Electronic Theses and Dissertations

2014

A Human-Centric Approach to Data Fusion in Post-Disaster Managment: The Development of a Fuzzy Set Theory Based Model

Mubarak Banisakher
University of Central Florida

STARS Citation

A HUMAN-CENTRIC APPROACH TO DATA FUSION IN POST-
DISASTERMANAGEMENT: THE DEVELOPMENT OF A FUZZY SET THEORY BASED
MODEL

by

MUBARAK BANISAKHER
M.S. Nova Southeastern University, 2004
MS. Tajik State University, 1997
BS. Bethune Cookman University, 2005

A dissertation submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy
in the Department of Industrial Engineering and Management Systems
in the College of Engineering and Computer Science
at the University of Central Florida
Orlando, Florida

Spring Term
2014

Major Professor: Pamela McCauley Bush

ABSTRACT

Disaster management is a multi-strategizing process which includes measures such as pre-disaster planning, implementation of the scheme, early warning, emergency response, and emergency relief to reduce the social impact of natural disasters or technological disasters (commonly known as man-made disasters). This dissertation focuses on how to provide an efficient and accurate information system in the post-disaster phase so that individuals are able to access and obtain the necessary resources in a timely manner. First, we studied the current post-disaster management system (PDMS) model and proposed a new data fusion method to process dataset for a geographic information system (GIS) based PDMS. Second, a human-centric approach for multi-source and multi-modal data fusion methods was specifically proposed to improve the dynamic metric of data collection. Third, a dynamic PDMS model was proposed; the dynamic knowledge database of PDMS was established by employing a rule-based fuzzy inference system (FIS). Then semi-supervised learning (SSL) based on graph and fuzzy set model were proposed to improve the presentation capacity of nodes in this GIS based PDMS. Simulation and comparative analysis on graph-based SSL and fuzzy set model were developed, and the results showed their effectiveness. We additionally designed a survey to collect data on the performance of hospitals' emergency abilities in post-disaster situations. The survey used fuzzy factor analysis to determine that the hospitals which performed best would be located by individuals in the simulation. The proposed methods offer a means of timely and effective decision-making for individuals in post-disaster situations and further aim to reduce casualties and property damage in post-disasters. The research context primarily focused on the following aspects:

(1) Studied the current PDMS architecture, design methods and applications, specially introduced some kind of data collection methods for PDMS. The general model of multi-input/multi-output dynamic system will also be introduced. System stability and reliability will be studied by researching PDMS applications.

(2) Studied the human-centric approach for data fusion technology and convert human-centric approach, which represents the human needs, to Human Factors (HF). To do this we will analyze human needs using factor analysis tools such as Principal Component Analysis (PCA); proposed a Gaussian-shaped optimizing method for data fusion of PDMS' emergency resources data, single source data fusion by Single Gaussian Model (SGM) and multi-source data fusion by Gaussian Mixed Model (GMM). These processes were combined by HF- a human-centric approach customized specially for PDMS as a density parameter in fusion process.

(3) Introduced a multi-mode data fusion method; it includes giving a formal expression of multi-mode data after deleting invalid or redundant attributes, normalization methods using density function, and finally, exploring their advantages and disadvantages.

(4) Compared unsupervised and semi-supervised learning models using fuzzy set models, and illustrated how targeted individual involvement needs intervention under the semi-supervised learning algorithms in order to meet the needs of decision-making information.

(5) Based on the PDMS knowledge presentation and its reasoning mechanisms, a rule-based fuzzy inference system model will be applied for knowledge presentation of source information of post-disaster. The individual needs vector information will be stored as an inputs to the system.

(6) Simulation and comparative analysis methods included Gaussian fusion for map pictures and compared with PCA, graph-based SSL using Gaussian dataset and compared with

non-Gaussian dataset, and the rule-based knowledge presentation system by fuzzy mean

compared with a clustering algorithm.

(7) A survey was designed to collect data from several hospitals located in a small city in

Volusia County, Florida, USA. The survey data was analyzed using fuzzy factor analysis under

different levels to measure each hospital's performance ability during post-disaster emergencies.

Evaluation factors were categorized into nine groups with 72 sub-factors. Nine hospitals were

involved in this survey and approximately 53 records were stored in database. Each factor was

weighted to determine the final result in addition to a simulation which required individuals to

locate the best performance on emergency ability of hospitals in emergency situations.

ACKNOWLEDGMENTS

I would like to express my deepest appreciation to my committee chair Dr. Pamela McCauley-Bush, whose expertise and understanding were indispensable in my work. She was a source of constant support and her advisement kept my focused during my research. The research freedom she provided, with valuable guidance made my research experience continually exciting. Without her guidance and persistent help, this dissertation would not have been possible.

I would like to thank my committee members, Dr. Gene Lee for all the help and support he provided me. He responded graciously to my call when I asked him to be part of my success. Dr. Christopher Geiger I thank, not only for being the most helpful professor during my years of study at UCF, also during my research. He guided me during my preparation for the candidacy and the final presentations. I would like to thank Dr. Cliff Zou for being incredibly nice to me, for without his help I could have lost my status as an active student at the University in one of the semesters. I also want to thank Dr. Fuqian Shi for responding to my questions throughout my research.

I dedicate this thesis to my parents, my family, my wife, and sons and I would like to thank everybody who has supported me throughout the whole process, both by keeping me harmonious and helping me to stay focused.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

1.1 Introduction

The environmental impact of natural disasters is increasing; there are many factors leading to the occurrence of this phenomenon, such as the increase of population growth density, population migration and unplanned urbanization, environmental degradation, and global climate change. The political measures on modern social risks of natural disasters for socio-economic issues have been changed.

In the last 20 years, a greater proportion of the population died due to natural and unnatural disasters in the 1980s (86328 / a) than in the 1990s (7525 / a). However, the affected population from the 1980s to the 1990s increased from an average of 147 million per year to 211 million people per year. The number of geological disasters remained fairly stable while the number of hydro-meteorological disasters was increased (caused by water and weather disasters). In the 1990s, more than 90% of the population died due to natural disasters in hydro-meteorological accidents such as droughts, storms and floods. Of the population affected by natural disasters, floods affected more than two-thirds, but compared with other disasters, it resulted in a lower proportion of death, accounting for only 15% of the total deaths (Munichi Re Group, 2001).

The occurrence of natural disasters has had a huge impact on the affected regions; the affected area of the economy, culture and all aspects of society were hit hardly. As per Archibald and McNeil (2012), approximately $10 billion in damages is caused by hurricanes in the U.S, every year. Aside from this, they possess the potential to cause heavy catastrophic losses. For instance Hurricane Katrina caused $108 billion in property damage and killed 1,200 people in

the year 2005 on the Gulf Coast of the U.S. Since populations around the coastal areas are expanding, hurricanes, floods and tsunamis are becoming a common cause of concern for people. Evacuation is the most effective measure used to protect people pre and post-disaster. This disaster management tactic encourages or forces visitors and residents in designated danger areas to relocate. Evacuations are exceptionally useful in the event of hurricanes given that officials are able to facilitate evacuees with advance warnings of disasters.post-disasterFrom 1997 to 2007, the annual loss caused by natural disasters in the United States was fell between $500 billion to $ 1,000,000,000 per week (IDNDR 1999a) (Munichi Re Group, 2001).

Natural disasters not only cause significant damage to the economy but also heavy casualties. During the period 2004-2008, tens of thousands of persons in the United States were affected by natural disasters each year. According to the statistics of the world's eight major countries, the annual average was more than 10,000 individuals affected by natural disasters, of which, 80% of casualties were affected in a secondary stage after the disaster, so secondary prevention interventions were discussed (Ruggiero et al., 2012; Widener et al., 2011). An effective and established disaster rescue plan improved for disaster information infrastructure systems and facilities could be of great value in post-disaster relief. At present, the majority of post-disaster management systems (PDMS) only focus on disaster warning and information collection, however emergency response mechanisms do not give full play to the role of information technology and communications technology. Additionally, it did not take full advantage of the existing network infrastructure and mobile equipment; the previous PDMS biased in favor of the specific relief measures. In fact, they would be quite ineffective for the disaster communication mechanism if individual autonomy in disaster relief did not apply (Lee et al., 2011).

PDMS not only has the basic geographic database facilities, but also needs to constantly be updated on the infrastructure in pre-disaster including hospitals, storage/warehouses, and traffic conditions. In fact, a dynamic PDMS plays a crucial role for individuals. PDMS can be used to efficiently provide information to individuals in post-disaster situations, but can also effectively be updated by the same individuals (Liang et al., 2009). PDMS operates more efficiently when managed by the autonomous behavior of the thousands of individuals involved in post-disaster relief. For example, satellite navigation system guilders for improving the traffic environment relies on individuals' active feedback to the system. Through the PDMS establishment with scalable interface, the use of communication technologies to collect individual's needs, and the establishment of emergency response methods based on the current effective resource supply capacity, individuals will fully get the current disaster situation and the availability of surrounding resources in real-time and also receive more effective path information for decision-making (Ahmed, 2013; Brandon, 2011; Raju et al., 2013).

There are many problems associated with PDMS having self-learning ability; these problems include considering system stability and redundancy. From the point of view of system design, the system is full of incomplete disaster information by a distributed database system with a large number of customer applications, especially in post-disaster where poor communication and damage to network infrastructure exist. Following the PDMS's dynamic ability, our target is to discuss how to maintain the integrity and validity of the information, that is, how to filter the valid information to enable the system in forwarding the development of a knowledge convergence process. Secondly, we need to know how to carry out real-time calculation of the PDMS to allow individuals to get effective decision-making support. There have been some worldwide disaster information systems, for example the 1995 Hyogoken-

Nambu (Kobe) earthquake had a major devastating impact which triggered the need for an information system that will be at least capable of handling the following aspects (Iwai et al., 1998):

- Damage distribution of lifelines and buildings.

- Lifeline recovery and refugees in emergency shelters.

- Damage assessment of buildings.

- Seismic upgrading of existing houses.

Pakistan's PDMS was a well-developed system, adequately and effectively acting as a source of information for the system by fully integrating Remote Sensing (RS), Geographical Information Systems (GIS), Global Positioning System (GPS), and communication technology. Mubushar, et al. (2005) and Wu et al. (2012) developed an evacuation logistics variables system for the some activities resources being reached a safe location and remain there until it was safe to return during Hurricanes Katrina and Rita. Based on the individual ability to access to decision-making support in post-disaster situations, Japanese scholars developed a PDMS by using digital pens and tabletops as user interfaces (Hidemi et al., 2012). In the 1980s, Herbert N. Wigder raised the possibility of using mobile phones for communication between the individual and PDMS (Wigder et al., 1989). The key challenges and strategies in post-disaster management and the main measures that will prove to be helpful in responding, mitigating and improving the disaster situations. Kevin E. Nufer (Nufer et al., 2003) also established some patterns that might help in disaster planning and proposed different medical needs in hurricane and other disaster events that will be benefit the research for decision-making in post-disaster management. It will conclude the pertinent role that post-disaster management techniques play in responding the disasters. The constantly escalating disasters and their heavy impact on society and the

environment have made it obligatory to manage and evaluate the data that has been generated in disaster management sessions.

This survey will present a wide array of information with respect to data analysis in disaster management as it will offer an amalgamation of all prominent information, which until now has yet to be submitted anywhere else. It is noteworthy that there are currently considerable problems with accessibility, availability, and usage of trustworthy, up-to-date and precise data for disaster management. This dimension is, however, extremely crucial to disaster rejoinder as timely, current, and accurate spatial data unfolding the contemporary situation is principal to effectively countering back to an emergency.

The previous literature showed that the information obtaining from a PDMS was very important. This prompted us to focus on ways to ensure the efficient use of individual PDMS. If we consider the design of disaster information system to effectively solve the individual information to a sufficient demand, meanwhile, it will allow individuals to make the maximum extent for the support of information systems; it also will bring a very important role in post-disaster relief works. Bottleneck of this study will be the efficient integration of information systems, data sources, effective communication of the system's self-learning ability as well as the individual and the system. Through this study, the bottleneck to resolve these problems is to improve the current PDMS model by introducing a dynamic PDMS model, by doing this improvement, we are expressing individual needs in post-disaster that will greatly enhance the completeness, real-time and efficient of the PDMS, thus, improve the post-disaster emergency relief, reduce disaster losses, stable the victims' psychological and emotional problem during post-disaster (Pidgeon et al., 2000).

Disaster service information system has been investigated by previous scholars in recent years; Urakami et al., (2009) proposed a disaster service information system so that disaster information can be shared automatically amongst shelters. It is vital to consider that collecting disaster information in the microcomputer is essential, which itself has constraint in data volume. Additionally, it is also necessary that information related to disasters must be efficiently transmitted via a relatively low speed wireless network. A data format and conversion algorithm can be taken into account in order to disseminate the disaster information efficiently. However, it is necessary to take into consideration that the contents of information related to disaster that will be stored in the system are strongly dependent on the storage capability of microcomputers. In this context, a converting method is necessary that will compress the data volume.

Floods are other major disasters that are frequent in countries like China and India and cause relentless damages. In this context, it is necessary to do risk analysis in flooding so that appropriate measures can be determined in order to control and mitigate floods. Although, there are a wide number of approaches that assist in analyzing the risk of floods, for instance, system modelling and simulation, hydrology-hydraulics model, RS and GIS as well as number of factors like flooding and relations between them, but it is difficult to configure and develop the process of flooding and make accurate forecasts that help in the management of floods.

Duan et al. (2011) have considered the Songhuajiang River Basin as their test area in order to suggest an advanced regional risk assessment model and in this context they have taken into account various factors, like rainfall, population density, topography, water density, Gross Domestic Product (GDP), etc. Risk assessment is crucial in flooding as it is a procedure of combining diverse factors. However, the model presented by Duan et al is limited in data

acquisition as not all factors have been taken into consideration, such as disaster prevention capacity, land use, and other significant factors.

Yamamoto et al., (2011) presented an application that can not only help in surveying the disaster area, but also in the diffusion of real-time video images to headquarters situated at near distance. This application is called an unmanned flying observation robot, which will aid in the observation and forwarding of disaster situations. The same will be done with the help of the usual internet networks, and thus, this system does not require any special or expensive equipment. This application is useful enough as it assists in delivering several observations from the disaster area on an hour to hour basis. It is considerable that during disasters like earthquakes, tsunamis, typhoons, etc. helicopter cannot typically be operated because of adverse weather, but unmanned helicopters are able to better cope with the situation as it is relatively faster, more accurate, safer and cheaper.

<div align="center">1.2 Problem Statement</div>

Typically, PDMS includes basic database construction in pre-disaster, the foundation of network communications platform, and application system of client side. In which, in the pre-disaster construction phase, the underlying database needs to be studied focusing on the distribution and status of the platform architecture, including GIS, emergency information about resource distribution (such as hospitals, warehouses, road infrastructure, hedge places) and post-disaster data collection and treatment. In order to establish a system model with dynamic self-learning capability, it is necessary that the system provide efficient decision-making information to individuals in post-disasters. In our dissertation research, we will address the following issues:

- The need to study the current PDMS's problems and how to put forward corresponding measures regarding the current shortcomings. We will also study the parametric model of

PDMS, including data acquisition method, data fusion methods, and fuzzy set based model. We need to analyze the existing disaster information management system model is unified, parameterized, and further to give the corresponding of its description.

- Data pre-processing of PDMS and the needs to find a solution to remove redundant data which includes how to propose an effective integration method for multi-source data. For multimode data preprocessing, we also need to find a solution to normalize and discover how to import the entire fusion process applying human-centric approach-human factors.

- Present the knowledge of PDMS under formal expression situation, where we need to find a solution of how to present the knowledge-base, and how to build knowledge inference system and make it practically efficient in PDMS.

- Create a dynamic PDMS model which includes the use of semi-supervised learning model with fuzzy set and ensuring individuals' need as a semi-supervised guidance factor to improve the effect of the system.

- Create a simulation environment to validate the proposed model in order to obtain the performance measures that will show the effectiveness of the model including semi-supervised model with fuzzy set and rule-based knowledge presentation using test dataset.

- Designed a survey to collect the data of the performance of emergency ability of hospitals which will be benefit all individuals in post-disaster to find the nearest and best support resources.

1.3 Organization of the Dissertation

The remainder of this dissertation proceeds as follows. Chapter 2 introduced a literature review of the PDMS general model and existing cases; various data fusion technology, and

human-centric approach were also introduced; fuzzy inference model for knowledge representation system, as well as semi-supervised learning method and disaster information system are introduced. Chapter 3 presents a formal model of PDMS, human-centric data fusion methods including a human factor control technology applied on multi-source, multi-mode data fusion algorithm. Rule-based knowledge representation methods of PDMS were proposed. Studying and comparing with semi-supervised, fuzzy clustering algorithm. Chapter 4 described the simulation process for the proposed model and algorithms, mainly for semi-supervised learning with fuzzy set and rule-based presentation in PDMS; discussed and analyzed the results. In Chapter 5, we designed a survey form and distributed them to some typical hospitals in a small city of USA; the aim of this survey was to collect data to show the emergency ability of hospitals in post-disaster situations. The survey was comprised of nine groups and 72 questions. Through fuzzy factor analysis, we determined the overall evaluation of each hospital, and then simulated and found the best or most suitable resources for each individual to locate using fuzzy model proposed in this dissertation. Finally in Chapter 6, we present conclusions and a summary of the entire dissertation research and future works.

CHAPTER 2: LITERATURE REVIEW

2.1 Post –Disaster Management System

Disasters have a devastating effect on the environment of the human beings and human survival. Disaster does not usually refer to only to local effect, but due to expansion and development, disasters continue to have global effects. In accordance with the cause of some disasters, they are classified as man-made disaster and natural disasters. Leidner, Pan and Pan (2009) mention the prominence of information technology (IT) in managing disasters. In this context, considering the case study of Singapore's response to severe acute respiratory syndrome (SARS) and Asian Tsunami disasters, the authors have explain the role of IT in crisis response. After analyzing the case from different dimensions, the authors make it clear that the infrastructure of information technology, combined with leadership networks, and present capabilities, such as the ability to construct and apply IT, the aptitude to identify signals, and the aptitude for a broad view, are all crucial aspects of disaster management and crisis response. IT and informational structures help in a vital manner in countering back to the crisis. In order to follow the movement of patients' contact and health conditions numerous IT applications were utilized.

However, Widener and Horner (2011) use a hierarchical approach for managing hurricane disasters and the distribution of goods to affected people in highly populated areas. It is essential to generate effective plans to provide aid after hurricanes, as well as after other severe weather events by utilising Geographic Information Systems (GIS) and Spatial Optimization Strategies. Although not all storms cause serious damage to human life, all of them do in some

manner impact the urban systems, which in turn leave the population with lack of food, water and other necessities.

Curtis and Mills (2012) highlight the importance of suitable related disaster data and its appropriate fusion in order to do spatial analysis. It is mentioned that fine scale response and recovery data related to disasters for the purpose of doing spatial analysis is still very rare. It is important to make a note that this is unlucky as deep information with respect to spatial models of recovery is of high importance in envisaging the restoration of homes, streets and neighbourhoods. In this context, authors make recommendations to collect fine scale geographic data in real-time for the transitional phase between response and recovery.

This information or data is helpful as it can be utilized for assessing and analyzing the degree of damage, and at the same time, for creating a baseline for consequent monitoring of recovery. In this manner, suggestion regarding a spatial video system has been made that has been used for collecting data from the post-disaster landscape of Tuscaloosa, which was affected severely by a large tornado. After processing, this video can be viewed easily within a Geographic Information System that amalgamates street-level images with exact location. Additionally, such data can also be exploited for supporting the in progress recovery. Spatial video technology is very useful in the different visible dimensions of recovery as it helps in capturing, mapping, geo-referencing and analyzing, which in turn is helpful for planners to plan more effective and sustainable recovery after disasters.

In the viewpoint of Hristidis et al. (2010) the research area of disaster management is very significant due to the rate at which it is gaining growing attention. The role of computer scientists is major during the situations of disasters as they can aid in formulating ways that will

11

aid in managing and evaluating the data generated at the time of disasters. It is essential to manage disaster related data and analyze it for evaluation, because with this help its planners can devise ways for responding and mitigating disasters. Disaster related data can be held in numerous forms, such as satellite imagery data, newspaper articles, reports, remote sensing data, etc. However, there are a number of challenges associated with the management of this data, for instance, it has a wide number of users, time sensitivity, and it relies on the level of credibility, its assorted format, and so on. Despite this, there are several data analysis technologies which can be of great help, such as information extraction (IE), which helps store assorted data in a common structure. IE further aids in processing and information retrieval (IR) which will help in investigating disaster-related information. Additionally, information filtering (IF), data mining and decision support are other helpful technologies. (Witteveen et al., 2012) In relation to the importance of technology in disaster management, Lee et al. (2012) bring to light a multi-agency disaster management system. In this manner, the authors extend the literature that defines the importance of information systems and advanced technologies in disaster management of large scale and high intensity. It is significant to note that quality of information and quality of the system both are important obstacles in effective and competent multi-agency disaster management, and thus, they are critical precursors for the success of information systems. The model Lee et al. (2012) use illuminates the purpose of IS usage as a contemplative determinant of IS success in the public sector domain of disaster management (DM). It is considerable that models of IS that have been suggested previously for business environments will fall short due to their expounding power and applicability in severely impulsive disaster environments that demand rapid response from different organizations.

The number of terrible disasters has increased at a vast level in the last few years and is causing a threat to general people's life. However, spatial information technology can be applied to natural disasters for mitigating them in near future. This shows immensely helpful data for post-disaster management as it makes it clear that spatial information technology aids in the examination and determination, risk assessment and zoning, loss analysis, distribution and law, forecasting and also in active monitoring of the application in natural disasters. Spatial information technology is an inclusive incorporation of geographic information systems, remote sensing, global positioning systems, communication and network technology. This technology allows for qualitative and quantitative evaluation of disaster regions and providing detailed information. Additionally, it helps in assessing and interpreting remote sensing image data and can also aid in the decision-making and managing of natural disasters (Liu et al., 2010; Pidgeon et al., 2000); it also can be classified as geological disaster, weather and environmental disaster, chemical and biological disaster, and marine disasters, location and mechanism of the delineation. The continuous occurrence of worldwide disasters has a huge impact on human society and a destructive of power resource. There are many huge disasters took place throughout the year of 2011 as seen below:

- Feb 22, New Zealand's second largest city-Christchurch: M6.3 earthquake killed many people and due to the magnitude of the earthquake, the focal depth was shallow which caused very serious damage. The public reflects the people trapped under collapsed buildings. It was filled with the smell of gas in the air of the entire city, Christchurch Cathedral partially destroyed.

- The Northeast United States suffered a big snowstorm which resulted in many deaths. Land and air transportation were seriously affected in parts of Washington, with snow thickness over 1 foot (approximately 0.3 m), and blocked people the way back home. Since the beginning of winter, New York City's snowfall record was nearly 140 cm, and reached 96 cm in February.

- On the afternoon of March 11, the Tohoku and Kanto region domestic observed a M8.8 earthquake, the largest in the history of the Richter scale; the source was located near Miyagi Sanriku coastal. The earthquake triggered a tsunami in a wide range of areas from Hokkaido to Okinawa up to (10 m). Japan's defense ministry reported that, in Minamisoma and Fukushima, about 1800 people were suffered by a devastating blow to the deceased.

- May 25, 11 states in the Eastern United States suffered from the largest tornado hit in recent history. In Joplin city, Missouri, people stood on the ruins of the tornado-destroyed houses; statistics show that tornado disasters led to the death of 122 people.

- June 4, in Southern Chile, Puyehue Cordon Caulle volcanic eruptions, profuse volcanic ash with the emergence of the phenomenon of volcanic lightning. A large number of hot ash and stones on the red sky leading to flights cancellations, following the timely emergency information service, more than 3,500 people evacuated.

Floods, tsunami, earthquakes and other natural disasters also brought up the number of deaths, and the number of people indirectly affected is even larger. The analysis of post-disaster casualties to the number of people affected directly by the world's major disaster situation in several countries are illustrated in Figure 2-1, Figure 2-2. The number of people affected and deaths are also illustrated (provided by datamarket.com). The analysis indicated that the majority

14

of the people affected need a Post-disaster Management System (PDMS) to provide a strong and

efficient support.
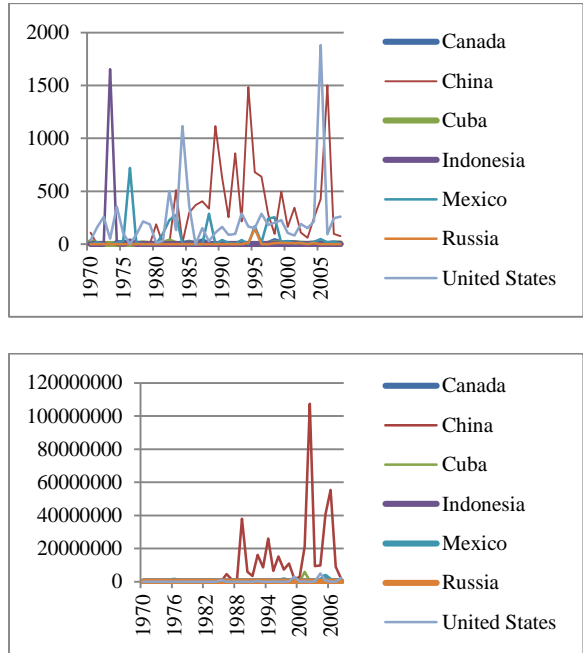
Figure 2-1 Eight Storms Killed and Affected Population of Eight Countries From 1970 to 2008
Source: DataMarket Inc.
http://datamarket.com/data/list/?q=Tsunami&ref=search
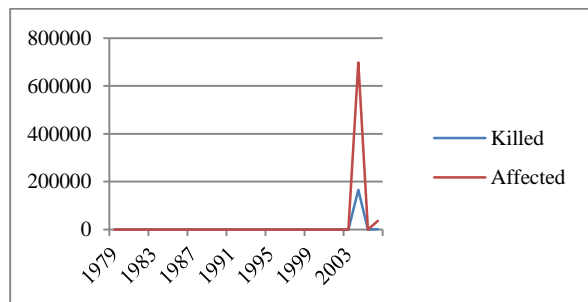
Figure 2-2 Tsunami Killed and Affected Population of Indonesia From 1979 To 2008
Source: DataMarket Inc.
http://datamarket.com/data/list/?q=Tsunami&ref=search

15

Although, it is meaningful that an earthquake cannot be stopped from occurring, or a volcano from exploding, or the sea from erupting, the technical know-how and scientific knowledge can be used to offer early warnings, to increase the earthquake and wind resistance of houses, bridges, poles, and also to arrange appropriate community response to disasters. Over the last few years, technological expertise, technologies for confronting disasters and scientific knowledge in terms of the intensity and allocation of natural hazards have grown to a large extent. It will be fruitful to the survey, an array of technologies and technologically advanced that can be used in managing, mitigating and responding to catastrophic disasters and it will also be useful in post-disaster management. In this context, the classification scheme that has been generated in this study is to classify the post-disaster management related discussion to the advanced technologies and scientific knowledge and identify the role that novel technologies and technologically advanced methods play in disaster response (Shahid Hamid et al., 2010). The study of post-disaster literature shows that there are many issues involved in PDMS like food distribution (Widener et al., 2011), the economic and social impact in post-disaster situations (Tsai et al., 2010; Tsai et al., 2011). For this Shahid Hamid (Shahid Hamid et al., 2010) introduces an improved Monte Carlo model to estimate the average annual loss from hurricane wind damage to residential properties in the state of Florida, and psychological impact (Witteveen et al., 2012; McLaughlin et al.,2010) which are the indispensable works in recovery efforts segments. In general, pre and post-disaster management includes at least five phases (Tsai et al., 2011) (Shown in Figure 2-3).
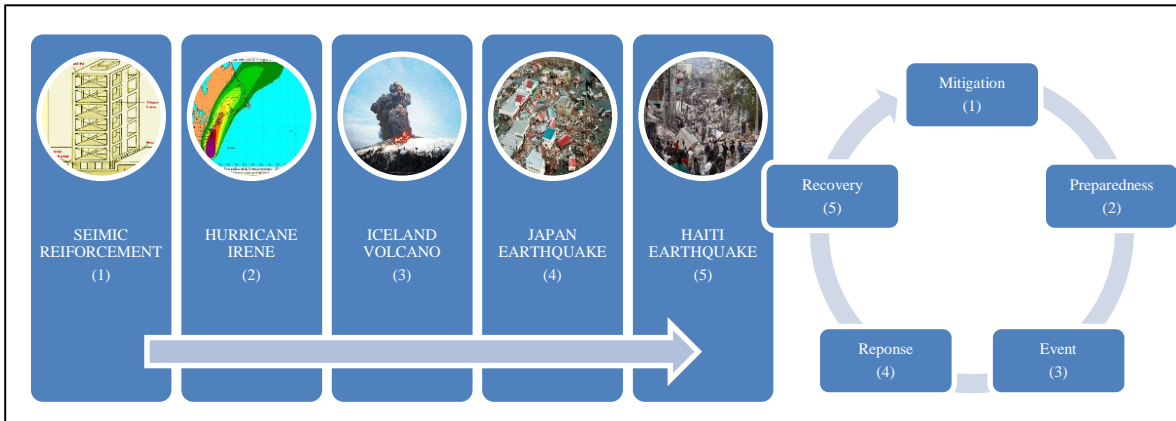
Figure 2-3 Five Phases of Pre and Post-Disaster Management and Cases

Despite these five phases, this dissertation research work focuses on the ability of an individual to obtain necessary resources from PDMS and how to provide a high-quality platform and intelligent models to acquire the most efficient information for decision-making in post-disaster.

Information technology (IT) applications have been applied effectively in post-disaster management, such as, GIS and Google map. In our system we will improve the data fusion technology by supplying network for PDMS. Hristidis et al. proposes a data collection method from an effective complement in pre-disaster including previous activities of the disaster repository through open access of GIS or GPS data (Hristidis et al., 2010). The functions from pre-disaster will benefit the post-disaster management systems as follows (Curtis et al., 2012):

- Early detection

- Prevention or advance warning

- Problem analysis and assessment of scope

- Public notification system and appropriate authorities

17

- Mobilization of a response

- Containment of damage

- Relief and medical care service

On the other hand, spatial video dataset will also assist in creating the primary database of PDMS. Li et al. (2011) show that scale geographic data was collected effectively for the real-time intermediate phase between response and recovery.

However, for the essential database, the problems of data conflicts, multi-source and multi-modal data need to be researched further; also the knowledge-base of the system with a dynamic ability is a key factor for PDMS which should provide more efficient decision-making for individuals. The recent progress in the science and technology of natural disasters and associated management has made it possible over the period to introduce pertinent changes in the incorporated approach to the problems of natural disasters. The surveyed techniques that will be used in the paper with the above classification scheme, are technologies that aid in comprehending the mechanism of atmospheric, biological, geological, and other related origins, and to evaluate that how such disasters occur and what makes them transfer develop from hazards into disasters. The surveyed techniques that will be focused on the study regarding post-disaster management are technologies that not only assist in managing disasters, but also in supporting data collected after disasters to support post-disaster management (Preece et al., 2013).

The importance of PDMS also leads to the worldwide organizations to develop a more effective model to address this issue. The space-based disaster information management and emergency response system, hosted by United Nations, is a platform to promote space technology in disaster management and emergency response, and a program led by the United

18

Nations office for outer space affairs. The platform uses existing space technology, such as, Earth observation, meteorological satellites, and communication and navigation satellites, to provide accurate and timely information for decision-makers in support of disaster management. This plays an important role in knowledge acquisition, processing and transfers they are the core elements of the activity monitoring.

Following the above mentioned, we know that IT plays a critical role in PDMS, especially for a GIS based PDMS. But the current PDMS only provide some static information for every individual in post–disaster that will influence the timely response for getting effective information from PDMS, which is the reason that we need to improve this manner and benefit individuals in post-disaster to acquire "quick response" and "effective response" from the system.

2.2 PDMS Data Collection and Analysis

In order to establish an effective collaboration between PDMS and individuals for data sharing, appropriate agreements with a timely response for the need of these individuals must be adopted. Such agreements will increase their willingness to participate in this process. Additionally establishing an appropriate relationship between government, private and academic sectors will lead to better utilization of their capabilities (e.g. national mapping agencies) and allow PDMS to acquire plenty of data from the disaster response.

Data fusion regarding post-disaster management will help in gaining the current available knowledge and experience, which in turn will be helpful in declining the overall cost of these activities. Mansourian et al. (2006) illustrate that a disaster is a severe disruption in the working of a community, which causes losses that go beyond the recovery ability of the affected community. Disasters result in heavy monetary losses, as well, because financial resources allocated to development will be diverted to disaster issues and in helping people to recover their

19

lives. It is noteworthy that proper management of disasters is a necessity, since they impact sustainable development of society. In this context, disaster management can be regarded as a series of tasks comprising of improvement, awareness, retort and recovery.

Ikuo Yamamoto et al. (2011) establishes a formal structure using GIS and spatial data affairs at group and organizational level of disaster management community. He shows the benefits of GIS data, particularly how spatial data application in PDMS as it appears in Figure 2-4.
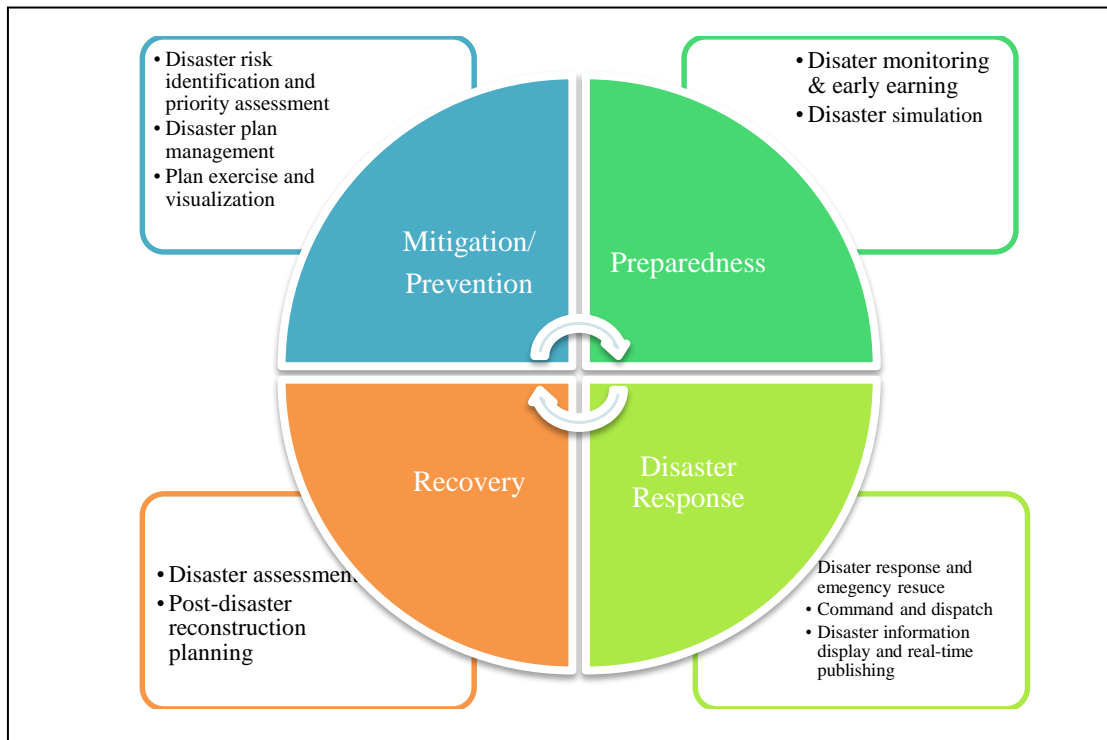
Figure 2-4 GIS Benefits in Disaster Management

Data fusion can substantially assist in disaster management as information collected has a spatial component and without it disaster management cannot be done effectively and efficiently. Although, disaster management is rapidly gaining attention from different researchers, it is difficult to collect information related to data management and evaluation of disaster

management. Since accurate and timely information of the disaster area is very important,

Fajardo et al. utilizes delay tolerant network (DTN) architecture to make the phones of people

serve as sensing nodes in post-disaster situations more effective, as it shown in Figure 2-5)

(Ishimaru et al., 2010).
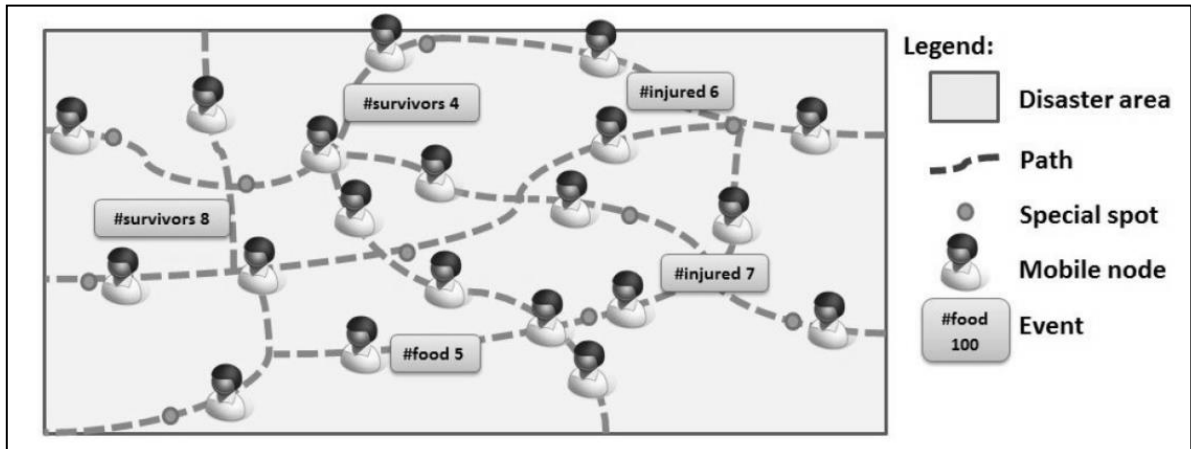
Figure 2-5 Individuals' Situation in Post-Disaster and Their Relations

Data fusion in post-disaster management helps in analyzing, discovering, and organising

all the pertinent information. From the research of Fajardo et al. (2012), it is apparent that since

2002 to 2011, 4130 natural disasters have been recorded, which in turn have caused

approximately more than 1 million deaths and property loss worth billions (Fajardo et al. 2012).

However, authors highlight that fact that telecommunications infrastructure have also

been damaged as well as several buildings and other infrastructure. It is significant that

responders depend on information regarding the affected areas for assessing the situation, and at

that time, even with the unavailability of communication infrastructure, information must be

collected with greatest coverage of the area so as to deal with the critical need (Holt-Giménez,

2002). This makes it clear that different authors have relied on various technologies or

techniques for the purposing of data, and thus, assessing the different data collection methods

that researchers have utilized in post-disaster management will help in providing a rich blend of data.

Some scholars focused on human-interactive interface of PDMS through Geospatial Semantic Web (GSW) technologies and natural language interfaces, in which, it will search geospatial features automatically from multiple semantically heterogeneous source interfaces (Zhang et al., 2010). Zhou et al. (2009) also introduced a GIS-based approach including representation, organization and accessing of PDMS, including logical data models presentation of disastrous events, database implementation techniques, and database queries and report generating methods.

A huge project will help to create basic database for data collection of PDMS- weather bug system, this GIS-based weather-bug system will also benefit disaster risk management, focus on observations from a real-time emergency management disaster scenario (Zerger et al., 2003). Some PDMS data collection works were based on Earth Network Project (ENP). Figure 2-6 shows the weather site distribution of North America and Figure 2-7 shows the worldwide sites distribution.
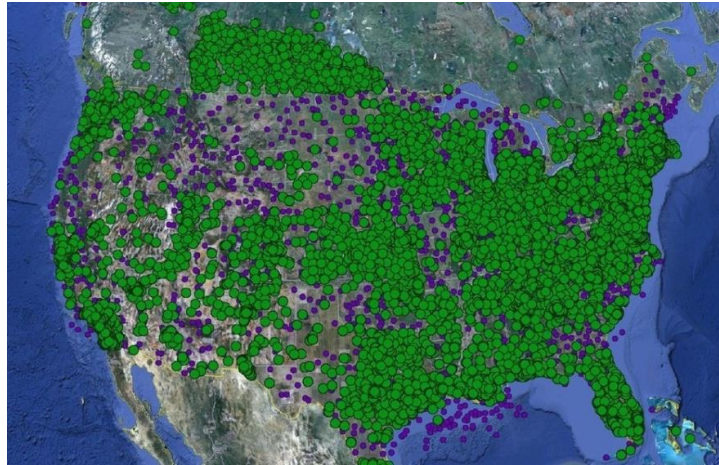
Figure 2-6 Weather Network and Sites of Unites States
Source: Weatherbug
 http://www.Weatherbug.com

Figure 2-7 Earth Network Project-Weather Site Distribution in the World
 Source: Weatherbug
 http://www.Weatherbug.com

Remote sensor and wireless sensor networks also collect necessary data based on the

current GIS system (Duan et al., 2011; Pultz et al., 2002), such as satellites. In 1998, Russia used

this type of satellite during severe spring floods and summer forest fires in Northern Russia

(Alexey et al., 1999). Furthermore, Radars and mobile devices also can be accessed in post-

disasters; so, it is important to study the mechanism of information sharing (IS) and simplicity

presentation method in the area affected by disaster (Urakami et al., 2009). Figure 2-8 shows a project on radar data collection for weather prediction system in Orlando, FL.

Contrary to this, Pultz and Scofield (2002) apply remotely sensed data for prediction and monitoring, and thus, his survey technique will be helpful for the data fusion in post-disaster management. The utilization of remote sensing data as input to flood management applications will help in managing disasters and their use will continue to be increased because of the increasing number of satellites and data providers. It is worth to consider that this will enhance accessibility to information, which in turn will augment the exploitation of remote sensing as input in flood management applications. The use of Earth observation satellites for managing flood disasters have been used by authors, along with the, flash flood analysis and prediction, as well as by the user community (Chen et al. 2009). The prominent aspect is the remote sensing management cycle that encompasses prevention where history, corporate memory, and climatology are vital, alleviation that protects life or property, pre-flood, that is the preparing and forecasting phase, where remote sensing is a necessity, a response that considers the actions and at last, a recovery that helps in assessing the damage (Kim & Choi, 2013).

Saha and Matsumoto (2002) (Saha et al., 2007) on the other hand suggest a different framework for post-disaster management and use wireless networks for disaster management. In this manner, they add significant data in post-disaster management by using hybrid networks of sensor networks and cellular in disaster management. Nevertheless, it is of due consideration that these base stations of cellular networks also get damaged during disasters as in the 1995 earthquake in Japan and in Nigatta in 1999, base station were not accessible. In light of these shortcomings, the technique suggested by the authors is an updated framework that will help in

collecting data from sensor networks and in managing disaster situations (Tsai et al., 2010; Tsai et al., 2011).

In order to disseminate, data from sensor networks during disasters Adhoc Relay Station (ARS) can be used in which cells are mainly divided into three parts. Disaster management requires a real-time efficient framework, and due to this an updated hybrid framework is essential. A wireless sensor network for disaster management (WSNDM) is vital as it disseminates the entry over the network by using its multi-hoping routing technique along with customized hybrid networks.

Singh, et al., (2008) put forth an imaging technique for managing natural disasters so that they can be mitigated or responded back in coming years. The technique they have used is synthetic aperture radar (SAR), which is capable enough to produce fine-resolution pictures of the earth terrain unimpeded by weather and illumination situations. Although observing natural disasters by using radar images is a difficult task, the authors in their research work have dealt with the scrutinizing and detecting of the alterations that takes place in subsidence because of natural disasters. In order to examine this technology, the City of New Orleans in the USA was considered as the testing area. It is pertinent to mention that on raw SAR data minimum ratio detector (MRD) is applied, while on complex data, D-InSAR is applied so that change can be observed. Different forms of satellite data sets were used for evaluation. The disaster information suggested by authors is highly significant in the post-disaster management because of the fact that it will help in sharing all forms of disaster information automatically, whether related to the number of injured or the number of people who died. This system will include the safety disaster information, as well as a mass of information in a wireless base satiation apparatus and this will be exchanged with the information stored in the other microcomputer. Therefore, the basic data

acquisition will be based on a variety of open-systems, although some stations or communication devices are destroyed by the impact of the disasters, the incomplete data can be recovered. Lu and Yang (Lu et al., 2011) reveal the characteristics of online social networks and then propose socio-technical design principles to address the communication challenges under uncertain emergency.
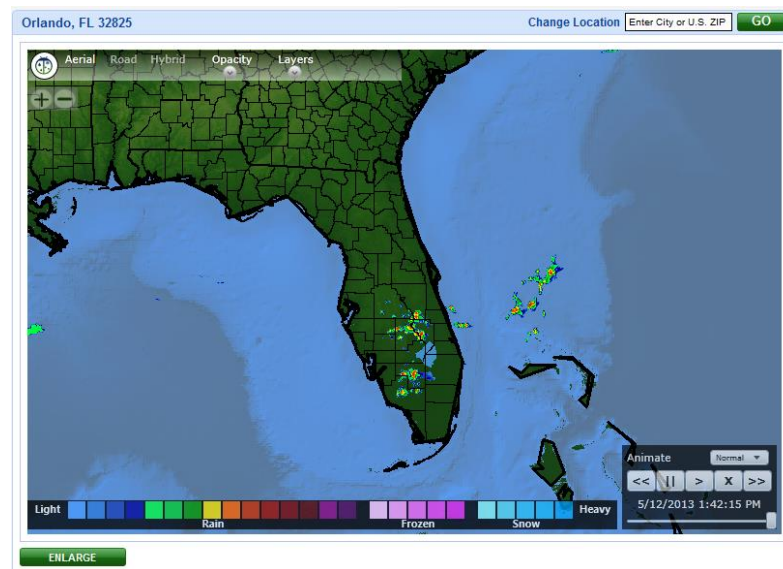
Figure 2-8 Radar Data Based Weather Prediction System
Source: NOAA, National Weather Service Radar Stations
http://radar.weather.gov/radar.php?rid=mlb

GIS dataset is a big data that we can acquire from public website or some open system such as Federal Emergency Management Agency (FEMA) and Google maps; the dataset from different sources needs to be preprocessed. As we know, what is of importance is not how to get the dataset but how to deal with those big dataset in an effective manner, such as how to scale the dataset under minimal loss. This is why we introduce data fusion technologies in Section 2.3.

2.3 Data Fusion Technologies in PDMS

The data of the PDMS comes from different sources and also appears in different modal, the data fusion technologies need to be applied in PDMS. Although the concept of data fusion began in the early 1970s, the real technical progress and development started in 1980's; with the increase in complexity in the PDMS, it is clearly considered limited to relying on a single data. Therefore, in the fault diagnosis system that using a multi-source technology will be able to monitor a range of features quantity making the data fusion operations improves the accuracy and reliability of the system. Earlier applications of PDMS ran on wearable PC's (Kim et al., 2008). Normally, data fusion technology is the use of computer chronological observation information, certain criteria is automatically analyzed and synthesized to complete the necessary decision-making, evaluation tasks, and information processing.

Data fusion in post-disaster management aids in analyzing the rich amount of data that different scholars have submitted. This survey will in turn be helpful in understanding the impact of disasters in regard to the physical, intellectual, emotional, psychological and other facets of human lives. Besides this, data fusion in post-disaster management will also submit distinctive approaches that authors have used to manage post-disaster situations and ways they have suggested to manage future disasters situations or areas that can be affected by disasters (Platz et al. 2007).

Data fusion acquires information from multiple sensors; it also can be observed as facts of information from multiple sensors. Under the inference engine, knowledge features and knowledge-base are matched to make the fault diagnosis decisions available to users. Fault diagnosis system based information fusion can be added to the self-learning module. The fault decision gives feedback to the knowledge-base for updates through the modification of the

27

corresponding confidence factor. At the same time, the dynamic response between knowledge-base and users' self-learning module creates a reasoning process. Expert system achieves self-learning functions, such as the ability to: acquire new knowledge, sum up new experiences, and constantly expand the knowledgebase. The three layers of fusions are as follows:

- Data level fusion: where to collect the original data, with the integration of the synthesis and analysis of the data in the original forecast of a variety of sources, without keeping it pretreatment. Data fusion generally uses centralized fusion system for the fusion process. This is a low-level of integration, such as imaging sensors, which is the integration of data layers containing a blurred image of a pixel image processing to confirm the target attribute.

- Feature level fusion: the middle level of integration; at first it takes feature extraction of the original information from the sources. Second, a comprehensive analysis of the characteristics of information is processed. The advantage of feature level fusion is to achieve a considerable compression by conducive real-time processing and decision analysis based on the extracted features; thus fusion results provide decision analysis with more elements of information. Feature level fusion is generally a distributed or centralized fusion system which can be divided into two categories: one is the integration of the target state; the other is the target characteristics of fusion.

- Decision level fusion: for observing the same goal by different types of sources. Basic processing of each source is done locally including preprocessing, feature extraction, recognition or judgment to establish the preliminary conclusions of the observed target,

and related processing decision fusion judgment and ultimately obtain the mutual outcome.

In our dissertation, we are going to apply a human-centric guided Gaussian fusion operation for data level fusion where Table 2-1 illustrates the methods of different level fusions.

Table 2-1 Levels Fusion and Relative Methods

Level	Data Level	Feature Level	Decision Level
Method	• Least Squares • Maximum Likelihood Estimation • Kalman Filter • Neural Networks	• Parameters based classification Statistics • Classical Inference, Bayesian Inference • D-S Theory • Information Theory Neural Networks • Clustering Analysis • Logical Template Methods • Recognition model based Voting • Fuzzy Set Voting	• Bayesian Inference • D-S Theory • Neural Networks • Fuzzy Logic

Using data level fusion, Christopher Stiller integrated various automotive sensors such as radar and video, to improve safety and traffic efficiency. Simultaneously, it is supported by Stiller, León and Kruse (2011) the causes of disasters can be divided into four major areas, noted as: by human blunders and technological malfunctions, by deliberate malevolence, by steps of nature (earthquakes, tornados, hurricanes, floods, and droughts) and by amalgamation of some or all the preceding ways. Disasters can be regarded as events of generally low probability and high outcome that results in the deterioration of life and property. However, post-disaster management helps in collecting data related to the loss of life and property during a disaster or the way it

impacts the society. Post-disaster management is a significant way of predicting, mitigating and recovering from disasters in some manner (Thompsona et al. 2011). Disasters are serious events that are of huge magnitude and pose a great threat to the society, as well as to the environment. It is considerable that any form of disaster brings adverse impact to the society, negatively affects the community, causes significant loss of human life and does a lot of material damage. Jotshi et al. (2009) applied a special data fusion technology to develop a robust methodology for the dispatching and routing of emergency vehicles (EVs) in a post-disaster environment. They also considered this in an earthquake scenario situation with a large number of casualties in need to obtain medical attention (Jotshi et al., 2009). J. Vega (2008) developed a pattern-based data retrieval method from fusion databases.

Considering the fusion techniques in a multisource problem, G. Nachouki introduced xml-based multi-source data description and developed a formalized language for online markets indices tracking system (Nachouki et al., 2008).

Some research focused on remote sensing field such as multi-sensor remote data fusion. B. Khaleghi introduced the remote sensing data fusion problems and its trends in remote sensing image fusion, which is still a very immature technology. He also recommends that the following key issues need to be further addressed:

- Space registration model
- Establishment of a unified mathematical fusion model
- Improvement of the accuracy of the data pre-processing process
- Improvement of the accuracy and credibility

With the development of computer technology, communication technology, new theories and methods, remote sensing image data fusion technology is transferred from theoretical

research to practical wider range of applications. It will eventually become intelligent and real-time direction development in conjunction with GIS, real-time dynamic fusion for updating and monitoring (Gigli et al., 2007; Khaleghi et al., 2013).

Y. Shkvarko, and B.V. Dasarathy, developed intelligent technology for end-user-oriented environmental resource management based on a quality of remote sensing images (Dasarathy B.V., 1998; Shkvarko et al., 2006).

Multi-source fusion technologies will be applied in this dissertation for data fusion aspects of PDMS, it includes the study of multi-source data fusion theory and methods of processing. This aspect is divided as follows:

- Hypothesis testing based data fusion: a data-based test using the principle choice of optimal data fusion to perform multi-sensor data hypothesis and obtain a comprehensive conclusion.

- Filtering target tracking based data fusion: to use Kalman filter by a single sensor using a detection network to a plurality of sensors, then a united Kalman filter algorithm performs the multi-sensor filter tracking processing.
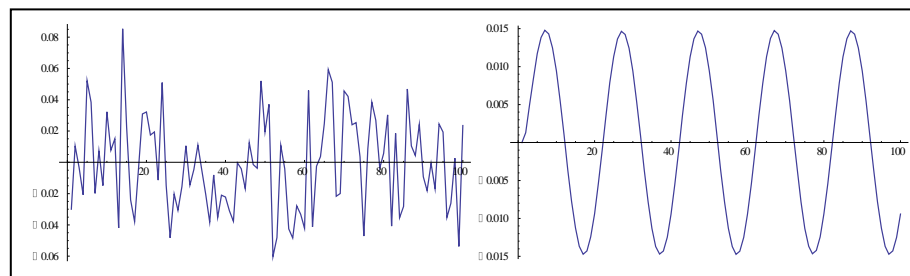
Figure 2-9 Kalman Filter of Dataset Sin[PI/10]

- Cluster analysis based data fusion: a statistical clustering analysis or principle of fuzzy clustering analysis on multi-objective, multi-sensor case with a large number of observation data samples, so that it can naturally aggregate data samples from the same target to different target.

- Pattern recognition based data fusion: the use of statistical pattern recognition or the principle of fuzzy pattern recognition, usually a single sensor pattern recognition criterion on the basis of multi-target multi-sensor pattern recognition with sentencing guidelines which establish minimum risk through fusion target classification and identification.

- Artificial intelligence based data fusion: uses artificial intelligence techniques to multi-sensor data fusion. It has a great advantage for imprecise and uncertain information processing, and therefore the direction of development of intelligent fusion method is divided into Expert System (ES) based and Artificial Neural Network (ANN) based. This intelligence fusion has metrics of a high degree of parallelism and nonlinear processing function, self-learning, self-organizing capacity, and distribution storage and fault tolerance (Saporta, 2002).

In our dissertation, we also used multimodal fusion techniques for PDMS dataset.

C. Rousseau introduced an intelligent multimodal presentation of information using several communication modalities to produce the most relevant user outputs, and proposed a conceptual model for intelligent multimodal presentation of information (Rousseau et al., 2006). B. Srinath Reddy introduced a semantic fusion method on different input modalities by using transferable belief models (Reddy et al., 2010). A multimodal crisis management system (XISM) was employed for processing of natural gesture and speech commands elicited by a user to efficiently manage complex dynamic emergency scenarios on a large display (Krahnstoever et

al., 2002). Multimodal fusion also was applied in PAD-based (Gilroy et al., 2009), support vector machines based hidden Markov models (SVMHMM) (Liu et al., 2009), multimodal function optimization (MMFO) by particle swarm optimization (PSO) (Li et al., 2012) and biometric issues (Poh et al., 2010a; Poh et al., 2010b). Xie et al, applied multimodal fusion on video search using a query-dependent fusion strategy, and G. Papandreou (Papandreou et al., 2009) introduced it for audiovisual speech recognition (Xie et al., 2007).

Homogeneous and non-homogeneous configurations for different data need more intelligence methods, such as Bayesian Fusion (BF) (Pearl, 1988), and Gaussian Fusion (GF) (Weckenmann et al., 2009).

To date, data fusion with Gaussian processes (GPs) has been widely applied for multimodal data processing. Gaussian model uses Gaussian probability density functions (normal distribution curve) to quantify things accurately, and decomposes the objective into a number of the model based on the Gaussian probability density functions (PDF) (Amir et al., 2003; Han et al., 2000; Hauptmann et al., 2004).

T. Deselaers presented a new technique based on support vector machines (SVMs) and Gaussian mixture densities (GMDs) to create a generative/discriminative object classification technique using local image features (Deselaers, 2010), and H.L. Kennedy (2012) introduced a probabilistic framework for fusing location estimates, using Gaussian mixture models. In measurement, for assigning an appropriate PDF to a certain kind of knowledge about a quantity, the principle of maximum information entropy (PME) was introduced (Jaynes E.T., 2003). A fuzzy multi-sensor data fusion based on the Kalman model was used to help reduce integrated vehicle health maintenance system (IVHMS) failure risk in the works of James A. Rodger (2012) and Sun et al. (2010).

We will follow the general steps of fusion process but will develop a Gaussian based fusion algorithm instead of PCA or Fuzzy operators. Based on the references we studied above, most works of the previous are based on the filtering, clustering and artificial methods such as ANN, so we will focus on the preprocess of dataset making each dataset following the Gaussian distribution that will benefit the process of the next steps.

2.4 Human-Centric Approach for Data Fusion

Human-centric Approach (HCA) rooted from user-oriented design and marketing and for advertising purposes; Romaniuk J introduced a model for HCA where it is shown in Figure 2-9 (Romaniuk et al., 2010). Data fusion is concerned not only with information collected from physical sensors, but also with insights of the human behavior. While traditional data fusion systems focus on observing physical targets by physical sensors, evolving applications are moving towards characterization of non-physical targets such as small groups, organizations, or cyber-attackers (Onwubiko et al., 2002). A human-centric integrated approach to web information search and sharing was introduced to estimate users' interests and to construct user profiles to reflect those interests in applying them for information acquisition in an online collaborative information seeking context (Shtykh et al., 2011).
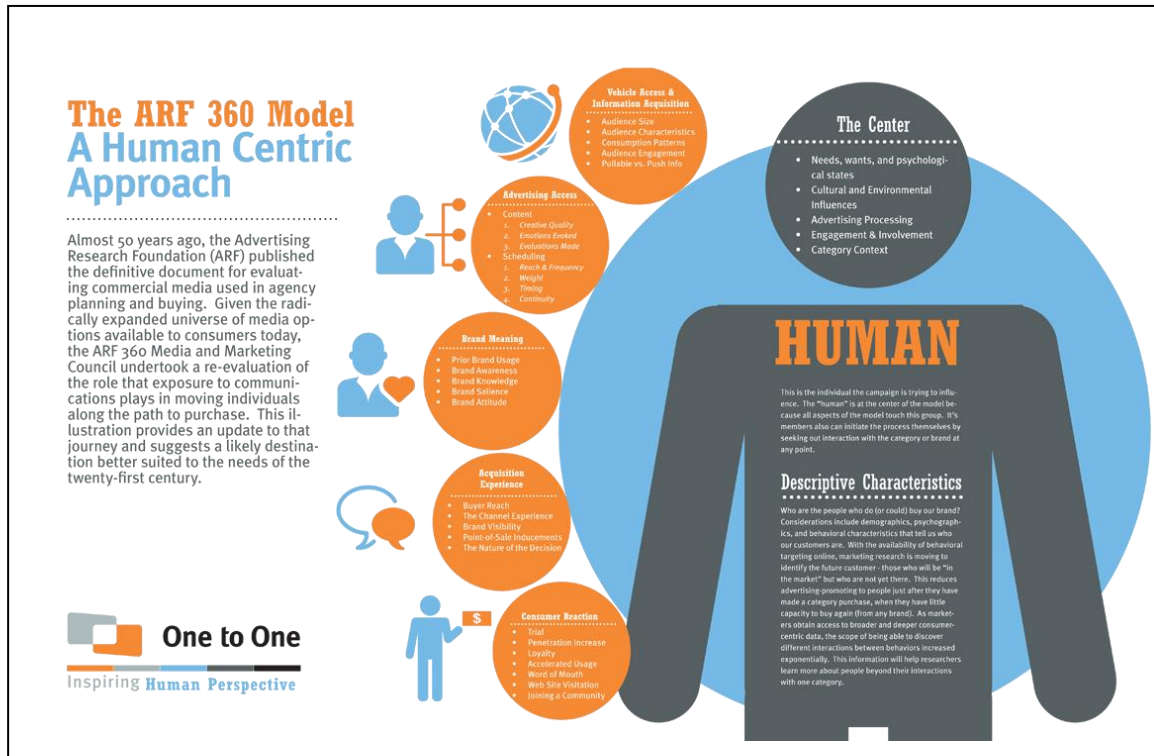
Figure 2-10 The Advertising Research Foundation 360-Model for HCA
Source: Advertising Research Foundation
http://m.thearf.org/news?newsid=208

Human-centric adaptation for multimodal interfaces process included two or more combined input modes, which is different with rule-based approaches. Kong (Kong et al., 2011) quantifies the preference of a modality under an interaction context and handles adaptation as an optimization problem. Specifically, a user can report a preference score of a modality based on their personal favors. The customized quantification makes the adaptation algorithm select modalities that fit user's personal needs. Lee (Lee et al., 2003) developed a human-centric interactive environment allowing designers to use natural gesture, voices, and advanced HCI technologies for a media-rich and comprehensive design presentation; where Zhu (Zhu et al., 2012) considered humans to be an essential element of the services computing of user factors. N. Ghadiri applied Computing with Words (CWW) for handling uncertainty human-centric

35

approach to group context-awareness (Ghadiri et al., 2011). On the other hand, HCA was also widely applied in product design and manufacturing (Mavrikios et al., 2007), program understanding (Buse, 2010) and geospatial data fusion (Levin et al., 2010). In our dissertation, we mainly focused on Human Factors (HF) which is related to negative factors such as information overload, confusion, and distraction data fusion layer with HF considerations (Atrey et al., 2010; Wagh et al. 2011). HF for data fusion was adopted in this research and factor analysis will be applied in factor operations. In which, the Principal component analysis (PCA) is a basic method for this issue, PCA is a mathematical procedure that uses an orthogonal transformation to convert a set of observations of possibly correlated variables into a set of values of linearly uncorrelated variables called principal components. The number of principal components is less than or equal to the number of original variables.

Factor analysis is a powerful tool for social studies, when study variable changes, the number of factors have to change. In addition, the interpretation of the actual meaning of each factor is not absolute. However, exploratory factor analysis (EFA) has some limitations. First, it assumes that all factors will affect the measure items (rotation). In the actual study, we tend to assume that there is no causal relationship between factors, it may not affect another factor measure items. Second, exploratory factor analysis assumes that the measure item residuals are independent of each other, but in fact, the residuals of the measure items because a single deviation, related sub-factors and other factors. Third, EFA will force all as an independent factor, although this is the plan for solving a number of factors. The most obvious is that the variables and the dependent variable should be related rather than independent. These limitations require a more flexible modeling approach. In exploratory factor analysis, a test model (such as orthogonal factor) is not the researcher's theory of exact model. Strengths of the confirmatory

36

factor analysis allow researchers to clearly describe the details of a theoretical model. Description of this relationship is also known as measure model which is a necessary step before hypothesis testing. Confirmatory factor analysis is often solved using maximum likelihood estimation (MLE) method.

Human-centric method is a new and promising method in product design. In our dissertation, we will apply this method in fusion process which is an innovation method for data fusion. Based on the research of the previous references, HCM will be feasible for data fusion and make the fusion process more reasonable for a GIS based PDMS.

2.5 Fuzzy Evaluation Model and Semi-Supervised Method

Fuzzy Set Theory (FST) was utilized in conjunction with classical methods to enhance the reliability and utility of disaster efforts while we focus on various forms of natural or man-made disasters (Esogbue, 1996), also there are some Fuzzy Set methods applied in PDMS, such as, fuzzy quantities (Facchinetti et al., 2006), fuzzy logic programming (Straccia et al., 2012), and fuzzy relation concepts for risk management (Iliadis, 2005). He and Chen, introduced a chaotic differential evolution algorithm to solve a fuzzy clustering iterative model for evaluating flood disaster (Chen et al., 2011; He et al., 2011), while Sun et al. (2008) and Zou et al. (2012) represented the diffused-interior-outer-set model (DIOSM) to obtain the possibility-probability distribution (PPD) for risk calculation.

As we know that Fuzzy evaluation model was applied in supervised learning process for PDMS and compared with non-supervised process, supervised or semi-supervised process, it will result in much more accurate results through evaluation the outputs and continue to adjust the inputs.

37

On the other hand, unsupervised learning is based on training samples of the category unknown (not marked) to solve various problems in pattern recognition, such as clustering analysis. Sometimes, the lack of prior knowledge is difficult to manually label the category while artificial category label is too costly. Naturally, we hope that the computer on our behalf (partially) will help us to complete these tasks. These works include selecting some representative samples of the classifier from a huge collection of data (Leng et al., 2013; Lin et al., 1996; Nojun et al., 1999). First of all, samples automatically are divided into different categories marked by humans. The most common unsupervised learning based automatic classification is calculated based on the similarity.

The similarity measure also can be defined based on human experience. By clustering objective function of the squared error, such as various types of samples to the class mean vector distance and minimum variance criteria, it can be divided into:

- K-means algorithm

- Fuzzy K-means algorithm (K-means variant 1)

- Iteration K-means algorithm (K-means variant 2)

- Consolidation Act (also known as clustering)

- Secession Law (also known as decomposition clustering)

The earliest use of the notion of semi-supervised learning algorithm should be self-training (Chapelle et al., 2006), this also appeared in some of the literature of the 1960s and 1970s (eg, 1965 Scudder literature) (Scudder, 1995), 1967 the Fralick literature (Fralick, 1967), and 1970 Agrawala literature (Agrawala, 1970).

Since the basic idea of the training methods is to use supervised learning techniques to learn the marked data, unlabeled data and then use the result of learning to get it marked, then the

new tag data is added to the marked data learning process. However, the performance of this method depends on the supervised learning technology, and when to use the 0-1 cost function empirical risk minimization learning, unlabeled samples will lapse (Scudder, 1965).

Semi-supervised learning was widely used in knowledge discovery (Klose et al., 2005), human action recognition (Zhang et al., 2011), kernel density estimation (Du et al., 2013), fuzzy clustering algorithm (FCA) (Grira et al., 2008; Yin et al., 2012), neural networks (Chen et al., 1997), and Bayesian network (Chakraborty, 2011).

Chang et al. (2012) introduced the two-teachers-one-student (2T1S) method for SSL which is a multi-view approach blending the concepts of co-training and consensus training. Based on the reduced support vector machine (RSVM), which is different from other existing approaches, they selected multi-view in the represented kernel feature space rather than in the input space (Lee et al., 2001; Lee et al., 2007). Tian et al. (2012) investigated the benefit of combining both cluster assumption and manifold assumption underlying most of the semi-supervised algorithms using the flexibility and the efficiency of multiple kernel learning.

Following a brief introduction of semi-supervised learning are two common assumptions. The clustering hypothesis refers to sample point in the same cluster are likely to have the same category tags. This assumption can be expressed by another equivalent, that is, through the decision boundary regions, it will be a relatively sparse area of data points, and if the decision boundary through the data points is in a relatively dense area, then it is likely to be sample points in a cluster where it is divided into different categories and clustering contradicts the assumption.

Flow assumption is the existence of low-dimensional characteristics of high-dimensional data and representation of another species is an example of a small local neighborhood having similar properties (Zhou, 2007). No strict proof of the equivalence of both high-dimensional data

39

in low-dimensional characteristics of SSL. Data is reflected through the local neighborhood similarity, such as a two-dimensional tape in a high-dimensional data global distance metric is no due to excessive dimensions, but for a partial range of the distance metric, there will be a certain sense. The difference is that SSL method must have the training set and test samples. In the training set to find the law and the test sample using this law, rather than monitor learning training set, only one set of data in the data sets to find the law. The purpose of supervised learning method is to identify things; the result of recognition performance is identified data with labels. Training sample set must be composed by a labeled sample, rather than SSL methods which only analyze the data set itself. If data sets show some kind of aggregation, it can be classified according to the nature of the aggregation, but it is not with a pre-classification label for this purpose (Pang et al., 2009; Weng et al., 2008; Zhang et al., 2010).

Semi-supervised learning combined with fuzzy set model was adopted in this dissertation. Some fuzzy set model based algorithms for semi-supervised learning were developed and improved by previous scholars, such as, fuzzy Petri net (Konar et al., 2005), and active fuzzy constrained clustering (AFCC) (Maraziotis, 2012; Waegeman et al., 2011). Yan et al. (2013) and Xue et al. (2010) proposed a new heuristic semi-supervised fuzzy co-clustering algorithm (SS-HFCR) for categorization of large web documents and fuzzy rough semi-supervised outlier detection (FRSSOD) approach with the help of some labeled samples and fuzzy rough C-means clustering. These methods will benefit the development of fuzzy set model of SSL the PDMS.

From the above research, fuzzy evaluation makes the knowledge presentation more exact, but the shortcomings of fuzzy evaluation are in the complexity of calculation. So a fuzzy inference system needs to be developed to reduce the complexity. Combining with SSL for

PDMS is an innovative method that will make the PDMS provide more effective information for decision-making of individuals in post-disaster.

CHAPTER 3: METHODOLOGY

3.1 Overview

This innovative approach includes two parts; the first is a Human Factors (HF) dominated

fusion algorithm. As we know the data processing method of PDMS in this dissertation is multi-

source and multi-modal, so HF based Gaussian fusion needs to be developed. The second is the

fuzzy sets based semi-supervised algorithm for PDMS model; we included and compared various

PDMS systems' knowledge presentation, and clarify the efficiency of the theoretical model of

semi-supervised algorithm combining with fuzzy set model and rule-based system. Related

analysis will be complete in this chapter. Multi-source, multi-mode data pre-processing and

fuzzy set model in semi-supervised model are also crucial. The framework of this dissertation is

illustrated in Figure 3-1.
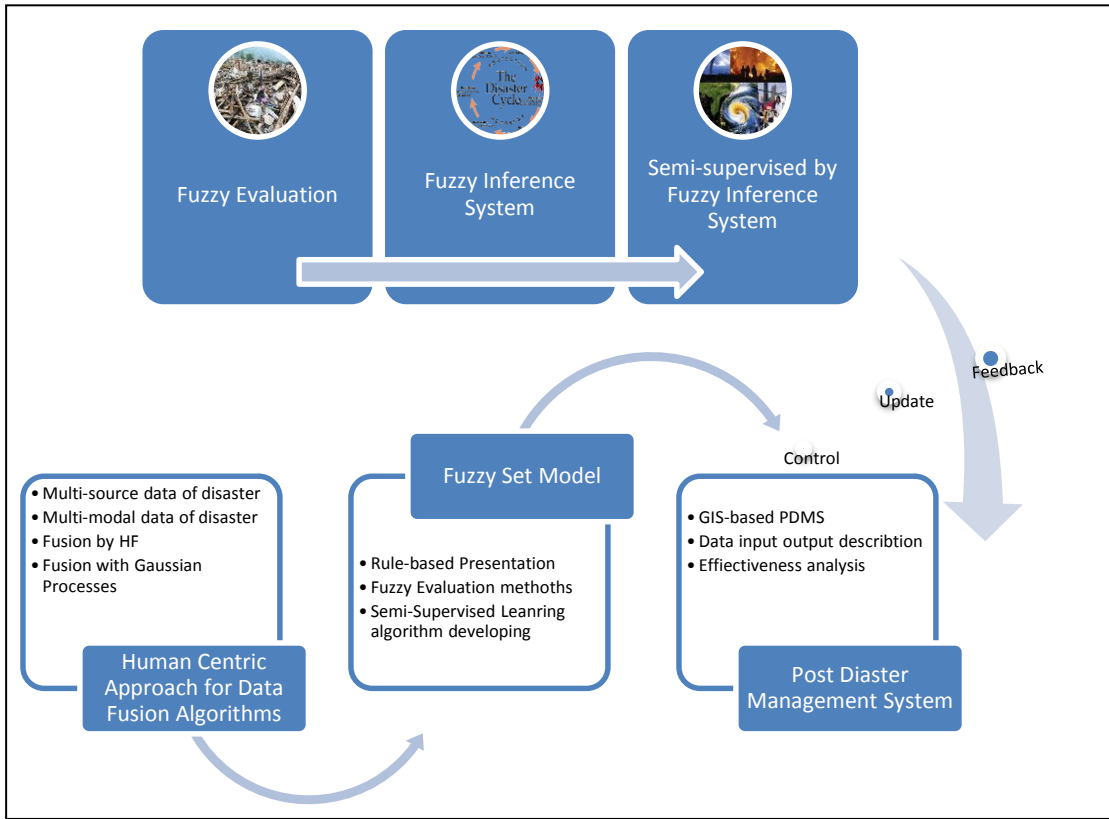
Figure 3-1 The Research Framework of the Proposed Models in Dissertation
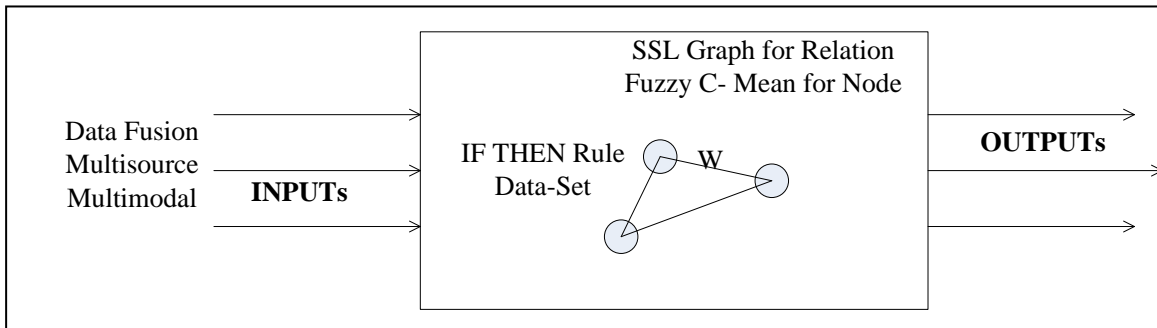
Figure 3-2 Algorithms for Model of the PDMS in Dissertation

Data fusion is an integration method on data, before this data is input into the system, we need to simplify it. Based on different type of data, different fusion algorithms need to be

developed, we developed a multisource data fusion algorithm based Gaussian model and improved super kernel based multi modal fusion algorithm.

The PDMS is organized as a graph with each node presented by IF THEN rules that is a form of knowledge presentation. The node's presentation will be clustered by semi-supervised fuzzy c-mean algorithm and the graph is organized by the SSL graph special algorithm that we improved. The supervised algorithm is a feedback system; the parameters of the system can be adjusted based on the evaluation of the OUTPUTs. This is why we call this PDMS dynamic and can provide more effective information to individuals in post-disaster situations. Through this system, the individual can receive a "quick response" from PDMS.

3.2 Definitions and Problem Description

3.2.1 *Definitions of PDMS Formalization*

The critical feature of GIS-based PDMS, in essence, is the calculation on diagram. This dissertation is based on such an assumption. The disaster supply resources are listed on the map, and it was also assumed that the node can act as a resource provider, and the path is for the transport value. The node is presented by a vector called resource vector. Figure 3-3 and 3-4 show the circumstance, and all resource nodes are presented as a matrix noted by N, and transportation value as a weight matrix W.
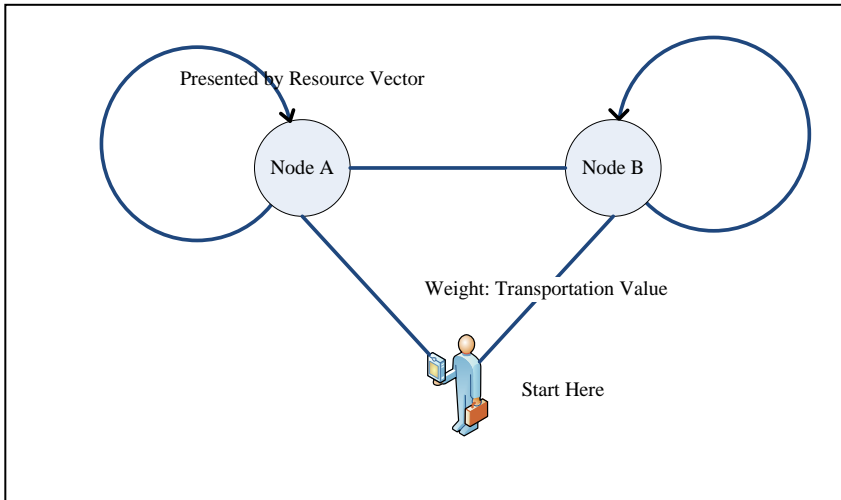
Figure 3-3 Resource Nodes and Their Values on Map
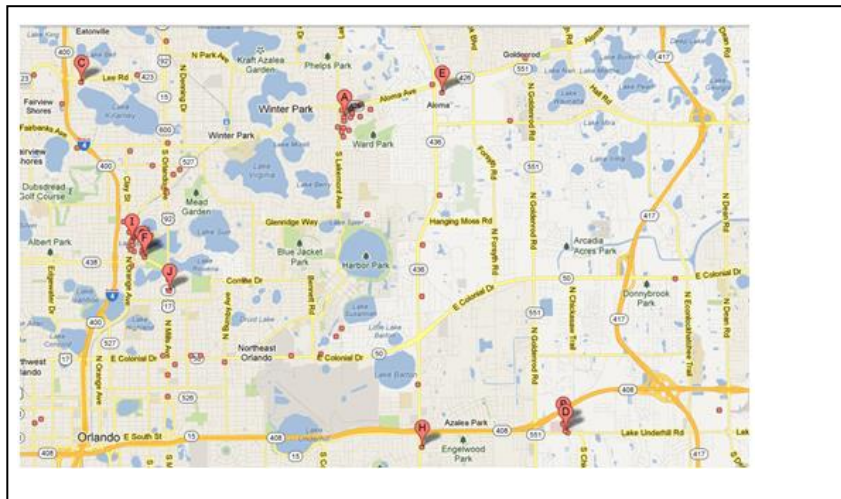
Figure 3-4 Hospitals Listed for a Graph and Presented as Vectors
Source: Google Map Tool
https://support.google.com/maps/

By mathematical description, adjacent matrix can be used for a matrix to represent various points and the relationships of each edge.

Definition 3.1: For undirected graph, $G(p \times q)$, p is the number of nodes, q is the number of edges. b_{ij} denotes the relationship of the node i and edge j in adjacent matrix, where as if the node i and edge j are connected, then $b_{ij} = 1$ or else $b_{ij} = 0$. Figure 3-4 shows the

45

undirected graph and its adjacent matrix. We can get location from map and convert it to a map (directed or undirected).
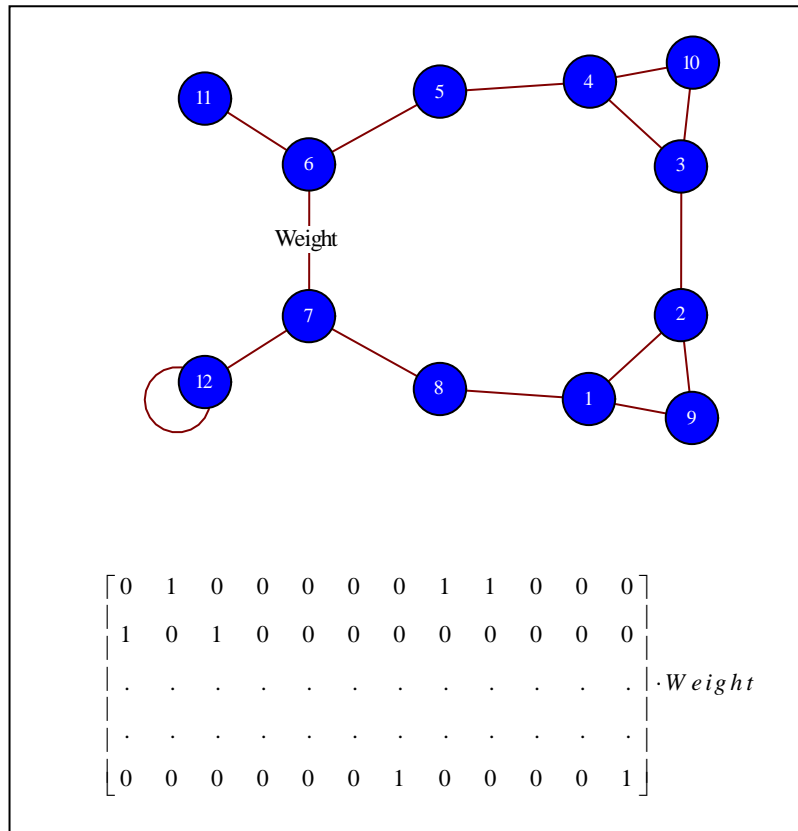
Figure 3-5 Undirected Graph and its Adjacent Matrix

Definition 3.2: For a directed graph, if $b_{ij} = 1$, it indicates that edge j left from node i. If $b_{ij} = -1$, then it means that edge j entry node i. If $b_{ij} = 0$, then it means that j and i are not associated (Shown in Figure 3-5).
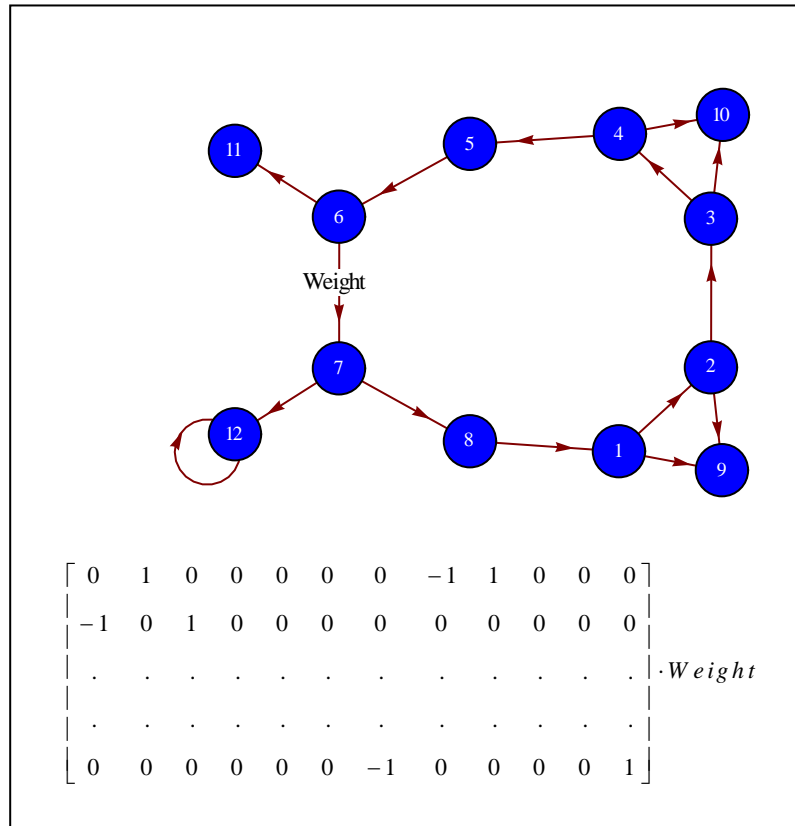
Figure 3-6 Directed Graph and its Adjacent Matrix

The key issue of adjacent matrix is to determine the relative importance of each

evaluation index degree, weight- w_j , and given the subject based on the evaluation of the

evaluation scale to determine the value of program estimation assessed on the amount.

Adjacent matrix is named the entire program as a matrix array, which is applied widely

in order to get multi-target system solutions from a comprehensive assessment of the merits of a

number of factors starting level approach. It is a combination of quantitative and qualitative

evaluation methods, which is represented in matrix form evaluation of each alternative on the

evaluation value, and then the weighted sum of the value of program evaluation is calculated to

determine that the weighted evaluation values and the largest program is the optimal solution.

Adjacent matrix of the basic starting point is to establish a hierarchy of evaluation and analysis in determining the weights, and whether the correlation matrix can be more simple and workable. It is based on specific evaluation system that uses a matrix to determine the system evaluation and its corresponding weight, and then evaluate the system's various programs to calculate the comprehensive evaluation value - the value of each evaluation project of weighted sum.

Adjacent matrix method is characterized to make it easy to accept the evaluation of complex system problems by mathematical thinking process, through the multi-objective problem into two indicators comparing the degree of importance, so that the evaluation process is simplified and clear. This method lies in determining the weight of each index and w_i given by the evaluation of the main yardstick for evaluation to determine the value of program evaluation assessed on the amount of v_{ij}.

The general steps are as follows:

*Step 1***:** Determine the indicators

Content indexing is assessed quantitatively by the basic requirements. Evaluation index system is a hierarchical structure. General assessment scale by the two to three levels of indicators form as module indicator (i.e., assessment scale different options module content can be different based on the evaluation contents coverage differences, index module may also need to be divided into different modules), level 1 index - Project Indicators, and level-2 indicators, which is an indicator module further subdivided derived from level 1 index.

Step 2: Determine the weight system

The indicator system of the various indicators for the program (evaluation body) is different, and this difference in the degree of importance of each indicator needs to assign different weights to reflect it as such. A set of evaluation index weights corresponding to the reorganization of the system of weights has become more important.

Suppose that any set of weights $w_i (i = 1, 2, \cdots n)$ system must satisfy the following two conditions:

$$0 < w_i \leq 1 (i = 1, 2, \cdots, n)$$
$$\sum_{i=1}^{n} w_i = 1 \tag{3.1}$$

And for more levels index, we have that

$$\sum_{i=1}^{n} \sum_{j=1}^{m} w_i w_j = 1 [0,1] \tag{3.2}$$

Step 3: Individual evaluation

Usually uses the following two methods:

(1) Expert assessment method: the expert scores by the arithmetic mean value while removing the lowest score and the highest score.

(2) Letter of inquiry Delphi method: the use of expert knowledge and long experience to reduce the impact of the authority.

Step 4: Comprehensive assessment

In a layer of the index system, the assessor is that if the corresponding system of weights $w_{ij} (i = 1, 2, \cdots n, j = 1, 2, \cdots m)$ satisfies:

(1) $\quad 0 \le w_{ij} \le 1, i = 1, 2, \cdots n, \ j = 1, 2, \cdots m$

(2) $\displaystyle \sum_{j=1}^{m} w_{ij} = 1$

(3) $\displaystyle v_i = \sum_{j=1}^{n} w_i \cdot v_{ij}$

, where v_i presents the comprehensive evaluation, v_{ij} presents individual evaluation, and

w_j presents its weight.

In level 2 index system, the evaluation value is calculated by

$$v_i = \sum_{k=1}^{n} w_k \cdot \sum_{j=1}^{m} w_{ij} \cdot v_{ij} \qquad (3.3)$$

We listed the corresponding correlation matrix as follows:

Table 3-1 Levels Evaluation Value Calculation by Weights

	x_1	x_2	...	x_n	v_i
	w_1	w_2	...	w_n	
a_1	v_{11}	v_{12}	...	v_{1n}	$v_1 = \sum_{j=1}^{n} v_{1j} \cdot w_j$
a_2	v_{21}	v_{22}	...	v_{2n}	$v_2 = \sum_{j=1}^{n} v_{2j} \cdot w_j$
...	...	...	...		
a_i	v_{i1}	v_{i2}	...	v_{in}	$v_n = \sum_{j=1}^{n} v_{nj} \cdot w_j$

3.2.2 *Factors Analysis for Resource Matrix*

Principal component analysis (PCA) was used for a human-centric approach and converted as Human Factors (HFs) after vectoring operators of PDMS system. It is a statistical analysis method of using more variables by linear transformation to elect a smaller number of key variables multivariate. In order to fully analyze the problem, it often puts forward many variables associated with factors and each variable is in varying degrees; it reflects some of the information on this topic. PCA was first introduced by K. Pearson (Pearson, 1901) for non-random variables, and later H. Hotelling (Hotelling, 1935) extended this method to the case of random vectors. The size information used is usually squared deviations or variances to be measured; we have the following steps for HF analysis in PDMS:

Step 1: The composition of the sample matrix X. Suppose that there are m variables to be analyzed, in fact, it is equivalent to an m-dimensional data in the coordinates of each dimension. Our goal is to ensure the similarity between comparative data undistorted premise and to describe the data to minimize the dimensions of $l(l < m)$. Matrix $X = [X_1, X_2, \cdots X_n]$ is of n data (these data are in the form of a column vector), so $X = [X_1, X_2, \cdots X_n]_{m \times n}$.

Step 2: Calculate the mean of sample x of the m-dimensional mean as follows:

$$u(i) = \frac{1}{n} \sum_{j=1}^{n} X[i, j] \tag{3.4}$$

Step 3: Calculate the deviation of the mean and the observed value, in each dimension, with the current value of $X[i, j]$ and minus $u(i)$, a matrix operation is as follows:

$$\begin{aligned} B &= X - uh \\ h[j] &= 1 \quad for \quad j = 1, 2, \cdots, n \end{aligned} \tag{3.5}$$

Step 4: Calculate the covariance matrix, b_i represents the i-th row of B, then the covariance matrix is,

$$C = E[B \otimes B] = E[B \cdot B^*] = \frac{1}{n} \sum B \cdot B^* \tag{3.6}$$

And then we have that,

$$c[i, j] = \frac{1}{n} < b_i, b_j > \tag{3.7}$$

Step 5: Compute the covariance matrix C 's eigenvalues and eigenvectors.

Step 6: Sort. The eigenvalues are sorted descend, and according to the eigenvalues to adjustment eigenvectors arrangement.

Step 7: Calculate the total energy and select the greatest one.

If $V \cdot$ is a diagonal matrix of C , then the total energy is the sum of all the eigenvalues on the diagonal, noted S . Refer to the Step 6, V has been reordered, so while the previous eigenvalues of $V \cdot$ are greater than or equal to 90% of S , that these eigenvalues can be used to "characterize" the current matrix, assume the characteristic values of l .

Step 8: Calculate basis-vector matrix W , in fact, W is the first L columns of the matrix V , so W is the size of $m \times l$.

Step 9: Calculate z -scores (optional)

$$z[i, j] = \frac{B[i, j]}{\sqrt{D[i, j]}} \tag{3.8}$$

Step 10: Calculate the new sample for dimension reduction matrix,

$$Y = W^* \cdot Z = KLT\{X\} \tag{3.9}$$

W^* represents the conjugate transpose matrix W , the size is $l \times m$, and Z is $m \times n$, so the size of Y is $l \times n$ that dimension reduction of the n for l -dimensional vectors was finished.

3.3 Multisource Fusion Under a Human-Centric Approach

Referring to Roux et al. (1995), Ye et al. (2009), Mao et al. (2010) and Peng et al., (2012), we developed a new model for multisource data fusion. The new fusion model developed a

probabilistic density function (PDF) based Gaussian fusion employing HFs, whereas HFs will be assigned as a probabilistic density.

HF is initialed as fitness between the user, equipment and their environments, and also is the scientific application of knowledge about the capacities and limitations of users with the aim of making products, systems, services and environments safe, efficient and easy to use; while growing the complexity of telecommunications services and equipment makes the human element more important. HF is a key factor for the commercial success of any telecommunications product or service. In this dissertation, HF is used as an evaluation on dataset for data fusion of PDMS, which guided the fusion results under HFs' value. Multisource or multimodal data have different   HF values; it depends on the questionnaire system or fields expert and is defined as,

Definition 3.3: For source data X , $HF = \delta(X)$ has the following properties,

(1) $0 \leq \delta(x) \leq 1$

(2) $\delta(X) = 1$

(3) $\delta(X_1 \cup X_2) = \delta(X_1) + \delta(X_2)$, if $X_1 \cap X_2 = \varnothing$

54
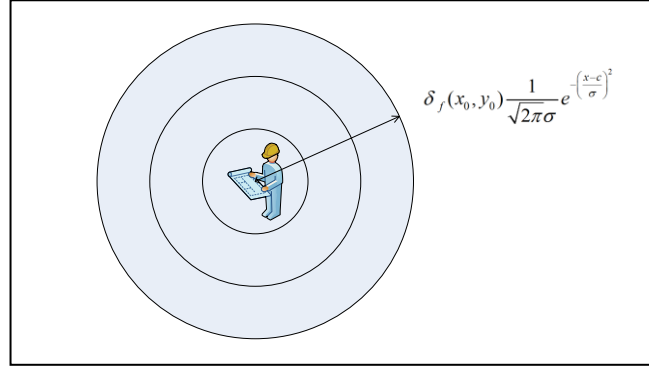
Figure 3-7 Human Factors in Gaussian Fusion Process

3.3.1 *Multisource Data Fusion Single Gaussian Model*

Given a data has a δ defined by Definition (1), as we know that δ is a density function, it reflects how concentrated the data is. So, we defined δ as the data - x 's PDF. In Gaussian fusion processing, we have that,

$$f(x,\mu,\sigma^2) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \tag{3.10}$$

The parameter μ is the mean or expectation of data - $x \in X$, and σ^2 is the deviation. SGM is applied to induct the density function of the proposed data X and we define,

$$\delta(X,\mu,\Delta) = \frac{1}{\sqrt{(2\pi)^3 |\Delta|}} e^{-\frac{1}{2}(X-\mu)^T \Delta^{-1}(X-\mu)} \tag{3.11}$$
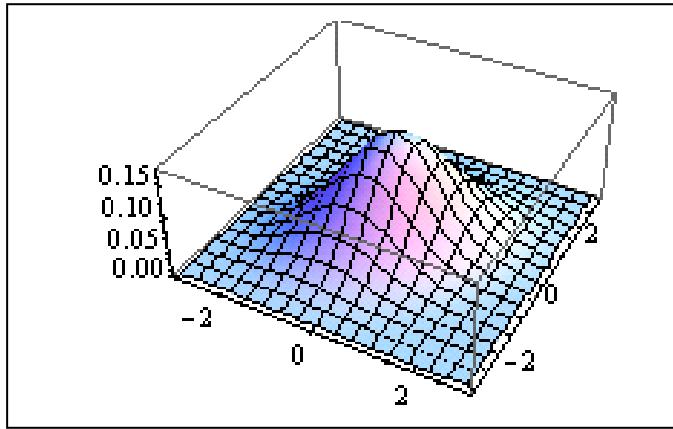
Figure 3-8 Gaussian Model in R^2

, where $X \in R^n$ represents data, Δ is the n-dimensional covariance matrix, and μ is the mean value of the density function δ , it is relative to HF. As we know, the density function's properties are determined by (Δ, μ) , so this is a parameter estimation problem (Denoeux, 2011). For any point $P_i \in R^n$, its PDF is $\delta(P_i, \mu, \Delta)$, and if, for any data x_k , each x_k is regarded as an independent event, then the PDF of x_k is:

$$\delta_k = \delta(x_k, \mu, \Delta) = \prod_i \delta(P_i, \mu, \Delta) \tag{3.12}$$

Maximum likelihood estimation can be used to estimate the parameters (Δ, μ) under (3.12). Taking the logarithm of (3.12), we have:

$$O(\mu,\Delta) = \ln(\prod_i \delta(x_i, \mu, \Delta))$$

$$= \sum_i \ln(\delta(x_i, \mu, \Delta))$$

$$= \sum_i [-\frac{3}{2}\ln(2\pi) - \frac{1}{2}\ln|\Delta| + \frac{1}{2}(x_i - \mu)^T \Delta^{-1}(x_i - \mu)$$

$$= -\frac{3n}{2}\ln(2\pi) - \frac{n}{2}\ln|\Delta| - \frac{1}{2}\sum_i [x_i - \mu)^T \Delta^{-1}(x_i - \mu)]$$

Taking the partial derivative w.r.t. μ of $O(\mu,\Delta)$ and setting it to 0, we obtain:

$$\partial_\mu(O(\mu,\Delta)) = -\frac{1}{2}\sum_i [-2\Delta^{-1}(x_i - \mu)] = \Delta^{-1}\sum_i [(x_i - \mu)] = \Delta^{-1}\sum_i x_i - n\mu = 0 \tag{3.13}$$

This gives $\hat{\mu} = \frac{1}{2}\sum_i x_i$. Similarly, for Δ, we can obtain $\hat{\Delta} = \frac{1}{n-1}\sum_i (x_i - \hat{\mu})(x_i - \hat{\mu})^T$.

Thus, if the density of each point in x_k is $\delta(P, \hat{\mu}, \hat{\Delta})$, our estimation of the parameter μ is:

$$\hat{\mu} = (\frac{1}{n}\sum_i e_{1i}, \frac{1}{n}\sum_i e_{2i}, \cdots \frac{1}{n}\sum_i e_{ni}) \tag{3.14}$$

, where e_{li} is the coordinate of x_i in R^n.

The covariance $\hat{\Delta}$ is converted to

$$\hat{\Delta} = \frac{1}{n-1}\sum_i [e_{1i} - \hat{\mu}_1, e_{2i} - \hat{\mu}_2, \cdots, e_{ni} - \hat{\mu}_n] \begin{bmatrix} e_{1i} - \hat{\mu}_1 \\ e_{2i} - \hat{\mu}_2 \\ \cdots \\ e_{ni} - \hat{\mu}_n \end{bmatrix} = \frac{1}{n-1}\sum_{j=1}^{n}\sum_{i=1}^{n}(e_{ji} - \hat{\mu}_j)^2 \tag{3.15}$$

If $n = 1$, the density appears as a typical one-dimensional Gaussian distribution (see Fig. 3-8).
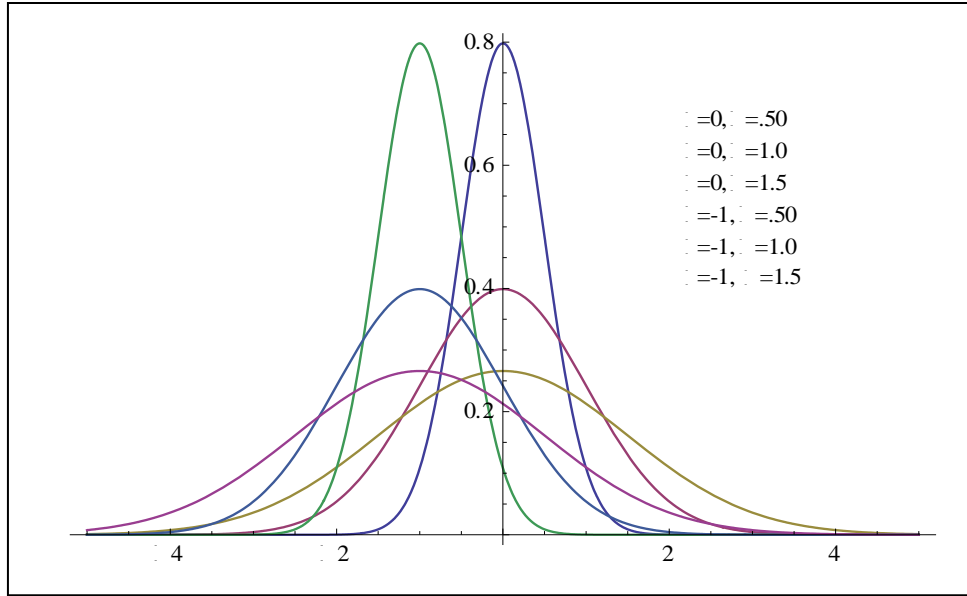
Figure 3-9 One-Dimensional Gaussian Density Functions

If $n = 2$, then the PDF is a two-variable Gaussian density function. The calculation of the

covariance matrix and the correlation coefficient r is extremely complex; therefore, suppose that

there is no correlation between the two information structures. A single Gaussian density

function is calculated recursively as follows:

(1) If x_1, x_2 have the same mean μ and standard deviation σ with factors weights,

then we find that,

$$\delta_{(x_1, x_2)} = \frac{HF}{\sqrt{2\pi}\,\sigma\,\sqrt{(1-r^2)}}\, e^{-\frac{1}{2}[(\frac{(x_1-\mu)^2 - 2r(x_1-\mu)(x_2-\mu) + (x_2-\mu)^2}{2\sigma^2})]} = \frac{HF}{\sqrt{2\pi}\,\sigma}\, e^{-\frac{1}{2}(\frac{(x_1-\mu)^2 + (x_2-\mu)^2}{2\sigma^2})} \qquad (3.16)$$

58

(2) If the mean and standard deviation are different, then the joint density function is defined as,

$$\delta_{(x_1,x_2)} = \frac{HF}{\sqrt{2\pi\sigma_1\sigma_2}\sqrt{(1-r^2)}} e^{-\frac{1}{2}[(\frac{(x_1-\mu_1)^2-2r(x_1-\mu_1)(x_2-\mu_2)+(x_2-\mu_2)^2}{\sigma_1^2+\sigma_2^2})]} = \frac{HF}{\sqrt{2\pi\sigma_1\sigma_2}} e^{-\frac{1}{2}(\frac{(x_1-\mu_1)^2+(x_2-\mu_2)^2}{\sigma_1^2+\sigma_2^2})} \quad (3.17)$$

3.3.2 *Multisource Data Fusion by Gaussian Mixed Model and PE*

Gaussian mixed model (GMM) is among the most statistically mature methods of clustering for density estimation (Shi, 2012). S. Vasudevan (2012) used multiple sets of heterogeneous sensor data by Gaussian processes (GPs) and presented an experiment on large scale terrain modeling in mining automation. Additionally Gaussian sum probability hypothesis density (GSPHD) filter by Yin, et al. (2008), Bayesian Gaussian neural network by Ye et al. (1999) and applications on speak recognition by R. Kumari (Kumari et al., 2012) were referenced in this section.

For multisource data fusion, we need to calculate all of data- x_k 's density functions and calculate the new density function. For m multisource data, let $\delta_{fusion} = \sum_{i=1}^{m} \alpha_i \delta(X, \mu_i, \Delta_i)$, for a normalized weight parameter α that $\sum_{i} \alpha_i = 1$. And $\alpha_i = \delta_f(x_{0i}, y_{0i})$, where is (x_{0i}, y_{0i}) the i-th factors To calculate and simplify the covariance matrix Δ, let

$$\Delta = \begin{bmatrix} \sigma^2 & 0 & \cdots & 0 \\ 0 & \sigma^2 & \cdots & 0 \\ 0 & \cdots & \cdots & 0 \\ 0 & 0 & \cdots & \sigma^2 \end{bmatrix} = \sigma^2 \vec{I} \quad (3.18)$$

59

From the SGM, we have that

$$\delta(X, \mu, \sigma^2 \vec{I}) = \frac{1}{\sqrt{(2\pi)^3}} \sigma^{-1} e^{-\frac{(X-\mu)^T(X-\mu)}{2\sigma^2}} \tag{3.19}$$

Calculate:

$$\partial_\mu[\delta(X, \mu, \sigma^2 \vec{I})] = \frac{1}{\sqrt{(2\pi)^3}} \partial_\mu (\sigma^{-1} e^{-\frac{(X-\mu)^T(X-\mu)}{2\sigma^2}})$$

$$= \frac{1}{\sqrt{(2\pi)^3}} \sigma^{-1} e^{-\frac{(X-\mu)^T(X-\mu)}{2\sigma^2}} \partial_\mu (-\frac{(X-\mu)^T(X-\mu)}{2\sigma^2})) \tag{3.20}$$

$$= \delta(X, \mu, \sigma^2 \vec{I})(\frac{X-\mu}{\sigma^2})$$

and

$$\partial_\Delta[\delta(X, \mu, \sigma^2 \vec{I})] = \frac{1}{\sqrt{(2\pi)^3}} ((-1)\sigma^{-2} e^{-\frac{(X-\mu)^T(X-\mu)}{2\sigma^2}}) + \frac{1}{\sqrt{(2\pi)^3}} \sigma^{-1} e^{-\frac{(X-\mu)^T(X-\mu)}{2\sigma^2}} [\frac{(X-\mu)^T(X-\mu)}{\sigma^3}]$$

$$= \delta(X, \mu, \sigma^2 \vec{I})(\frac{(X-\mu)^T(X-\mu)}{\sigma^3} - \frac{1}{\sigma^2})$$

$$\tag{3.21}$$

Then, for $\Delta = c\vec{I}$, $c \in R$, GMM is defined as $X_{fusion} = \sum_i \alpha_i \delta(X, \mu_i, \sigma_i)$, $i = 1, 2, \cdots, m$.

The number of parameters for estimation is $3m$. If we let

$\theta = [\alpha_1, \alpha_2, \cdots, \alpha_m, \mu_1, \mu_2, \cdots \mu_m, \sigma_1^2, \sigma_2^2, \cdots, \sigma_m^2]$, the object is that,

$$L(\theta) = \ln[\prod_i X_{fusion}] = \sum_i \ln(X_{fuson})) = \sum_i \ln(\sum_{j=1}^m \alpha_1 \delta(x_i, \mu_j, \sigma_j^2)) \tag{3.22}$$

, which can be differentiated w.r.t. μ_j and σ_j. Thus, we have that:

60

$$\partial_{\mu_j}(L(\theta)) = \sum_i \frac{\alpha_j \delta(x_i, \mu_j, \sigma_j^2)}{\sum_{j=1}^{3} \alpha_i \delta(x_i, \mu_j, \sigma_j^2)} \frac{x_i - \mu_j}{\sigma_j^2} \tag{3.23}$$

Let $\varphi_j(P_i) = \dfrac{\alpha_j \delta(x_i, \mu_j, \sigma_j^2)}{\sum_{j=1}^{3} \alpha_j \delta(x_i, \mu_j, \sigma_j^2)}$, and we have that,

$$\partial_{\mu_j}(L(\theta)) = \sum_i \varphi_j(x_i)(\frac{x_i - \mu_j}{\sigma_j^2}) \tag{3.24}$$

Continuously, we have that,

$$\partial_{\sigma_j}(L(\theta)) = \sum_i \frac{\alpha_j \delta(x_i, \mu_j, \sigma_j^2)}{\sum_{j=1}^{3} \alpha_i \delta(x_i, \mu_j, \sigma_j^2)} [\frac{(x_i - \mu_j)^T (x_i - \mu_j)^T}{\sigma_j^3} - \frac{1}{\sigma_j^2}]$$
$$= \sum_i \varphi_j(x_i)[\frac{(x_i - \mu_j)^T (x_i - \mu_j)^T}{\sigma_j^3} - \frac{1}{\sigma_j^2}] \tag{3.25}$$

Let (3.24) and (3.25) be o , we have that,

$$\hat{\mu}_j = \frac{\sum_i \varphi_j(x_i) x_i}{\sum_i \varphi_j(x_i)} \tag{3.26}$$

$$\hat{\sigma}^2 = \frac{1}{3} \frac{\sum_i \varphi_j(x_i)(x_i - \mu_j)^T (x_i - \mu_j)}{\sum_i \varphi_j(x_i)} \tag{3.27}$$

For α_j , under the constraint $\sum_j \alpha_j = 1$, use Lagrange multipliers to re-define the object as,

$$J = L(\theta) + \lambda(1 - \sum_{i=1}^{m} \alpha_i) = \sum_i \ln(\sum_j \alpha_j \delta(x_i, \mu_j, \sigma_j^2)) + \lambda(1 - \sum_{i=1}^{m} \alpha_i) \qquad (3.28)$$

Differentiating this new object w.r.t. α_j, we have that:

$$\partial_{\alpha_j} J = \sum_i \frac{\delta(x_i, \mu_j, \sigma_j^2)}{\sum_{j=1}^{3} \alpha_j \delta(x_i, \mu_j, \sigma_j^2)} - \lambda = \frac{1}{\alpha_j} \sum_i \varphi_j(x_i) - \lambda = 0 \qquad (3.29)$$

$$[\hat{\alpha}_1, \hat{\alpha}_2, \cdots, \hat{\alpha}_m] = [\frac{1}{\lambda} \sum_i \varphi_1(x_i)], \frac{1}{\lambda} \sum_i \varphi_2(x_i), \cdots, \frac{1}{\lambda} \sum_i \varphi_m(x_i)] \qquad (3.30)$$

$$\hat{\alpha}_1 + \hat{\alpha}_2 + \cdots + \hat{\alpha}_m = \frac{1}{\lambda}(\sum_i (\varphi_1(x_i) + \varphi_2(x_i) + \cdots + \varphi_m(x_i)) = 1 \qquad (3.31)$$

So we have that,

$$[\hat{\alpha}_1, \hat{\alpha}_2, \cdots \hat{\alpha}_m] = [\frac{1}{m} \sum_i \varphi_1(x_i), \frac{1}{m} \sum_i \varphi_2(x_i), \cdots, \frac{1}{m} \sum_i \varphi_m(x_i)] \qquad (3.33)$$

, where φ is also a function of parameters, and we can resolve this using the following

iteration:

Step 1: Let $\theta = [\alpha_1, \alpha_2, \cdots \alpha_m, \mu_1, \mu_2, \cdots \mu_m, \sigma_1^2, \sigma_2^2, \cdots \sigma_m^2]$. Given an initial value to achieve

convergence, $\mu_1, \mu_2, \cdots \mu_m$, it may be calculated by the cluster method.

Step 2: Calculate $\varphi_j(x_i)$.

Step 3: Calculate $\tilde{\mu}_j = \frac{\sum_i \varphi_j(x_i) x_i}{\sum_i \varphi_j(x_i)}$.

Step 4: Calculate $\sigma_j = \dfrac{1}{m} \dfrac{\sum\limits_i \varphi_j(x_i)(x_i - \tilde{\mu}_j)^T (x_i - \tilde{\mu}_j)}{\sum\limits_i \varphi_j(x_i)}$.

Step 5: Calculate $\alpha_j = \dfrac{1}{m} \sum\limits_i \varphi_j(x_i)$.

Step 6: Let $\hat{\theta} = [\hat{\alpha}_1, \hat{\alpha}_2, \cdots \hat{\alpha}_m, \hat{\mu}_1, \hat{\mu}_2, \cdots \hat{\mu}_m, \hat{\sigma}_1^2, \hat{\sigma}_2^2, \cdots \hat{\sigma}_m^2]$. If $\left\| \theta - \hat{\theta} \right\| < \xi$, for a given

threshold ξ , then stop the process, otherwise go to *Step 2*.

In actuality, the density function of data fusion under this special structure is a product of

the fusion of SGMs. For all multisource data x_k and their SGM densities $\delta(x_k)$,

$\delta(X_{fusion}) = \prod\limits_k \delta(x_k)$, and so we have that:

$$
\begin{aligned}
\prod_k \delta(x_k) &= \prod_k \frac{1}{\sqrt{(2\pi)^3}\,\sigma_k} e^{-\frac{1}{2\sigma_k^2}(X-\mu)^T(X-\mu)} \\
&= \frac{1}{\sqrt{(2\pi)^{3k}} \prod_k \sigma_k} e^{\sum_k -\frac{1}{2\sigma_k^2}(X-\mu)^T(X-\mu)} \\
&= \frac{1}{\sqrt{(2\pi)^{3k}} \prod_k \sigma_k} e^{\alpha X^2 + \beta X + \gamma} \\
&= \frac{1}{\sqrt{(2\pi)^{3k}} \cdot C \cdot \prod_k \sigma_k} e^{-\frac{1}{2\sigma_k^2}(X'-\mu')^T(X'-\mu')} \\
&= C'\delta
\end{aligned}
\tag{3.34}
$$

, where C is undetermined const. In particular, if $\sigma = 1$, then we have that,

$$
\delta(I_{fusion}) = \frac{1}{\sqrt{(2\pi)^{3k}} \prod_k \sigma_k} e^{\alpha x^2 + \beta x + \gamma} = \frac{1}{\sqrt{(2\pi)^{3k}}} e^{\alpha x^2 + \beta x + \gamma} = \frac{1}{\sqrt{(2\pi)^{3k}}} e^{x'^2 + \beta' x' + \gamma'} = C \frac{1}{\sqrt{(2\pi)^3}} e^{(x'-\mu)^2} \tag{3.35}
$$

This is a linear transformation of the basic Gaussian function. Thus, for any two multisource data x_i, x_j, the fusion result is $x_{fusion} = < \alpha\, x_{ij}, \beta\mu_{ij}, \gamma\delta_{ij} >$, where α, β and γ are undetermined coefficients. Figure 3-10 shows some results on EM of Gaussian Mixed model.

	μ	σ
0.130982	−7.03521	6.59786
0.869018	1.40553	4.78557
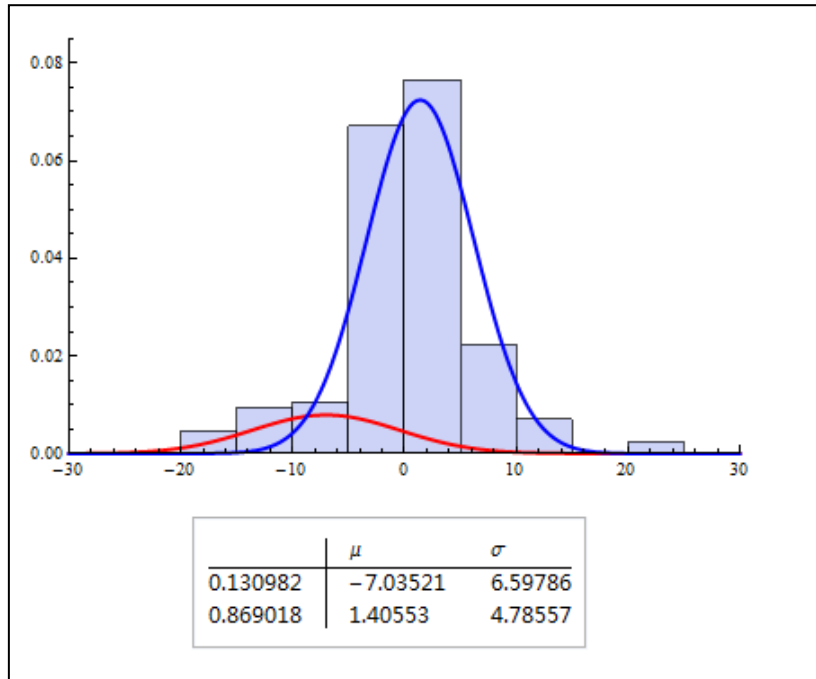
Figure 3-10 Expectation Maximization for Gaussian Mixture Distributions (by Anthony Fox, See Appendix- B)

3.4 Multimodal Fusion by Human Factors

Multi-modal data is describing an object whose data is collected by different methods or devices called data modality. For example, in multi-modal face recognition, multi-modal data may be caused by a human face 2D image and 3D shape models constitute the two modes (Vybornova et al., 2008). In a multi-modal video mining, the video can be broken down into subtitles, audio and images modal; pages of text and images can also be present as a distinct modes, and then described from a different point view to express the information pages (Liang et al., 2011; Loza et al., 2010).

The difference between multi-modal data mining and traditional data mining is that the former will carry out excavation work to the utilization of information between multiple modes, and tap the potential links between them (Fisher III et al., 2000; Hershey et al., 2001).

At present, there are two key questions in mining multi-modal data:

(1) How to mine and describe the different modal correlation between the information effectively. The biggest difference between multi-modal data and traditional data is multiple correlations between modes, so this correlation is the traditional data mining excavation that is not considered an important issue.

(2) How different the results of modes based integration are, even though data mining systems in each mode are given in good performance. How to combine the results of these minings effectively is still a very complex issue.

3.4.1 *Definitions on Preprocessing of Multimodal Data*

For different modal data that come from different device providers, such as images from satellites and cameras, and text from weather stations or from telecommunication stations in PDMS we have pre-definitions as below,

Definition 3.3 Parametric model based joint distribution (JD) of multimodal data. It is a method for probability density estimating on eigenspace space. This method will give a model by parameters and then suppose that the data's JD is subjected to a parametric model, so it is parameter estimation (PE) problem.

Definition 3.4 Non parametric model and eigenspace transformation; to overcome the complexity of PE on parametric model based JD, non parametric model was developed so that the first step is to map the data to a low dimensional space.

65

Let $v_i \xrightarrow{f} V \in R^{N_v}$ and $a_i \xrightarrow{f} A \in R^{N_a}$ be a sampled map of features in n-dimensional space, $f_v : R^{N_v} \mapsto R^{M_v}$ and $f_a : R^{N_a} \mapsto R^{M_a}$ be mapping functions from high dimensional to low dimensional space. The parameters of f_v and f_a are α_v and α_a respectively; α_v and α_a satisfy,

$$\{\alpha'_v, \alpha'_a\} = \arg\max I(f_v(V, \alpha_v), f_a(A, \alpha_a)) \tag{3.36}$$

, which $I(a,b)$ is the mutual data of a and b. The maximal operation of I will reduce the complexity of JD.

3.4.2 *Super Kernel for Multimodal Fusion*

Depend on the difference fusion process; we developed HFs based fusion in product operator, linear operator and nonlinear operator.

(1) Suppose that each modal was independent, and the posterior probability $P(d_i \mid h)$ of each modal was estimated exactly, let D be the multimodal data, we have that,

$$P(D \mid h) = \frac{P(h)\prod_{i=1}^{m} P(d_i \mid h)}{P(D)} \tag{3.37}$$

(2) Let each classification's output be $C_1, C_2, \cdots C_n$, then the linear fusion by HFs is

$$C = \sum_{i=1}^{n} C_i HF_i \tag{3.38}$$

Subjected to $\sum_{i=1}^{n} HF_i = 1$.

66

(3) There are two steps for nonlinear fusion of multimodal data. One is independent modality analysis and another is super kernel algorithm (Shown in Fig. 3-11)
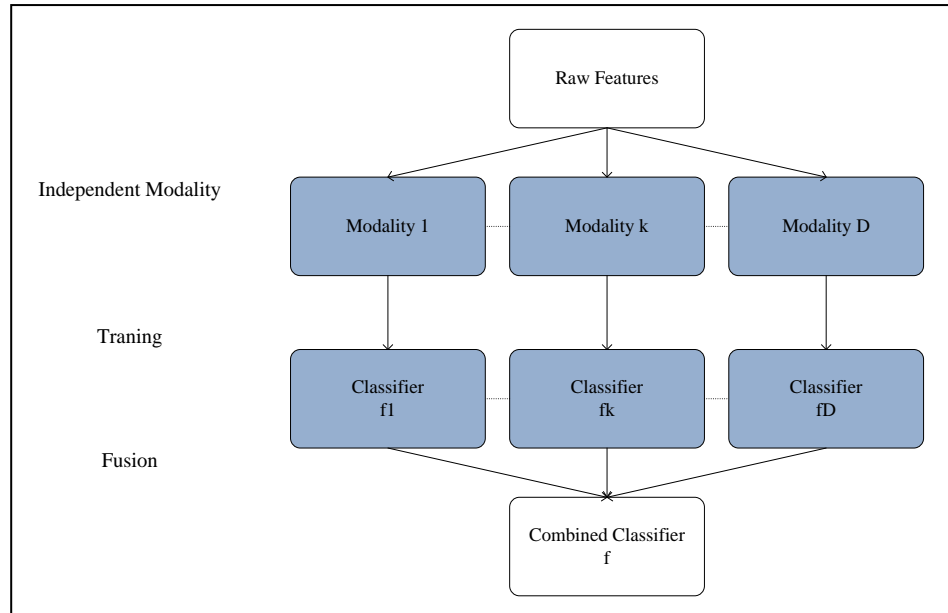
Figure 3-11 Multimodal Fusion in Independent Condition

The steps are as follows:

Step 1: Input matrix $X_{m \times n}$, n is the number of training samples, m is the number of multimodal source providers.

Step 2: Apply PCA (introduced by Section 3.2.2) to de-noisy and de-dimensionality of X

Step 3: Using independent component analysis to acquire an estimation of features of each modality.

Step 4: Using independent modality grouping to get $M_1, M_2, \cdots M_D$ which are independent modalities.

Step 5: Super kernel integration on $M_1, M_2, \cdots M_D$ as following algorithms [Wu, et al., 2004].

Figure 3-12 shows the peosycode of super kernel algorithm for independent multimodal fusion.

Algorithm **Super kernel Fusion**

Input:

X=[X$_1$, X$_2$, ..., X$_n$]; /* *training data*

Y =[Y$_1$, Y$_2$, ..., Y$_n$]; /* *Labels of X*

Output:

f; /**Class-prediction function*

Variable:

f$_1$,...f$_D$; /**A set of discriminative functions*

m$_1$,..., m$_D$; /**A set of* n * |mD| *matrices*

K; /* *Super kernel matrix with dimension of* n* D

Function calls:

fd(x^d_i); /* *Prediction score of* x^d_i *from* fd

Train(K; Y); /* *Train a discriminative function*

IMA(X); /* *Independent modality analysis*

Prob(s); /**Convert an SVM score to probability*

Begin:

1) (m$_1$,...,m$_D$)<- IMA(X);

2) **for** each d = 1,...,D

3) f$_d$ <-Train(m$_d$; Y);

4) **for** each data x$_i$ of X

5) **for** each discriminative function f$_d$

6) K(i,d) Prob(f$_d$(x^d_i));

7) f <- Train(K, Y);

8) return f;

End

Figure 3-12 Super Kernel Algorithm for Independent Multimodal Fusion
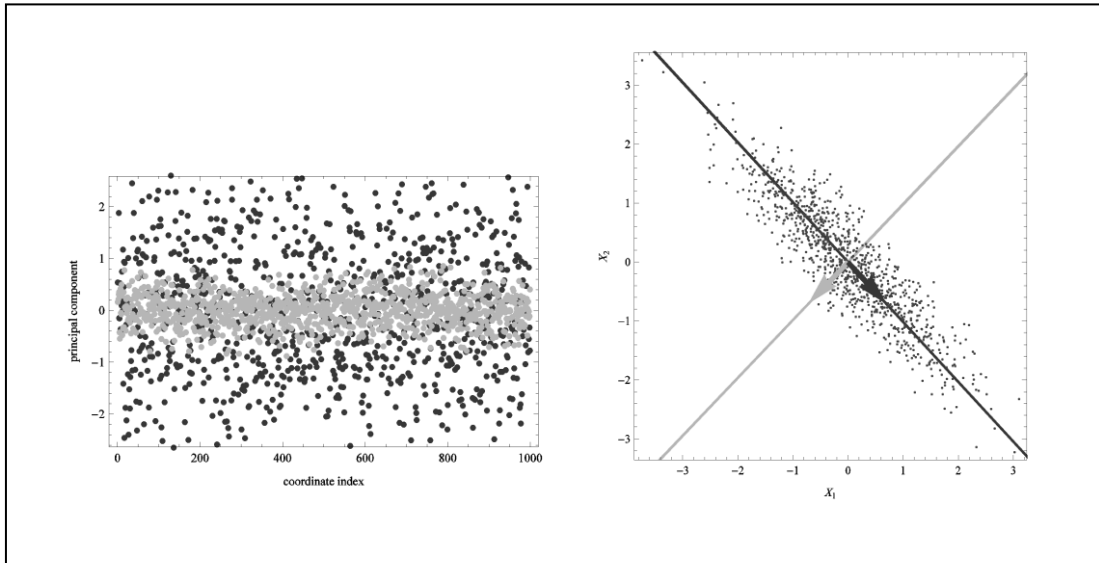
Figure 3-13 Multimodal Fusion by PCA

3.5 Fuzzy Inference System in PDMS

3.5.1 *Definitions on FIS*

Approximate reasoning is viewed as a process of approximate solution of a system of relational assignment equations. In human thinking, reasoning process is often approximate and such imprecise reasoning cannot use the classical two-valued logic value. LA Zadeh (1975a; 1975b; 1975c) in 1975 first proposed the synthesis fuzzy inference rules and the conditional statement "If x is A, then y is B" is converted to fuzzy relations rules. Since then JF Baldwin, RR Yager and others have also used their true value with fuzzy logic. And Zadeh (2005) also introduced Linguistic Variable (LV) which translates natural language to a membership of fuzzy set (Shown in Figure 3-14).
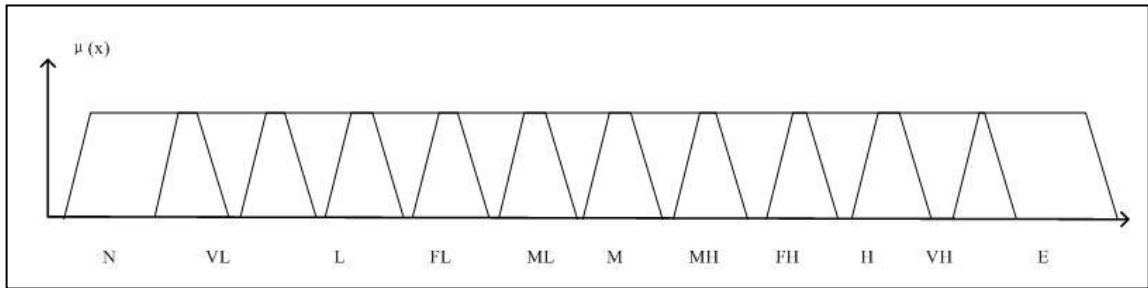
Figure 3-14 Linguistic Variables and Their Membership Functions

Fuzzy sets have been studied following the fuzzy systems development rapidly. In this dissertation, Fuzzy Implication Operators were used in knowledge presentation of PDMS for decision-making (See Figure 3-15).

Since the selection of implication operators and fuzzy reasoning are closely related to the effect, in particular, associated triangular norms and implication operators studies on fuzzy reasoning and fuzzy logic combined with important and broad implications are mutual. So the purpose is to accompany each other based on triangular norms and implication operators to establish a new form of fuzzy propositional calculus system.
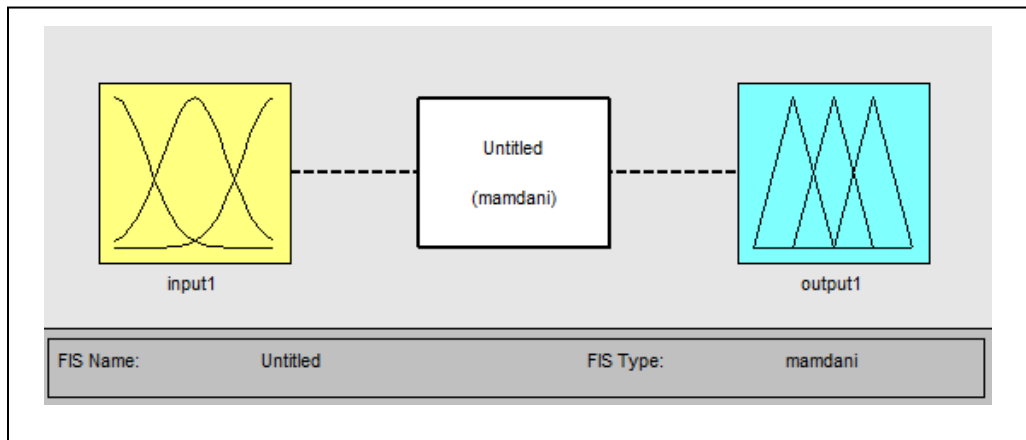
Figure 3-15 Fuzzy Implication for Knowledge Presentation by IF-THEN Rule

3.5.2 *Rule-Based FIS for Knowledge Presentation of PDMS*

In the study of human cognition model, Newell and Simon (1972) developed a rules-based knowledge presentation system. At present, rule-based representation has become the most widely used in artificial intelligence, especially in expert systems. Many successful expert systems are based on this knowledge representation model. The basic form of production $P \to Q$ or "IF P THEN Q", P is a prerequisite for production, also known as front piece, which gives the possibility to use the production prerequisite by the fact that the logical combination to form; Q is a group conclusions or operation, also known as post-production pieces

By the Gaussian fusion, HFs as the density functions, so the FIS by HF will be the fusion of any two "IF-THEN" rules integration. Suppose that,

RULE: IF "x is A", THEN "y is B", the assertion "x is A" satisfies the Gaussian distribution,

$$f(P, A, \sigma) = \frac{1}{\sqrt{2\pi}\sigma_A} e^{-\frac{1}{2}[\frac{(P-A)^2}{\sigma_A^2}]} \tag{3.39}$$

Such as by Mamndini, we have that,

$$Mamndini_{FIS} = Min(\frac{1}{\sqrt{2\pi}\sigma_A} e^{-\frac{1}{2}[\frac{(P-A)^2}{\sigma_A^2}]}, \frac{1}{\sqrt{2\pi}\sigma_B} e^{-\frac{1}{2}[\frac{(P-B)^2}{\sigma_B^2}]}) \tag{3.40}$$
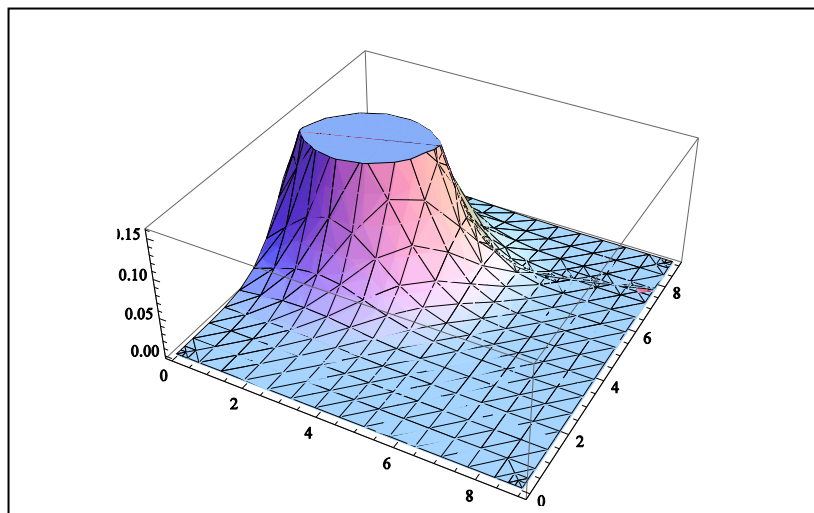
Figure 3-16 IF-THEN Rule by Mamndini Gaussian Implication

3.5.3 *FIS for Fusion and Decision-Making*

Suppose that the m knowledge rules in v_k under Gaussian model (see formula (3.10))

conclude a particular assertion at the δ_k level, we have that:

$$\text{IF } x_1 \text{ is } v_1 \text{ and } x_2 \text{ is } v_2 \text{ and} \cdots, \text{ and } x_m \text{ is } v_m \text{ THEN } I_o \text{ is } v_o, \varpi \qquad (3.41)$$

, where ω is the result, this can be simplified to,

$$\text{IF } v_1 \text{ and } v_2 \text{ and,} \cdots, \text{ and } v_m \text{ THEN } \varpi \, (v_1, v_2, \cdots, v_m) \qquad (3.42)$$

and

$$\text{IF } x_1 \text{ and } x_2 \text{ and,} \cdots, \text{ and } x_m \text{ THEN } x_1 \wedge x_2 \wedge \cdots \wedge x_m \qquad (3.43)$$

$$\text{IF } d_1 \text{ and } d_2 \text{ and,} \cdots, \text{ and } d_m \text{ THEN } d_1 \wedge d_2 \wedge \cdots \wedge d_m \qquad (3.44)$$

$$\text{IF } \rho_1 \text{ and } \rho_2 \text{ and,} \cdots, \text{ and } \rho_m \text{ THEN } \varpi \, (\rho_1, \rho_2, \cdots, \rho_m) \qquad (3.45)$$

Now, we only discuss the density function δ_k in our FIS. Letting $\Delta = [\delta_1, \delta_2, \cdots, \delta_n]$, we have that,

$$\text{IF } \Delta \text{ THEN } \varpi(\Delta) \tag{3.46}$$

From the previous section, we know that $\varpi(\Delta)$ is a Gaussian density function, so the rule is re-labeled as "IF X THEN $f(X)$", and for this rule set, we have

$$R_i: \text{IF X THEN } f_i(X) \tag{3.47}$$

However, as $f(X)$ is a nonlinear function, it is difficult to find its minimum point under the Mamdani model, so we need to linearize $f(X)$ and use the nonlinear conjugate gradient algorithm to optimize the parameters of $f(X)$.

3.6 Semi-Supervised Methods by Fuzzy Set Model

3.6.1 *Definitions and Problem Formalization*

Graph-based SSL was widely used in structural parameterized learning (Lin 1995), clustering (Biswas, 1983) and IF-THEN rules (Klose et al., 2005). We developed two methods for this GIS based PDMS. Firstly, the GIS-based PDMS' resource map was converted to a directed graph that presented by matrix (See Section 3.2.1) and then the decision information will be conducted from the PDMS by two ways, one is clustering operation, and another is graph-based semi-supervised optimization process. In this dissertation, we develop a semi-supervised fuzzy c-mean for knowledge clustering and a graph minimization method for PDMS resource calculation.

3.6.2 *Semi-Supervised Learning Process with Fuzzy C- Mean*

Fuzzy c-mean [Bezdek, et al., 1981] is one of classical clustering algorithms; its objective is to find a fuzzy separating on a given data-set $\{X_1, X_2, \cdots X_N\}$, $X_j \in R^D$, $j = 1, 2, \cdots, N$ to minimize the cost function:

$$J(U, A) = \sum_{i=1}^{c} \sum_{j=1}^{N} u_{i,j}^m D_{ij}, (i = 1, 2, \cdots, c, j = 1, 2, \cdots, N) \tag{3.48}$$

$U = [u_{i,j}]$ - Membership matrix.

$u_{i,j} \in [0,1]$ - The j -th membership of the i -th class.

$A = [A_1, A_2, \cdots A_c]$ - The matrix of cluster center.

$m \in [1, \infty)$ -fuzzy parameter and assigned by $1.5 \square 2.5$.

D_{ij} -the distance of A_i and X_j.

Minimized the formula (3.48), we get the iteration,

$$u_{i,j} = \frac{1}{(\sum\limits_{l=1}^{c} (D_{lj} / D_{ij})^{\frac{1}{1-m}})} \tag{3.49}$$

$$A_i = \frac{(\sum\limits_{j=1}^{N} u_{i,j}^{m} X_j)}{\sum\limits_{j=1}^{N} u_{i,j}^{m}} \tag{3.50}$$

Distance-based semi-supervised learning algorithm was adopted for PDMS, suppose that data-set $X = \{X_1, X \cdots, X_n\}$ for each X_i is a d-dimensional vector. $S \subset X \times X$ is the similarity pair points and $D \subset X \times X$ is the rest, so $(X_P, X_Q) \in S$ or $(X_P, X_Q) \in D$. We defined indicators as,

$$ML_i \subset (X_{m_p}, X_{m_q}) \mid X_{m_p}, X_{m_q} \in X, l_{m_p} = l_{m_q} = i \tag{3.51}$$

The algorithm is (Lai et al., 2011; Murty et al., 1999; Pedrycz, 1985; Zhang et al., 2004):

(1) For the training data, a special distance function was used for each class as,

$$f_{A_i}(X_P, X_Q) = [(X_P - X_Q)^T A_i (X_P - X_Q)]^{\frac{1}{2}} \tag{3.52}$$

$A_i (i = 1, 2, \cdots, k)$ is a positive diagonal matrix for every class, and k is the number of classification. So the minimized objective function (3.48) was converted as

$$g(A_i) = \sum_{(X_P, X_Q) \in ML_i} \|X_P - X_Q\|_{A_i}^2 - lb \left(\sum_{(X_P, X_Q) \in \Omega / ML_i} \|X_P - X_Q\|_{A_i} \right) \tag{3.53}$$

(2) For test data-set, the objective function is,

$$J = \sum_{i=1}^{k} \sum_{j=1}^{n} f_{A_i}(x_j, c_i) U_{ij}^{b} \qquad (3.54)$$

So the cluster center and membership were calculated as,

$$c_i = \frac{\sum_{j=1}^{n} u_{ij}^{b} x_j}{\sum_{j=1}^{n} u_{ij}^{b}} \qquad (3.55)$$

$$u_{ij} = \frac{1}{\sum_{h=1}^{k} \left(\frac{\left\| c_i - x_j \right\|^2}{\left\| c_q - x_j \right\|^2} \right)^{\frac{1}{b-1}}} \qquad (3.56)$$

We know that there are some objectives for semi-supervised fuzzy clustering algorithms listed in Table 3-2 (Endo et al., 2009; Liu et al., 2008; Pedrycz et al., 1997).

Table 3-2 Some Objective Functions And Their Prosperities

Algorithm s	(1) Pedrycz-97 (2) Li-08 (3) Endo-09
Objective	$(1)\ J_k = \sum\limits_{i=1}^{c} \sum\limits_{k=1}^{N} u_{ik}^{p} d_{ik}^{2} + \alpha \sum\limits_{i=1}^{c} \sum\limits_{k=1}^{N} (u_{ik} - f_{ik} b_k)^{p} d_{ik}^{2}$ (2) $J_m = \sum\limits_{i=1}^{c} \sum\limits_{k=L+1}^{N} u_{ik}^{m} d_{ik}^{2} + (1-\alpha) \sum\limits_{i=1}^{c} \sum\limits_{k=1}^{L} u_{ik}^{m} d_{ik}^{2} + \alpha \sum\limits_{i=1}^{c} \sum\limits_{k=1}^{L} (u_{ik} - f_{ik})^{m} d_{ik}^{2}$ $(3)\ J = \sum\limits_{i=1}^{c} \sum\limits_{k=1}^{N} u_{ik} \left\| x_k - v_i \right\|^2 + \lambda^{-1} \sum\limits_{i=1}^{c} \sum\limits_{k=1}^{N} (u_{ik} - \bar{u}_{ik}) \log(u_{ik} - \bar{u}_{ik})$
Centroid	$(1)\ v_i = \dfrac{\sum\limits_{k=1}^{N} u_{ik}^{2} x_k}{\sum\limits_{k=1}^{N} u_{ik}^{2}}$ $(2)\ v_i^{(l)} = \dfrac{\sum\limits_{k=1}^{N} (u_{ik}^{(l)})^{2} x_k}{\sum\limits_{k=1}^{N} (u_{ik}^{(l)})^{2}}$ $(3)\ v_i = \dfrac{\sum\limits_{k=1}^{N} u_{ik} x_k}{\sum\limits_{k=1}^{N} u_{ik}}$

Algorithm	(1) Pedrycz-97
s	(2) Li-08
	(3) Endo-09

Partition Matrix	
	$(1)\ u_{ij} = \dfrac{1}{1+\alpha} \left\{ \dfrac{1 + \alpha\,(1 - b_j \sum_{l=1}^{c} f_{lj})}{\sum_{l=1}^{c} (\dfrac{d_{ij}}{d_{lj}})^2} + \alpha\, f_{ij} b_j \right\}$
	$(2)\ u_{ik} = (1-\alpha) \left(\dfrac{1}{\sum_{l=1}^{c} (\dfrac{d_{ij}}{d_{lj}})^2} \right) + \alpha\, f_{ik} = (1-\alpha) u_{ik}^{FCM} + \alpha\, f_{ik}$
(labeled)	
	$(3)\ u_{ik} = \overline{u}_{ik} + \dfrac{e^{-\lambda d_{ik}}}{\sum_{j=1}^{c} e^{-\lambda d_{jk}}} (1 - \sum_{j=1}^{c} \overline{u}_{jk})$

Distance Metric	
	(1) Mahalonobis
	(2) Mahalonobis
	(3) Euclidean

3.6.3 *Semi-Supervised Learning Process with Graph for PDMS*

Graph-based SSL uses a graph representation of the data with a node for each labeled and unlabeled example. The algorithm is to create an optimization objective function or minimization cost function; this cost function consisted of loss function and regularize. We will adopt the alternative minimization method for PDMS map calculation introduced by Wang et al. (2008). Let it start from a simple function as Gaussian fields and harmonic functions (Zhu et al., 2003).

(1) Let the objective function be,

$$\sum_{i \sim j} w_{ij} (f(x_i) - f(x_j))^2, i, j = 1, 2, \cdots, n \tag{3.57}$$

, where f is harmonic function that satisfying Δf is 0 on unlabeled node and $f(x_i) = y_i$ for $i = 1, 2, \cdots, l$ on labeled node. Rewrite the objective (3.57) by Laplacian operator (see Appendix I-(1)), we have that,

$$\sum_{i \sim j} w_{ij} (f(x_i) - f(x_j))^2 = f^T \Delta f \ , \ i, j = 1, 2, \cdots, n \tag{3.58}$$

So the objective is,

$$\min_f \sum_{i \sim j} w_{ij} (f(x_i) - f(x_j))^2 = f^T \Delta f \tag{3.59}$$

By matrix transform, we can get the matrix as $W = \begin{bmatrix} W_{ll} & W_{lu} \\ W_{ul} & W_{uu} \end{bmatrix}$ on labeled node arranged before the l -row and u -col. And the same operating on matrix D and Δ , let $f = \begin{bmatrix} f_l \\ f_u \end{bmatrix}$, f_i denotes the value of the node i . So we have that,

$$f_u = (D_{uu} - W_{uu})^{-1} W_{ul} Y_l = -\Delta_{uu}^{-1} \Delta_{ul} Y_l \tag{3.60}$$

But this method cannot deal with noisy data.

(2) Local and global consistency

This method sets the objective function as,

$$
\min_{F} \frac{1}{2} \left(\mu \sum_{i=1}^{n} \|F_i - Y_i\|^2 + \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} \left\| \frac{F_i}{\sqrt{d_{ii}}} + \frac{F_j}{\sqrt{d_{jj}}} \right\|^2 \right)
\tag{3.61}
$$

, where parameter $\mu > 0$, $F = \left[F_1^T, F_2^T, \cdots, F_n^T \right]^T$ is a $n \times c$ matrix, c is the number of classifications, the label of node x_i is $y_i = \arg\max_{j \leq c} F_{ij}$, and $Y_{ij} = 1$ if $y_i = j$ otherwise 0. The objective function will be acquire by the following steps

Let

$$
O = \frac{1}{2} \left(\mu \sum_{i=1}^{n} \|F_i - Y_i\|^2 + \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} \left\| \frac{F_i}{\sqrt{d_{ii}}} + \frac{F_j}{\sqrt{d_{jj}}} \right\|^2 \right)
\tag{3.62}
$$

Differential on (3.62) and let it be 0, w.r.t. we have that,

$$
\left. \frac{\partial O}{\partial F} \right|_{F = F^*} = L F^* + \mu (F^* - Y) = 0
\tag{3.63}
$$

$$
(L + \mu I) F^* = \mu Y
\tag{3.64}
$$

$$F^* = \mu(L + \mu I)^{-1}Y$$

$$= (\frac{1}{\mu}L + I)^{-1}Y$$

$$= \frac{\mu}{1+\mu}(I - \frac{1}{1+\mu}(I-L))^{-1}Y \tag{3.65}$$

$$= \frac{\mu}{1+\mu}(I - \frac{1}{1+\mu}(D^{-\frac{1}{2}}WD^{\frac{1}{2}}))^{-1}Y$$

(3) Graph-based semi-supervised learning algorithm

For each data point to its k nearest neighbors or to examples within some distance ε , the weight w_{ij} of an edge between x_i and x_j is $e^{\frac{-\|x_i - x_j\|^2}{\varepsilon}}$. By (2) the graph serves as a proxy for the manifold. A term is added to the standard Tikhonov regularization problem (see Appendix I-(2)) to enforce smoothness of the solution relative to the manifold as well as relative to the ambient input space. The minimization problem becomes,

$$\arg\min_{f \in H} \left(\frac{1}{l}\sum_{i=1}^{l} V(f(x_i), y_i) + \lambda_A \|f\|_H^2 \, \lambda_I \int_M f(x)\|\nabla_M f(x)\|^2 \, dp(x) \right) \tag{3.66}$$

, where H is a reproducing kernel Hilbert space and M is the manifold on which the data lies. λ_A and λ_I are the smoothness parameters. The graph is used to approximate the intrinsic regularization term. Defining the graph Laplacian $L = D - W$, we have that,

$$f^T L f = \sum_{i=1}^{l+u}\sum_{j=1}^{l+u} W_{ij}(f_i - f_j)^2 \approx \int_M f(x)\|\nabla_M f(x)\|^2 \, dp(x) \tag{3.67}$$

3.7 Summary and Conclusion

From the above mentioned, a graph-based semi-supervised learning algorithm is used to construct a diagram to achieve semi-supervised learning, where the graph nodes represent data, and the weighted edge represents the similarity between data. Thus, the graph structure for the performance of the algorithm will have a greater impact. We need a good map to reflect the semi-supervised learning assumptions (such as smoothness assumptions, clustering hypothesis, and manifold assumptions), and it can reflect people's domain knowledge in this field. However, people do not necessarily understand a particular field of knowledge, so researchers generally use some common methods to build, such as E- neighborhood graph and k-nearest neighborhood (KNN) graph. This figure is derived from a directed graph, although there is literature research for directed graphs for semi-supervised classification. While disaster management system maps can be assumed as a fully connected graph, in which, all the edges between the nodes are connected with the figure of merit is relatively easy to construct without having to consider parameter selection, and because all edges have a weighted value (i.e., a derivative weighting function), so the algorithm for solving this is greatly convenient. Its disadvantages are also obvious, and to calculate the volume increases all the edges need to be addressed using Minkowski distance or Mahalanois distance. In addition, many researchers have proposed data flow characteristic shape or density to measure the distance between data points.

We get the distance between the data and get the distance from the weight of the edge. Calculation of the weight usually uses various kernel functions; mostly it is Gauss radial basis function (RBF).

However, semi-supervised learning algorithm considers only the similarity between the data points, but no mention of dissimilarity. In fact, we often encounter similar information, such

as two data can not belong to the same category that has different labels, and these are often also provided by useful information domain knowledge. Therefore, the training classifiers incorporating such information would help to improve the performance. But for such information to be simply used directly in a variety of algorithms is not feasible. Because of the similarity between data points generally expressed with edge weights, weight usually non-negative, then the data points which are indicated of the similarity between the information do not need to represent negative weights, then these negative weights that added to the learning algorithm will lead to the optimization of the original target making it becomes unsolvable.

CHAPTER 4: SIMULATION RESULTS AND DISCUSSION

4.1 Overview

In this chapter, we simulated mixed Gaussian based fusion operators with fuzzy c-Mean method on disaster dataset; Graph-based hospital distribution dataset was also simulated by comparing it with a current GIS based project on android apps for PDMS. We also discussed the effectiveness of Gaussian fusion dataset based semi-supervised learning by comparing it with other dataset types, such as moons type. Continuously, an evaluation dataset of hospitals emergency capacity in post-disaster management was applied in the proposed semi-supervised methods by rule-based presentation.

4.2 Simulation for Data Fusion of Disaster Map

Kalman filter on the target disaster map for scaling and decreasing the population of PDMS so that will also decrease the complexity of calculation. We downloaded the disaster map from FEMA and apsnet.com and combined with declaration map of FEMA. One is a Florida Hurricane map and another is a post hurricane map.

4.2.1 Gaussian Fusion on Disaster Map Using Discrete Wavelet Transformation

Histogram, a graphical representation of the distribution of data, to plot the density of data, and often for density estimation: estimating the probability density function of the underlying variable.
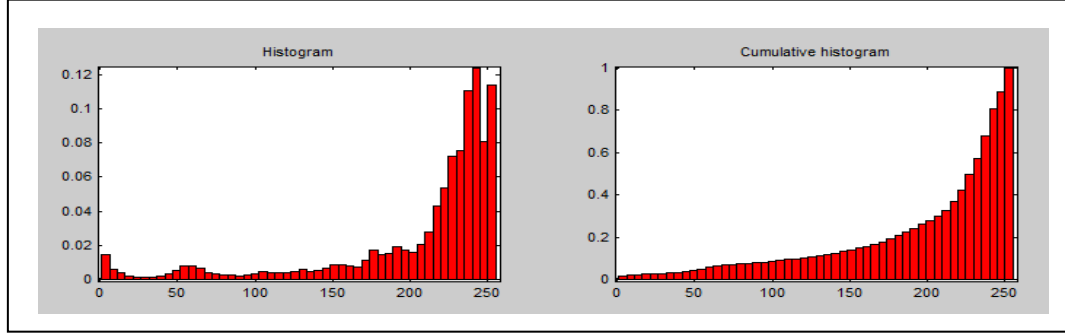
Figure 4-1 Histogram of Pictures

SWT, a wavelet transform algorithm, was designed to overcome the lack of translation-invariance of the discrete wavelet transform (DWT). Before fusion, all images will be preprocessed by DWT, and as we know that the continuous wavelet transform (CWT) is defined by:

$$X_{WT}(\tau, s) = \frac{1}{\sqrt{|s|}} \int x(t) \cdot \psi^* \left(\frac{t-\tau}{s}\right) dt \tag{4.1}$$

But the DWT, will calculate the sub-band coding to generate a fast computation of wavelet transform, so we need to convert continuous WTs to discrete WTs by the following transformation:

$$\psi_{s,\tau}(x) = \frac{1}{\sqrt{|s|}} \psi \left(\frac{x-\tau}{s}\right) \tag{4.2}$$

$$\psi_{s,\tau}(x) = \frac{1}{\sqrt{|s_0^j|}} \psi \left(\frac{x - k\tau_0 s_0^j}{s_0^j}\right) \tag{4.3}$$

$$\Psi_{s,\tau}(x) = A \Psi(s_0^j x) e^{-2j\pi f(k\tau_0^4 s_0^j)} \tag{4.4}$$

86

The procedure starts with passing this signal; the convolution operation in discrete time is defined as follows:

$$x[n] \times h[n] = \sum_{k=-\infty}^{+\infty} x[k] \cdot h[n-k]$$

(4.5)

The procedure can be expressed mathematically continuous as:

$$y[n] = \sum_{k=-\infty}^{+\infty} h[k] \cdot x[2n-k]$$

(4.6)

$$y_h[k] = \sum_n x[n] \cdot g[2k-n]$$

(4.7)

$$y_l[k] = \sum_n x[n] \cdot h[2k-n]$$

(4.8)

iDWT is the inverse process of DWT after fusion operation. In contrast to decomposition, the reconstruction process is comprised of up-sampling and then filtering. The filters are determined by the wavelet chosen.

And also we used fusion algorithms for comparative analysis in this study. Suppose there have two map pictures, initial map and disaster map, denoted by matrices A and B must be of same size, the fusion method has,

Mean-Mean and Max-UD (Upper-Down) fusion algorithm were composed of the calculation process as below that we introduced the code:

```
        'max' : D = abs(A) >= abs(B) ; C = A(D) +
B(~D)

        %      - 'min' : D = abs(A) <= abs(B) ; C =
A(D) + B(~D)

        %            - 'mean' :  C  =  (A+B)/2  ;  D  =
ones(size(A))


        Up-Down fusion, with paramMETH >= 0

        %              x = linspace(0,1,size(A,1));

        %              P = x.^paramMETH;

        %              Then each row of C is computed
with:

        %                  C(i,:) = A(i,:)*(1-P(i)) +
B(i,:)*P(i);

        %                  So C(1,:)= A(1,:) and C(end,:)=
A(end,:)
```

While the fusion operation was finished, the normal evaluation of the result picture needed to be further developed. As we know, the entropy and spatial frequency (SF) were the normal evaluation on the quality of the pictures, but for exact evaluation of the fusion results in the comparative analysis we adopted entropy, SF and image quality index (IQI) that Wang and Bovik introduced this indices in 2002 (Wang et al., 2002). Now we list the calculation methods as below:

A. Entropy is a logarithmic measure of the number of states with significant probability of being occupied as

$$E = -c \sum_i p_i \ln p_i \tag{4.9}$$

, where c is the Boltzmann constant, equal to 1.38065×10^{-23}

B. Spatial Frequency (SF) is a characteristic of any structure that is periodic across position in space and used to measure the overall activity level of an image.

$$SF = \sqrt{\left(\frac{1}{MN}\sum_{i=1}^{M}\sum_{j=2}^{N}[I(i,j)-I(i,j-1)]^2\right)^2 + \left(\frac{1}{MN}\sum_{j=1}^{N}\sum_{i=2}^{M}[I(i,j)-I(i-1,j)]^2\right)^2} \tag{4.10}$$

C. Image Quality Index (IQI): it was used as image quality distortion measure, which is mathematically defined by modeling the image distortion relative to the reference image as a combination of some factors as (4.11).

$$Q_0(x,y) = \frac{4\sigma_{xy}\overline{xy}}{(\overline{x}^2 + \overline{y}^2)(\sigma_x^2 + \sigma_y^2)} \tag{4.11}$$

Above mentioned indices will be compared with the proposed method in this study and the results will be shown in simulation section.

4.2.2 The Frameworks of Gaussian Fusion for Disaster Map

Suppose that there are two maps, one is the initial map that was acquired from map providers including some emergency resources in post-disaster, and another is disaster map which was published by special faculty like FEMA or NOVA. These two maps will be integrated to one map that will keep some emergency resources and also embed some disaster information that will benefit individuals in receiving accurate information from PDMS and allow a quick response from the system, the framework was shown in Fig 2.
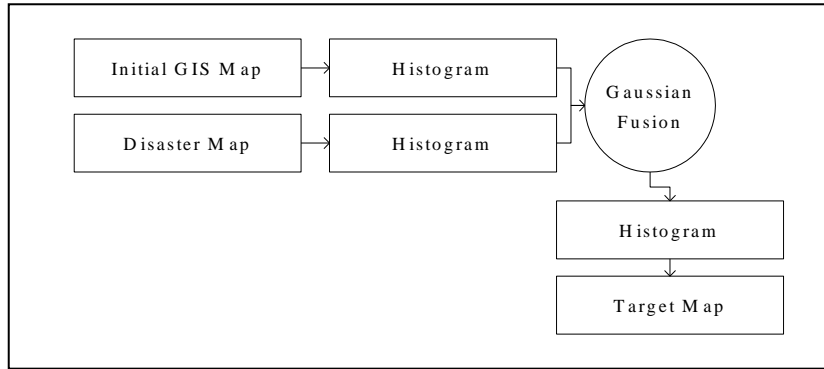
Figure 4-2 The Frameworks of Gaussian Fusion With Two Maps and Target Map Located and Gaussian Transform was Applied in the Maps.

The expectation map required was the target map; it scales the map in PDMS to further decrease the calculation complexity.

4.2.3 Results and Discussion

A) The DWT for maps

The map was by DWT first and then applied the inverse DWT. Fig 3 showed the transform on map.
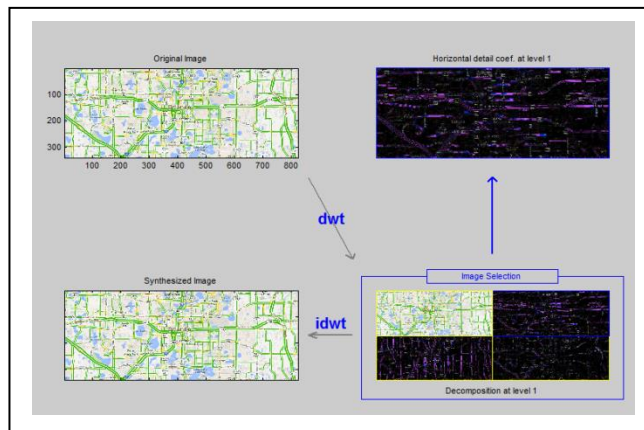
Figure 4-3 D DWT and iDWT on Map

We also can compare this to another transform, Stationary Wavelet Transform which is a stable transform for image normally, but this transform SWTlost much information. Fig 4 shows the result of the SWT on the map.
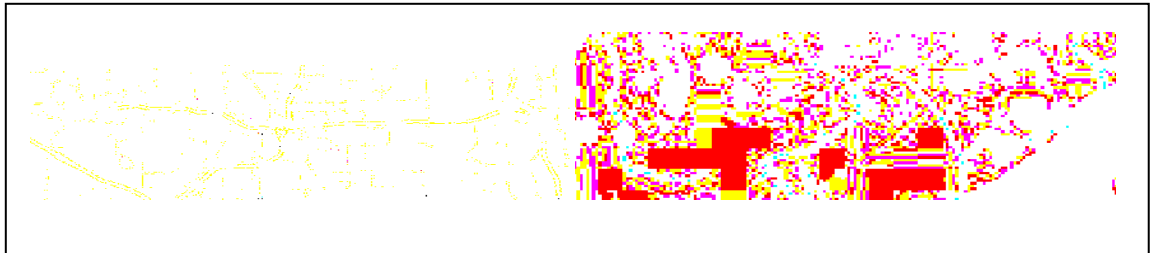
Figure 4-4 SWT on the Map

B) Gaussian fusion by DWT images

Now we simulated the images after DWT applied and then using the fusion method. Figure 4-5 showed the fusion result using Gaussian fusion algorithm.
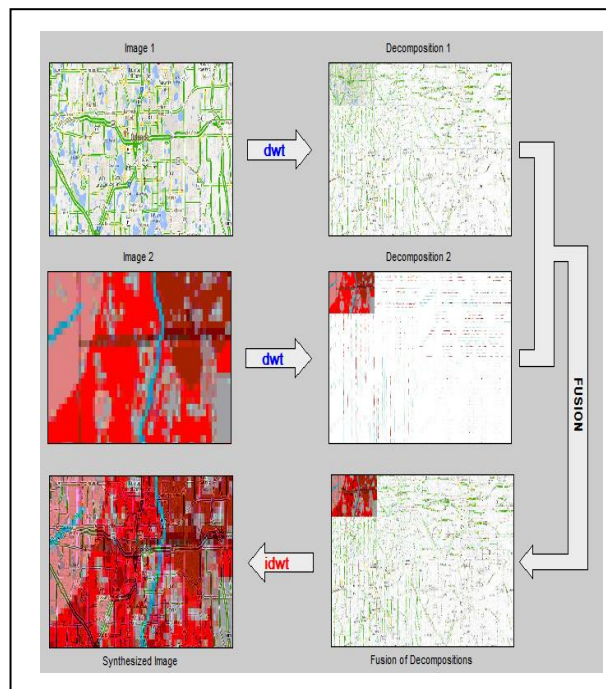
Figure 4-5 Gaussian Fusion After DWT

C) Compare with Fusion by Mean-Mean and Max-UD

91

To show the effectiveness of the proposed method in this study, we simulated two other fusion algorithms called mean-mean and max-UD. The mean-mean and Max-UD we simulated in Matlab 20012a. Fig 6 and Fig 7 showed the results of mean-mean fusion and max-UD fusion that we introduced in the calculation methods in Section 2. For these results pictures, we calculated the entropy, SF and IQI that we introduced in Section 2 and formula (9), (10), (11), we acquired the Figures 8, 9, and 10 to show their indices. The evaluation indices of initial and disaster maps before fusion were listed in Table 4-1. And we simulated 20 pairs initial and disaster maps and calculated the evaluation indices separately which were shown in Table 4-2.
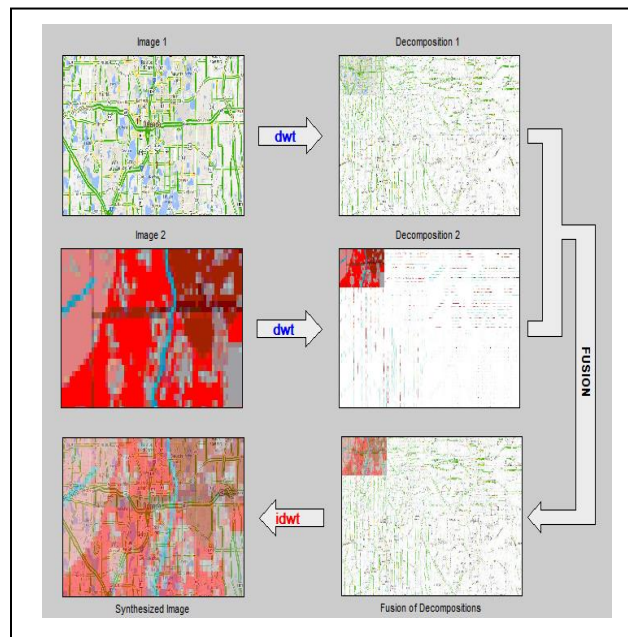
Figure 4-6 Mean-Mean Fusion on Map

Figure 4-7 Max-UD Fusion

From the above mentioned Figures, we know that Gaussian fusion appeared to have better performance in Times and IQI. To show the effectiveness of the proposed method, we compared the calculation results in Table 4-3.

Figure 4-8 Gaussian Fusion by Evaluation Indices

93

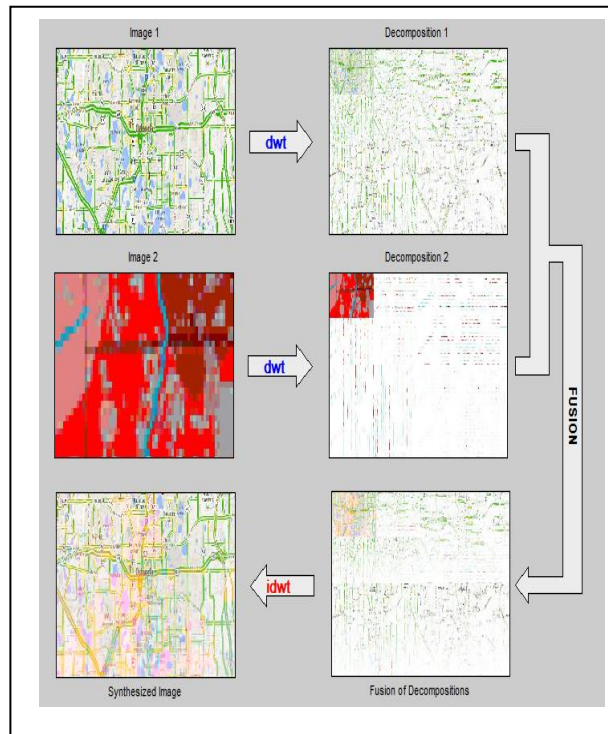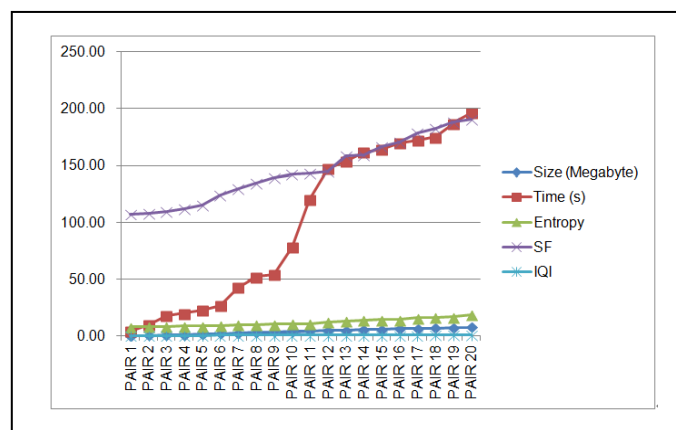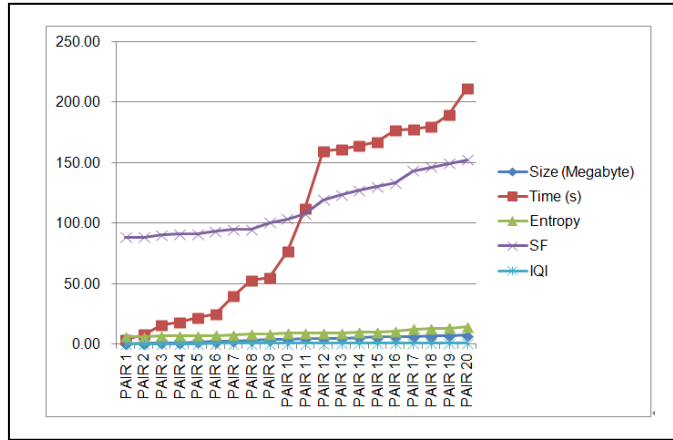Figure 4-9 The Evaluation Indices of Mean-Mean Fusioin
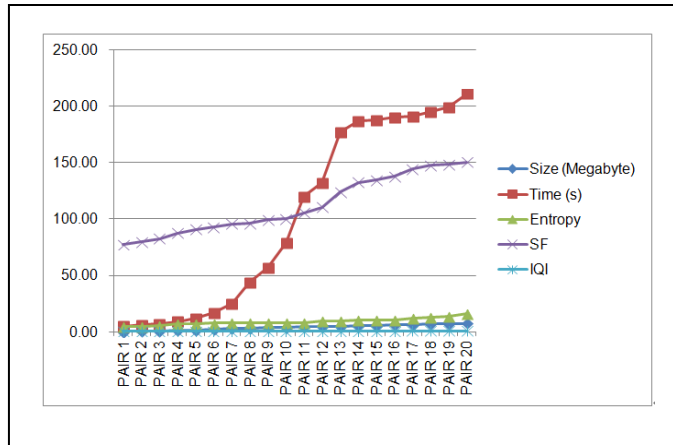
Figure 4-10 The Evaluation Indices of Max-UD Fusion

Table 4-1 The Initial and Disaster Maps Evaluation Indices

Imagery	Evaluation	Laplacian	Gradient	Contract	Ratio	Morph	SWT	DWT	aDWT
Initial	Entropy	7.11	7.12	7.02	6.75	7.09	7.03	7.01	7.19
	SF	139.4	148.5	149.4	90.02	145.3	146.7	138.4	188.2
	IQI	0.69	0.7	0.65	0.47	0.64	0.66	0.62	0.66
Dmap	Entropy	7.02	7.01	6.88	6.48	6.76	6.99	7.02	7.2
	SF	159.2	160.2	162.2	88.2	142.2	158.7	143.2	177.6
	IQI	0.71	0.68	0.6	0.55	0.62	0.56	0.54	0.63

Table 4-2 Pairs Simulation Results and Their Evaluation Indices

Gaussian Fusion Label	Size (Megabyte)	Time (s)	Entropy	SF	IQI
PAIR 1	0.08	4	7.94	107.3	0.35
PAIR 2	0.5	10	8.27	107.9	0.45
PAIR 3	0.8	18	8.59	109.2	0.44
PAIR 4	1.02	20	8.61	112.2	0.45
PAIR 5	1.5	23	8.7	114.9	0.56
PAIR 6	2.02	27	9.29	123.9	0.57
PAIR 7	2.6	43	9.64	129.9	0.59
PAIR 8	3.2	52	10.15	134.6	0.61
PAIR 9	3.6	54	10.39	139.3	0.65
PAIR 10	3.9	78	10.49	142.2	0.66
PAIR 11	4.5	120	10.84	143.4	0.67
PAIR 12	4.7	147	12.39	144.5	0.68
PAIR 13	4.9	154	13.11	158.5	0.72
PAIR 14	5.5	162	14.12	158.9	0.72
PAIR 15	5.8	164	14.4	166.5	0.74
PAIR 16	6.2	170	14.42	170.4	0.75
PAIR 17	6.5	172	15.9	178.5	0.76
PAIR 18	6.7	175	16.15	182.3	0.8
PAIR 19	7.01	187	16.81	188.5	0.88
PAIR 20	7.5	196	18.47	190.3	0.89

Table 4-3 The Evaluation Indices Of Gaussian, Mean-Mean and Max-UD Fusion

Evaluation of Fusion	Entropy	SF	IQI	Time
Gaussian *	7.27	188.9	0.68	23
Mean-Mean	6.82	175.2	0.62	20
Max-UD	6.87	158.8	0.72	27
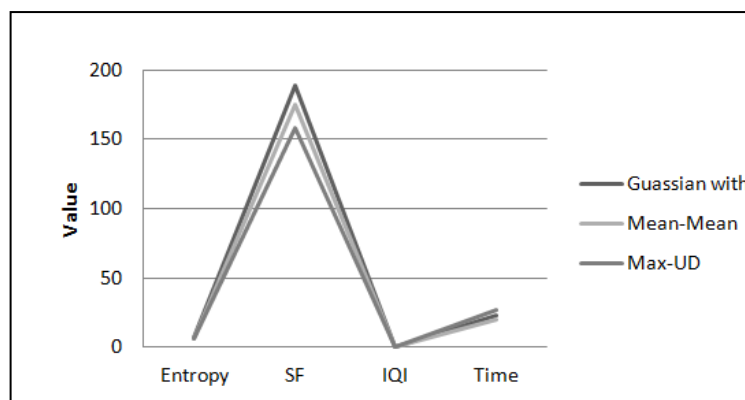
Figure 4-11 The Performance of Gaussian Fusion

In this case, map based PDMS was improved by scaling the size of map to decrease the calculated cost of each individual getting timely and effective information from the post-disaster system. Normaly, map based PDMS did not scale the map and the calculation was based on the whole map, and resulted in a system the with an output of abundent information. We proposed a Gaussian fusion method to keep disaster information and emegency resources in map in PDMS; through the fusion results, the evaluation indices including entropy, SF and IQI were calculated to show the effectiveness of the proposed method with Mean-Mean and Max-UD fusion algorithms. Next, in order to improve the algorithm of Gaussian fusion in parameter estimation and to properly acquire the map dataset. The shorcomings of this method are that the tradeoff of

the Gaussian transform need to be considered and second, the descrete transform will be time

cost that we need to improve in future works.

4.3 Simulation for Dataset of Fuzzy C-mean and Graph-Based Hospital Distribution

GIS is a directed graph, node value (multidimensional vector-based calculation), we can

acquired the hospitals information in the pre-post stage from map based product. Figure 4-1 is a

map for hospitals distribution, while we applied the method in Section 3.2.1. We can conduct a

map and store them as a matrix shown as Figure 4-2.
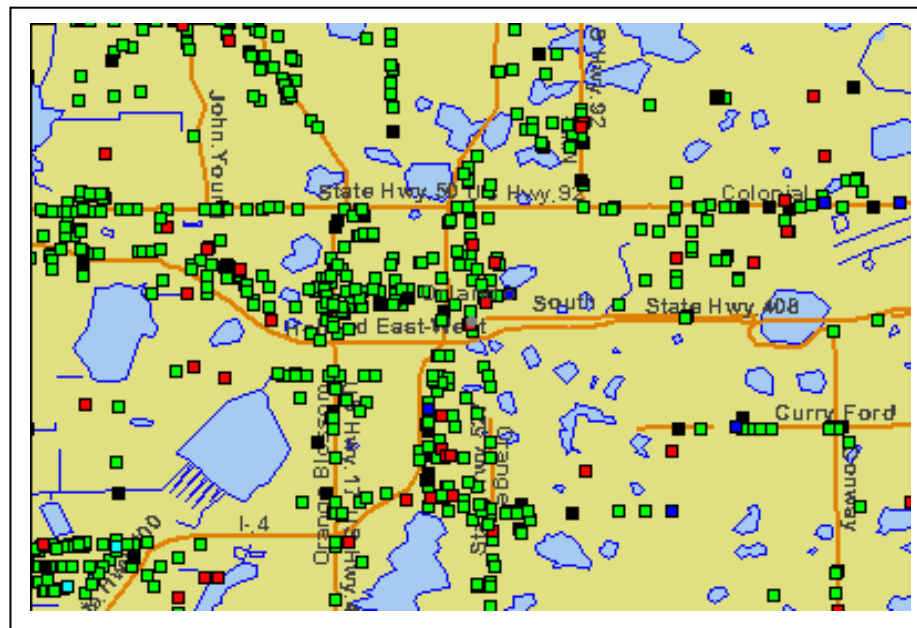
Figure 4-12 Hospitals Distribution Map Dataset
Source: City Data Generation Dataset
http://www.city-data.com/city/Orlando-Florida.html

97

Emergency Hospitals in GIS as undirected graph
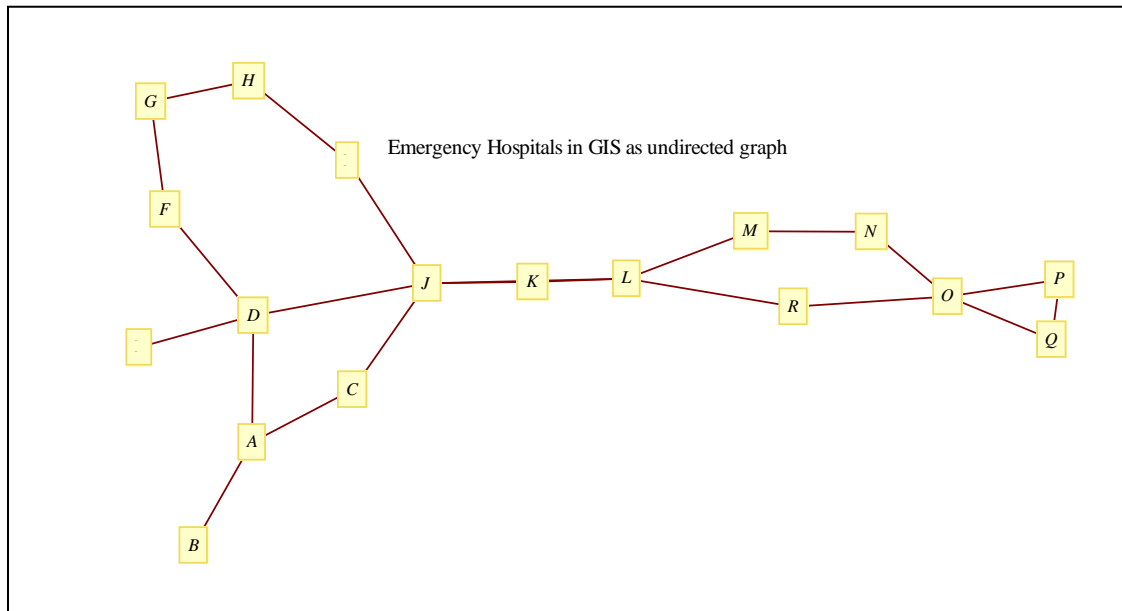
Figure 4-13 Graph Converted From Hospitals Distribution Map

For such cases it is much more appropriate to use fuzzy clusters where each node can belong to multiple clusters, too. By fuzzy clusters, they are described by their cluster centers, we can print out the representative of each cluster. It shows the relationship of nodes in the graph-based system introduced in Section 3.6.2. We know that fuzzy c-mean method can be applied in the dataset processing of PDMS, and making it clear for decision-making in post-disaster systems. GPS dataset point (x) was collected from a GIS based hospitals location dataset, which is assigned by the clustering number, fuzziness, and accuracy; we simulated and ran the results in Figure 4-8
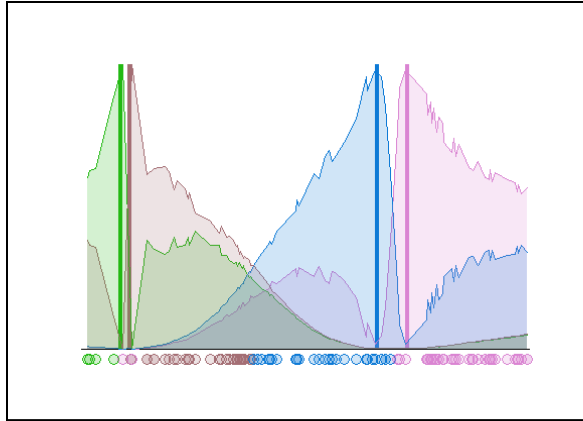
(a) [DataN=100, ClusterN=3, Fuzziness=2, Accuracy=0.3]
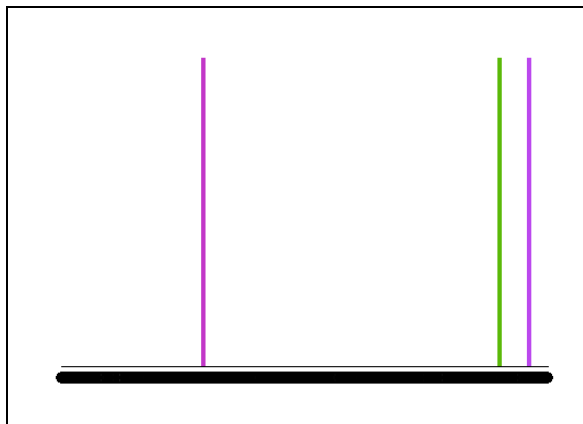
(a) Initialed: [DataN=1000, ClusterN=3, Fuzziness=2, Accuracy=0.3]
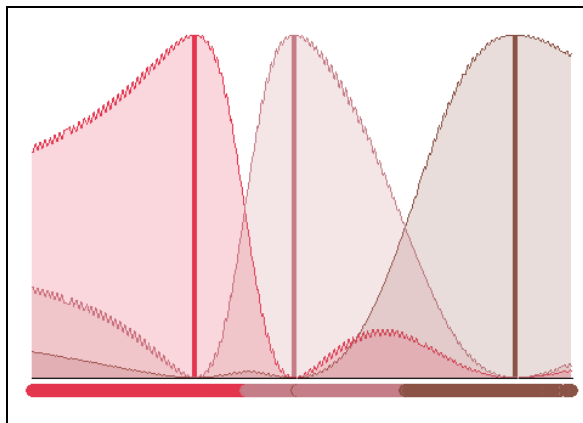
(b) Steps: [DataN=1000, ClusterN=3, Fuzziness=2, Accuracy=0.3]

Figure 4-14 Fuzzy C-mean Using GIS Based Hospitals Location Dataset

Suppose that we can identify the position as a net like Figure 4-14 (a), so we can start the algorithm from a node; it represents the hospital which the individuals are looking for in post-disaster shown in Figure 4-14 (b).

We can compare the Genetic Arithmetic based application from the Philippines; the researchers developed an android application based PDMS, based on its special geographical location and natural disasters. MyDisasterDroid's main function is to offer the optimal combination of routes, supplying the rescue team with the shortest and most efficient route to so they can aid as many victims as possible. The application calculates the route using the famous traveling salesman problem, in which given the limited number of travel destinations as well as the distance between each other, traveling salesman most choose the route to make the trip to meet a cost and time minimum (see Figure 4-15).
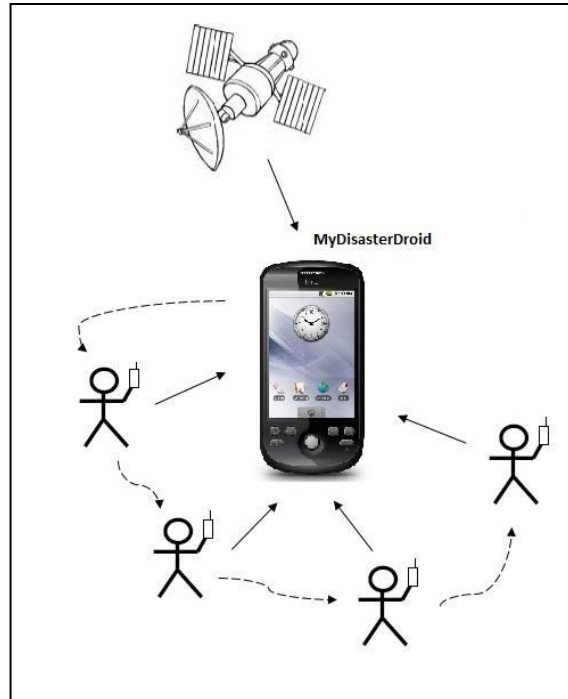
Figure 4-15 Frameworks of Shortest-Path Searching Based MyDisasterDriod Project
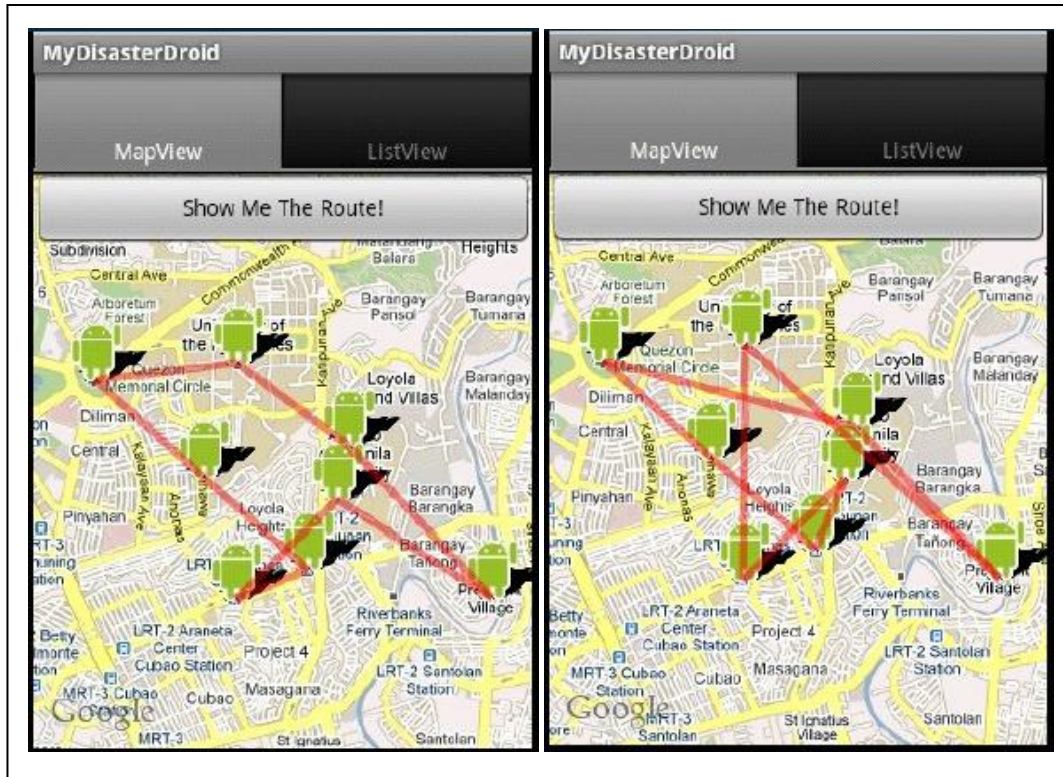
Figure 4-16 The Main Interaction of MyDisasterDriod Project by Path Calculation
Source: Mobile Android based System.
Http:// www.wseas.us/e-library/transactions/communications/2010/89-544.pdf

(a) Mutation = 200 (b) Mutation =2

Figure 4-17 Path Calculation of MyDisasterDriod Project by GA Algorithm
Source: Mobile Android based System.
Http:// www.wseas.us/e-library/transactions/communications/2010/89-544.pdf

In addition to this standard shortest route outside, MyDisasterDroid passion for each location will be used as an important indicator of the number, and gives different weights in order to get the best rescue effect.

Another project is VeRSiert is a full project rather than simply an android application. The project started in May 2008 and was hosted by the German government, "Public Safety Research" under the framework of the subprojects. It mentions government programs to respond to disasters and crises, but it is positioning itself as a response to large-scale public events management software. VeRSiert focuses on how to provide users with added value, including

micro-blogging, geographic positioning, multicast and broadcast information, public transport, electronic ticketing, mobile payment which is a static system and cannot calculate dynamically for post-disaster management.

4.4 Effectiveness of Graph-Based Semi-Supervised Learning by Gaussian Fusion Dataset

Based on the FCM algorithm mentioned above, now we discuss the effectiveness of semi-supervised learning method on Gaussian fusion based dataset that was introduced in Section 3.3; we only adopt 3 Gaussian mixed models that we discussed in Section 3.3.2 Multisource Gaussian Mixed Model

To show the Gaussian model's effectiveness, we initialed some moons type dataset for semi-supervised learning, shown as in Figure 4-18.

(a) Input (Data Type= 2 Moons)   (b) Graph Generated

(c) Outputs By SSL

Figure 4-18 2 Moon Type Dataset Simulation Results by SSL Algorithm

If we adopt 2- Gaussian Mixed dataset we have the results shown as Figure 4-10.

(a) Input (Data Type= 2 Gaussian Mixed)   (b) Graph Generated

(c) Outputs By SSL

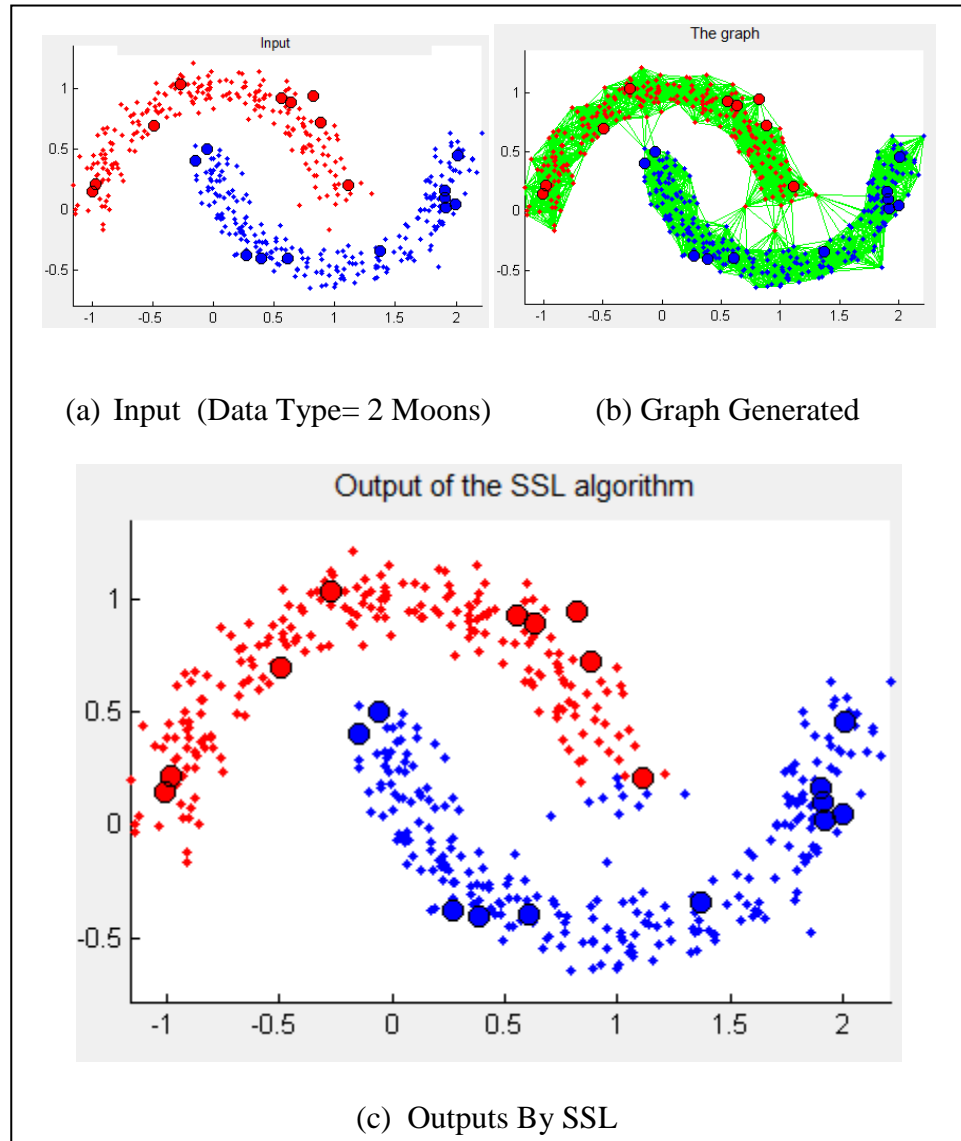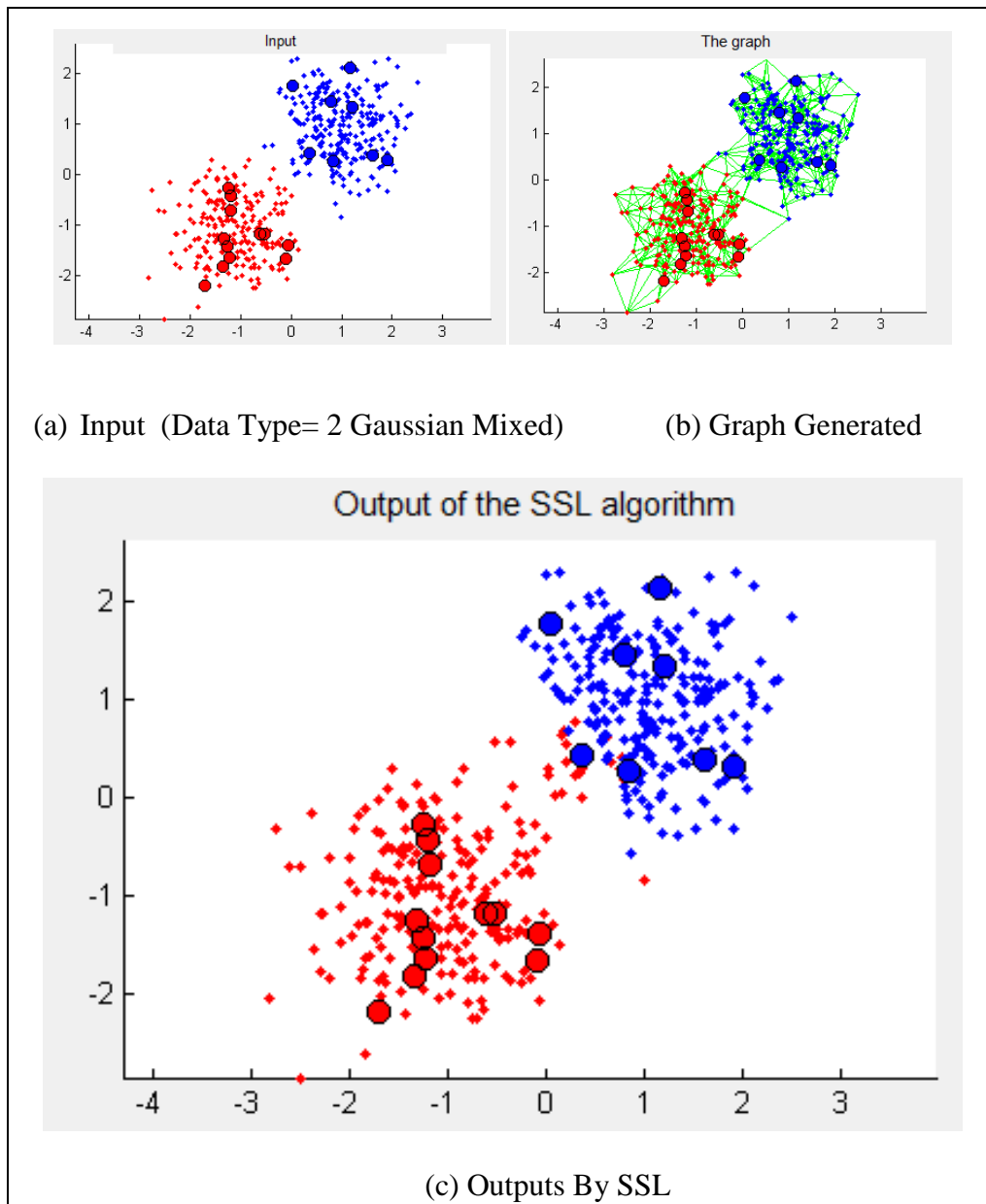Figure 4-19 2- Gaussian Fusion Type Dataset Simulation Results by SSL Algorithm

If we adopt 3- Gaussian Mixed dataset we have the results shown as Figure 4-20.

(a) Input (Data Type= 3 Gaussian Mixed)   (b) Graph Generated

(c) Output By SSL
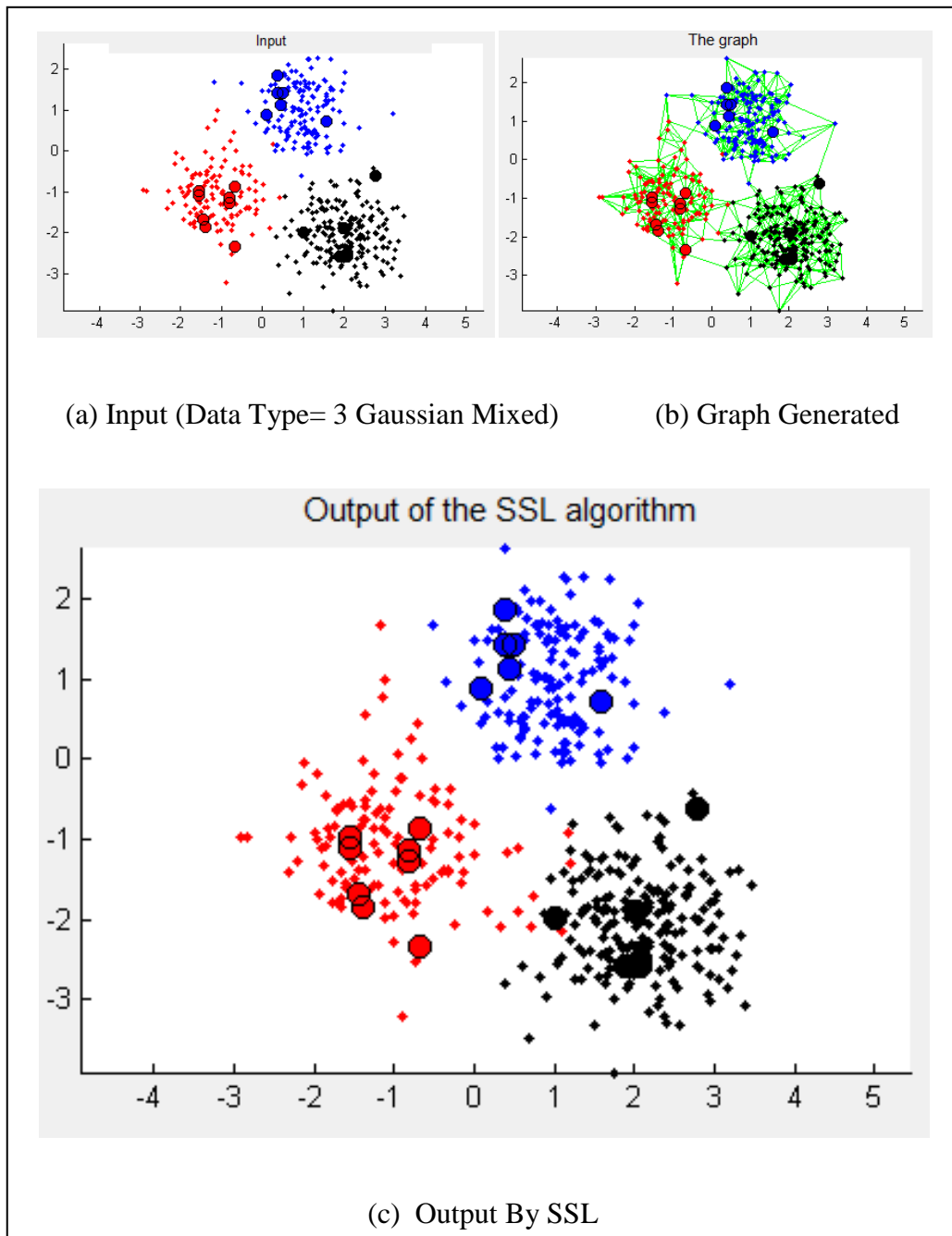
Figure 4-20 3-Gaussian Fusion Type Dataset Simulation Results by SSL Algorithm

The results show in Table 4-4.

Table 4-4 2 Parameters of Moon 2-Gaussian and 3-Gaussian Datasets Under SSL

Parameters	2-Moon dataset	2-Gaussian	3-Gaussian
Dimension	3	3	3
Number of Labeled Points	20	20	20
Graph Type	Sym. KNN	Sym. KNN	Sym. KNN
Number of Neighbors	15	15	15
Kernel Width	2	2	2
Graph Statistics			
Number of Edges between Classes	45	13	98
Number of Edges within Classes	7138	1609	4826
Weights between Class	44.29	8.92	89.87
Weights within Class	7097.13	1485.98	4596.67
Connected Components	1	1	1
SSL Regularization Parameter	1	1	1
SSL Results			
Training Error	5.00%	0.50%	1.00%
Test Error	2.30%	4.60%	2.10%
Not Labeled	0.00%	0.00%	0.00%

4.5 Discussion

Determined the quantities of constituents found in each of the three types of dataset and some optical properties. Altogether, there are 19 input attributes, all of which are numerical. The goal predicate is Boolean categorical with three different classes/labels corresponding to the three hospitals. Accordingly, we use fuzzy predicates induced by appropriate fuzzy sets on the domains of the numerical input attributes. Although the goal predicate is Boolean categorical, the use of fuzzy predicates for the input predicates is not meaningless. The reason is that using fuzzy sets allows us to more easily model regions of overlapping goal classes and in a more natural way than by splitting the numerical attributes into Boolean classes.

4.6 Summary and Conclusions

In this chapter we simulated some algorithms introduced in Chapter III. All of the algorithms will improve the current PDMS including disaster data-set collection, scaling the disaster map that will benefit the continuous calculation for individuals to find the best resources. Fuzzy c-mean method was applied in improvement of hospital resource distribution that will be useful for shortest path finder algorithm to locate the resource in emergence conditions. Additionally, the effectiveness of the Gaussian shaped dataset in fusion process by comparing with other types dataset is shown. So all important algorithms including Gaussian fusion, fuzzy models (Fuzzy c-mean and SSL by fuzzy model) were simulated in this chapter show that all these algorithms will be more important to improve the current PDMS system, especially for GIS based PDMS. This dissertation proposed some improved algorithms in different phases of disaster, data collection, data processing, and data organization; the algorithms proposed in this dissertation must be integrated in an application system to show an overall performance of the new PDMS that the nest future works we need to do. Summarized all the works in this chapter, we have that,

(1)   We simulated the data fusion algorithm combining Gaussian operations on disaster map. The target map and resource map were combined and decreased the population of the disaster map. It showed the integrated map will scale an area to aid individuals in locating more suitable resources (hospitals).

(2)   For different dataset coming from different sources, multisource fusion was developed and simulated, and for multimodal fusion, it is only needed to normalize the data and then go to the multisource fusion process, that the reason we must not to simulate this algorithm in this

chapter. For an emergency node, evaluation data must be integrated to show an overall

performance in post-disaster, so fuzzy c-mean was simulated for shortest path to locate the

resource, we also introduced a project to show the effectiveness and feasibility of the algorithm

to present the data of each node in graph.

(3)　　Effectiveness of the Gaussian dataset also was simulated by Matlab 6.0 program. The

results showed that Gaussian dataset will be more stable and there will be fewer edges between

classes, weights between class, and less training error under the same initial conditions.

Several aspects of PDMS' algorithms were improved by simulation in this chapter, as we

know that all these algorithms improved the system in different phases which will potentially

improve the application system in future developing works. There were additional shortcomings

in these algorithms, such as, the SSL organized graph needed high maintenance of works before

disaster, and Gaussian dataset also cost more for preprocessing.

CHAPTER 5: SURVEY DESGIN, FUZZY ANALYSIS AND SIMULATIONS FOR EMERGENCY HOSPITAL PERFORMANCE IN POST-DISASTER

5.1 Introduction

Fuzzy factors analysis on matrix type dataset is to conduct a sub-dataset indexing by some indices using fuzzy transformation; the survey dataset in this chapter will be processed by fuzzy factors analysis, now we started the steps as follows.

In order to conduct sub-dataset indexing by some indices using fuzzy transformation, it is necessary to use fuzzy factor analyst on matrix type dataset. The survey dataset in this chapter were processed by fuzzy factor analysis, including the following steps:

5.1.1 Steps of Evaluation

Let the universe set be the P evaluation indices, and noted that $u = \{u_1, u_2, \cdots \cdots, u_p\}$

Let the evaluation level be $v = \{v_1, v_2, \cdots \cdots, v_p\}$, each level relative to a fuzzy subset.

A) Establish the fuzzy relation matrix R

Quantify u_i $(i = 1, 2, \cdots \cdots, p)$, i.e. it is to calculate the fuzzy membership of each index for the evaluation object $(R \mid u_i)$ and to continue to get fuzzy relationship matrix:

$$R = \begin{bmatrix} R \mid u_1 \\ R \mid u_2 \\ \cdots \\ R \mid u_p \end{bmatrix} = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1m} \\ r_{21} & r_{22} & \cdots & r_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ r_{p1} & r_{p2} & \cdots & r_{pm} \end{bmatrix}_{p.m} \qquad (5.1)$$

,where the i -th row and j -col element, r_{ij} , in matrix R , denotes an object was evaluated by factor u_i of level v_j 's membership and an object's aspect in u_i was calculated by fuzzy vector $(R \mid u_i) = (r_{i1}, r_{i2}, \cdots \cdots, r_{im})$. In other algorithms, the evaluation was calculated by only one factor, so the fuzzy factors evaluation needs more information from the matrix.

B) Evaluation factors are weighted

In fuzzy comprehensive evaluation process, the weight vector $A = (a_1, a_2, \cdots \cdots, a_p)$ is needed where the element a_i in A is the membership of factor u_i for evaluation objects. For a multi-level evaluation process, analytic hierarchy was used in order to sort the importance of factors and to decide the weights. Let the normal weights be

$$\sum_{i=1}^{p} a_i = 1 , a_i \geq 0 , i = 1, 2, \cdots \cdots, n .\qquad(5.2)$$

C) Result vector by fuzzy synthetic evaluation

A and R of object were synthesized using given operator to get fuzzy synthetic evaluation B , i.e.

$$A \square R = (a_1, a_2, \cdots \cdots, a_p) \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1m} \\ r_{21} & r_{22} & \cdots & r_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ r_{p1} & r_{p2} & \cdots & r_{pm} \end{bmatrix} = (b_1, b_2, \cdots \cdots, b_m) = B \qquad(5.3)$$

, where b_1 is calculated by A and the j -th Col of R which denotes the membership of the level set v_j on the evaluation object.

D) Analysis on vector of fuzzy comprehensive evaluation

The maximum membership principle is the most common method in practice, however in certain cases it will not work effectively and will lose much of the information, potentially leading to inaccurate evaluation results. Due to this, the weighted average method was proposed for seeking membership level for many objects, followed by the evaluation according to their sorted levels.

5.1.2 The Weights by Analytic Hierarchy

It is important to use a comprehensive evaluation to find the weights. Fuzzy analytic hierarchy method is a proven and effective method of determining the weight coefficients in practice. It is particularly suitable for difficult to obtain quantitative indicators to analyze complex problems. There are many interrelated and layered factors involved in considering principles based on the objective reality of the fuzzy judgment and the relative importance of each level while still maintaining a quantitative representation. Mathematical methods can be used to determine the weights of all elements in the relative order of importance. The steps were introduced as follows:

A) Let object's evaluation index be $u = \{u_1, u_2, \cdots \cdots, u_p\}$.

B) Establish the judgment matrix: the element value of judgment matrix reflects the individuals' understanding of the relative importance of each element, generally as 1-9 and their reciprocal scaling method. But the importance of factors can be compared with each other with a meaningful description of the ratio, the value of the corresponding element of judgment matrix and then take the ratio $S = \left(u_{ij}\right)_{p \times p}$.

C) Calculate the judgment matrix: calculate the maximal eigenvalue λ_{max} of S and eigenvector A; actually, eigenvector A is the distribution of weights.

D) Consistency check: set consistency index $CI = \dfrac{\lambda_{max} - n}{n - 1}$, and average consistency random index RI. This method is used to construct a random matrix by a plurality of samples; the random structure and inverse scaling of the sample fills the upper triangular matrix. The value of the main diagonal is always one. In correspondence to the position of entry transpose, the random number uses the inverse of the corresponding position. Then in order for each random sample matrix to calculate the consistency index values obtained for these values, the average random consistency index values RI. While the random consistency ratio

$CR = \dfrac{CI}{RI} < 0.10$, the sort that results in satisfactory consistency, namely the distribution of weights is reasonable; other else to adjust the value judgment matrix elements of redistribution of weight coefficient values.

5.2   Survey Design and Data Collection

The survey was designed for the emergency ability of hospitals in Central Florida Area, Volusia County and Flagler County. About ten hospitals were considered for involvement in this survey and seven of them were selected for feedback through the survey forms. A self-administered questionnaire was used to collect data. The surveys included nine indexes and each index included eight questions that required every respondent to answer. Some questions must not be marked; the survey questionnaire was randomly distributed to people, and the questionnaire was independently completed, and each questionnaire was validated. From the 70 dispersed surveys, 59 were returned, resulting in a return rate of 84%; 54 forms were validated;

the validation rate was 91.5%. The respondents included doctors, nurses, patients and administrators of varying ages. The nine group indexes are listed in Table 5-1.

Table 5-1 Nine Indexes for the Emergency Ability of Hospitals in Post-Disaster

Group	Factors
GP-1	Decision-making system responsible for command and control
GP-2	Clear, accurate and timely communication of hospital in post-disaster
GP-3	Well-developed safety and security procedures
GP-4	The abilities for maintaining patient triage operations of hospital
GP-5	Surge capacity in post-disaster
GP-6	The ability of the continuity of essential services
GP-7	Human resources system
GP-8	Logistics and supply management for the hospital in post-disaster
GP-9	The ability for post-disaster recovery

The eight questions referred to as sublevel factors are listed in Table 5-2. Table 5-2 only listed the first group factor GP-1. The rest refers to the Appendix D which listed all factors and levels.

Table 5-2 Sub-Factors of Group 1 Including Eight Factors (Questions)

Indexes	Sub-Factors
G11	Incident command team ability to respond
G12	Hospital command center ability
G13	The basic qualifications of disaster managers and employees
G14	Members of emergency response teams are sufficient trained
G15	Coincided with the World Health Organization (WHO) standards.
G16	Focal point continuity assurance capabilities.
G17	Strict compliance with the basic principles and accepted strategy
G18	Ensuring proper management and coordination of activities
GP-9	The ability for post-disaster recovery

For each question, the respondents needed to mark a score by 5- Likert Scale (1, 3, 5, 7, 9); A semantic for understanding using 4 levels can also be set: Good, median, normal and bad which is shown in Table 5-3.

Table 5-3 Quantitative Evaluation of Grading Standards

Score	Evaluation	Level
$x_i > 8.5$	Good	$E1$
$7.5 < x_i \leq 8.5$	Median	$E2$
	Normal	$E3$
$6.5 < x_i \leq 7.5$	Bad	$E4$
$x_i \leq 6.5$		

With the survey data, fuzzy comprehensive evaluation was applied to calculate the performance of the emergency ability of hospitals in post-disaster. Evaluation of the object is to determine the factors of the evaluation set. But the GP-3, GP-4, and GP-7 received few responses throughout the survey process, leaving for consideration onlyGP-1, GP-2, GP-5, GP-6, GP-8,

and GP-9 which was shown in Table 5-4. Table 5-4 averaged all the evaluation data set

of 54 records.

Table 5-4 The First Level Factors for the Emergency Ability of Hospitals in Post-Disaster

GP-1	GP-2	GP-5	GP-6	GP-8	GP-9
7	7	7.25	7.25	7.25	7
6.25	6.6	8	8.333	8.2	7.667
7.857	7	7.667	8.333	8.2	8.667
7	8	7	9	7.4	7.667
7.857	8.5	8.333	5.667	8.2	9
7	7	7	7.667	9	8
8.143	9	8.333	9	6.6	8
8.429	7.5	7	6.667	9	7.667
9	9	8.667	6.667	6.6	7.667
8.143	9	6	7.333	5.4	5.667
8.143	7	8	8.333	7	9
8.714	9	8.667	8.333	9	7.333
7.286	6.5	7.333	8	8.6	8
6.429	7.5	8	8.333	9	9
6.25	6.6	8	8.333	8.2	7.667
7.857	7	7.667	8.333	8.2	8.667
7	8	7	9	7.4	7.667
7.857	8.5	8.333	5.667	8.2	9
7	7	7	7.667	9	8
8.143	9	8.333	9	6.6	8
8.429	7.5	7	6.667	9	7.667
9	9	8.667	6.667	6.6	7.667
8.143	9	6	7.333	5.4	5.667
9	8.5	9	8.333	9	7.667
6.25	6.6	8	8.333	8.2	7.667
7.857	7	7.667	8.333	8.2	8.667
7	8	7	9	7.4	7.667
7.857	8.5	8.333	5.667	8.2	9
7	7	7	7.667	9	8
8.143	9	8.333	9	6.6	8

GP-1	GP-2	GP-5	GP-6	GP-8	GP-9
8.429	7.5	7	6.667	9	7.667
9	9	8.667	6.667	6.6	7.667
8.143	9	6	7.333	5.4	5.667
9	8.5	9	8.333	9	7.667
9	9	8.667	6.667	6.6	7.667
8.143	9	6	7.333	5.4	5.667
9	8.5	9	8.333	9	7.667
6.25	6.6	8	8.333	8.2	7.667
7.857	7	7.667	8.333	8.2	8.667
7	8	7	9	7.4	7.667
7.857	8.5	8.333	5.667	8.2	9
7	7	7	7.667	9	8
8.143	9	8.333	9	6.6	8
8.429	7.5	7	6.667	9	7.667
9	9	8.667	6.667	6.6	7.667
8.143	9	6	7.333	5.4	5.667
8.143	9	8.333	9	6.6	8
8.429	7.5	7	6.667	9	7.667
9	9	8.667	6.667	6.6	7.667
8.143	9	6	7.333	5.4	5.667
9	8.5	9	8.333	9	7.667
9	9	8.667	6.667	6.6	7.667
8.143	9	6	7.333	5.4	5.667
9	8.5	9	8.333	9	7.667

By expert and the confidence dataset from the survey form collection, we listed the weighted values on each second level factor in Table 5-5 and the first level factors' weights were calculated by the steps in Section 5.3.

Table 5-5 Two Grades of Evaluation Factors of Hospital Emergency Ability and Weighting

Factors	Evaluation Factors	Weights	Factors	Evaluation Factors	Weights
GP-1	G11	0.143	GP-6	G61	0.113
0.212	G12	0.121	0.176	G62	0.224
	G13	0.178		G63	0.135
	G14	0.221		G64	0.115
	G15	0.13		G65	0.119
	G16	0.121		G66	0.203
	G17	0.178		G67	0.214
	G18	0.213		G68	0.132
GP-2	G21	0.213	GP-8	G81	0.103
0.204	G22	0.123	0.216	G82	0.12
	G23	0.132		G83	0.12
	G24	0.225		G84	0.125
	G25	0.21		G85	0.23
	G26	0.211		G86	0.2
	G27	0.303		G87	0.178
	G28	0.129		G88	0.189
GP-5	G51	0.121	GP-9	G91	0.116
0.178	G52	0.213	0.165	G92	0.165
	G53	0.222		G93	0.167
	G54	0.122		G94	0.125
	G55	0.121		G95	0.168
	G56	0.204		G96	0.21
	G57	0.233		G97	0.227
	G58	0.121		G98	0.154

5.3 Index Weights Fuzzy Analysis Steps

The following steps are used to get each factor's weight: 1) Determine the evaluation

object set P =Emergency Abilities of Hospitals

2) Structural evaluation factors set: $u = \{u_1, u_2, \cdots \cdots, u_6\} = \{\text{GP-1}\ldots \text{GP-9}\}$

3) Determine the domain level reviews: $v = \{v_1, v_2, \cdots \cdots, v_4\} = \{$Good, Median, Normal,

Bad$\}$

4) Weight calculation First level index

Construct judgment matrix of six factors:

$$S = \left(u_{ij} \right)_{p \times p} = \begin{bmatrix} 1 & \frac{4}{3} & \frac{5}{4} & 1 & \frac{9}{5} & \frac{6}{5} \\ \frac{3}{4} & 1 & \frac{9}{10} & \frac{8}{9} & \frac{7}{5} & \frac{8}{9} \\ \frac{4}{5} & \frac{10}{9} & 1 & \frac{4}{5} & \frac{3}{2} & 1 \\ 1 & \frac{9}{8} & \frac{5}{4} & 1 & 2 & \frac{5}{4} \\ \frac{5}{9} & \frac{5}{7} & \frac{2}{3} & \frac{1}{2} & 1 & \frac{4}{6} \\ \frac{5}{6} & \frac{9}{8} & 1 & \frac{4}{5} & \frac{6}{4} & 1 \end{bmatrix} \tag{5.4}$$

The maximal eigenvalue of the judgment matrix is $\lambda_{\max} = 6.00589$ by Mathematica 9.0

and continue to calculate the $CI = \dfrac{\lambda_{\max} - n}{n - 1} = \dfrac{6.00589 - 6}{6 - 1} = 0.001178$ and $RI = 1.24$ 。 So the

ratio is $CR = \dfrac{CI}{RI} = \dfrac{0.001178}{1.24} = 0.00095 < 0.10$, so it was regarded that the results of the weighted

processing were reasonable. The relative eigenvector of $\lambda_{\max}$ is

$$A_0 = \left(1.21372, 0.935715, 0.9911, 1.21138, 0.634379, 1.0 \right) \tag{5.5}$$

Normalized and we have that,

$$A = \left(0.202, 0.156, 0.165, 0.202, 0.109, 0.166 \right) \tag{5.6}$$

119

5) Calculate the weights of sub-level index.

Similarly, we still use the AHP method to find the index weights. Two indicators were constructed for each of their own judgment matrix, and then calculate the maximum eigenvalue of Mathematica and consistency test. Come to a reasonable weight coefficient. The GP-1 weighted vector is

$$\{0.143, 0.121, 0.178, 0.221, 0.130, 0.121, 0.178, 0.213\} \tag{5.7}$$

Normalized and we have that,

$$\{0.110, 0.093, 0.136, 0.169, 0.100, 0.093, 0.136, 163\} \tag{5.8}$$

, and continue to calculate GP-2, GP-5, GP-6, GP-8, GP-9.

6) Multi-level fuzzy comprehensive evaluation

Synthesized using M $(\circ, \oplus)$ the weighted average fuzzy operator A and R will be synthesized and foot fuzzy comprehensive evaluation result vector. Commonly used fuzzy comprehensive evaluation is used to take large or small algorithm in many factors, each share of the weight factors is often very small. In the fuzzy synthetic operation, a lot of information is lost, the results are not easily distinguishable and often leads to irrational (i.e. model failure) situations. So, for the above mentioned problem, the weighted average type fuzzy synthesis operator is used. The formula is:

$$b_i = \sum_{i=1}^{p} \left(a_i \cdot r_{ij} \right) = \min \left(1, \sum_{i=1}^{p} a_i \cdot r_{ij} \right), j = 1, 2, \cdots, m \tag{5.9}$$

So we have that $A_1 = a \circ R = (0.130, 0.320, 0.307, 0.167)$ and by normalization we have that $(0.141, 0.346, 0.332, 0.187)$, so we get the overall evaluation weight for GP-1 is:

$V_{GP-1} = 4 \times 0.141 + 3 \times 0.346 + 2 \times 332 + 1 \times 181 = 2.12$, and continue to calculate the GP-2,

GP-5, GP-6, GP-8 and GP-9 which were also listed in Table 5-5.

So the overall performance of emergency ability of each hospital in post-disaster was presented by fuzzy set,

$$E = \frac{v(GP-1)}{W_1} + \frac{v(GP-2)}{W_2} + \frac{v(GP-5)}{W_5} + \frac{v(GP-6)}{W_6} + \frac{v(GP-8)}{W_8} + \frac{v(GP-9)}{W_9} \quad (5.10)$$

Or simply calculated as,

$$E = \sum_{i=1}^{9} W_i v(GP-i), \ W_3 = W_4 = W_7 = 0 \quad (5.11)$$

The most commonly used method in practice is the principle of maximum degree, but this method use is conditional as there are issues of validity and it may draw unreasonable evaluation results. According to the principle put forward, the weighted average method of seeking membership level, the principles for the use of a weighted average of these levels of evaluation to analyze the results of the evaluation. The results of this method with the principle of maximum degree the results obtained by the method is quite different, but the results were more in line with the actual situation.

5.4 Simulation on IF-THEN Rules Presented Graph Using Survey Data-Set

We labeled seven hospitals to be {H1, H2, H3, H4, H5, H6, H7}; the value of each node (hospitals) is the performance data which was calculated in Section 5.3 and it was {7.35, 7.65, 6.88, 8.20, 8.54, 7.12, 8,82}, the distance of each hospital will be labeled on the available road on the graph. The IF-THEN will draw the relationship between each hospital, the processing was

described in Section 3.5.2, and the fuzzy operation results were the value of emergency for a

suitable resource such as "IF Hospital A is good THEN Hospital B is better, Cost is X". The best

results are calculated by each node one by one and finally, the totally cost and the best node were

located for individuals. Node A is the start node with the cost calculated by the value of the

performance of each hospital and the distance as

$$Cost = v(GP - i) \cdot a + \frac{15}{dis} \cdot (1 - a) \tag{5.12}$$

We now let $a = 0.9$, and calculated all the costs between two nodes whereas "-" is

unknown or maximal, the cost matrix was.

$$\text{Cost=} \begin{bmatrix} 0 & - & 6.5 & - & - & - & 15 \\ - & 0 & 8 & 7.5 & - & - & - \\ 6.5 & 8 & 0 & - & 7 & 4.5 & - \\ - & 7.5 & - & 0 & 4.5 & - & - \\ - & - & 7 & 4.5 & 0 & 4.5 & - \\ - & - & 4.5 & - & 4.5 & 0 & 2.5 \\ 15 & - & - & - & - & 2.5 & 0 \end{bmatrix} \tag{5.13}$$

122

And the rule set is:

IF H1 THEN H7, Cost is15
IF H2 THEN H3, Cost is 8
IF H3 THEN H5, Cost is 7
IF H4 THEN H5, Cost is 4.5
IF H5 THEN H6, Cost is 4.5
IF H1 THEN H3, Cost is 6.5
IF H3 THEN H6, Cost is 4.5
IF H4 THEN H2, Cost is 7.5
IF H6 THEN H7, Cost 2.5
IF P THEN H6, Cost is 12

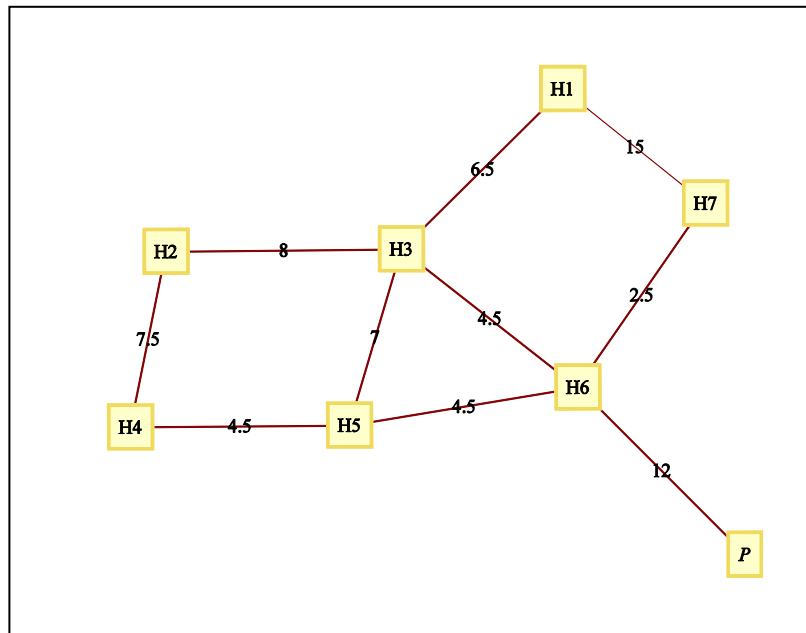We draw the non vector graph by these rules. The node is the location of hospitals and the value on the edge was the cost shown in Fig 5-1

Figure 5-1 The Graph for Indivuduals and Hospitals Based on the Cost Calculation
The process starts at node P to the resource location H2 in Fig 5-2.
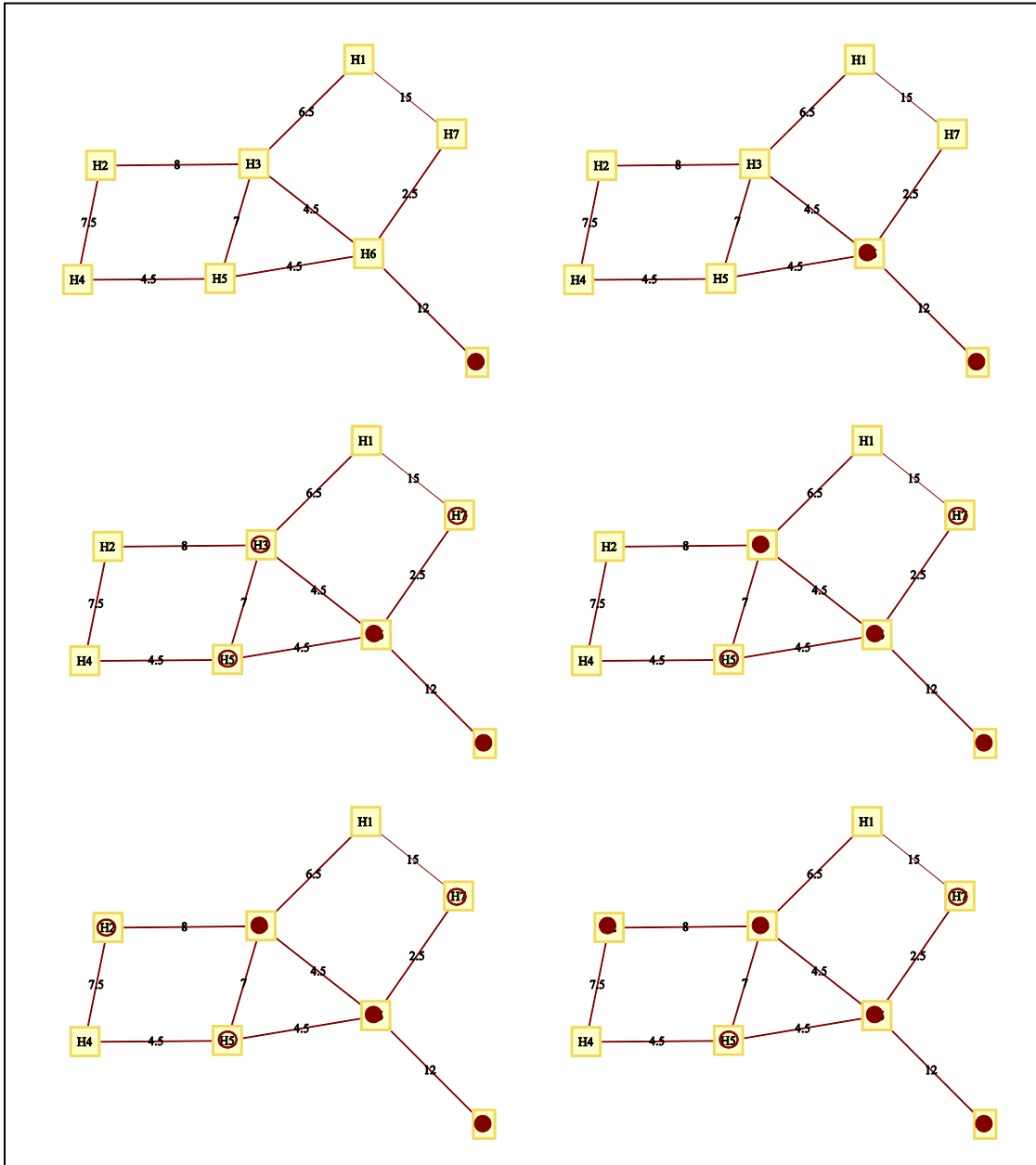
Figure 5-2 Calculation Process Using Shortest Path Method Between P and H2

"○"is the template node in the path, and "●" is the confirmed node in the path.

The results showed that following the cost calculation, it was easier for the individual to locate the resource using shortest path method.

5.5   Summary and Concluding Remarks

This chapter collected the evaluation data from hospitals using a multi-level index survey system; the indexes were converted to factors which were evaluated by the Likert Scale nine point method. By using multi-level fuzzy factor analysis, the weights were determined; the weights decided the importance of the factor in the evaluation process for hospitals' performance in post-disaster. By a matrix operation, the final overall evaluation dataset was acquired and also was applied in the simulation works. In simulation process, we assigned seven hospitals to be involved in the system for individuals to locate that was also consistent with the survey results. Of course some problems in the survey existed; the next work is to collect the more data from hospitals and to involve more hospitals. The weight system will be improved and for a more exact evaluation result.

CHAPTER 6: SUMMARY, CONCLUSIONS AND FUTURE SCOPE

6.1 Summary and Conclusions

Disasters are hazardous events that are the result of the uncertainties of nature and when such disasters intersect with human society then it results in tremendous destruction, asserting innumerable casualties, and also becomes a key driver of unfortunate consequences. Urban areas are the most disaster prone areas and there is a higher possibility in such areas to expand into disasters because of the huge and greatly concentrated population. Urban areas are the most disaster prone areas and there is a higher possibility in such areas to expand into disasters because of the huge and greatly concentrated population. Additionally, the composite layout of transportation in urban areas and multi-story buildings make it more difficult and complex to offer aid to victims and to reach for them in affected locations. In this manner, the situation in such areas can be exacerbated and the probability of causalities in such areas is much higher. However, urban areas are not the only ones affected. If disasters occur in small areas and places, then also it becomes catastrophic as disasters always result in the loss of life and property, whether moderate or extreme. It can be concluded from the above discussion that data fusion in context of post-disaster management is essential and will be of much importance as it helps in generating information that different researchers have submitted as the result of their research work.

 In this dissertation, current post-disaster management systems were analyzed, and in each component of the PDMS, potential improvement of algorithms for data collection of disaster, resource distribution, and disaster map scaling were developed in order to improve the current systems. In these algorithms, we focused on human-centric approaches and showed its

promised application in the future post-disaster system using fuzzy models which were specialized for all PDMS algorithms.

6.1.1 *Gaussian Fusion Algorithms*

Research Gap: The need to verify the effectiveness of fusion algorithm based on GIS dataset especially for the features of dataset.

Research Finding: A Gaussian fusion based human-centric approach was introduced for scaling a disaster map that contributed the computing complexity while locating the resource map. And for multisource dataset, Gaussian Mixed Model was applied for the preprocessing and integration of all data set into a Gaussian model using fuzzy operators. Theoretic inference was also finished using expectation maximum algorithm. And for original data set, PCA was also applied to make data in the same size. This method loses much useful information and makes the final results unbelievable, but we can increase the population of the data set to overcome the shortcomings of this method. Gaussian fusion algorithm will also make the calculation more complex; we need to take the balance between the cost for data integrating and the calculation for the disaster scaling.

Conclusion: We integrated the resource map and disaster map using Gaussian fusion algorithms and human factors as the parameters throughout of the process of the Gaussian fusion as density functions in calculation. Simulation works were finished and showed this method's effectiveness by comparing it with other traditional methods such as mean-mean. We also introduced multi-modal fusion algorithms for disaster dataset, and this algorithm is a novel algorithm for normal dataset, it was only needed to be normalization operation, and then converted to multisource data fusion algorithm.

128

<center>6.1.2 *Disaster Map Systems*</center>

Research Gap: The need to know how individuals received emergency resources or provider response in the post disaster stage. There were two technical improvements on the aspects. One is how to organize an effective map (node, edge and weight) to make the algorithms more suitable and output better results. Another is how to present the node that means how to formally present or mathematically present the knowledge of the resource, and furthermore to make the fuzzy model more feasible in the inference system.

Research Finding: Semi-supervisor learning based graph embedded in the disaster map system was introduced which will help to improve the algorithm in this dissertation, especially for the Gaussian type dataset. The simulation results also showed the effectiveness by comparing with other dataset such as 2-moons. And for presentation of the knowledge of each node in disaster map, "IF-THEN, x" rules were applied by using fuzzy operations. We introduced the rules integrating and inference processing, and also simulated the fuzzy inference system for the algorithm that we proposed in this dissertation, especially the fuzzy c-mean. This algorithm will make the presentation more centric.

*Conclusion***:** SSL will be more effective in presenting the graph system, but the organizing processing cost is high. For improving the interaction between the system and individuals in post-disaster situations, the server needs to take more computing task for the SSL based graph system that increase the speed at which individuals receive a response. The algorithm could potentially be improved by collecting plenty of data set in pre-disaster which will cut the cost of calculation of the disaster map in post-disaster stage.

6.1.3 *Post-Disaster Management System Improvement*

Research Gap: The need to develop and improve the current PDMS as a potential application and to show all simulation works effectively for the proposed algorithms.

Research Finding: In order to better understand the methods proposed in this dissertation, we continued to design a survey form for collecting the dataset for evaluating the emergency ability of hospital in post disaster and then we applied the data set in a graph-based inference system, fuzzy inference system was also applied for individuals to locate the emergency resource in post disaster. The survey form dataset firstly was preprocessed by different level matrix operating to get the weight for every factor which was used for evaluation the performance of the emergency ability of hospitals in post disaster, and then the weighted system was used for calculating the overall evaluation on the emergency ability using fuzzy presentation. The values of the evaluation were used in a graph system presented by "IF-THEN, x" fuzzy model. Simulation work also was done and showed the feasibility and effectiveness of the proposed method.

*Conclusion***:** All proposed algorithms were indicated and illustrated clearly in theoretic and simulation works in the dissertation. All of the algorithms will be applied in the future PDMS developing program, although there are many improvements which need to be considered.

6.1.4 *Benefits of Research*

The benefits of the research will be extended by both the scientific community and the practitioners in PDMS systems.

They are:

6.1.5 *Fusion Algorithms for GIS Dataset in PDMS*

- Scientific Knowledge: it was proved that Gaussian fusion was effective for the single data fusion and Gaussian Mixed fusion for multisource dataset fusion for the GIS dataset. An initial algorithm process was developed which will be benefit the current PDMSs

- Practitioners: PDMS system will be more effective when organizing a disaster map system using the proposed algorithms for GIS dataset-Gaussian Fusion and Gaussian Mixed Fusion.

6.1.6 *Disaster Map Systems*

- Scientific Knowledge: it was proven that SSL organized map and Fuzzy c mean inference node presentation method will be effective to show the disaster map system, also it was illustrated for each individual in post-disaster, the organized disaster map will shorten the response for the system and the individuals.

- Practitioners: current PDMS will be improved by the proposed SSL and FCM organized disaster map systems.

6.2 Future Works

This research is a first step in exploring the PDMS, as we know that our earth is an ambiguous and massive system, which has highly composite dynamics and interaction and over the period. The procedure of pleasant development between humans and the environment is constantly impacted by natural disasters like global earthquakes, tsunami, floods, droughts, landslides, rainstorms, volcanic erosion, soil erosion and various other natural disasters. However, the survey conducted has identified the fact that monitoring, early warning, mitigation and countering back is the biggest demand for the state and society. The survey has identified the trend that different authors have recognized, that natural disasters are attributable to serious harm

131

to society, individuals, life and property and are responsible for hampering the social and economic sustainable development. Thus, different researchers and authors have highlighted the importance of modern technology and information systems to assist in investigating natural disasters, monitoring them, prevention and control against them. In this context, it has been identified that most of them have highlighted the importance of spatial information technology and spits exceptional advantages in the examination, determination, mitigation and response to natural disasters, distribution and risk relate to disasters.

6.2.1 *The GIS Dataset Collection for PDMS*

The technologies manage and collect data at the time of disasters in distinct forms and are also useful in presenting data in an evaluative manner so that emergency planners and people associated with disaster management can easily interpret that data for designation and planning ways with respect to disaster management. Moreover, the role of technologies is pertinent in sorting data apart from its fusion, as they present the distinguishing data in common forms generated from different sources so that it will be easy to interpret it. Techniques like information extraction (IE), information retrieval (IR), information filtering (IF), data mining and decision support will be of complete help and facilitate emergency planning in a significant manner.

6.2.2 *Management System for Disaster Response Organizations*

It is substantial to underline that information overload and mismanagement is a challenge for disaster response organizations because irrelevant information and can distract and result in incorrect outcomes, and thus, through advanced technology it's filtering or formatting is necessary. If information systems are exploited for filtering and formatting disaster related

132

information, then it can possibly diminish the perception of users for its support in their tasks and planning. In addition to this, it has been identified that large numbers of researchers have proposed their own formulated models and strategies for the purpose of coping with the disasters and managing it in an effective and efficient manner. With the help of research and testing in areas that are more disaster prone or have suffered any disasters, researchers have formulated models for assessing risk of the disasters to issue warnings regarding them and to devise ways in which timely warning of serious disasters can be offered, which will offer due to government and people, as well to cope up with the situation.

6.2.3 *Social Networks and Internet Issues for PDMS*

Social networks and the internet will be of enormous help in collecting and distributing information at the time of disaster occurrence and updated information about the situation. In this manner, such networks are highly useful in generating factual information related to the disaster affected areas and people. It is noteworthy that if systems are designed carefully and used effectively then these will be of great help in disaster management, and this information will be helpful in future disaster management. Overall, it can be stated that the trends that have been identified from the surveyed techniques are assisted by recently generated science and technology that have made life easier for those responsible for responding to disaster and planning disaster management. It is clear from the discussion that information systems and technologies are exceedingly supportive in not only managing the disasters, but also in managing the data related to disasters.

6.2.4 *Evaluation Issues on Risks of Disasters*

Furthermore, another recognized trend is the formulation of models in terms of different technologies that aid in assessing risks of disasters. It is the advances in technologies that the authors have highlighted which have not strengthened before time warning capabilities for lessening natural hazards, but also utilize their aftermath for future management. The trend has been identified that expansion of spatial technologies, global communication and new information technologies has extended the availability and accessibility of information on natural disasters.

6.2.5 *Dataset Fusion Algorithms Improvement*

In future works, there are other aspects to be focused on and also the algorithms need to be improved. The integrating level for all of the algorithms needs to be improved. The algorithms in this dissertation were introduced and developed separately; we hope to integrate the dataset preprocessing algorithms and the inference systems. The next step of work which needs to be further explored is to develop an application system in a different way such as mobile computing based PDMS. Cloud computing also needs to be considered to enhance the performance of the server system for the PDMS.

6.2.6 *Big Data Issues on PDMS*

On the other hand, big data always being connected to mobile computing (MC) is an issue that makes data more effective for application in disaster management. MC involves mobile communications, the internet, databases, distributed computing technology; it enables a computer or other information intelligent terminal equipment in the wireless environment to achieve data transfer and resource sharing. Its role is to be useful, accurate, and make timely

information available at any time, any place and to any customers. This will dramatically change the way people communicate in post-disaster situations. We also introduced some improvement in the disaster management field for the future. For example, emergency data collection methods for post-disaster needs to be considered; big data fusion technology applied for constructing a database, visualization tools developed for mobile client, output for providing information to emergency management or first responders, and how general public access the information.

APPENDIX A: BASIC MATRIX TRANSFORM

Source: http://www.wikipedia.org/

(1) Laplacian Matrix Transform (Wiki)

In the mathematical field of graph theory the Laplacian matrix, sometimes called admittance matrix, Kirchhoff matrix or discrete Laplacian, is a matrix representation of a graph. Together with Kirchhoff's theorem it can be used to calculate the number of spanning trees for a given graph. The Laplacian matrix can be used to find many other properties of the graph; see spectral graph theory. Cheeger's inequality from Riemannian geometry has a discrete analogue involving the Laplacian Matrix; this is perhaps the most important theorem in Spectral Graph theory and one of the most useful facts in algorithmic applications. It approximates the sparsest cut of a graph through the second eigenvalue of its Laplacian.

Given a simple graph G with n vertices, its Laplacian matrix $L := (\ell_{i,j})_{n \times n}$ is defined as,

$$L = D - A$$

That is, it is the difference of the degree matrix D and the adjacency matrix A of the graph. In the case of directed graphs, either the indegree or outdegree might be used, depending on the application.

From the definition follows that:

$$\ell_{i,j} := \begin{cases} \deg(v_i) & if \quad i = j \\ -1 & if \quad i \neq j \ and \quad v_i \ is\ adjacent \quad to\ v_j \\ 0 & otherwise \end{cases}$$

, where $\deg(v_i)$ is degree of the vertex i.

The normalized Laplacian matrix is defined as:

137

$$\ell_{i,j} := \begin{cases} 1 & if \quad i = j \quad and \quad \deg(v_i) \neq 0 \\ \dfrac{-1}{\sqrt{\deg(v_i)\deg(v_j)}} & if \quad i \neq j \, and \quad v_i \, is \, adjacent \quad to \, v_j \\ 0 & otherwise \end{cases}$$

Here is a simple example of a labeled graph and its Laplacian matrix.

Labeled graph	Degree matrix	Adjacency matrix	Laplacian matrix

$$\begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\begin{pmatrix} 2 & -1 & 0 & 0 & \cdot \\ -1 & 3 & -1 & 0 & \cdot \\ 0 & -1 & 2 & -1 & \\ 0 & 0 & -1 & 3 & \cdot \\ -1 & -1 & 0 & -1 & \\ 0 & 0 & 0 & -1 & \end{pmatrix}$$

For a graph G and its Laplacian matrix L with eigenvalues $\lambda_0 \leq \lambda_1 \leq \cdots \lambda_{n-1}$

L is always positive-semidefinite ($\forall i, \lambda_i \geq 0; \lambda_0 = 0$).

The number of times 0 appears as an eigenvalue in the Laplacian is the number of connected components in the graph.

L is an M-matrix.

λ_0 is always 0 because every Laplacian matrix has an eigenvector $V_0 = [1,1,\cdots,1]$ that, for each row, adds the corresponding node's degree (from the diagonal) to a "-1" for each neighbor so that $LV_0 = 0$

The smallest non-zero eigenvalue of L is called the spectral gap.

If we define an oriented incidence matrix M with element M_{ev} for edge ℓ (connecting vertex i and j, with $i > j$) and vertex v given by

$$M_{ev} = \begin{cases} 1 & if \quad v = i \\ -1 & if \quad v = j \\ 0 & otherwise \end{cases}$$

Then the Palladian matrix L satisfies $L = M^T M$, where M^T is the matrix transpose of M.

The second smallest eigenvalue of L is the algebraic connectivity (or Fiedler value) of G The Laplacian is an operator on the function of g. When G is k-regular, $\ell = I - A/k$, where A is the adjacency matrix of G and I is an identity matrix. All matrices are $n \times n$ where n is the number of vertices in G.

(2) Tikhonov regularization, named for Andrey Tikhonov, is the most commonly used method of regularization of ill-posed problems. In statistics, the method is known as ridge regression, and, with multiple independent discoveries, it is also variously known as the Tikhonov–Miller method, the Phillips–Twomey method, the constrained linear inversion method, and the method of linear regularization. It is related to the Levenberg–Marquardt algorithm for nonlinear least-squares problems.

When the following problem is not well posed (either because of non-existence or non-uniqueness of x, $Ax = b$, then the standard approach is known as ordinary least squares and seeks to minimize the residual, $\|Ax - b\|^2$, where $\|\cdot\|$ is the Euclidean norm. This may be due to the system being over determined or underdetermined (A may be ill-conditioned or singular). In the latter case this is no better than the original problem. In order to give preference to a

particular solution with desirable properties, the regularization term is included in this minimization:

$$\left\| A x - b \right\|^2 + \left\| \Gamma x \right\|^2$$

for some suitably chosen Tikhonov matrix, Γ, In many cases, this matrix is chosen as the identity matrix $\Gamma = I$, giving preference to solutions with smaller norms. In other cases, low pass operators (e.g., a difference operator or a weighted Fourier operator) may be used to enforce smoothness if the underlying vector is believed to be mostly continuous. This regularization improves the conditioning of the problem, thus enabling a numerical solution. An explicit solution, denoted by $\hat{x}$, is given by:

$$\hat{x} = (A^T A + \Gamma^T \Gamma)^{-1} A^T b$$

The effect of regularization may be varied via the scale of matrix Γ. For $\Gamma = 0$ this reduces to the unregularized least squares solution provided that $(A^T A)^{-1}$ exists.

APPENDIX B: MATHEMATICA CODE FOR EM OF GAUSSIAN MIXED MODEL

```
Manipulate[
 Module[{d1, d2, sim, updates},
  SeedRandom[sr];
  d1 = NormalDistribution[μ1, σ1];
  d2 = NormalDistribution[μ2, σ2];
  sim = With[{musAndsigmas = RandomChoice[{p1, 1 - p1} → {{μ1, σ1}, {μ2, σ2}}, numsamples]},
    musAndsigmas[[All, 2]] RandomReal[NormalDistribution[], numsamples] +
     musAndsigmas[[All, 1]]];
  updates = NestList[em[♯, sim] &, {{-10, 5}, {5, 5}, {.2, .8}}, numiterations];
  Grid[
   {{Show[{Histogram[sim, Automatic, "PDF", PlotRange → {{-30, 30}, {0, Automatic}}],
       Plot[{updates[[-1, 3, 1]] * PDF[Apply[NormalDistribution, updates[[-1, 1]]], x],
         updates[[-1, 3, 2]] * PDF[Apply[NormalDistribution, updates[[-1, 2]]], x]},
        {x, -30, 30}, PlotStyle → {{Red, Thick}, {Blue, Thick}}, PerformanceGoal → "Speed",
        PlotRange → All]}, ImageSize → {400, 275}, AxesOrigin → {-30, 0}]},
     {Grid[
       {{Panel[TableForm[updates[[-1, 1 ;; 2]],
          TableHeadings → {updates[[-1, 3]], {"μ", "σ"}}]]}}]}}]
  ],
 Style["first distribution", Bold],
 {{μ1, 8}, -20, 20, ImageSize → Tiny, Appearance → "Labeled"},
 {{σ1, 5}, 1, 10, 1, ImageSize → Tiny, Appearance → "Labeled"},
 Delimiter,
 Style["second distribution", Bold],
 {{μ2, -12}, -20, 20, ImageSize → Tiny, Appearance → "Labeled"},
 {{σ2, 3}, 1, 10, 1, ImageSize → Tiny, Appearance → "Labeled"},
 Delimiter,
 Style["probability", Bold],
 {{p1, .3, "p"}, .1, .9, ImageSize → Tiny, Appearance → "Labeled"},
 Delimiter,
 Style["number of samples", Bold],
 {{numsamples, 100, ""}, 50, 200, 10, ImageSize → Tiny, Appearance → "Labeled"},
 Style["number of iterations\nof EM algorithm", Bold],
 {{numiterations, 3, ""}, 1, 20, 1, ImageSize → Tiny, Appearance → "Labeled"},
 Delimiter,
 {{sr, 326, "random seed"}, 1, 9999, 1, ImageSize → Tiny, Appearance → "Labeled"},
 ControlPlacement → Left,

 Initialization :→ (

  problists[σ_, y_] := Block[{probs, totalprobs},
    probs = Table[σ[[3, i]] * Map[PDF[Apply[NormalDistribution, σ[[i]]], ♯] &, y], {i, 2}];
    totalprobs = Total[probs];
    Map[♯ / totalprobs &, probs]];
  pi[j_, p_] := Mean[p[[j]]];
  emu[sim_, j_, p_] := Total[sim * p[[j]]] / Total[p[[j]]];

  emstd[sim_, j_, u_List, p_] := Total[(sim - u[[j]])^2 * p[[j]]] / Total[p[[j]]];

  em[params_, sim_] := Module[{theprobs, mus, vars, probs},
    theprobs = problists[params, sim];
    mus = Map[emu[sim, ♯, theprobs] &, {1, 2}];
    vars = Map[emstd[sim, ♯, params[[1 ;; 2, 1]], theprobs] &, {1, 2}];
    probs = Map[pi[♯, theprobs] &, {1, 2}];
    Append[Transpose[{mus, vars^.5}], probs]];

  )]
]
```

Figure A- Appendix B: Mathematica Code for Em of Gaussian Mixed Model
Source: Mathmatica code developed by Anthony Fox
http://demonstrations.wolfram.com/ExpectationMaximizationForGaussianMixtureDistributions/

APPENDIX C: FUSION METHOD

```
%
%    - simple ones, METHOD is
%         - 'max'  : D = abs(A) >= abs(B) ; C = A(D) + B(~D)
%         - 'min'  : D = abs(A) <= abs(B) ; C = A(D) + B(~D)
%         - 'mean' : C = (A+B)/2 ; D = ones(size(A))
%         - 'rand' : C = A(D) + B(~D); D is a boolean random matrix
%         - 'img1' : C = A
%         - 'img2' : C = B
%
%    - parameter-dependent ones, METHOD is of the following form
%      METHOD = struct('name',nameMETH,'param',paramMETH) where nameMETH
%      can be:
%         - 'linear'    : C = A*paramMETH + B*(1-paramMETH)
%                           where 0 <= paramMETH <= 1
%         - 'UD_fusion' : Up-Down fusion, with paramMETH >= 0
%                           x = linspace(0,1,size(A,1));
%                           P = x.^paramMETH;
%                           Then each row of C is computed with:
%                           C(i,:) = A(i,:)*(1-P(i)) + B(i,:)*P(i);
%                           So C(1,:)= A(1,:) and C(end,:)= A(end,:)
%         - 'DU_fusion' : Down-Up fusion
%         - 'LR_fusion' : Left-Right fusion (columnwise fusion)
%         - 'RL_fusion' : Right-Left fusion (columnwise fusion)
%         - 'userDEF'   : paramMETH is a string 'userFUNCTION' containing
%                           a function name such that:
%                           C = userFUNCTION(A,B).
%
%    In addition, [C,D] = WFUSMAT(A,B,METHOD) returns the boolean
%    matrix D when defined or an empty matrix otherwise.
```

Figure B- Appendix C: Fusion Method

APPENDIX D: SURVEY FORM FOR THE EMERGENCY ABILITY OF HOSPITALS IN
POST-DISASTER

University of
Central Florida

DISSERTATION TITLE: A HUMAN CENTRIC APPROACH TO DATA FUSION IN POST-DISASTERMANAGEMENT: THE DEVELOPMENT OF A FUZZY SET THEORY BASED MODEL SURVEY-OFFICIALS

----FOR HOSPITAL EMERGENCY RESPONSE ABILITY IN POST DISASTER----

BACKGROUND
NAME:
ORGANIZATION:
EMAIL ADDRESS:
By Likert Scale: 1-very poor; 3-poor; 5-fair; 7-good; 9-very good

Group 1: Decision-making system responsible for command and control

	very poor	poor	fair	good	very good
1.1 In the post-disaster, how to evaluate the hospital's incident command team or other special ad hoc body's ability to respond?	☐	☐	☐	☐	☐
1.2 How is the speed of hospital command center bringing together or coordinating all the activities of the hospital's emergency	☐	☐	☐	☐	☐
1.3 The basic qualifications of disaster managers and employees in all hospitals.	☐	☐	☐	☐	☐
1.4 How to evaluate whether members of emergency response teams are sufficient trained and how are the command system structure and function?	☐	☐	☐	☐	☐
1.5 The core internal and external consultation paper is coincided with the World Health Organization (WHO) standards.	☐	☐	☐	☐	☐
1.6 Evaluation on focal point continuity assurance capabilities.	☐	☐	☐	☐	☐
1.7 How is the strict compliance with the basic principles and accepted strategy	☐	☐	☐	☐	☐
1.8 Overall evaluations on ensuring proper management and coordination of activities related to the response.	☐	☐	☐	☐	☐

Group 2: Clear, accurate and timely communication of hospital in post-disaster

	very poor	poor	fair	good	very good
2.1 The ability of public information coordinator hospital spokesman communicating with the public.	☐	☐	☐	☐	☐
2.2 The capabilities of split / waiting areas and emergency response to command center.	☐	☐	☐	☐	☐
2.3 How to evaluate hospital administration, department / unit heads and staff of the facility information exchange mechanism between?	☐	☐	☐	☐	☐
2.4 All patients' priority decisions communicated to all relevant staff and stakeholders with timely.	☐	☐	☐	☐	☐
2.5 How about the appropriate and timely collection, processing and reporting of information on regulatory stakeholders?	☐	☐	☐	☐	☐
2.6 The speed of draft brief key massages for target audiences being prepared for the most likely disaster scenarios.	☐	☐	☐	☐	☐
2.7 How is the reliable and sustainable primary and backup communications system availability?	☐	☐	☐	☐	☐
2.8 Evaluation on brief hospital staff roles and responsibilities	☐	☐	☐	☐	☐

Group 3: Well-developed safety and security procedures

	very poor	poor	fair	good	very good
3.1 The speed to identify areas where increased vulnerability is anticipated in post disaster.	☐	☐	☐	☐	☐
3.2 The speed to establish a reliable model of hospital staff, patients and visitors.	☐	☐	☐	☐	☐
3.3 The effective mechanism of escort emergency medical personnel and their families and patient care.	☐	☐	☐	☐	☐
3.4 Clearly defined crowd control and effective	☐	☐	☐	☐	☐
3.5 Harmful substances for the prevention and control of infection.	☐	☐	☐	☐	☐
3.6 Security of confidential information collection, storage and reporting are clearly defined.	☐	☐	☐	☐	☐
3.7 Create an area of radioactive, biological and chemical purification and isolation.	☐	☐	☐	☐	☐
3.8 Overall evaluation on pairs of hospital security team responsible for all hospital safety and security activities.	☐	☐	☐	☐	☐

Group 4: The abilities for maintaining patient triage operations of hospital

	very poor	poor	fair	good	very good
4.1 The speed of specify experienced personnel to supervise all shunt operation.	☐	☐	☐	☐	☐
4.2 How is the potential environmental harm for receiving patients and waiting for the security field?	☐	☐	☐	☐	☐
4.3 How about the triage area's providing the necessary personnel, medical supplies and critical care services?	☐	☐	☐	☐	☐
4.4 How is the clear of the entrance and exit routes to/from the triage area?	☐	☐	☐	☐	☐
4.5 Evaluation the waiting area for wounded patients able to walk.	☐	☐	☐	☐	☐
4.6 How is the mass-casualty triage protocol based on severity of illness/injury, survivability and hospital capacity that follows internationally accepted principles and guidelines?	☐	☐	☐	☐	☐
4.7 How to identify triage methods, including adequate supply triage tag?	☐	☐	☐	☐	☐
4.8 Assessment of hospital admission, discharge, referral and surgical suites protocols.	☐	☐	☐	☐	☐

Group 5: Surge capacity in post-disaster

	very poor	poor	fair	good	very good
5.1 How is the maximal capacity required for patient admission and care based on availability of human and essential resources and the adaptability of facility space for critical care?	☐	☐	☐	☐	☐
5.2 How to evaluate the available planning assumptions and tools based on the demand for hospital services?	☐	☐	☐	☐	☐
5.3 Assessment on expanding the capacity of hospitalization	☐	☐	☐	☐	☐

Figure C- Appendix D: Survey Form for the Emergency Ability of Hospitals in Post-Disaster

5.4 Evaluation on regulatory policy for the patient in case of emergency overflow. ☐ ☐ ☐ ☐ ☐

5.5 How is the ability for potential gaps in the provision of medical services, emphasizing the key and surgical care in a hospital emergency room? ☐ ☐ ☐ ☐ ☐

5.6 Evaluation prioritizes or cancels nonessential services for your hospital in post disaster. ☐ ☐ ☐ ☐ ☐

5.7 Assessment of adaptation hospital admission and discharge criteria, and according to the available processing power and needs to give priority to clinical intervention. ☐ ☐ ☐ ☐ ☐

5.8 Evaluation your hospital's contingency plan for post mortem care with the appropriate partners. ☐ ☐ ☐ ☐ ☐

Group 6: The ability of the continuity of essential services

	very poor	poor	fair	good	very good
6.1 Evaluation on the continuous provision of essential medical services throughout the community.	☐	☐	☐	☐	☐
6.2 Evaluation the arrangements for appropriate assessment basic lines including water, electricity, oxygen availability.	☐	☐	☐	☐	☐
6.3 How about hazardous waste and other hospital waste collection and disposal emergency mechanism?	☐	☐	☐	☐	☐
6.4 How is the adequate supply availability in most likely disaster events on hospital food and water supplies?	☐	☐	☐	☐	☐
6.5 How is the systematic and deploy an emergency evacuation plan that designed to protect the continuity of existence of intensive care?	☐	☐	☐	☐	☐
6.6 How is the resources needed to ensure the continuity of essential hospital services?	☐	☐	☐	☐	☐
6.7 Overall evaluation of your hospital's essential hospital services maintenance.	☐	☐	☐	☐	☐
6.8 Overall evaluation on the ranking mechanism for all hospital services.	☐	☐	☐	☐	☐

Group 7: Human resources system

	very poor	poor	fair	good	very good
7.1 How is the human resources practitioners' professional skill including knowledge inspection qualified and training coverage?	☐	☐	☐	☐	☐
7.2 How do you assess hospital estimates and continue to monitor staff absenteeism and employee sick leave policy clear	☐	☐	☐	☐	☐
7.3 How is the hospital providing a contingency plan for the provision of food, water and living space for hospital personnel?	☐	☐	☐	☐	☐
7.4 How to provide hospitals with training and exercises may increase the clinical needs of the field?	☐	☐	☐	☐	☐
7.5 Evaluation on pairs of hospital evaluation to ensure adequate translation rotation and self-care support clinical staff morale and reduce medical errors.	☐	☐	☐	☐	☐
7.6 The availability of multidisciplinary psychosocial support teams.	☐	☐	☐	☐	☐
7.7 Assessment personnel to deal with the epidemic-prone diseases of the respiratory system, the provided national law with the appropriate vaccinations.	☐	☐	☐	☐	☐
7.8 Overall evaluation of hospital domestic support measures, including travel, child care, care for a sick or disabled family member, the shift to mobilize employees' flexibility and longer working hours.	☐	☐	☐	☐	☐

Group 8: Logistics and supply management for the hospital in post disaster

	very poor	poor	fair	good	very good
8.1 Assess how emergency pre-hospital transport strategies and networks and transport services to ensure continued patient transferred to hospital	☐	☐	☐	☐	☐
8.2 How are in the home or in the alternative treatment sites providing medicines for patients receiving treatment How effective?	☐	☐	☐	☐	☐
8.3 How are the essential supplies and pharmaceuticals in accordance with national guidelines avoiding loss due to expiration?	☐	☐	☐	☐	☐
8.4 How is your hospital's uninterrupted cold chain for essential items requiring refrigeration	☐	☐	☐	☐	☐
8.5 Evaluate the quality emergency projects.	☐	☐	☐	☐	☐
8.6 Evaluation on the availability for rating agencies and central reserves, and with local suppliers and national and international aid agencies, the availability of emergency protocol.	☐	☐	☐	☐	☐
8.7 How is the consumption of essential supplies and pharmaceuticals using the most likely disaster scenarios?	☐	☐	☐	☐	☐
8.8 Evaluation on the policy of developing and maintaining update all equipment, supplies and medicines inventory shortages early warning mechanism to establish in your hospital's policy evaluation.	☐	☐	☐	☐	☐

Group 9: The ability for post-disaster recovery

	very poor	poor	fair	good	very good
9.1 The speed for arranging a disaster recovery officer responsible of response operations.	☐	☐	☐	☐	☐
9.2 How is the guaranty of a comprehensive structural integrity and safety assessment?	☐	☐	☐	☐	☐
9.3 The speed that a team of hospital staff to carry out a post-action hospital inventory assessment.	☐	☐	☐	☐	☐
9.4 The service that professionally conducted debriefing for staff within 24-72 hours after the occurrence of the emergency.	☐	☐	☐	☐	☐
9.5 The speed of establishing employee assistance program in disaster recovery.	☐	☐	☐	☐	☐
9.6 How about all employees appropriate recognition in the disaster phase.	☐	☐	☐	☐	☐
9.7 The time that resources needed to complete repairs and replacements in disaster recovery phase.	☐	☐	☐	☐	☐
9.8 Evaluation on the timely that an action after an evaluation report will be sent to the hospital in time management, emergency managers and man appropriate stakeholders.	☐	☐	☐	☐	☐

CONFIDENCE: ☐ ☐ ☐ ☐ ☐

Private Policy Declare: Do you want to public this questionnaire including your name, affiliation and contact email?
☐ Yes, I agree. ☐ No, I disagree.

Figure D- Appendix D: Survey Form for the Emergency Ability of Hospitals in Post-Disaster

LIST OF REFERENCES

Agrawala, A. K. (1970). Learning with a probabilistic teacher. *IEEE Transactions on Information Theory*, 16, 373-379.

Ahmed, Z. (2013). Disaster risks and disaster management policies and practices in Pakistan: A critical analysis of Disaster Management Act 2010 of Pakistan. *International Journal of Disaster Risk Reduction,* in press.

Alexey, V., & Vitaliy, K. (1999). Disaster monitoring using remotely sensed data from Russian Earth observation satellites. *Geoscience and Remote Sensing Symposium, 1999. IGARSS '99 Proceedings. IEEE 1999 International, IEEE Press*, 1, 191-193.

Amir, A., Iyengar, H. W. G., Lin, C. Y., et al. (2003). *IBM research TRECVID-2003 system*, NIST Text Retrieval Conf. (TREC).

Pradeep, K. A, Abdulmotaleb, E., & Mohan S. K., (2010). Effective multimedia surveillance using a human-centric approach, *Multimed Tools Appl, Springer*.

Bezdek, J. C., (1981). *Pattern Recognition with Fuzzy Objective Function Algorithms*. Norwell, MA: Kluwer Academic Publishers

Biswas, P. (1983). A Supervised Learning Algorithm for Hierarchical Classification of Fuzzy Patterns. *Information Sciences*, 31, 91-106.

Brandon, P. S. (2011). Extreme Management in Disaster Recovery, The Twelfth East Asia-Pacific Conference on Structural Engineering and Construction. *Procedia Engineering*, 14, 84-94.

Buse, R. P. L. (2010). *A Human-Centric Approach to Program Understanding* (Phd Thesis). Virginia University, Virgina, USA.

Chakraborty, S. (2011). Bayesian semi-supervised learning with support vector machine, *Statistical Methodology*, 8, 68-82.

Chang, C., Pao, H., & Lee, Y. (2012). An RSVM based two-teachers–one-student semi-supervised learning algorithm. *Neural Networks*, 25, 57-69.

Chapelle, O., Schölkopf, B., & Zien, A. (2006). *Semi-Supervised Learning*. Cambridge, MA: The MIT Press, London, England.

Chen, K., Yang, L., Yu, X., et al. (1997) A self-generating modular neural network architecture for supervised learning. *Neurocomputing*, 16, 33-48.

Chen, J., Zhao, S., & Wang, H. (2011). Risk Analysis of Flood Disaster Based on Fuzzy Clustering Method, IACEED2010. *Energy Procedia*, 5, 1915-1919.

Curtis, A., & Mills, J. W. (2012). Spatial video data collection in a post-disaster landscape: The Tuscaloosa Tornado of April 27th 2011. *Applied Geography*, 32, 393-400.

Dasarathy, B. V. (1998). Intelligent learning techniques for multi-source information fusion environments, Decision and Control, 1998. Proceedings of the 37th IEEE Conference on, *IEEE Press*, 221-226.

Denoeux, T. (2011). Maximum likelihood estimation from fuzzy data using the EM algorithm. *Fuzzy Sets and Systems*, 183, 72-91.

Deselaers, T. (2010). GeorgHeigold, HermannNey, Object classification by fusing SVMs and Gaussian mixtures. *Pattern Recognition*, 43, 2476-2484.

Du, R., Wu, Q., He X., et al. (2013). MIL-SKDE: Multiple-instance learning with supervised kernel density estimation. *Signal Processing*, 93, 1471-1484.

Duan, G., Zhao, W., Hu, Z., et al. (2011). *An improved model of regional flood disaster risk assessment based on remote sensing data*. Paper presented at Geoinformatics, 2011 19th International Conference (pp. 1-6). IEEE Press.

Endo, Y., Hamasuna, Y., Yamashiro, M., et al. (2009). *On semisupervised fuzzy c-means clustering.* Paper presented at 2009 IEEE International Conference on Fuzzy Systems (pp. 1119-1124). USA: IEEE Press.

Esogbue, A. O. (1996). Fuzzy sets modeling and optimization for disaster control systems planning. *Fuzzy Sets and Systems*, 81, 169-183.

Facchinetti, G., & Pacchiarotti, N. (2006). Evaluations of fuzzy quantities. *Fuzzy Sets and Systems*, 157, 892- 903.

Fajardo J. T. B., Yasumoto K., Shibata N. et al. (2012). *DTN-based data aggregation for timely information collection in disaster areas, Wireless and Mobile Computing, Networking and Communications (WiMob)* (pp. 333-340). Paper presented at 2012 IEEE 8th International Conference on IEEE Press.

Fisher, J. W. III, Darrell, T., Freeman, W., et al. (2000). Learning joint statistical models for audio-visual fusion and segregation. Advances in Neural Information Processing Systems 13. MIT Press, Cambridge MA.

Fralick, S. C. (1967). Learning to recognize patterns without a teacher. *IEEE Transactions on Information Theory*, 13:57–64.

Ghadiri, N., Baraani-Dastjerdi, A., Ghasem-Aghaee, N., et al. (2011). A Human- Centric Approach to Group-based Context- Awareness. *International Journal of Network Security & Its Application*, 3, 1, 47-66.

Gigli, G., Bosse´, E., Lampropoulos, G. A. (2007). An optimized architecture for classification combining data fusion and data-mining. *Information Fusion*, 8, 366-378.

Gilroy S.W., Cavazza M., Niiranen M., et al. (2009). *PAD-based multimodal affective fusion*, Paper presented at Affective Computing and Intelligent Interaction and Workshops, 2009. ACII 2009. 3rd International Conference on (pp. 1 – 8). IEEE Press.

Grira, N., Crucianu, M., & Boujemaa N. (2008). Active semi-supervised fuzzy clustering. *Pattern Recognition*, 41, 1834 -1844.

Hamida, S., Kibria, B. M. G., Gulati, S., Powell, M., Annaned, B., et al., (2010). Predicting losses of residential structures in the state of Florida by the public hurricane loss evaluation model. *Statistical Methodology*, 7, 552-573.

Han, J. W., & Kamber M. (2000). *DATA MINING: Concepts and Techniques*. Morgan Kaufmann Publishers.

Hauptmann, A. G., & Christel, M. G. (2004). *Successful approaches in the trec video retrieval evaluations.* In Proc. ACM Multimedia.

He, Y., Zhou, J., Kou, P., et al. (2011). A fuzzy clustering iterative model using chaotic differential evolution algorithm for evaluating flood disaster. *Expert Systems with Applications*, 38, 10060-10065.

Holt-Giménez, E. (2002). Measuring farmers' agroecological resistance after Hurricane Mitch in Nicaragua: a case study in participatory, sustainable land management impact monitoring. *Agriculture, Ecosystems and Environment*, 93, 87-105.

Hristidis, V., Chen, S. C, Li, T., Luis, S., & Deng, Y (2012). Survey of data management and analysis in disaster situations. *The Journal of Systems and Software*, 83, 1701-1714.

Hershey, J., Movellan, J. (2001). Using audio-visual synchrony to locate sounds. *Advances in Neural Information Processing Systems 12*, MIT Press, Cambridge MA.

Hidemi, F., Kazue, K., Kenji, S., et al. (2012). Development and Evaluation of Disaster Information Management System Using Digital Pens and Tabletop User Interfaces. *Journal of Japan Association for Earthquake Engineering, 2012,* 3, 3001-30020.

Hotelling, H., (1935). *The most predictable criterion, J. Ed.Phych*. (26.) 139-142.

Hristidis, V., Chen S., Li, T., et al. (2010). Survey of data management and analysis in disaster situations. *The Journal of Systems and Software*, 83, 1701-1714.

Hussain, M., Arsalan, M. H., Siddigi, K., et al. (2005). *Emerging Geo-Information Technologies (GIT) for Natural Disaster Management in Pakistan: An Overview, RAST 2005*. Paper presented at Proceedings of the 2nd International Conference on Recent Advances in Space Technologies.

Iliadis, L. S. (2005). Fundamental Fuzzy Relation Concepts of a D.S.S. for the Estimation of Natural Disasters' Risk (The Case of a Trapezoidal Membership Function), *Mathematical and Computer Modelling*, 42, 747-758.

Ishimaru, Y., Sun, W., Yasumoto, K., & Ito, M. (2010). *Dtn-based delivery of word-of-mouth information with priority and deadline.* Paper presented at 5th International Conference on Mobile Computing and Ubiquitous Networking, IEEE Press, USA.

Iwai, S., Kameda, H. (1998). *Post-earthquake reconnaissance and disaster information management methodologies*. Paper presented at Eleventh European Conference on Earthquake Engineering.

Jaynes, E. T. (2003). *Probability Theory: The Logic of Science.* Cambridge University Press.

Jianjun, Y., Jianqiu, Z., & Zesen, Z. (2008). Gaussian Sum PHD Filtering Algorithm for Nonlinear Non-Gaussian Models. *Chinese Journal of Aeronautics*, 21, 341-351.

Jotshi, A., Gong, Q., Batta, R. (2009). Dispatching and routing of emergency vehicles in disaster mitigation using data fusion. *Socio-Economic Planning Sciences*, 43, 1-24.

Kim, K. N., & Choi, J. (2013). Breaking the vicious cycle of flood disasters: Goals of project management in post-disaster rebuild projects. *International Journal of Project Management*, 31, 147-160.

Kennedy, H. L. (2012). Fusion of possibly biased location estimates using Gaussian mixture models. *Information Fusion*, 13, 214-222.

Khaleghi, B., Khamis, A., Karray, F. O., et al. (2013). Multisensor data fusion: A review of the state-of-the-art. *Information Fusion*, 14, 28-44.

Kim, J. H., Hong, K. S. (2008). Distributed-computing-based multimodal fusion interface using voiceXML and KSSL for wearable PC. *Electronics Letters*, 44, 1, 58 - 60.

Klose, A., & Kruse, R. (2005). Semi-supervised learning in knowledge discovery. *Fuzzy Sets and Systems*, 149, 209-233.

Konar, A., Chakraborty, U. K., & Wang, P. P. (2005). Supervised learning on a fuzzy Petri net. *Information Sciences*, 172, 397-416.

Kongn, J., Zhang, W. Y., Yu, N., et al. (2011). Design of human-centric adaptive multimodal interfaces. *Int. J. Human-Computer Studies*, 69,854-869.

Krahnstoever, N., Schapira, E., Kettebekov, S., et al. (2002). *Multimodal human-computer interaction for crisis management systems*. Applications of Computer Vision*, 2002*. (WACV 2002). Proceedings. Sixth IEEE Workshop (pp. 203 – 207). IEEE Press.

Kumari, R. S. S., Nidhyananthan, S. S., & Anand, G (2012). Fused Mel Feature sets based Text-Independent Speaker Identification using Gaussian Mixture Model, *International Conference on Communication Technology and System Design 2011*, Procedia Engineering, 30, 319-326.

Lai, D. T. C., & Garibaldi, J. M. (2011). *A Comparison of Distance-based Semi-Supervised Fuzzy c-Means Clustering Algorithms*. Paper presented at 2011 IEEE International Conference on Fuzzy Systems (FUZZ) (pp.1580 – 1586). USA: IEEE Press.

Lee Y. J., & Mangasarian, O. L. (2001). *RSVM: reduced support vector machines*. Paper presented at the first SIAM international conference on data mining.

Lee, J. K., Bharosa, N., Yang, J., Janssen, M. & Rao, H. R. (2012). Group value and intention to use — A study of multi-agency disaster management information systems for public safety. *Decision Support Systems,* 50, pp: 404–414.

Lee, C. H., & Jeng, T. (2003). A Context Manager for tangible media design presentation A human-centric interaction approach. *Automation in Construction*, 12, 487-493.

Lee, Y. J., & Huang, S. Y. (2007). Reduced support vector machines: a statistical theory. *IEEE Transactions on Neural Networks*, 18, 1–13.

Lee, J. K., Bharosa, N., Yang, J., et al. (2011). Group value and intention to use -A study of multi-agency disaster management information systems for public safety. *Decision Support Systems*, 50, 404-414.

Leidner, D. E., Pan, G., & Pan, S. L. (2009). The role of IT in crisis response: Lessons from the SARS and Asian Tsunami disasters. *Journal of Strategic Information Systems*, 18, 80-99.

Leng, Y., Xu, X., & Qi, G. (2013). Combining active learning and semi-supervised learning to construct SVM classifier. *Knowledge-Based Systems*, 44, 121-131.

Levin, E., Zarnowski, A., Cohen, C. A., et al. (2010). *Human-centric Approach to Inhomogeneous Geospatial Data Fusion and Actualization*. Paper presented at ASPRS Annual Conference, San-Diego.

Li, Q. G., Kang, L., Tang, D. Q., et al. (2011). Applications on Spatial Information Technology in Natural Disasters, 2011 3rd International Conference on Environmental Science and Information Application Technology (ESIAT 2011), *Procedia Environmental Sciences* 10, 1396 -1400.

Li, M., Lin, D., & Kou, J. (2012). A hybrid niching PSO enhanced with recombination-replacement crowding strategy for multimodal function optimization. *Applied Soft Computing*, 12, 975-987.

Liang, C. C., Wang, C. H., Luh, H., et al. (2009). Disaster avoidance mechanism for content-delivering service. *Computers & Operations Research*, 36, 27-39.

Liang, Y., & Leung, K. S. (2011). Genetic Algorithm with adaptive elitist-population strategies for multimodal function optimization. *Applied Soft Computing*, 11, 2017-2034.

Lin, C. T. (1995). A neural fuzzy control system with structure and parameter learning. *Fuzzy Sets and Systems*, 70, 183-212.

Lin, C. T., & Lu, Y. C. (1996). A Neural Fuzzy System with Fuzzy Supervised Learning. *IEEE Transactions on Systems, Man, and Cybernetics-Part B: Cybernetics*, 26, 5, 744-763.

Li, C., Liu, L., & Jiang, W. (2008). Objective function of semi-supervised fuzzy c-means clustering algorithm. *IEEE International Conference on Industrial Informatics*, pp. 737–742.

Liu, Y., Zheng, F., Cai, K., et al. (2009). Cross-Media Retrieval Method Based on Temporal-spatial Clustering and Multimodal Fusion. *Internet Computing for Science and Engineering (ICICSE)*, 2009 Fourth International Conference (pp. 78 – 84). IEEE Press.

Liu, X., Zhang, J., Cai, W., et al. (2010). Information diffusion-based spatio-temporal risk analysis of grassland fire disaster in northern China. *Knowledge-Based Systems*, 23, 53-60.

Loza, A., Bull, D., Canagarajah, N., et al. (2010). Non-Gaussian model-based fusion of noisy images in the wavelet domain. *Computer Vision and Image Understanding*, 114, 1, 54-65.

Lu, Y., & Yang, D. (2011). Information exchange in virtual communities under extreme disaster conditions. *Decision Support Systems*, 50, 529-538.

Mansourian, A., Rajabifard, A., Valadan Zoej, M. J., et al. (2006). Using SDI and web-based system to facilitate disaster management, Computers & Geosciences, 32, 303-315.

Mao, Z., Yang Z., Wu, X., et al. (2010). *Robust adaptive coregistration of InSAR images using multisource information and joint data fusion.* Paper presented at Audio Language and Image Processing (ICALIP), 2010 International Conference (pp.1700 – 1704). IEEE Press.

Maraziotis, I. A. (2012). A semi-supervised fuzzy clustering algorithm applied to gene expression data. *Pattern Recognition*, 45, 637-648.

Mavrikios, D., Karabatsou, V., Pappas, M., et al. (2007). An efficient approach to human motion modeling for the verification of human-centric product design and manufacturing in virtual environments. *Robotics and Computer-Integrated Manufacturing*, 23, 533-543.

McLaughlin, K. A., Fairbank, J. A., Gruber, M. J., et al. (2010). Trends in Serious Emotional Disturbance Among Youths Exposed to Hurricane Katrina. *Journal of the American Academy of Child & Adolescent Psychiatry*, 49, 10, 990-1012

Munichi Re Group (2001). *Annual Report on Risk*. Retrieved from http://www.munichre.com/en/homepage/default.aspx

Jain, A. K., Murty, M. N., & Flynn, P. J. (1999). Data clustering: A review. *ACM Comput. Surv.*, vol. 31, pp. 264–323.

Nachouki, G., & Quafafou, M. (2008). Multi-data source fusion. *Information Fusion*, 9, 523-537.

Nojun, K., & Choi, C. H. (1999). *Improved mutual information feature selector for neural networks in supervised learning*. Neural Networks, 1999. Paper presented at IJCNN '99. International Joint Conference (pp.1313 – 1318). IEEE Press.

Nufer, K. E., Wilson-Ramirez, G., & Crandall, C. S. (2003). Different Medical Needs Between Hurricane and Flood Victims. *Wilderness and Environmental Medicine*, 14, 89-93

Onwubiko, C., & Owens, T. J. (2002). Situational Awareness in Computer Network Defense: Principles, Methods and Applications. *Information Science Reference (an imprint of IGI Global),* USA.

Pang, Y., Chen J., & Wang, F. (2009). <u>*A Special Supervised Learning Algorithm and Its Applications.*</u> Hybrid Intelligent Systems, 2009. HIS '09. Paper presented at Ninth International Conference (pp. 462 – 466). IEEE Press.

Papandreou, G., Katsamanis, A., Pitsikalis, V., et al. (2009). Adaptive Multimodal Fusion by Uncertainty Compensation With Application to Audiovisual Speech Recognition, *IEEE Transactions on Audio, Speech and Language Processing*, 17, 3, 423-435.

Pearl, J. (1988). *Probabilistic Reasoning in Intelligent Systems*. San Francisco, CA: Morgan Kaufmann Publisher.

Pearson, K. (1901). On lines and planes of closest fit to systems of points in space. *Philosophical Magazine*, 2, 559-572.

Pedrycz, W. (1985). Algorithms of fuzzy clustering with partial supervision. *Pattern Recognition Letters*, 3, 1, 13–20.

Pedrycz, W., & Waletzky, J. (1997). Fuzzy clustering with partial supervision, *IEEE Transactions on Systems, Man and Cybernetics*, 27, 5, 787–795.

Peng, W., Jia, L., Tang, J., et al. (2012). *The traffic information fusion method based on the multisource detectors.* Paper presented at Control and Decision Conference (CCD*C)*, 2012 24th Chinese (pp.3860 – 3864). IEEE Press.

Pidgeon, N., & O'Leary, M. (2000). Man-made disasters: why technology and organizations (sometimes) fail. *Safety Science*, 34, 15-30.

Platz, E., Cooper, H. P., Silvestri, S., & Siebert, C. F. (2007). The Impact Of A Series Of Hurricanes On The Visits To Two Central Florida Emergency Departments. *The Journal of Emergency Medicine*, 33, 1, 39-46.

Poh, N., Kittler, J., & Bourlai, T. (2010). Quality-Based Score Normalization With Device Qualitative Information for Multimodal Biometric Fusion. *IEEE Transactions on System, Man, and Cybernetics-Part A: System and Humans*, 40, 3, 539-554.

Poh, N., Windridge, D., Mottl, V., et al. (2010). Alexander Tatarchuk, and Andrey Eliseyev, Addressing Missing Values in Kernel-Based Multimodal Biometric Fusion Using Neutral

Point Substitution. *IEEE Transactions on Information Forensics and Security*, 5, 3, 461-469

Preece, G., Shaw, D., & Hayashi, H. (2013). Using the Viable System Model (VSM) to structure information processing complexity in disaster response. *European Journal of Operational Research*, 224, 209-218.

Pultz, T. J., & Scofield, R. A. (2002). *Applications of Remotely Sensed Data in Flood Prediction and Monitoring: Report of the CEOS Disaster Management Support Group Flood Team.* Paper presented at Geoscience and Remote Sensing Symposium, 2002. IGARSS '02. 2002 IEEE International. (2, pp.768 – 770) IEEE Press.

Raju, E., & Becker, P. (2013). Multi-organizational coordination for disaster recovery: The story of post-tsunami Tamil Nadu, India. *International Journal of Disaster Risk Reduction*, in press.

Reddy, B. S., & Basir, O. A. (2010). Concept-based evidential reasoning for multimodal fusion in human–computer interaction. *Applied Soft Computing*, 10, 567-577.

Rodger, J. A. (2012). Toward reducing failure risk in an integrated vehicle health maintenance system: A fuzzy multi-sensor data fusion Kalman filter approach for IVHMS. *Expert Systems with Applications*, 39, 9821-9836.

Romaniuk, J., & Gugel, C. (2010). The ARF 360 model: update to a human-centric approach. *Journal of Advertising Research*, 50, 3,.334-343.

Rousseau, C., Bellik, Y., Vernier, F., et al. (2006). A framework for the intelligent multimodal presentation of information. *Signal Processing*, 86, 3696-3713.

Roux, L., & Desachy, J. (1995). *Information fusion for supervised classification in a satellite image. Fuzzy Systems, 1995.* Paper presented at International Joint Conference of the Fourth IEEE International Conference on Fuzzy Systems and The Second International Fuzzy Engineering Symposium., Proceedings of 1995 IEEE Int. (pp.1119 – 1124). IEEE Press.

Ruggiero, K. J., Resnick, H. S., & Paul, L. A. (2012). Randomized controlled trial of an internet-based intervention using random-digit-dial recruitment: the Disaster Recovery Web project, Contemporary Clinical Trials, 33, 237-246.

Saha, S., & Matsumoto, M., (2007). *A Framework for Data Collection and Wireless Sensor Network Protocol for Disaster Management.* Paper presented at Communication Systems Software and Middleware, 2007. COMSWARE 2007. 2nd International Conference (pp. 1-6). IEEE Press.

Saporta, G. (2002). Data fusion and data grafting. *Computational Statistics & Data Analysis*, 38, 465-473.

Scudder, H. J. (1965). Probability of error of some adaptive pattern-recognition machines. *IEEE Transactions on Information Theory*, 11:363-371.

Scudder, H. J. (1995). Probability of error of some adaptive pattern-recognition machines. *IEEE Transactions on Information Theory*, 11, 363-371.

Shi, F., & Xu, J. (2012). Employing Emotional Cellular Models and its Similarity Computing in Valence - Arousal based Semantic Inference. *Information: An International Interdisciplinary Journal*, 15, 2, 467-482.

Shkvarko, Y., Montiel, J. L. L., & Garibay, R. B. (2006). *End-User-Oriented Multisource Information Fusion for Resource Management with Remote Sensing Imagery*. Paper presented at Geoscience and Remote Sensing Symposium, 2006. IGARSS 2006. IEEE International Conference (pp.2163-2166). IEEE Press.

Shtykh, R. Y., & Jin, Q. (2011). A human-centric integrated approach to web information search and sharing. *Human-centric Computing and Information Sciences*, 1, 2, 1-37.

Singh, D., Chamundeeswari, V.V., Singh, K., et al. (2008). *Monitoring and Change Detection of Natural Disaster (Like Subsidence) Using Synthetic Aperture Radar (SAR) Data.* Recent Advances in Microwave Theory and Applications, 2008 (pp.419-421). MICROWAVE 2008. International Conference on IEEE Press.

Stiller, C., León, F. P., & Kruse, M. (2011). Information fusion for automotive applications - An overview. *Information Fusion*, 12,244-252.

Straccia, U., & Madrid, N. (2012). A top-k query answering procedure for fuzzy logic programming. *Fuzzy Sets and Systems*, 205, 1-29.

Sun, Y., Huang, R., Chen, D., et al., (2008). Fuzzy Set-Based Risk Evaluation Model for Real Estate Projects. *Tsinghua Science and Technology*, 13, SI, 158-164.

Sun, X. J., Gao, Y., Deng, Z. L., et al., (2010). Multi-model information fusion Kalman filtering and white noise deconvolution. *Information Fusion*, 11, 163-173.

Tian, X., Gasso, G., & Canu, S. (2012). A multiple kernel frame work for inductive semi-supervised SVM learning. *Neurocomputing*, 90, 46-58.

Tsai, C. H., & Chen, C. W. (2010). An earthquake disaster management mechanism based on risk assessment information for the tourism industry-a case study from the island of Taiwan. *Tourism Management*, 31, 470-481.

Tsai, C. H., & Chen, C. W. (2011). The establishment of a rapid natural disaster risk assessment model for the tourism industry. *Tourism Management*, 32, 158-171.

Urakami, M., Fujii, K., Kota, M., et al. (2009). *A Design of Reduced Data Representation for Information Sharing in the Disaster-Affected Area.* Paper presented at Advanced Information Networking and Applications Workshops, 2009. WAINA '09. International Conference (pp.856-861). IEEE Press.

Vasudevan, S. (2012). Data fusion with Gaussian processes. *Robotics and Autonomous Systems*, 60, 1528-1544.

Vega, J. (2008). JET-EFDA Contributors, Intelligent methods for data retrieval in fusion databases. *Fusion Engineering and Design*, 83, 382-386.

Vybornova, O., Mendona, H., Lawson, J. Y. L., et al. (2008). *High Level Data Fusion on a Multimodal Interactive Applications Platform, Multimedia, 2008. ISM 2008*. Paper presented at Tenth IEEE International Symposium (pp.493-494). IEEE Press.

Waegeman, W., Verwaeren, J., Slabbinck, B., et al. (2011). Supervised learning algorithms for multi-class classification problems with partial class memberships. *Fuzzy Sets and Systems*, 184, 106-125.

Wagh, A., Li, X., Wany, J., et al., (2011). Human-centric Data Fusion in Vehicular Cyber-Physical Systems. *The First International Workshop on Cyber-Physical Networking Systems*, IEEE Press, USA.

Wang, J., Jebara, T., & Chang, S. F. (2008). Graph transduction via alternating minimization. Paper presented at Twenty- Fifth International Conference on Machine Learning (307, pp. 1144-1151) Helsinki, Finland, ACM, 2008.

Wang Z., & Bovik, A. C. (2002). A universal image quality index. *IEEE Signal Processing Letters*, 9, 3, 81–84

Weckenmann, A., Jiang, X., Sommer, K. D., et al. (2009). Multisensor data fusion in dimensional metrology. *CIRP Annals - Manufacturing Technology*, 58, 701-721.

Weng, P. H., Liu, F. T., Chen, Y. J., et al. (2008). *A Hybrid Supervised Neural Network Learning Algorithm. Innovative Computing Information and Control, 2008*. Paper presented at ICICIC '08. 3rd International Conference (pp.294-297). IEEE Press.

Widener, M. J., & Horner, M. W. (2011). A hierarchical approach to modeling hurricane disaster relief goods distribution. *Journal of Transport Geography,* 19, 821-828.

Wigder, H. N., Fligner, D. J., Rivers, D., et al. (1989). Microcomputer-assisted transmission of disaster data by cellular telephone Original Research Article. *The Journal of Emergency Medicine*, 7, 6, 651-655.

Witteveen, A. B., Bisson, J. I., Ajdukovic, D., et al. (2012). Post-disaster psychosocial services across Europe: The TENTS project. *Social Science & Medicine*, 75, 1708-1714.

Wu, Y., Chang, E. Y., Chang, K. C. C., et al. (2004). *Optimal multimodal fusion for multimedia data analysis.* In Proc. ACM Multimedia.

Wu, H. C., Lindell, M. K., & Prater, C. S. (2012). Logistics of hurricane evacuation in Hurricanes Katrina and Rita. *Transportation Research Part F* 15,445–461

Xie, L., Natsev, A., & Tesic, J. (2007). Dynamic Multimodal Fusion in Video Search, *Multimedia and Expo, 2007*, IEEE International Conference on IEEE Press, pp.1499 - 1502.

Xue, Z., Shang, Y., & Feng, A. (2010), Semi-supervised outlier detection based on fuzzy rough C-means clustering, Mathematics and Computers in Simulation, 80, 1911-1921.

Yamamoto, I., Inagawa, N., Tsuji, T., et al. (2011). *Research and development of the multifunctional observation and data transfer system for disaster prevention*. Image and Signal Processing (CISP). Paper presented at 2011 4th International Congress (pp.1661-1665). IEEE Press.

Yan, Y., Chen, L., & Tjhi, W. C. (2013). Fuzzy semi-supervised co-clustering for text documents. *Fuzzy Sets and Systems*, 215, 74–89.

Ye, H., & Ni, W. (1999). Nonlinear system identification using a Bayesian-Gaussian neural network for predictive control. *Neurocomputing*, 28, 21-36.

Ye, Y., Wang, B., Shao, M., et al. (2009). *Design of Mineral Feeding Supervisory Control System Based on Multi-source Information Fusion*. Intelligent Human-Machine Systems and Cybernetics, 2009. Paper presented at IHMSC '09. International Conference (pp.82 – 85). IEEE Press.

Yin, X., Shu, T., & Huang, Q. (2012). Semi-supervised fuzzy clustering with metric learning and entropy regularization. *Knowledge-Based Systems*, 35, 304-311.

Zadeh, L. A. (1975a). The Concept of a Linguistic Variable and its Application to Approximate Reasoning-I. *Information Sciences*, 8: 299-249.

Zadeh, L. A. (1975b). The Concept of a Linguistic Variable and its Application to Approximate Reasoning-II. *Information Sciences*, 8: 301-357

Zadeh, L. A. (1975c). The Concept of a Linguistic Variable and its Application to Approximate Reasoning-III. *Information Sciences*, 9: 43-80

Zadeh, L. A. (2005). Toward a generalized theory of uncertainty (GTU)-an outline. *Information Sciences*, 172: 1-40.

Zerger, A., & Smith, D. I. (2003). Impediments to using GIS for real-time disaster decision support. *Computers, Environment and Urban Systems*, 27, 123-141.

Zhang, D., Tan, K., & Chen, S., (2004). Semi-supervised kernel-based fuzzy c-means. *Lecture Notes in Computer Science: Neural Information Processing*, 3316, 1229–1234.

Zhang, C., Zhao, T., & Li, W. (2010). Automatic search of geospatial features for disaster and emergency management. *International Journal of Applied Earth Observation and Geoinformation*, 12, 409-418.

Zhang, Y., Yu, X., Chen, X., et al. (2010). *Unmanned hybrid electric vehicles FNN control based on self-organized learning algorithm and supervised learning algorithm.* Computer, Mechatronics, Control and Electronic Engineering (CMCE). Presented at 2010 International Conference (pp. 555 – 558). IEEE Press.

Zhang, T., Liu, S., Xu, C., et al. (2011). Boosted multi-classsemi-supervised learning for human action recognition. *Pattern Recognition*, 44, 2334-2342.

Zhou, Z. H. (2007). *Co-training Paradigm in Semi-supervised Learning*. Presented at Chinese Workshop on Machine Learning and Applications, Nanjing, China.

Zhou, Y., Liu, G. J., Fu, E. J., et al. (2009). An object-relational prototype of GIS-based disaster database. The 6th International Conference on Mining Science & Technology. *Procedia Earth and Planetary Science*, 1, 1060-1066.

Zhu, Y., Shtykh, R. Y., & Jin, Q. (2012). A human-centric framework for context-aware flowable services in cloud computing environments. *Information Sciences*, doi:10.1016/j.ins.2012.01.030.

Zhu, X., Ghahramani, Z., & Lafferty, J. (2003). *Semi-supervised learning using Gaussian fields and harmonic functions.* Presented at Twentieth International Conference on Machine Learning, Washington DC, USA, 2003 (pp. 912-919): AAAI Press.

Zou, Q., Zhou, J., Zhou, C., et al.. (2012). Fuzzy risk analysis of flood disasters based on diffused-interior-outer-set model, *Expert Systems with Applications*, 39, 6213-6220.