Trinity University
Digital Commons @ Trinity

Computer Science Honors Theses | Computer Science Department

4-19-2013

Social Network Analysis Using a Multi-agent System: A School System Case

Lizhu Ma
Trinity University, lma@trinity.edu

Social Network Analysis Using a Multi-agent System: A School System Case

Lizhu Ma

Abstract

The quality of k-12 education has been a major concern in the nation for years. School systems, just like many other social networks, appear to have a hierarchical structure. Understanding this structure could be the key to better evaluate student performance and improve school quality. Many researches have been focusing on detecting hierarchical structure by using hierarchical clustering algorithms. Compared to existing methods, we design an interaction-based similarity measure to accomplish hierarchical clustering in order to detect hierarchical structures in social networks (e.g. school district networks). This method uses a Multi-agent System for it is based on agent interactions. With the network structure detected, we also build a model, which is inspired by the MAXQ algorithm, to decompose funding policy task into subtask and then evaluate these subtasks by using funding distribution policies from past years and looking for possible relationships between student performances and funding policies. For experiment, we use real school data from Bexar county's 15 school districts. The first result shows that our interaction based method is able to generate meaningful clustering and dendrogram for social networks. And our policy evaluation model is able to evaluate funding policies from past three years in Bexar County and conclude that increasing funding does not necessarily have a positive impact on student performance and it is generally not the case that the more spend the better.

Acknowledgments

I would like to thank my thesis advisor, Dr. Zhang for her incredible guidance and help throughout the work on this thesis.

I would like to thank the members of my thesis committee, Dr. Myers and Dr. Hibbs for their support and comments.

Finally, I would like to thank my parents for their support and encouragement.

Social Network Analysis Using a Multi-agent System: A School System Case

Lizhu Ma

A departmental senior thesis submitted to the Department of Computer Science at Trinity University in partial fulfillment of the requirements for graduation with departmental honors.

April 19, 2013

Thesis Advisor

Department Chair

Associate Vice President
for
Academic Affairs

Social Network Analysis Using a Multi-agent System: A School System Case

Lizhu Ma

Contents

List of Tables

List of Figures

Chapter 1
Introduction

1.1 Motivation

The quality of K-12 education has been a very big concern for years. Many researches have been conducted in the field. Some of them focus on school district performance. For example, Färe et al. (1989) built a model to analyze individual school district achievement. Bidwell and Kasarda (1975) examined determinants of school district organizational effectiveness and concluded that student achievement could be affected by school district organization. Some of them studied school choice (Bettinger 1999; Lubienski & Lubienski 2006). Generally, students in each school district face at least two school choices – public and private. Some also have charter school as a third option. Some researchers have studied whether and how these three types of schools affect each other. For example, Hoxby (1994) studied whether private schools provide competition for public schools. These are just some examples. There are also many others factors have been studied, such as, school size (Slate & Jones 2005), teacher quality (Rockoff 2004; Harris & Sass 2007), school/school district administrator quality (Meier et al. 2003; Clark et al. 2009), funding (Crampton 2009; Anderson 2011), etc.

As it can be seen from above, previous researches in this field mostly studied the impact of one or two of those factors on school performance. And the approaches they used are mostly statistical methods (Bohte 2002; Meier et al. 2003). Only a few have used computational simulations (Sklar et al. 2004).

1.2 Background

A social network is a set of people (or organizations or other entities) connected by a set of socially-meaningful relationships (Wellman 1996). School system, which is a set of many different actors such as students, teachers, etc., is a social network. There might be underlying community structure within a network, which is the division of network nodes into groups within which network connections are dense (Newman and Girvan 2003). Thus finding community structure is very important for understanding inherent structures for complex networks (Wakita and Tsurumi 2007).

Social network analysis has been an emerging field in recent years. It views social relationships in terms of nodes and edges (ties). Researches have shown that social networks play a critical role in determining the way problems are solved, organizations are run, etc. (SNAMAS-09 2009).

A multi-agent system (MAS) is a set of autonomous and interactive entities called agents (Guessoum et al. 2003). Multi-agent system and social network analysis share some similarities (e.g. agents, relationships, etc.). Many researches have successfully combined these two together (Grant 2009; Ma et al. 2009).

Social networks often have an underlying hierarchical structure. Thus hierarchical clustering algorithms can often be useful and appropriate methods to detect the multilevel structure of the network (Fortunato 2010).

Based on how the hierarchical dendrogram is formed, hierarchical clustering algorithms are divided into two classes: agglomerative algorithms and divisive algorithms. Agglomerative or bottom-up algorithms start with each node in its own singleton cluster, and at each step merge these clusters into larger ones until all clusters are merged into one big cluster (Schaeffer 2007).

Reinforcement learning algorithms address the problem of how an agent can learn to take actions that maximize reward while interacting directly with its environment (Dietterich 2000). In order to meet the need for large-scale and complex problems, hierarchical reinforcement learning has been introduced. Hierarchical reinforcement learning (HRL) is a technique in which reinforcement learning methods are augmented with prior knowledge about the high-level structure of behavior (Marthi et al. 2005).

1.3 Research Goals

The major purpose of this work is to study social networks with a focus on social interactions by using a multi-agent approach. There are mainly two goals:

The first one is to detect hierarchical community structure in social networks by using an agglomerative hierarchical algorithm. Existing agglomerative hierarchical algorithms usually calculate similarity or dissimilarity between two clusters by using some measure of distance between pairs of observations. We, however, develop a method that calculates similarity based on social interactions, which is ideal for social networks.

The second goal is to study how policies can possibly affect organization performance. In the school system network, how funding is distributed in school system is a very important matter; however, researches in the field have not been able to draw any sound conclusions yet. Inspired by the MAXQ method developed by Dietterich (2000), we develop a model that study funding policies from past years. We also implement school system as an example.

1.4 Our Approaches

Our goal is to study social networks (e.g. school network) and focus on agent interactions within the networks. This work mainly has two parts. The first one is to detect hierarchical community structure in social networks by using an interaction-based agglomerative hierarchical clustering algorithm. We use interaction between two agents to be the similarity measure for clustering. This algorithm has been applied to several school districts in Bexar County, and it provides satisfying results on generating the hierarchical structure of school district.

The second part is on funding policy evaluation. We study funding policies for 15 Bexar county school districts in 3 years and evaluate these policies. This model first decomposes the whole funding distribution task into several subtasks and then evaluates these subtasks separately.

This thesis has multiple contributions. First, we design a hierarchical clustering method that is suitable for interaction based social networks. Second our funding evaluation system helps to evaluate policies in a complex social network system by decomposing a complex task into several subtasks.

In addition, this research contributes to the field of Multi-agent system by showing how a multi-agent system can help in social network structure detection.

Chapter 2
Related Work

There are two problems I would like to focus. The first one is to find underlying hierarchical structure of school networks. The second one is to study funding policies in the past years and look for optimal policy. So I look into literature on social network analysis and reinforcement learning. They are presented in the two following sections.

2.1 Social Network Analysis

Social network analysis has been a fast growing field in recent years. It helps to provide explanations for social phenomena or problems, from individual creativity to corporate profitability (Borgatti et al. 2009). Many of the social networks that have been studied appear to be very complex. Examples of such include World Wide Web (Barabasi et al. 2000; Wellman 2001), citation network (Newman 2001; Rangeon et al. 2010), email exchange network (Creamer et al. 2009), etc.

Social networks are often represented by graphs, which are structures formed by a set of nodes and a set of edges that connect pairs of nodes. Nodes represent agents and edges are connections between agents. There might be underlying community structure within a network, which is the division of network nodes into groups within which network connections are dense (Newman and Girvan 2003).

Thus finding community structure is very important to understanding inherent structure for complex networks (Wakita and Tsurumi 2007).

2.1.1 Clustering

Clustering is a widely used way to detect potential structure within a network. Clustering is the process of grouping objects into a set of classes, called clusters, so that objects within a class have high similarity to each other (Jiang et al. 2004).
Graph clustering is the task of grouping the nodes into clusters in such a way that there should be more edges within each cluster than between the clusters (Schaeffer 2007). Graph clustering, therefore, can used to detect communities in a network.

Clustering algorithms can generally be divided into two categories: hierarchical and partitional. Hierarchical clustering algorithms recursively find nested clusters either in a top-down mode or a bottom-up mode. Partitional clustering algorithms find all the clusters simultaneously as a partition of the data (Jain 2009). Because the former is good at finding hierarchical structure in a network, this review will focus on it.

Hierarchical clustering algorithms produce multi-level clustering. A hierarchical clustering process is generally constructed by generating a sequence of partitions or grouping, where each subcluster belongs to one supercluster in its entity. The root cluster contains at most all of the nodes, and each of the leaf clusters contains at least one node. The process can be can be graphically represented by a tree, called dendrogram (Schaeffer 2007). The branches of a dendrogram not only record the formation of the clusters but also indicate the similarity between the clusters. An example of dendrogram is shown below:

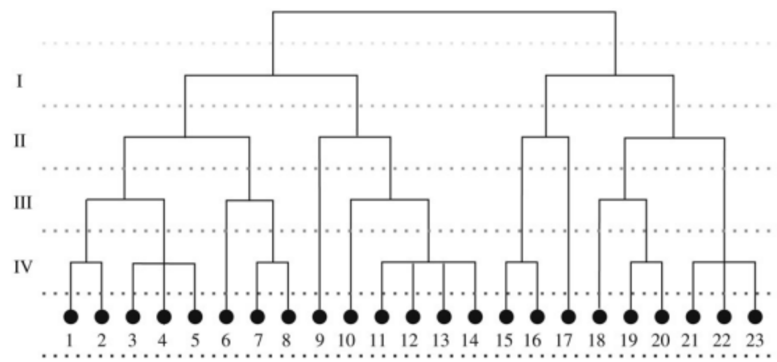

Figure 1 Dendrogram example. This is an example dendrogram that groups 23 elements into clusters (Schaeffer 2007).

Hierarchical clustering is a good way to represent communities in social network because it has several advantages. For instance, it is a discrete method that provides a partition of the nodes into clusters. The procedure is explicit and clear and there exist a wide range of programs and applications for hierarchical clustering (Wasserman and Faust 1994).

The starting point of any hierarchical clustering algorithm is to define a similarity measure between nodes. After a measure is chosen, the similarity for each pair of nodes is computed. Then at each step clusters are either merged together or split, which depends on using top-down or bottom-up method, by optimizing a certain criterion on the data set. A stopping condition may be imposed on the algorithm to select the best clustering with respect to a quality measure on the current cluster set (Schaeffer 2007; Fortunato 2009).

Based on how the hierarchical dendrogram is formed, hierarchical clustering algorithms are divided into two classes: agglomerative algorithms and divisive algorithms.

2.1.2 Agglomerative algorithms

Agglomerative or bottom-up algorithms start with each data element (node) in its own singleton cluster, and at each step merge these clusters into larger ones until all clusters are merged into one cluster.

For agglomerative approaches, different criteria of cluster similarity provide various merge strategies. They can further be divided into two kinds. The first one is that of linkage methods (e.g. single linkage, complete linkage, average linkage, etc.). They focus on calculating similarity between clusters. The second kind are methods which allow the cluster centers to be specified (e.g. as an average of the member nodes of the cluster). This kind includes the centroid, median and minimum variance methods (Schaeffer 2007).

Müllner (2011) says that, among them, the seven most common methods are single, complete, average (UPGMA), weighted (WPGMA, McQuitty), Ward, centroid (UPGMC) and median (WPGMC) linkage. Agglomerative clustering has received many attentions since the 1960s. Some recent surveys on it include (Murtagh and Contreras 2011), (Müllner 2011), and (Xu & Wunsch 2005).

2.1.3 Divisive algorithms

Divisive or top-down or algorithms start with one cluster containing all elements and split the cluster iteratively into smaller clusters.

For divisive approaches, the essential problem is to decide how to split clusters at each step. Some are based on heuristic methods such as the deterministic annealing algorithm, while many others are based on the graph theoretical methods (Jiang et al. 2004). In order to decide which cluster to be split, some combined Bi-Section k-means with divisive clustering together (Savaresi et al. 2002; Steinbach et al. 2003; Cimiano et al. 2004). They initiate Bi-Section-KMeans with the overall cluster containing all nodes. Then the cluster with the largest variance is selected and KMeans is called to split this cluster into two subclusters (Cimiano et al. 2004).

Divisive algorithms were rarely used in the past but they have becoming popular in recent year because Girvan and Newman proposed their famous divisive algorithm in 2002, which is regarded as very important and the beginning of a new era in the field of community structure detection (Fortunato 2009). Their method split clusters by removing edges that has low “betweenness”, which is a variable expressing the frequency of the participation of edges to a process (Fortunato 2009). Their method has provided some very good results on a variety of networks (Boccaletti et al. 2006).

2.1.4 Comparisons

Surveys on the comparison of the two algorithms provide detailed reviews into them. Gutierrez-Osuna’s review (2002) believes that divisive clustering has received much less attention because divisive algorithm is a computationally intensive task. The reason is that it must tentatively split all clusters to decide which cluster to be split, although it is believed that divisive clustering actually is more likely to produce meaningful results than agglomerative methods for small number of clusters (Ripley 1996).

However, the complexity of divisive clustering can be reduced if there are good ways to select which cluster to be split. Cimiano et al. (2004) compared divisive and agglomerative algorithms. They compared hierarchical agglomerative clustering algorithm and Bi-Section-KMeans as an instance of a divisive algorithm. They found that the time complexity of naive implementations of hierarchical agglomerative clustering algorithms is $O(n^3)$ where n is the number of terms. Optimized implementations can achieve $O(n^2 \log n)$. The time complexity of Bi-Section-KMeans algorithms is $O(nk)$ where n is the number of terms and k is the number of clusters.

2.2 Hierarchical Reinforcement Learning

Reinforcement learning algorithms address the problem of how an agent can learn to take actions that maximize reward while interacting directly with its environment (Dietterich 2000). In order to meet the need for large-scale and complex problems, hierarchical reinforcement learning has been introduced. Hierarchical reinforcement learning (HRL) is a technique in which reinforcement learning methods are augmented with prior knowledge about the high-level structure of behavior (Marthi et al. 2005).

Barto and Mahadevan (2003) explained HRL using the idea of a “macro-operator”, which is a sequence of operators or actions that can be invoked by name as if it were a primitive operator or action. Macros form the basis of hierarchical specifications of action sequences because a macro-operator can “call" other macros. Most of the current research on hierarchical RL focuses on action hierarchies that follow roughly the same semantics as hierarchies of macro.

Barto and Mahadevan (2003) reviewed three major approaches to hierarchical RL: the Options by Sutton et al., the hierarchies of abstract machines (HAMs) approach by Parr and Russell, and the MAXQ method by Dietterich.

2.2.1 Options

Sutton et al. (1998) formalized an approach to including activities in RL with their notion of an option, which are closed-loop policies for taking action over a period of time. Starting from a Markov decision process (MDP), a set of options defined over an MDP constitutes a semi-Markov decision process (SMDP). The simplest kind of option consists of a policy, a termination condition, and an input set. If the option is executed, then actions are selected until the option terminates. When the option terminates, the control goes back to the agent and another option can be selected. The policy learned for an option depends upon the rewards in the option’s SMDP (Uther 2002).

2.2.2 Hierarchies of Abstract Machines (HAMs)

Parr (1998) developed an approach to hierarchically structuring MDP policies called Hierarchies of Abstract Machines or HAMs. The emphasis is on simplifying complex MDPs by restricting the class of realizable policies rather than expanding the action choices. The idea of the HAM approach is that with the MDP, a user provides a series of state machines. These state machines can refer to each other, and hence form a hierarchy. The state machines partially specify a policy. Only policies consistent with the state machines are allowed. Then the original MDP can be turned into a new and smaller MDP, which can be solved using traditional methods (Uther 2002).

2.2.3 MAXQ

Dietterich (2000) developed another approach of hierarchical RL called the MAXQ Value Function Decomposition, or MAXQ. Unlike options and HAMs, the MAXQ approach does not rely directly on reducing the entire problem to a single SMDP. Instead, a hierarchy of SMDPs is created whose solutions can be learned simultaneously. The MAXQ approach starts with a decomposition of a core MDP M into a set of subtasks. All the tasks form a task graph hierarchically. As the task graph decomposes the action space of the problem, the MAXQ graph decomposes the value function of the problem (Uther 2002). Dietterich also proposed two ways to achieve optimal rewards. The first one is that a hierarchically optimal policy is a policy that achieves the highest cumulative reward among all policies consistent with the given hierarchy. The second is a recursively optimal policy that for each subtask M_i, the corresponding policy is optimal for the SMDP defined by the set of states, the set of actions, the state transition probability function, and the reward function.

2.2.4 Other Approaches

Andre (2002) developed ALisp, which is a language based on Lisp to write partial program that coupled with an environment results in a semi-Markov decision process over the joint choice states, and finding the optimal policy in this SMDP is equivalent to finding the optimal completion of the partial program in the original MDP (Andre 2002; Marthi et al. 2005). Several other approaches have been developed in recent years. For instance, Hengst (2002), inspired by MAXQ, developed HEXQ, which is an algorithm that automatically attempts to decompose and solve a model-free factored MDP hierarchically. Dethlefs and Cuayahuitl (2011) combined hierarchical reinforcement learning and Bayesian networks together for natural language generation in situated dialogue. Cao and Ray (2012) incorporated Bayesian priors in the MAXQ framework for hierarchical reinforcement learning (HRL).

Chapter 3
The Approach

Community detection in social network analysis has attracted many attentions in recent years. The idea is to divide network nodes into groups within which the network connections are dense, but between which they are sparse (Newman & Girvan 2003). Because social networks are usually represented by graphs, community defection algorithms for graphs can often be applied to social networks. Many algorithms have been introduced. For instance, there are modularity-based methods (e.g. modularity optimization), clustering methods (e.g. partitional clustering), dynamic algorithms (e.g. random walk), etc. (Fortunato 2009). Among them, hierarchal clustering has been regarded as an effective way to detect community structure in social networks because social networks often have a hierarchical structure (Fortunato 2009). Therefore hierarchical clustering algorithms can reveal the multilevel structure of the graphs.

The major method for hierarchical clustering is the agglomerative approach (bottom-up) (Fortunato 2009). The basic idea of agglomerative algorithm is that it iteratively merges the two nodes or clusters with highest similarity, until there is only one big cluster left. So at the end of the process, the root cluster contains at most all of the nodes, and each of the leaf clusters contains at least one node. The process can be can be graphically represented by a tree, called dendrogram (Schaeffer 2007).

Most existing algorithms either use real physical distance or some shared property between two nodes to calculate similarity. Here I present an agglomerative clustering algorithm called interaction-based similarity measure clustering algorithm and introduce a method that use social interaction to calculate similarity between pairs of nodes or clusters.

3.1 Overall Algorithm

The overall algorithm for the model is that it first takes the whole network as the input. Then by using the hierarchical clustering algorithm, the hierarchical structure of the network is detected. Once the hierarchical structure is obtained, the funding evaluation algorithm is called to decompose funding distribution task into subtasks and study policies for each subtask by calculating reward for each policy.

Algorithm 1 Hierarchical Structure Detection and Analysis.

```
function run ()
        s = similarity_matrix (number_of_agents, interactions)
        structure = hierarchical_clustering (s)
        policy_evaluation (policies, tasks, student_performance)
end
```

These two functions (hierarchical_clustering and policy_evaluation) will be introduced in 3.2 and 3.3.

3.2 Interaction-based Similarity Measure Algorithm

3.2.1Input

Suppose there is a set of S of N nodes. The input to the algorithm can be defined as a similarity matrix (Müllner 2011; Day 1984).

Definition 1 Similarity Matrix. A $d \times d$ *similarity matrix* D on a set S is $d(a, b)$ where $a, b \in S$. $d(a, b)$ is a quantitative measure of the similarity between two nodes a and b. $d(a, b)$ is both reflexive and symmetric, i.e. we have $d(a, a) = 0$ and $d(a, b) = d(b, a)$ for all $a, b \in S$.

If the set S has N nodes, there should be $\binom{N}{2}$ pairwise similarities.

3.2.2 Similarity Measure

The network is essentially built up by a multi-agent system, where agents interact with each other and these interactions have different levels because some agents interact more closely while others have more loose interactions. So first we define what interaction is in a social network:

The first step of the interaction-based similarity measure algorithm is to calculate similarity between each pair of nodes. This similarity measure depends on specific problem. For instance, in biological studies, it could be real physical distance between two nodes. In film actors' collaboration networks, the similarity could be calculated by how many films actors have appeared in together (Marchiori & Latorav 2000).

In interaction-based social networks, it would be different because there is no real physical distance measure in the system.

Therefore we propose a way to define similarity $d(a, b)$. Because nodes in our system are agents and agents have interaction with each other, we use their interaction level to be the similarity measure. Agents (nodes) that share more frequent interaction have higher similarity while agents that share less frequent interaction have lower similarity. For instance, in our school system, student-teacher pair should have higher similarity than student-principal pair.

Definition 2 Social Interaction (Znaniecki 1965). Social interaction between two agents A and B occurs usually when (1), A initiates it by performing a social action intended to influence B; then B 'reacts', i.e., performs an action in consequence of A's action. Or when (2) each of them independently starts a social action bearing upon the other, and each reacts to the other's action.

Definition 3 Similarity Measure. The similarity $d(a, b)$ between nodes a and b is the interaction level between a and b.

3.2.3 Output

The output of a hierarchical clustering procedure is a dendrogram.

Definition 4 Dendrogram (Müllner 2011). Given a finite set S_0 (initial set) with cardinality $N = |S_0|$, a dendrogram is a list of $N - 1$ triples (a_i, b_i, δ_i), $i = 0, \ldots, N - 2$, such that $\delta_i \in [0, \infty)$ is the distance between a_i and b_i and $a_i, b_i \in S_i$, where S_{i+1} is recursively defined as $(S_i \setminus \{a_i, b_i\}) \cup n_i$, and $n_i \notin (S_i \setminus \{a_i, b_i\})$ is a label for a new cluster.

The set S_0 are the initial data points. In each step, n_i is the new cluster which is formed by joining the clusters $a_i$and $b_i$at the distance δ_i. The order of the clusters within each pair (a_i, b_i) does not matter. "$\setminus$" represents relative complement. $S_i \setminus \{a_i, b_i\}$ is the set of elements in S_i but not in $\{a_i, b_i\}$. The procedure contains $N - 1$ steps, so that the final state is a single cluster which contains all N initial nodes.

The dendrogram represents a recursive procedure where at each step a new cluster n_i is formed from two clusters a_i and b_i based on their distance δ_i. In every step, a new cluster is added to the set and the previous two individual nodes that formed the cluster are eliminated from the set.

3.2.4 Algorithm

Algorithm 2 Interaction-based Similarity Measure Algorithm.

1. Initially each node $a, b, \ldots$ is in its own cluster $a_i, b_i, \ldots$, where $i = 0, \ldots, N - 2$, and N is the size of S.

2. Iteratively merge the two clusters with highest similarity d, say a_i and b_i, until there is only one cluster left.

After the step that nodes are clustered into their first cluster, which uses node similarity d, we deal with clusters directly. We use average linkage, which is the average similarity between all pairs of nodes in the two clusters, to calculate cluster similarity:

$$\delta(a_i, b_i) = \frac{\sum a \in a_i, \sum b \in b_i,\ d(a, b)}{|a_i| * |b_i|} \quad (1)$$

where a_i and b_i are the cluster pair and a and b are nodes within them.

3.3 Funding Evaluation

Some of the problems of reinforcement learning tend to be very large in scale. So the hierarchical approach to reinforcement learning has been developed to decompose complex problems by using temporal abstraction and hierarchical control (Barto & Mahadevan 2003). Among hierarchical reinforcement learning algorithms, the MAXQ approach is considered one of the most effective methods (Mirzazadeh et al. 2007). Here I present the MAXQ approach as a way to decompose complex reinforcement learning problems and the MAXQ-Q approach as a learning algorithm (Dietterich 2000).

3.3.1 MAXQ Method

> **Definition 5 MAXQ Task**. The overall MAXQ task is to solve a Markov Decision Problem (MDP) M, which is defined over a set of states S and actions A with reward function $R(s'|s,a)$ (the reward received upon entering state s' after performing action a in state s) and transition probability function $P(s'|s,a)$ (the probability of entering state s' as a result of performing a in s).

The basic idea of the MAXQ approach is that it decomposes the whole task into a set of subtasks, which may further be decomposed into smaller subtasks (Mirzazadeh et al. 2007). This structure forms a hierarchy tree whose leaves are primitive actions. Primitive actions return the rewards for performing the actions. Each subtask has some termination conditions, which are the conditions that once fulfilled the control of program returns to the parent subtask.

We say that hierarchical policy for a MAXQ graph is a set of policies $\pi = (\pi_0, \pi_1, \ldots, \pi_n)$, one for each subtask.

As the action space of the problem is decomposed by the task graph, we decompose the action-value function $Q\ (p, s, a)$, i.e. the total expected reward of performing action a in subtask p and then following the hierarchical policy $\pi = (\pi_0, \pi_1, \ldots, \pi_n)$ into two components (Mirzazadeh et al. 2007).

The first component $V\ (a, s)$ is the expected total reward received while executing action a in state s, and the second component, which is called as the completion function, $C\ (p, s, a)$ is the expected total reward of completing parent task p following policy π after a has returned. Thus we have the MAXQ decomposition equations:

$$Q^{\pi}(p, s, a) = V^{\pi}(a, s) + C^{\pi}(p, s, a) \qquad (2)$$

where $V^{\pi}(a, s) = \begin{cases} Q^{\pi}(a, s, \pi_i(s')) & \textit{if } i \textit{ is composite subtask} \\ \sum_{s'} P(s'|s, i)\, R(s'|s, i) & \textit{if } i \textit{ is a primitive action} \end{cases}$

and $C^{\pi}(p, s, a) = \sum_{s',\ n} P(s', n\,|s, a)\gamma^n Q^{\pi}(p, s', \pi(s'))$.

Note: n here stands for time steps and γ is the time discount factor which determines the importance of future rewards. s' is the new state that is entered after performing action a instate s. $V(a, s)$ is the expected total reward received while executing action a in state s. The completion function $C(p, s, a)$ is the expected total reward of completing parent task p following policy π after a has returned.

3.3.2 Policy Evaluation Model

The MAXQ algorithm is further developed into MAXQ-Q by combining Q-learning together. However, because of our current limitation on data, I am not able to implement a learning algorithm in the model.

Instead, I build a funding policy evaluation system by using the idea of MAXQ decomposition.

Funding is usually distributed into several areas therefore the model breaks the major task into several subtasks. Then each subtask can be accomplished by several primitive actions. Subtasks could be different funding distribution areas, which can be accomplished by different primitive actions (distribution methods). A reward that is evaluated by student performance is returned to each finished action. Then once different funding policies are inputted into the model, it will evaluate them, compare the rewards and finally return the one with highest reward.

Chapter 4
Experiments and Results

In order to test the model presented in Chapter 3, I implement a school system network in the experiment by using real school data provided by Texas Education Agency[1].

The quality of k-12 education has been a big concern in the nation for years. There are many factors that may affect school performance, such as school size (Slate & Jones 2005), school choice (public, private, or charter) (Hoxby 1994; Bettinger 1999; Lubienski & Lubienski 2006), teacher quality (Rockoff 2004; Harris & Sass 2007), school/school district administrator quality (Meier et al. 2003; Clark et al. 2009), funding (Crampton 2009; Anderson 2011), etc. Previous researches in this field mostly studied the impact of one or two of those factors on school performance. And the approaches they used are mostly statistical methods (Bohte 2002; Meier et al. 2003). Only a few have used computational simulations (Sklar et al. 2004). Ours differs from previous ones because ours observe the emergence of school system performance based on a comprehensive list of agents and the interactions between them.

School funding is a very important matter surrounding education. There is conflicting evidence on whether or not an increase in school funding will truly

[1] http://ritter.tea.state.tx.us/perfreport/snapshot/

produce an increase in student achievement or whether it is not the amount of funding but how funding is spent that is truly important.

The general understanding is that schools are underfunded and require more funding in order to function properly. Crampton (2009) indicates this when he states that investment in the human, social, and physical capital of a school can have a strong positive impact on student achievement. However, Anderson (2011) thinks that the effects of increasing per-pupil spending on student achievement are not necessarily positive. Although school funding has been steadily increasing over the years since 1970, American schools are still falling behind schools in other countries. There are many concerns these days on school quality. For one, there are concerns that money is not being spent in the right ways that schools really need. Also, people have been wondering whether funding is a key feature issue in the failing education system or not (Anderson 2011).

Therefore our system is designed to be a decision aid system that analyzes funding distribution records and their relations to student performance in the past year based on a hierarchical decomposition.

4.1School System Structure

Most of the school districts in the United States are composed of public schools, private schools, and charter schools. For each school district, there is a school board, which appoints a superintendent. Each school has its own principal and teachers.

We designed a multi-agent system to model school system. Our system includes many actors/agents such as students, teachers, principals, superintendent, etc. Our method is different from existing ones because ours observe the emergence of school system performance based on a comprehensive list of agents and the interactions between them.

The data is obtained from the Texas Education Agency. Bexar County is the geographical area we focus on. According to the snapshot report provided by the Texas Education Agency, there are 15 public school districts in Bexar as of 2011. Bexar County is used because it is a representative area. It consists of mixed type school districts. In the snap shot of year 2011, among the 15 school districts, 11 of them were rated as "academically acceptable" and the rest 4 were rated as "recognized".

An abstract representation of one school can be presented as the following graph:

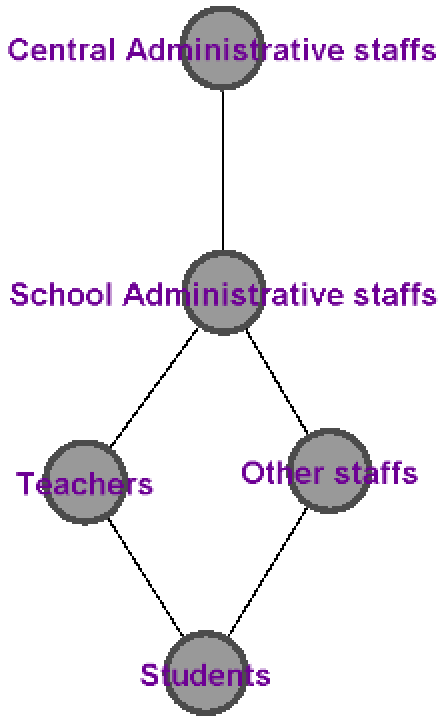

Figure 2 School example.

Note: this graph is generated by Gephi[2], an open source graph visualization and manipulation software. This graph only presents one school as an example of the organization because of space limit. These should be other schools also connected to the central administrative staff in the graph.

Each node in the graph represents one type of agents. Edges represent interaction between nodes. Students are the largest group in the school system. They interact with each other. They also learn from the teachers. Teachers teach students and report to school administrative staffs (e.g. principal). School administrative staffs report to central administrative staffs (e.g. superintendent).

The simulation is based on individual agents. So here I also present a graph of a sample school district, where there are 10 students, 2 teachers, 1 other staff, 1

[2] https://gephi.org/

central administrative staff, and 1 school administrative staff. The graph is generated by a free and open-source application called NodeXL[3].

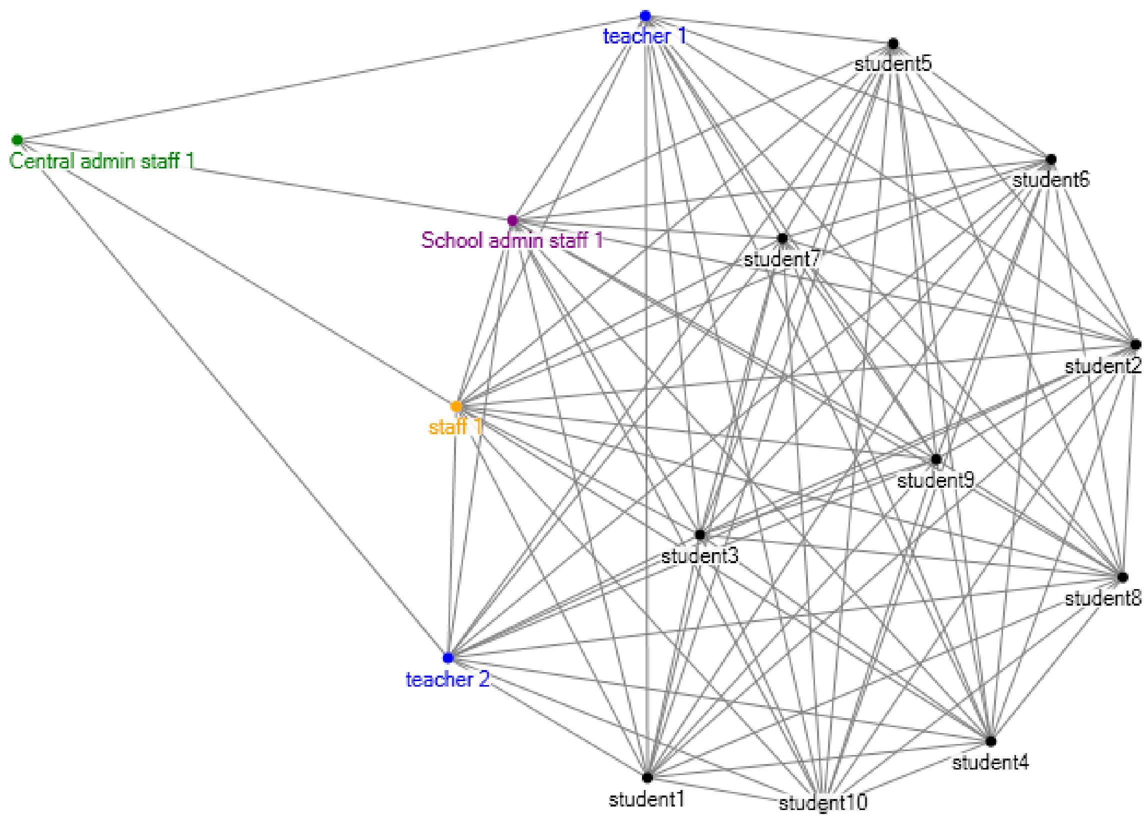

Figure 3 School district agents and interactions figure.

The following table shows the number of students, teachers, central administrative staffs, school administrative staffs and other staffs in the 15 school districts as of October 29, 2010[4].

---

3 http://nodexl.codeplex.com/

4 http://ritter.tea.state.tx.us/perfreport/snapshot/2011/itemdef.html

DISTRICT NAME	# of schools	Total students	Total teachers	Total Central Admin	Total School Admin	Total other staffs	Total agents
Alamo Heights	6	4744	335	6.04	12.08	253.68	5350.8
Harlandale	30	14846	970	20.86	62.58	1022.14	16921.58
Edgewood	22	11904	788	0	34.82	905.32	13632.14
Randolph Field	3	1167	86	5.04	3.36	72.24	1333.64
San Antonio	99	54894	3437	76.31	152.62	3968.12	62528.05
South San Antoni	18	9860	687	29.36	44.04	719.32	11339.72
Somerset	7	3779	258	5.89	17.67	318.06	4378.62
North East	73	66364	4377	0	171.14	3936.22	74848.36
East Central	15	9571	573	12.06	24.12	603	10783.18
Southwest	14	11779	735	15.02	45.06	720.96	13295.04
Lackland	2	985	80	1.68	3.36	80.64	1150.68
Ft Sam Houston	2	1427	118	2.31	6.93	103.95	1658.19
Northside	107	94632	6117	0	255.5	6387.5	107392
Judson	30	22016	1437	0	89.37	1459.71	25002.08
Southside	9	5310	375	7.49	22.47	352.03	6066.99

Table 1 School district profiles 2011.

4.2 Experiment Settings I: clustering

The experiment is composed of two parts: hierarchical structure detection and funding policy evaluation. In the first part, I use a hierarchical clustering algorithm to detect hierarchical structure in the network. Then in the second part I use the structure detected to help to evaluate previous funding policies.

As shown in Figure 3, the system has many agents connecting with each other. However the structure of the network cannot be seen directly from the graph visualization, so it needs to be found. Therefore here I use the agglomerative hierarchical clustering algorithm presented in Chapter 3 to detect underlying hierarchical community structure in the school system.

Students and teachers are further divided into 10 groups: regular education students, special education students, ESL education students, career education students, gifted education students, regular education teachers, special education teachers, ESL education teachers, career education teachers, and gifted education teachers. The following table shows the number for each group as of October 29, 2010:

DISTRICT NAME	TOTAL STUDENTS	Special student	ESL student	Career student	Gifted student	Regular student	TOTAL TEACHER	Regular teacher	Special teacher	ESL teacher	Career teacher	Gifted teacher
Alamo Heights	4744	284.64	237.2	284.64	759.04	3178.48	335	234.5	30.15	16.75	3.35	33.5
Harlandale	14846	1336.14	2375.36	3711.5	890.76	6532.24	970	727.5	116.4	58.2	38.8	29.1
Edgewood	11904	1190.4	2261.76	2261.76	952.32	5237.76	788	543.72	94.56	94.56	23.64	23.64
Randolph Field	1167	81.69	0	93.36	93.36	898.59	86	72.24	6.02	0	0.86	4.3
San Antonio	54894	5489.4	8783.04	10978.8	3293.64	26349.12	3437	2199.68	378.07	343.7	137.48	240.59
South San Antoni	9860	788.8	1479	1972	394.4	5225.8	687	377.85	68.7	116.79	27.48	68.7
Somerset	3779	340.11	377.9	718.01	151.16	2191.82	258	196.08	25.8	20.64	10.32	5.16
North East	66364	6636.4	5972.76	9954.6	4645.48	39154.76	4377	3107.67	437.7	218.85	131.31	393.93
East Central	9571	1052.81	861.39	1627.07	574.26	5455.47	573	338.07	63.03	40.11	17.19	40.11
Southwest	11779	1295.69	1413.48	2473.59	471.16	6125.08	735	536.55	73.5	73.5	29.4	22.05
Lackland	985	98.5	19.7	157.6	49.25	659.95	80	61.6	6.4	0.8	3.2	0.8
Ft Sam Houston	1427	171.24	57.08	199.78	85.62	913.28	118	92.04	12.98	2.36	2.36	5.9
Northside	94632	11355.84	6624.24	17033.76	9463.2	50154.96	6117	3914.88	795.21	489.36	183.51	428.19
Judson	22016	2201.6	1541.12	5504	1761.28	11008	1437	1120.86	129.33	28.74	71.85	57.48
Southside	5310	531	637.2	1274.4	318.6	2548.8	375	258.75	41.25	18.75	15	18.75

Table 2 Student and teacher groups data.

Note: for “special student”, “ESL student”, “career student”, and “gifted student”, the number at each cell represents the number of students who receives this kind of education. All students should receive regular education. The number of “regular education” is intended to represent the number of students who only received regular education, and it is calculated by total number of students minus number of students who receive non-regular education. However, because some students might receive more than one kind of non-regular education, there should exist some inaccuracies on the number.

Before we start, similarity between each pair of nodes must be calculated. Because nodes in the system are agents, and agents have interactions with each other, the similarity measure is set to be their interaction levels. The higher the interaction level, the higher the similarity is. We define the interaction level as a function of the time length of agents’ interaction:

$interaction\_level = f(time\_length\,)$ (3)

So the more time the two agents spend on interaction, the higher the interaction level is.

The following table represents interaction levels between all kinds of agents except interactions within and between students and teachers, which are shown in the next table. Interaction levels are represented by a scale from 0 – 9. 0 means lowest interaction, while 9 means highest interaction level.

	Students	Teachers	Other staffs	School admin staffs	Central admin staffs
Students	/	/	5	2	0
Teachers	/	/	5	4	2
Other staffs	5	5	9	4	2
School admin staffs	2	4	4	9	8
Central admin staffs	0	2	2	8	9

Table 3 Interaction levels (1).

Central and school administrative staffs generally do not interact with students directly very often, but they can still have an effect on student performance and school quality (Meier et al. 2003; Clark et al. 2009). However, they should work with each other and teachers more directly (through recruiting, training, and rewarding high-quality principals and teachers) (Meier et al. 2003).

The following table shows interaction within and between students and teachers. All students should receive regular education from regular education teachers, so they interact with each other frequently. There are also 4 other kinds of education: special, ESL, gifted, and career education. Students who also receive these 4 kinds of education interact with teachers of these 4 kinds of education directly.

	res	ses	ees	ges	ces	ret	set	eet	get	cet
res	9	7	7	7	7	9	6	6	6	6
ses	7	9	7	7	7	8	9	6	6	6
ees	7	7	9	7	7	8	6	9	6	6
ges	7	7	7	9	9	8	6	6	9	6
ces	7	7	7	7	9	8	6	6	6	9
ret	9	8	8	8	8	9	7	7	7	7
set	6	9	6	6	6	7	9	7	7	7
eet	6	6	9	6	6	7	7	9	7	7
get	6	6	6	9	6	7	7	7	9	7
cet	6	6	6	6	9	7	7	7	7	9

Table 4 Interaction levels (2). Note: res - regular education student, sed - special education student, ees - esl education student, ges - gifted education student, ces - career education student, ret - regular education teacher, set - special education teacher, eet - esl education teacher, get - gifted education teacher, and cet - career education teacher.

This table is implemented as the input matrix for the clustering algorithm. The output is a dendrogram that shows out the underlying structure. I write the simulation code in Python and use the *fastcluster* library and its interface to Python to accomplish the clustering process because it has proven to be performing well in terms of complexity[5]. Because *fastcluster* uses dissimilarity rather than similarity between nodes as the input, the matrix for the simulation program is calculated by $1 - d$. d is the similarity measure presented in Table 3 and 4.

4.3 Experiment Settings II: funding policy evaluation

The second setting part of this chapter is on evaluating funding policy. The snapshot report provides data on annual funding expenditure for regular education, special education, bilingual education, career and technical education, and gifted education.

[5] http://math.stanford.edu/~muellner/fastcluster.html

The following figure represents the funding allocation:

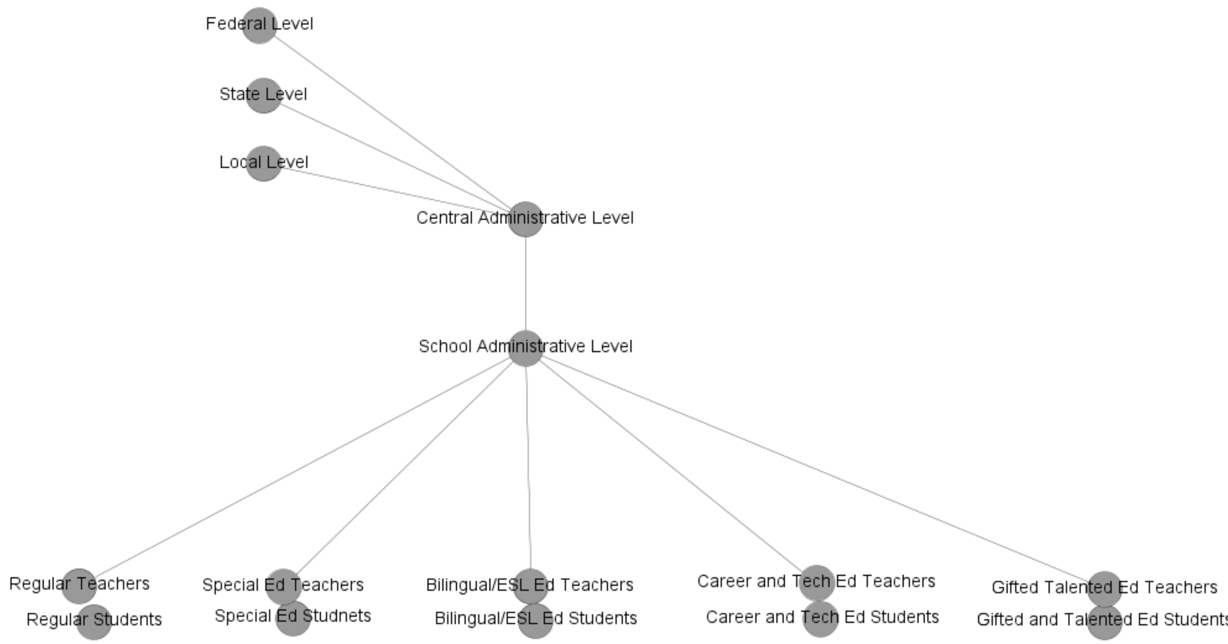

Figure 4 Funding allocation.

Inspired by the hierarchical reinforcement learning algorithm presented in chapter 3, we design a model that evaluates funding policies from past years.

The MAXQ algorithm introduces the idea of task decomposition. The root task is divided into subtasks. Then each subtask can be accomplished by several primitive actions.

In our experiment, the task is decomposed into several subtasks, which are regular education area, special education area, bilingual education area, career and technical education area, and gifted education area. "Primitive actions" are the different funding distribution methods. After each subtask chooses an action, the reward for that action is returned. Each subtask will iterate over all actions and rewards will be compared in the end and the action that returns highest reward will be chosen.

It should be noted that not every school district has all five subtasks. The following is the task decomposition graph:

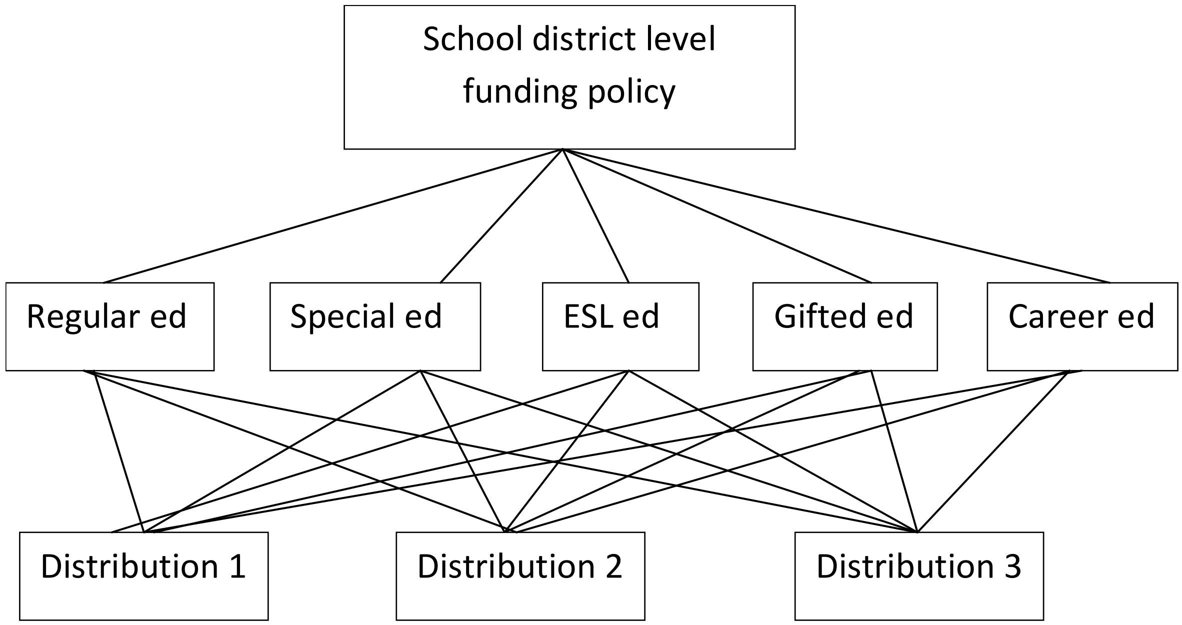

Figure 5 Task decomposition.

The reward that is returned for each action is a function of student performance and agent interaction quality.

$$r = student\_performance * \frac{1-\alpha^{\overline{f(time\_length)}+1}}{1-\alpha} \quad (4)$$

The student performance is evaluated by the TAKS tests passing out rate each year as shown in Table 5. $\frac{1-\alpha^{\overline{f(time\_length)}+1}}{1-\alpha}$ is inspired by Joseph et al. (2013) and Du et al. (2009), which represents the limit in the amount of capital gained by agents on an interaction. α is set to 0.8 in our experiment. And $\overline{f(time\_length)}$ is the average interaction level between agents involved in this subtask.

The following table shows TAKS tests passing out rate for students enrolled as of October 29, 2010 and specific amount of money spent on each area for the year 2009-2010:

DISTRICT NAME	TAKS:	TOTAL EXPENDITURE	Regular educati	Special educa	ESL education	Career educatic	Gifted educatio
Alamo Heights	87	43715556	29726578.08	6994488.96	0	437155.56	874311.12
Harlandale	71	127324400	68755176	16552172	8912708	3819732	0
Edgewood	59	111248406	50061782.7	13349808.72	7787388.42	3337452.18	0
Randolph Field	91	12701043	9525782.25	1397114.73	0	127010.43	508041.72
San Antonio	62	504207177	226893229.7	90757291.86	40336574.16	15126215.31	5042071.77
South San Antoni	63	93898291	41315248.04	14084743.65	11267794.92	3755931.64	1877965.82
Somerset	65	33487386	15739071.42	4688234.04	1674369.3	1339495.44	669747.72
North East	81	559721134	347027103.1	123138649.5	5597211.34	16791634.02	5597211.34
East Central	70	74311079	41614204.24	14862215.8	2229332.37	2229332.37	1486221.58
Southwest	67	99087846	59452707.6	14863176.9	990878.46	3963513.84	0
Lackland	87	11806204	7674032.6	2361240.8	118062.04	236124.08	118062.04
Ft Sam Houston	75	19322253	11013684.21	6376343.49	0	386445.06	193222.53
Northside	80	761630471	472210892	167558703.6	0	22848914.13	7616304.71
Judson	69	174658801	104795280.6	33185172.19	5239764.03	6986352.04	0
Southside	68	46849510	26704220.7	6090436.3	2810970.6	1873980.4	0

Table 5 Funding data for the year of 2010. TAKS is the total number of students who passed all the TAKS tests they attempted expressed as a percentage of the total number of students who took one or more tests.

Data for 2008 and 2009 can be found in Appendix A.

4.4 Results

4.4.1 Hierarchical Clustering

The following is the output dendrogram for Lackland school district. I choose this one to present here because it is one of the best-performing school districts in Bexar County, which will be discussed in the next section. The dendrogram is generated by Python’s matplotlib library.

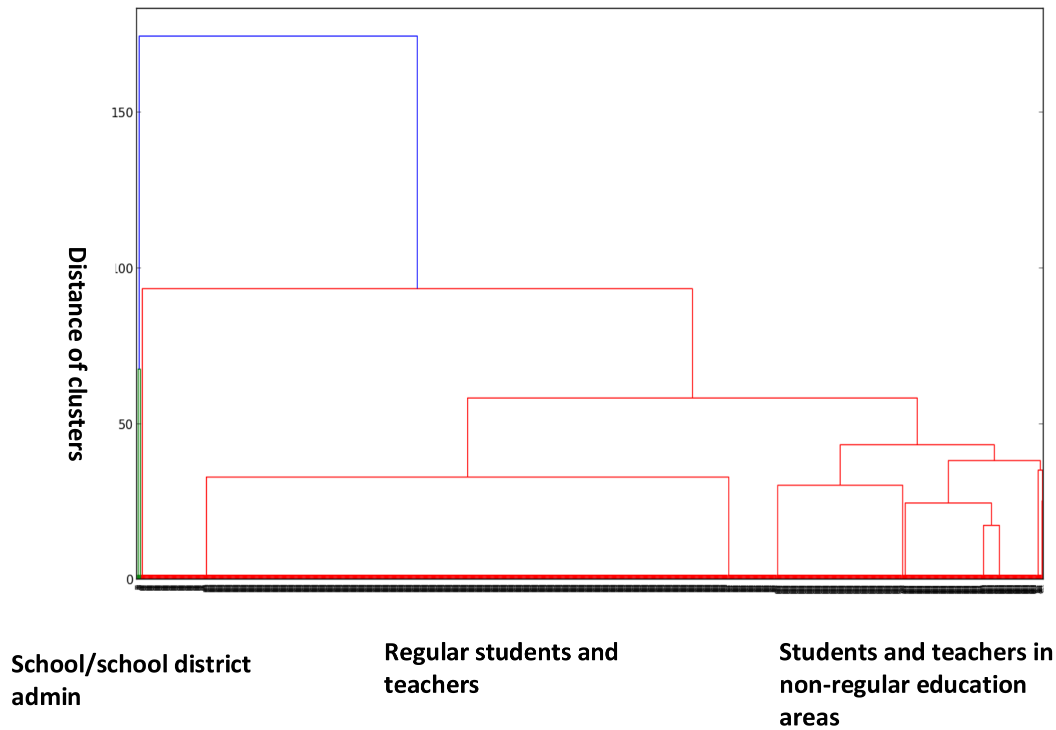

Figure 6 Lackland (1). The number on the y axis shows the distance of the two clusters that are formed together.

Students and teachers are first clustered into each subgroups. Then it is clustered together with other staffs (these are shown in red). School and school district staffs are clustered together as shown in green on the left. Then they are clustered into one final cluster as shown in blue. Because of the large size of data, it is not easy to tell how the individual cluster looks like. Therefore here I present another dendrogram generated for Lackland school district whose agent numbers are all divided by 10 for the sake of display (after the dividing, for all the numbers between 0-1, they will be rounded up to 1 rather than 0):

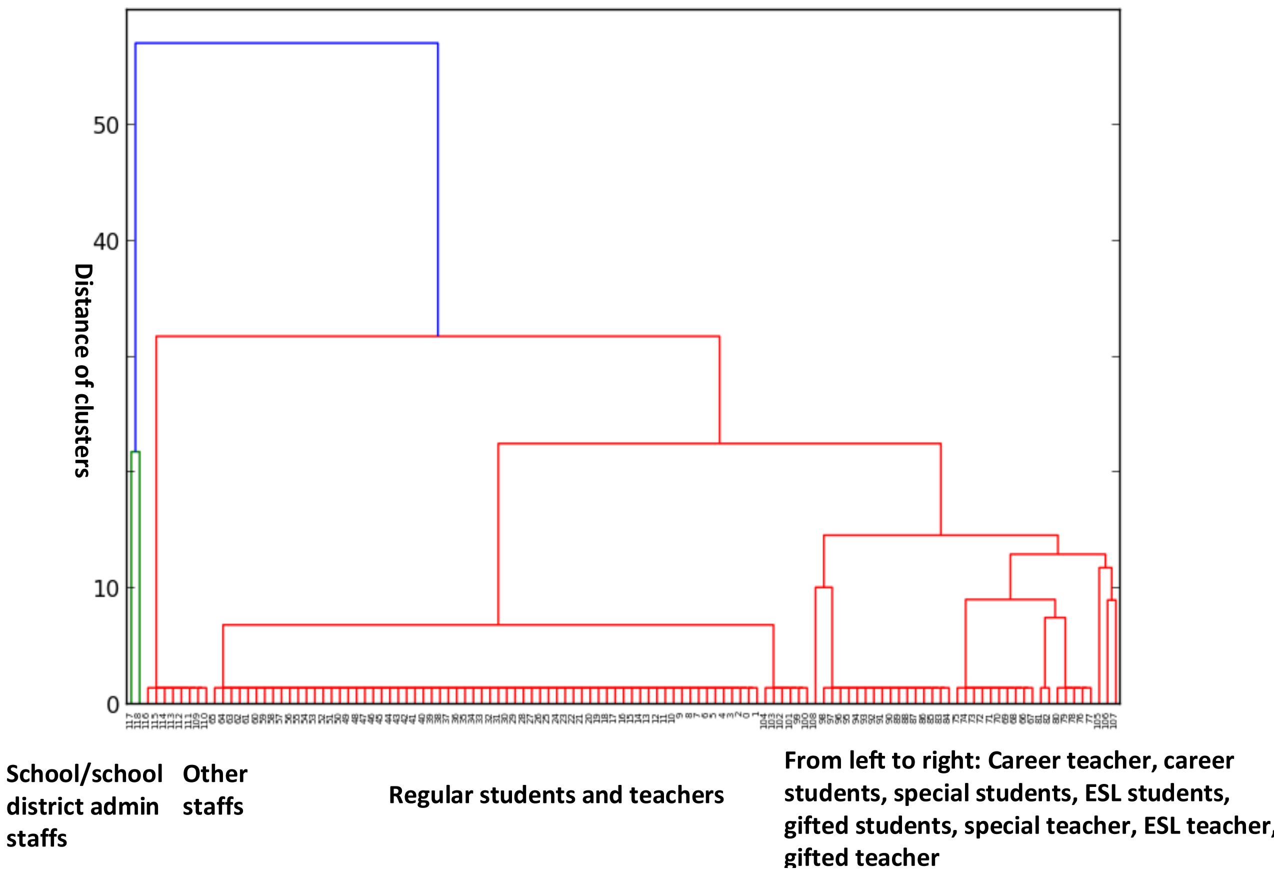

Figure 7 Lackland (2). Note: each group of nodes has its description underneath it.

As the graph shows, agents are first clustered into its own groups. Then each student group is clustered together with its corresponding teacher group. Following that, all students and teachers are formed into on cluster. Then it is clustered with other staffs. These clustering processes are in color red. School and school district staffs are clustered together as shown on far left in green. The cluster that forms all agents together is done in the end, which is shown in color blue.

4.4.2 Funding Evaluation

The data we obtained only contains student performance for the whole school district. It does not provide specific performance evaluation on subgroups (e.g. special education group). Because of this, the result outputted from our model would be the same for all five subareas for each school district.

Our funding policy evaluation model shows the following result:

School district	Most effective funding policy year
Alamo Heights	2009
Harlandale	2010
Edgewood	2008
Randolph	2008
San Antonio	2009
South San Antonio	2010
Somerset	2009
North East	2008
East Central	2010
Southwest	2009
Lackland	2009
Ft Sam Houston	2008
Northside	2009
Judson	2009
Southside	2010

Table 6 Funding evaluation result.

In order to compare the results, here I provide a table of total expenditure per pupil of 2008, 2009 and 2010.

	2008	2009	2010
Alamo Heights	8644	9138	9230
Harlandale	8710	8494	8809
Edgewood	9039	9451	9050
Randolph Field	14787	9309	10620
San Antonio	8745	8743	9153
South San Anton	8581	8978	9426
Somerset	8185	8196	9036
North East	8328	8544	8582
East Central	7830	8033	8028
Southwest	8155	8179	8650
Lackland	13179	11946	12640
Ft Sam Houston	11514	11849	13135
Northside	7943	8028	8327
Judson	7864	7886	8051
Southside	7936	8830	9023

Table 7 Total Expenditure per pupil.

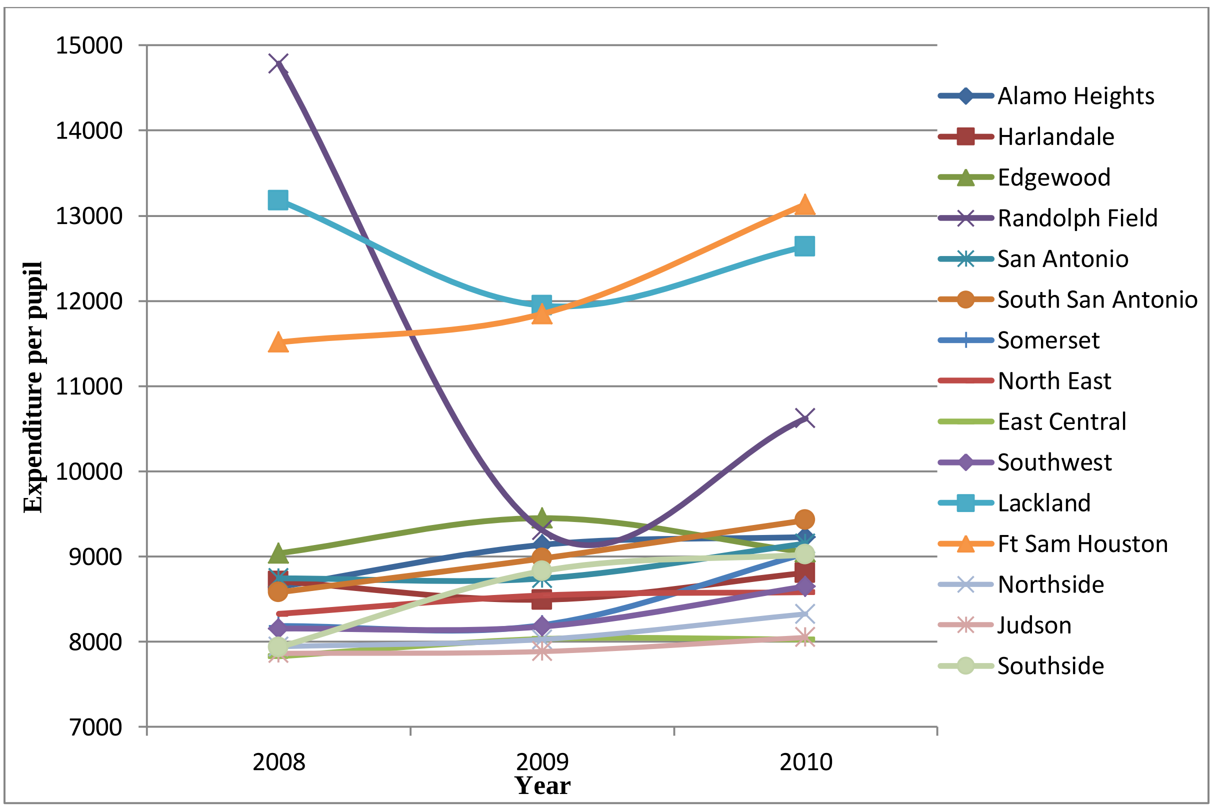

Figure 8 Total Expenditure per pupil.

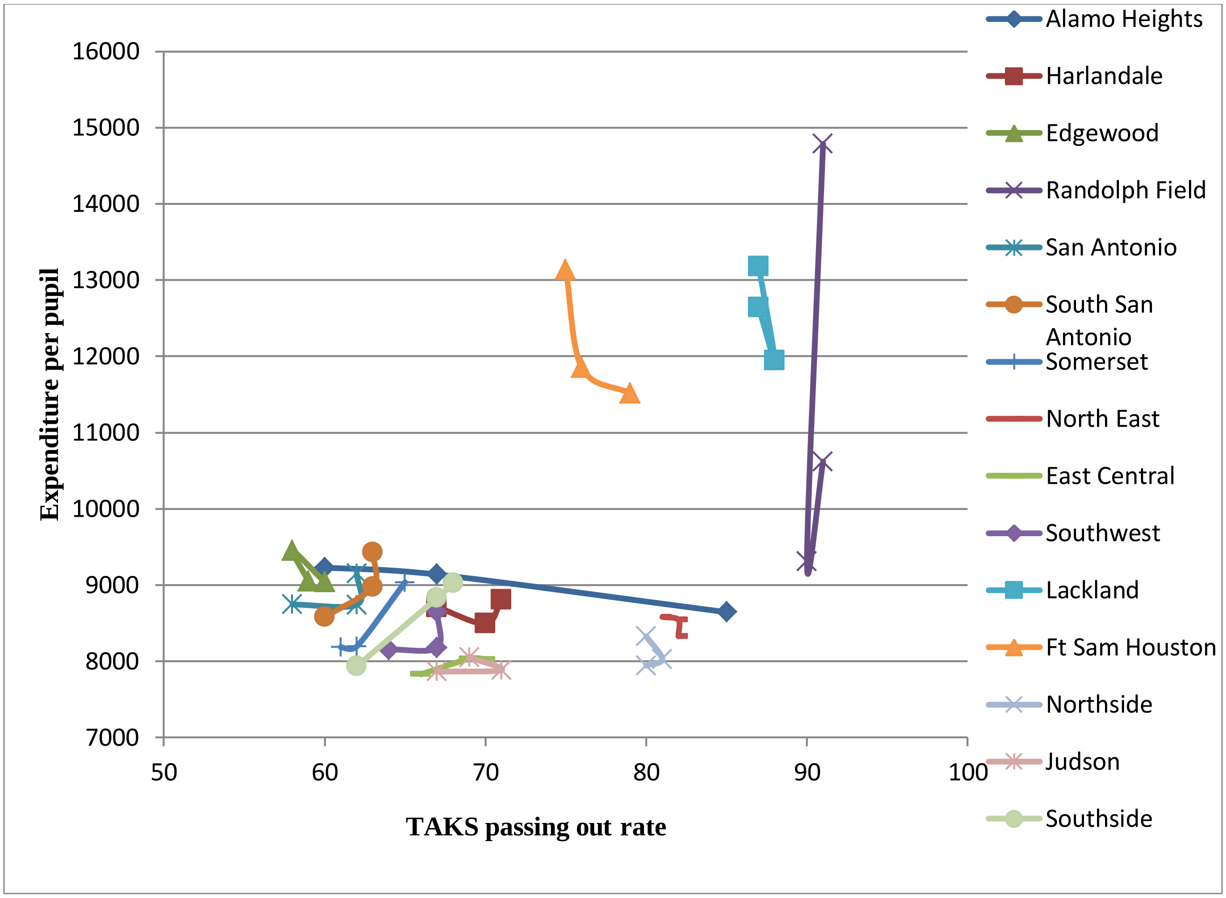

Figure 9 Relations between TAKS passing out rate and expenditure per student of all three years. Note: the three points on each line represents the three years: Left point – 2008, middle point – 2009, right point – 2010.

From the above figure 8 and 9, one can make the following observations: first, for some school districts, when expenditure per pupil increases, they show some decrease on student performance (e.g. Northside and Southwest). But for some other districts, when there is a decrease on expenditure per pupil, their student performances tend to increase (e.g. Alamo Heights and Ft Sam Houston). There is also another kind of school districts, whose students perform better while their expenditures increase (e.g. Somerset and Southside). Therefore we can conclude that it is not necessarily the case that the more spend the better. The data on Bexar County does not show a consistent relation between changes in funding and student performance.

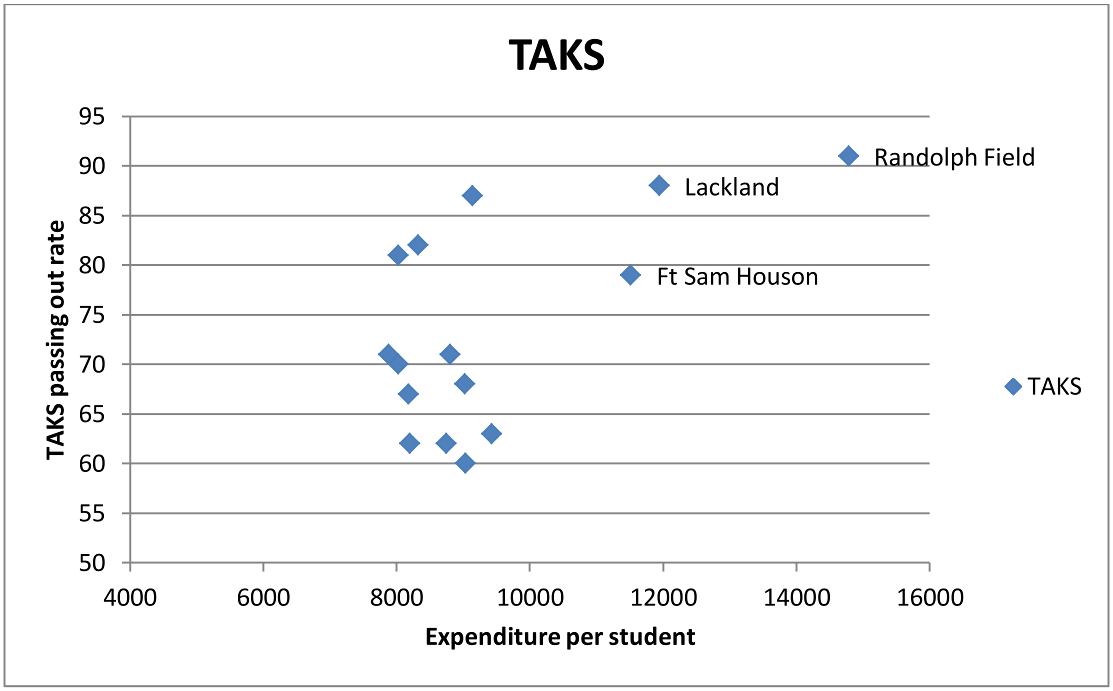

Figure 10 Relations between best TAKS passing out rate in all three years and expenditure per student.

Figure 10 shows the results of comparing all school districts together. For school districts that have their total expenditure per student under 10000, more funding does not seem to make them perform better than other schools. However, Ft Sam Houston, Lackland and Randolph field districts have relatively high expenditure per pupil and they do seem to perform a lot better. Most of their expenditures are above 11000. Randolph field, which is the best performing district, even has expenditure per pupil above 14000. However, these three all appear to be military base school districts, which might be the reason for their possibility of high expenditure. This could not be the case for all school districts.

Therefore it can be concluded that our model shows that increasing funding does not necessarily have a positive impact on student performance and there is not a very consistent relation between student performance and increasing funding. However three school districts show that when the expenditure per pupil reaches a certain high level, it does appear to have a very positive impact on student performance.

Chapter 5
Conclusion and Future Work

5.1 Conclusion

In recent years, many efforts have been devoted into Multi-agent system and social network analysis. Many works have proved that Multi-agent system is a good tool for social network analysis.

This work contributes to both social network analysis and Multi-agent system. We focus on modeling social networks using multi-agent system with a focus on social interactions. By implementing a school network, even with the limited data, this work has shown some very promising results. With the hierarchical structure generated, we are able to evaluate funding polices for past 3 years for Bexar county school districts and conclude that increasing funding does not necessarily have a positive impact on student performance. However for some types of school district, when the expenditure per pupil reaches a certain high level, it does appear to have a very positive impact on student performance.

5.2 Further Work

There are many potential works could be done in the future.

First our current similarity measure use interaction level between agents and the simulation data is defined by us. If one use real data (e.g. using questionnaire to gather interactions information from actual human agents in the network), then the model should provide a more real-world result. In addition, the hierarchical clustering algorithm could be further revised to reduce computational complexity.

In the funding policy evaluation part, the current reward function is limited by the data we could get. We only have data for student performance of the whole school district. The ideal situation would be to have performance data on each “subtask” (e.g. special education). In addition, the reward function could also be re-designed depending on what kinds of data are available.

The current policy evaluation model is applied to funding policy only. One could also use this model to study other kinds of policies. Our model is inspired by the MAXQ algorithm, which is a hierarchical reinforcement learning technique. However, due to the limitation of data, the reinforcement learning part is not fully implemented. If one has more future actions available, one could further develop the model by implementing a learning algorithm.

Bibliography

Anderson, Nicole. "Per Pupil Spending: How Much Difference does a Dollar make?" (2011). *All Graduate Reports and Creative Projects.* Paper 20. Web.

Andre, David, and Stuart Jonathan Russell. "State Abstraction for Programmable Reinforcement Learning Agents". *Proceedings of the National Conference on Artificial Intelligence.* Menlo Park, CA; Cambridge, MA; London; AAAI Press; MIT Press; 1999 , 2002. 119-125. Web.

Barabási, Albert-László, Réka Albert, and Hawoong Jeong. "Scale-Free Characteristics of Random Networks: The Topology of the World-Wide Web." *Physica A: Statistical Mechanics and its Applications* 281.1 (2000): 69-77. Web.

Barto, Andrew G., and Sridhar Mahadevan. "Recent Advances in Hierarchical Reinforcement Learning." *Discrete Event Dynamic Systems* 13.4 (2003): 341-79. Web.

Bettinger, Eric P. "The Effect of Charter Schools on Charter Students and Public Schools." *Economics of Education Review* 24.2 (2005): 133-47. Web.

Bidwell, Charles E., and John D. Kasarda. "School District Organization and Student Achievement." *American Sociological Review* (1975): 55-70. Web.

Boccaletti, Stefano, et al. "Complex Networks: Structure and Dynamics." *Physics reports* 424.4 (2006): 175-308. Web.

Bohte, John. "School Bureaucracy and Student Performance at the Local Level." *Public administration review* 61.1 (2001): 92-9. Web.

Borgatti, Stephen P., et al. "Network Analysis in the Social Sciences." *Science* 323.5916 (2009): 892-5. Web.

Cao, Feng, and Soumya Ray. "Bayesian Hierarchical Reinforcement Learning". *Advances in Neural Information Processing Systems 25.* 2012. 73-81. Web.

Cimiano, Philipp, Andreas Hotho, and Steffen Staab. "Comparing Conceptual, Divise and Agglomerative Clustering for Learning Taxonomies from Text." (2004). Web.

Clark, Damon. "School Principals and School Performance." (2010). Web.

Crampton, Faith E. "Spending on School Infrastructure: Does Money Matter?" *Journal of Educational Administration* 47.3 (2009): 305-22. Web.

Creamer, Germán, et al. "Segmentation and Automated Social Hierarchy Detection through Email Network Analysis." *Advances in Web Mining and Web Usage Analysis.*Springer, 2009. 40-58. Web.

Day, William, and Herbert Edelsbrunner. "Effective Algorithms for Agglomerative Hierarchical Clustering Methods." *Journal of Classification* 1 (1984): 7-24. Print.

Dethlefs, Nina, and Heriberto Cuayáhuitl. "Combining Hierarchical Reinforcement Learning and Bayesian Networks for Natural Language Generation in Situated Dialogue". *Proceedings of the 13th European Workshop on Natural Language Generation.* Association for Computational Linguistics , 2011. 110-120. Web.

Dietterich, Thomas G. "Hierarchical Reinforcement Learning with the MAXQ Value Function Decomposition." *arXiv preprint cs/9905014* (1999). Web.

Dietterich, Thomas G. "An Overview of MAXQ Hierarchical Reinforcement Learning." *Abstraction, Reformulation, and Approximation.*Springer, 2000. 26-44. Web.

Du, Nan, et al. "Large Human Communication Networks: Patterns and a Utility-Driven Generator". *Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining.* ACM , 2009. 269-278. Web.

Färe, Rolf, Shawna Grosskopf, and William L. Weber. "Measuring School District Performance." *Public Finance Review* 17.4 (1989): 409-28. Web.

Fortunato, Santo. "Community Detection in Graphs." *Physics Reports* 486.3 (2010): 75-174. Web.

Grant, Tim. "Modelling Network-Enabled C2 using Multiple Agents and Social Networks". *Social Networks and Multiagent Systems Symposium (SNAMAS), Edinburgh (Scotland).* 2009. Web.

Guessoum, Zahia, et al. "Dynamic and Adaptive Replication for Large-Scale Reliable Multi-Agent Systems." *Software Engineering for Large-Scale Multi-Agent Systems.*Springer, 2003. 182-198. Web.

Gutierrez-Osuna, Ricardo. "Pattern Analysis for Machine Olfaction: A Review." *Sensors Journal, IEEE* 2.3 (2002): 189-202. Web.

Harris, Douglas N., and Tim R. Sass. "Teacher Training, Teacher Quality and Student Achievement." *Journal of public economics* 95.7 (2011): 798-812. Web.

Hengst, Bernhard. "Discovering Hierarchy in Reinforcement Learning with HEXQ". *MACHINE LEARNING-INTERNATIONAL WORKSHOP THEN CONFERENCE-*. 2002. 243-250. Web.

Hoxby, Caroline Minter. *Do private schools provide competition for public schools?* (1994). Web.

Jain, Anil K. "Data Clustering: 50 Years Beyond K-Means." *Pattern Recognition Letters* 31.8 (2010): 651-66. Web.

Jiang, Daxin, Chun Tang, and Aidong Zhang. "Cluster Analysis for Gene Expression Data: A Survey." *Knowledge and Data Engineering, IEEE Transactions on* 16.11 (2004): 1370-86. Web.

Joseph, Kenneth, Wei Wei, and Kathleen M. Carley. "An Agent-Based Model for Simultaneous Phone and SMS Traffic Over Time." *Social Computing, Behavioral-Cultural Modeling and Prediction.*Springer, 2013. 65-74. Web.

Lubienski, Christopher, and Sarah Theule Lubienski. "Charter, Private, Public Schools and Academic Achievement: New Evidence from NAEP Mathematics Data." *Occasional Paper*.111 (2006): A12. Web.

Ma, Jie, et al. "Colony Evolution in Social Networks Based on Multi-Agent System". *Natural Computation, 2009. ICNC'09. Fifth International Conference on.* IEEE , 2009. 594-597. Web.

Marchiori, Massimo, and Vito Latora. "Harmony in the Small-World." *Physica A: Statistical Mechanics and its Applications* 285.3 (2000): 539-46. Web.

Marthi, Bhaskara, et al. "Concurrent Hierarchical Reinforcement Learning". *Proceedings of the national conference on artificial intelligence.* Cambridge, MA; 2005. Web.

Meier, Kenneth J., Laurence J. O'Toole Jr, and Holly T. Goerdel. "School Superintendents and School Performance: Quality Matters." *Retrieved September* 23 (2003): 2005. Web.

Mirzazadeh, Farzaneh, Babak Behsaz, and Hamid Beigy. "A New Learning Algorithm for the Maxq Hierarchical Reinforcement Learning Method". *Information and Communication Technology, 2007. ICICT'07. International Conference on.* IEEE , 2007. 105-108. Web.

Müllner, Daniel. "Modern Hierarchical, Agglomerative Clustering Algorithms." *arXiv preprint arXiv:1109.2378* (2011)Web.

Murtagh, Fionn, and Pedro Contreras. "Algorithms for Hierarchical Clustering: An Overview." *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* 2.1 (2012): 86-97. Web.

Newman, Mark EJ, and Michelle Girvan. "Finding and Evaluating Community Structure in Networks." *Physical review E* 69.2 (2004): 026113. Web.

Newman, Mark EJ. "Scientific Collaboration Networks. II. Shortest Paths, Weighted Networks, and Centrality." *Physical review E* 64.1 (2001): 016132. Web.

Parr, Ronald, and Stuart Russell. "Reinforcement Learning with Hierarchies of Machines." *Advances in neural information processing systems* (1998): 1043-9. Web.

Rangeon, Sandrine, Wade Gilbert, and Mark Bruner. *Mapping the world of coaching science: A citation network analysis* (2010).Web.

Ripley, Brian D. *Pattern Recognition and Neural Networks*. Cambridge university press, 2008. Web.

Rockoff, Jonah E. "The Impact of Individual Teachers on Student Achievement: Evidence from Panel Data." *The American Economic Review* 94.2 (2004): 247-52. Web.

Savaresi, Sergio M., et al. "Cluster Selection in Divisive Clustering Algorithms." *Proceedings of the 2nd SIAM* 1401 (2002): 299-314. Web.

Schaeffer, Satu Elisa. "Graph Clustering." *Computer Science Review* 1.1 (2007): 27-64. Web.

Sklar, Elizabeth, Mathew Davies, and Min San Tan Co. "SimEd: Simulating Education as a Multi Agent System". *Proceedings of the Third International Joint Conference on Autonomous Agents and Multiagent Systems-Volume 3.* IEEE Computer Society , 2004. 998-1005. Web.

Slate, John R., and Craig H. Jones. "Effects of School Size: A Review of the Literature with Recommendations." *Essays in Education* 13 (2005): 1-22. Web.

SNAMAS-09. "Social Networks and Multi-Agent Systems Symposium (SNAMAS-09) Introduction." *Proceedings of the Symposium* (2009). Web.

Steinbach, Michael, et al. "Discovery of Climate Indices using Clustering". *Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining.* ACM , 2003. 446-455. Web.

Sutton, Richard S., Doina Precup, and Satinder Singh. "Between MDPs and Semi-MDPs: A Framework for Temporal Abstraction in Reinforcement Learning." *Artificial Intelligence* 112.1 (1999): 181-211. Web.

Uther, William TB. *Tree based hierarchical reinforcement learning* (2002). Web.

Wakita, Ken, and Toshiyuki Tsurumi. "Finding Community Structure in Mega-Scale Social Networks". *Proceedings of the 16th international conference on World Wide Web.* ACM , 2007. 1275-1276. Web.

Wasserman, Stanley, and Katherine Faust. *Social Network Analysis: Methods and Applications*. 8 Vol. Cambridge university press, 1994. Web.

Wellman, Barry. "Computer Networks as Social Networks." *Science* 293.5537 (2001): 2031-4. Web.

Wellman, Barry. "An Electronic Group is Virtually a Social Network." *Culture of the Internet* (1997): 179-205. Web.

Xu, Rui, and Donald Wunsch. "Survey of Clustering Algorithms." *Neural Networks, IEEE Transactions on* 16.3 (2005): 645-78. Web.

Znaniecki, Florian. *Social Relations and Social Roles*. Irvington Pub, 1986. Web.

Appendix A

DISTRICT NAME	TAKS:	TOTAL EXPENDITURE	Regular educati	Special educa	ESL education	Career educatic	Gifted educatioi
Alamo Heights	87	43715556	29726578.08	6994488.96	0	437155.56	874311.12
Harlandale	71	127324400	68755176	16552172	8912708	3819732	0
Edgewood	59	111248406	50061782.7	13349808.72	7787388.42	3337452.18	0
Randolph Field	91	12701043	9525782.25	1397114.73	0	127010.43	508041.72
San Antonio	62	504207177	226893229.7	90757291.86	40336574.16	15126215.31	5042071.77
South San Antoni	63	93898291	41315248.04	14084743.65	11267794.92	3755931.64	1877965.82
Somerset	65	33487386	15739071.42	4688234.04	1674369.3	1339495.44	669747.72
North East	81	559721134	347027103.1	123138649.5	5597211.34	16791634.02	5597211.34
East Central	70	74311079	41614204.24	14862215.8	2229332.37	2229332.37	1486221.58
Southwest	67	99087846	59452707.6	14863176.9	990878.46	3963513.84	0
Lackland	87	11806204	7674032.6	2361240.8	118062.04	236124.08	118062.04
Ft Sam Houston	75	19322253	11013684.21	6376343.49	0	386445.06	193222.53
Northside	80	761630471	472210892	167558703.6	0	22848914.13	7616304.71
Judson	69	174658801	104795280.6	33185172.19	5239764.03	6986352.04	0
Southside	68	46849510	26704220.7	6090436.3	2810970.6	1873980.4	0

Table 5 Funding data for the year of 2010 (This is the same table as the one shows on page 33).

DISTRICT NAME	TAKS:	TOTAL EXPENDITU	Regular educatio	Special educatior	ESL education	Career educa	Gifted edu
Alamo Heights	87	42201000	28696680	6752160	0	422010	844020
Harlandale	70	121900873	68264488.88	17066122.22	8533061.11	3657026.19	0
Edgewood	58	109710502	54855251	14262365.26	9873945.18	2194210.04	0
Randolph Field	90	11198699	8287037.26	1343843.88	0	111986.99	447947.96
San Antonio	62	475712598	209313543.1	90385393.62	38057007.84	14271377.94	4757126
South San Antonio	63	89483843	40267729.35	12527738.02	9843222.73	3579353.72	1789676.9
Somerset	62	28349875	14174937.5	4252481.25	1700992.5	1417493.75	283498.75
North East	82	539863507	345512644.5	118769971.5	10797270.14	16195905.21	5398635.1
East Central	69	72919493	42293305.94	13854703.67	2187584.79	2916779.72	1458389.9
Southwest	67	92692917	57469608.54	13903937.55	926929.17	2780787.51	0
Lackland	88	11504425	7477876.25	2300885	115044.25	230088.5	115044.25
Ft Sam Houston	76	17038811	10052898.49	5452419.52	0	340776.22	170388.11
Northside	81	708054393	460235355.5	148691422.5	0	21241631.79	7080543.9
Judson	71	167623416	105602752.1	28495980.72	5028702.48	6704936.64	1676234.2
Southside	67	44867400	27817788	5384088	2243370	1794696	0

Table 8 Funding data for the year of 2009.

DISTRICT NAME	TAKS:	TOTAL EXPENDITURES	Regular edu	Special edu	ESL educati	Career edu	Gifted edu
Alamo Heights	85	39519675	27268575.8	6323148	0	395196.75	790393.5
Harlandale	67	122823357	70009313.5	17195270	8597635	3684700.7	0
Edgewood	60	106068921	51973771.3	14849648.9	9546202.9	2121378.4	0
Randolph Field	91	17137802	13881619.6	1542402.18	0	171378.02	514134.06
San Antonio	58	478572444	220143324	90928764.4	38285796	14357173	4785724.4
South San Antoni	60	84130593	37858766.9	11778283	9254365.2	3365223.7	2523917.8
Somerset	61	29063485	15113012.2	4650157.6	1453174.3	1162539.4	581269.7
North East	82	515611617	329991435	113434556	10312232	20624465	5156116.2
East Central	66	68671384	41202830.4	12360849.1	2060141.5	2060141.5	1373427.7
Southwest	64	88978363	53387017.8	14236538.1	889783.63	2669350.9	0
Lackland	87	11320981	7585057.27	2377406.01	113209.81	226419.62	113209.81
Ft Sam Houston	79	16407048	9680158.32	5086184.88	0	328140.96	164070.48
Northside	80	679468526	441654542	142688390	0	20384056	6794685.3
Judson	67	162002348	106921550	27540399.2	1620023.5	6480093.9	1620023.5
Southside	62	39610696	23766417.6	4753283.52	1980534.8	1584427.8	0

Table 9 Funding data for the year of 2008.