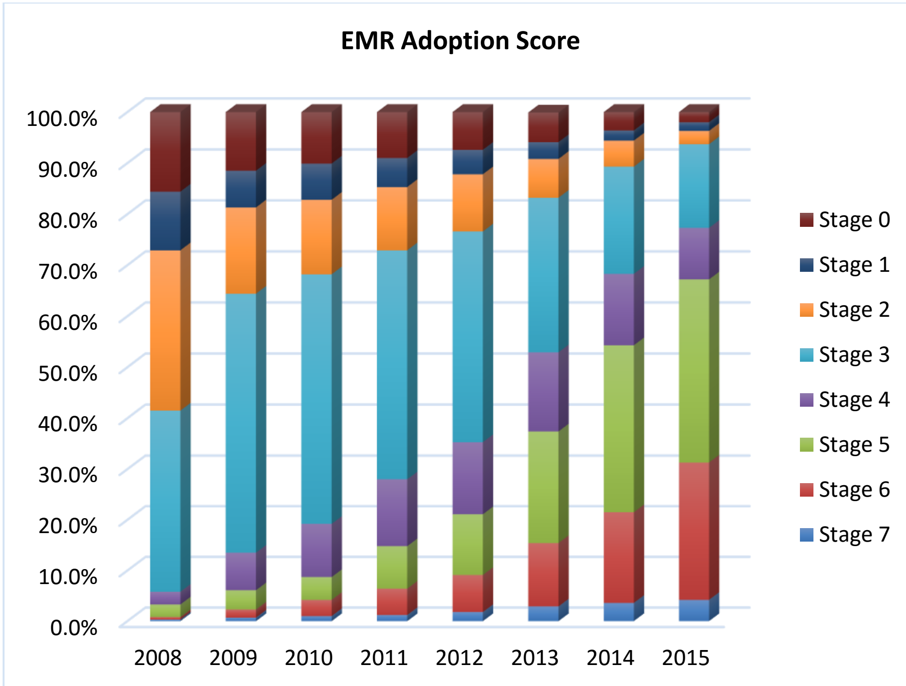


Figure 1-1 Summary of EMR Adoption Score of U.S. hospitals, 2008-2015. Adapted from HIMSS Analytics (2015)

By unifying the fragmented data and applications (Angst et al. 2010), the EMR is believed to offer the promise of more cost-effective, efficient and safer health care at higher quality that blends evidence-based clinical practice guidelines or protocols to manage complexity (James 2005; Kumar and Aldrich 2010). It is shown that the adoption of single or multiple component EMR technologies have enhanced the delivery of health care by improving the accuracy and completeness of the problem list (Galanter et al. 2010), preventing serious medication errors (Bates et al. 1998), facilitating evidence-based prescriptions by physicians (Davis et al. 2007), increasing physician's time efficiency (Pizziferri et al. 2005) and improving medication and non-medication health outcomes (Yu et al. 2009). A systematic review of articles published since 1998 further reveals that the adoption of EMR have positively impacted healthcare through improvement in communication among clinicians, care of certain diseases (cancer, sexually transmitted disease and diabetes), and cost savings as a result of better data management (Holroyd-Leduc et al. 2011).

Since the EMR involves storing information on a patient's health, medical history, conditions, tests, treatment, referrals, medications, demographic information and other non-clinical information (Kumar and Aldrich 2010) and employing evidence-based protocols in medical practices, the transition from paper to the automated system could take place by breaking the massive undertaking into a series of small steps (James 2005). Thus it is observed that the implementation of EMR usually proceeds from automated billing and schedule, to automated laboratory and imaging work and automation in pharmacy, to the adoption of an “electronic file cabinet, to an interoperable decision-support system (James 2005). At each step of the EMR adoption, adopters can be affected and restricted by different technical, organizational, social and policy contexts. As a result, it is reported that U.S. hospitals adopt the EMR at different rates (Angst et al. 2010; Jha et al. 2009; Kazley and Ozcan 2007).

1.2 RESEARCH QUESTIONS

The overall purpose of the study is to investigate the adoption of EMR by hospitals with the presence of network. As will be reviewed in following sections, network exists in many social and economic systems. The structures and properties of the networks determine how entities in the networks behave. The role of network in the diffusion of innovations in medical and health domain has long been studied (Coleman et al. 1966). Network-based interventions is well perceived by scholars and practitioners as tools to encourage the adoption of new medical practices (Fennell and Warnecke 1988). A survey of literature suggests that the notion of network has not been well studied on the adoption of EMRs. Thus, this study sets out with three research questions.

First, *what is the structure of the networks among hospitals' adoption of EMR in the US?* This question concerns the hospitals as connected by their hospital system affiliation and spatial proximity. The network of hospitals has been studied such as patient sharing networks (Lee et al. 2011). But studies of this kind are usually conducted at regional or local level due to the computational capacity to process the patient sharing data. Instead, this study focuses on two sources that construct the network among hospitals. The first is the linkages among hospitals by their hospital system affiliations. The second is the network constructed by their spatial proximity. The rationale about this construct is that information and influence, in the form of uncertainty reduction and peer pressure, is transmitted at both social and physical distance. Under this general research question about network structure, below three specific questions are asked:

- *RQ1-1: What is the network typology of hospitals' EMR network, organizational and spatial?*
- *RQ1-2: What are the hierarchical roles of hospitals? Who are the highly*

connected hospitals?

- *RQ1-3: What does the hospital network look like?*

Second, *how does the presence of network contribute to the diffusion of EMR in hospitals?* As Valente (2005) pointed out, studies in the diffusion of innovation originates from two general interests. One is the mechanisms underlying the effective dissemination of information about the innovations; the other is the explanation to the heterogeneity in individuals' adoption behavior – some adopt while some do not; some adopt sooner while some wait longer. Network methods and models allow one to examine the diffusion of innovation via the relational influence among individuals, since social factor is perceived a more important determinant to adoption (Ryan and Gross 1943). It is especially the case in EMR adoption. Because EMR involves major changes in the delivery of health care at different divisions of hospitals and by personnel of various skill and experience levels, the uncertainty associated is high. As a result, prior adopters serve as major source of uncertainty reduction and the adoption decision depends on how the relationship between adopting and non-adopting hospitals regulates the uncertainty (Angst et al. 2010). The influence can be transmitted through direct information and persuasion or indirect comparison or competition (Burt 1987). And two additional questions need to be addressed:

- *RQ2-1: Are hospitals with more connections more likely to adopt EMR?*
- *RQ2-2: Does the network exposure, direct and indirect, affect hospital's adoption of EMR?*

Finally, *what network-based policies can accelerate the EMR adoption?* The purpose of identifying individuals with different roles in a network is to introduce strategies targeting them in order to bring in behavioral change. Valente (2010) suggested that network interventions have been developed such as identifying opinion leaders/key players, groups and leaders within groups, rewiring the networks and strengthening critical nodes and links. The purpose of these interventions is to increase the non-adopters' exposure to innovation and change the construct of the network to maximize certain network properties. The interventions can be realized by policies such as providing incentives to central hospitals and/or targeted groups of hospitals, or facilitating communication between hospitals and health systems. Thus two additional questions are introduced:

- *RQ3-1: Can policies targeting hospitals with high centrality accelerate the diffusion of EMR?*
- *RQ3-2: What strategies can be used to alter the network structure in order to facilitate the adoption of EMR?*

1.3 RESEARCH DATA AND METHODOLOGY

The data for hospital's EMR adoption were obtained from Healthcare Information and Management Systems Society (HIMSS) Analytics database (HIMSS Analytics 2015). The HIMSS Analytics database, and its predecessor Dorenfest Integrated Healthcare Delivery Systems Plus (IHDS+), surveyed U.S. hospitals[1] since 1980s about their health information technology implementation status. The original purpose of the database was

[1] The complete HIMSS Analytics database reports EMR adoption at health facilities including ambulatory, free standing data center, home health, IDS/RHA, in-hospital data center, sub-acute and hospital. For the purpose of this study, only data of hospitals were obtained.

to provide vendors information about EMR market. Thus the database contains information about hospital's demographic information (e.g. name, address, parent hospital system, parent system type, population served), its EMR implementation plans and the software, hardware and infrastructure installed across all facilities at the hospital. As HIMSS has identified several component technologies within EMR, the data provides the installation status of each technology[2] and their vendor information. This data source has been widely used by scholars studying the adopting of EMR in hospitals (Angst et al. 2010; Furukawa, Raghu, and Shao 2011; N. J. Zhang et al. 2013; Fareed et al. 2015).

Access to the database was granted by HIMSS Analytics through an online application. The database is open to applicants with research and education purpose. The Dorenfest IHDS+ database covers results from 1986 to 2003. Data for 2004 onward is updated by the HIMSS Analytics through a donation from Dorenfest Institute for Health Information. For the purpose of this study, data will be obtained for years 2005 to 2013 because data of previous years only contains a small number of hospitals. The 2005 data included 4,010 hospitals. The number of hospitals being surveyed grows over time and as of year 2013, 5,419 hospitals were included in the survey. The HIMSS database does not contain data on governmental hospitals. Therefore, the analysis performed in this dissertation considers only non-federal, public or private hospitals. The original data set is only available in Microsoft Access format, thus it requires data transformation and cleaning

[2] HIMSS Analytics records the implementation status of an EMR component technology as under one of the following categories: 1) Contracted/ Not Yet Installed, 2) Installation in Process, 3) Live and Operational, 4) Not Automated, 5) Not Reported, 6) Not Yet Contracted, and 7) To be Replaced.

into the appropriate format.

Table 1-2 Summary of Data

Type	Data	Note
Demographic Information	Hospital Name	
	HAEntityID	Identifier across different data tables with same year
	UniqueID	Identifier across different years
	ParentID	Identifier of hospital system affiliation
	Longitude/Latitude	
EMR Elements	Laboratory	Basic EMR
	Pharmacy	
	Radiology	
	Clinical Data Repository (CDR)	
	Nursing Documentation (DOC)	Intermediate EMR
	Electronic Medication Administration Records (EMAR)	
	Clinical Decision Support (CDS)	Comprehensive EMR
	Computerized Physician Order Entry (CPOE)	
Characteristics	Type	
	Ownership Status	
	Number of Full-time Equivalent (FTE)[3]	
	Number of Beds	

This study followed the methodology suggested by Furukawa et al. (2010) and Fareed et al. (2015) and coded hospitals' EMR adoption status into three stages based on the implementation of eight core EMR element technologies at the hospitals: 1) *Basic*

[3] As discussed in Chapter 4, Number of FTE was not included in the final statistical analysis.

EMR, the hospital has implemented information systems at pharmacy, radiology and laboratory, as well as Clinical Data Repository (CDR), 2) *Intermediate* EMR, in addition to all the EMR element technologies in *Basic* stage, the hospital has implemented Nursing Documentation (DOC) and Electronic Medication Administration Records (EMAR), and 3) *Comprehensive* EMR, in additions to all applications in *Basic* and *Intermediate*, the hospital has completed implementing Clinical Decision Support (CDS) and Computerized Physician Order Entry (CPOE). It should be noted that the methodology to categorize hospitals' EMR capabilities is subject to the source of data, as organizations that collect the data employ different survey methods. For example, the American Hospital Association (AHA) Annual Survey is another source of EMR data. However, the EMR elements and terminologies being adopted in the AHA data is different from those in the HIMSS database. Thus, categorization methods derived from the AHA data are not applicable to studies using the HIMSS data (see, for example, Adler-Milstein et al. 2014).

In response to the three sets of research questions, the study proceeds in three sequential studies: a network analysis of the organizational-spatial network of hospitals, regression analysis on the role of network in the diffusion of EMR in hospitals, and an agent based model to explore network-based policy scenarios (Figure 1-2).

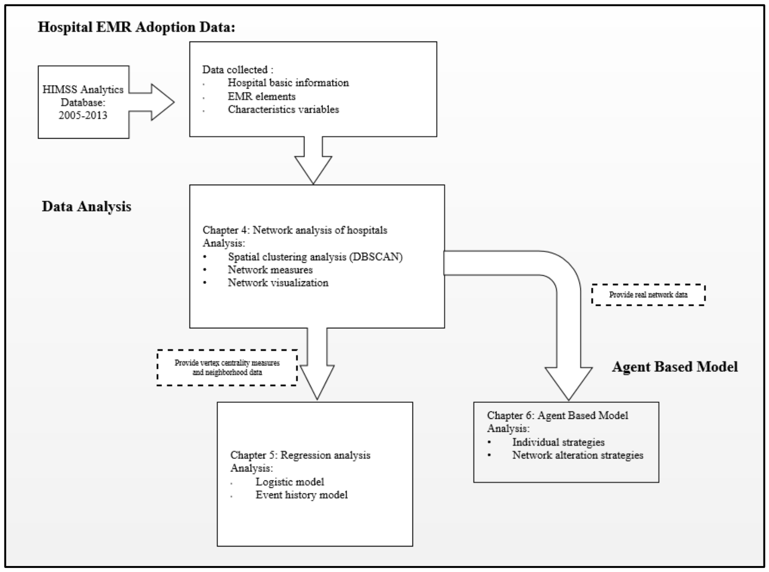


Figure 1-2 Flowchart of Methodology

1.4 ROADMAP OF THE DISSERTATION

The rest of this dissertation is organized as follows. Chapter 2 reviews literature of topics of empirical studies of EMR adoption, network methods and models, the diffusion of innovation and agent based modeling. The purpose of the literature review is to identify research gaps, as well as analytical and theoretical opportunities to explore the adoption of EMR by hospitals. Chapter 3 reports the network analysis of hospitals' spatial-organizational network. Chapter 4 addresses the role of network in the diffusion of EMR among hospitals with regression analysis. Chapter 5 explores possible policy scenarios using agent based models. Since each of the three studies uses a separated methodology, a methodology section is also included in each chapter. Chapter 6 summarizes the findings and discusses policy implications of this dissertation.

CHAPTER 2 LITERATURE REVIEW

The purpose of this chapter is to 1) identify research gaps in existing empirical studies on EMR adoption and 2) determine theories and analytical tools related to the issues studied in this dissertation. To that end, the literature is assessed under four categories: existing models of EMR adoption, network analysis, diffusion of innovation and agent based modeling. The literature search strategy being adopted in this chapter is summarized in Appendix 2-1.

2.1 EMPIRICAL STUDIES ON EMR ADOPTION

Increasing the adoption of EMR allows the healthcare system to leverage the effectiveness and benefits, and improve the quality of health care (Otto and Nevo 2013). However, regardless of efforts being made over the past two decades, U.S. hospitals and physicians still have not fully adopted the EMR and have been progressing the implementation at a rate lower than expected (Otto and Nevo 2013; DesRoches et al. 2013; Gan and Cao 2014; Sherer, Meyerhoefer, and Peng 2016). Accelerating the EMR adoption has thus become a crucial policy issue (Angst et al. 2010; Jha et al. 2006). To identify the policy schemes that can help to facilitate the adoption process, the wealth of literature in EMR and HIT has been conducted and devoted to unveiling the barriers and enablers of the adoption (see, for instance, Holroyd-Leduc et al. 2011; Jha et al. 2006; Kazley and Ozcan 2007).

Many of the studies in the field are based on surveys conducted on health professionals, managers and patients of their perceived barriers and facilitators to their acceptance of the EMR. Two systematic review studies were found that surveyed literature on this subject[4]. McGinn et al. (2011) reviewed articles published between 1999 and 2009 and found that across the literature being examined, nearly all factors were considered barriers by some and facilitators by others but more barriers were identified. The ten most common factors are: 1) design or technical concerns, 2) privacy and security concerns, 3) cost, 4) lack of time and workload, 5) motivation to use EMR, 6) productivity, 7) perceived ease of use, 8) patient and health professional interaction, 9) interoperability, and 10) familiarity and ability with EMR. Amongst the ten factors, financial, time-related and technical barriers were suggested in the literature as the most-cited barriers to EMR adoption. The authors further noted that especially in studies on physicians, health professionals, and managers, the importance of organizational factors, including practice size, change in tasks, IT support, training, management, administration-health professional relationship, the choice of EMR system and inter-organizational relations are highlighted. When technical support and training are in place, these factors are usually considered as facilitators, whereas when there is a lack of IT support or training, these factors tend to be perceived as barriers. Boonstra and Broekhuis (2010) also conducted a systematic review but focused on the barriers to adoption as perceived by physicians. The authors suggested the barriers fall into eight categories, including 1) financial, 2) technical, 3) time, 4) psychological, 5) social, 6) legal, 7) organizational, and 8) challenges in the change process.

[4] The two systematic reviews are conducted by scholars outside the U.S. However, the majority of articles reviewed in each of the study are from U.S. researchers. For example, 28 out of 52 studies in McGinn et al. (2011) took place in the U.S. The rest of the studies were conducted in countries social-economically comparable to the U.S., such as Canada and European countries. Therefore, the findings are generalizable to the U.S. contexts.

According to the authors, the identified barriers are not separated from each other; instead, some have heavy reliance on others. Special attention should be paid to the ***organizational*** barriers and challenges in the change as they determine the relative importance of the other barriers and mediate the barriers during the implementation process.

One research gap in the existing literature is that being an innovation by itself, the EMR has not been well studied from a diffusion of innovation perspective. These studies also view the decision to adopt as a functional of its internal resources and neglecting the fact that adoption is a reflection of other entities within a social system (Angst et al. 2010). It is the mutual influence between adopters and non-adopters that decides how information about the innovation is transmitted, filtrated and learned, and drives the diffusion processes. To this end, Angst et al. (2010) investigated U.S. hospitals' adoption of EMR via a social contagion lens. The authors argued that 1) as the "great connector" of health care system, the EMR offers significant network externalities; 2) the outlets for communicating the innovations are extensive so that influence is inevitable; and that 3) transparency and information sharing is a key component of health care. The authors thus employed a social contagion model which considers a hospital's adoption of the EMR as a function of its susceptibility to the influence of prior adopters, its spatial and social proximity to adopters, and the potency of influence exerted by the adopting hospitals. With the social contagion model, the authors were able to show that hospital characteristics such as size and age and the social proximity between adopting and non-adopting hospitals do play a role in a hospital's decision to adopt the EMR. Similarly, Sherer, Meyerhoefer and Peng (2016) took an approach based upon institutional theory and argued that the development of structures in an organization is strongly shaped by the institutional environment and the effects of institution are dispersed through mimetic, normative, and coercive isomorphism, through copying, learning and pressures. The author highlighted that the healthcare environment in the U.S. is highly institutionalized, the institutional forces will influence health providers in making decision to adopt EMR. Their study of physicians across 2008 to 2012 found that institutional forces do have impacts on the adoption of EMR by health providers.

2.2 NETWORK ANALYSIS

Networks of relationships usually vary by shape and size. As a result, analyzing the properties of the networks can be as complex as they look like. In order to simplify the complexities, some mathematical and statistical methods are used to describe the network properties and capture the network structure. The networks can be further explored in network models to find out how the properties and structure will affect the wider behavior of the system (Newman 2010).

2.2.1 Network Measures and Models

To begin with, a network is a set of relational *nodes* joined by *links*. The *nodes* are also referred to as *vertices, actors,* or *sites* and the links as *edges, ties* and *bonds* by different disciplines. The *network* is also called *graph* in mathematical term. A network can be a *directed* graph, where the link has a direction and point from the first node to the second without the second connecting to the first, or an *undirected* graph, where the nodes are joined by links running two opposite directions. Both directed and undirected graphs exist in our everyday networks. An intuitive example of directed network is the links over the World Wide Web – the hyperlinks can direct from one web page but not necessarily point back. Undirected networks are more prevalent in real world (Jackson 2010), and they can be found in networks such as friendships and partnerships.

In a network, the nodes can connect to one another through direct links between each other or, if they are not immediately connected, indirect ties by running through links of nodes in-between them. Such indirect interaction is captured in *path*, a sequence of links between two distinct nodes across a network. A node can have a *neighborhood*, the set of nodes it has direct links with. A neighborhood can also be obtained for a set of nodes, by grouping the neighbors of each member nodes. The total amount of nodes in one's neighborhood is called the *degree* of a node. By dividing the average degree of all nodes by n-1, n denoting the total number of nodes, the *density* or *connectance* of a network is obtained. Density is usually used to describe the connectivity of nodes in a network. The network is considered *dense* if the *density* tends to a constant when n goes to infinite and *spare* if the density tends to 0. As Newman (2010) has noted, the condition that n approximates to infinite does not work in most practical networks because the nature of the networks cannot easily change. But for some others, the size of the networks do change which allows one to measure at different sizes and decide if they are sparse or density. Thus Newman (2010) suggested networks such as the World Wide Web and friendship are spare networks.

Several network measures and methods are developed that use the above-mentioned concepts and mathematics to capture the structure and characteristics of network and its elements. The measures and methods examine networks at multiple levels. At macro level, the networks can be studied by the global patterns. At micro level, the positions of the vertices are concerned using centrality measures and with their neighborhoods. The networks can also be examined at meso level. At meso level, the focus is on the sub-networks that might form inside the networks. To study these local and segregation patterns, a variety of methods are employed to detect and search for groups in the networks. Table 2-1 provides a summary of the network measures and methods at each level.

Table 2-1 Network Measures and Methods

Level of Analysis	**Measure/Method**	**Description**	**Significance**	**Advantage/Disadvantage**
Macro	Degree distribution	Distribution of the relative frequencies of nodes with different degrees.	Indicate network typology (scale-free).	Not applicable
	Average path length	The mean of all shortest paths between pairs of nodes in the network.	Indicate network typology (small world).	
	Clustering coefficient	The average of each node's network density.	Suggest the existence of sub-graphs.	
Micro	Degree centrality	$\frac{d_i}{n-1}$, where *d* is the degree of node *i* and *n* is the total number of nodes in the network.	Identify central nodes based on degrees.	Advantage: simplistic measure Disadvantage: cannot capture position of nodes in the network.
	Closeness centrality	$\frac{n-1}{\sum D_{ij}}$,where D_{ij} is the length of geodesic path (the shortest path) node *i* with *j* *(j=n-1)* other nodes.	Suggest critical mediating nodes.	Advantage: incorporate the location of nodes Disadvantage: measure sensitive to fluctuations.

	Betweenness centrality	$\frac{\frac{g_{ij}p_k}{g_{ij}}}{(n-1)(n-2)}$,where $g_{ij}p_k$is the number of shortest paths node$p_k$lies and $g_{ij}$is the total number of geodesic paths in the network.	Suggest critical bridging nodes, the removal of which can affect information transmission in the network.	Advantage: capture strategic positions in the network Disadvantage: time consuming to calculate.
Meso	Partitioning clustering	For a given network of *n* nodes, create an initial partitioning, given *k* , and uses an iterative relocation technique to improve the partitioning so that the *k* clusters are created where nodes are same enough if they are in the same cluster and far apart if they belong to different clusters.	Identify clusters with a pre-defined *k*.	Advantage: work well with compact and separated clusters, linear complexity and computational attractiveness Disadvantage: reliance on initial choice of starting points, prior knowledge about the number of clusters, does not distinguish outliers and cannot identify irregular shapes.
	Hierarchical clustering	Agglomerative (bottom-up) or divisive (top-down) clustering using single-link, complete-link or average-link clustering methods.	Identify clusters of nodes based on their similarity.	Advantage: no prior knowledge about the number of clusters, versatility, and allows multiple partitions Disadvantage: no back-tracking capability, arbitrary selection of merge or split points and inability to scale.
	Density-based clustering	Consider clusters as maximal sets of density-connected points. Retrieve all nodes densitiy-reachable from node *p* and form a cluster if *p* is considered a core point.	Identify clusters based on the density-reachability of points	Advantage: no need to identify the number of clusters a priori, notion of arbitrary shapes and noise Disadvantage: arbitrary selection of parameters,

	Divisive clustering	Remove edges with high edge-betweenness – the number of shortest paths between a pair of nodes running along it.	Detect inter-community edges based on their betweenness	Advantage: notion of information spread along the edges Disadvantage: no overlapping clusters
	Modularity-based clustering	Determine the quality of clusters with modularity measure - by comparing the number of actual edges falling within groups with the expected density if the edges within the network are placed randomly and regardless of community structure.	Identify *good* clusters based on the quality functions	Advantage: effective estimate of the goodness of a clustering Disadvantage: modularity optimization is NP-complete

The above-mentioned network measures are tools to understand the properties of networks. To further investigate the effects of these properties on the behavior of elements in the network, mathematical models that could represent the network structure are needed. These models provides us with tractable means to look at the structure and dynamics in the networks and how. In the research of innovation diffusion, the structure of networks has been one of the most intensively researched topics because it allows one to examine the role played by word-of-mouth communication through relational influence among individuals (Brown and Reingen 1987; Kiesling et al. 2012). The three most commonly used models are random, small-world and scale-free networks.

Compared to random network, small-world network typology provides a more accurate representation of real-world networks (Wakolbinger, Stummer, and Gunther 2013).). It is a departure from regular network where nodes are highly ordered and random network where nodes are connected by a random fashion. The small-world topology suggests that real-life networks are as clustered as regular networks, but path length[5] is as small as it is a random network. Watts and Strogatz (1998) was able to show that the small-world phenomenon is common in sparse networks with many vertices. As a result, small-world topologies are included as it is believed that it provides a better illustration than traditional random networks. Some studies that compare small-world networks with random and/or regular topologies suggest it is a more favorable approach for the diffusion of innovation (Sebatiano A. Delre, Jager, and Janssen 2007; Kocsis and Kun 2008), because small-world facilitates the influence at local level in sub-communities. Especially in the early stages of innovation diffusion, small-world topology helps to reach the critical mass, which may not easily be realized in random or regular networks (Choi, Kim, and Lee 2010; Alkemade and Castaldi 2005). But since the diffusion in small-world networks mostly occurs within local clusters, the diffusion process can be slowed if the communication to other regions is not effective (Rahmandad and Sterman 2008).

But the small-world networks do not quite differentiate the role played by different individuals in the network. In contrast, scale-free network, which has a notion of the heterogeneity in the degree of the nodes, provides another alternative for modeling the social structure, especially when it comes to the influence of opinion leaders (Wakolbinger, Stummer, and Gunther 2013). Scale-free network signifies the role played by certain individuals. Scale-free network, introduced by Barabasi and Albert (1999) noted that the

[5]Averaged shortest path length over all pairs of vertices

probability that a vertex in the network connect with others can be expressed in a power-law, decay function. Following a scale-free network topology, there are certain individuals in the social network that have substantially more connections to others than the rest of the individuals. These individuals are seen as hubs, and in the innovation diffusion context, they are the opinion leaders, the information and persuasion from whom can heavily affect other consumer's decision making. As a result, the role of opinion leader is studied invariably in the models with scale-free network topologies (Sebastiano A. Delre et al. 2010; Janssen and Jager 2003; van Eck, Jager, and Leeflang 2011). These models all suggest the positive roles of opinion leader in advancing the communication of information through their social influence. Respectively, these results suggest the use of marketing strategies that target individual consumers who may play a role of opinion leader in the society and employ their influence to promote the innovations. However, the influence of opinion leader is not always positive and to encourage the diffusion, a scenario was studied by Moldovan and Goldenberg (2004) about the role of “resistance leader” who disseminates negative word-of-mouth and initiate the contagion process. Their results suggest that the dissemination of negative word-of-mouth can render the influence of other (positive) opinion leaders obsolete. The authors acknowledged that it is difficult to prevent the emergence of opposition because the oppositions might arise because the innovation represents a deviation from accepted social norm or threatens the opinion leader's expertise by requiring new knowledge and technique. To that end, Moldovan and Goldenberg (2004) suggested supporting opinion leaders' expertise status and increasing market acceptance through positive word-of-mouth as potential strategies.

In terms of the role played by different network models in the diffusion of innovation, literature suggests that the advantages of scale-free and small-world topologies are more salient before a critical mass is reached than after (Delre et al. 2010; Choi, Kim, and Lee 2010; Alkemade and Castaldi 2005). Thus this tipping point is a watershed to compare the contribution of different typologies to innovation diffusion. Some studies have shown that although critical mass is easier to reach with scale-free and small-world networks, once the critical mass is reached random network serves as a better approach for the diffusion (Kiesling et al. 2012). This can be explained by the emergence of network effects once upon reaching the critical mass. Whereas at early stages the diffusion of new technologies mainly relied on word-of-mouth from social networks or external influence such as marketing, once the information cascades have occurred, the network effect becomes significant and the value of adopting an innovation is in direct relation to the size of the network, which results in the swapped effects from social networks.

Understanding the role of network structure also indicates the means to induce behavior change. In the field of innovation diffusion, the characteristics of different network models allow one to identify network intervention schemes to foster the adoption of innovations targeting different groups and/or individuals in the network. In addition to using opinion leaders focusing on the changes brought by particular individuals, the changes can also happen at network levels by strategically change the presence and position of certain links and nodes at different stages of the diffusion. Many of the assumptions about the behavior change might find it difficult to test using traditional statistical tools, especially when empirical data for such analysis is hard to collect. As a result, the issues are usually investigated through computer simulation. Agent-based modeling offers one such tools to illustrate and investigate the dynamics in diffusion networks (Jackson 2010; Valente 2010).

2.2.2 Network Analysis in Health and Medical Studies

The study by Angst et al. (2010), as mentioned above, draws on the role of networks in health organizations. In their work, the authors measured how social proximity – whether or not the hospitals belong to the same health system- influence the adoption processes. This notion of network in the medical domain can find it roots as early as half a century ago when Coleman et al. (1966) probed how networks affected doctors' adoption of tetracycline. The authors found that physicians' adoption of the new drug is affected by the number and types of social networks they have with other physicians. The study provided an insightful conclusion that physician learn and evaluate the innovation through interactions with other physicians. Inspired by Coleman et al. (1966), later studies have broadened the concept and considered networks as the linkages among physicians, clinics, hospitals and other related institutions (Fennell and Warnecke 1988). The networks, thus, are perceived as the vehicles for the spread and use of new medical innovations. As the benefits and effectiveness grow with the size of the network, due to network externalities, the end result is the improved quality of health care and decreased health expenditures (Fennell and Warnecke 1988).

Simply put, network analysis refers to the technique "to analyze the interpersonal communication in a social system" (Valente 1995, 2). It has been applied in health and medical studies in three approaches: 1) transmission networks, which can be the transmission of disease or information, 2) social network, with a focus on the social structure to promote or influence health and health behavior, and 3) organizational networks, the structures and types of relationships existing in public health systems (Luke and Harris 2007). As will be dicussed shortly, the diffusion of innovation in the health domain suggests it falls into the first category – transmission networks. Using this approach, studies often analyze how information dissemination is influenced by different network structure; for instance, entities connected in a network(group) are more likely to share information with another and reach common understanding (Valente 1995) and the centrality of adopters can affect the speed and breath of the information being transmitted to non-adopters(Luke and Harris 2007).

In the health and medical domain, network has been examined mostly by looking at the role of individuals' network ties in facilitating adoption of innovative health and medical practices. For example, Boulay et al. (2002) explored how indirect exposure to family planning campaigns from one's interpersonal communication channels, as compared to direct exposure from radio program, influenced women's adoption of family planning in six Nepalese villages. The authors were able to find that whereas direct exposure was associated with family planning knowledge, indirect exposure was more likely to affect the actual use of contraception. Similarly, Valente and Saba (2001) studied the same issue in the context of Bolivia and found the knowledge and use of contraception was associated with campaign exposure and interpersonal communication. Together, the results of the studies highlighted the importance of the interaction of mass media and interpersonal communication in facilitating health innovations. Individuals' social network have also been utilized in the study of AIDS prevention. Broadhead et al. (1998) conducted a study on this subject to compare the effectiveness of the traditional outreach model that relies on professional outreach workers to that of a peer-driven intervention approach. The study revealed that the peer-driven intervention model outperformed the traditional model in factors of the number of injecting drug users recruited, the ethnic and geographic representativeness of the recruits and the effectiveness of HIV prevention education.

2.3 DIFFUSION OF INNOVATION

As noted by Greenhalgh et al. (2008) in their extensive review of the diffusion of innovation in health organizations, several theoretical approaches have been employed in research on the subject, such as diffusion research, medical sociology, communication studies, marketing and economics, development studies, organizational studies and health promotion. The diffusion theory, which has a notion of contagion, mimicry and learning (Strang and Soule 1998), has been seen as complementing network analysis in the research of innovations. Network analysis supports the diffusion of innovation research by specifying who influences whom during the diffusion process and is enhanced by the diffusion research with a real-world application to compare and clarify network models (Valente 1995).

The subject of innovation diffusion has been studied via a number of different perspectives and can find its roots in anthropology, sociology, geography, political science, economics, marketing, and history (Kiesling et al. 2012; Hall 2004). The wealth of literature into modeling the diffusion of innovation began in the 1960s as pioneered by scholars such as Fourt and Woodlock (1960), Floyd (1962), Rogers (1962) and Bass (1969). In his seminal book *Diffusion of Innovation*, Rogers (1962) defined the term "diffusion" as the process in which an innovation is communicated through certain channels over time among the members of a social system" (p.5). The definition points to the four main elements in the diffusion process as he identified: the *innovation* itself, the *communication channels* through which people create and share information, the *time* it takes for innovation to diffuse and the *social system* where innovation occurs and which affects the diffusion of innovation with its structures, norms, members, decisions and consequences. Observing the S-shaped cumulative curve of the rate of adoption[6], he argued that adopters can be divided into five categories: 1) innovators, 2) early adopters, 3) early majority, 4) late majority, and 5) laggards, under the assumption that the adopters are normally distributed. Also, he noted that the innovativeness of adopters vary by their socioeconomic status, personality values and communication behavior.

Whereas Rogers probed the subject with heterogeneity of adopters and described the theoretical underpinning of diffusion, the Bass model captured the dynamics of population (Meade and Islam 2006) in a mathematical model. Having its origins in epidemiology (Putsis and Srinivasan 2000), Bass (1969) suggested that adopters of innovations can be specified as innovators and imitators. The decision to adopt an innovation made by imitators are influenced by the decisions of other members of the social system. He developed the model based upon the assumption that "the probability that an

[6] See Appendix 2-3 for illustration.

initial purchase will made given that no purchase has yet been made is a linear function of the number of previous buyers"(Bass 1969, p.216). And the individual decision to adopt is the result of joint influence from a desire to innovate and a desire to imitate. The Bass model provides a parsimonious way to look at the whole market and interpret its behavior (Kiesling et al. 2012), and thus allows further efforts to extend and refine its framework. Model modifications have been done in the areas of: 1) the introduction of marketing variables in the parametrization of the models, 2) the generalization of the models to study innovation at different stages and in different countries, 3) capturing the diffusion of successive generations of technology and 4)allowing the effects of competition and networks (Meade and Islam 2006; Chandrasekaran and Tellis 2007) .

Although the Bass model was successful at describing the diffusion of innovation at aggregate level, it was also suggested that the Bass model and the models derived from it cannot fully capture the heterogeneity of adopters, as heightened in Rogers(1962), or their complex and dynamic social structure (Angst et al. 2010). In this respect, new research frameworks have been explored by researchers with enhanced understanding of human behavior and techniques. In their review of diffusion of innovation literature over the past 40 years, Peres, Muller, and Mahajan (2010) observed several shifts in the focus of research interests on the subject. In terms of the driving force of the diffusion of innovation, the focus has shifted from *word-of-mouth* to *interdependencies* among consumers. Previously, the diffusion process was seen as *monotonically increasing*; now *turning points and irregularities* in the curve have been embraced. Moreover, the diffusion is examined from not only its *temporal* but also *spatial* aspects. With regard to scope of analysis, the focus has shifted from *industry-level* down to *brand-level* analysis and extended to cover diffusion of *services* rather than just *products*. As literature has suggested that the network among individuals is not *fully connected*, current efforts have been made to include *small-world and partially connected networks* into the discussion. As a result, the research is increasingly extending its focus from *aggregate* models to models capturing the *heterogeneity of individuals*. Among them, the heterogeneity of individuals and consumer interdependencies, which are considered the two major drivers of new product diffusion (Peres, Muller, and Mahajan 2010), have been most extensively studied.

With regard to heterogeneity of adopters, both Rogers (1962) and Bass(1969) acknowledged that the individuals are heterogeneous in their innovatiaveness and restrained by internal states, external environment and rules of decision making; as a result, heterogeneity exists in their propensity to adopt a new product(Peres, Muller, and Mahajan 2010). The heterogeneity is usually reflected in the time it takes to adopt an innovation and could be affected by factors such as needs and price. For instance, the adoption of a new product can be affected by the interaction of an individual's income and price (Horsky 1990), by their physical proximity (Goldenberg, Libai, and Muller 2010),by their socioeconomic status, experience and learning (Dekimpe, Parker, and Sarvary 2000) and by the tension between the short-term investment and long-term return of benefit (Brynjolfsson and Hitt 2000).

In addition, the interdependency among consumers decides how social influences, whether that be interpersonal communication, network externalities or other types of social signals, drive the diffusion process (Peres, Muller, and Mahajan 2010). And the interdependency is usually reflected in the structure and dynamics of their networks. The underlying rationale for applying network analysis into the study of our social and economic systems is that relationships do matters. An individual's ideas, opinion, attitudes, beliefs and behaviors are a function of those of his or her social networks'(Valente 2010). Networks describe, in a society or organization, who talk to whom and to what extent. As a result, understanding the composition of the networks helps to uncover how innovations interact with one's position and affiliation to diffuse across different linkages or ties in networks.

As noted by Valente (2010) and Jackson (2010) using network methods and models in the study of diffusion has been growing in recent years for several reasons. First, networks change behavior. Attribute theories of behaviors that explain behaviors as a result of attitudes towards behaviors cannot provide knowledge about how to change the attitudes. Network analysis provides good explanations of the influence mechanism about how attitudes and behaviors change. Second, understanding networks can explain behavioral change. Networks have, inevitably, permeated to every aspect of our social and economic lives. Networks have influenced human behaviors through networks of information, disease, transportation, trade and so on. Therefore understanding the structures of our networks is critical and necessary. Third, understanding network structures can inform us about the means to change behavior. Network research offers the lens to not only understand what type of networks emerge in our society but also identify significant individuals, ties and groups to induce behavior change. And finally, the techniques and computational programs to conduct network research have been greatly improved in the last several decades.

2.4 AGENT BASED MODELING

The neoclassical belief that profit maximization strategies will always prevail in market selected has been doubted by the complexity of real-world systems. Agent based modeling (ABM) that embraces the bounded rationality of individual agents has thus appealed scholars in economics, politics and many other social sciences. It does not assume linear equilibrium; instead, it captures the micro-behavior of entities - the agents, in systems. ABM simulates systems with autonomous, endogenously interacting agents. Agents have their internal states and own set of rules governing their decision-making. The modeling allows agents to execute different behaviors so as to interact with others and with the environment. Agents can represent heterogeneous characteristics. The interactions between agents can create subsequent generations of agents that inherit the characteristics of existing ones while reflecting the new environment. Through repetitive interactions, aggregate structure emerges so that not only micro-behavior is captured but also macro-behavior can be studied in the system.

2.4.1 Agent Based Models of the Diffusion of Innovation

The ABM have been employed by scholars to simulate the variance in *innovation*s, *agents*, their *interactions*, and the *environment* (Nan, Zmud, and Yetgin 2013). Recall that Rogers (Rogers 2003) suggested the four fundamental elements of innovation diffusion: innovation, communication channel, time and social system. We can find the four elements being reflected in the attributes examined by ABM. In terms of variance in *innovations*, Nan et al. (2013) reviewed prior literature and indicated that the variance is attributed to the relative advantage of innovations, the presence of network externalities and arduousness between innovations and consumers. In addition, an *agent* in a social system can play a role either as adopter or influencer in the diffusion process. As an adopter, the agent's innate innovativeness decides the likelihood that he/she will be aware of and adopt an innovation as well as interact with other agents in the system. As an influencer, the variance comes from how the agent will exert its power through regulative, cognitive or normative influence (Scott 2008). The interactions among agents take place in the social network that they construct. As a result, the various configurations of the relationships decides how agents' interaction will result in different means in the information flow of innovation-related messages.

The above-mentioned variance can be simulated in ABM through rules and configuration assigned to agents and system. Surveys of existing ABMs on the diffusion of innovation over the past two decades have observed that ABMs on the subject have primarily proceeded with different strategies to model consumer behavior, social influence, marketing practices and government policies (Kiesling et al. 2012; Wakolbinger, Stummer, and Gunther 2013). In particular, a variety of deterministic and stochastic approaches have been employed to model the decision rules by consumers. The rules range from the simplest simple decision rules (Goldenberg et al. 2000), to utilitarian approaches (Delre, Jager, and Janssen 2007), state transition approaches (Goldenberg et al. 2007), opinion dynamics (Martins, Pereira, and Vicente 2009), and to the most sophisticated social psychological and econometric approaches (Schwarz and Ernst 2009). In terms of the social influence, models have been developed to operate at micro (Moldovan and Goldenberg 2004), meso (Janssen and Jager 2001) and macro (Deroian 2002) levels, use different interaction typologies (Erdos and Renyi 1960) and/or incorporate the qualitative modeling of social influence (Kim et al. 2011). Whereas the social influence schemes capture the impacts of word-of-mouth in the diffusion process, marketing practices are also incorporated into ABMs in the form of targeting (Gunther et al. 2011), pricing strategies (van Vliet et al. 2010), timing of marketing activities (Gunther et al. 2011) and product characteristics(Ma and Nakamori 2005). Finally, the role of government policies such as regulations (Schwarz and Ernst 2009), taxes (Schwoon 2006) and subsidies (Cantono and Silverberg 2009) are investigated to see which kind of policy serves as facilitator or hurdle to the diffusion of innovations.

2.4.2 Implications from Agent Based Models

Agent based modeling provides powerful tools to unveil the many inquiries about human behavior in the diffusion of innovation. The assumptions about the factors attributable to the changes in the diffusion process are examined in the form of rules, variables and models. Existing agent based models of the diffusion of innovation have invariably investigated model features, areas of study, modeling and calibrating issues and challenges (Kiesling et al. 2012; Wakolbinger, Stummer, and Gunther 2013). The following paragraphs will summarize findings from the literature regarding the rules and strategies that can be utilized to accelerate the adoption of innovations. In specific, these strategies can be classified under 1) marketing, 2) policy, 3) social networks and dynamics and 4) geographic influence.

The *marketing* activities have been implemented in ABMs through changing the information about new products available to agents or subsets of agents, adjusting the timing of introducing the information to agents, and targeting agents with special characteristics to see at what levels the marketing activities can help to facilitate the diffusion of innovation. Gunther et al. (2011) studied the diffusion of a novel biomass fuel. They found that timing of marketing activities is crucial to the diffusion of the innovation. Intermittent mass communication (introducing the marketing activities in the form of adjusting information level in intervals) leads to an earlier takeoff and a faster increase in diffusion rate than the continuous one (constantly offering information). The authors noted that the difference might be due to the network externalities involved as the information distribution is reinforced through word-of-mouth during inactive intervals in the intermittent scheme. They also found that targeting opinion leaders (agents with a higher influence level) can help to enhance the information flow. Thus it was suggested that targeting opinion leaders such as experts and providing them with sufficient information about the new product can accelerate the diffusion process. Similarly, Bohlmann et al. (2010) studied the structure of the market and communications links between innovator and follower market segments. Their simulation suggested that an early emphasis on innovator adoption rather than cross-section communication between innovators and followers can better facilitate the adoption process. Thus marketing activities such as mass media campaign and organizational communication should take advantage of influential adopter (agents with a greater number of links to others) to accelerate the adoption of an innovation. Schwarz ad Ernst's study of water-saving innovations (2009) signified the importance of information campaign in the process. The authors implemented the information campaign through a rise in the importance of environmental issues of all agents. The simulation results showed that among the four scenarios (baseline, information campaign, subsidy and regulation) on three innovations (showerhead, toilet flush and rain harvesting system) examined in the study, information campaign can accelerate the diffusion of all three technologies and provide the second-highest increase among the scenarios. The study thus indicated that information campaign that increases potential adopter's awareness of the critical issue involved can be employed to facilitate the diffusion. In addition, Laciana and Rovere (2011) adopted an Isling model originated from Physics and highlighted the significance of "seeding" - the distribution of early adopters, to the successful diffusion of innovations. According to the authors, the seeding strategies should be closely related to the newness of innovation and the geographic distribution of potential adopters - when there is a clear advantage over the old product/service, the innovation is adopted at a higher rate when adopters are spatially dispersed. Thus marketing activities should target a concentrated set of early adopters of the innovation when the innovation does not show clear advantage over old ones and target a broader spatial distribution of early adopters if it is with a clear advantage. Finally, Toole et al. (2012) studied the diffusion of Twitter in the presence of mass media, word-of-mouth and geospatial networks of adopters. They found that mass media influence does not take place in the early stages of adoption because the spreading occurs primarily through word-of-mouth. In this period, the adoption is only correlated with demographical covariates. However, in later stages, mass media starts to influence potential adopters and can result in a two to four fold increase in the number of adopters.

The *social network* typologies have been simulated by most of the existing ABMs. Two of the findings from the social network simulations that could shed light on mechanisms to facilitate innovation diffusions are the role played by hubs or well-connected vertices and the network externalities emerged from the diffusion network. For instance, Gunther et al. (2011) and Schwarz and Ernst (2009) as discussed before both identified the crucial role of opinion leaders in the diffusion process. Similarly, van Eck et al. (2011) studied the role of opinion leaders in children's adoption of free Internet games. They found that social networks with active opinion leaders transmit information faster, diffuse innovation more quickly, and result in higher penetration rate of innovation. Opinion leaders can better judge the product quality and exert both normative and informational influence in the social networks thus the more innovative the opinion leader, the higher adoption rate of the innovation. Delre et al. (2007) also supported the positive roles of network hubs, the VIPs, in the market penetration of new products and suggested marketing activities to targeting the VIPs as they have many connection with the consumers. However, the authors noted that the positive effects of VIPs may not be found in some markets. An example given by the authors of such case is the pharmaceutical market. In the pharmaceutical market, physicians are the network hubs but they only have a limited number of patients so that they do not have the information power that VIPs have. As a result, marketing activities that target physicians can only inform a relatively small number of consumers. Thus the authors suggested for markets alike, advertising directly targets consumers can be utilized to stimulate the information spread about the innovation. Considering the role of networks in facilitating the diffusion of innovations, it was recommended that policy makers should invest in the formation of networks in order to improve the information spreading (Deroian 2002).

In addition to the role played by certain individuals in the social network, the network externalities surfaced from the diffusion network have also been studied extensively in ABMs. Many have captured the positive externalities gained during the diffusion process. Tran (2012) modeled the diffusion of energy innovations and found that a combination of adopter's personal preference and network externalities can generate the highest adoption rate. According to him, network externalities can facilitate the adoption of energy innovations even to individuals without a personal preference because they can be influenced directly or indirectly from their friends' purchasing behavior or friends' social networks. Demarco et al. (2009) modeled the adoption of EMR by hospitals and examined the externalities at three levels: global, global vendor specific and neighboring medical facility specific network externalities. Their model suggested that the network externalities arisen from global vendor specifics can push the adoption rate the most. Local adoption as a result of neighboring facility-specific network externalities can also help to increase the effects of global adoption.

By adjusting the pricing strategies to new and old products and imposing behavioral constraints to innovation adopters and producers, ABMs can incorporate *policy* in the forms of subsidies, taxes and regulations. Subsidies can be modeled as a reduction in the price of the innovation; the magnitude and length of the subsidy have been invariably tested in order to identify a better subsidy mechanism. Cantono and Silverberg's model (2009) of environmentally friendly energy technologies found that short-term subsidies can trigger a self-sustained process of diffusion. The effect of the subsidy policy remains strong even after the subsidies have phased out. They also noted that subsidy policies can accelerate the diffusion of new technologies but their success depends on the level and length of subsidies. Faber et al. (2010) also discussed the magnitude of subsidies and highlighted that the influence of subsidies only takes effect within a certain range. In their model, no significant effect was found in subsidies when the level is too low and no additional effects is obtained when too high. Also they argued that the cost-effectiveness of a subsidy policy should be considered. The cost-effectiveness decreases as the subsidy per unit is higher. Echoing Cantono and Silverberg (2009) they suggested phasing out the subsidies when the effects is reached. Ferro et al. (2010) modeled generally the influence of policy incentives that vary in timing, total amount, size (of incentives to each adopter) and time span. Their results suggested that at the beginning of diffusion, the maximum policy effectiveness is obtained by giving incentives of little size to as many as potential adopters. In contrast, at the end of the process, it is worthwhile to have a large budget to ensure that the incentives are provided to almost all remaining potential adopters. With regard to the source of subsidies, Zhang and Nuttall (2011) simulated four scenarios (government-financed competition, government-financed monopoly, electricity supplier-financed competition, and distribution network operator-financed monopoly) of the roll-out of smart metering in UK. Their policy implications are that if the government is to bear the cost of subsidies it is worthwhile to impose an obligation on electricity suppliers the roll out the devices through competition and if the government is unable to provide subsidies, it is more effective for distribution network operators to bear the cost and roll out as monopoly through re-bundling the services.

Since ABM allows modelers to incorporate the *location* of agents, some ABMs have been done to identify the impacts of agents' location on their adoption behavior. Gunther et al. (2011) compared different targeting strategies and suggested targeting smaller regions can result in a faster diffusion than targeting larger regions. Toole et al. (2012) noted that individual's spatial-social network plays a crucial role in the early adoption of Twitter as knowledge about the new technology was spread primarily through word-of-mouth communications.

2.5 SUMMARY

This chapter summarizes previous empirical studies on the diffusion of EMR. The review of literature suggests that being an innovation itself, EMR has not been studied through the lens of the diffusion of innovation. Previous empirical studies mainly viewed the decision to adopt as a function of the decision-maker's internal resources, neglecting the fact that adoption is also affected by other entities within a same social system. Some recent works have embraced this approach, as discussed in this chapter. The chapter also elaborated on the tools and theories in network analysis, diffusion of innovation and agent based modeling that can be utilized to better understand the influence mechanisms behind hospitals' EMR adoption. Network analysis and the diffusion of innovation theories provides the analytical tools and theoretical support to explain how innovations are diffused in social systems. They have been found to complement each other in the study of innovation. Furthermore, agent based modeling provides powerful computational tools which allows us to investigate the variance in innovation, individuals, the environment using computational simulations.

CHAPTER 3 NETWORK ANALYSIS

This chapter examines the first research question regarding the structure of the network among US hospitals. The network being studied in this chapter is the one constructed based on hospitals' organizational connection and spatial proximity. In other words, hospitals are considered connected through either the same hospital system affiliation or spatial cluster. This dissertation follows the terminology by the American Hospital Association (AHA) which defines ***hospital system*** as "…either a multihospital or a diversified single hospital system. A multihospital system is two or more hospitals owned, leased, sponsored, or contract managed by a central organization. " (The American Hospital Association 2016)[7]. Therefore, we use "***hospital system***" to represent the organizational affiliation of hospitals. In this chapter and following sections of this dissertation, we also introduce the network of hospitals constructed based on their hospital system affiliation and spatial proximity; we denote this network as the "***hospital network***". Three specific questions are addressed in this chapter. First, what is the network typology of hospitals' EMR network, organizational and spatial? Second, what are the hierarchical roles of hospitals? Who are the highly connected hospitals? And third, what does the hospital network look like? Using network measures and techniques, this chapter unveils

---

[7] Another confounding terminology is hospital network. According to the AHA, "Network is a group of hospitals, physicians, other providers, insurers and/or community agencies that work together to coordinate and deliver a broad spectrum of services to their community." (The American Hospital Association 2016) This definition is beyond the scope of the network affiliation being examined in this dissertation. Also, this terminology does not equal to the "hospital network" discussed in this dissertation.

the organizational-spatial network of US hospitals, and provides network data for analysis in following chapters.

It should also be noted that throughout the years, hospitals in the US have been undergoing a series of strategic structural changes, such as merger, acquisition and close of business. In addition, a number of new hospitals are being built every year. As a result, the network being studied in this dissertation is not static. Thus, network analysis is performed for each of the nine years (2005 to 2013).

Table 3-1 Number of Hospitals Which Underwent Management or Structural Changes

Year	Merger and Acquisition	Construction*	Deletion**
2006	170	121	47
2007	128	326	56
2008	130	430	73
2009	109	268	42
2010	81	215	32
2011	101	205	31
2012	241	309	78
2013	626	425	128
Total* **	1586	2299	487

Note:
*Construction: Construction of new hospital or hospital facilities.
**Deletion: Close of business, integrated with other health system, or included in the license for other hospital.
***Number only reflects hospitals included in the HIMSS data

3.1 METHODOLOGY

3.1.1 Spatial Clustering

In order to construct the spatial properties of the network, this chapter employs the Distance-based Spatial Clustering of Application with Noise (DBSCAN) method to perform the spatial clustering analysis. DBSCAN is a clustering algorithm developed by Ester et al. (1996), which is based on local connectivity and density functions of points. Compared to other clustering methods, as discussed in Appendix 2-1, DBSCAN has several advantages, as in particular to the purpose of this study: first, there is no need to specify the number of clusters before conducting the clustering; second, it is able to identify clusters of irregular shapes; and third, it has a notion of noise (Figure 3-1) so that points with low connectivity will not be included into any cluster (Ester et al. 1996). This study uses DBSCAN under a few assumptions. First, the structures of the spatial clusters may vary across the years and the number and size of the clusters will be different. Second, the shape of the clusters may not be regular and it is possible to have some clusters located within other clusters. Third, the spatial disparity – some hospitals are located in populated, urban areas whereas some others are located in unpopulated, rural areas – makes it possible that some hospitals are "isolated" from any cluster. Thus, it is not ideal to include such isolated hospitals into any cluster. These assumptions make DBSCAN a proper method for the spatial clustering analysis.

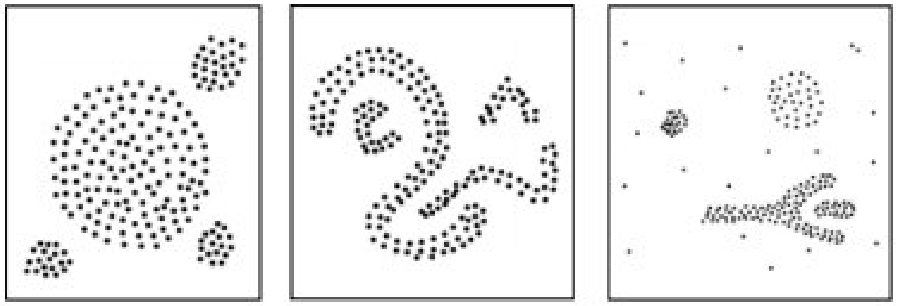


Figure 3-1 Illustration of DBSCAN (Ester et al. 1996)

DBSCAN operationalizes the connectivity and density among points based on their distance with each other. An adjacency matrix of distance is required by the algorithm to perform the analysis. To calculate the adjacency matrix, each hospital's longitude and latitude data was used and the great-circle distance was calculated[8]. The notion of cluster, as in DBSCAN, is regarded as a maximal set of density-connected points. The algorithm requires two parameters to be defined prior to the analysis: the maximum radius of the community, ε, and the minimum number of points in an ε-neighborhood of any point, m. The algorithm operates by first arbitrarily selecting a point p. It then retrieves all points density-reachable[9] from p with regard to ε and m. A cluster is formed if p is a core point. Otherwise, no points are density-reachable and the algorithm visits another point of the database. The process continues until all of the points in the database have been processed. Thus, the selection of the parameters ε and m is very important as it determines the overall

[8] Note that only 2012 and 2013 data contains hospital's longitude and latitude data. For network analysis performed on 2011 and before, their longitude and latitude data was looked up in the 2012 and 2013 by matching common identifier, UniqueID.

[9] A point p is density-reachable from q wrt.ε and $m$if there is a chain of points $p_{1,\dots}, p_n, p_1 = q, p_n = p$such that $p_{i+1}$is directly density-reachable from p_i

number and size of the clusters. Before performing the DBSCAN, a sensitivity test was conducted in order to select the most appropriate parameters. The general purpose is to identify a pair of ε and m with which there are not too many points being isolated, nor are there too many points in a single cluster. After performing the sensitivity test, $\varepsilon = 20$ miles and $m = 3$ was selected. Table 3-2 shows a sample result from the sensitivity test on 2005 data.

Table 3-2 Sample Result of DBSCAN Sensitivity Test (2005 data)

$\boldsymbol{\varepsilon}$	$\boldsymbol{m}$	Number of Clusters	Number of Isolated Points	Number of Points in Largest Cluster
20	3	216	986	430
22	3	193	838	455
25	3	143	612	503
27	3	121	494	1002
30	3	88	398	1711
20	4	155	1287	361
22	4	150	1094	387
25	4	126	818	460
27	4	99	673	534
30	4	80	515	1582
20	5	110	1577	342
22	5	109	1415	356
25	5	102	1105	446
27	5	104	871	459
30	5	75	661	1024

Note:
ε : maximum radius of the community
m: minimum number of points in an ε-neighborhood of any point

3.1.3 Network Construction

It takes two steps to construct the network. First, an edge is created between any two hospitals if they are from the same hospital system. This is processed using the hospital system identifier, ParentID. Second, an edge is plotted between any two hospitals within the same spatial cluster. In order to avoid a fully-connected graph, an additional condition was superimposed that the edge will only be plotted when the distance between the two hospitals is less than 100 miles. The rationale to add this condition is that DBSCAN only takes into account the radius of any point's neighborhood, ε , rather than the radius of the spatial cluster. As a result, for populated regions it is possible to have a spatial cluster spanning a large geographic area with many hospitals. However, hospitals at the periphery of the spatial cluster may be far away from each other; thus the spatial influence is minimal. Finally, combine the edges created in the first steps. If a pair of hospitals is connected both spatially and via the hospital system, only one edge is kept. In other words, the network is constructed among hospitals that are either connected by ***either*** hospital system ***or*** spatial cluster. Figure 3-2 illustrates the network construction process.

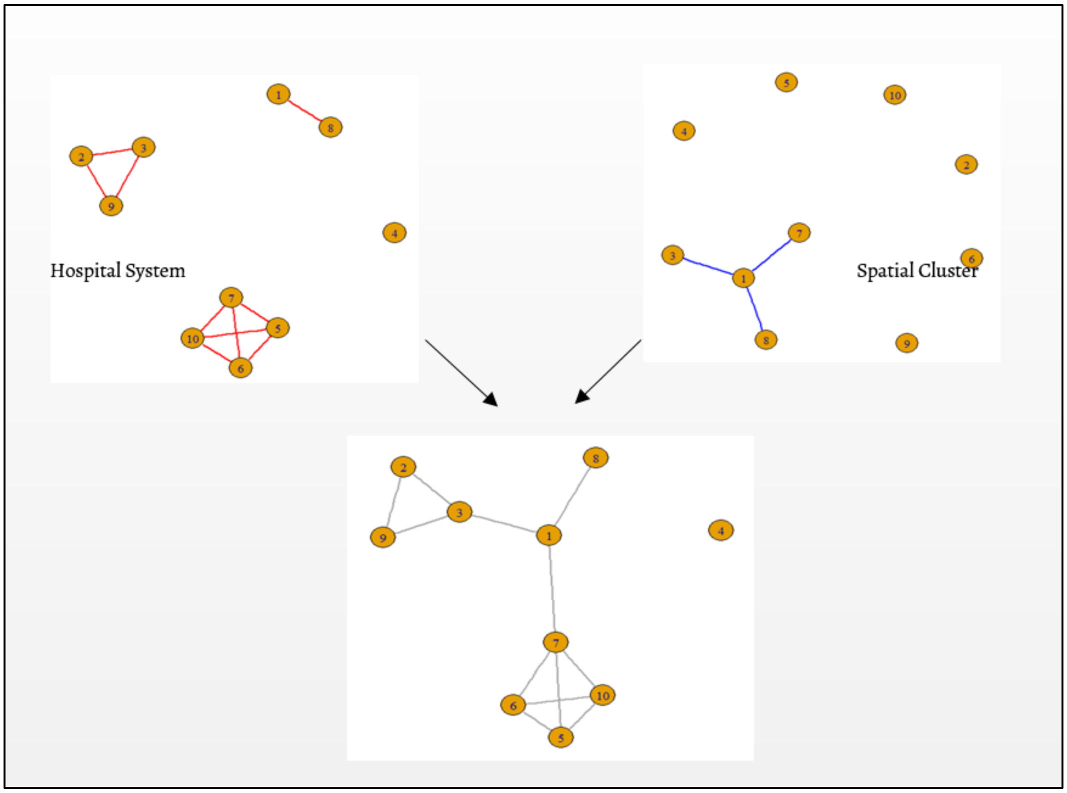


Figure 3-2 Illustration of Network Construction

3.1.3 Network Analysis

After the network is constructed, several analyses are performed to unveil the network structure. To create network visualization, NodeXL(M Smith et al. 2010) is employed to generate node subgraphs and Gephi (Bastian, Heymann, and Jacomy 2009) is selected to present network visualization with geographic aspects. The igraph package for R (Csardi and Nepusz 2006) is used to calculate network measures, including micro-level node degree centrality and macro-level measures for the network. To assess the network structure of real hospital network, simulated random graphs are implemented to compare key statistics. This analysis also uses the igraph package and was performed in R. Note that the study in this chapter, as well as the rest of the dissertation, are performed on hospitals located in the lower fifty states. Thus hospitals in Alaska, Hawaii and Puerto Rico which were originally included in the HIMSS database are excluded from the study and eliminated during data processing.

3.2 RESULTS

Figure 3-4 presents some sample subgraphs of nodes of the 2013 data, with 1.5 levels of adjacent vertices to include in each subgraph (see Figure 3-3 for illustration). Red nodes in the subgraphs indicate thumbnails of the focal node.

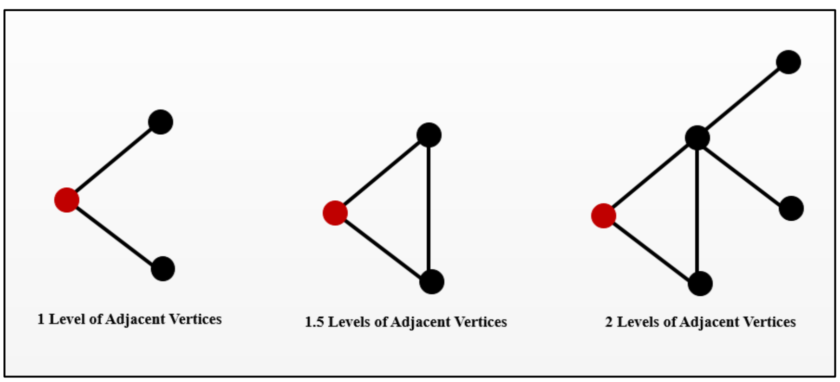


Figure 3-3 Illustration of Levels of Adjacent Vertices

Sample A illustrates a scenario when the hospital is located in a populated area and from a small hospital system (only three in total). Meanwhile, the two other hospitals (nodes in green) is spatially connected to partial of the hospital in the spatial cluster. Sample B denotes a hospital located in a less populated area with only one spatial neighbor (node in green) and from a small hospital system (seven hospital in total). However, the other hospital does not have any spatial connection with any of the hospitals in the cluster, indicating that these two hospitals of a same system may be located apart from each other. Sample C presents a scenario when the hospital is located in a populated area but from a relatively small hospital system. But they do not have any spatial connection to other hospitals in the spatial cluster. Sample D is similar to Sample C; the two other hospitals from the same system is located apart from the hospital in red. Sample E illustrates a hospital located in a populated area (nodes in top right) and also from a relatively large hospital system (nodes in bottom left). Meanwhile, some other hospitals from the same hospital system are located in the same spatial cluster, as there are several overlapping edges between the spatial cluster and the hospital system subgraph. Sample F is a hospital from a moderate-sized hospital system. It only has one spatial neighbor; at the same, this spatial neighbor has an addition spatial connection to another hospital in the hospital system. Figure 3-5 presents subgraphs with 2 levels of adjacent vertices.

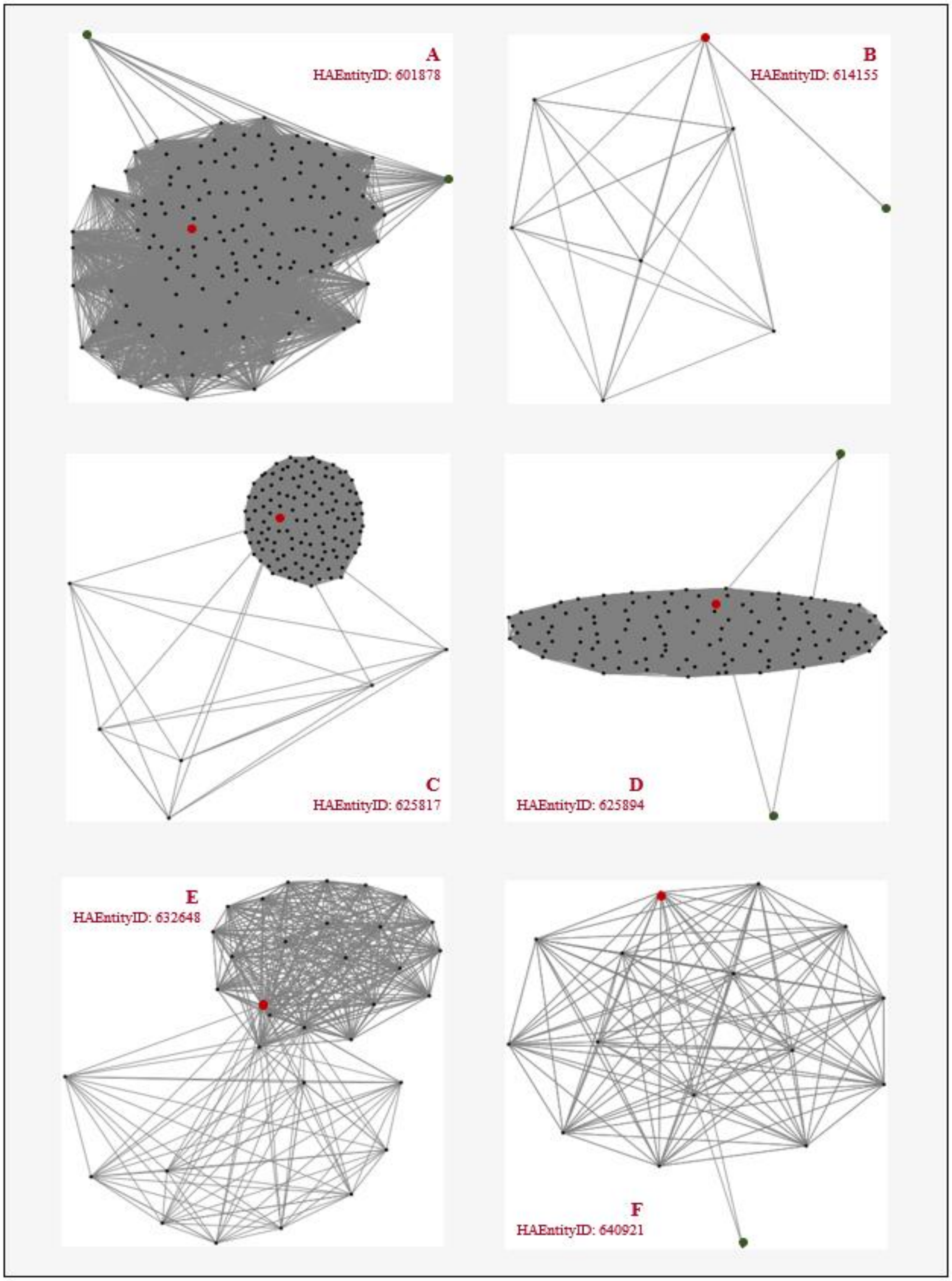


Figure 3-4 Sample Node Subgraphs, 1.5 Levels of Adjacent Vertices

Note: Red points indicate focal nodes.

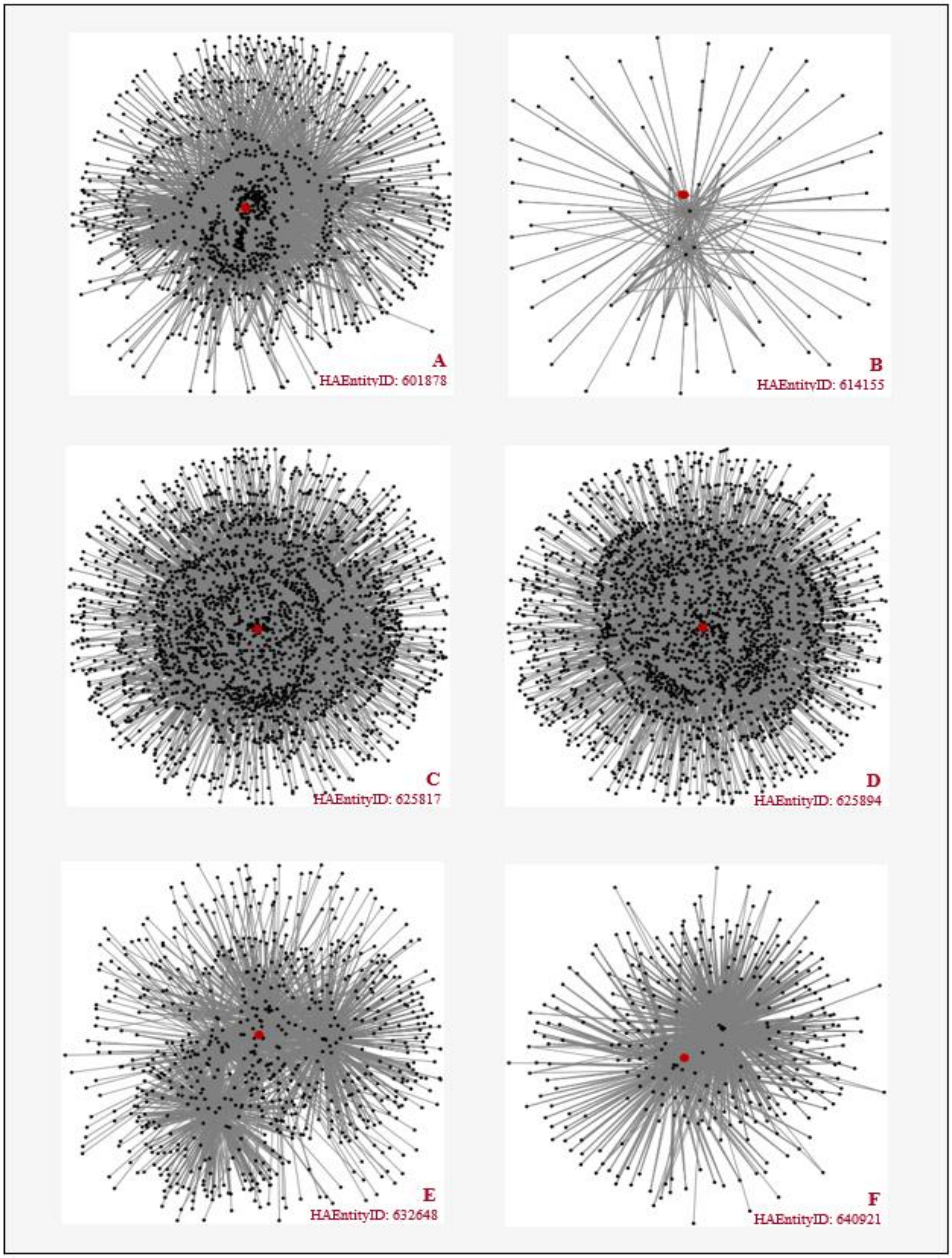


Figure 3-5 Sample Node Subgraphs, 2 Levels of Adjacent Vertices

Note: Red points indicate focal nodes.

Although the total of number of hospitals included in the data remains growing (Table 3-3) and a portion of hospitals underwent structural changes (Table 3-1), the overall appearance of the network does not differ from each other for the time period being studied. For comparison purpose, network visualization of 2005 and 2013 are shown (Figures 3-6 and 3-7) to illustrate the network at the beginning and end of the time period. The network visualization is plotted over the US map to indicate the location of hospitals. Due to limits in computing capacity, only screenshots were taken. For each visualization, two maps with darker and lighter shades are provided.

Table 3-3 Summary of Network Data

Year	Number of Hospitals
2005	3692
2006	4761
2007	4807
2008	4965
2009	5089
2010	5137
2011	5224
2012	5379
2013	5419

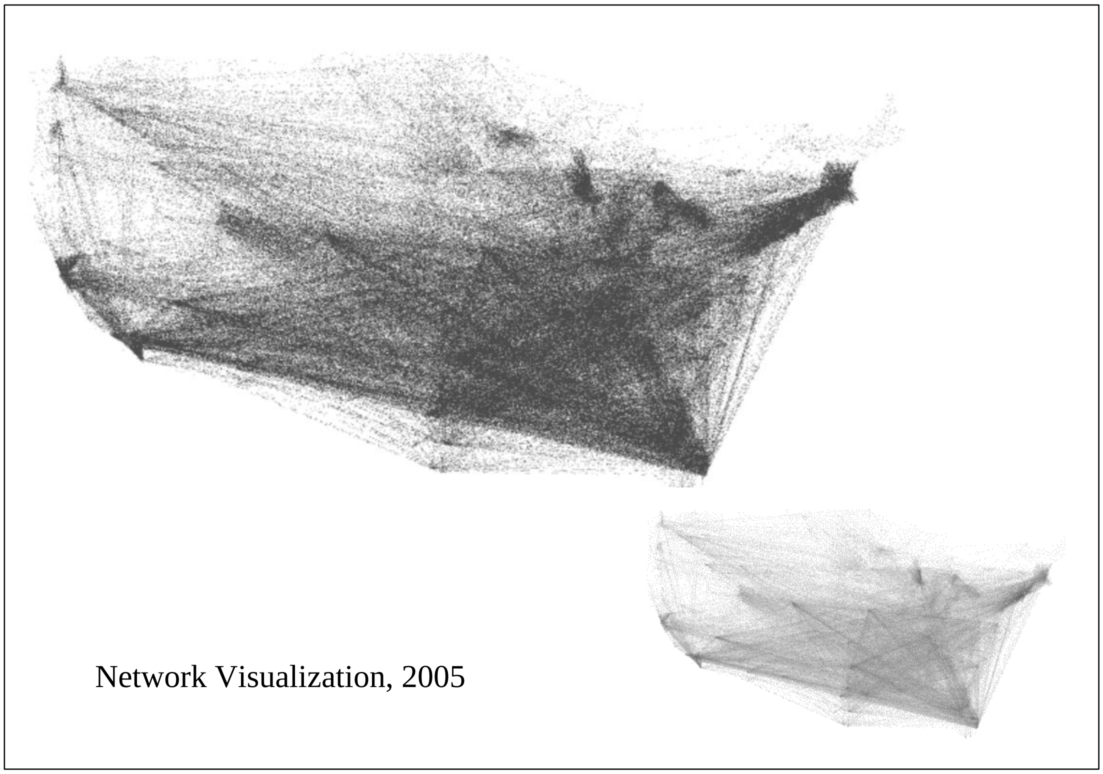


Figure 3-6 Network Visualization, 2005

As each edge represents a hospital's spatial-organization network connection, the shade of the graph indicates where the highly connected hospitals are location. Figures 3-6 and 3-7 show that, generally speaking, these hospitals are located in the New England area, Florida, Texas, and California. Hospitals in the Midwest have more connection than those in the West. We can also identify the existence of several hospitals that "bridge" the network across different regions as represented by edges spanning across the map.

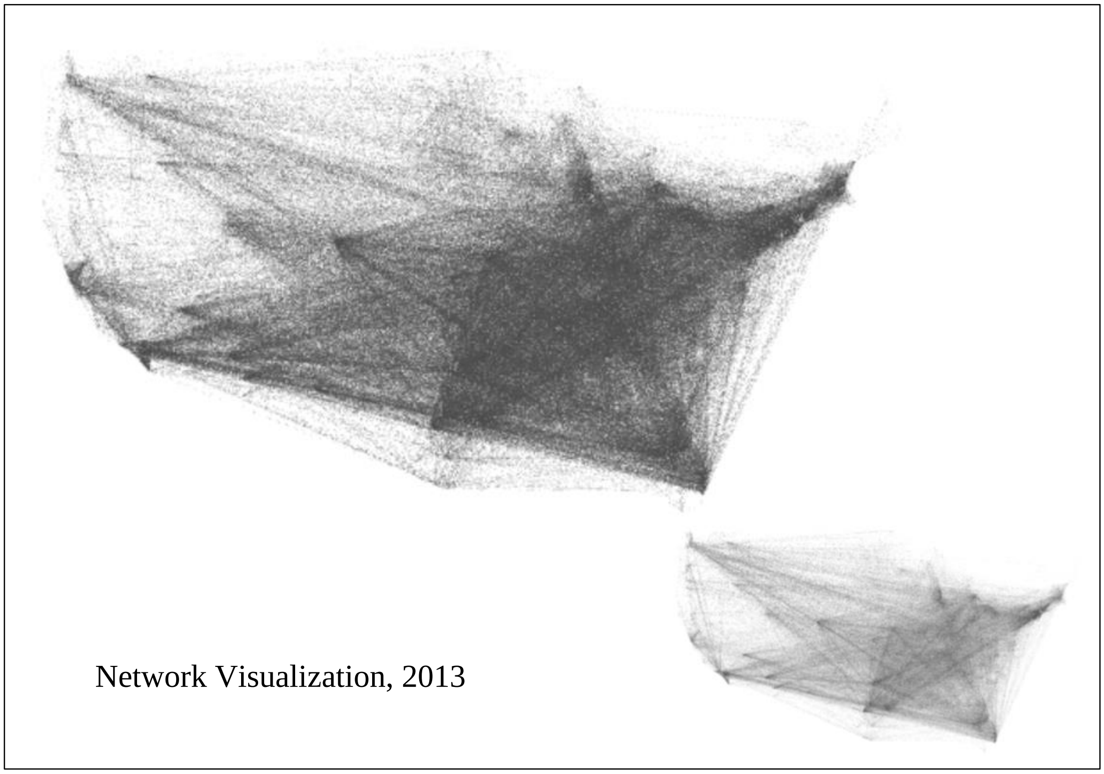


Figure 3-7 Network Visualization, 2013

As a node's connectivity can be calculated by network measure - degree centrality[10], Figures 3-8 and 3-9 report the distribution of top 200 hospitals of highest degree centrality by state in 2005 and 2013 data, respectively. Florida dominates in terms of number of hospitals as there are 40 and 34 hospitals in 2005 and 2013 that are among the top 200 most connected hospitals. Echoing the observation from network visualization, Texas, California and Pennsylvania are also home of many highly connected hospitals. The number of states home to top 200 hospitals in 2013 is lesser than that in 2005, indicating

[10] See Chapter 2 for definition and formula of degree centrality.

that highly connected hospitals are consolidating geographically over time. Meanwhile, the mean degree centrality of the top 200 hospitals is greater in 2013 than in 2005. This may has to do with the fact that more hospitals are included in the database over time.

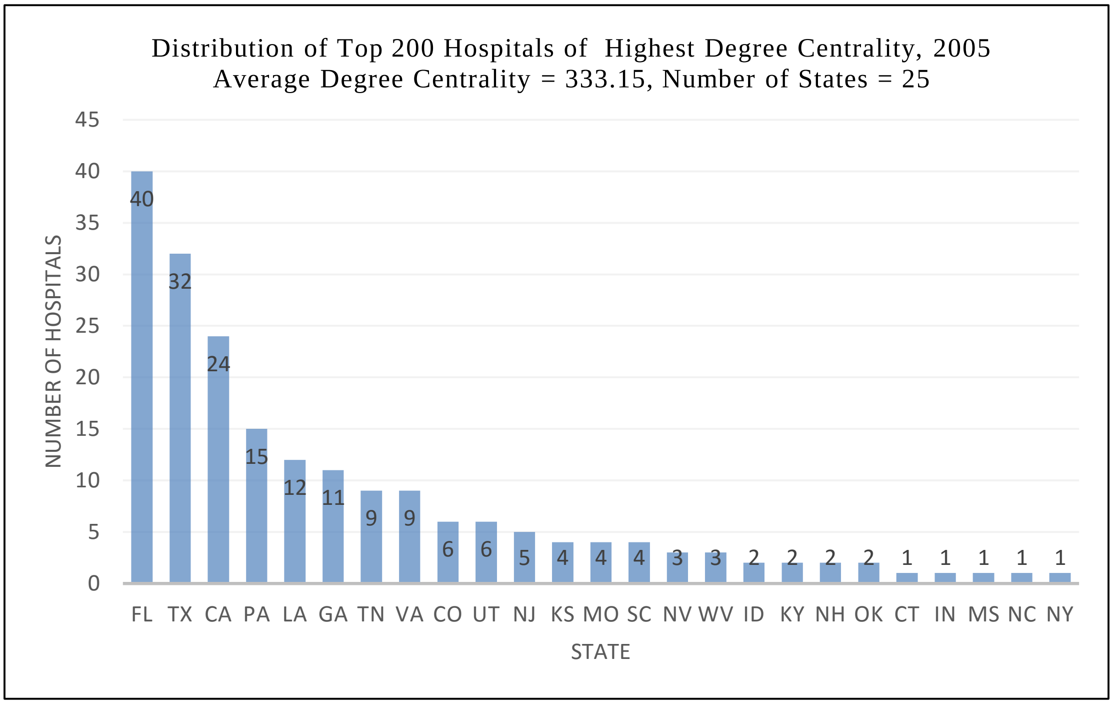


Figure 3-8 Distribution of Top 200 Hospitals by Degree Centrality, 2005

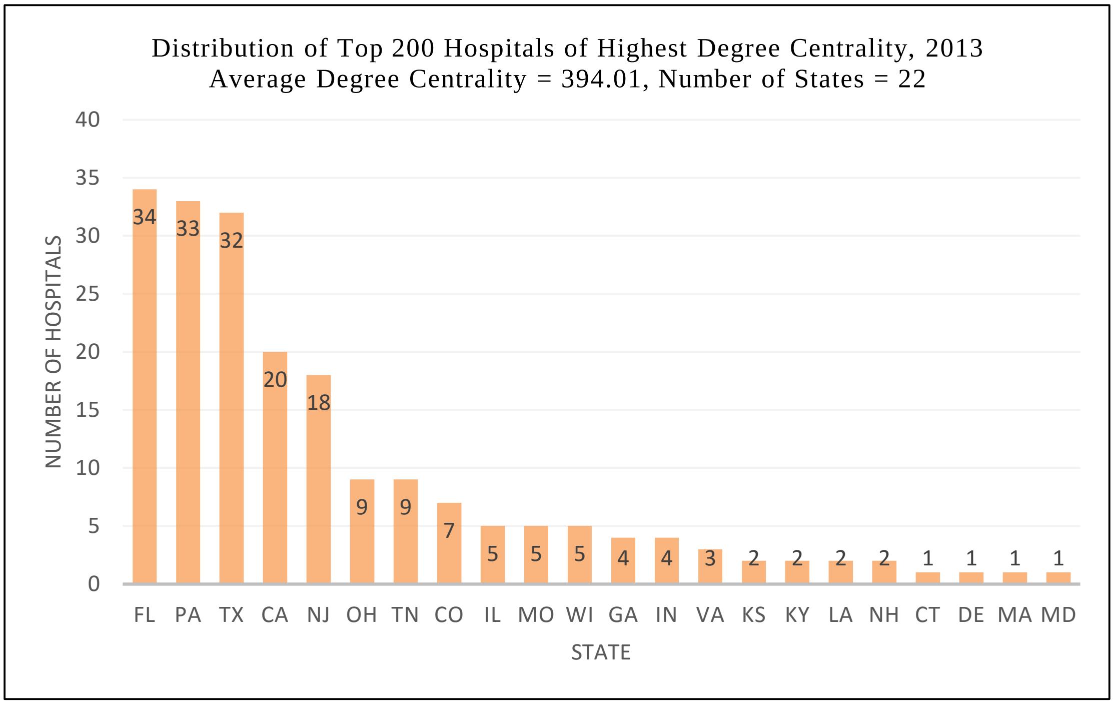


Figure 3-9 Distribution of Top 200 Hospitals by Degree Centrality, 2013

Figure 3-10 reports the distribution of hospital's degree centrality on a linear scale and how it evolves from 2005 to 2013. The two vertical lines indicate the mean and median, as shown in graph. It is observable from the graph that across the nine years, the overall pattern of the distribution is relatively constant – on the one hand, a majority of hospitals fall on the lower end of the distribution and roughly fifty percent of the hospitals (left to the median line) have a degree centrality of 50 or less; on the other hand, some hospitals appear to be very well connected. A small number of hospitals can be identified that have a degree centrality over 400. The skewness of the distribution is attributable to the spatial clustering performed earlier. First, DBSCAN can identify "isolated" nodes so that some may have zero spatial connection. And second, in parameter setting a minimum number of points $m = 3$ was selected thus many nodes requires as few as two other neighbors to construct a spatial cluster.

The distribution of degree centrality of the hospitals indeed shows that it has a long tail – where the majority of hospitals have low degree centrality and the average is brought up by the presence of a few nodes. One may wonder if the typology of the hospital network is related to the well-known scale-free network, as discussed in Chapter 2. Barabasi and Albert (1999) noted that in a scale free network, the degree distribution follows a power law[11]. This degree distribution can be illustrated by transforming the x and y axes in Figure 3-10 into log-log format. Figure 3-11 summarizes the degree distribution of 2005 to 2013 data. It suggests that the distribution is not close to that of a scale-free network. Figure 3-12 shows the degree distribution of 2013 with the corresponding power law fit plotted. It can be seen that the degree distribution does not follow the power law fit.

[11] According to Barabasi and Albert, in scale free networks, "the probability *P(k)* that a vertex in the network interacts with *k* other vertices decays as a power law following $P(k) = k^{-r}$." (Barabási and Albert 1999, 510)

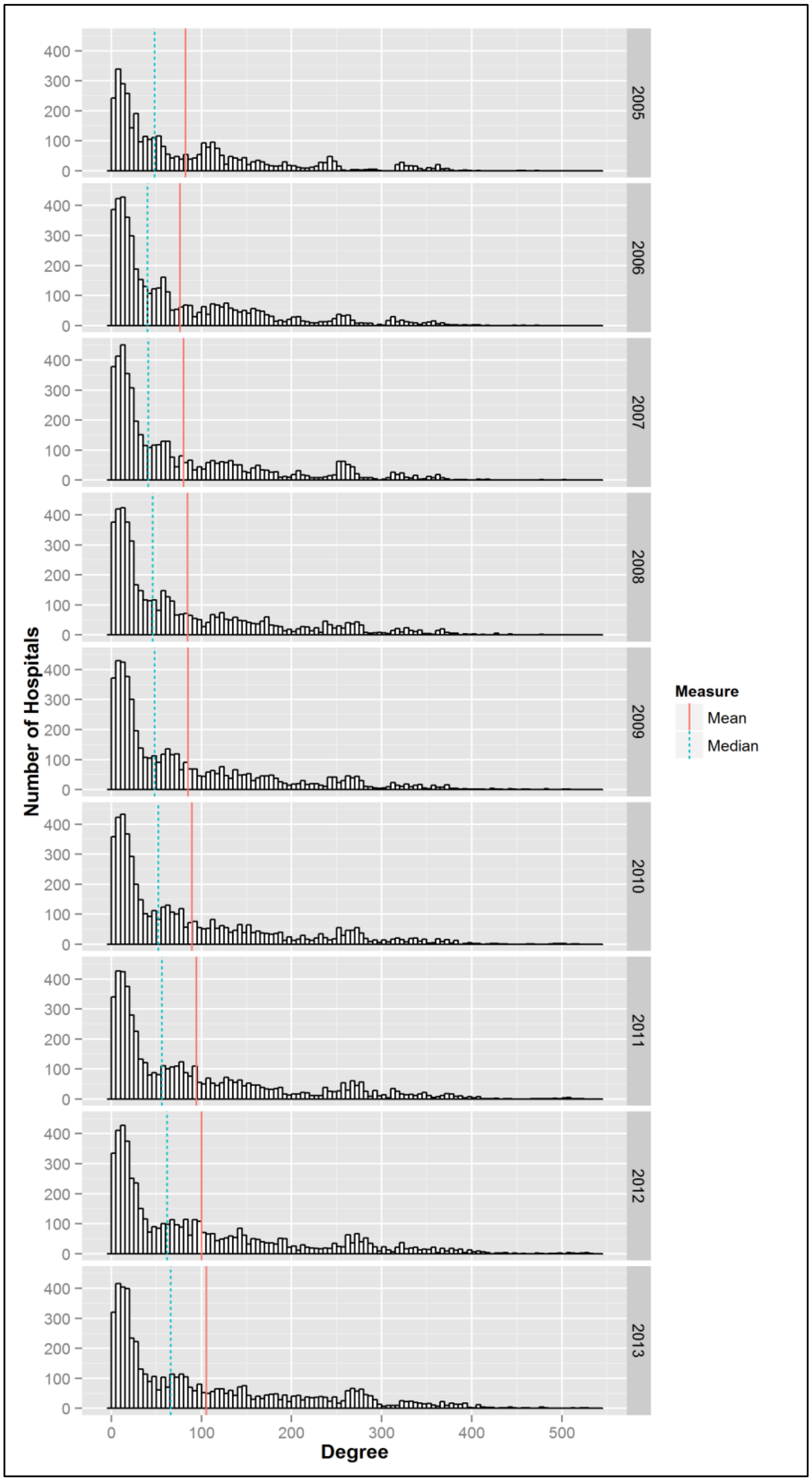


Figure 3-10 Histogram of Degree Centrality

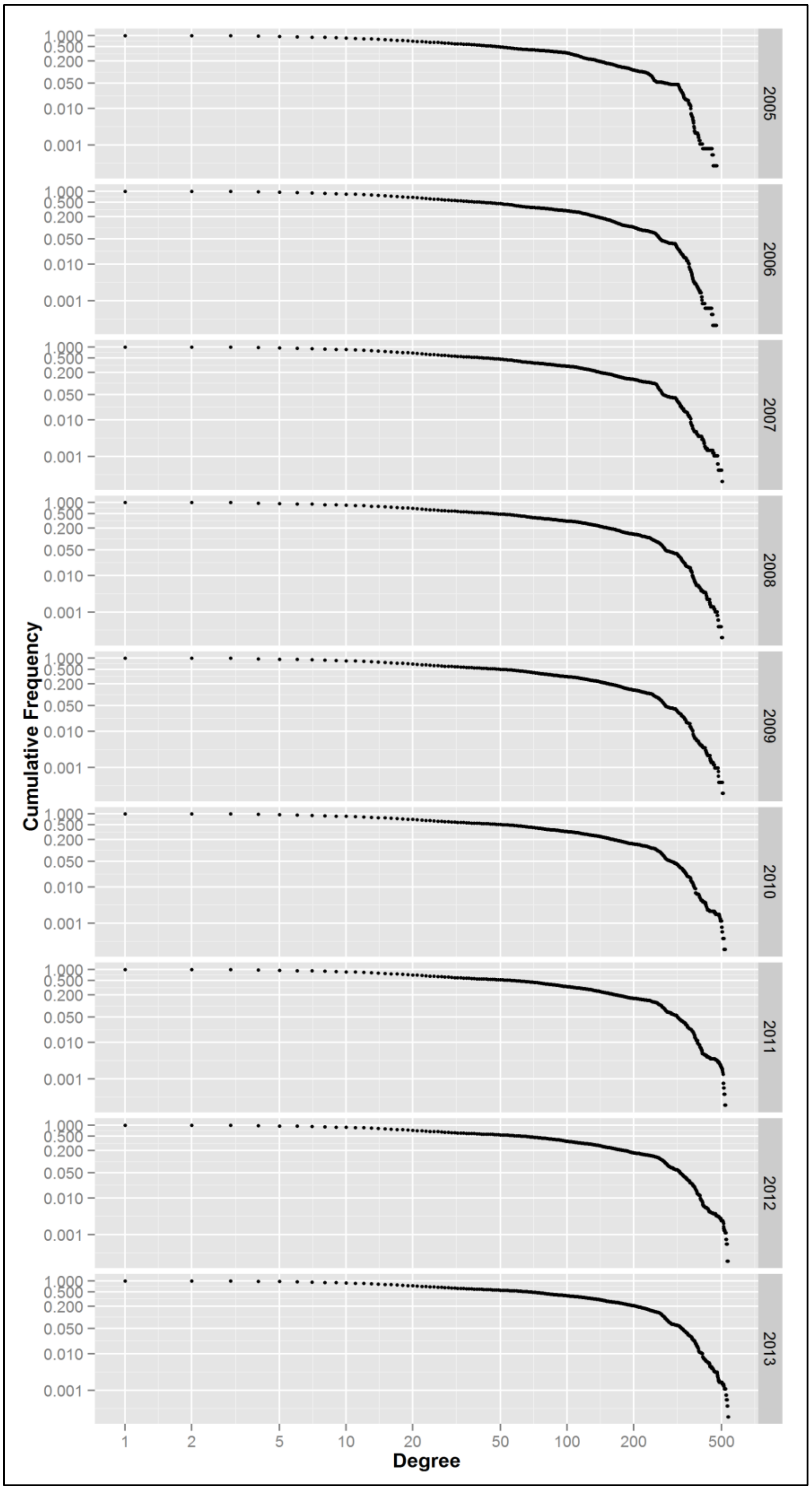


Figure 3-11 Degree Distribution

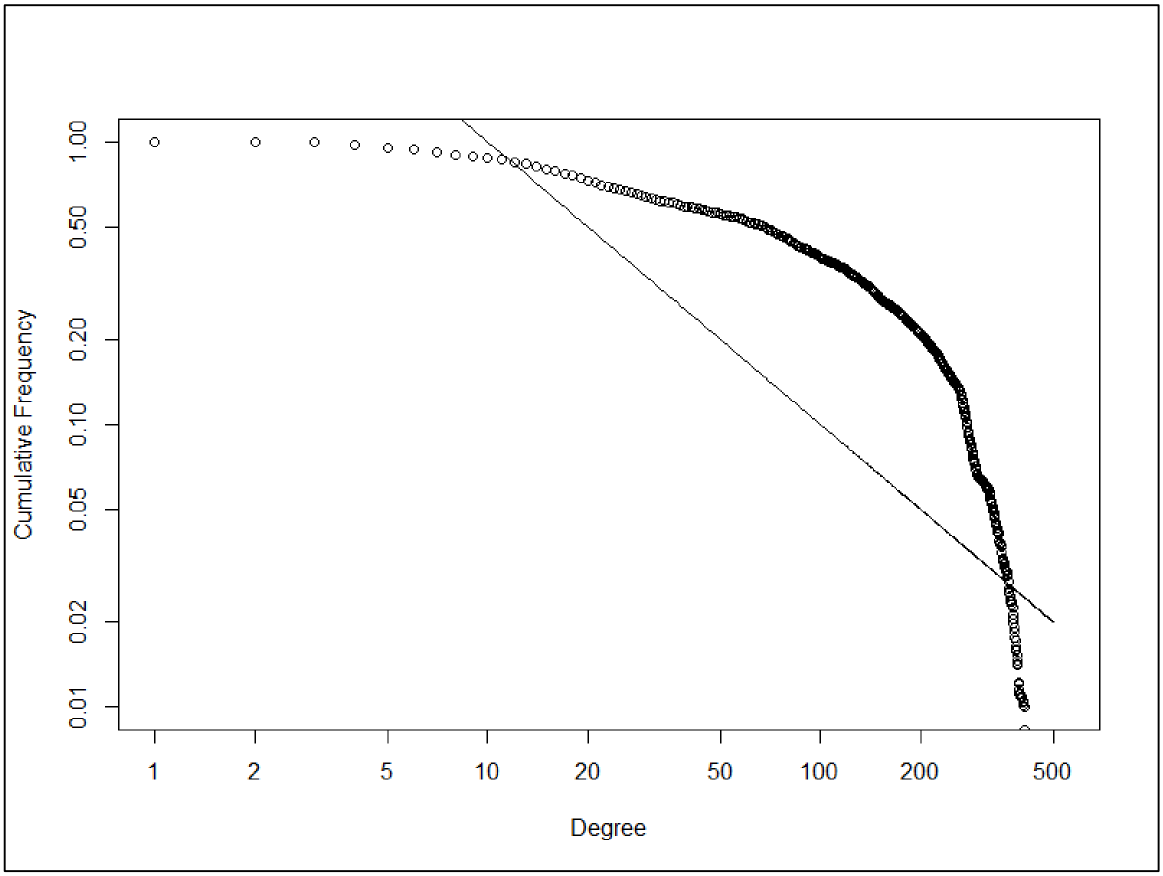


Figure 3-12 Degree Distribution with Power Law Fit, 2013

An equally well-known network typology is the small-world networks. This chapter also assesses the hospital network's small-world properties. Watts and Strogatz (1998) noted that this can be realized by comparing the empirical network's observed clustering coefficient and average path length with those of a random graph with the same number of vertices and average number of edges. The empirical network is deemed having have small-world properties if the observed clustering coefficient is noticeably higher than that of a random network while their average path length remain close. Thus simulated random networks were implemented with the number of vertices and average number of edges of 2005 and 2013 data. Table 3-4 lists the results of summary statistics of simulated random network, after 1000 simulations and those of the real hospital network. The hospital network is more clustered as the clustering coefficient is evidently higher than those of the simulated random network. Meanwhile, the average path length of the hospital network is also higher. These suggest that instead of being a small world network, the hospital network exhibits more of a lattice-like structure than a random network.

Table 3-4 Summary Statistics of Simulated Random Network and Real Network

Network	Measurement	Min	Median	Mean	Max
Simulated Random Network, w/ 2005 Statistics	Clustering Coefficient	0.02212	0.02225	0.02227	0.02229
	Average Path Length	2.134	2.134	2.134	2.134
Simulated Random Network, w/ 2013 Statistics	Clustering Coefficient	0.01942	0.01951	0.01952	0.01965
	Average Path Length	2.105	2.105	2.105	2.105
Empirical Hospital Network, 2005-2013	Clustering Coefficient	0.7123	0.7466	0.7439	0.7843
	Average Path Length	3.528	3.729	3.736	3.929

3.3 SUMMARY

This chapter employed network analysis and network models to 1) present the organizational-spatial network among US hospitals and 2) analyze the network structure of the hospital network and assess the network properties. The overall pattern of the network is relatively constant from 2005 to 2013. Florida, Texas, Pennsylvania and California are among the largest states home to highly connected hospitals. Meanwhile, with the addition of hospitals included into the database and the fact that a number of hospital underwent structural change over the nine years, some small structural changes did take place. Over the time, the average degree centrality of hospitals is rising. The number of hospitals with degree centrality at hundreds high keeps growing.

The observed hospital network is also assessed for its scale-free and small-world properties. The results does not support the existence of scale-free or small-world properties in the hospital network. As discussed in Chapter 2, existing literature on the diffusion of innovation and network models have always discussed the role of different network models in the diffusion process. Studies suggest that innovations diffuse faster in ideal small-world and scale-free networks before a critical mass is reached at early stages of the diffusion (Delre et al. 2010; Choi, Kim, and Lee 2010; Alkemade and Castaldi 2005). Once the tipping point is reached, the effect of social network is swapped and the diffusion performs better in random networks(Kiesling et al. 2012). The analysis from this chapter shows that the observed hospital network exhibits more like a lattice than a classical random network. This dissertation will assess its performance and compare with ideal network models with computational simulations in Chapter 5.

CHAPTER 4 EVENT HISTORY MODEL

This chapter aims to explore the impacts of the presence of network on the diffusion of Electronic Medical Records among US hospitals. In previous chapter, a network analysis was performed that constructed the organizational-spatial network of the hospitals. The network was not built ungrounded – studies in EMR adoption based themselves on the theories of the diffusion of innovation (Angst et al. 2010) and institutional theory (Fareed et al. 2015) have all pointed out the presence of organizational and spatial pressure during hospital's decision-making process.

The network constructed earlier provides two valuable inputs for the analysis in this chapter. First, it calculates each hospital's degree centrality. This captures the hospital's susceptibility and infectiousness during the diffusion process. Second, it records each hospital's neighborhood – group of hospitals having direct ties to it. This data can help to explain how hospitals are exposed to information about EMR adoption. These two inputs serve as the main explanatory variables of network presence.

This chapter studies the occurrence of the adoption of EMR by a hospital. Among the plethora of statistical models available, event history model is considered most helpful. An event history states the longitudinal record of when the event happened to a sample of individuals (Allison 2014). Event history models have been widely applied in the study of innovations(Strang and Tuma 1993). In this study, the Cox proportional hazard model is selected to estimate the impacts of network variables. The model also incorporates several hospital characteristics variables.

4.1 METHODOLOGY

4.1.1 Cox Proportional Hazard Model

Event history models are based on ***hazard*** rate, the conditional probability that the event of interest happens at a given time to a given individual, given that the event has not already happened (Allison 2014). For this chapter, the Cox (1972) proportional hazard model is employed. The model can be written as

$$h_i(t) = h_0(t)\exp(b_1 x_{i1} + b_2 x_{i2} + \cdots + b_k x_i) \tag{4.1}$$

or, by taking the logarithm,

$$log h_i(t) = a(t) + b_1 x_{i1} + b_2 x_{i2} + \cdots + b_k x_i \tag{4.2}$$

where $a(t)$ may be any function of time. The model can also allow for time varying explanatory variables, for example

$$log h_i(t) = a(t) + b_1 x_{i1} + b_2 x_{i2} + b_3 x_{i3}(t) + \cdots + b_k x_i \tag{4.3}$$

where, b_1 and b_2 are coefficients for time-constant explanatory variables and b_3 is the coefficient for time-varying explanatory variable. The Cox model used in this chapter examines both time-constant and time-varying variables, and is stated as

$$log h_i(t) = a(t) + b_1 TYPE_i + b_2 OWNERSHIP_i + b_3 SIZE_i + b_4 DEGREEx_i(t) + b_5 DIREXPOx_i(t) + b_6 SYSTEMEXPOx_i(t) \tag{4.4}$$

where

$h_i(t)$ denotes the hazard at time *t*,

b_1 is the coefficient for categorical variable hospital type,

b_2 is the coefficient for categorical variable hospital ownership status,

b_3 is the coefficient for hospital size,

b_4 is the coefficient for the hospital's degree centrality at time t,

b_5 is the coefficient for the hospital's direct network exposure at time t, and

b_6 is the coefficient for the hospital's network exposure from the entire system.

4.1.2 Model and Variable Specification

In the Cox model (4.4) illustrated above, we consider hospital characteristics variables – type, ownership status and size, as time-constant. However, it is possible that some underwent structural change and the so do the characteristics variables. The function used to implement the Cox model is the *coxph* function included in the *survival* package (Therneau 2015) installed in R. This function handles time-dependent covariates if the data is reshaped into the required format – each time period for an individual is shaped as a separate row (observation) (Fox and Weisberg 2011). As a result, one does not need to specify whether the variables are time-varying as long as the data is correctly formatted. Thus, even though in Model (4.4) the three characteristics variables are stated as time-constant – and they indeed are in most of the hospitals, the model does allow the possibility that structural change took place to some hospitals and their type, ownership status, or size varied over time.

As mentioned in Chapter 1, this study followed the methodology by Furukawa et al. (2010) and Fareed et al. (2015) and coded hospitals' EMR adoption status into three stages based on the implementation of eight core EMR element technologies at the hospitals: *Basic*, *Intermediate*, and *Comprehensive* EMR. In this chapter, the event of an EMR adoption is counted as the fully implementation of ***Basic*** EMR in the hospital - information systems at pharmacy, laboratory and radiology as well as Clinical Data Repository (CDR) are all fully implemented and operational. The Basic status captures the inception of EMR implementation at the hospitals. It is modeled as a dichotomous event – 1 if adopted and 0 if not adopted.

The first network variable is ***degree centrality*** of the hospitals at time t. The degree centrality is often used to examine the infectiousness and susceptibility of central individuals in the diffusion process (Strang and Tuma 1993; Valente 2005). However, to distinguish between infectiousness and susceptibility, the analysis usually are conducted on directed graphs thus use in-degree centrality for infectiousness and out-degree centrality for susceptibility. Since the network in this study is an undirected network, we will use degree centrality to indicate both infectiousness and susceptibility.

The second network variable is the ***network exposure*** from the hospital's direct ties. This is calculated as the proportion of adopting hospitals in each hospital's neighborhood. The neighborhood data was produced in the network analysis in previous chapter. Valente (2005) stated that network can exert influence through individual's exposure to information from his or her personal network. Individuals are exposed to the information about the innovation from their network. As Valente (2005) noted, the exposure can come from direct transmission of contacts or from comparison and competition with structural equivalent network members. In this model, we consider the network exposure as a result of information communication with direct ties.

This model also introduces a system variable, the ***system exposure*** to the innovation. This variable is calculated as the total rate of adoption at each time *t* in the entire population multiplied by the hospital's degree centrality. It is included to capture the pressure to adopt EMR by hospitals from the health care system across the nation. The variable accounts for the hospital's susceptibility to the adoption of all other hospitals, even outside its own network.

Besides the network covariates, this model also include several hospital characteristic statistics. The first concerned is hospital ***type***. Fareed et al. (2015) noted that the constitution of different types (general and special) of physicians can affect the hospital's likelihood of complying with the institutional pressures. The type variable in this chapter is a categorical variable indicating the hospital's specialty. The second characteristic indicator is hospital's ***ownership status***. Literature suggests that hospital's ownership status can affect how the decision makers weigh the investment versus return and pressures from stakeholders (Boonstra and Broekhuis 2010; Fareed et al. 2015). The ownership covariate is a categorical variable where hospitals are owned, managed or leased. Finally, this model incorporates hospital's ***size*** variable. In the studies of the diffusion of innovation in general, as well as EMR adoption in specific, size has invariably been found as a significant predictor of an organization's response to innovation(Damanpour 1991; Angst et al. 2010). Size captures not only the hospital's financial capability to adopt the EMR but also how it deals with the public anticipation and scrutiny(Fareed et al. 2015). The size variable is computed from ***number of beds*** at the hospital. Several indicators are available from the HIMSS data to count for size, such as number of beds, number of staffed beds, and number of full-time equivalent. We selected number of beds for several reasons. First, this indicator yields the less number of missing values as compared to number of staffed beds and the two indicators are highly correlated (Appendix 4-2). Second, although the number of full time equivalent does not have any multicollinearity issues with the other variables in the model (Appendix 4-3), it also creates a lot of missing value by eliminating nearly 25 percentage of the events in the Cox model (Appendix 4-4). Therefore, only the number of beds was selected as the size indicator. Because the distribution of the number of beds is skewed, as will be shown shortly in the descriptive analysis, its logarithm format is taken and included into the model.

4.2 RESULTS

4.2.1 Descriptive Statistics

Figure 4-1 summarizes the adoption of EMR by US hospitals in the HIMSS database from 2005 to 2013. Although the adoption status of Intermediate and Comprehensive EMR are not counted in the statistical analysis in this chapter, they are also included in the graph. Note that the x-axis in the graph is the actual number of hospitals at each EMR stage at a given year, instead of percentage. It is reported in this manner because the number of hospitals included in the HIMSS had been growing for the time period being studied (please refer to Table 3-3) and reporting the percentage can generate a misleading interpretation of the growth. By the year 2013, 4,705 hospitals in the US have had a Basic EMR capability, this has almost doubled the statistic (2,228) in 2005 (Table 4-1). The number of hospitals with Intermediate and Comprehensive EMR capabilities has also been growing, while showing different patterns. The growth of Basic EMR was fastest between 2006 and 2009 but seemingly slowed down after 2009. In addition, the growth of Intermediate EMR has been growing steadily despite a little slow down between 2009 and 2011. Meanwhile, Comprehensive EMR has been experiencing the greatest growth since 2010. The time tick between 2009 and 2010 is an important milestone worth noting. In 2009, the Health Information Technology for Economic and Clinical Health (HITECH) Act, under the American Recovery and Reinvestment Act (ARRA), was enacted. The Act was introduced to provide incentive mechanisms to hospitals and health providers for the meaningful use of EMR technologies. Studies have found that the enactment of the Act has shown as effective means of accelerating the adoption(Sherer, Meyerhoefer, and Peng 2016). As a result, we see in Figure 4-1 that the Intermediate and Comprehensive EMR have shown growth after the enactment of the Act. In contrast, the Basic EMR has been growing slowly since 2009. In 2009, the number of hospitals having Basic EMR capabilities was 3969, which is nearly 80 percent of the total hospitals in the United States. It barely grew seven percent by 2013. The slow growth in the Basic EMR, since it represents the inception of EMR implementation at the hospital, may indicate that there are factors restricting hospital's EMR adoption even when federal incentives are available. The obstacles to these hospitals await further investigation.

Table 4-1 Basic EMR

Year	2005	2006	2007	2008	2009	2010	2011	2012	2013

Basic EMR	2228	2412	2975	3577	3969	4043	4242	4483	4705

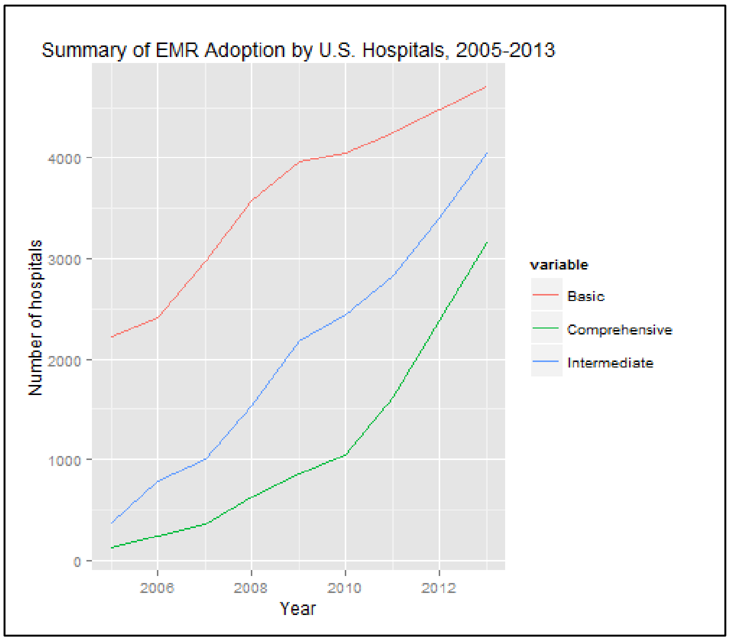


Figure 4-1 Summary of EMR Adoption, 2005-2013

Figure 4-2 summarizes hospital characteristic variable ***type***. General medical and surgical constitute the largest category, followed by critical access, academic, long term acute and pediatric. Fewer hospitals fall into other types. The number of hospitals in each type also vary with time. An evident increase is observed in the number of hospitals in critical access from 2005 and 2006. This is attributable to the data collection by HIMSS when more critical access hospitals were included into the database in 2006. The number of long term acute hospitals keeps growing while the number of academic hospitals remains decreasing. The number of hospitals under “other specialty” has been growing fast recently[12].

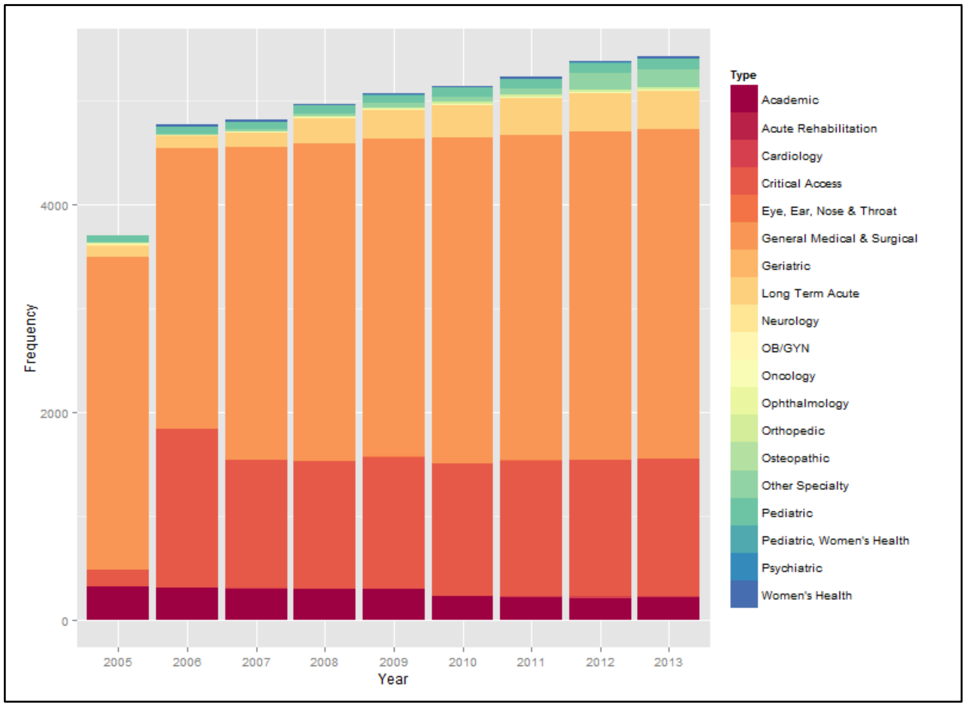


Figure 4-2 Distribution of Hospital by Specialty and Category

The data used in this study does not distinguish between public and private hospitals[13]. For the purpose of this dissertation, the following three ***ownership*** statuses categories are considered: 1) ***owned***, the hospital is owned by a central organization, 2) ***managed***, the management of the hospital’s daily operation is contracted to another organization, and 3) ***leased***, the hospital gives another entity the right to manage the facility

[12] For actual number, please refer to Appendix 4-5.
[13] See Appendix 4-1 for illustration of hospital ownership.

and acquire benefit. A majority of the hospitals in the data are owned, and the number of hospital under this ownership has been growing from 2005 to 2013. Hospitals that are managed or leased are comparatively few in number, but the number has also been growing though not dramatically.

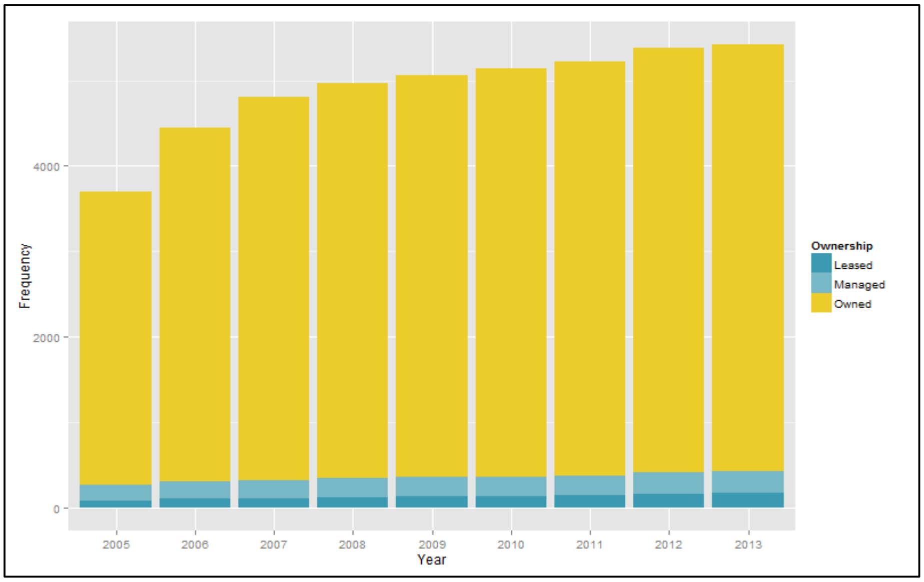


Figure 4-3 Hospital Ownership Status

The distribution of variable ***number of beds*** is left skewed (Figure 4-4). The mean number of beds ranges between 162.9 and 210.7 from 2005 to 2013. But the largest hospital has up to 1,868 beds while the smallest has as low as two beds. The mean also keeps decreasing throughout the years. It may has to do with the addition of new, small hospitals. Additionally, the number of hospitals with bed number between 20 and 30 has been growing dramatically and a dominant group of all hospital across 2005 to 2013 (Appendix

4-6 and 4-7). To deal with the skewness in the distribution, this variable is transformed by taking the logarithm of number of beds.

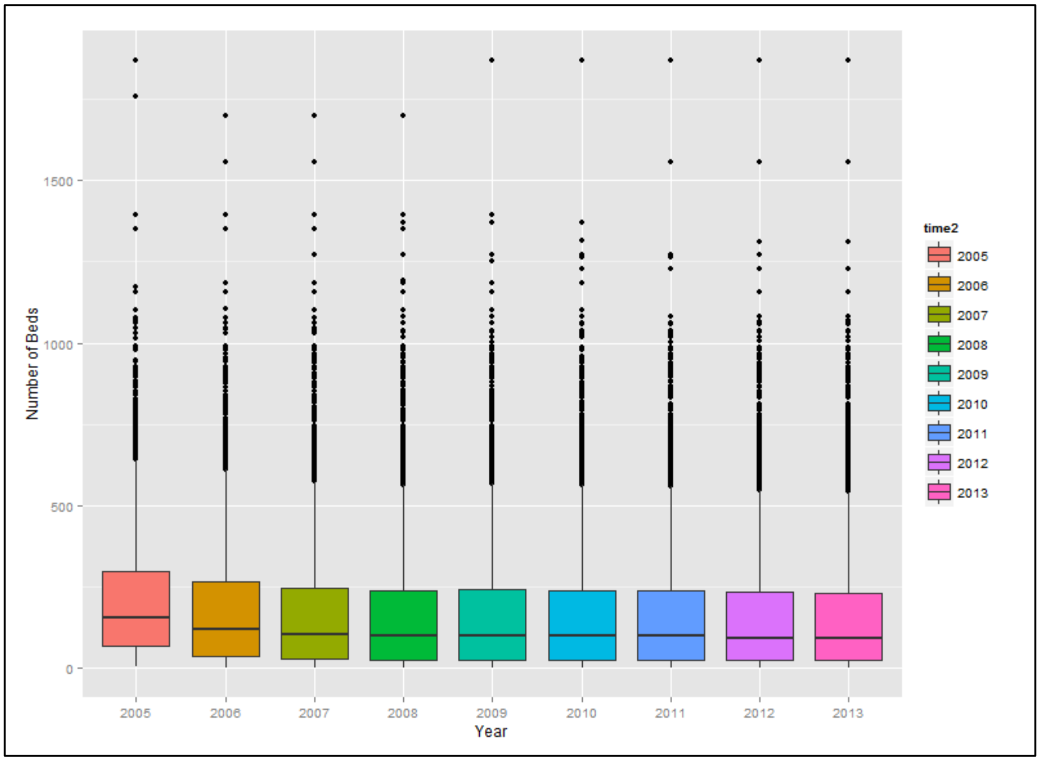


Figure 4-4 Number of Beds, Boxplot

Table 4-2 Summary Statistics of Number of Beds

Year	Mean	Median	Min	Max
2005	210.6725	154.5	6	1868
2006	171.7	103	2	1868
2007	171.2744	103	2	1700
2008	168.4328	100	2	1700
2009	167.5348	100	2	1868
2010	166.4594	99	2	1868
2011	165.3132	98	2	1868
2012	163.1067	93	2	1868
2013	162.9421	92	2	1868

The distribution of ***degree centrality*** has been reported in Figure 3-10, Chapter 3. It can be seen that it is also left skewed. Thus its logarithm is also taken. The distribution of degree centrality has been elaborated in Chapter 3 so it will not be discussed here.

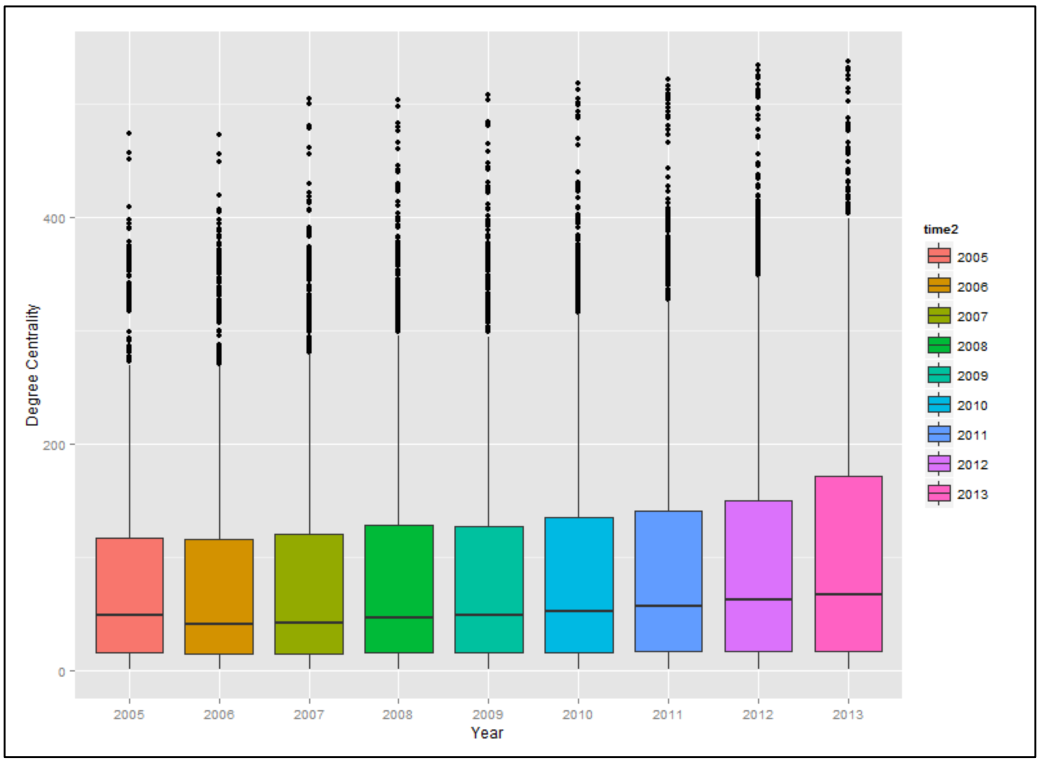


Figure 4-5 Degree Centrality, Boxplot

The variable direct ***network exposure***, as computed by the proportion of adopting network neighbors of each hospital, has been growing despite a slight drop in 2006. This may has to do the addition of new hospitals in the database. The growth in direct network exposure is very intuitive as more hospitals are adopting the EMR. Its distribution shows a slight but not significant left skewness; thus the original data was used in the model, instead of logarithm format. The variable ***system exposure*** is an interaction term calculated by rate of system adoption rate of EMR multiplied by the hospital's degree centrality. The change in adoption rate of the entire system is reflected in Figure 4.1.

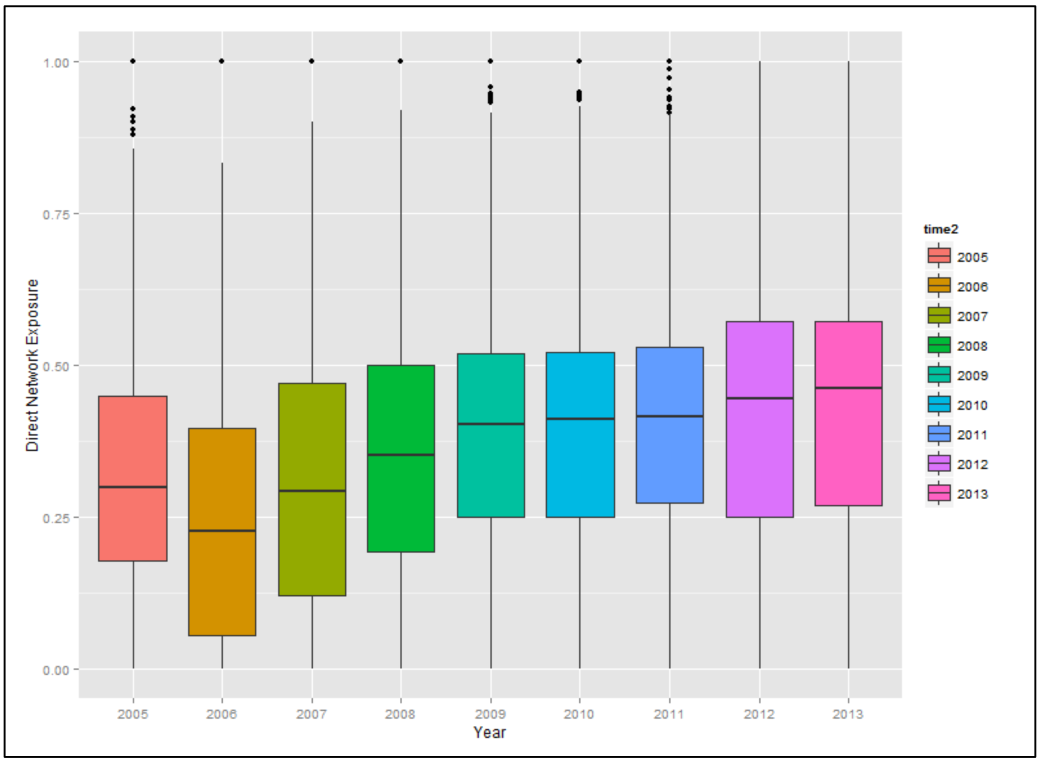


Figure 4-6 Direct Network Exposure, Barplot

4.2.2 Preliminary Regression Analysis

Allison (2010) noted that although evaluating multicollinearity in event history models is not necessary, it may as well be able to pose a potential problem. He suggested to do a preliminary check with a linear regression as event history models do not provide a multicollinearity test. Thus before proceeding to the statistical analysis with the Cox model, a logistic regression was performed. The dependent variable is the probability that a hospital will adopt the EMR – the binary adoption status of EMR. The independent variables are the same as in the Cox model (4.4), plus the system adoption rate. The model is written as

$$ln\frac{p_i}{1-p_i} = a_0 + b_1TYPE_i + b_2OWNERSHIP_i + b_3SIZE_i + b_4DEGREEx_i + b_5DIREXPOx_i + b_6SYSTEMEXPOx_i + b_7SYSTEMADOPT \quad (4.5)$$

The unit of analysis in the logistic regression is hospitals at each year. Besides testing multicollinearity, the purpose of this logistic regression is to provide some preliminary understanding about the covariates to be tested in the Cox model.

The Generalized Variance Inflation Factors (GVIF) were calculated for each of the variables in the logistic regression (Table 4-3). The GVIF is below 3 for all variables and the measure GVIF^(1/2*Df) is below 2[14], which indicates that the multicollinearity issue should not be a concern for this model.

Table 4-3 Multicollinearity Test

Variable	GVIF	Df	GVIF^(1/(2*Df))
Log Number of Beds	2.670610	1	1.634200
Type	2.892065	15	1.029939
Ownership Status	1.050826	2	1.012471
Log Degree Centrality	1.241195	1	1.114089

[14] For GVIF, a common threshold is 5 – variables with GVIF less than 5 do not pose a multicollinearity problem. For measure GVIF^(1/(2*Df)), the threshold is 2.

Direct Network Exposure	1.159765	1	1.076924
System Adoption Rate	1.123654	1	1.060025

The result of the logistic regression is summarized in Table 4-4. Among the three hospital characteristic variables, hospital ***size*** is a significant indicator of hospital's ERM adoption. One percentage increase in the number of beds is associated with 0.362 increase in the log odds of the hospital adopting the EMR while holding other variables constant. Consistent with previous empirical studies (Angst et al. 2010; Fareed et al. 2015), the size of the hospital proves to have a positive impact. The hospital type is a categorical type; in this logistic regression, the reference model is for academic hospitals. Compared to academic hospitals, critical access (a decrease in log odds by 0.370), eye, ear, nose and throat (a decrease in log odds by 0.934) and other specialty (a decrease in log odds by 1.385) hospitals are less likely to adopt the EMR, while holding other variables constant. Significance is not found in the coefficients of other types. With regard to the ***ownership*** status of the hospitals, the reference model is for hospitals that are leased. Compared to leased hospitals, managed hospitals are less likely to adopt the EMR (a decrease in log odds by 0.246), while owned hospitals are more likely to adopt with an increase in log odds by 0.153.

The coefficients for the network covariates is very interesting. To the contrary of theories and other network studies(Valente 2005), ***degree centrality*** of the hospitals is found to be negatively associated with the probability of a hospital adopting the EMR. A percentage increase in degree centrality decreases the log odds of them adopting the EMR by 0.059. However, the two network exposure variables are shown to have significantly positive impacts on the adoption. A unit change in the ***direct network exposure*** – adoption rate by network neighbors, is associated with 2.661 increase in the log odds of adoption. In addition, a unit change in the ***system adoption*** – total adopting rate among all hospitals, can bring up the log odds by 4.101. Interestingly enough, although the degree centrality is found to be negatively associated with the probability to adopt, the interaction term ***system exposure*** we introduced, to measure hospital's susceptibility to the national trend of EMR adoption, has shown to have a positive impact on the adoption. The coefficient suggests that when there are more adopting hospitals throughout the nation, hospitals with higher degree centrality are more likely to adopt the EMR (a unit change in the interaction term is associated with 0.003 increase in the log odds).

Table 4-4 Logistic Regression Results for Hospital EMR Adoption

Logistic Regression Coefficients of Variables			
Dependent Variable: Probability of Adopting the EMR			
Variable	Coef.	Std.Err.	Pr(>|z|)
(Intercept)	-4.242	0.178	0.000
Hospital Size			
Log Number of Beds	**0.362*** **	0.018	0.000
Hospital Type (Reference: Academic)			
Cardiology	0.390	0.335	0.245
Critical Access	**-0.370*** **	0.084	0.000
Eye, Ear, Nose and Throat	**-0.934** **	0.427	0.028
General Medical & Surgical	-0.089	0.068	0.187
Geriatric	-1.333	1.418	0.347
Long Term Acute	-0.003	0.094	0.971
Neurology	11.930	138.3	0.931
OB/GYN	0.378	0.595	0.525

Oncology	0.330	0.299	0.270
Ophthalmology	-0.548	1.479	0.711
Osteopathic	10.71	197.0	0.957
Other Specialty	**-1.385*****	0.120	0.000
Pediatric	-0.166	0.119	0.166
Pediatric, Women's Health	-0.246	0.245	0.133
Psychiatric	-11.02	197.0	0.955
Women's Health	-0.246	0.245	0.316
Hospital Ownership Status (Reference: Leased)			
Managed	**-0.246*****	0.087	0.005
Owned	**0.153****	0.073	0.037
Network Covariates			
Log Degree Centrality	**-0.059*****	0.018	0.001
Direct Network Exposure	**2.661*****	0.066	0.000
System Adoption Rate	**4.101*****	0.115	0.000
System Exposure (Degree Centrality * System Adoption Rate)			
System Exposure	**0.003*****	0.0004	0.000
Other statistics			
Number of observations		42252	
Pseudo R^2		0.163	
*** Significant at .01 level ** Significant at .05 level * Significant at .1 level			

4.2.3 Cox Proportional Hazard Model

The regression analysis using the Cox model (4.4) is performed using two separate models: one with the interaction term - system exposure, and one without. The results are listed in Table 4-5.

Model 1 tested the Cox model ***without*** the interaction term. To make the results comparable to the regression model, we also selected academic and leased as the reference model for the two categorical variables, type and ownership status, respectively. Similar to the logistic regression, the hospital characteristic variables are significant indicators of hospital's EMR adoption. A percentage increase in the number of beds will enhance the hazard of adopting by six percentage while holding other variables constant[15]. Comparing to academic hospitals, cardiology, long term acute and other specialty hospitals are more likely to adopt; the hazards are brought up by 106.2 percent, 51.2 percent, and 32.5 percent, respectively. In contrast, critical access and general medical and surgical hospitals are less likely to adopt, the hazard was decreased by 26 percent and 12 percent respectively. Note that because the number of hospitals in type neurology and acute rehabilitation is small, the model cannot distinguish between event and non-event in the calculation in this two categories. As a result, an NA was computed. As for the hospital's ownership status, compared to a hospital which is leased, an owned hospital is more likely to being an adopter; the hazard is increased by 39.6 percent. The coefficients of the characteristic variable suggest that hospital's size, type and ownership are predictors of the hospital's EMR adoption in some cases. In this model, we only included two network variables, log of the hospital's degree centrality, and the direct network exposure, calculated as the rate of the hospital's network neighbor who are adopters at each time. The two indicators have shown to be have positive, significant impacts on the hazard of adoption. A percentage change in the hospital's degree centrality is associated with 2.9 percent increase in the hazard. Meanwhile, a unit change in the direct network exposure is associated with 228.8 percent increase in the hazard. Note that the direct network exposure is the rate of neighbor, ranging between 0 and 1; the change in the variable will not exceed 1. We can interpret

[15] The coefficients of covariates in the Cox model are usually interpreted using exp(coef.), as multiplicative effects on the hazard. For example, an exp(coef.)=1.06 means that the hazard is brought up by the variable by six percent (6%=0.06=1.06-1).

this coefficient by dividing the variable into appropriate unit. For example, a percentage change (0.01) in the direct network will bring up the hazard by 2.288 percent.

Table 4-5 Cox Model for Hospital EMR Adoption

Cox Proportional Hazard Model Coefficients of Variables						
Dependent Variable: Hazard of Adopting the EMR						
Variable	**Model 1 Without Interaction Term**			**Model 2 With Interaction Term**		
	Coef.	Exp(coef.)	Pr(>|z|)	Coef.	Exp(coef.)	Pr(>|z|)
Hospital Size						
Log Number of Beds	**0.059*** **	1.060	0.002	**0.058*** **	1.060	0.002
Hospital Type (Reference: Academic)						
Cardiology	**0.724*** **	2.062	0.008	**0.702** **	2.018	0.011
Critical Access	**-0.301*** **	0.740	0.001	**-0.296** **	0.743	0.000
Eye, Ear, Nose and Throat	-0.506	0.603	0.477	-0.470	0.625	0.509
General Medical & Surgical	**-0.128** **	0.880	0.038	**-0.126** **	0.881	0.040
Geriatric	-0.786	0.456	0.434	-0.751	0.472	0.454
Long Term Acute	**0.414*** **	1.512	0.000	**0.401*** **	1.494	0.000
Neurology	NA	NA	NA	NA	NA	NA
OB/GYN	0.275	1.317	0.352	0.275	1.316	0.352
Oncology	0.380	1.462	0.219	0.369	1.446	0.232
Ophthalmology	-0.309	0.734	0.757	-0.288	0.749	0.773
Osteopathic	0.861	2.367	0.390	0.899	2.458	0.369
Other Specialty	**0.282** **	1.325	0.036	0.217	1.241	0.114
Pediatric	-0.123	0.885	0.335	-0.104	0.901	0.413
Pediatric, Women's Health	-0.716	0.489	0.313	-0.716	0.489	0313
Psychiatric	-10.05	0.00004	0.974	NA	NA	NA
Women's Health	-0.295	0.745	0.476	-0.299	0.741	0.469
Acute Rehabilitation	NA	NA	NA	NA	NA	NA
Hospital Ownership Status (Reference: Leased)						
Managed	0.055	1.056	0.660	0.062	1.064	0.621
Owned	**0.333*** **	1.396	0.002	**0.323*** **	1.382	0.002

Network Covariates						
Log Degree Centrality	**0.028****	1.029	0.021	-0.024	0.976	0.257
Direct Network Exposure	**1.187*****	3.288	0.000	**1.160*****	3.189	0.000
System Exposure (Degree Centrality * System Adoption Rate)						
System Exposure				**0.001*****	1.001	0.002
Other statistics						
Number of Events	4810			4810		
R^2	0.038			0.038		
Likelihood ratio test	598.6, p=0			607.7, p=0		
Wald test	612.5, p=0			624.1, p=0		
Score (logrank) test	624.7, p=0			637.2, p=0		
*** Significant at .01 level ** Significant at .05 level * Significant at .1 level						

Model 2 tests the Cox model (4.4) using the all-inclusive model ***with*** the system exposure interaction term. Similar to model 1, the statistical significance has been found in the hospital characteristic variables. One percent increase in the number of beds at the hospital is associated with 6 percent increase the hazard. Compared to academic hospitals, cardiology and long term acute are more likely to adopt; the hazards are enhanced by 101.8 and 29.4 percent. Critical access and general medical surgical hospitals are less likely to be adopters and the hazards are decreased by 25.7 and 11.9 percent, respectively. Still, owned hospitals are more likely, than leased hospitals, to adopt the EMR. The hazard is increased by 38.2 percent. Interestingly, when including the interaction term into the model, network variable hospital's degree centrality is no longer a significant indicators. In fact, the coefficient shows a slight decreasing effect, echoing the results in the logistic regression, although not statistically significant. Still, the direct network exposure is shown to have a positive impact on the adoption; a percentage increase in the adopting rate by the hospital's direct neighbors increases the hazard by 21.89 percent. Although in this model, the degree centrality of the hospitals does not suggest any statistical significance, the interaction term measuring the hospital's susceptibility to the national trend of EMR adoption is found to be significant and positive. A unit change in the interaction term is associated with 0.1 percent increase in the hazard. The results from this model reveals that hospitals are susceptible to the exposure from the changes in the entire health system, and this susceptibility is varied by how connected they are. When there are more hospitals across the nation is adopting the EMR, hospitals with more organizational and spatial connections are more susceptible to this change and more likely to be an adopter.

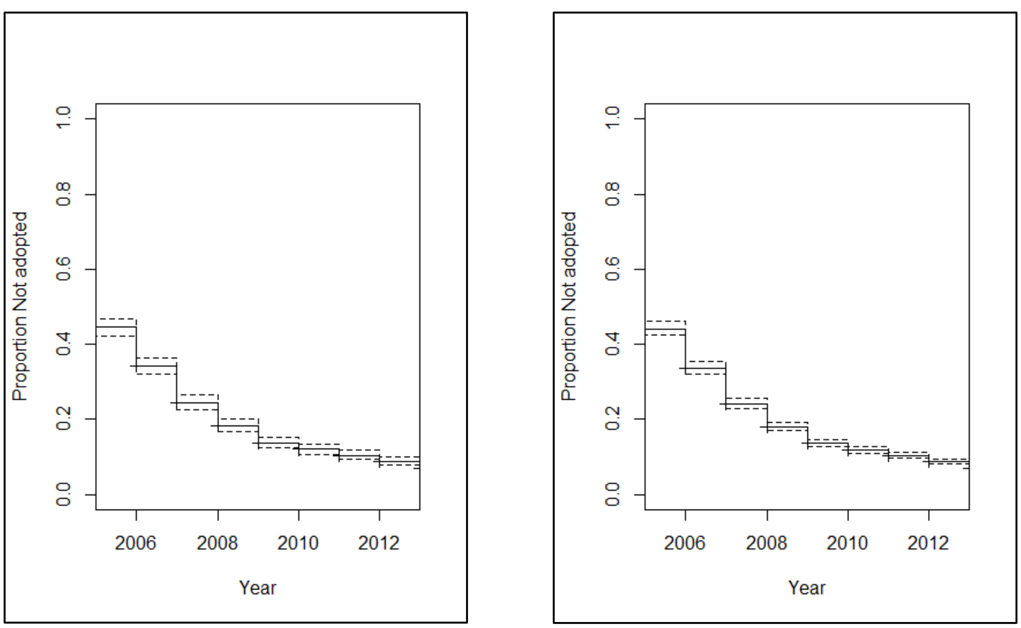


Figure 4-7 Estimated Survival Function
Left: Model 1, Right: Model 2

Figure 4-7 presents the estimated survival function for the Cox model (4.4). The dash lines in the graphs suggest 95-percent confidence interval. The only difference between the two models using the Cox proportional hazard model is the inclusion of the interaction term in Model 2. The inclusion of this variable swapped the statistical significance of variable degree centrality and also alters its direction of influence. However, the R-squared for the two models remains at the same value for the two models.

Model 1 and 2 are performed on reference model with hospital type as academic and ownership status as leased. To ensure that the results is applicable to other types and ownership status, Model 3 and 4 are introduced by reducing the number of levels in the categorical variables hospital type. To do that, the hospital type variable is collapsed into five categories – the four largest categories in the original data: academic, critical access, general medical and surgical and long term acute, and an "other" category which includes all other types. Model 3 and 4 are implemented on reference model with hospital type as all other type and ownership status as owned. As shown in Table 4-6, the result is consistent with findings in Model 1 and 2. When the interaction term is included, the degree centrality variable is no longer significant. Hospital characteristic variables continue to show statistically significant impacts on hospital's EMR adoption.

Table 4-6 Cox Model for Hospital EMR Adoption, Reduced Dummies

Cox Proportional Hazard Model Coefficients of Variables						
Dependent Variable: Hazard of Adopting the EMR						
Variable	**Model 3 Without Interaction Term**			**Model 4 With Interaction Term**		
	Coef.	Exp(coef.)	Pr(>|z|)	Coef.	Exp(coef.)	Pr(>|z|)
Hospital Size						
Log Number of Beds	**0.063*** **	1.065	0.001	**0.063*** **	1.065	0.001
Hospital Type (Reference: All Other Types)						
Academic	-0.154	0.858	0.103	-0.142	0.868	0.133
Critical Access	**-0.413*** **	0.661	0.000	**-0.394*** **	0.675	0.000
General Medical & Surgical	**-0.246*** **	0.782	0.000	**-0.231*** **	0.793	.001
Long Term Acute	**0.297*** **	1.346	0.001	**0.298*** **	1.347	0.000
Hospital Ownership Status (Reference: Owned)						
Managed	**-0.277*** **	0.758	0.000	**-0.258*** **	0.773	0.000
Leased	**-0.328** **	0.720	0.002	**-0.316*** **	0.729	0.002
Network Covariates						
Log Degree Centrality	**0.029** **	1.029	0.019	-0.030	0.970	0.149
Direct Network Exposure	**1.173*** **	3.232	0.000	**1.146*** **	3.146	0.000
System Exposure (Degree Centrality * System Adoption Rate)						
System Exposure				**0.002*** **	1.002	0.001
Other statistics						
Number of Events	4810			4810		
R^2	0.037			0.038		
Likelihood ratio test	582.6, p=0			594.3, p=0		
Wald test	594.0, p=0			608.4, p=0		
Score (logrank) test	604.9, p=0			620.3, p=0		
*** Significant at .01 level ** Significant at .05 level * Significant at .1 level						

4.3 SUMMARY

This chapters explores the adoption of EMR by US hospitals through the lens of their spatial-organizational networks. Using the Cox proportional hazard model, we found that the presence of network does help to facilitate the adoption of EMR. In addition to hospital characteristic factors which have been consistently found as significant indicators, this chapter reveals that hospitals which have a central role (higher degree centrality) and those with more of their neighbors being adopters are more likely to be an EMR adopter. Moreover, central hospitals are more likely to be affected by the national trend of adoption. The network variables in this model captures their infectiousness, susceptibility and communication through the network. So far, it is shown that network indeed play a role in the diffusion of EMR among hospitals. The next question is how to utilize this feature and maximize the network benefit. This question is investigated in next chapter using agent based models.

CHAPTER 5 AGENT BASED MODELING

Chapter 4 demonstrates that during the diffusion of EMR among US hospitals, the existence of network does have an impact on hospital's adoption of the technologies. But it is still unclear as to how networks can be utilized to accelerate the adoption. This chapter serves as an exploratory analysis, using agent based models, to explore policy scenarios which can take advantage of the network properties of hospitals. Previous literature has summarized that legislations on HIT adoption generally fall under four categories: 1) consumer empowerment and data transparency, 2) creation of health networks, 3) financial support and 4) data confidentiality and security (Angst, Desai, and Wulff 2006). These legislations can take advantage of the network in several means. For example, policies on creating health networks aim to foster the interoperability and data exchange among hospitals; in network analysis, this indicates building links among nodes that are not connected before. In medical realm, the notion of network was used to accelerate the diffusion of medical innovation through awarding funding and contracts to certain health facilities to develop network programs (Fennell and Warnecke 1988). Thus the purpose of the ABMs implemented in this chapter is to identify any policy scenarios which strategically deploy policy incentives by maximizing the network properties of the hospitals.

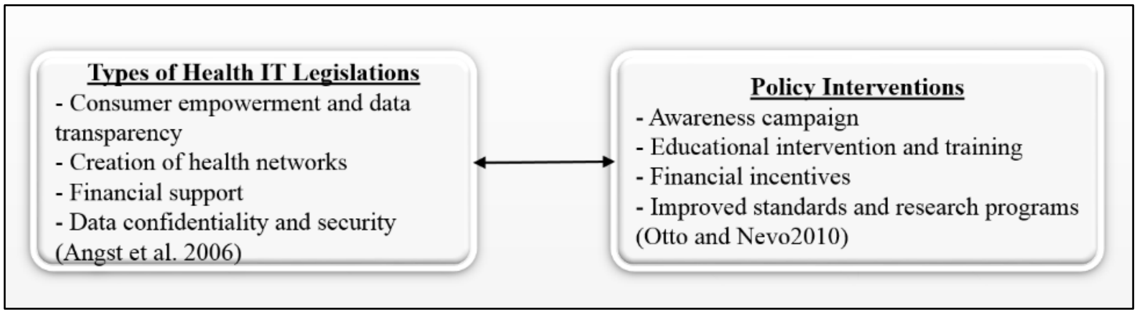


Figure 5-1 HIT Legislations and Policy Interventions

It is suggested that the best way to demonstrate network influences on adoption is to design behavior change interventions (Valente 2005). Valente (2012) defined ***network interventions*** as "… purposeful efforts to use social networks or social network data to generate social influence, accelerate behavior change, improve performance, and/or achieve desirable outcomes …" (49). In this paper, he reviewed studies on network interventions and noted four categories of network intervention: 1) individuals, by selecting nodes based on their network properties (e.g. centrality), 2) alteration, by purposefully changing the network structure, 3) segmentation, by selecting group of nodes, and 4) induction, by activating peer-to-peer interaction to achieve information cascade. This chapter follows Valente's discussion and tested network interventions of the first two categories: individuals and alteration.

It should be noted that the main focus of this chapter is to look at the potential offered by network to develop interventions and polices to affect the adoption of EMR. Thus it does not address 1) the correlation between the closeout of certain hospitals (as reflected in real data) and their EMR adoption, and 2) the ranking of the effectiveness of different policy interventions. Otto and Nevo (2010) summarized that policy interventions on EMR adoption generally fall into four types (Figure 5.1): 1) awareness campaign, 2) educational intervention and training, 3) financial incentives and 4) improved standards and research programs. The policy interventions considered in this chapter are not distinguished by their type and the corresponding effectiveness; rather they are "generic" policy interventions introduced to certain hospitals to have them implement the EMR.

5.1 METHODOLOGY

5.1.1 Model

The agents being studied are hospitals. Each agents can be at one of the two states – adopted or not adopted. For each network intervention, simulations are separately performed on two networks, the real hospital network and an ideal network (scale-free network), in order to assess the performance of network interventions in networks with different structures. As discussed in Chapter 2, the structure of hospital network remains comparatively constant despite the fact that a number of hospital experienced structural changes through the years. Thus, the real hospital network is constructed based on the 2013 network as the selection will not make a big difference on the result.

Adapted from the models by Delre et al. (2007) and van Eck et al. (2011), this chapter develops a model where an agent's decision to adopt is a joint influence of information and social influence. The informational influence is the agent's perceived quality of the innovation, whereas social influence comes from the agent's network neighbors. The decision to adopt depends on the utility from the innovation and its threshold utility, and the agent i will adopt the innovation only when the utility received from the adoption $U_{i,t}$ exceeds a threshold $U_{i,min}$:

$$U_{i,t} \geq U_{i,min} \tag{5.1}$$

The utility$U_{i,t}$depends on two components: the agent's individual preference $y_{i,t}$ and the social influence from *i*'s social network, $x_{i,t}$:

$$U_{i,t} = \beta_i x_{i,t} + (1 - \beta_i) y_{i,t} \tag{5.2}$$

$$q \geq p_i \Rightarrow y_{i,t} = 1 \tag{5.3}$$

$$q < p_i \Rightarrow y_{i,t} = 0 \tag{5.4}$$

$$x_{i,t} = \frac{adoptingneighbor_{i,t}}{totalneighbor_{i,t}} \tag{5.5}$$

where β_i represents the importance of social influence and weights the importance of the two components. q is the quality of the innovation and p_i is the individual preference of agent i. $x_{i,t}$ denotes the proportion of agent i's neighborhood that have adopted the innovation at time t. Thus the agent's decision to adopt as a result of social influence is in proportion to the number of adopting neighbors – if feels more pressured as more neighbors have adopted the innovation.

5.1.2 Parameter Setting

In line with Delre et al. (2007) and van Eck et al. (2011), the fixed parameters in this model include the utility threshold $u_{i,min}$, product quality q, quality threshold p_i, and a probability of adoption resulting from external marketing efforts e. In this model, we considered the adoption as a result of the combined information and normative influence.

However, it is also possible the external marketing efforts can convince the agents to adopt regardless their thresholds (Delre, Jager, and Janssen 2007). Thus the non-adopting agents will adopt the innovation with the probability e. From network analysis of empirical data in Chapter 2, it is noted that a number of hospitals underwent structural changes (roughly 4% of the total hospitals each year, on average). Therefore, several additional parameters are introduced to take into account the endogenous changes among the agents: $p.add$, probability of adding new agents at the end of each time step, $p.del$, probability of an agent being removed at the end of each time step, and rr, rate of rewiring. Table 5-1 summarizes the parameter setting.

Table 5-1 Parameter Setting

Parameter	Type	Variable	Value
N_{real}	Fixed	Initial number of agents in real network	5407
N_{ideal}	Fixed	Initial number of agents in scale-free network	5000
$U_{i,min}$	Fixed	Utility threshold	U(0,1)
q	Fixed	Product quality	0.5
p_i	Fixed	Quality threshold	U(0,1)
β_i	Varied	Importance of social network	$\bar{\beta}$ (0.1,0.9)
e	Fixed	Probability of adoption due to external marketing efforts	0.0001
$i.seed$	Varied	Number of adopters at time 0*	(0,1)
$p.add$	Fixed	Probability of nodes being added at each time step**	0.01
$p.del$	Fixed	Probability of nodes being removed at each time step**	0.01
rr	Varied	Rewiring rate**	(0.01,0.1)
lr	Varied	Rate of nodes to add links	(0.01,0.1)

* i.seed varied at (0.1,1) to test individual interventions. i.seed at 0 in all other models.
** Empirical data suggests on average 4% of the total hospitals underwent structural changes. Thus p.add and p.del are set at 0.01 to reflect the change at each time step (quarter).
***rr varied at (0.01,0.1) to test alteration interventions. rr at 0.01 in all other models.

5.2 RESULTS

Simulations are implemented with the baseline model (no network intervention introduced), individual model and alteration model. For individual and alteration model, variation in relevant parameters are introduced to test different strategies. The simulation are performed on both real and ideal networks. In all ABMs performed in this chapter, results of 50 time steps are reported. We consider each time step as a quarter of a year, thus the simulation considers a time-span of a little over twelve years. Unless otherwise notified, the results reported are mean value taken after 20 realizations.

5.2.1 Baseline Model

Figure 5-2 and Figure 5-3 present the adoption curve in scale-free and real networks, respectively. As β reflects the importance of social influence in agent's decision making process, its value is also varied to account for the interaction between network structure and social influence. We can see that when network intervention is absent, in a scale-free network model, a higher β is always associated with a lower uptake, as vice versa. In other words, the adoption rate is negatively associated with the importance of social influence. In contrast, in the real network model, while at the early time steps curves with lower β tend to dominate the graph, the pattern appears to reverse where the speed of growth is slowed for lower β while expedite for higher β . And by time step 35, the uptake is positively associated with β value. Furthermore, the final uptake in the real hospital network is higher in the real network model than in the scale-free network model.

As the baseline model does not introduce any network intervention, the fact that high β led to a lower adoption rate in the ideal network may suggest that in such a model, the influence from social influence is minimal since β weighs the importance of social influence versus informational influence. However in the real hospital network, social influence does not have impact only until first several time steps and when it is present, it is able to bring up the final uptake to a much higher value.

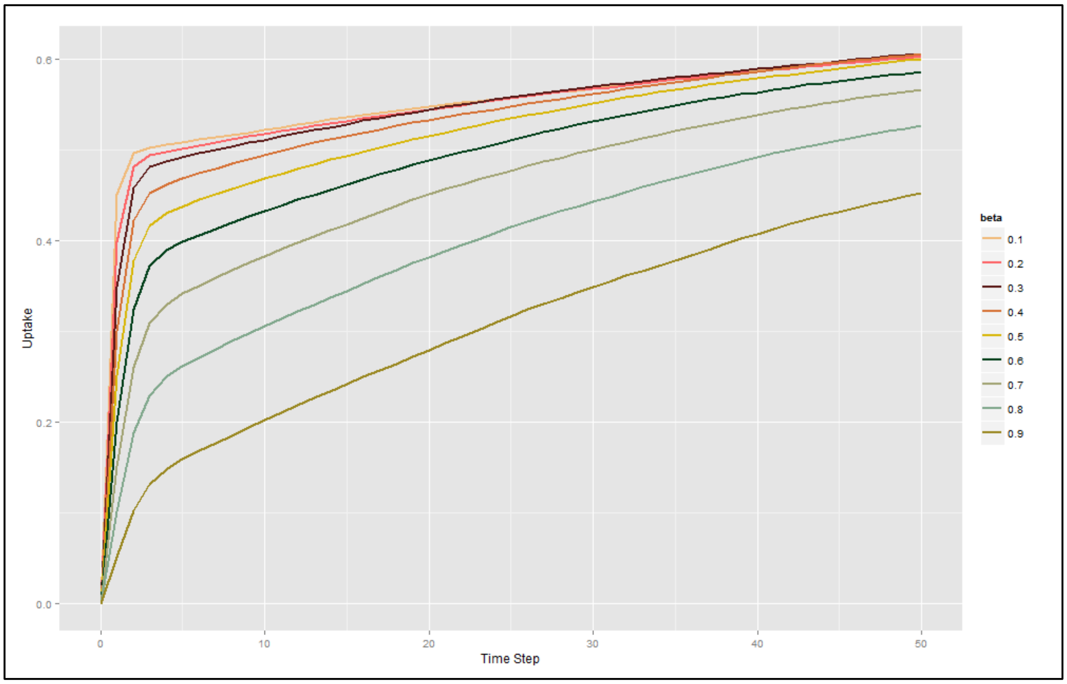


Figure 5-2 Baseline Model, Scale Free Network

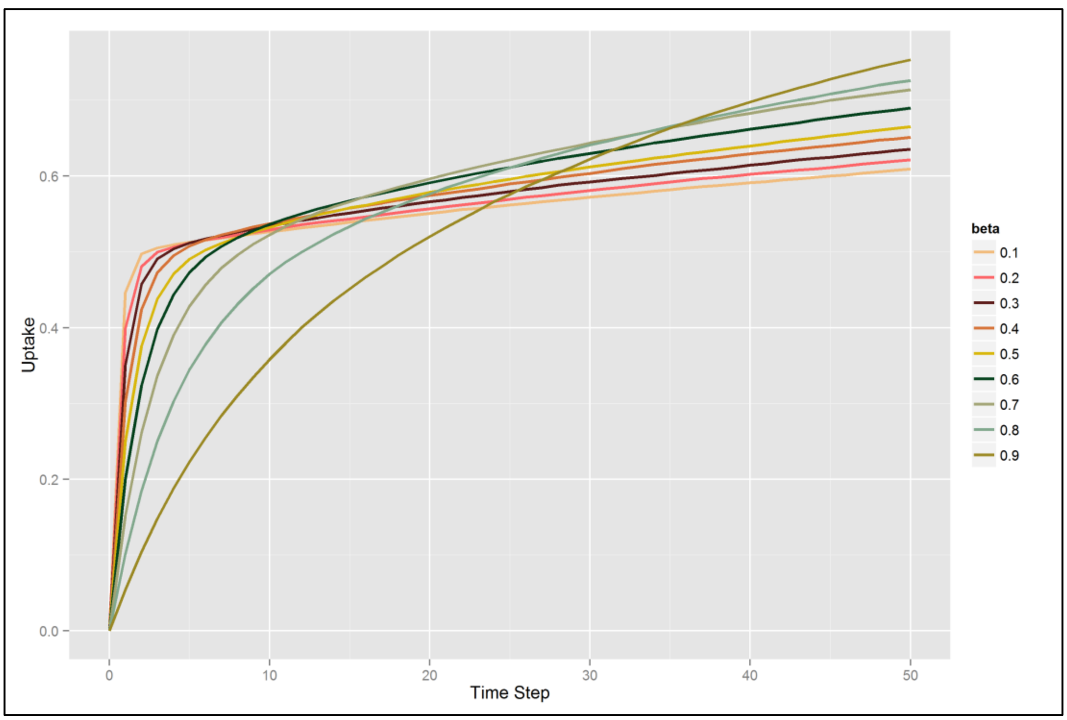


Figure 5-3 Baseline Model, Real Network

5.2.2 Individual Interventions

The individual intervention model, as noted above, considers the selection of agent based on their network properties. Here we consider the selection for seeding – individuals provided with incentives to adopt at the beginning. It is expected that by strategically selecting the individuals for seeding, the influence of network is maximized. Thus it is crucial to identify these important individuals. Valente (2012) mentioned several network measures often used by scholars to determine these individuals. In this model, we tested and compared two measures: degree centrality and closeness centrality. At the beginning of the simulation, agents with highest degree or closeness centrality are selected and seeded as adopters. The number of agents being seeded varies from 1 to 10 percent of the total amount of agents. To compare the results, seeding strategy by randomly selecting the initial adopter is also implemented.

Figure 5-4 and Figure 5-5 presents the results of the individual intervention model on the real hospital network and the scale-free network, respectively. In addition to varying the seeding strategy – degree centrality, closeness centrality and random selection (rows), the simulations were also implemented by varying the value of β (columns). As seen in the baseline model, β, the social importance factor, does introduce difference to the adoption. Thus we may expect an interaction between β and the seeding strategies, and in different network structures.

The two graphs show that when β is low (0.2 and 0.4), the pattern of the adoption curve does not differ much by the selection strategy. Also, the variance in the uptake between highest and lowest seeding rate is also small. This result is very intuitive as when social influence is less important, the decision to adopt is mostly dependent on the information influence the agent has received. Thus, varying the strategies based on the network and the agents' network roles will not introduce much change to their decision making. Also note that the patterns are very close in both real and ideal network.

In contrast, when β is high (0.6 and 0.8), we could identify several noticeable differences. First, in the real hospital network model, a higher β is associated with higher uptake (as compared to lower β), across all three strategies. In addition, the difference in the seeding rate generates greater variance with higher β. However, this is not the case in the scale-free network model. The difference in seeding rate does introduce a greater variance, but the higher β does not bring up the uptake too much and the different is not very noticeable.

Second, in the scale-free network model, when the seeding rate is lower, the final uptake is much lower with higher β. This may have to do with the fact than when the initial adopters are fewer but agent put more reliance on the social influence, their adoption is greatly slowed as there are not a sufficient amount of adopters in the network. However, the same pattern cannot be found in the real network model. The uptake grows simultaneously with the β.

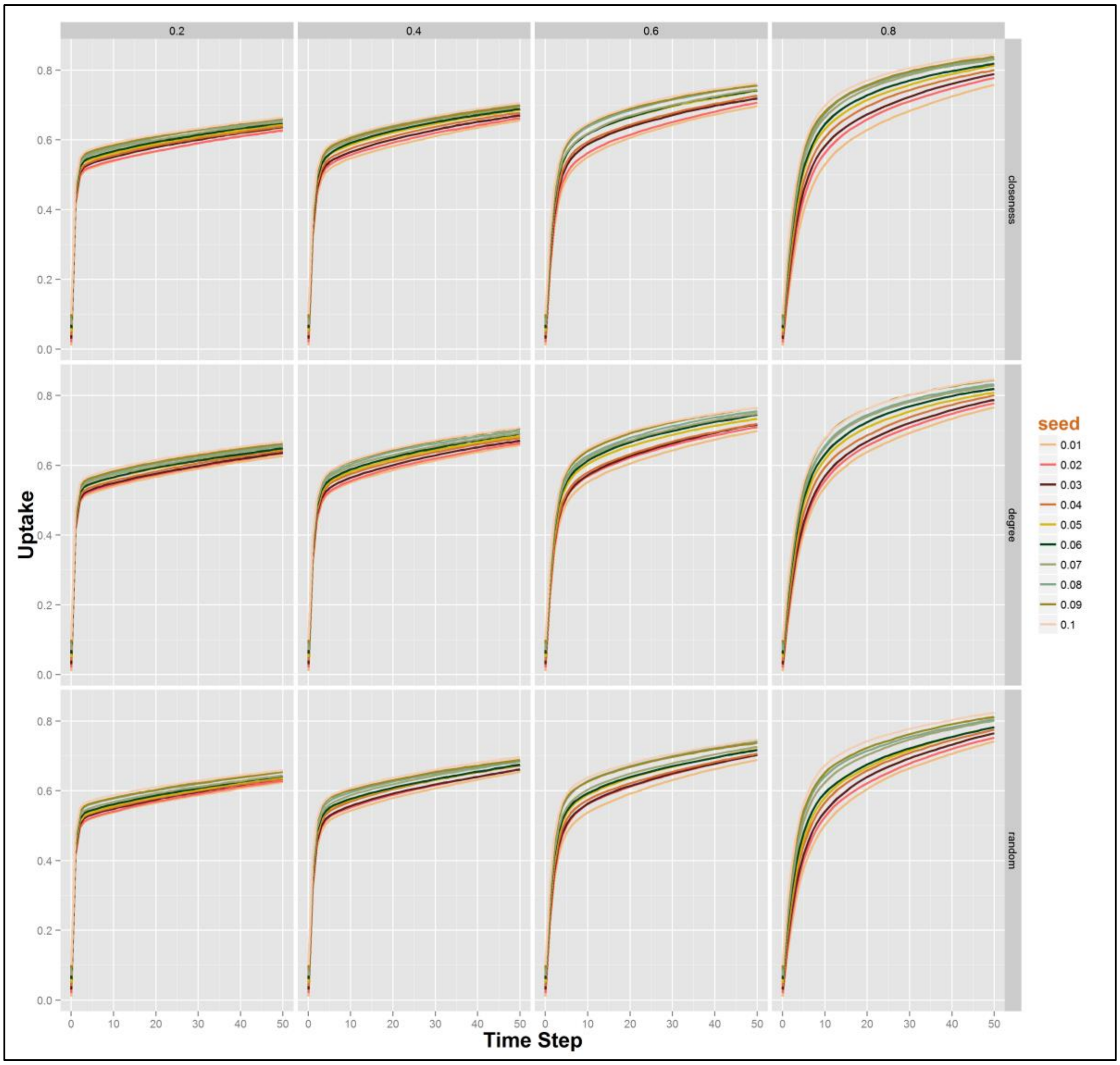


Figure 5-4 Individual Intervention, Real Network

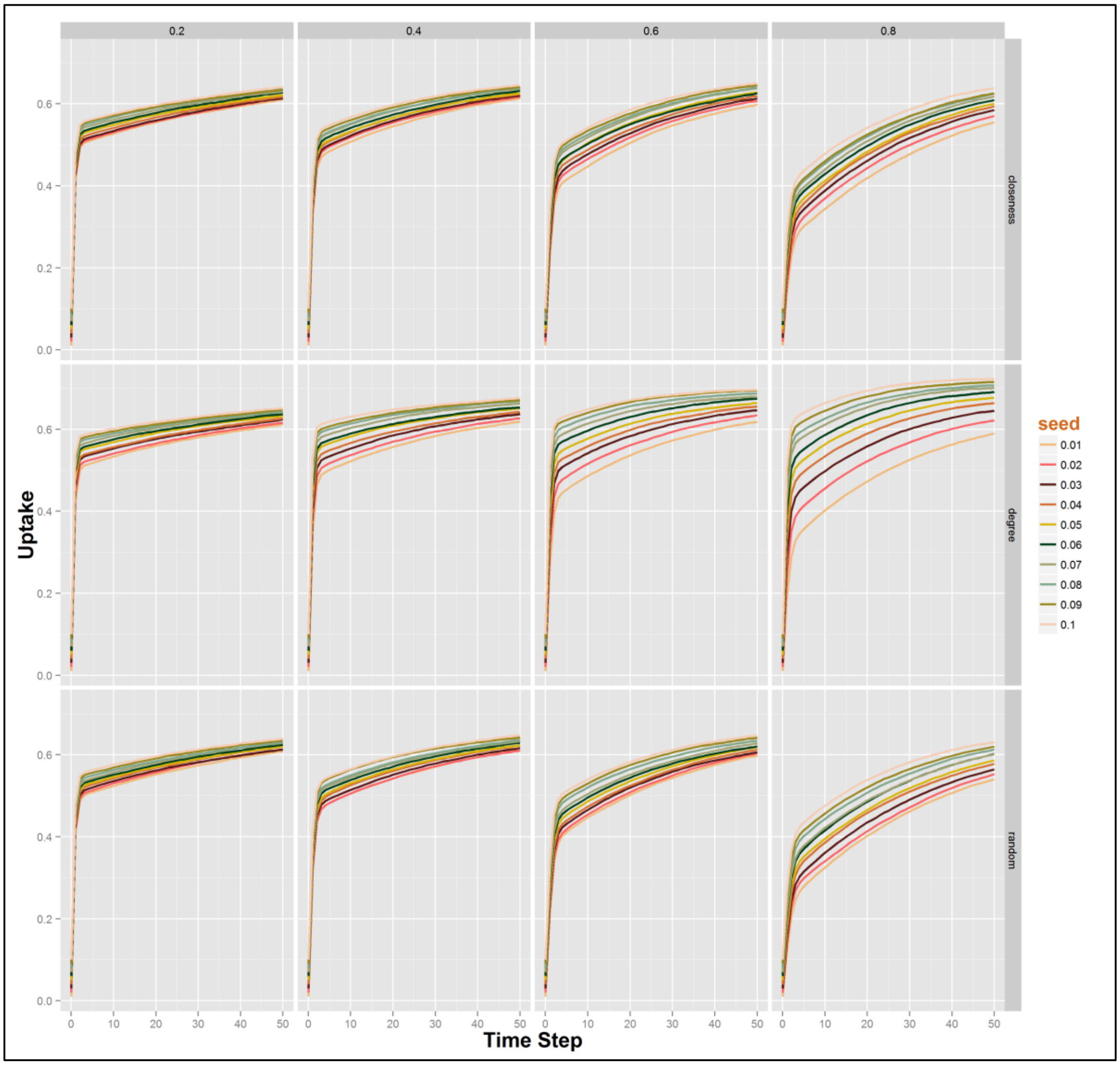


Figure 5-5 Individual Intervention, Scale Free Network

Third, the three seeding strategies do not differ from each other too much, in the real hospital network model, in terms of the final uptake (Figure 5-6), even though the shape of the adoption curve shows some small variance. Meanwhile, in scale-free network, the shape and final uptake are distinct to each seeding strategy (Figure 5-7). The results show that seeding by degree centrality dominates the two other strategies by bringing up the final uptake dramatically.

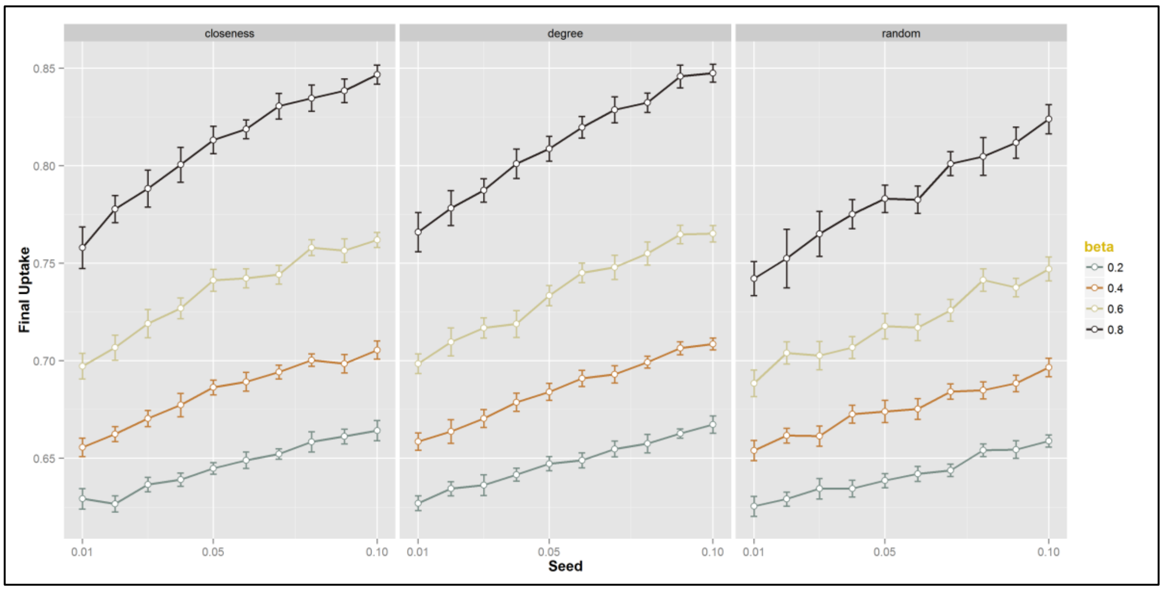


Figure 5-6 Final Uptake of Individual Interventions, Real Network

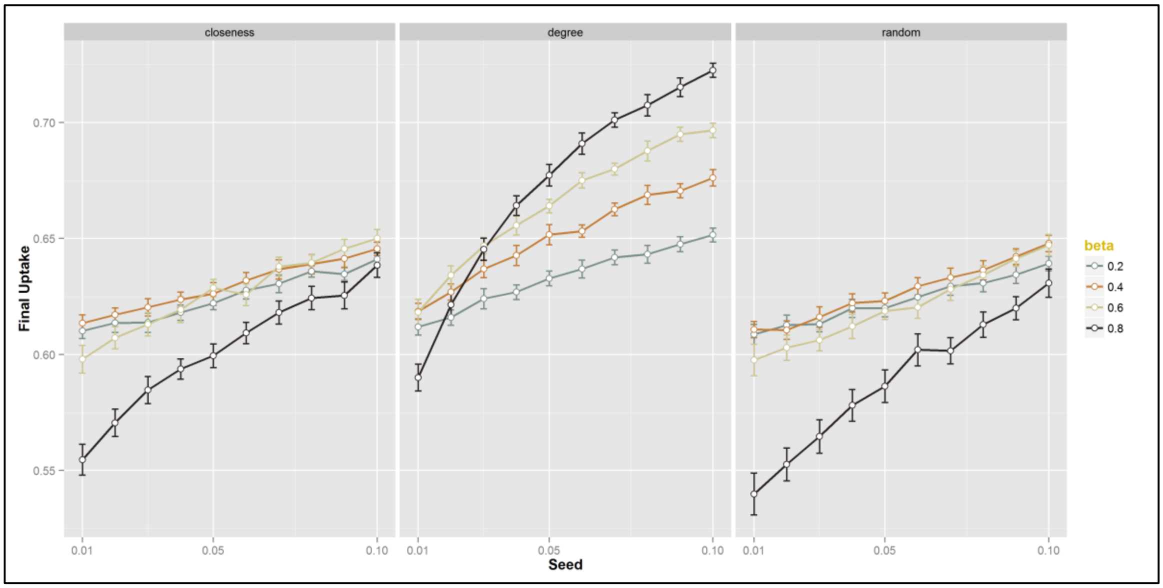


Figure 5-7 Final Uptake of Individual Interventions, Scale Free Network

In order to compare the differences in the two network models, in parallel, the difference in the average uptake at each time step is calculated, and also faceted by seeding strategy and varying β value (Figure 5-8). The x axes indicate the uptake difference – the uptake in real hospital network model subtracted by the uptake in scale-free network model. Consistent to previous findings, the difference is very minimal when β is at 0.2 and 0.4, but when its value grows, the difference is more noticeable. Most of the time, the diffusion performs much better in real hospital network model, at the difference is a positive value. The only exception is during the first several time steps when the scale-free network dominates but suddenly surpassed by the real network - the "cliff" observed in some of the subgraphs.

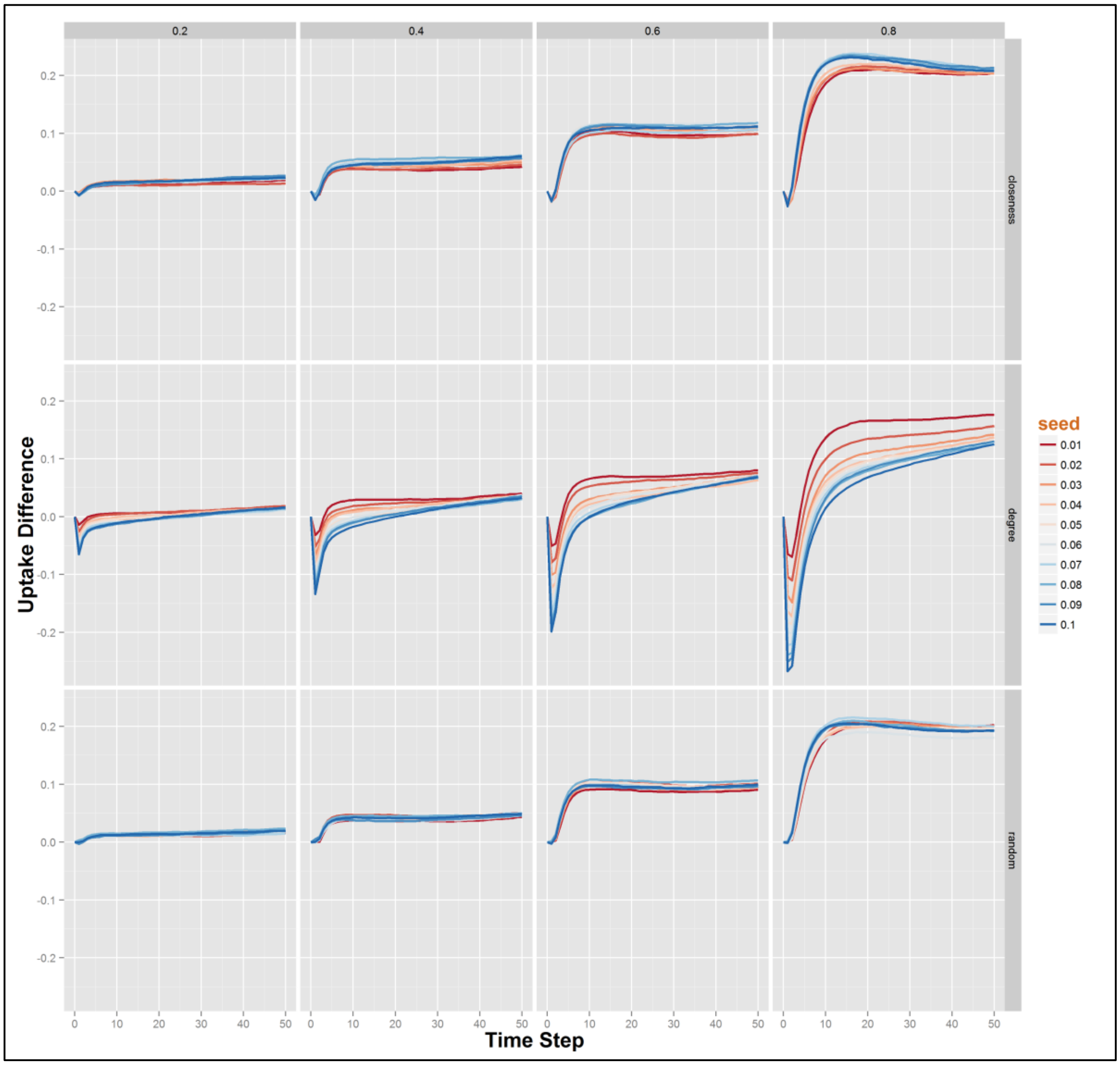


Figure 5-8 Uptake Comparison, Individual Interventions

5.2.2 Alteration Interventions

This chapter also looks at the network interventions through alternation by purposely changing the network structure. Policy interventions via rewiring or adding links can purposefully enhance some properties of the network so as to foster the diffusion of innovation in the network. Fennell and Warnecke (1988) indicated that in the medical realm, the notion of network was embraced by rewarding network based programs to health facilities encourage the adoption of medical innovations. This rationale is also applicable to our policy intervention scenarios of EMR adoption – incentives can be sent to hospitals to let them build network programs with other hospitals, especially with those previously not connected. To test the alteration interventions, simulations were implemented by varying the rewiring rate at the end of each time step and by randomly adding edges among the agents. The rewiring rate, *rr*, and rate of adding links, *lr*, were varied at (0.01, 0.01). Same as in the individual intervention models, the simulations were performed on both real hospital network and the ideal network – scale-free network (Figure 5-9 and Figure 5-10).

In contrast to the results from the scale-free network model where a greater uptake variance is observed as a result of the variation in rates of rewiring or adding links, in the real hospital network model the variance is very little. In fact, a very minimal variance in the uptake is found in the strategy of adding links although the final uptake is greater when β is at higher values. Meanwhile, increasing the rewiring rate is able to bring up the adoption rate (Figure 5-11). The variance in the final uptake as due to difference in the rewiring rate and β is also greater. The patterns observed in the real network model is different from those in the scale-free network model. The variance in the final uptake as a result of varying the rates of rewiring or adding links reaches approximately 30 percent in rewiring models and 20 percent in adding link models (Figure 5-12). The variance also grows in proportion to the value of β. Also, in the rewiring models, the system reaches a stable state very quickly, especially when the rewiring rate is high; but the adding links model has not reached any seemingly stable state by end of time step 5. Neither is the stable state identified in any of the strategies in the real hospital model.

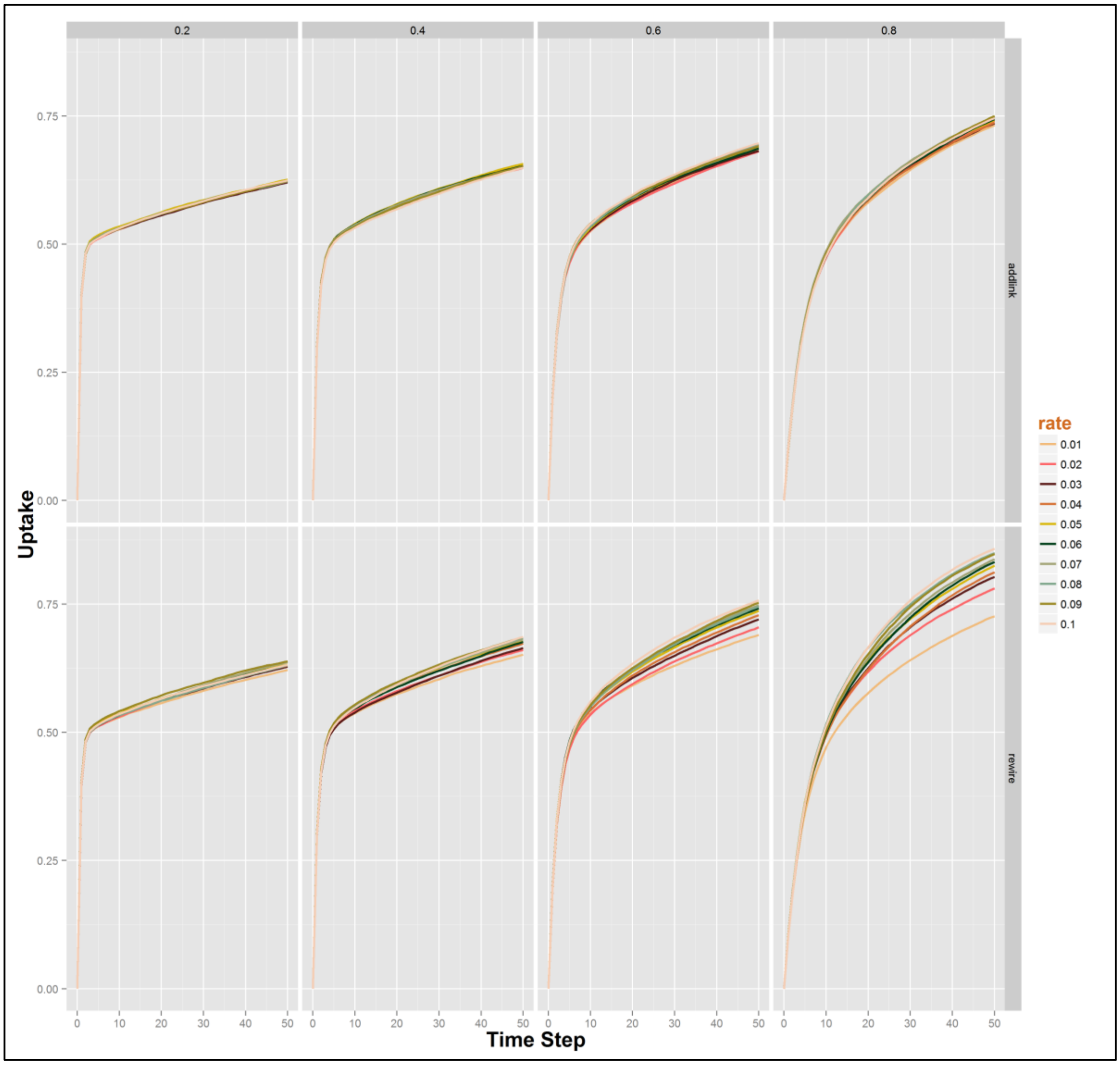


Figure 5-9 Alteration Interventions, Real Network

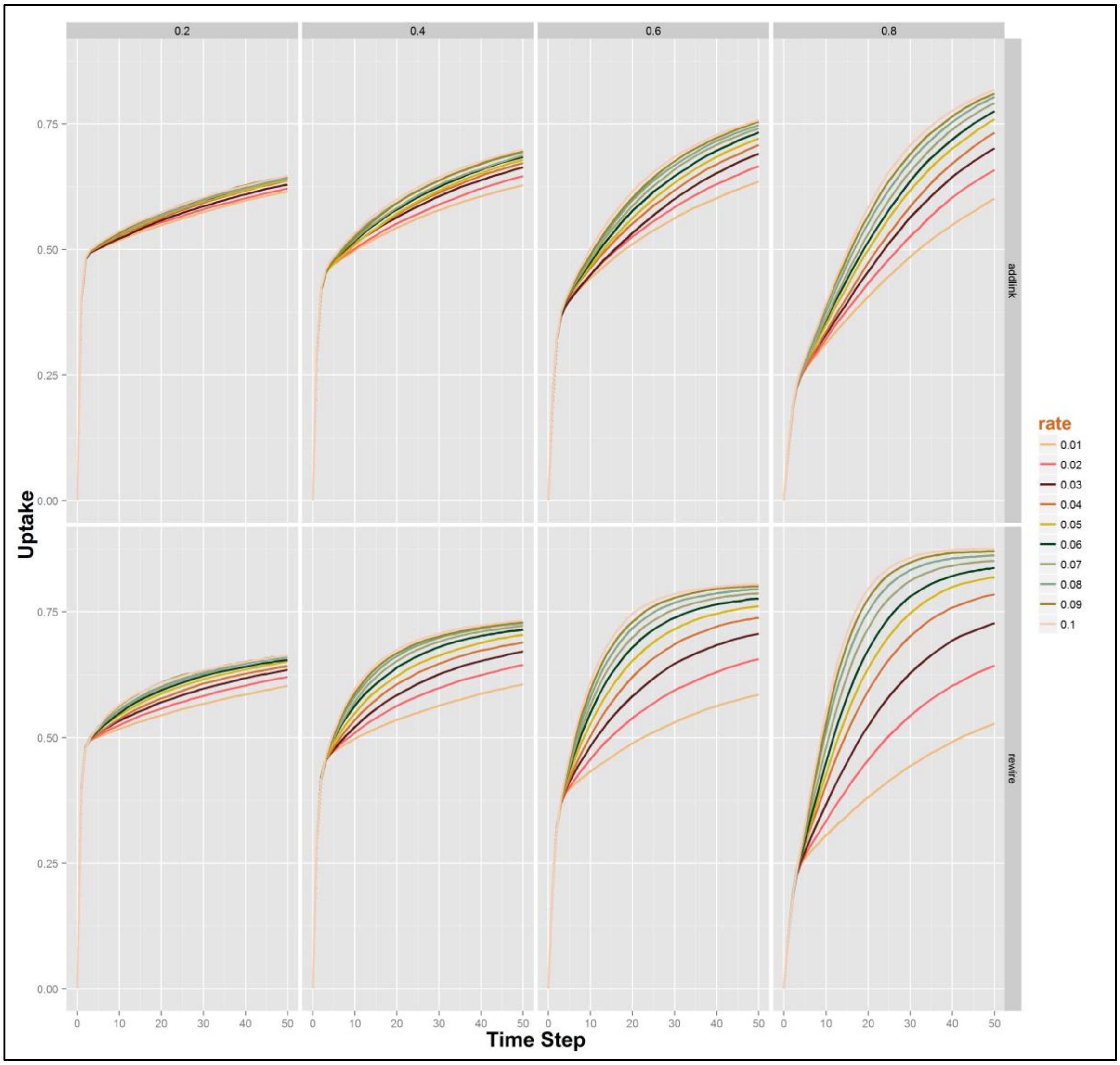


Figure 5-10 Alteration Interventions, Scale Free Network

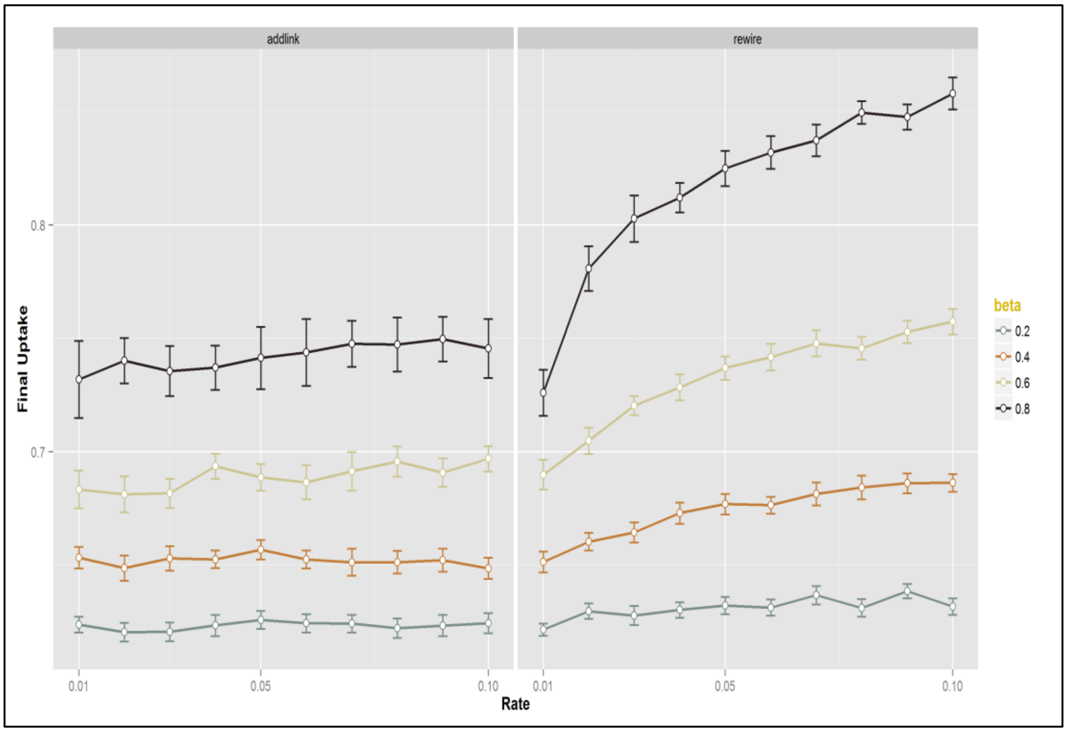


Figure 5-11 Final Uptake of Alteration Interventions, Real Network

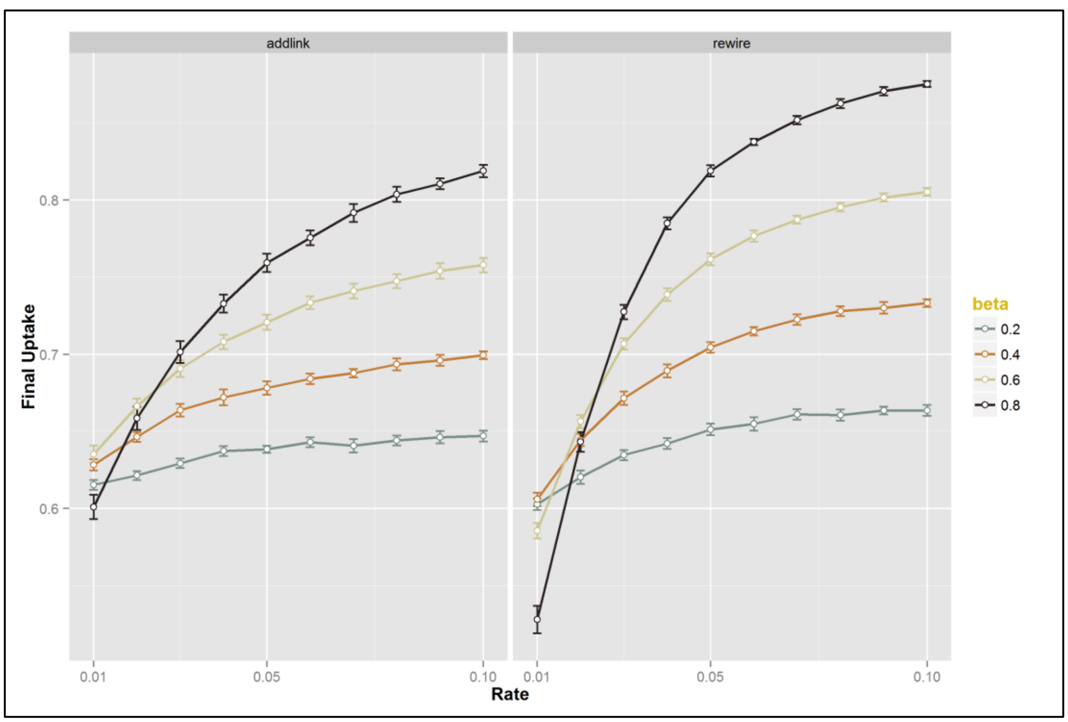


Figure 5-12 Final Uptake of Alteration Interventions, Scale Free Network

In contrast to individual interventions where the strategies perform better in the real hospital network than in the ideal network model (Figure 5-8), only in several cases does the real hospital network model dominate the ideal network model by means of the uptake difference (Figure 5-13). It shows that when the rate of rewiring or adding links is less than 0.05, the difference is positive. When the rate is greater than 0.05, the difference is negative in most of the time. The "cliff" noted in the individual intervention models is also observed here and lasts for several time steps longer. It indicates that at the beginning of the diffusion, varying the rate of rewiring or adding links can bring a lot of changes to the network, however after a few time steps, this advantage is swapped. Combining with the results earlier that the alteration strategies do not bring as much variance in the final uptake in the real hospital network as in the ideal network, it may has to do with the fact that the number of edges in the real hospital network far exceeds the that in the scale-free network (Table 5-2). Compared to the scale-free network, the hospital network is more saturated with links thus introducing rewiring or new links only brings significant change to the structure of the network when this type of change is of small scale.

Table 5-2 Network Properties, Real vs. Ideal

Network	Number of Edges	Number of Vertices
Real Hospital Network	5417	286213
Scale-free Network	4000	4999

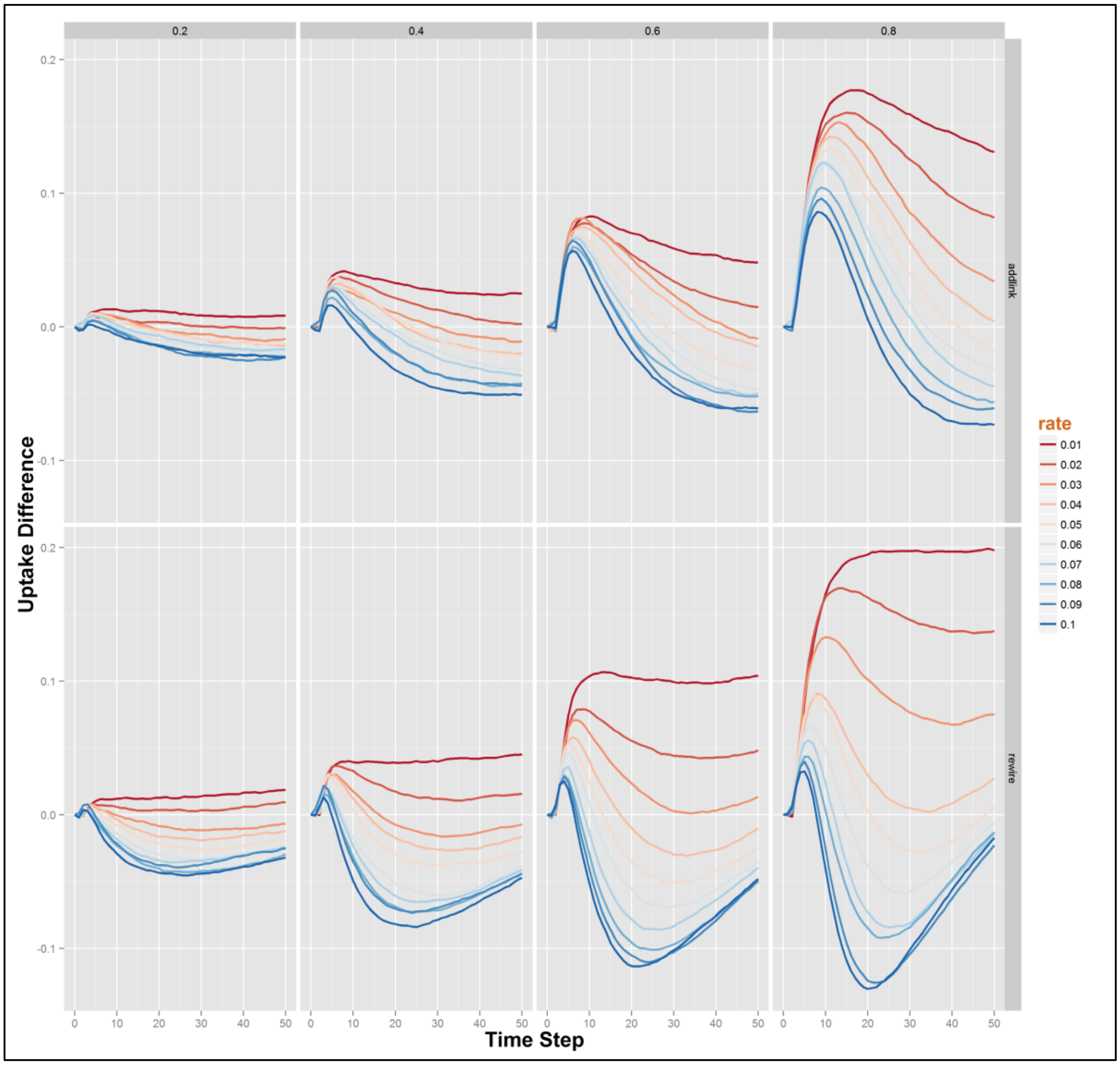


Figure 5-13 Uptake Comparison, Alteration Interventions

5.3 SUMMARY

This chapter investigated the network among hospitals using computational models in search of policy interventions that can maximize the value of the network. To that end, we built an agent based build and tested several network inventions as inspired by other network intervention studies in the past (Valente 2012). Two groups of interventions were examined: individual and alteration; and several different strategies were implemented under each group. Unlike previous studies of agent based model which were built primarily on ideal network models (for instance, see Cantono and Silverberg 2009; Gunther et al. 2011; Pegoretti, Rentocchini, and Marzetti 2012), the model implemented in this chapter used real network data and compared the results with scale-free – an ideal network.

The results suggest that for a network with some special properties, assumptions made from ideal networks may not be applicable. For example, in the model of alteration strategies, it is found that although the interventions can introduce a lot of changes in a scale-free network, it is not necessarily the case in the hospital network; the variance observed in the real hospital network models is not very significant, especially in the case of adding links.

The findings from this chapter also indicate that some strategies that do not act well in the ideal network are able to bring positive changes to the real network. The results from the individual intervention models showcase that whereas the three seeding strategies do not help to raise the uptake of the innovation, they do allow the adoption of the innovation to grow when more agents are selected for seeding at the beginning of the diffusion. Also, while the three strategies performed differently in the scale-free network, their final uptake in the real network are very close to each other. As a result, when implementing policy interventions to networks like the real hospital network, the question of importance may not be “who to select” but rather “whether or not”. The simulation shows that as long as there is intervention, the uptake will be brought up.

But the above discussion does not mean that network-based interventions are futile. In both groups of models, the introduction of network interventions bring up the final uptake, as compared to the baseline model, in all but one of the strategies. The only exception is the adding link model. It should also be noted that in the ABMs implemented in this chapter, the simulation were only implemented for 50 time steps. But the diffusion does not stop here. The uptake still has a tendency to grow and it grows faster as the rate of seeding or rewiring is greater in the real hospital networks. In contrast, the diffusion in the scale-free network tends to remain stable in most of the models by the end of time step 50. Thus, there is great potential in the network interventions to bring continuous changes to the network for a longer period of time.

CHAPTER 6 CONCLUSION

This dissertation set out to explore the adoption of Electronic Medical Records by hospitals in the United States. It integrates theories in the diffusion of innovation, methods and models in network analysis and analytical approach from agent based modeling. It addresses the presence of network among hospitals and empirically examines its role in the diffusion of EMR among US hospitals. The network considered in this dissertation is of organizational and spatial nature. We construct this network as inspired by previous literature which suggest that the health care industry is a highly institutionalized one and hospitals will comply with the organizational environment to seek support and legitimacy (Sherer, Meyerhoefer, and Peng 2016; Scott 2008; Fareed et al. 2015). To investigate the role of network in the diffusion of EMR among hospitals, it incorporates the hospital's structural roles in the network and applied the theories in the diffusion of innovation to study how the network has to do with the hospital's infectiousness, susceptibility and communication with others. With a significant role played by the network properties being found, this dissertation extended its scope by implementing agent based model to test different policy scenarios which could maximize the network properties among the hospitals. To my knowledge, this dissertation conducts the first empirical model on hospital data of such a long time spread (nine years), and the first agent based model of network interventions applied to the real hospital network, other than theoretical ideal networks.

6.1 SUMMARY OF FINDINGS

Chapter 3 performed the network analysis on hospitals. The purpose of this chapter was to analyze the structure of hospital networks, identify hospital's network roles using centrality measures and provide the network variables for the empirical studies in Chapter 4. To construct the network, this chapter also employed a clustering method – DBSCAN, to determine spatial clusters among hospitals. The analysis reveals the visualization of the organizational-spatial network among the hospitals. The network of hospitals presented in this dissertation shows that hospitals with more connections are located in certain states and regions across the United States. Although it is observed that a certain amount of hospital underwent structural changes during the study period, the overall appearance of the network changes only slightly. A disparity in terms of the hospital's centrality measures is also revealed. The distribution of the centrality measure is left skewed, indicating that on the one hand, there are many hospitals with relatively fewer connections, and on the other, a few hospitals are found to have greater number of organizational and spatial neighbors. The network typology of the hospital network constructed in this dissertation was also examined and compared with some notable network models. It is found that the network structure is not close to either small world or scale-free network. This special feature of the network is further explored in the following chapters.

In Chapter 4, the network was further probed through empirical models. This chapter addressed the research question about the role of the presence of network in the diffusion of EMR among hospitals. To that end, a Cox proportional hazard model was introduced to examine if the hazard of adopting the EMR is associated with the hospital's infectiousness, susceptibility and network exposure as captured by several network measures. The results from this chapter suggested that in addition to demographic variables which have been consistently found as significant indicators of the hospital's EMR adoption, the network variables have also shown their statistical significance. Hospitals with more connections, and more direct connections with adopted neighbors are more likely to be adopters. In addition, the regression analysis found that when there are more hospitals throughout the nation have adopted the EMR, the hazard is higher for hospitals with more connections (higher degree centrality), which is captured by an interaction term introduced into the model. The findings from the Cox model is consistent with results from a preliminary logistic regression analysis.

With Chapter 4 indicating that the network properties does distinguish on their propensity to adopt the EMR, Chapter 5 set out to identify any policy intervention schemes that can utilize the network properties of the hospitals. In this chapter, a set of agent based model were implemented. Basically, the model assumes that the decision to adopt the EMR is the results of a joint influence of information and social influence. Whereas information influence is impacted by the adopter's internal judgement on the quality of the innovation, social influence is the external communication and influence from peers. The policy interventions, therefore, were introduced aiming to maximize the social influence. Network-based policy interventions being tested in this chapter are inspired by previous studies in research fields including, but not limited to, economics and health. In particular, computational simulations were conducted to test network interventions focusing on individuals and network alteration. The simulations were also separately performed on the real hospital network and an ideal network - scale-free network – to compare the results.

It was shown that for a special network structure like the hospital network, assumptions made from ideal networks may not be applicable. It also indicates that some network interventions not found to bring significant change in the ideal network are able to act well to the real hospital network. The chapter have also established a primitive model for future exploratory of hospital networks.

6.2 LIMITATIONS OF THIS STUDY

This dissertation has several limitations. First, the organizational-spatial network constructed here does not distinguish between the organizational and spatial links. In other words, the organizational and spatial influence among the hospitals are treated equally. No weight was given to the organizational or spatial edges when plotting the network. This may pose a problem as the impacts from two types of influence are not equivalent.

Second, the EMR adoption was treated as a single occurrence of event in this dissertation. However, as noted in Chapter 1 and 3, the adoption of EMR is not the implementation of one single technology but rather a cumulative process of implementing a number of information technologies with different capabilities across multiple units of the hospital. In this dissertation, we selected the adoption of a Basic EMR as the event and examined the inception of EMR at hospitals. The study has not looked at the adoption of Intermediate or Comprehensive EMRs and how the network of interest affect hospitals' adoption of EMR with more advanced capabilities. The influencing mechanisms can work differently since a higher level of EMR has a stricter requirement on interoperability, financial and technological readiness. Also note that due to the discrepancy in data collection by different data providers, the methodologies available to measure the EMR capability at hospitals also vary. This dissertation employed the categorization methodology commonly adopted by scholars using the HIMSS database; others base their study on data sources such as the AHA may found this method contradictory. The methodological difference may render the interpretation of the findings in this study restricted. Also, the regression analysis in this dissertation did not include the influence of hospital's Medicare and/or Medicaid participation or geographic disparity (urban/rural) in the analys.

Third, to examine the potential offered by network to develop incentives and policies to affect adoption of EMR by hospitals, this dissertation does not evaluate the effectiveness of different EMR-related policy interventions. It is true that different types of intervention – financial incentives, standard setting, awareness campaign and education – can influence the hospital's adoption of EMR differently. The agent based model in this dissertation only considers a type of policy intervention that will eventually lead to the implementation of EMR at the hospital and does not address the impacts of different types of interventions on hospitals' propensity to adopt the EMR.

Fourth, the agent based model and the network interventions implemented in this dissertation so far provide only some framework for using computational models to study the diffusion of innovations among hospitals. The decision making rules are still very preliminary and have not been adapted to account for the influence mechanisms among hospitals. Explanations to the intervention strategies also requires further development. For example, in the alteration strategies, it was found that rewiring strategies serve at an effective way to foster the adoption of EMR and the uptake of the innovation was brought up. However, theoretically rewiring means that randomly delete and/or add links among certain vertices. In real world cases, this is not realizable as one cannot cut off connections among hospitals. Thus, better explanations should be developed to further validate the findings from computational models.

6.3 POLICY IMPLICATIONS

In February 2009, the Health Information Technology for Economic and Clinical Health (HITECH) Act were written into law, as part of the American Recovery and Reinvestment Act (ARRA). The Act was enacted in response to the lagged development in the adoption of HIT and aimed to foster the adoption of HIT by health providers who were facing financial and technological barriers. In addition to provide incentive payments to eligible hospitals and professionals through the Center for Medicare and Medicaid Services (CMS), the Act also benchmarks the “meaningful use” in three EMR stages. By “meaningful use”, it indicates that providers must prove that certified EMR technologies have been used in ways that can be measured significantly in quality and in quantity(Health Resources and Services Administration 2016). By January 2016, 4,450 hospitals in the United States have been registered in the incentive program(Center for Medicare and Medicaid Services 2016).

Although recent literature has suggested that the introduction of the HITECH Act has positively contributed to the adoption of EMR (Sherer, Meyerhoefer, and Peng 2016), findings from the empirical model in this dissertation suggest that disparities in hospital demographics still play a role in hospital’s propensity to adopt the EMR. For example, the size of hospitals proves to be a consistent significant indicators across all models. Big hospitals have the financial capability and technological infrastructure to implement the EMR technologies; in contrast, small ones may have been restricted by their financial and technological readiness. In Chapter 4, the results from both logistic regression and the Cox model both suggest that critical access hospitals are less likely to adopt the EMR. Note hospitals of this type are small and rural hospitals. Previously literature has also highlighted the gap between critical access hospitals and other types in terms of the propensity to adopt the EMR (Adler-Milstein et al. 2014). As a result, when introducing incentive policies to hospital, those with barriers in initial financial investment should not be neglected.

The theme of this dissertation is network; it hopes to provide policy recommendation through the lens of networks among hospitals. In the empirical study, it was found that hospitals with more connections, and among those who are adopters, are more likely to adopt the EMR. Can we increase their network connections by building new networks among the hospitals? The answer is probably yes. In this dissertation, we only considered the network of hospital system and spatial proximity; in the realm of health and medical studies, network building is nothing new. As mentioned in previous sections, the notion of network has a long tradition. It was used to accelerate the diffusion of medical innovation through awarding funding and contracts to certain health facilities to develop network programs (Fennell and Warnecke 1988). In health and medical domain where networks of member organization, conferences and special interest group is everywhere (Angst et al. 2010), policy interventions could look at connecting hospitals from large and small systems, populated and underserved area, large and small. As Otto and Nevo (2010) noted, besides financial incentives, policy interventions on health IT often center on awareness campaign, educational interventional and training and improved standards and research programs. Policies of this kind should incorporate the notion of network building in their plans.

This study found that when introducing policy incentives, selecting early adopters based on their network properties may serve as a feasible approach. The results from the agent based models showed that selecting early adopters by their degree or closeness centrality measures can raise the adoption rate. Targeting individuals with higher centrality measures can help to facilitate the information transmission by effectively connecting to the rest in the network. This is not to suggest that the financial incentives should be prioritized towards the ones with more connections; rather, when educational, training or awareness campaign were introduced, they could be taken place at hospitals with greater structural importance in the network in order to maximize the influence.

6.4 IMPLICATIONS FOR FUTURE RESEARCH

This dissertation has several implications to future research. First, the study presents the network structure and properties among U.S. hospitals. In addition to findings in existing literature about the network based on micro-level, patient exchange network among hospital, we consider a network of more organizational and spatial nature. The network investigated here can be informative to studies with a focus on the organizational and spatial characteristics of hospitals.

Second, as discussed earlier in this study the adoption of EMR is only studied as a single occurrence - the three fundamental EMR technologies being adopted. But the EMR implementation is also a continuous process. There are EMR technologies of higher capabilities, the adoption of which based on previous adoption of simpler ones. Even though we do not include the adoption of higher capability EMR technologies here, the network analysis and its results obtained from here can serve as good exploratory tools for future study of the subject. Models of the diffusion generational innovations can be included to further explore the issue.

And third, in agent based models the policy interventions using the network properties of the hospitals is tested. The results can provide knowledge about the critical nodes and links among hospitals. This study only set out a preliminary computational model of hospitals and adoption rules, it can be expanded to better account for the influence mechanisms among hospital. Also, in his article Valente (2012) noted four types of network interventions: individual, alteration, segmentation and induction. This dissertation only examined two of them (individual and alteration). Future models can be developed to incorporate the other two strategies.

APPENDIX

Appendix 2-1 Literature Search Strategy

Research Area	Search Strategy	Keywords
Electronic Medical Records	Introduction to EMR	Electronic medical records
	Empirical studies of EMR adoption	Electronic medical records/EMR, empirical, adoption, model
Network Analysis	Network measures	Network, measure, macro, micro, meso
	Network analysis	Network, clustering, method
	Network models	Network, typology, small world, scale-free
	Network analysis in health and medical studies	Network, analysis, model, health, medical
	Diffusion theory	Diffusion, innovation
Diffusion of Innovation	Network in diffusion of innovation	Network, diffusion, innovation
	Organizational and spatial network of EMR adoption	Organizational, spatial, EMR, network, adoption, model, empirical
	Agent based modeling of innovation diffusion	Agent based model, diffusion, innovation
Agent Based Modeling	Networks in agent based modeling	Agent based model, network, typology
	Opinion leaders in diffusion of innovation	Agent based model, diffusion, innovation, opinion leader
	Spatial network of agents	Agent based model, spatial

Appendix 2-2 Network Measures and Methods

Macro Measures

One of the most fundamental and defining characteristics of network structure is *degree distribution*. Degree distribution describes the relative frequencies of nodes with different degrees and thus provides important information about how nodes with different types of neighborhood are linked in the network. To make degree distribution easy to interpret, it is usually presented by plotting the distribution as a function of degree.

By looking at the degree distribution, an important network property can be obtained – the network topology. Network topology captures the layout of nodes and links over a network. Networks with different natures are embedded with different arrangement of the elements. A network can be *regular*, where all nodes have the same degree, *random*, where the links are placed between paired nodes at random, *scale-free*, where most nodes have very small degree whilst a few have very large degree. Introduced by Erdos and Renyi (1960) ,the random graph topology has been stimulating the research on the subject for decades. But it was later realized that many real world networks do not present the properties of random graphs. In contrast, nodes are indeed heterogeneous in their degrees and exhibit a power law. Such scale-free network structure, as pioneered by Barabasi and Albert (1999), can be found in many of networks we experience everyday such as the World Wide Web and some social networks. Because scale-free network typology suggests the different sizes of neighborhoods that one might have, it thus has several implications for the heterogeneity in network components. Ideas such as hubs and opinion leaders are important to understand how certain individuals can exert their influence and will be elaborated in next section. As a result, degree distribution provides a simple description of the structure of the network and helps to distinguish between different networks.

The global patterns of network can also be captured by the paths between nodes. *Average path length*, the mean of all shortest paths between nodes, is another measure about the network structure. In addition to the scale-free network structure found in our daily networks, another significant phenomenon is the so-called small world effect. A famous example of the small world network is the experiment conducted by Stanley Milgram in the 1960s who asked randomly selected participants in Wichita, Kansas and Omaha, Nebraska to send mails to recipients located in Boston, Massachusetts. If the participant does not know the recipient personally, the participant was asked to forward the mail to someone he or she knows might be more likely to know the target contact. The results of the experiments showed that of the letters finally being delivered to the target contact, the average path length approximates to six, a number echoing the famous "six degree of separation". If a network is a highly-ordered regular network, it tends to have large average path length and highly clustered, as reflected by clustering coefficient[16]. If the nodes of a network are connected by a random fashion, the network usually have small average path length and clustering coefficient. But what we have observed in many real life networks is that they are clustered but with small path length. Thus, the small-world networks depart from these two typologies and offer another perspective to study network properties in our social and economic networks. As Watts and Strogatz (1998) showed,

[16] Clustering coefficient will be discussed in detail later.

small-world network is supported by many empirical examples such as film actors and power grid. Average path length thus provides another indication of the structure of the networks. Used alone or combined with other indicators, the measure tells information about the cohesiveness of the network and allows one to draw conclusions about the typologies.

Watts and Strogatz (1998) introduced another measure of the macro level structure of network, the *clustering coefficient*. Clustering depicts how nodes in a network tend to cluster into certain groups. A network can have high clustering or low clustering, depending on the nature of the network. As mentioned above, our social networks in most cases have high clustering, because people tends to share similar social networks with their friends. But if a network connected in a random fashion, individuals might now cluster with each other and, as a result, the network has a low clustering. Clustering thus provides a tool to measure the structure of the network with the groups and positions nodes. Clustering coefficient is calculated by averaging each node's network density – the actual links in a node's neighborhood divided by the total of allowable links. A network with high clustering coefficient indicates that the network is "clumpy" and there might be some sub-graphs existing in the network. In contrast, a network with low clustering coefficient suggests that two nodes both connected to a third node do not link with each other and nodes are not more likely to cluster.

Meso Analysis

Clustering coefficient, from the large scale structure of the network, tells about the groups or communities in the network and indicates that the different values in the coefficient in different types of network has to do with the formation of the groups and communities. If a network presents high clustering, the next step of exploration is to identify the clusters and compare the different clusters and their member nodes. In the studies of diffusion, the identification of groups or clusters is beneficial to understand adoption at group-level and identify specific hierarchical positions in the network(Valente 2010). Many clustering methods have been developed over the past several decades using different induction principles. Early clustering algorithms are primarily partitioning or hierarchical methods; later methods are developed that incorporate cluster-based, modularity-based and divisive models.

Partitioning clustering method indicates that for a given network of n nodes, partitioning clustering would classify them into k clusters, which satisfy the requirements that 1) a least one data point exists in each cluster and 2) each object can only exist in one cluster (Han and Kamber 2011). The partitioning method then creates an initial partitioning, given k, and uses an iterative relocation technique to improve the partitioning so that the k clusters are created where nodes are same enough if they are in the same cluster and far apart if they belong to different clusters. The techniques to perform the partitioning clustering employ heuristic methods(Han and Kamber 2011; Fortunato 2010). The most widely used algorithm is *k-means* clustering. It starts by selecting k data points, each by themselves represents a cluster mean. The remaining data points are assigned to the cluster where the cluster mean is closest to the object. The system is then updated by obtaining a new mean for each cluster and iterates the process until the means are stable and the clusters no longer change. A square-error criterion is used to determine if there is no change.

In addition to partitioning clustering, *hierarchical clustering* is another popular and classic method for community detection. It operates the clustering based on the assumption that the structure of the network is hierarchical. And compared to partitioning clustering that requires the specification of number of clusters in advance, hierarchical clustering allows the fact that little is known about the community structure and can probe that in a top-down or bottom-up fashion. And therefore the clustering can be performed with either agglomerative (bottom-up) or divisive (top-down) algorithms. Agglomerative hierarchical clustering starts from each vertices as the initial cluster and merge the singleton clusters into larger and larger clusters until a desired condition is achieved. In contrast, divisive hierarchical clustering starts from the all vertices as one cluster and subdivides the cluster into smaller ones until certain termination conditions are met. The agglomerative and divisive clustering are performed based on the similarity among the vertices. Several methods are employed to measure the similarity among vertices and they thus further divide the hierarchical clustering methods. The main methods include single-link clustering, complete-link clustering, and average-link clustering.

One shortcoming of k-means partitioning clustering method is that it does not allow convex, arbitrary shaped communities. An approach to discover the existence of arbitrary shaped communities was developed on a *density-based* notion. The method performs clustering based on the dense region that are separated by low density noisy points(Han and Kamber 2011). One major density-based method, Density-Based Spatial Clustering of Applications with Noise (DBSCAN), was developed by Ester et al. (1996) which identifies clusters based on the local connectivity and density function of points. The notion of cluster, as in DBSCAN, is regarded as a maximal set of density-connected points. The algorithm requires two parameters to be defined prior to the analysis: the maximum radius of the community, ε, and the minimum number of points in an ε-neighborhood of any point,m. The algorithm operates by first arbitrarily selecting a point p. It then retrieve all points density-reachable[17] from p with regard to$\varepsilon$and$m$. A cluster is formed if p is a core point. Otherwise, no points are density-reachable and the algorithm visits another point of the database. The process continues until all of the points in the database have been processed.

In addition, *divisive clustering* is a method that identifies the edges linking nodes of different communities and remove them. The Girvan and Newman (2002) model is a salient example under this category. It suggested a divisive algorithm to identify edges that are between the communities. The algorithm employs the notion of "betweenness" proposed by Freeman(1977). The betweenness of a node is defined as the number of shortest paths between other pairs of nodes passing through it. Betweenness thus suggests the influence of a node over the flow of information between other nodes(Girvan and Newman 2002).Girvan and Newman (2002) generalized Freeman's definition of "betweenness" and applied that to edges betweenness, in order to identify edges in a network connecting other pairs of nodes. To this end, they defined the edge betweenness of an edge as the number of shortest paths between a pair of nodes running along it. The community detection can then be performed by identifying high edge-betweenness, inter-community edges who connect the sub-communities of the whole network.

---

[17] A point p is density-reachable from q wrt.ε and $m$if there is a chain of points $p_{1,...,}p_n, p_1 = q, p_n = p$such that $p_{i+1}$is directly density-reachable from p_i

In order to effectively identify *good* clusters, quality functions are adopted to assess the clustering results. The most commonly used quality function is *modularity* developed by Newman and Girvan(2004). The measure is developed by comparing the number of actual edges falling within groups with the expected density if the edges within the network are placed randomly and regardless of community structure, as a random graph is not supposed to present any cluster structure. The modularity then equals to the number of edges falling within groups minus the expected number of edges when the edged are randomly placed in the network (a null model). The modularity can be either positive or negative, and a positive modularity indicates the presence of community structure. The modularity thus provides an effective way to estimate the goodness of a clustering. And the pursuit of good clustering can be achieved by employing algorithms that optimize the modularity.

The clustering methods as discussed above each has its advantages and disadvantage. The selection of an appropriate method to conduct clustering analysis depends on the purpose of the study and nature of the network being studied. For example, the error-minimizing k-means partitioning methods works well when the clusters are compact and separated from one another. It also has the advantages of linear complexity and thus computational attractiveness, ease of interpretation, implementation simplicity and adaptability to sparse data (Rokach and Maimon 2005). However, the k-means partitioning method also has the several features that pose some of its disadvantages. First, it involves the selection of starting partition points so that the identification of clusters depends on the initial choice of the starting points. Second, it requires the specification of the number of clusters prior to the analysis. Third, it does not distinguish if an object is an outliers and thus outliers are forced to join one of the clusters (Milligan and Cooper 1987). And finally, it cannot identify clusters with nonconvex or irregular shapes(Han and Kamber 2011). Compared to partitioning clustering method, hierarchical clustering is attractive in that it does not call for a preliminary knowledge of the number of clusters. It also has the strength of being versatile and allowing for multiple partitions (Rokach and Maimon 2005). However, the method has no back-tracking capability so that one cannot undo the clustering once a merge or split has been completed. In addition, the method requires the selection of merge or split points, which sometimes can be arbitrary, and since it can never undo the process, the decision becomes critical. And finally, a major disadvantage of hierarchical clustering method is its inability to scale due to its non-linearity and the computational complexity involved(Milligan and Cooper 1987; Han and Kamber 2011; Fortunato 2010). The distance-based clustering method is able to identify clusters of arbitrary shapes and does not have to specify the number of clusters a priori. It can also detect points with low connectivity and exclude them from any cluster(Ester et al. 1996). The concept of betweenness employed in divisive clustering render it advantageous to other methods to model information spread by its computational complexity and representation of centrality (Fortunato 2010).But it does not consider overlapping clusters since each vertex is assigned to a single cluster. As a result, we can see that the size of the network, distribution of vertices and the purpose of clustering all affect the choice of the method, and sometimes a compromise has to be reached between accuracy and running time (Danon et al. 2005).

Micro Measures

With meso level analysis, one is able to identify groups of nodes that are more likely to cluster with each other. But even within a same cluster, nodes can play different roles. The positions of the nodes are different in the network so that some might be more central and some are comparatively peripheral. Because the connectivity of nodes determines how information can be transmitted through them, it is usually believed that nodes with different positions will thus access to information and exert their influence differently. As a results, measures that are concerned with the centrality of individuals allows one to examine the flow of information and influence related to the heterogeneity of individuals. Many centrality measures have been developed and widely used in the area of network analysis and are based on different aspects of the concept of centrality. Some of the most popular measures of centrality include *degree centrality*, *closeness centrality*, and *betweenness centrality*.

Degree centrality is most straight forward measure to identify the central nodes in a network. It is calculated through dividing the *degree* of a node by n-1, n being the total number of nodes in the network. Because in a network of n nodes, a node can have as many as n-1 and as few as 0 connections to others, degree centrality therefore falls between 0 and 1, and larger scores are associated with more central nodes. Degree centrality provides a simplistic measure of how connected the nodes are; but in some cases the importance of the nodes are also associated with their locations in the network. Especially in networks where information or influence being transmitted could decay with distance, degree centrality might seem insufficient to capture the changes.

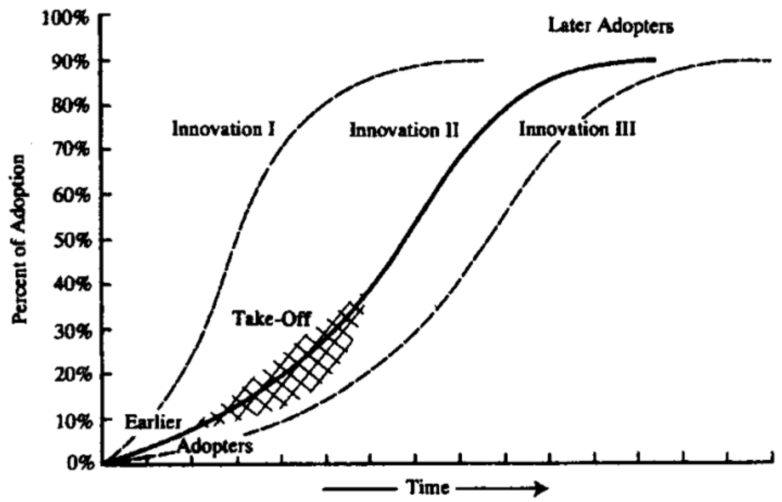
Closeness centrality, another centrality measure that incorporates the average distance of a node to the others in a network, serves an alternative tool. The average distance measure is calculated as the total of shortest distance one nodes has with (n-1) others divided by n-1. Since the measure suggests that low values for more central nodes and high values for more peripheral ones, average distance is then taken its inverse value which is called the normalized *closeness centrality*. The inverse measure is more easily understandable because it denotes higher value for more central nodes and vice versa. Compared to degree centrality, closeness centrality is able to capture the importance of nodes that are not connected to many others but may serve as mediating hubs. It can also capture the centrality of nodes in non-Euclidian social space and distinguish the links directed to and from a node(Valente 2010). But it also has some problems. As Newman (2010)pointed out, the range of the value of closeness centrality is very small and thus the measure is very sensitive to fluctuations – even the smallest change in the network structure can the order of the values substantially.

Another measure of centrality that looks at the position of nodes is *betweenness centrality*. It indicates the importance of the node in relation to how it lies on the shortest paths between other nodes. The betweenness centrality is calculated by first obtaining the ratio of the number of shortest paths one nodes lies on over the total number of shortest paths in the network. The ratio is then divided by the maximum possible links a network can has without the node, (n-1)(n-2) for directed or (n-1)(n-2)/2 for undirected networks. This notion of betweenness allows one to examine the nodes not only in terms of centrality but also bridging (Valente 2010). Because in networks the power and influence are reflected in how nodes can affect the transmission of information, the central nodes in this context can control how information is transmitted and/or filtered. Thereby nodes with high betweenness can derive higher power and the removal of such nodes involves structural changes of the network(Newman 2010).

The micro-level measures of centrality examine the structure of network through their positions and indicate their power in the diffusion of information and products. In real world networks, these central nodes are often perceived as opinion leaders or key players. With regard to the diffusion of innovation, these central individuals of the network are more likely to hear about the innovation earlier than those on the periphery(Valente 1995) and adopt early(Liu, Madhavan, and Sudharshan 2005).In addition, their positive or negative comments of the innovation can diffuse faster and broader across the network than those from others. Because the identification of the central players are critical for selecting of targets in the interventions or campaigns to foster the diffusion(Iyengar, Van den Bulte, and Valente 2010), the centrality measures are often used to identify the opinion leaders and key players[18] and test the effectiveness of the interventions or campaigns targeting different populations of the network across the diffusion process.

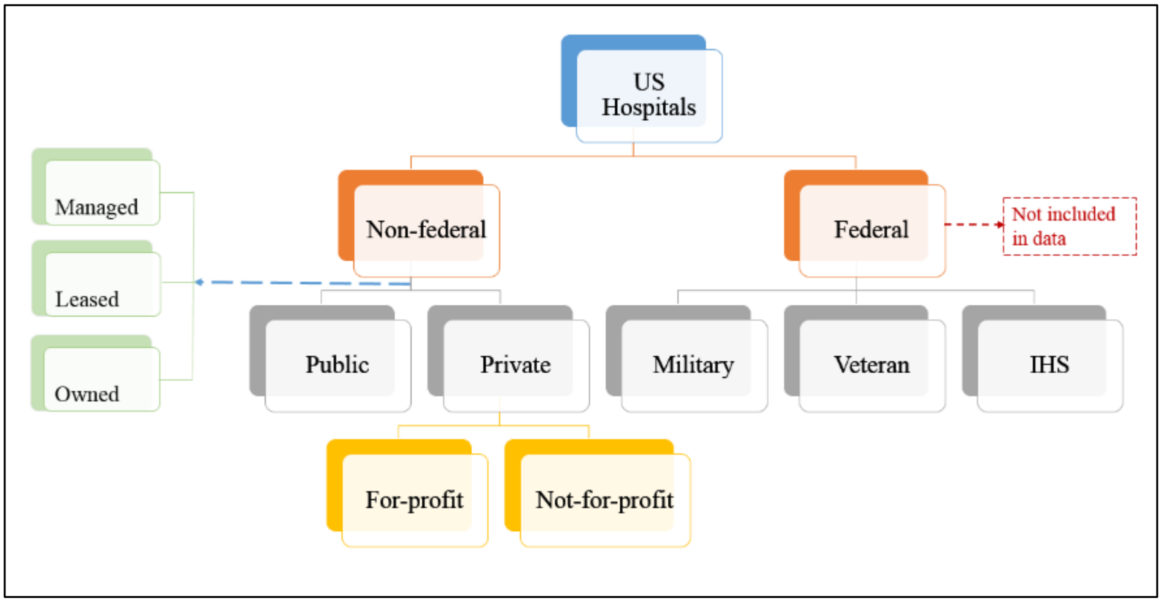
[18] The identification techniques also include self-reported leadership and key informant technique. But these techniques are not obtainable from the statistics and metrics from the network structure. As a result, we will skip the discussion of them here.

Appendix 2-3 S-Curve of the Diffusion of Innovation (Rogers 2003, 11)



Diffusion is the process by which (1) an *innovation* (2) is *communicated* through certain *channels* (3) over *time* (4) among the members of a *social system*.

Appendix 4-1 US Hospital Types



Appendix 4-2 Correlation Coefficient, Number of Beds and Number of Staffed Beds

	NofBeds	**NofStaffedBeds**
NofBeds	1.0000000	0.9726312
NofStaffedBeds	0.9726312	1.0000000

Appendix 4-3 Multicollinearity Test including Number of FTE

	GVIF	Df	GVIF^(1/(2*Df))
Degree Centrality	1.248234	1	1.117244
Direct Exposure	1.08003	1	1.039245
Number of Beds	3.466474	1	1.861847
Number of FTE	2.94981	1	1.717501
Type	2.422264	15	1.029929
Ownership Status	1.057206	2	1.014004
System Adoption Rate	1.15423	1	1.074351

Appendix 4-4 Cox Model with and without variableNumber of FTE

With NofFTE

```
> summary(model.1)
Call:
coxph(formula = Surv(time, time2, basicStatus) ~ logdeg + directPer +
    logbed + NofFTE + Type.general + Type.critical + Type.acute +
    Type.other + Type.arehab + Type.cardio + Type.eent + Type.Geriatric +
    Type.neuro + Type.obgyn + Type.osteo + Type.ped + Type.pw +
    Type.psych + Type.wh + Type.onco + Type.op + owned + managed +
    inter, data = totalnew)

  n= 12983, number of events= 3767
   (3909 observations deleted due to missingness)
```

Without NofFTE

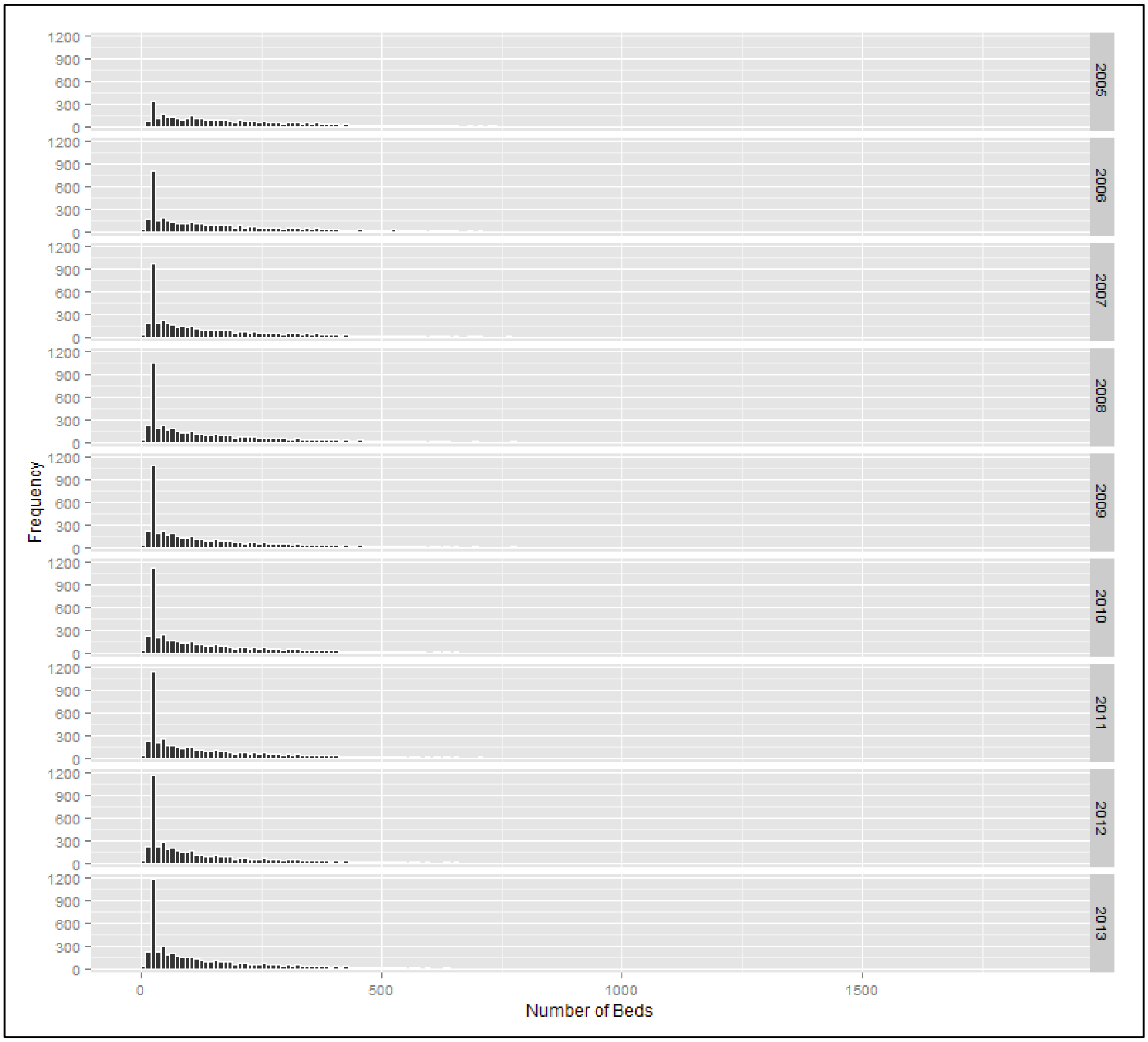
```
> summary(model.1)
Call:
coxph(formula = Surv(time, time2, basicStatus) ~ logdeg + directPer +
    logbed + Type.general + Type.critical + Type.acute + Type.other +
    Type.arehab + Type.cardio + Type.eent + Type.Geriatric +
    Type.neuro + Type.obgyn + Type.osteo + Type.ped + Type.pw +
    Type.psych + Type.wh + Type.onco + Type.op + owned + managed +
    inter, data = totalnew)

  n= 15542, number of events= 4810
   (1350 observations deleted due to missingness)
```

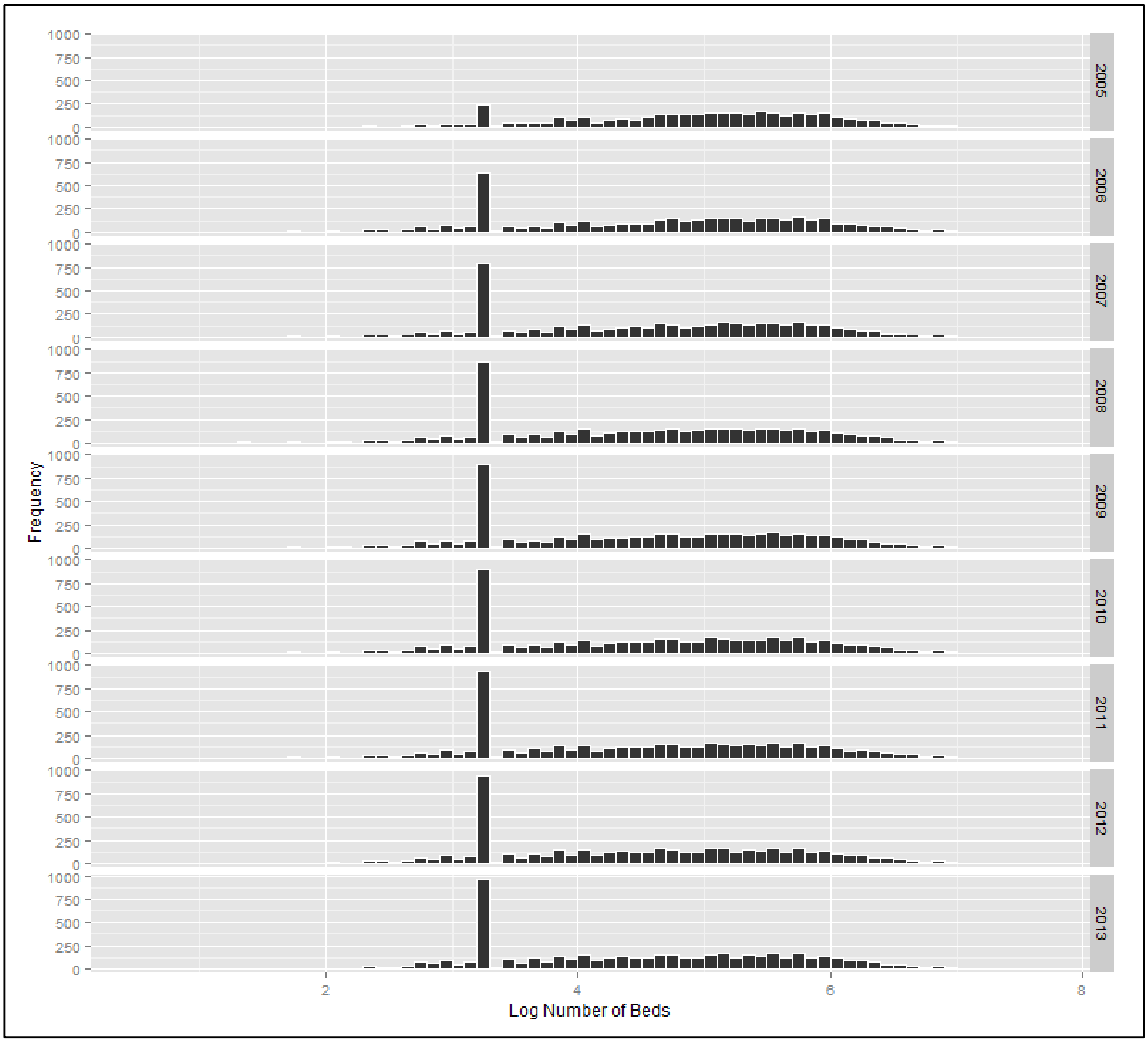
Appendix 4-5 Summary of Variable Hospital Type

	2005	2006	2007	2008	2009	2010	2011	2012	2013
Academic	313	305	299	291	293	222	219	206	210
Acute Rehabilitation	0	0	0	0	0	0	0	0	1
Cardiology	5	6	8	7	7	7	11	15	15
Critical Access	158	1523	1226	1223	1262	1273	1298	1311	1321
Eye, Ear, Nose & Throat	0	2	2	2	4	5	5	5	5
General Medical & Surgical	3013	2695	3012	3057	3059	3129	3132	3162	3167
Geriatric	2	0	0	0	0	0	0	0	0
Long Term Acute	108	125	140	245	280	319	359	369	373
Neurology	1	1	0	0	0	0	0	0	0
OB/GYN	16	0	0	0	0	0	0	0	0
Oncology	2	3	7	8	10	10	10	12	12
Ophthalmology	2	0	0	0	0	0	0	0	0
Orthopedic	3	5	8	11	14	16	20	20	22
Osteopathic	1	0	0	0	0	0	0	0	0
Other Specialty	0	7	13	26	39	50	57	164	174
Pediatric	68	70	74	77	79	86	91	93	94
Pediatric, Women's Health	0	5	5	5	5	5	6	6	7
Psychiatric	0	1	0	0	0	0	0	0	0
Women's Health	0	13	13	13	13	15	16	16	17

Appendix 4-6 Distribution of Degree Centrality, Actual Number



Appendix 4-6 Distribution of Degree Centrality, Log Format



REFERENCES

Adler-Milstein, Julia, Catherine M. DesRoches, Michael F. Furukawa, Chantal Worzala, Dustin Charles, Peter Kralovec, Samantha Stalley, and Ashish K. Jha. 2014. “More Than Half of US Hospitals Have At Least A Basic EHR, But Stage 2 Criteria Remain Challenging For Most.” *Health Affairs* 33 (9): 1664–71. doi:10.1377/hlthaff.2014.0453.

Agency for Healthcare Research and Quality. 2006. *Costs and Benefits of Health Information Technology*. Evidence Reports/Technology Assessments 132. Rockville, MD: Agency for Healthcare Research and Quality.

Alkemade, Floortje, and Carolina Castaldi. 2005. “Strategies for the Diffusion of Innovations on Social Networks.” *Computational Economics* 25 (1-2): 3–23. doi:10.1007/s10614-005-6245-1.

Allison, Paul D. 2010. “Survival Analysis.” In *The Reviewer’s Guide to Quantitative Methods in the Social Sciences*, edited by Greg Hancock and Ralph Muler, 1 edition, 413–24. New York: Routledge.

———. 2014. *Event History and Survival Analysis*. Second Edition edition. SAGE Publications, Inc.

American Hospital Association. 2007. “Continued Progress Hospital Use of Informational Technology.” Washington D.C.: American Hospital Association.

Angst, Corey M., Ritu Agarwal, V. Sambamurthy, and Ken Kelley. 2010. “Social Contagion and Information Technology Diffusion: The Adoption of Electronic Medical Records in U.S. Hospitals.” *Management Science* 56 (8): 1219–41.

Angst, Corey M., Anuj Desai, and Jennifer Wulff. 2006. “STATE-LEVEL HIT ACTIVITY AND THEMES.” CHIDS Research Briefings. Center for Health Information and Decision Systems, Robert H. Smith School of Business, University of Maryland.

Barabási, Albert-László, and Réka Albert. 1999. “Emergence of Scaling in Random Networks.” *Science* 286 (5439): 509–12. doi:10.1126/science.286.5439.509.

Bass, Frank M. 1969. “A New Product Growth for Model Consumer Durables.” *Management Science* 15 (5): 215–27.

Bastian, Mathieu, Sebastien Heymann, and Mathieu Jacomy. 2009. “Gephi: An Open Source Software for Exploring and Manipulating Networks.” In .

Bates, DW, LL Leape, DJ Cullen, and et al. 1998. “EFfect of Computerized Physician Order Entry and a Team Intervention on Prevention of Serious Medication Errors.” *JAMA* 280 (15): 1311–16. doi:10.1001/jama.280.15.1311.

Berner, Eta, Don E. Detmer, and Donald Simborg. 2005. “Will the Wave Finally Break? A Brief View of the Adoption of Electronic Medical Records in the United States.” *Journal of the American Medical Informatics Association* 12 (1): 3–7.

Bodenheimer, Thomas, and Kevin Grumbach. 2003. “Electronic Technology: A Spark to Revitalize Primary Care?” *JAMA* 290 (2): 259–64.

Bohlmann, Jonathan D., Roger J. Calantone, and Meng Zhao. 2010. “The Effects of Market Network Heterogeneity on Innovation Diffusion: An Agent-Based Modeling Approach.” *Journal of Product Innovation Management* 27 (5): 741–60. doi:10.1111/j.1540-5885.2010.00748.x.

Boonstra, Albert, and Manda Broekhuis. 2010. “Barriers to the Acceptance of Electronic Medical Records by Physicians from Systematic Review to Taxonomy and Interventions.” *BMC Health Services Research* 10 (1): 231. doi:10.1186/1472-6963-10-231.

Boulay, Marc, J. Douglas Storey, and Suruchi Sood. 2002. “Indirect Exposure to a Family Planning Mass Media Campaign in Nepal.” *Journal of Health Communication* 7 (5): 379–99. doi:10.1080/10810730290001774.

Broadhead, Robert S., Douglas D. Heckathorn, David L. Weakliem, Denise L. Anthony, Heather Madray, Robert J. Mills, and James Hughes. 1998. “Harnessing Peer Networks as an Instrument for AIDS Prevention: Results from a Peer-Driven Intervention.” *Public Health Reports (1974-)* 113 (June): 42–57.

Brown, Jacqueline Johnson, and Peter H. Reingen. 1987. “Social Ties and Word-of-Mouth Referral Behavior.” *Journal of Consumer Research* 14 (3): 350–62.

Brynjolfsson, Erik, and Lorin M. Hitt. 2000. “Beyond Computation: Information Technology, Organizational Transformation and Business Performance.” *The Journal of Economic Perspectives* 14 (4): 23–48.

Burt, R. S. 1987. “Social Contagion and Innovation: Cohesion versus Structural Equivalence.” *American Journal of Sociology* 92: 1287–1335. doi:10.1086/228667.

Cantono, Simona, and Gerald Silverberg. 2009. “A Percolation Model of Eco-Innovation Diffusion: The Relationship between Diffusion, Learning Economies and Subsidies.” *Technological Forecasting and Social Change* 76 (4): 487–96. doi:10.1016/j.techfore.2008.04.010.

Center for Medicare and Medicaid Services. 2016. “EHR Registrations By State and Program Type.” https://www.cms.gov/Regulations-and-Guidance/Legislation/EHRIncentivePrograms/Downloads/January2016_RegistrationsbyStateandProgramType.pdf.

Chandrasekaran, Deepa, and Gerard Tellis. 2007. “A Critical Review of Marketing Research on Diffusion of New Products.” In *Review of Marketing Research*, edited by N.K. Malhorta, 39–80. Armonk, NY: M.E. SHarpe.

Chassin, MR, and GW Galvin. 1998. “The Urge Need to Improve Health Care Quality.” *JAMA* 280: 1000–1005.

Choi, Hanool, Sang-Hoon Kim, and Jeho Lee. 2010. “Role of Network Structure and Network Effects in Diffusion of Innovations.” *Industrial Marketing Management*, Case Study Research in Industrial Marketing, 39 (1): 170–77. doi:10.1016/j.indmarman.2008.08.006.

Coleman, James Samuel, Elihu Katz, Herbert Menzel, and Columbia University Bureau of Applied Social Research. 1966. *Medical Innovation: A Diffusion Study*. Bobbs-Merrill Co.

Cox, D. R. 1972. “Regression Models and Life-Tables.” *Journal of the Royal Statistical Society. Series B (Methodological)* 34 (2): 187–220.

Csardi, Gabor, and Tamas Nepusz. 2006. “The Igraph Software Package for Complex Network Research.” *InterJournal* Complex Systems: 1695.

Damanpour, Fariborz. 1991. “Organizational Innovation: A Meta-Analysis of Effects of Determinants and Moderators.” *The Academy of Management Journal* 34 (3): 555–90.

Davis, Robert L, Jeffrey Wright, Francie Chalmers, Linda Levenson, Julie C Brown, Paula Lozano, and Dimitri A Christakis. 2007. “A Cluster Randomized Clinical Trial to Improve Prescribing Patterns in Ambulatory Pediatrics.” *PLoS Clinical Trials* 2 (5). doi:10.1371/journal.pctr.0020025.

Dekimpe, Marnik G., Philip M. Parker, and Miklos Sarvary. 2000. “Multimaret and Global Diffusion.” In *New Product Diffusion Models,* edited by Vijay Mahajan, Eitan Muller, and Yoram Wind, 49–73. Boston: Kluwer Academic Publishers.

Delre, Sebastiano A., Wander Jager, Tammo H. A. Bijmolt, and Marco A. Janssen. 2010. “Will It Spread or Not? The Effects of Social Influences and Network Topology on Innovation Diffusion.” *Journal of Product Innovation Management* 27 (2): 267–82. doi:10.1111/j.1540-5885.2010.00714.x.

Delre, Sebastiano A., Wander Jager, and Marco A. Janssen. 2007. “Diffusion Dynamics in Small-World Networks with Heterogeneous Consumers.” *Comput Math Organiz Theor* 13 (2): 185–202.

DeMarco, D.M., S. Kovela, L.M. Smith, J.T. Verella, G.P. Learmonth, and S.D. Patek. 2009. “Agent-Based Simulation Model for Predicting Adoption Rates of Electronic Medical Records.” In *Systems and Information Engineering Design Symposium, 2009. SIEDS ’09.*, 7–12. doi:10.1109/SIEDS.2009.5166146.

Deroian, Frederic. 2002. “Formation of Social Networks and Diffusion of Innovations.” *Research Policy* 31 (5): 835–46.

DesRoches, Catherine M., Dustin Charles, Michael F. Furukawa, Maulik S. Joshi, Peter Kralovec, Farzad Mostashari, Chantal Worzala, and Ashish K. Jha. 2013. “Adoption Of Electronic Health Records Grows Rapidly, But Fewer Than Half Of US Hospitals Had At Least A Basic System In 2012.” *Health Affairs*, July, 10.1377/hlthaff.2013.0308. doi:10.1377/hlthaff.2013.0308.

Erdos, P, and A Renyi. 1960. “On Th Evolution of Random Graphs.” *Pub Math Inst Hung Acad Sci* 5 (17): 17–61.

Ester, Martin, Hans-peter Kriegel, Jörg S, and Xiaowei Xu. 1996. “A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise.” In , 226–31. AAAI Press.

Faber, Albert, Marco Valente, and Peter Janssen. 2010. “Exploring Domestic Micro-Cogeneration in the Netherlands: An Agent-Based Demand Model for Technology Diffusion.” *Energy Policy* 38 (6): 2763–75. doi:10.1016/j.enpol.2010.01.008.

Fareed, Naleef, Gloria J. Bazzoli, Stephen S. Farnsworth Mick, and David W. Harless. 2015. “The Influence of Institutional Pressures on Hospital Electronic Health Record Presence.” *Social Science & Medicine* 133 (May): 28–35. doi:10.1016/j.socscimed.2015.03.047.

Fennell, Mary L., and Richard B. Warnecke. 1988. *The Diffusion of Medical Innovations: An Applied Network Analysis*. Plenum Press.

Ferro, Enrico, Brunella Caroleo, Marco Cantamessa, and Maurizio Leo. 2010. “Policy Incentives for Innovation Diffusion: An Agent-Based Simulation.” In *Electronic Government and the Information Systems Perspective*, edited by Kim Normann Andersen, Enrico Francesconi, Åke Grönlund, and Tom M. van Engers, 166–73. Lecture Notes in Computer Science 6267. Springer Berlin Heidelberg. http://link.springer.com/chapter/10.1007/978-3-642-15172-9_17.

Floyd, A. 1962. “Trend Forecasting: A Methodology for Figure of Merit.” In *Technological Forecasting for Industry and Government: Methods and Applications*, edited by James Rieser Bright, 95–105. New Jersey: Prentice Hall.

Fourt, Louis A., and Joseph W. Woodlock. 1960. “Early Prediction of Market Success for New Grocery Products.” *Journal of Marketing* 25 (2): 31–38.

Fox, John, and Sanford Weisberg. 2011. “Cox Proportional-Hazards Regression for Survival Data in R.” https://socserv.socsci.mcmaster.ca/jfox/Books/Companion/appendix/Appendix-Cox-Regression.pdf.

Furukawa, Michael F, T S Raghu, and Benjamin B M Shao. 2010. “Electronic Medical Records, Nurse Staffing, and Nurse-Sensitive Patient Outcomes: Evidence from California Hospitals, 1998-2007.” *Health Services Research* 45 (4): 941–62. doi:10.1111/j.1475-6773.2010.01110.x.

Furukawa, Michael F., T.S. Raghu, and Benjamin B.M. Shao. 2011. “Electronic Medical Records, Nurse Staffing, and Nurse-Sensitive Patient Outcomes: Evidence From the National Database of Nursing Quality Indicators.” *Medical Care Research and Review* 68 (3): 311–31.

Galanter, William L., Daniel B. Hier, Chiang Jao, and David Sarne. 2010. “Computerized Physician Order Entry of Medications and Clinical Decision Support Can Improve Problem List Documentation Compliance.” *International Journal of Medical Informatics* 79 (5): 332–38. doi:10.1016/j.ijmedinf.2008.05.005.

Gan, Qiwei, and Qing Cao. 2014. “Adoption of Electronic Health Record System: Multiple Theoretical Perspectives.” In *2014 47th Hawaii International Conference on System Sciences*, 0:2716–24. Los Alamitos, CA, USA: IEEE Computer Society. doi:10.1109/HICSS.2014.352.

Goldenberg, Jacob, Barak Libai, Sarit Moldovan, and Eitan Muller. 2007. “The NPV of Bad News.” *International Journal of Research in Marketing* 24 (3): 186–200.

Goldenberg, Jacob, Barak Libai, and Eitan Muller. 2010. “The Chilling Effects of Network Externalities.” *International Journal of Research in Marketing* 27 (1): 4–15. doi:10.1016/j.ijresmar.2009.06.006.

Goldenberg, Jacob, B Libai, S Solomon, N Jan, and D Stauffer. 2000. “Marketing Percolation.” *Physica A: Statistical Mechanics and Its Applications* 284 (1-4): 335–47.

Greenhalgh, Trisha, Glenn Robert, Paul Bate, Fraser Macfarlane, and Olivia Kyriakidou. 2008. *Diffusion of Innovations in Health Service Organisations: A Systematic Literature Review*. John Wiley & Sons.

Gunther, Markus, Christian Stummer, Lea M. Wakolbinger, and M. Wildpaner. 2011. “An Agent-Based Simulation Approach for the New Product Diffusion of a Novel Biomass Fuel.” *Journal of Operational Research Society* 62: 12–20.

Hall, Bronwyn H. 2004. “Innovation and Diffusion.” 10212. NBER Working Paper.

Health Resources and Services Administration. 2016. “What Is ‘Meaningful Use’?” U.S. Department of Health and Human Services. http://www.hrsa.gov/healthit/meaningfuluse/MU%20Stage1%20CQM/mu.html.

Heisey-Grove, Dawn, Lisa-Nicole Danehy, Michelle Consolazio, Kimberly Lynch, and Farzad Mostashari. 2014. “A National Study of Challenges to Electronic Health Record Adoption and Meaningful Use.” *Medical Care* 52 (2): 144–48. doi:10.1097/MLR.0000000000000038.

Hillestad, Richard, James Bigelow, Anthony Bower, Federico Girosi, Robin Meili, Richard Scoville, and Roger Taylor. 2005. “Can Electronic Medical Record Systems Transform Health Care? Potential Health Benefits, Savings, And Costs.” *Health Affairs* 24 (5): 1103–17.

HIMSS Analytics. 2015. “HIMSS Analytics Database.” http://www.himssanalytics.org/data/index.aspx.

Holroyd-Leduc, Jayna M., Diane Lorenzetti, Sharon E. Straus, Lindsay Sykes, and Hude Quan. 2011. “The Impact of the Electronic Medical Record on Structure, Process, and Outcomes within Primary Care: A Systematic Review of the Evidence.” *Journal of the American Medical Informatics Association* 18 (6): 732–37. doi:10.1136/amiajnl-2010-000019.

Horsky, Dan. 1990. “A Diffusion Model Incorporating Product Benefits, Price, Income and Information.” *Marketing Science* 9 (4): 342–65.

Jackson, Matthew O. 2010. *Social and Economic Networks*. Princeton, NJ: Princeton University Press.

James, Brent. 2005. "E-Health: Steps On The Road To Interoperability." *Health Affairs*. doi:10.1377/hlthaff.w5.26.

Janssen, Marco A., and Wander Jager. 2001. "Fashions, Habits and Changing Preferences: Simulation of Psychological Factors Affecting Market Dynamics." *Journal of Economic Psychology* 22 (6): 745–72.

Jha, Ashish K., Catherine M. DesRoches, Eric G. Campbell, Karen Donelan, Sowmya R. Rao, Timothy G. Ferris, Alexandra Shields, Sara Rosenbaum, and David Blumenthal. 2009. "Use of Electronic Health Records in U.S. Hospitals." *The New England Journal of Medicine* 360 (16): 1628–38. doi:http://dx.doi.org.mutex.gmu.edu/10.1056/NEJMsa0900592.

Jha, Ashish K., Timothy G. Ferris, Karen Donelan, Catherine M. DesRoches, Alexandra Shields, Sara Rosenbaum, and David Blumenthal. 2006. "How Common Are Electronic Health Records In the U.S.? A Summary of the Evidence." *Health Affairs* 25 (6): 496–507.

Kazley, Abby Swanson, and Yasar A. Ozcan. 2007. "Organizational and Environmental Determinants of Hospital EMR Adoption: A National Study." *J Med Syst* 31: 375–84.

Kiesling, Elmar, Markus Gunther, Christian Stummer, and Lea M. Wakolbinger. 2012. "Agent-Based Simulation of Innovation Diffusion: A Review." *Central European Journal of Operations Research* 20 (2): 183–230.

Kim, Shintae, Keeheon Lee, Jang Kyun Cho, and Chang Ouk Kim. 2011. "Agent-Based Diffusion Model for an Automobile Market with Fuzzy TOPSIS-Based Product Adoption Process." *Expert Systems With Applications* 38 (6): 7270–76.

Kumar, Sameer, and Krista Aldrich. 2010. "Overcoming Barriers to Electronic Medical Record (EMR) Implementation in the US Healthcare System: A Comparative Study." *Health Informatics* 16 (4): 306–18.

Laciana, Carlos E., and Santiago L. Rovere. 2011. "Ising-like Agent-Based Technology Diffusion Model: Adoption Patterns vs. Seeding Strategies." *Physica A: Statistical Mechanics and Its Applications* 390 (6): 1139–49. doi:10.1016/j.physa.2010.11.006.

Lee, Bruce Y., Sarah M. McGlone, Yeohan Song, Taliser R. Avery, Stephen Eubank, Chung-Chou Chang, Rachel R. Bailey, et al. 2011. "Social Network Analysis of

Patient Sharing Among Hospitals in Orange County, California." *American Journal of Public Health* 101 (4): 707–13. doi:10.2105/AJPH.2010.202754.

Luke, Douglas A., and Jenine K. Harris. 2007. "Network Analysis in Public Health: History, Methods, and Applications." *Annual Review of Public Health* 28 (1): 69–93. doi:10.1146/annurev.publhealth.28.021406.144132.

Martins, Andre C.R., Carlos de B. Pereira, and Renato Vicente. 2009. "An Opinion Dynamics Model for the Diffusion of Innovations." *Physica A: Statistical Mechanics and Its Applications* 388 (15): 3225–32.

Ma, Tieju, and Yoshiteru Nakamori. 2005. "Agent-Based Modeling on Technological Innovation as an Evolutionary Process." *European Journal of Operational Research* 166 (3): 741–55. doi:10.1016/j.ejor.2004.01.055.

McCullough, Jeffrey S. 2008. "THE ADOPTION OF HOSPITAL INFORMATION SYSTEMS." *Health Economics* 17: 649–64.

McGinn, Carrie A., Sonya Grenier, Julie Duplantie, Nicola Shaw, Claude Sicotte, Luc Mathieu, Yvan Leduc, France Légaré, and Marie-Pierre Gagnon. 2011. "Comparison of User Groups' Perspectives of Barriers and Facilitators to Implementing Electronic Health Records: A Systematic Review." *BMC Medicine* 9 (1): 46. doi:10.1186/1741-7015-9-46.

Meade, Nigel, and Towhidul Islam. 2006. "Modelling and Forecasting the Diffusion of Innovation – A 25-Year Review." *International Journal of Forecasting* 22 (3): 519–45.

Moldovan, Sarit, and Jacob Goldenberg. 2004. "Cellular Automata Modeling of Resistance to Innovations: Effects and Solutions." *Technological Forecasting & Social Change* 71 (5): 425–42.

Nan, Ning, Robert Zmud, and Emre Yetgin. 2013. "A Complex Adaptive Systems Perspective of Innovation Diffusion: An Integrated Theory and Validated Virtual Laboratory." *Computational and Mathematical Organization Theory*, 1–37.

Newman, M.E.J. 2010. *Networks: An Introduction*. 1 edition. Oxford ; New York: Oxford University Press.

Otto, Peter, and Dorit Nevo. 2013. "Electronic Health Records. A Simulation Model to Measure the Adoption Rate from Policy Interventions." *Journal of Enterprise Information Management* 26 (1-2): 165–82. doi:http://dx.doi.org.mutex.gmu.edu/10.1108/17410391311289613.

Pegoretti, Giovanni, Francesco Rentocchini, and Giuseppe Vittucci Marzetti. 2012. "An Agent-Based Model of Innovation Diffusion: Network Structure and Coexistence under Different Information Regimes." *Journal of Economic Interaction and Coordination* 7 (2): 145–65. doi:10.1007/s11403-012-0087-4.

Peres, Renana, Eitan Muller, and Vijay Mahajan. 2010. "Innovation Diffusion and New Product Growth Models: A Critical Review and Research Directions." *International Journal of Research in Marketing* 27 (2): 91–106.

Pizziferri, Lisa, Anne F. Kittler, Lynn A. Volk, Melissa M. Honour, Sameer Gupta, Samuel Wang, Tiffany Wang, Margaret Lippincott, Qi Li, and David W. Bates. 2005. "Primary Care Physician Time Utilization before and after Implementation of an Electronic Health Record: A Time-Motion Study." *Journal of Biomedical Informatics*, Human-Centered Computing in Health Information Systems. Part 2: Evaluation HCI: Evaluation, 38 (3): 176–88. doi:10.1016/j.jbi.2004.11.009.

Putsis, Williams R. Jr., and V. Srinivasan. 2000. "Estimation Techniques for Macro Diffusion Models." In *New Product Diffusion Models,* edited by Vijay Mahajan, Eitan Muller, and Yoram Wind, 263–91. Boston: Kluwer Academic Publishers.

Rahmandad, Hazhir, and John Sterman. 2008. "Heterogeneity and Network Structure in the Dynamics of Diffusion: Comparing Agent-Based and Differential Equation Models." *Management Science* 54 (5): 998–1014. doi:10.1287/mnsc.1070.0787.

Rogers, Everett M. 1962. *Diffusion of Innovation*. New York: The Free Press.

———. 2003. *Diffusion of Innovations*. 5th ed. New York, NY: Free Press.

Ryan, B., and N. C Gross. 1943. "The Diffusion of Hybrid Seed Corn in Two Iowa Communities." *Rural Sociology* 8 (1): 15–24.

Schwarz, Nina, and Andreas Ernst. 2009. "Agent-Based Modeling of the Diffusion of Environmental Innovations — An Empirical Approach." *Technological Forecasting & Social Change* 76 (4): 497–511.

Schwoon, Malte. 2006. "Simulating the Adoption of Fuel Cell Vehicles." *Journal of Evolutionary Economics* 16 (4): 435–72. doi:10.1007/s00191-006-0026-4.

Scott, W. Richard. 2008. *Institutions and Organizations : Ideas, Interests and Identities*. SAGE.

Sherer, Susan A., Chad D. Meyerhoefer, and Lizhong Peng. 2016. "Applying Institutional Theory to the Adoption of Electronic Health Records in the U.S." *Information & Management*. doi:10.1016/j.im.2016.01.002.

Smith, MF. 2004. “E-Health: Roadmap for 21st Century Health Care Consumers.” Paris: Organization for Economic Co-Operation and Development Forum 2004: Health of Nations.

Smith, M, N Milic-Frayling, B Shneierman, E Mendes Rodrigues, J Lekovec, and C Dunne. 2010. “NodeXL: A Free and Open Network Overview, Discovery and Exploration Add-in for Excel 2007/2010.” Social Media Research Foundation. http://nodexl.codeplex.com/.

Strang, David, and Sarah A. Soule. 1998. “Diffusion in Organizations and Social Movements: From Hybrid Corn to Poison Pills.” *Annual Review of Sociology* 24: 265–90.

Strang, David, and Nancy Brandon Tuma. 1993. “Spatial and Temporal Heterogeneity in Diffusion.” *American Journal of Sociology* 99 (3): 614–39.

The American Hospital Association. 2016. “Fast Facts on US Hospitals.” http://www.aha.org/research/rc/stat-studies/fast-facts.shtml.

Therneau, Terry M. 2015. “A Package for Survival Analysis in S. Version 2.38.” http://CRAN.R-project.org/package=survival.

Toole, Jameson L., Meeyoung Cha, and Marta C. González. 2012. “Modeling the Adoption of Innovations in the Presence of Geographic and Media Influences.” *PLoS ONE* 7 (1): e29528. doi:10.1371/journal.pone.0029528.

Tran, Martino. 2012. “Agent-Behaviour and Network Influence on Energy Innovation Diffusion.” *Communications in Nonlinear Science and Numerical Simulation* 17 (9): 3682–95. doi:10.1016/j.cnsns.2012.01.016.

Valente, Thomas W. 1995. *Network Models of the Diffusion of Innovations*. Cresskill, N.J: Hampton Press.

———. 2005. “Network Models and Methods for Studying the Diffusion of Innovations.” In *Models and Methods in Social Network Analysis*, edited by Peter J. Carrington, John Scott, and Stanley Wasseran. Structural Analysis in the Social Sciences. New York: Cambridge University Press.

———. 2010. *Social Networks and Health: Models, Methods, and Applications*. 1 edition. Oxford ; New York: Oxford University Press.

———. 2012. “Network Interventions.” *Science (New York, N.Y.)* 337 (6090): 49–53. doi:10.1126/science.1217330.

Valente, Thomas W., and Walter P. Saba. 2001. “Campaign Exposure and Interpersonal Communication as Factors in Contraceptive Use in Bolivia.” *Journal of Health Communication* 6 (4): 303–22. doi:10.1080/108107301317140805.

van Vliet, Oscar, Bert de Vries, André Faaij, Wim Turkenburg, and Wander Jager. 2010. “Multi-Agent Simulation of Adoption of Alternative Fuels.” *Transportation Research Part D: Transport and Environment* 15 (6): 326–42. doi:10.1016/j.trd.2010.03.006.

Wakolbinger, Lea M., Christian Stummer, and Markus Gunther. 2013. “Market Introduction and Diffusion of New Products: Recent Developments in Agent-Based Modeling.” *International Journal of Innovation and Technology Management* 10 (5): 1340015(19).

Watts, Duncan J., and Steven H. Strogatz. 1998. “Collective Dynamics of ‘small-World’ Networks.” *Nature* 393 (6684): 440–42. doi:10.1038/30918.

Wu, Shinyi, Basit Chaudhry, Jerome Wang, Margaret Maglione, Walter Mojica, Elizabeth Roth, Sally C. Morton, and Paul G. Shekelle. 2006. “Systematic Review: Impact of Health Information Technology on Quality, Efficiency, and Costs of Medical Care.” *Annals of Internal Medicine* 144 (10): 742–52.

Yu, Feliciano B., Nir Menachemi, Eta S. Berner, Jeroan J. Allison, Norman W. Weissman, and Thomas K. Houston. 2009. “Full Implementation of Computerized Physician Order Entry and Medication-Related Quality Outcomes: A Study of 3364 Hospitals.” *American Journal of Medical Quality* 24 (4): 278–86. doi:10.1177/1062860609333626.

Zhang, Ning Jackie, Binyam Seblega, Thomas Wan, Lynn Unruh, Abiy Agiro, and Li Miao. 2013. “Health Information Technology Adoption in U.S. Acute Care Hospitals.” *Journal of Medical Systems* 37 (2): 1–9. doi:10.1007/s10916-012-9907-2.

Zhang, Tao, and William J. Nuttall. 2011. “Evaluating Government’s Policies on Promoting Smart Metering Diffusion in Retail Electricity Markets via Agent-Based Simulation*.” *Journal of Product Innovation Management* 28 (2): 169–86. doi:10.1111/j.1540-5885.2011.00790.x.

BIOGRAPHY

Yinyue Hu received her Bachelor of Arts in Communications from University of Shanghai for Science and Technology in 2008 and her Master of Arts in Communication, Culture and Technology from Georgetown University in 2010. She has worked as Graduate Research Assistant at the Center for the Studies of International Medical Practice and Policy, School of Policy, Government and International Affairs, George Mason University since 2010.