Towson University

Office of Graduate Studies

SITUATION AWARENESS IN MILITARY AND ASSISTIVE HEALTHCARE

by

Sheng Miao

A Dissertation

Presented to the Faculty of the Graduate School

of Towson University

in Partial Fulfillment of the Requirements for the Degree of

DOCTOR OF SCIENCE

Department of Computer and Information Sciences

TOWSON UNIVERSITY

Towson, Maryland, 21252

July 2017

DISSERTATION APPROVAL PAGE

This is to certify that the dissertation prepared by **Sheng Miao**, entitled "**Situation Awareness in Military and Assistive Healthcare**", has been approved by this committee as satisfactorily completing the dissertation requirements for the degree of **Doctor of Science** in Information Technology.

Dr. Ziying Tang Chair, Dissertation Committee	8/16/2017 Date
Dr. Robert J. Hammell II Member, Dissertation Committee	8/16/2017 Date
Dr. Jinjuan Feng Member, Dissertation Committee	8/16/2017 Date
Dr. Amanda Jozkowski Member, Dissertation Committee	8/16/2017 Date
Dr. Siddharth Kaza Chair, Department of COSC	8/18/2017 Date
Dr. Chao Lu Doctoral Program Director	8/16/2017 Date
Janet V DeLany Dr. Janet DeLany Dean of Graduate Studies	8-21-17 Date

Dedicated

to

My Parents

Table of Contents

Abstract

Situation Awareness in Military and Assistive Healthcare

Sheng Miao

Today Situation Awareness (SA) has become an attractive research topic which is widely used in various and complex systems. The concept of SA is rooted in the history of military theory and has been extended to commercial domains and utilized into the human factors domain in the last few decades. My dissertation focuses on studying how SA is used in the military context and in the assistive healthcare domain. We will study how to use SA to help commanders make decisions in the battlefield and support assistive healthcare systems to help users with various disabilities.

In the battlefield, SA techniques can help commanders make optimized decisions based on the intelligence from various sources. Some previous research has developed a fuzzy-based value of single piece of information determination system which delivers the most valuable information to commanders. Based on the existing system, we develop an adjustable VoI system which allows subject matter experts to tune special SA models for specific military contexts. Another research topic in military context is exploring the integration of complementary and contradictory information, since it is common to have multiple information describing a certain

event. We also introduce the utilization of human processing for military SA based on social media data.

In my dissertation, two special disability user groups are selected as targeted users, who are young adults with Autism Spectrum Disorder (ASD) and children with Spinal Muscular Atrophy (SMA). We employ SA techniques to implement assistive healthcare systems to help the targeted users to perform better daily activities, communicating with others, controlling their living environment, and even monitoring their needs and providing real-time help. The objective of this study is to improve their quality of life.

Acknowledgement

My deepest gratitude goes to my academic advisor, Dr. Ziying Tang. I would like to express my sincere thanks for her guidance, patience, encouragement and inspiration which support me during my doctoral study.

I would like to thank Dr. Robert J. Hammell II and Dr. Jinjuan Feng for their constant guidance over the past several years. Special thanks go to Dr. Amanda Jozkowski who provides professional guidance in my dissertation. Working with them has not only helped me mature as a scholar, but has also infused in me (for life) the value of diligence, attention to detail and perfection at every work.

I own my gratitude to Dr. Chao Lu, he introduced me to Towson University therefore I can have this great chance to do my doctoral program. I would like to thank all the faculty and staff in the Department of Computer and Information Sciences at Towson University for providing an enjoyable, supportive and friendly environment. My sincere thanks also go to my friends who bring me happiness, peace, and occasions to relax.

My deepest love goes to my parents, Dr. Qun Miao and Dr. Qing Yu, and all my family members. They support me both financially and emotionally throughout my doctoral study.

July, 2017

List of Tables

Page

List of Figures

Page

Chapter 1

Introduction

1.1 Background and Motivation

In the last few decades, complex and dynamic systems have attracted increased attention and focus as they can be used as effective, timely decision makers in substitution for human beings, helping eliminate human error. One of the most critical components in these decision making systems involves situation awareness (SA). The concept of SA is rooted in the history of military theory [3] and can also be traced to World War I, where it was recognized as a crucial component for aircrew in military aircraft to consider. Pilots need to know the many complex and dynamic events that occur simultaneously in flight, and the information is then diagnosed and used for guiding future actions. With the rapid development of information technology, the term SA has been extended from the original aircraft domain to the human factors domain. Some researchers have presented various definitions of SA:

1. SA is an abstraction that exists within our minds, describing phenomena that we observe in humans performing work in a rich and usually dynamic environment [4].

2. SA provides the primary basis for subsequent decision making and performance in the operation of complex, dynamic systems [5].

3. SA is the knowledge, cognition and anticipation of events, factors and variables affecting the safe, expedient and effective conduct of the mission [6].

4. SA is a continuous extraction of environmental information, integration of this information with previous knowledge to form a coherent mental picture, and the use of that picture in directing further perception and anticipating future events [7].

Today SA is generally defined as the perception of environmental elements and events with respect to time or space, the comprehension of their meaning, and the projection of their status after variables have changed, such as time, or a predetermined event.

Besides the original military aircraft function, SA also plays an important role in a variety of contexts. For example, commercial air traffic control systems [8, 9] rely on the SA of the system, which maintains up-to-date assessments of the rapidly changing locations of aircraft and their projected locations relative to each other. This kind of systems ensures safe and efficient landings and takeoffs in civil airports. Moreover, much research on SA in large and complex control systems has been completed, such as on refineries and nuclear power plants [10]. In order to achieve effective management, these systems keep tracking numerous system parameters and patterns that occur among them. Researchers have identified that SA can help to understand the current state of a system and might reveal clues as to the functioning of the system and future process state changes [11]. Another context where SA is effectively utilized includes emergency and disaster prediction. Some applications collect information from human sensors on social networks, such as Twitter, Facebook, and Blogs [12, 13]. SA techniques are employed in these applications to help identify if there is emergency or disaster nearby. With information aggregation and mapping, reports are created and posted online to help communities and people when disaster occurs.

In addition to these SA application domains, this dissertation will focus on studying how SA is used in the military context and in the assistive healthcare domain. We will study how to use SA to help commanders make decisions in the battlefield and support assistive healthcare systems to help users with various disabilities.

1.1.1 Situation Awareness in Military Context

SA helps assemble the perception and the understanding of the big picture in an event with many inputs and dynamically changing variables. This is especially important for the Armed Forces, who requires command and control operations and often face critical situations where SA is integral to optimal command and control. A key part of battlefield SA is the ability to transform the information being provided into knowledge and an attainable understanding of many variables as one big picture; the Army recognizes that this process must involve "a mixture of automation and human cognition" [14]. Furthermore, SA in military context refers to the product of applying analysis and judgment to the unit's SA to determine the relationships of the factors present and draw logical conclusions concerning threats to the force or mission accomplishment, opportunities for mission accomplishment, and gaps in information [15]. Commanders in the battlefield rely on SA to learn details about important persons, places and events within their area of operations so that they can address issues ranging from kinetic fights to legal disputes adjudication to economy recovery. Soldiers on the edge of conflict gather data to support their mission.

In the military context, some critical requirements are identified. The most important one is real-time. Since the battlefield environment is changeable, including enemy movement, weather and other factors, this requires the SA computation to respond agilely. Another critical requirement is accuracy. SA in the military context will support commanders so they can make operational decisions and develop courses of action. Because these decisions and courses of action are used for achieving military and civil-military objectives, failure or deviation could be fatal or cause serious consequences. Furthermore, comprehensiveness is the third critical factor in military SA. It is very common that intelligence acquired from different sources are conflicting or complementary with one another. As a result, SA in the military context should provide a complementary output based on different information.

Military operations are now supported by access to information from a myriad of sources that provide huge amounts of data which can be overwhelming. Indeed, a primary challenge nowadays is not a lack of information, but information overload [16], especially the Internet, which is flooded with intermingled information and more social media data becomes available. Another challenge is much of the data may be non-verbal in nature, such as numerical troop strength reports, GPS coordinates of important potential targets, video feeds from unmanned aerial vehicles, and readings from ground-based acoustic sensors. As a result, the military commander and staff are challenged not only by the established information overload dilemma, but more importantly with separating the important information from the routine [17]. The third challenge is that information collecting from different sources could be complementary or conflicting. In the battlefield, intelligences describing one event could come from multiple dissimilar array of sources; many times not always agreeing. One piece of intelligence maybe support and enhance the other one, which means they are complementary with each other. However, sometimes another piece of intelligence disagree with them totally, thus SA could be difficult in the military context.

1.1.2 Situation Awareness in Assistive Healthcare

Assistive healthcare is another domain which largely relies on SA techniques. With the fast development of information technology, various sensors and electronic health records systems, abundant kind of health information become available and analyzable in the healthcare realm. All these data can be used as input for SA in assistive healthcare system.

There are some assistive technologies and systems have been proposed in the medical and healthcare domain to support medical decisions and improve users' quality of life. For example, BioWatch [18] and BioSense [19] are used for detecting certain bioterrorism agents in the air and gathering syndrome surveillance data from hospitals. The Global Public Health Intelligence Network(GPHIN) mines global news for disease reports [20]. In addition, some SA systems are developed to track bed data and electronic health records are available within hospitals or clinicians' offices. Besides these traditional healthcare, living assistive systems for special disability user groups like Autism Spectrum Disorder (ASD) and Spinal Muscular Atrophy (SMA) can also benefit from SA techniques. The SA-based assistive systems can improve their quality of life largely and help them perform better daily activities, communicating with others, controlling their living environment, and even monitoring their needs and providing real-time help. SA in this context means understanding the status of users and performing automatic and interactive feedback based on information collected from assistive systems. For example, when a user with ASD cannot find eggs when he/she is cooking, this situation will be captured and interpreted by the system which understands that the user needs help and then the system will offer prompt help to remind the user where to find eggs.

Similarly, with the help of SA, the assistive system can recognize if a user with SMA needs to cough and inform his/her caregiver correspondingly.

Unfortunately, getting information to compute the SA in this setting is extremely challenging, as our targeted users here refer to people with special health conditions such as ASD and SMA. To get the situation or demanding data could be difficult because our targeted users probably have uncontrollable or unpredictable behaviors. For example, some users with ASD cannot keep their focuses on the current living task and run out of the experimental environment, which means the ambient sensors cannot track their activities at this time. Moreover, children with SMA, especially for the Type I group, are very lively and active. Therefore, they are likely to be impatient and get frustrated and make data collection challenging.

There are some requirements and challenges for these kinds of SA-based assistive applications. First, real-time feedback is required since users may need immediate response, and the delay in this context may cause serious results. For example, users with ASD can be very sensitive to different things and are easy to get frustrated, which means delayed response from assistive systems can make them impatient and even trigger emotional reactions. Similarly, for users with SMA, it can be extremely dangerous when there is an issue with a life-support system but the notification is not sent to caregivers on time. Secondly, user interface design is another challenge because the primary targeted users of systems can be users with various kinds of disabilities or very young children. Traditional interface design may not apply since users with SMA have very limit muscular functionality so that they cannot handle too many icons in the same page. And users with ASD may be confused with multi-level page navigation. Both graphic and interaction designs are critical factors for these kinds of users. With the well-designed user interface, targeted users will be able to understand and use the applications quickly. The third requirement is low-cost and customizable. To be able to widely support our targeted users, we require the assistive healthcare system to use inexpensive sensors and application platforms to achieve economic goals. And every single user is a special individual, which means that user interface should be customizable and multiple design solutions should be provided.

To understand the situations and needs of users better, multiple ambient sensors can be employed to collect information such as in-door location sensor [21] and distance sensor [22]. Moreover, user profile like preferences, functionality, and operation history are also included for SA. As a result, understanding the situation and making decisions based on this integrated information from various data sources provide significant challenges. Besides, we also have to think about the issue of personal privacy. To understand the situation better, we need to take personal sensitive user profile data, such as a user's behavioral functionalities, into consideration. Therefore, how to secure and protect this information is an aspect that we love to consider.

1.2 Research Method

1.2.1 Situation Awareness Model

According to Ensley's model, situation awareness (SA) has three levels: perception of elements in the environment; comprehension of the current situation; and projection of future status [23]. Considering the special conditions and requirements for user group, we extend this by adding one level and Figure 1.1 illustrates the SA architecture. Moreover, based on the requirements from military and healthcare contexts, Subject Matter Experts (SMEs) and healthcare professionals (HPs) should be included in this architecture as human computation.

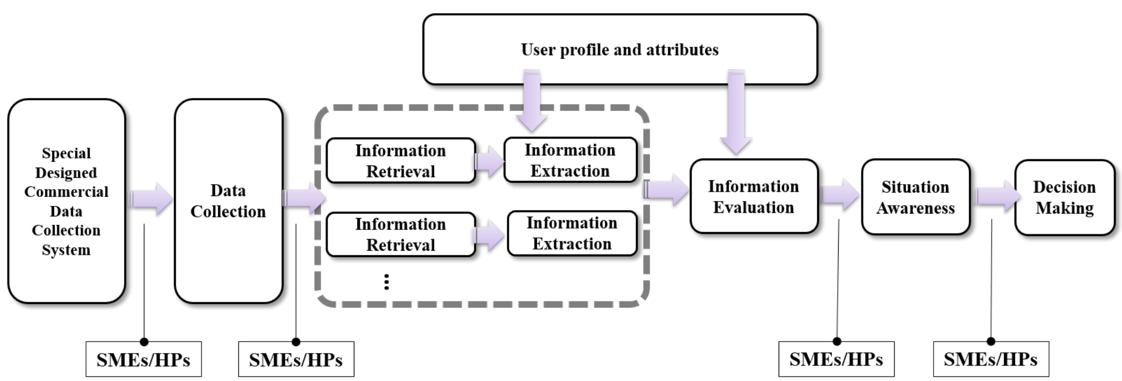

Figure 1.1: Architecture of Proposed Situation Awareness System.

After collecting raw data from sensors and other sources, we then need to start data processing such as information retrieval and extraction. The purpose of information retrieval is to search for the relevant information contained in a set of knowledge, while information extraction focuses on extracting relevant facts. Both of these procedures largely rely on attributes identification, which can impact the result of the whole system. This task cannot be accomplished by simple machine computation and SMEs/HPs involvement is mandatory and essential at this point.

The data processing step, shown as the dotted rectangle in Figure 1.1, is the task that may encompass information regarding user profiles and attributes, depending on the situation which we are trying to understand. It is also worth noticing that in this step, information retrieval and extraction may happen multiple times, as shown in the Figure. This is because when a new event happens, we may need to update the dataset and research focus and then process the media data again.

After data retrieval and information extraction, we then compute the Value of Information (VoI). This is called the information evaluation step in our architecture. User profiles and attributes can also make great contributions in this step, as demonstrated in Figure 1.1. When data are evaluated and values are given to the relevant information, we can then compute SA. In this step, SMEs/HPs involvement is recommended to optimize results. Furthermore, a group instead of an individual-level SA process is extremely helpful in this case.

After knowing the big picture, we then make decisions which project the future actions in the environment. In this step, human experts should be involved too. Decision making implies what are we going to do, based on knowing what is going on. Possible errors always exist and no SA system can guarantee accuracy all the time. Therefore, computer-aided decision making should only be used as a complementary module. SMEs/HPs must be the ones to determine the real actions, especially in the military and healthcare settings where consequences of a decision can be serious.

1.2.2 Research Questions

In this dissertation, we will focus on studying the following research questions:

1. How to acquire user information and situation data for users with special health condition?

2. How to compute SA based on noisy and large data from various data resources?

3. How to extract the most important information from dataset and compute SA?

4. How to integrate complementary and contradictory information when this system computes SA?

5. What is the human computing role in the SA computation?

1.3 Proposed Solution

1.3.1 Situation Awareness Application in Military Context

In today's military environment, vast amounts of disparate information are available. To aid situation awareness (SA), it is vital to have a way to judge the importance of information. A fuzzy-based system, which is used for assigning a Value of Information (VoI) determination for individual pieces of information, has been developed in recent research. In this study, we will present an investigation on the effect of using triangular and trapezoidal fuzzy membership functions with the fuzzy-based system. Moreover, the effect of integrating subsequent complementary and/or contradictory information into the VoI process will be studied in our research, some initial results will be also presented. Furthermore, we will give a detailed explanation on how social media data affect SA in military context, along with a discussion of involvement of Subject Matter Experts (SMEs) as human computation.

1.3.2 Situation Awareness System for Assistive Healthcare

Individuals with Autism Spectrum Disorder (ASD) can face great challenges when learning and maintaining basic living skills. This not only reduces the possibilities of independent living and employment, but also continuously brings social and financial burdens to their caregivers/mentors. Although research has been done to help autistic users, most research focuses on improving social and communication skills or offering help in case of emergency. In this study, we propose a novel portable context-aware assistive system to help the autistic users in their daily activities such as cooking and cleaning. To make it easily accessible and cost effective, we will employ mobile devices and cheap context sensors. Bayesian classification technique will be applied to implement automated prompting function in the system. This function can assist people affected by autism so they can improve their quality of life and lead an independent life.

Spinal Muscular Atrophy (SMA) is a rare genetic disease and it is the No. 1 genetic killer of infants and toddlers. SMA affects an individual's motor neurons in a range of ways depending on the particular type; people with SMA are often unable to walk and talk; many cannot sit up unassisted and are unable to eat and breathe, especially for those with SMA type I, the most severe type. This study will focus on exploring how to provide an accessible and responsive interactive way to communicate and control environment. By employing low-cost customizable, wearable devices, a multi-model system will be developed for type I SMA users to assist their daily activities.

1.4 Summary of Contribution

This dissertation studies the utilization of situation awareness (SA) in both military context and assistive healthcare. SA can provide decision supporting to commanders in the battlefield and improve quality of life for users with special health conditions. The detailed contributions are listed below:

1. Based on previous Value of Information (VoI) determination framework, we have developed two approaches using triangular and trapezoidal fuzzy membership functions. The distribution analysis reveals that both of two approaches have obvious differences thus they have their own strengths and weaknesses.

2. A customizable VoI system is designed for Subject Matter Experts (SMEs), which allows commanders and intelligence analyzers to tune customized fuzzy logic models for special military situations. This provides flexibilities to SMEs and helps to improve the accuracy of SA in the battlefield.

3. We have presented an investigation on determining combination of complementary and/or contradictory information. The initial results show that the proposed approach does not work.

4. A detailed explanation on how social media data affect SA is presented. The focus is set on the military environment and we have discussed it is necessary

and essential to involve SME based human computation in the SA architecture.

5. An ambient intelligence based context aware assistive system for young adults with ASD is presented. The system can help users to improve performing living activities and lead to certain independent life.

6. A naive bayes classifier based decision making module is designed for the assistive system for users with ASD. This model can perform like a caregiver who will send prompts when users get stuck in procedures of activities.

7. We have developed a multi-modal communication approach for young children with Spinal Muscular Atrophy (SMA). The system includes various reliable and low-cost sensors to increase interactivities for the users who cannot use regular input devices.

8. A comprehensive interactive system that utilizing the communication approach is designed. It can help young children with SMA to communicate with others, control smart devices in home, and ask for help. The system is web-based which means it is compatible with multiple platforms and devices.

1.5 Organization

The remainder of the dissertation is organized as follows: Chapter 2 studies situation awareness (SA) in the military context. Experiments and results of determining value of information are reported. Moreover, some initial results for integrating complementary and contradictory information are included in this chapter. Chapter 3 and Chapter 4 present approaches of SA in the healthcare context. Chapter 3 focuses on an ambient intelligent system for autism spectrum disorder users and Chapter 4 focuses on communication and environment control systems for spinal muscular atrophy users. These studies can help to improve the quality of life for special groups. Chapter 5 concludes the dissertation and introduces the future work.

Chapter 2

Situation Awareness Application in Military Context

2.1 Introduction

In March 2008, the United States Army Training and Doctrine Command (TRADOC) published its most recent version of the TRADOC Pamphlet 525-66, "Military Operations Force Operating Capabilities." [14] The purpose of this pamphlet is to identify "capabilities necessary of the Army to fulfill war fighting concepts." Eleven Force Operating Capability (FOC) areas are identified in the document, and two of these are Battle Command and Battlespace Awareness. Capability areas outlined in the Battle Command FOC include information and decision superiority. In a dynamic battlefield environment, timely and accurate information is critical to the goal of providing relevant information to commanders, improving situation awareness (SA), and increasing knowledge. A key part of the Battlespace Awareness FOC is the ability to transform the information being provided by sensors and other means into knowledge and understanding; the Army recognizes that this process must involve "a mixture of automation and human cognition." [14]

Figure 2.1 shows the military intelligence process. The diagram shows the dynamic nature of the process. It also conveys that the overall goal of the process is to provide the commander with SA and understanding within the context of a mission. Operations processes produce the intelligence requirements needed to achieve

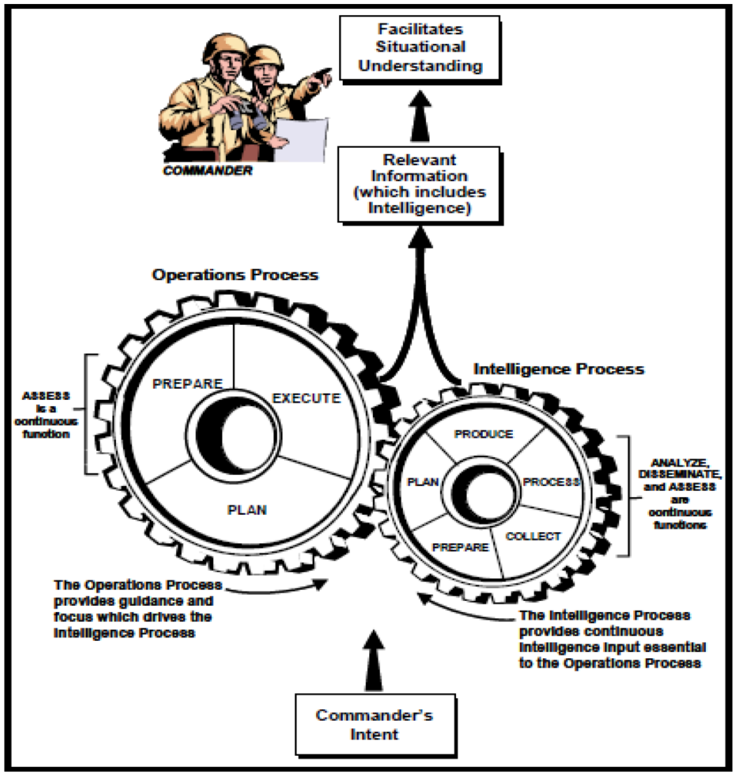

Figure 2.1: Intelligence Process [1]

mission objectives, while the intelligence process continually feeds operations. It is clear that relevant accurate, timely information is essential for proper SA and understanding.

Commanders rely on intelligence analysts to fulfill information requirements in order to define courses of actions. SA is required in order to make operational decisions. The intelligence analyst makes decisions regarding what information to send to the commander, and the commander seeks to develop multiple potential courses of actions and then act upon the most appropriate one.

SA is formally defined as "a person's perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future" [24]. The problems that routinely arise when humans are faced with trying to turn data and information into useful, actionable knowledge include information overload, complexity, and turbulence [25]. These same problems plague intelligence analysts. It is clear that automated tools, and perhaps process improvements, are needed to support effective decision making in highly complex military environments. These tools and process improvements are needed to assist in quickly determining important and accurate data in order to provide commanders with information and intelligence products.

From sophisticated unmanned ground acoustic sensors to open-source news feeds, the military commander and his staff are challenged not only by the familiar information overload dilemma, but also with the perhaps more critical problem of separating the important information from the routine. Calculating information importance, termed the value of information (VoI) metric, is a daunting task that is highly dependent upon its application to dynamic situations [24]. Solution flexibility is critical since VoI understanding must be readily applicable across a disparate range of information types. The process of assigning a VoI determination to a piece of information has historically been a multi-step, human-intensive exercise requiring intelligence collectors and analysts to make judgments within differing operational situations.

The cognitive processes behind assigning a VoI assessment resist codification with exact precision and offers an excellent opportunity to leverage a computational intelligence solution using fuzzy inference. Fuzzy systems have been shown to be effective at approximating reasoning where information is uncertain, incomplete, imprecise, and/or vague [25–28]. Recently, a Fuzzy Associative Memory (FAM) architecture was used to develop a system to calculate VoI in complex military environments based on the information's content, source reliability, latency, and the specific mission context under consideration [29, 30]. Military intelligence analysts were used as subject matter experts to provide the fuzzy association rules from which the system was constructed, and preliminary results from the system have been demonstrated and "validated" in principal and context [17, 31]. Efforts are continuing towards a more formal validation of the system and to empirically evaluate the effects of the system on intelligence analyst performance [32, 33]

This study presents an investigation on the effect of using two different membership functions within the fuzzy-based system and a comparative analysis of the differences between them. Moreover, we develops customizable system which provides flexibility for Subject Matter Experts (SMEs) to tune fuzzy logic models for specific battle contexts. Since military decisions require the ability to amalgamate information across a dissimilar array of sources: many times not always agreeing.

Based on this framework, we then study how analysts combine complementary and/or contradictory information in order to achieve the necessary SA to make sound decisions. Furthermore, we present a detailed explanation on how social media data affect SA and our focus is set on the military environment. The study reveals that SME based human computation is necessary during the SA process.

The remainder of this chapter is organized as follows: Section 2 presents background information related to VoI determinations within the military domain. A brief description of the recently developed fuzzy-based VoI system prototype is also provided in this Section. Section 3 reports on experiments and results of determining VoI for single piece of information. Section 4 discusses an approach taken to frame the decision space, the process being used to elicit knowledge, and some initial results for integrating complementary and contradictory information. In section 5, we study how to use social media data in military SA and human computation role during this process. Section 6 concludes this chapter.

2.2 Literature Review

2.2.1 What is Value of Information?

In order to turn large amounts of disparate information into useful knowledge, it is vital to have some way to judge the importance of individual pieces of information; the Value of Information (VoI) metric is used to do this. Ranking the "value" of information is a formidable task involving not only the sheer amount and diversity of information, but also the idea that the value of a piece of information will likely be influenced by the specific mission context to which it will be applied.

Before going further, it is useful to briefly address what is meant by information "value", and differentiate it from what could be meant by information "quality". One viewpoint is that "quality" refers to the fitness of data with respect to the inherent attributes of the data (accuracy, precision, timeliness, freshness, resolution, etc.) while "value" addresses the utility of the data within a specific application context [34]. The definition used in this study comes from that provided by [35]. Wilkins considers the practical importance of the information to the receiver, suggesting that information with value supports the receiver's ability to make informed decisions. The procedure for alphanumerically rating the "confidence" or "applicability" assigned a piece of information is essentially described in the annex to NATO STANAG (Standard Agreement) 2022 as well as in Appendix B of the US Army Field Manual FM-2-22.3 [1].

2.2.2 Value of Information Determination

U.S. military doctrinal guidance for determining VoI is vague at best [1] and does not address integrating the mission context into the decision. The guidance provides table 2.1 and table 2.2 for judging the "reliability" and "content" of a piece of data, with each characteristic broken into six categories. Reliability relates to the information source, and is ranked from A to F (reliable, usually reliable, fairly reliable, not usually reliable, unreliable, and cannot judge). Information content is ranked from 1 to 6 (confirmed, probably true, possibly true, doubtfully true, improbable, and cannot judge).

Doctrinal guidance does not provide any process for combining these determinations into a VoI metric. Additionally, it is obvious that combining only these two assessments of a piece of information would fall far short of representing all the critical aspects needed for a useful VoI determination.

Two other potential data characteristics include mission context and timeliness. Timeliness relates to how long ago the piece of information was collected, while mission context is set by the operational tempo of the military operation underway. The operational tempo relates to the decision cycle for the mission; that is, the time that can or will be used to plan, prepare, and execute the mission. Fast tempo operations may have a decision cycle measured in minutes to hours, while slower tempo operations may be measured in months or longer.

2.2.3 Value of Information System

It is interesting to observe that the information labels for source reliability and information content provided by the doctrinal evaluations are words, not numbers. The linguistic variables used as labels for information evaluations represent degrees of confidence (“usually reliable”, “probably true”, “improbable”). This arrangement

A	Reliable	**No doubt** of authenticity, trustworthiness, or competency; has a history of complete reliability
B	Usually Reliable	**Minor doubt** about authenticity, trustworthiness, or competency; has a history of valid information most of the time
C	Fairly Reliable	**Doubt** of authenticity, trustworthiness, or competency but has provided valid information in the past
D	Not Usually Reliable	**Significant doubt** about authenticity, trustworthiness, or competency but has provided valid information in the past
E	Unreliable	**Lacking** in authenticity, trustworthiness, and competency; history of invalid information
F	Cannot Judge	**No basis** exists for evaluating the reliability of the source

Table 2.1: Source Reliability [1]

1	Confirmed	**Confirmed** by other independent sources; **logical** in itself; **Consistent** with other information on the subject
2	Probably True	Not confirmed; **logical** in itself; **consistent** with other information on the subject
3	Possibly True	Not confirmed; **reasonably logical** in itself; **agrees with some** other information on the subject
4	Doubtfully True	Not confirmed; possible but **not logical; no other information** on the subject
5	Improbable	Not confirmed; **not logical** in itself; **contradicted** by other information on the subject
6	Cannot Judge	**No basis** exists for evaluating the validity of the information

Table 2.2: Information Context [1]

makes fuzzy logic an appropriate choice for computing the VoI [36]. A Fuzzy Associative Memory (FAM) model was chosen to construct the prototype fuzzy system. A FAM is a k-dimensional table where each dimension corresponds to one of the input universes of the rules. The ith dimension of the table is indexed by the fuzzy sets that comprise the decomposition of the ith input domain. Fuzzy if-then rules are represented within the FAM. For the prototype system, three inputs are used to make the VoI decision: source reliability, information content, and timeliness; the concept of varying mission contexts is also incorporated and will be explained shortly. The overall architecture of the prototype fuzzy system is shown in Figure 2.2. Instead of using one 3-dimentional FAM, two 2-dimensional FAMs were used. The reasoning behind this decision was presented in detail in [30] but essentially it provided a simpler knowledge elicitation process for the SMEs, decreased the total number of fuzzy rules, and provided a potential for the output of the first FAM to be useful on its own.

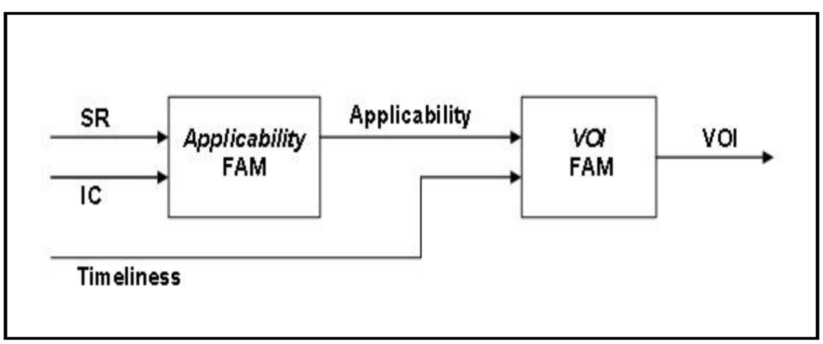

Figure 2.2: VoI System Architecture

As seen in Figure 2.2, two inputs feed into the Applicability FAM: source reliability (SR) and information content (IC); the output of the FAM is termed the information applicability decision. Likewise, two inputs feed into the VoI FAM: one of these (information applicability) is the output of the first FAM while the other input is the information timeliness value. The output of the second FAM, and the overall system output, is the VoI metric. The fuzzy rules represented in the FAMS capture the relationships between the input and output domains. The decomposition of the input and output domains into fuzzy sets defines the "language" of the rule base, which determines the antecedents and consequents that are used in the fuzzy rules. Since both FAMs have two inputs and one output, all the fuzzy rules in the system will be in the form "If X is A and Y is B, then Z is C", where A and B are fuzzy sets of the input domains and C is a fuzzy set over the output domain. For example, an actual rule in the Applicability FAM might be: "if Source Reliability is Usually Reliable and Information Content is Probably True, then Information Applicability is Highly Applicable." Knowledge elicitation from military intelligence SMEs was used to construct the fuzzy rules [31].

Within the Applicability FAM, the two input domains (source reliability and information content) are divided into five fuzzy sets following the guidance provided in [1]. The omission of the "cannot judge" category from both input domains is explained in [30]. The "information applicability" output domain was decomposed into nine fuzzy sets (ranging from not applicable to extremely applicable) while the VoI output domain utilized eleven fuzzy sets (ranging from not valuable to extremely valuable).

Up to this point, the contribution of mission context has not been mentioned. To account for differing mission tempos, three separate VoI FAMs were derived to represent three different tempos. Missions were characterized as either 'tactical' (high-tempo), 'operational' (moderate-tempo), or 'strategic' (slow-tempo). The system selects the correct VoI FAM based on the indicated mission context, thereby utilizing the appropriate fuzzy rule base to provide the VoI determination. The pairing of mission context within the VoI weight represents a new approach in transforming data into decisions.

More detailed descriptions of the FAMs, the fuzzy rule bases, the domain decompositions, and other implementation aspects of the system can be found in [17]. The series of surveys and interviews with SMEs that were used to integrate cognitive requirements, collect functional requirements, and elicit the fuzzy rules is presented in [31]. The VoI prototype system, the initial version and phase 2, has been demonstrated to the SMEs. Both versions of the prototype and its output have met SME expectations [32]. Note that there is no current system against which the results can be compared. As such, the system has not been tested thoroughly due to the human-centric, context-based nature of the problem and usage of the system.

Formal validation of the VoI system requires a comprehensive experiment which is currently under separate development.

2.2.4 Human Computation

The term human computation (HC) is defined and used as early as 1838 [37] in philosophy and psychology literature. It is introduced to the context of computer science theory afterwards [38] and its modern usage is inspired by von Ahn's 2005 dissertation titled "Human Computation" [39]. The definition from that thesis is "a paradigm for utilizing human processing power to solve problems that computers cannot yet solve”. Other authors use slightly different but compatible definitions [40, 41] and they all respect two critical statements which constitute human computation [42]:

- The problems fit the general paradigm of computation, and as such might someday be solvable by computers.
- The human participation is directed by the computational system or process.

Human computation has grown rapidly in both academy and industry. Various HC systems have been proposed for different purposes. For example, von Ahn and Dabbish introduce an interactive system to determine the contents of images by providing meaningful labels to users [43]. A similar "TagATune" system is presented by Law and von Ahn, which employs input-agreement as a quality control mechanism for human computation games [40]. Knowledge Collection from Volunteer Contributors is another kind of HC system which applies the collection method of aggregation [44]. Wolfers and Zitzewitz suggest an approach to handle market-generated forecasts by using HC [45]. HC has been used for search service too. For example, ChaCha [46] uses humans to interpret search queries and get the most relevant results.

2.2.5 Social Computing and Social Media

Social media is a virtual world governed by social norms and human behavior. It has become a hot topic in recent years. Obviously, humans are included in a social role where communication is mediated by technology. Social computing is defined as the interplay between persons' social behaviors and their interactions with computing technologies [47]. The key distinction between human computation and social computing is that social computing facilitates relatively natural human behavior that happens to be mediated by technology, whereas participation in a human computation is directed primarily by the human computation system [48].

Different research areas in social computing have grown quickly in recent decades [48]. For example, [49, 50] focus on social network theory and studying the structure of social networks to help stop the spreading of viruses, propagating information effectively, and checking robustness. Ranking and aggregating social media data are another recent research direction. [51] integrates an opinion identification toolkit into the retrieval process when processing social media data. [52, 53] propose different graph based ranking algorithms to analyze social data separately. [54] discusses how to use mediators strategies in social computing. Besides, web spam detection is another active area. [55] introduces an algorithm that exploits not only the structure of the Web graph but also page contents and features. In [56], a novel approach for identifying video spam is presented. [57] introduces a spam detecting system in social bookmarking.

There are also some other computational topics, such as query log processing, graph/link analysis and mining, and collaborative filtering. Review of those previous works can be found in [48].

2.3 Adjustable Method for Determining Value of Single Information

Recently, a Fuzzy Associative Memory (FAM) architecture was used to develop a system to calculate Value of Information (VoI) in military environments based on the source reliability (SR), information content (IC), and latency. However, since military contexts are various and complex, it is necessary to explore adjustable VoI determination to fit proposed fuzzy logic model to a specific battlefield. SMEs require flexible models for some special contexts thus increase the accuracy of VoI calculation. This section will present an adjustable method to help SMEs tune customized models and an investigation on the effect of using them.

2.3.1 Adjustable Value of Information Framework

A major factor in the design of any fuzzy system relates to the decomposition of the input and output domains into fuzzy sets. The "shape" of the fuzzy sets defines the membership functions for the system. While there are numerous shapes for fuzzy sets (triangular, trapezoidal, Gaussian, bell, and the like), triangular membership functions were used in the initial VoI system. To further facilitate computational efficiency, it was also required that the triangles were isosceles with bases of the same width; this triangular decomposition with evenly spaced midpoints has been referred to as a TPE system [58]. Figure 2.3(a) shows the TPE decomposition of a domain ranging from 1 to 5; Figure 2.3(b), (c), and (d) illustrate isosceles triangular decompositions of the same range that do not adhere to the restriction of having bases of the same width. Triangular decompositions, with and without bases of the same width, are included in our experimental framework.

In addition to using triangular membership functions, trapezoidal decompositions are another approach we would like to explore. Similar to the triangles, isosceles trapezoids both with and without bases of the same width are considered. Figure 2.4(a) shows the decomposition of a domain ranging from 1 to 5 using

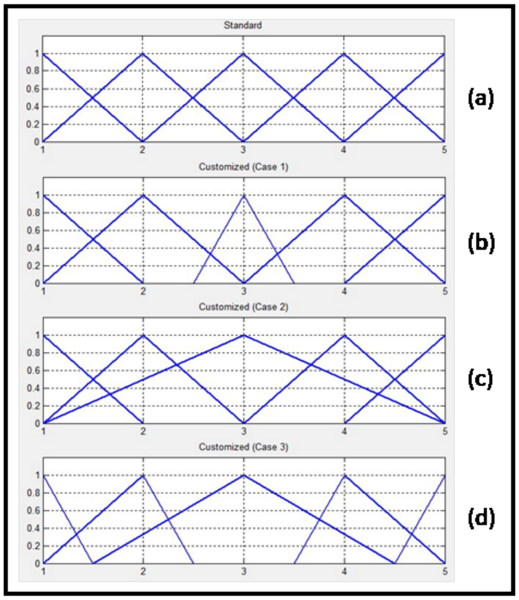

Figure 2.3: Standard and Customized Triangular Fuzzy Membership Functions

isosceles trapezoids with bases of the same width; Figure 2.4(b), (c), and (d) depict similar decompositions using isosceles trapezoids without the requirement for equally sized bases.

While we mentioned several forms of membership functions from which to choose, we selected trapezoidal and triangular fuzzy sets for two primary reasons. First, the membership degree calculations for both are linear, thereby facilitating high computational efficiency. This is significant since the purpose of the fuzzy VoI system is to help intelligence specialists find the most important information within a potentially large amount of data while frequently adhering to restrictive time constraints.

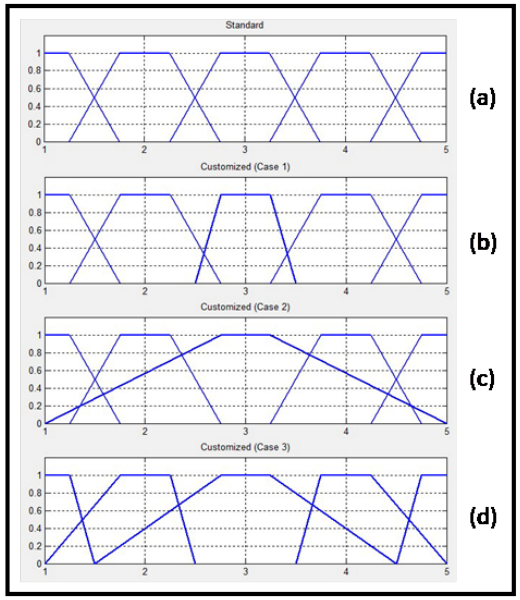

Figure 2.4: Standard and Customized Trapezoidal Fuzzy Membership Functions

The second reason is that these two forms can help in the data acquisition process. As implied earlier, significant knowledge elicitation efforts using intelligence specialists as Subject Matter Experts (SMEs) were required to construct the initial fuzzy rules; likewise, any membership function optimization will be determined by the SMEs. The triangular and trapezoidal functions are more visually understandable and provide an environment more conductive to human-in-the-loop knowledge acquisition. Based on these two reasons, trapezoidal and triangular membership functions are often used [59].

To facilitate the analysis of various domain decompositions using the triangular and trapezoidal fuzzy sets, we compare them from different aspects and display the results visually. Three categories of experiments are presented in the next section.

First, results using "standard" triangular and trapezoidal decompositions are compared, where "standard" means the use of isosceles shapes with bases of the same width (Figure 2.3(a) and Figure 2.4(a)). Next, "standard" triangular fuzzy membership functions are compared with "customized" triangular fuzzy membership functions, where "customized" means that the restriction for bases of the same width is removed (Figure 2.3(b), (c), and (d)). Finally, "standard" trapezoidal fuzzy sets are compared with "customized" trapezoidal fuzzy decompositions (Figure 2.4(b), (c), and (d)).

2.3.2 Experiment

This section provides the experimental results from comparing triangular and trapezoidal fuzzy set membership functions. Three subsections will be used to present the results. First, a comparison of the "standard" triangular and trapezoidal sets will be shown. Next, several "customized" triangular decompositions will be compared with the initial TPE fuzzy sets. Finally, several "customized" trapezoidal decompositions will be compared with the standard trapezoidal fuzzy sets.

2.3.2.1 Standard Triangular vs Standard Trapezoidal

Figure 2.5 compares the FAM outputs for the standard (TPE) triangular fuzzy sets (a, c) and the standard trapezoidal fuzzy sets (b, d). Figure 2.5(a) and (b) show the applicability FAM output for the two models; that is, the relationship between source reliability (x-axis) and information content (y-axis). The values of two inputs are from one to five, with the smaller value of one being "better" (better reliability/content) and five meaning "worse" (less reliability/content). The applicability output values vary from one to nine where the larger values represent better applicability; the colors vary from blue to red where the higher value is in blue (high applicability meaning reliable, probable information) and the lower value is in red (unreliable, improbable information).

Figure 2.5(c) and (d) show the VoI FAM output based on the two inputs of applicability and timeliness. Applicability is as mentioned above. Timeliness reflects the temporal age of the information, with values ranging from one to three: one means "recent" while three means "old". As with the applicability graphs, the VoI values are represented in the color shades within the graph. The numerical values for VoI range from zero to ten (blue meaning ten; red meaning zero) and the higher values represent higher VoI (more valuable information). The mission context is assumed to be "tactical".

Comparing results for the models, the output landscape of the triangular fuzzy models (a, c) looks smoother while the trapezoidal fuzzy models (b, d)produce some fairly well defined rectangles. To see why, note that when an input (in these standard models) has a membership value equal to 1 in a fuzzy set, the input belongs only to that fuzzy set (see Figure 2.3(a) and Figure 2.4(a)). For example, in the triangular fuzzy model, only the integer input values (1, 2, etc.) belong to just one fuzzy set; that is, there is only one input value in each triangular fuzzy set that will have a membership equal to one. For the trapezoidal fuzzy sets, however, there are several values in each set that have a membership equal to one and, thus, belong to only that fuzzy set. This creates areas within the color graphs that have the same calculated output values for applicability or VoI, thereby producing the more pronounced rectangles. Note these rectangles are seen within the color graph and at the four corners of the graph.

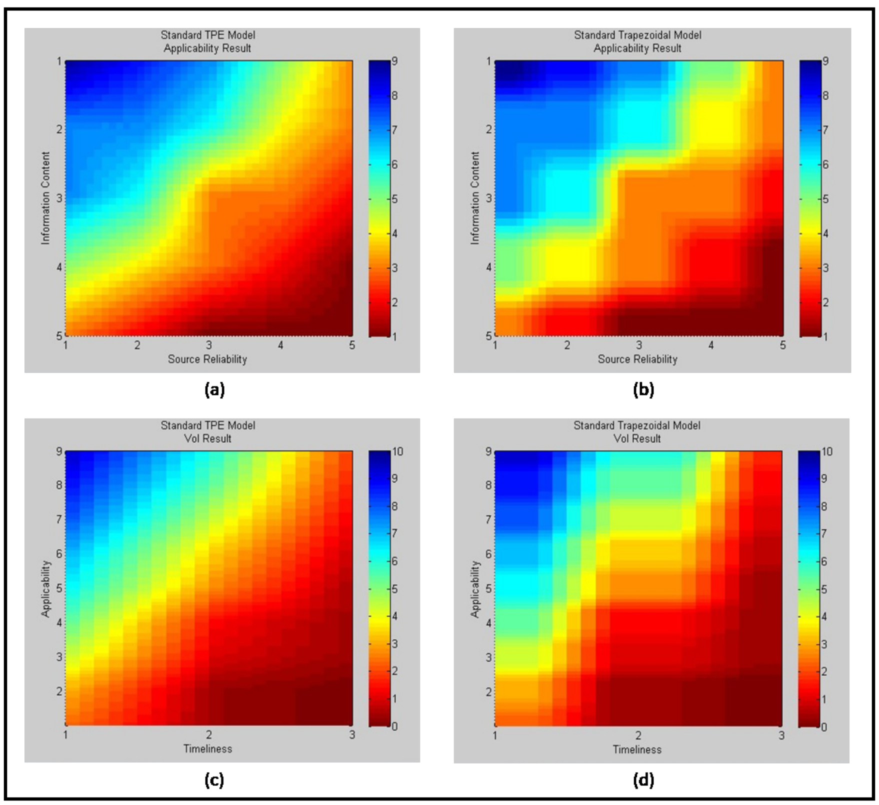

Figure 2.5: Applicability and VoI: Standard Triangular and Trapezoidal Fuzzy Sets

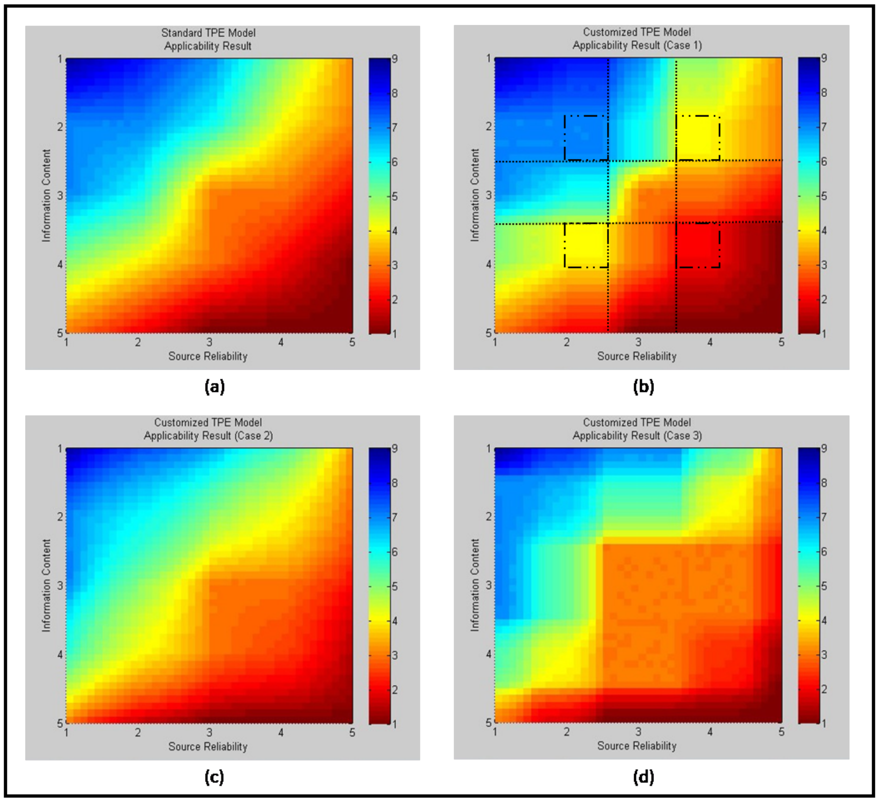

Figure 2.6: Applicability: Standard and Customized Triangular Fuzzy Sets

2.3.2.2 Standard Triangular vs Customized Triangular

In the experiments shown below, since the calculations of applicability and VoI are very similar, comparisons between the different models will be illustrated using the results of the applicability FAM only. Figure 2.3 and Figure 2.6 are used to compare the applicability values for the standard and customized triangular fuzzy models; Figure 2.3 shows the fuzzy set shapes for all domains in the standard model (2.3(a)) and the customized models (2.3(b), (c), and (d)), and Figure 2.6 provides the associated color graphs.

Case 1

In the standard model, both domains (source reliability and information content) are decomposed following the TPE restrictions as illustrated in Figure 2.3(a). The resulting color graph for the standard (TPE) model is shown in Figure 2.6(a). In the customized model, both inputs shrink the third fuzzy set from the standard 2 to 4 width (on the x-axis) to the customized width of 2.5 to 3.5 as shown in Figure 2.3(b); the corresponding color graph is Figure 2.6(b).

Compared with the applicability distribution of the standard model, the color graph for the customized model is much less smooth. It is also clear that two very similar color belts cross in the middle of the graph (as outlined); the edges are 2.5 and 3.5 (both vertically and horizontally). The middle of the graph for the customized model has similar color values; however, the outer edge of the color belt has smaller changes than that in the standard model and the inner edge has larger changes which cause the visible boundaries. Also, four rectangles in a solid color around the center are observable (and outlined) in Figure 2.6(b). The reason for the observed differences is that in the customized model, inputs between 2 to 2.5 and 3.5 to 4 only belong to one fuzzy set. This leads to smooth visualization and solid squares of the same color. On the other hand, input values between 2.5 to 3.5 belong to two fuzzy sets. The third fuzzy membership degree is changed faster (narrower triangle; slope is larger) than that in the standard model. As a result, this enhances the representation of boundaries.

Case 2

In this case, the third fuzzy set is assigned a wider range, encompassing the entire input domain, as shown in Figure 2.3(c). Again, in the standard model both domains are decomposed following the TPE restrictions (Figure 2.3(a)).

The values in the customized color graph in Figure 2.6(c) look smoother in the center with sharp variations occurring in red and blue at the corners; the red and blue corner values have a much smaller area than in the standard fuzzy triangular model. The reason is that the third fuzzy set in the customized model affects all the fuzzy membership degree calculations since it spans the entire input domain. For high value inputs, the third fuzzy set causes lower FAM values to join the calculation, resulting in a lower output value than that in the standard model.

The reverse occurs for the low value inputs; the middle fuzzy set contributes higher FAM values to the applicability result. Thus, the red and blue boundaries contract to the corners of the customized graph as compared to the standard triangular fuzzy model results.

Case 3

Considering that some users maybe prefer a wide range in the middle fuzzy sets (most IC and SR inputs would fall in the "middle") but smaller ranges at the edges (only extreme IC and SR inputs are considered "best" or "worst"), Figure 2.3(d) shows a fuzzy set pattern to provide such a system. In this model, the two ends are made narrower (range from 1 to 1.5 and 4.5 to 5), which means only a small range of inputs belong to these sets. The middle set has a wide input scope, which is from 1.5 to 4.5. Meanwhile, the input ranges of other two fuzzy sets are reduced appropriately.

Figure 2.6(d) shows the applicability distribution based on this customized model which is much more "blocky" than that of the standard TPE model shown in Figure 2.6(a). Because the middle fuzzy set is extended and covers numerous inputs, the resulting output has a number of areas in the middle values. The contraction of the other fuzzy sets causes much of the graph area to show up in orange and cyan colors, while only the corners have extreme high or low values corresponding to dark blue and dark red.

2.3.2.3 Standard Trapezoidal vs Customized Trapezoidal

Figure 2.4 and Figure 2.7 are used to compare the applicability values for the standard and customized trapezoidal fuzzy models; Figure 2.4 shows the fuzzy set shapes for all domains in the standard model (2.4(a)) and the customized models (2.4(b), (c), and (d)), and Figure 2.7 provides the associated color graphs.

Case 1

In the standard model, both domains (SR and IC) are decomposed as illustrated in Figure 2.4(a). The resulting color graph for the standard model is shown in

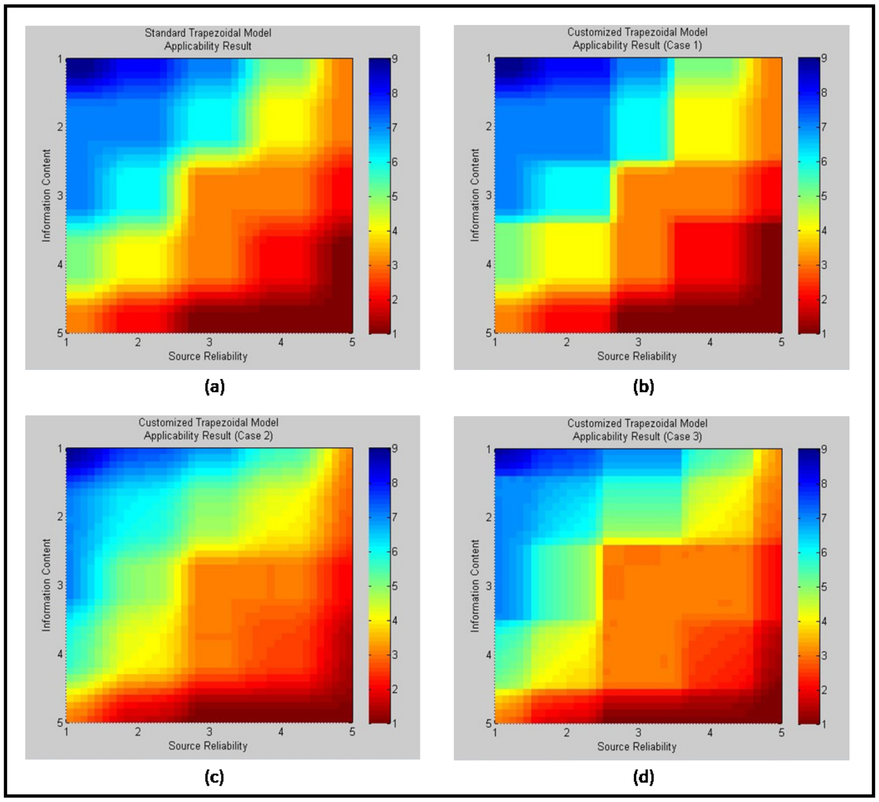

Figure 2.7: Applicability: Standard and Customized Trapezoidal Fuzzy Sets

Figure 2.7(a). In the customized case, the middle fuzzy set is still an isosceles trapezoid but the width is smaller than the other sets, as depicted in Figure 2.4(b). The left and right bottom points are 2.5 and 3.5; note the upper base is the same 2.75 to 3.25 as in the standard trapezoidal model. The corresponding color graph is shown in Figure 2.7(b).

As in Case 1 of the triangular fuzzy model, the color graph for this customized trapezoidal model illustrates two similar color belts crossing in the middle of Figure 2.7(b). Because the middle fuzzy set is narrower and more inputs belong to only the second or fourth fuzzy set, the edges corresponding to the middle SR and IC input values are smaller and more pronounced than those of the standard trapezoidal model shown in Figure 2.7(a). Also, the neighboring rectangles of solid color are larger than that in the standard model. Note that the areas associated with the four corners are similar in both the standard and customized color graphs.

Case 2

Considering that the opposite setup with the middle fuzzy set as shown in Figure 2.4(c), this case sets the middle fuzzy set to cover a wider input range, from 1 to 5. However, the upper base is still fixed from 2.75 to 3.25 and all other sets are the same as in the standard trapezoidal model.

Figure 2.7(c) illustrates the associated color graph. The result of the customized trapezoidal model reveals a similar trend as that of the corresponding triangular model; more areas in the middle values can be observed and a sharp variation happens in the corners as compared to the standard trapezoidal model in Figure 2.7(a). Nevertheless, the graph still presents the basic features of the trapezoidal fuzzy model - some rectangles in similar colors exist in the color graph which are not as obvious in the triangular fuzzy model.

Case 3

Based on the same scenario as with Case 3 for the triangular fuzzy sets, this customized trapezoidal model sets up a wide middle fuzzy set and narrower side sets as shown in Figure 2.4(d). In this setup, only very high or low value inputs are regarded as extreme conditions.

Figure 2.7(d) shows the applicability distribution based on the customized model. Compared with the result of the standard trapezoidal fuzzy model in Figure 2.7(a), similar results occur as with Case 3 for the triangular model. The areas in the middle values are larger than those in the standard trapezoidal model (Figure 2.7(a)) and only small sections of dark red and blue in the corners represent the extreme high and low applicability values. Moreover, the result of this customized model retains the features of a trapezoidal fuzzy model which produces larger areas of solid colors in the graph. However, one difference is that color boundaries between the rectangles are much narrower in the customized model. This makes the boundaries more pronounced and provides well-defined solid color rectangles.

2.3.3 Discussion

From the observation of triangular and trapezoidal fuzzy membership functions, the first major difference is that when using the triangular approach the results of the color graphs were strikingly different than those of the trapezoidal approach. Using the triangular fuzzy model produced graphs that were infinitely smoother in their transition between calculated values. The trapezoid models, on the other hand, produced plots that appeared "blockier", lending to larger areas of continual homogeneous values.

A second major difference observed is the increased flexibility for representing membership functions afforded by the trapezoidal representation. Using the trapezoids allowed an "interval of values" that maximized the individual membership functions (top of the trapezoids) as compared to the triangle representation that permitted only one. The introduction of the trapezoid dramatically increases the ability of the user to capture representations over the more simplistic triangular shape.

With this understanding, one might mistakenly choose one approach over the other, thinking on one hand the trapezoidal approach is inferior because of the “blocky effect”, or on the other hand, superior because of the added flexibility. The fact is both approaches have their own strengths and weaknesses. For example, depending on the context of the situation, the blocky effect might provide a better representation of the military function being modeled. An example of this effect can be seen when comparing a logistic battle function against that of a tactical combat battle function.

For the logistics operations, the fidelity of the information required for moving equipment can be significantly less critical than that required when conducting a combat cordon and search operation; as such, the logistical representation of VoI may very well be represented with larger areas of homogeneous values (the blocking effect).

Military contexts are extremely various in the battlefields thus they have different intelligence preferences. For example, a tactical operation requires information that is more fresh while a strategic military decision should be supported by information that has higher applicability. Commanders in the battlefield need different calculations of VoI which means various fuzzy logic models can be contributory. The experiments we have done provides two different shapes of membership function and SMEs are allowed to tune a customized model for a special military context. This will help to improve the accuracy of calculating VoI in a specific scenario and lead to better situation awareness. SMEs in U.S. Army have utilized our system to perform an assessment of efficacy to current military operations [60].

2.4 Integrating Complementary/Contradictory Information

While the prototype system has been shown to produce a reasonable Value of Information (VoI) determination based on the information characteristics reviewed above, it is clear that this VoI rating may change over time. One extremely simple case is that, since timeliness is one of the characteristics used to compute VoI, as time passes the VoI determination would be expected to change. But even without this obvious observation we can understand that the VoI evaluation for a piece of information might be impacted by the acquisition of new information.

As the information database constantly grows, a method for “re-prioritizing” the now-modified cache of information is needed. The VoI determinations, of course, are used to signify the relative priority amongst the pieces of information. Even within a constant mission context, however, new information might change previous VoI ratings by either (1) somehow revealing a change in the source reliability of a particular source, or (2) providing a piece of information that contradicts or complements previous information. A method for updating information priority based on new contradictory or complementary information is of particular interest to the military intelligence analysts.

Within the prototype system the information content characteristic is concerned with whether a piece of information is complementary or contradictory with other information. The literature shows much work has been done in this area; the situation of complementary or contradictory information is typically termed as 'data conflict'. Before describing our methodology for dealing with this issue we will briefly outline some of the significant work found in the literature.

2.4.1 Strategies for Handling Data Conflict

Data conflict can be divided into two types: one relates to the uncertainty about the attribute value caused by missing information, and the other is concerned with the difference in attribute values [61]. The uncertainties are generally easier to cope with while differing values are more complicated because it is between two or more different non-null values. Handing conflicting data has been an active research topic in the data fusion area for decades and this is the focus of our current problem. A recent survey provides us a detailed classification of strategies to handle inconsistent data, as shown in Figure 2.8; these strategies have been summarized in Table 2.3 [2].

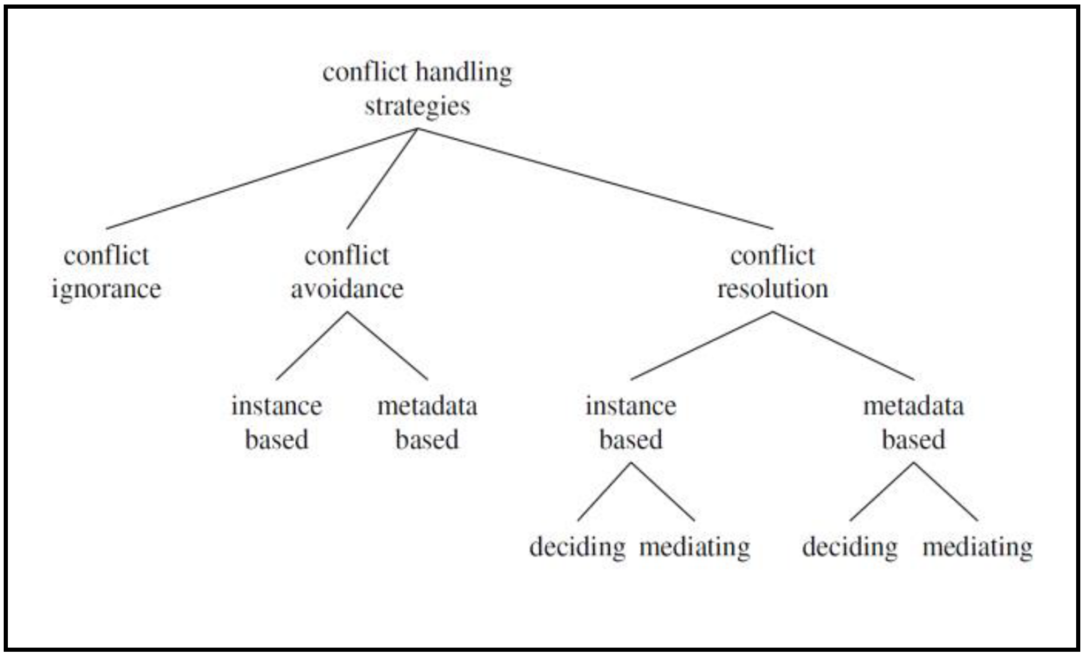

Figure 2.8: A classification of strategies to handle inconsistent data [2]

Basically, there are three classes of data conflict handling strategies; these are conflict ignorance, conflict avoidance, and conflict resolution. A number of commercial and prototypical data integrating systems in different classes have been proposed to handle complementary or contradictory information.

Strategies in the conflict ignoring class do not provide resolutions of uncertainties or contradictory information. Sometimes, they are not even aware of data

Strategy	Classification	Short Description
PASS IT ON	ignoring	escalates conflicts to user or application
CONSIDER ALL POSSIBILITIES	ignoring	creates all possible value combinations
TAKE THE INFORMATION	avoiding, instance based	prefers values over null values
NO GOSSIPING	avoiding, instance based	returns only consistent tuples
TRUST YOUR FRIENDS	avoiding, metadata based	takes the value of a preferred source
CRY WITH THE WOLVES	resolution, instance based, deciding	takes the most often occurring value
ROLL THE DICE	resolution, instance based, deciding	takes a random value
MEET IN THE MIDDLE	resolution, instance based, mediating	takes an average value
KEEP UP TO DATE	resolution, metadata based, deciding	takes the most recent value

Table 2.3: Conflict Resolution Strategies

conflicts. Two typical strategies in this class are 'Pass it on' and 'consider all possibilities'. 'Pass it on' presents all values to the user or application and defers resolution to them, while 'Consider all possibilities' creates all possible value combinations and transfers them to the front end. Pegasus [62], developed by HP Labs, is a widely-used conflict ignoring system, which focuses on multi-database aspects. It is based on the object-oriented data model and allows the user to access different heterogeneous distributed data sources. Based on Pegasus, a number of systems have been developed to extend it from different aspects. For example, Nimble, another integration system, uses the conflict ignoring strategy [63, 63]. It is the first XML-based commercial product and has been successfully implemented in several large companies. Both the Pegasus and Nimble systems use the 'Pass it on' strategy and escalate conflicts to users or applications.

The second class, conflict avoiding, includes strategies like 'Take the information', 'No gossiping', and 'Trust your friends'. All of them apply a unique decision equally to all data rather than performing individual resolution for each conflict. The first two strategies are decided by instance and the third one is based on metadata. Instance-based decisions mean the action is in regard to the actual data values while metadata-based decisions consider values based on metadata, such as freshness of data or the reliability of a source. TSIMMIS is one of the well-known conflict-avoiding systems [64]. Instead of using an existing data model it uses wrappers to extract data from sources and convert it into its own data model called an Object Exchange Model (OEM). The 'Trust your friends' strategy is used in the system and the final data representation is based on the preferred source. Besides TSIMMIS, Infomix [65] and HIPPO [66] are also classified as conflict-avoiding systems. Both systems follow the Consistent Query Answering approach(CQA) and use the 'No gossiping' strategy, which means they only return consistent tuples. The difference between them is that Infomix uses not only relational data but also XML data, HTML data or data from other data models, while HIPPO uses a special data structure called a conflict hypergraph. However, HIPPO takes advantage of the hypergraph model to reduce the computation of all repairs in the data.

The third class of handling inconsistent data is conflict resolution, which also includes instance-based and metadata-based methods; both of the methods can provide deciding and mediating strategies. Deciding strategies will return a preferred value among the existing values, while mediating strategies can produce an entirely new value, such as the average of a set of conflicting numbers. For the instance-based deciding method, 'Cry with the wolves' and 'Roll the dice' are two common strategies which provide the most often occurring value or a random value to the users or applications. 'Meet in the middle' is another instance-based method, but it is a mediating strategy and returns an average of the conflicting values. There is also a metadata-based deciding strategy called 'Keep up to date' which takes the most recent value as the output. Compared with conflict ignoring and avoiding systems, conflict resolution systems attract more attention. Hermes [67], Fusionplex [68], and HumMer [69] are three popular conflict resolution tools. All of them can provide complete and concise answers for queries. In particular, Hermes has a mediator that is specified by an expert in a declarative way. Fusionplex uses the 'Keep up to date' strategy only while the others use multiple approaches. In addition, HumMer provides semi-automatic virtual integration and the automated process continues as far as possible [70].

Based on these classified strategies, some advanced techniques have been proposed for conflict resolution. They focus on three different aspects: accuracy of sources, freshness of sources, and dependence between sources. Some data sources are more accurate and valuable than others and sometimes the difference between them can be very significant. In order to assess accuracy of sources, Dong et al. present a novel probabilistic model to compute the reliability of each source and return true values based on conflicting data [71]. Similarly, Zhao et al. present a Bayesian approach to score the source quality without supervision [72]. Wu and Marian [73] and Yin et al. [74] introduce a general assessment framework to calculate the source accuracy in a Web domain. Blanco et al. also apply a probabilistic model to the source assessment and enable several attributes working at the same time with Web data [75]. As the real world is dynamic, true values often evolve over time; some models have been proposed to consider the freshness of sources. A probabilistic model is presented in [76] to consider source freshness. This approach treats incorrect values and out-of-date values differently in truth discovery. Research on source dependence focuses on copy detection, which is another significant issue when integrating conflicting data. Data sources may copy from each other for all or some of the data, especially within the Web domain. Dong et al. and Berti-Equilleet al. propose some preliminary solutions and a probabilistic model to discover dependence between sources in [71, 77].

Although systems handling conflicting data have been proposed and developed, it is still a huge challenge to apply them in the military environment. Our problem dictates that we should not employ the strategies contained within the ‘conflict ignorance’ and ‘conflict avoidance’ classifications. The techniques present in the ‘conflict resolution’ class pose problems as well. Because of the special military intelligence analysis domain context, data training is often not possible. Further, many of the approaches return values (such as most recent, preferred, average or random) that may not be useful in our problem domain.

The accuracy of sources, freshness of sources, and dependence between sources are all critical in our context, so the techniques considering these are interesting. However, there is no general approach (or guidance) for how to combine these properties within our domain. Additionally, judgments related to sources are contained in the source reliability(SR) characteristic of the information; we hope to focus on the information context (IC) characteristic to update the VoI determinations based on new information.

2.4.2 Methodology and Initial Result

Based on the review of current data conflict handling methods proposed throughout the literature and the noted deficiencies relative to our domain, we propose a slightly different method to address the issue. Our approach is somewhat similar to the Hermes conflict resolution system in that we plan to use fuzzy rules provided by Subject Matter Experts (SMEs) as our "mediator". That is, knowledge from the military intelligence SMEs with respect to updating the VoI ratings will be translated into fuzzy rules and integrated into the prototype VoI system.

The necessary first step in augmenting the prototype system involves understanding how military intelligence analysts combine complementary and/or contradictory information in order to achieve the necessary situation understanding to make sound decisions. The following subsections detail the initial approach taken to frame the decision space, the process being used to elicit knowledge from the SMEs and some initial results.

2.4.2.1 Framing the Decision Space

In order to capture how analyzers combine information, there must first be an agreement to the abstract definitions that define complementary and contradictory information and the varying degrees to which a piece of information may or may not support the original premise. To frame the challenge the following definitions are given:

- The information is identical *totally complementary*
- The information supports the related premise *somewhat complementary*
- The information is unrelated *neutral*
- The information supports an alternative premise *somewhat contradictory*
- The information is totally opposite *totally contradictory*

As an example, suppose the original premise offered is: 'There is going to be a game at the stadium tonight' (original premise). Information that is *totally complementary* to this statement would state a fact that is in complete agreement. Whereas information that is *totally contradictory* would state the exact opposite: 'there is not going to be a game at the stadium tonight.' The definitions for *somewhat complementary* and *somewhat contradictory* are subtler though. In this example, the information that 'there is a backup at the exit from the stadium before the game' supports the related premise and therefore would be *somewhat complementary*. Where, the fact that there was 'not a backup at the stadium exit' would logically be *somewhat contradictory*. Neutral information is agnostic to the original premise.

2.4.2.2 Knowledge Elicitation

To assist in the knowledge elicitation process, a simple user prototype was designed to capture the *'perceived change in the VoI'* to an initial fact when it is combined with a new piece of information that either complements or contradicts it. The prototype allowed any combination of informational facts to be displayed. In addition to providing the scores associated with source reliability and information content, the additional fact given was also scored as to its degree of complementing or contradicting using the scale described above. The latency for both pieces of information was held constant (i.e., recent) for this exercise. The experiment was conducted by SMEs in U.S. Army and they gave the feedback which can help to improve the system.

An example of the user interface is shown in Figure 2.9. The initial fact, outlined in black, is scored as having an initial source reliability score of C and an information context score of 3. The additional fact given in this example has a source reliability score of A, an information content score of 2 and is designated as being *somewhat contradictory*. With that information, the user in this example adjusted the perceived VoI slider to a position that is very close to "*Significantly Increased*".

After the score is entered, the next two pieces of information are displayed.

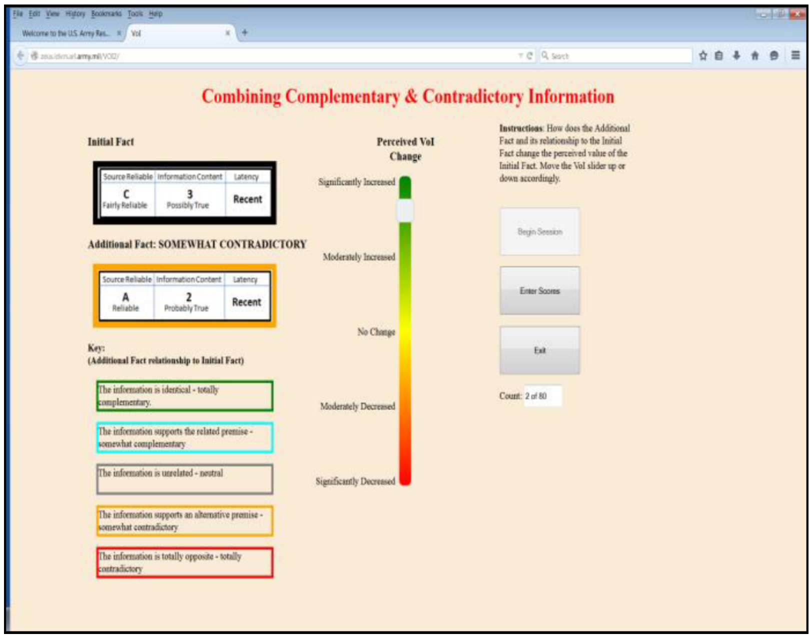

Figure 2.9: Complementary/Contradictory Capture Screen

For this exercise 80 different combinations of information were tested. The Initial Facts ranged from information classified as extremely valuable [A,2] to information of little value [E,4]. The Additional Facts offered varying degrees of complementing and contradicting information to the original fact from totally complementary to totally contradictory and were graded either better, the same or worse in informational value than the original fact.

2.4.2.3 Initial Results

During this exercise, four intelligence analysts rendered their opinions on the perceived change in VoI for 80 combinations of complementary and contradictory information. Since the military intelligence is classified and critical, SMEs in U.S. Army conducted analysis.

Anecdotal Summary of results:

1) The biggest change in perceived VoI occurred when combining total complementary data with a grade of [A, 2] with information graded above [C, 3]. In these cases, the change in perceived VoI was consistently rated with a *significant increase* rating. Information graded below [C, 3] saw changes close to moderately increase.

2) When the combining information was *somewhat complementary*, the rise in perceived VoI was directly related to the score of the original information. Restated, the higher the original information score the greater the potential change in perceived VoI. Original information with a lower rating saw the smallest change in perceived VoI; even when the additional fact was rated [A,2].

3) The results obtained when combining *total contradictory* information revealed the biggest between-subject differences. While the directions of the changes were generally the same, the magnitudes of the changes were obviously not aligned. Follow-up interviews will be conducted to determine the underlying cause.

4) When the combining information was *somewhat contradictory*, the decrease

SR \ APP	1	2	3	4	5	6	7	8	9
1			5		4		2.5		1
2		5		4		3	2	1	
3	5		3.5			2	1		
4	5	4	3	2	1				
5	4.5	3	1.5						

Table 2.4: Reverse Applicability FAM (APP: Applicability, SR: Source Reliability)

in perceived VoI followed the same pattern as that observed when combining somewhat complementary information, but it was less dramatic. Interestingly, the decrease was generally one whole point difference on the Likert scale.

2.4.2.4 System Modification

The knowledge elicitation effort detailed above is just the initial examination of if and how the process of combining information used by military intelligence analyst SMEs can be codified and automated. The preliminary results support the idea that a reasonable experiment can be developed to elicit the necessary fuzzy rules from the SMEs to represent the way they combine and reason over contradictory and complementary information. These new fuzzy rules will allow enhancements to the initial fuzzy-based VoI prototype so that changes in previous VoI determinations can be made to reflect the updated “value” of old information in the light of new information.

As mentioned previously, our methodology for handling data conflict is to modify the *information content* grade of a previously obtained piece of information to reflect the new VoI rating that it should now have based on some new piece of information. How the IC grade will be modified will be dictated by the fuzzy rules obtained from the SMEs. Based on the preliminary knowledge elicitation outlined in the section above, we examine the "backwards solving" to calculate the reverse Fuzzy Associative Memory (FAM) matrix given SR and applicability.

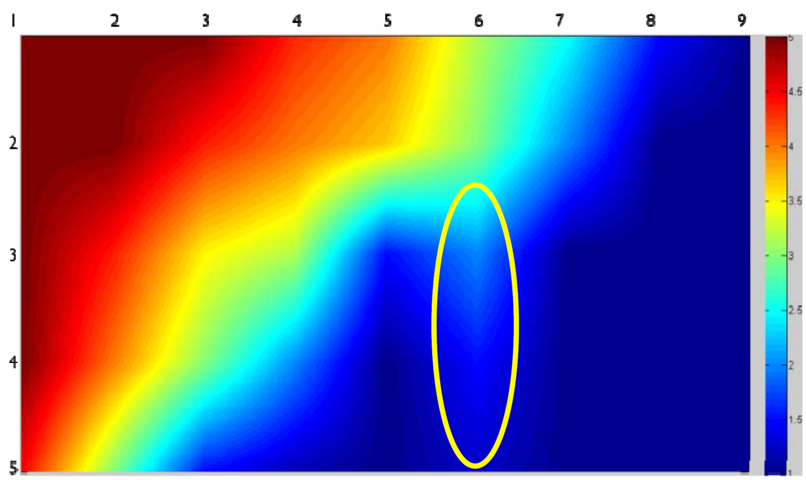

Figure 2.10: Reverse Applicability FAM-Region Growing

Table 2.4 shows a reverse applicability FAM using APP and SR as the inputs and IC as the output. Note there are some cells with no values which represent empty in the rule base. That is, there are SR and APP combinations for which no corresponding fuzzy rules were provided by the SMEs. To identify the empty rules based on knowledge provided by SMEs, we explore three approaches which are

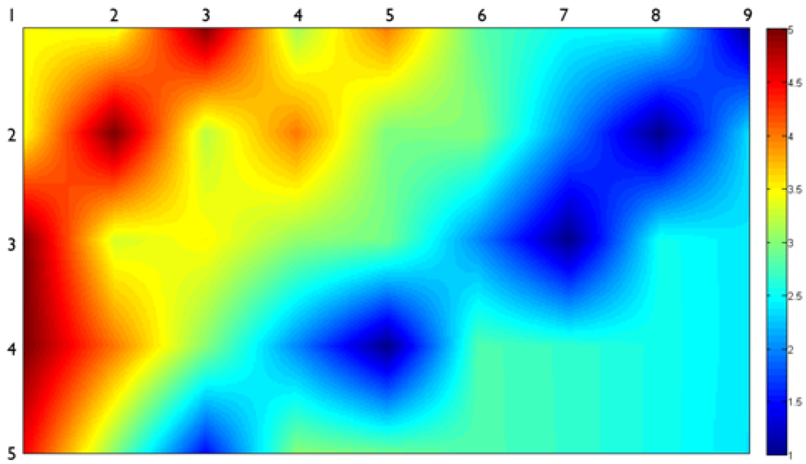

Figure 2.11: Reverse Applicability FAM-Weighted Averaging

region growing, weighted averaging, and rule base completion via training data. Figure 2.10, Figure 2.11, and Figure 2.12 illustrate the value distribution of these three approaches. The detailed empirical study is presented in [78].

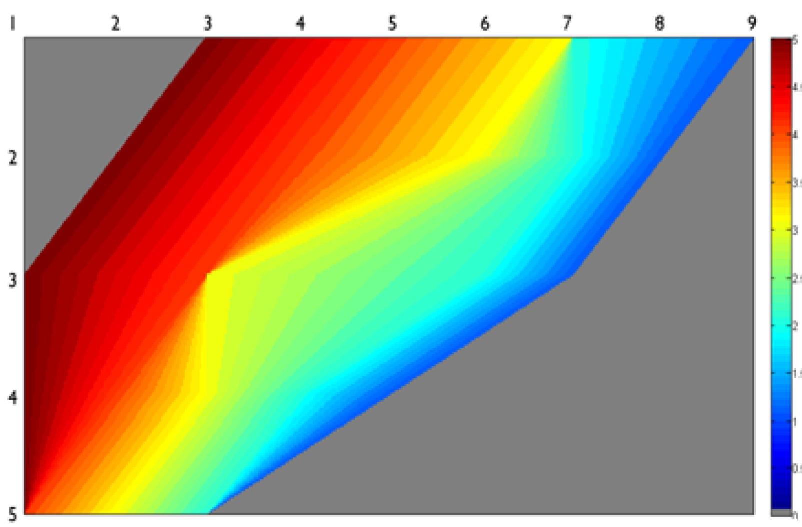

Figure 2.12: Reverse Applicability FAM-Training Data

From the study, none of the approaches resulted in a reasonable FAM. The region growing and weighted averaging methods did complete the empty cells in the FAM, but the resulting fuzzy rules were not consistent with the SMEs reasoning process illustrated by the original applicability FAM rule base. While rule base completion using training data produced resonable results for the cells that got filled in, over 35% of the FAM rules were still missing output values.

The final idea consisted of examining the reversed applicability FAM to see if the missing rule outputs were actually needed. That is, perhaps the reasoning process used by the SMEs to combine two pieces of information would not need the incomplete rules; this would imply that the reasoning process to combine information is consistent with the initial reasoning used to evaluate a single piece of information. Regrettably, data from an experiment using the SMEs to aggregate complementary and/or contradictory information revealed that the reverse applicability FAM (as well as the incomplete FAM that resulted from using the training data) was not sufficient to model the aggregation process. The proposed approach cannot work well for integrating complementary/contradictory information.

2.5 Utilizing Human Processing for Fuzzy-Based Military Situation Awareness Based on Social Media

Today social media such as Twitter, Facebook, YouTube, LinkedIn, and Flickr have attracted billions of individual users. A report from Netpop Research indicates that more than 76% of U.S. broadband users actively contribute to social media websites [79]. Social networking and related websites are among the top few online websites which have the highest Internet traffic [80]. Social media has changed the way that information is generated and propagated across communities and around the world. Users are able to collect, process, share and publish information anywhere and anytime, especially with the advances of Internet and mobile technologies. They can also easily find groups by interests. For example, movie fans can post their reviews on a Facebook group and sports players can upload videos to YouTube.

Data from social media can provide up-to-date information on topics which can cover almost every subject. They can be extremely useful for understanding the world for any individual user. From another perspective, social media has also changed the conditions in which the Armed Forces carried out military operations. Research on social media has been conducted in different directions. What we are interested in is how to use this for computing situation awareness (SA) in military settings.

Social networking has been considered as a means for emergency communications, and social media data have been utilized in emergency SA applications like disaster relief or natural hazards events [81–83]. Even though similar technologies may be effective and ready to be adopted in SA in regards to military events, it is not enough (and potentially unsecure) to only rely on machine computations. This is especially true as in the military setting we have command and control operations. The military often faces critical situations and the consequences of interpreting information incorrectly can be serious. Therefore, we think Subject Matter Experts (SMEs) based human computation is necessary here.

In this study, we explain the insights of social media and characteristics of the data they generate, how they change information strategies and affect SA from different aspects, illustrate why fuzzy based solutions are appropriate in this setting, and analyze why and how to encompass human computation in this procedure. Specifically, we suggest involving human experts in three stages when computing SA, namely during data processing, information evaluation and decision making. Our goal is to understand the relationship between SA and numerous individual social media users, and to explore a better way to support safety-critical SA, a socially cognitive understanding of “what is going on”.

2.5.1 Social Computing in Situation Awareness

Before discussing the importance of utilizing human processing power to solve SA challenges using social media data, it is necessary to study the characteristics of social media and the advantages and disadvantages it can bring to military SA.

2.5.1.1 Social Media and Related Strategic Change

It is obvious that with the advances of computer information and the development of mobile technology, social media and the data they generated have become part of our daily life. This not only means the speed and transparency of information have dramatically increased, but also implies the emergence of new information sources, as anyone is able to share his/her thoughts online without having to expose himself/herself.

There are several essential characteristics of social media we cannot ignore when considering it as the information generating and propagating platform:

- *Uncertainty*. Blogs, social networking web-sites and media-sharing sites allow users to share information anywhere and anytime. These high-volume social data sites can be noisy, vague, imprecise, and unclear, considering the barriers such as language and culture. Extracting useful information from these data sites are the keys for SA, yet it is also very challenging. In addition, the uncertainty feature of social data also includes the ambiguity of a user's true intention. This means the assumption that all information shared on social media implies the users' true intentions and that the attitude is not always correct, especially when the cost of generating and propagating information is so low on social networks.

- *Connectivity*. The relation among users within a group can be an important factor for information transferring. How to group or cluster users can also contribute to understanding the "big picture". But it is worth noticing that functional similarity or spatial closeness may not be enough for this process. Besides, one user may belong to multiple groups/clusters, and information sharing within groups can be highly event dependent. For example, one may forward an interesting picture or news to everyone in his/her contact list but choose to share his/her vacation photos with only his/her friends group. Therefore, the connectivity of a user and his/her groups should not be treated as a constant attribute. In addition, whether we should separately compute an individual user and his/her group and how much weight we should give are other aspects which influence SA and final decision making. Specifically, when a group of users has similar information, it is hard to decide whether we should only treat them as a whole unit since the simple copying of information can happen, or should we count it multiple times? Copy detection can be included here, although it is hard to know a user's true intentions when similar data appear in the group.

- *Collaboration.* The way users collaborate in social media offers another impact aspect for SA. One can positively support others in a discussion, or express negative objections on some topics. However, collaboration may not always be binary. Ambiguous information and unclear user attitudes are very common.

- *Inequality.* Similar to how information from authorities is more reliable compared with gossip newspapers, information from different resources/users in social media should not be treated the same. Thus, their contributions to the SA computing are not the same. However, although it may be easy to differentiate personal/individual social media users from business or social identities, it is challenging to evaluate reliability among several personal users or several business users. For example, information from Google or Amazon may be more convincing than that from a gun-shop manager; but how to compare information from a gun-shop manager and a banker is hard to decide. Some user attributes such as reputation, account created time and geolocation may be used to estimate the reliability of a user and therefore the value of his/her shared information, but those attributes themselves can be dynamic or not true by the nature of Internet.

- *Time-sensitivity.* The information from social media is highly time sensitive, which brings another important variable when understanding the environment. Depending on the social media type, some information is near-real-time, e.g. texts on Twitter, while others may be less time dependent, e.g. videos on YouTube. Therefore, if emergency SA such as natural disaster control is the application's goal, real-time data should be the focus. On the other hand, if a long-term strategic decision is the goal, we will need more data from a long-time period. In addition, information and a user's attitude can change too. For example, a user may support legalizing undocumented immigrants, but after a terror event, his/her option may change.

As technology and the Internet have changed the nature of conflicts, it is essential for the Armed Forces to observe and adjust the information strategies, which implies that the effective use of social media can be very helpful to understand the environment from the military point of view. How to utilize social media data effectively while avoiding overuse and basing decisions on misleading information is still a research question that has not been answered. Understanding the features of social media can assist us in extracting useful information and further understanding a situation.

2.5.1.2 Situation Awareness

SA seeks the perception and understanding of the "big picture" in an event with many inputs and dynamically changing variables. This is especially important for the Armed Forces who require command and control operations and often face critical situations. In the military context, the SA refers to the "product of applying analysis and judgment to the unit's SA to determine the relationships of the factors present and form logical conclusions concerning threats to the force or mission accomplishment, opportunities for mission accomplishment, and gaps in information" [15].

SA normally involves many actors/inputs. This becomes more complicated when social media is considered as a data input platform. According to Endsley's model, SA has three levels: perception of elements in the environment; comprehension of the current situation; and projection of future status [23]. We believe social media can bring great positive contributions for the first level. Social media users can be treated as human sensors which are similar as physical sensors to stimuli [84]. When an event occurs, human sensors respond by making posts and sharing information over social media like Facebook and Twitter. Below, we analyze the positive impacts social media can bring to SA.

- *Effective and real time.* With the development of mobile devices, network services and related social media applications, social media users/human sensors can provide real-time information on an event. In addition, information propagation through social media is also very quick. Thus, social media data can be very useful to provide up-to-date information when we need to know the current situation. This is also the reason social media data have been used for emergency SA in different areas, such as [81–83].

- *Informative.* As the social media data come from real human users with any background, the information can be on any topic and cover multiple knowledge domains. Besides, built-in devices and sensors such as GPS and cameras can also provide additional information like photos and videos. Social media data, when users share, post and process, may be synthetic, subjective/objective, or generative or derivative data. So, the advantage of social media data compared with data from traditional resources is obvious: it is all-purpose.

- *Cost effective.* Using humans as information sensors has little or no cost. Research on rewarding strategies to encourage and evaluate user contributions online have been an active direction in the social media realm. Many effective algorithms and techniques are proposed to stimulate social media users. Thus, compared with building physical sensors, the cost of human sensors is very low, especially when considering the performance and results they can generate.

- *Scalable and wide-ranging.* Social media users can be anywhere if the network service is accessible. This geographically distributed feature makes social media data highly scalable and can cover a wide range of locations. Thus, social media data have been utilized to compute disaster SA [85] and military and aviation operations [86]

Even though the advantages of using social media and human sensors are attractive and promising, they cannot replace physical sensors or recognized military agents. The reason is obvious: human sensors/social media users can be unreliable and pose potential security issues.

Unreliable or unsecure data found on social media can be divided into two categories, namely unintentionally wrong information and vicious information. The former one means a user propagates the information unintentionally or without truly understanding it; while the latter means, the user supports the opinion of the data and spreads the information purposely. The effect of these two kinds of data on SA and related decision making is different. How to separate those two kinds of data is still a remaining unresolved question. Depending on the content of the unsecured data, they may bring false damage or hazards for society. Therefore, identifying unreliable information, whether intentional, exaggerated or accidental is very necessary.

2.5.1.3 Fuzzy-Based Intelligence System

Today's military operations are supported by access to information from a myriad of sources that provide overwhelming amounts of data. Indeed, a primary challenge nowadays is not a lack of information, but instead the information overload [16]. Much of the data may be non-verbal in nature, such as numerical troop strength reports, GPS coordinates of important potential targets, video feeds from unmanned aerial vehicles, readings from ground-based acoustic sensors, and the like. However, much of it, including that from social media platforms, is linguistic in nature. Further, we may want to describe this data with additional linguistic attributes, such as “possibly true”, “true”, “improbable”, “usually reliable”, and so on. Also, as outlined above, data collected from social media may be uncertain, imprecise, incomplete, vague, and/or unequal in value.

Interacting with humans requires systems to address linguistic words and phrases; this is especially true in human computation systems. This situation leads naturally to the use of fuzzy logic in the construction of such systems. Fuzzy set theory, formally defined by Lotfi Zadeh in his seminal 1965 paper [28], provides the necessary multivalued framework for designing fuzzy inference systems. Fuzzy inference provides the ability to perform approximate, or fuzzy, reasoning. As defined in [87], approximate reasoning is the “process or processes by which a possibly imprecise conclusion is deduced from a collection of imprecise statements”.

A key concept in fuzzy inference is that of linguistic variables. A linguistic variable is primarily used to approximately characterize the values of variables as well as their relationships [36]. In some cases, the vagueness conveyed by using words may be on purpose since the situation may not require more precision. However, in many cases, the imprecision is an artifact of the lack of ability to be able to quantitatively describe the characteristics of an object [36]. In addition to providing the framework for fuzzy inference, fuzzy systems have been shown to be effective at approximate reasoning where information is uncertain, incomplete, imprecise, and/or vague [25, 26, 28, 88].

The literature contains numerous examples of the application of fuzzy-based system to the exploration of social media data. It is not within the scope of this paper to provide an exhaustive survey of such examples, but a few will be mentioned here to corroborate the applicability of fuzzy systems to the problem domain that is being considered.

The work presented in [89] proposes a fuzzy algorithm to mine multi-layered Twitter data. The goal is to generate indications of general public opinion by examining tweets. The results shown demonstrate that the method is both fast and high in predictive accuracy. A typical use of fuzzy inference is provided in [90]. This research uses an adaptive fuzzy inference method to understand the sentiment contained in social media data for both English and non-English languages.

The idea of clustering – whether it is applied to posts, documents, tweets, or users – is an important idea with respect to social data mining. As such, there are many examples in the literature of techniques that use fuzzy-based clustering methodologies; two examples are mentioned here. Fuzzy clustering is used in [91] to compute emotional and sentimental values from micro-blogs posted by mobile users. The use of a modified fuzzy C-means algorithm is presented in [92]. The fuzzy system is used to cluster user sessions to enable analysis based on the behavior of the users.

2.5.2 The Role of Human Processing

Social behaviors such as online discussion, collaborative projects and creative activities are very common in social media. The information generated in those social behaviors is not predetermined and primarily directed by participants' inspiration. Thus, using the social information to compute SA and make decisions is not an easy task, especially in the military setting where critical and emergency events often happen. Therefore, we believe human computation, a technique utilizing human processing power to solve problems that computers cannot yet solve [39], can be a good solution for this problem. Particularly, we suggest involving SMEs instead of general human users without military backgrounds and knowledge in this process. We suggest including human processing in three stages during the SA computation based on social media data, as illustrated in Figure 1.1.

After collecting raw social media data, we then need to start data processing such as information retrieval and extraction. Information retrieval searches for the relevant information contained in a set of knowledges, while information extraction is focused on extracting relevant facts. Both of those procedures largely rely on pre-defined keywords, which can impact the result of the whole system. How to decide on keywords is a task that cannot be accomplished by simple machine computation. Therefore, human involvement is mandatory and essential here. We suggest including a human computation step to determine the project goals before data processing and after data collection. For example, suppose we would like to study how the new refugee policy is propagated over the social media and find the public's opinion on both the east and west coasts of the U.S. Then we need to ask some SMEs to provide a list of important keywords such as "refugee(s)", "asylee(s)", "resettlement", "humanitarian protection", and so on.

The data processing step, shown as the dotted rectangle in Figure 1.1, is the task mainly for machine computation. This step may encompass information regarding user profiles and attributes, depending on the "situation" which we are trying to understand. For example, if we would like to research on how the new gun law and policy affects manufactures, those users who are handgun sellers or gun shop mangers may be our focus. It is also worth noticing that in this step, information retrieval and extraction may happen multiple times, as shown in the figure. This is because when a new event happens, we may need to update the keywords and research focus, and then process the media data again.

After information retrieval and extraction, we then compute the value of information (VoI). This is called the information evaluation step in our architecture. User profiles and attributes can have an important contribution in this step too, as demonstrated in Figure 1.1. Considering the inequality and connectivity features of social media, we evaluate the VoI from each relevant user or user group based on his/her/their user information such as roles, importance for this event, reliability, reputation and so on. Different methods to compute the VoI have been proposed already; one such fuzzy-based system is presented in [93].

When data are evaluated and values are given to the relevant information, we can then cluster them into groups based on opinions or facts and rank information in each cluster for computing global SA. The clustering methods like [91, 92] can be adopted here. In this step, we suggest involving human computation to optimize results. Furthermore, we think a group instead of an individual-level SA process is necessary here, similar to that presented in [86]. Specifically, a group of SMEs who are related to the event are given the results of ranked data and related user information, and then they need to comprehend the environment individually. The SMEs' knowledge and experiences are significant for computing global SA, especially on new and emergency events, which is not reproducible by computers alone. For example, social media users can report and discuss the fact that large amounts of illegal immigrations appear in some European countries. Regular users may just interpret this as an impact factor for the local job market. SMEs might understand this is a potential security problem which may also include diplomatic and cultural conflict and then make corresponding reactions and decisions. Note that we suggest each SME evaluate the environment individually. This is because each expert has his/her opinion on understanding of “what is going on”, which is decided by his/her own experiences. Then the final SA result can be generated based on the aggregated evaluations.

After seeing the big picture of the current situation, we then make decisions which project future actions in the environment. In this step, human experts should be involved too. Decision making implies “what we are going to do”, based on knowing “what is going on”. Possible errors always exist and no SA system can guarantee correctness all the time. Therefore, computer-aided decision making should only be used as a complementary module. Human experts must be the ones to determine the real actions, especially in the military setting where consequences of a decision can be serious.

2.6 Conclusions

The types and amount of information available to the military community have seen an unparalleled increase over the last several decades. As a result, military commanders and staff have the daunting task of not only separating the important information from the routine, but also the challenge of combining information that complements and contradicts the standing premise. Two approaches for codifying the contextual underpinnings (framework) and cognitive interpretation for capturing Value of Information (VoI) utilizing source reliability, information content and latency based on triangular and trapezoidal fuzzy membership functions. While both of these approaches for capturing VoI are intuitively simple to comprehend and computationally easy to calculate, differences are observed. Both approaches have their own strengths and weaknesses.

Moreover, an approach to determining how analysts combine complementary and/or contradictory information in order to achieve the necessary situational understanding to make sound decisions has been examined. The experimental knowledge elicitation exercise tested 80 different combinations of "new" and "old" information. The Initial Facts ranged from information classified as extremely valuable [A,2] to information of little value [E,4]. The Additional Facts offered varying degrees of complementing and contradictory information to the original fact from totally complementary to totally contradictory, and were graded either better, the same, or worse in informational value than the original fact. The approach of "backwards solving" to update Information Content does not work since there are incorrect or missing roles. We will work with SMEs in U.S. Army to develop other approach to reason integrated VoI.

We have also discussed how to inquire environment understanding when considering social media as the data input platform. Our emphasis is set on military related events and requirements, based on which we analyze the features of social media data, the relationship between social media and situation awareness, and the necessity of including Subject Matter Experts (SMEs) during this process. We have demonstrated how to involve human computation at three different stages in the SA architecture, and also explained that adopting fuzzy logic technologies for knowledge discovery is suitable for our focus setting. Even though it is clear that including SMEs benefits SA computing, it is also an expensive procedure which is hard to evaluate. This means we not only need to think about how to motivate subject experts when bringing them extra works, but also how to aggregate a group of experts' decisions, as human skills and experiences are hard to estimate.

Chapter 3

Assistive Healthcare System for Young Adults with Autism Spectrum Disorder

3.1 Introduction

Autism spectrum disorder (ASD) and autism are both general terms for a group of complex disorders of brain development. These disorders are characterized, in varying degrees, by difficulties in social interaction, verbal and nonverbal communication skills, and repetitive behaviors [94]. U.S. Centers for Disease Control and Prevention (CDC) have published autism statistics which show that around 1 in 68 American children are on the ASD – a ten-fold increase in prevalence in the last 40 years. Over 3 million individuals in the U.S. and tens of millions worldwide are affected by ASD and these numbers are fast-growing.

Unfortunately, there was no medical detection or cure for autism till now. Some individuals with ASD may have to deal with their symptoms lifelong even after treatment and interventions, and autism can cost a family $ 60,000 a year on average [95]. In addition to social impairment and communication issues, autistic individuals may face significant challenges from the following three aspects. First, they need considerably more time to acquire a skill to fluency and often have difficulties maintaining the skill [96,97]. Second, individuals with ASD often lack of executive functioning skills such as time management [98,99]. Third, they also have sensory-based challenges that can cause significant distraction, pain or negative sensory experiences that can result in debilitation or panic. Therefore, individuals with ASD normally need constant prompts and continuous support to assist their skill development and acquisition, as well as skill maintenance and generalization, even for basic living skills such as eating and cleaning [96,100]. Moreover, continuous interventions are needed to help them stay focused, even when completing familiar tasks. Because of the intensive level of prompts and personal attention needed by individuals with ASD, many cannot have an independent and self-conducted life. It not only increases stress and personal fatigue for the caregivers, but also brings growing social and financial burdens.

In the past few decades, information technology and various sensors are fast developed, which makes the Ambient Intelligent (AmI) System popular and also implemented extensively. Some researchers have proposed different methods and technologies to provide living assistance to people with disabilities such as autism [101–105]. However, most of them are for autistic children and they focus on improving their social, academic or organization skills. There is also research for other populations such as deaf children [106] and older adults [107], but it does not satisfy the requirements from targeted users.

In this chapter, an ambient intelligence based context aware assistive system is

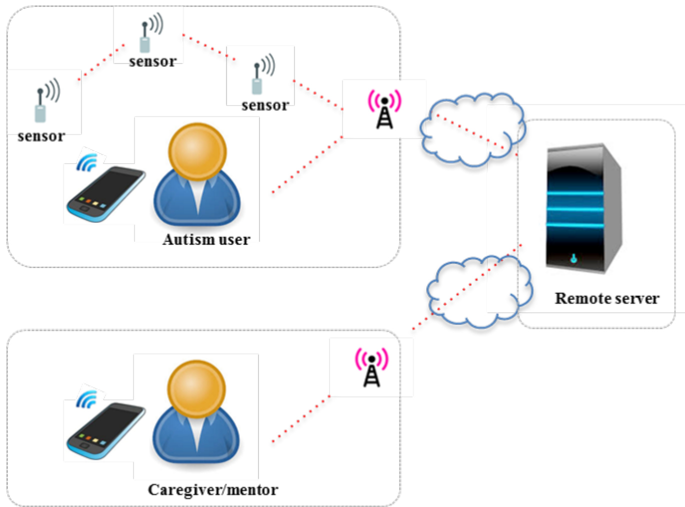

Figure 3.1: The proposed assistive system

proposed for young adults with ASD to improve their living skills, as demonstrated in Figure 3.1. The system uses various context information including location, time, schedule, personal preference, functionalities and so on, to automatically provide customized and real-time support for individuals with ASD in their daily activities. To provide a cost effective, flexible and easily accessible method, various low cost sensors are combined with mobile phones to build the system. Since previous research [96, 98] has identified that not many individuals with ASD have unique communication needs, the prompts are delivered in a variety of forms including text, image, audio, and video.

Moreover, the Bayesian classification technique [108] is employed to implement automated prompting functions in the system. An attractive feature about that technique is that it is relatively simple to implement. Researchers have shown that it performs well when used as a classification tool [109]. For computational convenience, most implementations of the Bayesian classification technique require the assumption of conditional independence of all the predictive attributes given the outcome variable. In this study, this assumption is relaxed in a selective manner to eliminate inappropriate conditional independence assumptions and thereby minimize the impact of such assumptions on the belief revision process. This is accomplished by grouping interrelated variables for the assessment of posterior probabilities. Sets of variables considered as part of the same group, which are called composite attributes, are used to partition the entire set of predictive attributes into disjointed sets.

Hopefully, this new intervention approach can assist autistic users to learn and maintain basic living skills so that they can improve their life quality, augment their ability to live an independent life and relieve the workload, stress and financial costs of society and autistic families.

The reminder of this chapter is organized as follows: Section 2 provides a comprehensive literature review related to ambient intelligent system, ASD, and naive bayes classification. Section 3 presents a novel context-aware support system that can help young adults with ASD. Section 4 introduces experiment of this research. Some initial results and discussions are also provided in this section. Section 5 concludes this chapter.

3.2 Literature Review

3.2.1 Ambient Intelligent System and Autism Spectrum Disorder

Using context information in ambient intelligent (AmI) systems to support people with cognitive disabilities has been an active research direction for the past few decades. Most of the existing research on Autism Spectrum Disorder (ASD) focuses on improving social and communication skills and targets children with ASD. For example, [101] discusses the development and evaluation of a tablet-based application to teach communication skills for children with ASD . [102] discusses how individuals with ASD desire to use technology to improve social, academic, and organization skills. In addition, most of the previous works are designed for the academic setting. For example, [103] studies how to use vSked, an interactive system, to improve student independence in classrooms and reduce prompts initiated by the educators. [104] proposes a mobile tool for autistic children to practice social skills. [105] develops a tablet-based application for activity schedules that addresses classroom routines and verbal communications.

There is research on supporting everyday activities and workplace tasks, but most of them are in the conceptual, exploratory stage without actual implementation or user involvement [110, 111]. Studies on using context information to provide services to people with cognitive disabilities have also been explored. [107]

describes a context aware device that uses a camera to assist people with dementia in independent living tasks such as using the bathroom. Similarly, [112] uses location-based information to provide prompts to individuals with cognitive disabilities for catering tasks. However, all these studies only explore one type of context information. [113] and [114] separately examine a much broader range of context information (location and motion information collected through sensors, visual information collected through cameras) for the purpose of providing automatic prompting for people with cognitive disabilities. However, their studies focus on workplace tasks and the use of cameras raises privacy concerns in the independent living environment, which can be very sensitive for our user population.

3.2.2 Naive Bayes Classifier

Bayes networks are one of the most widely used graphical models to represent and handle uncertain information [115, 116]. Two major components constitute Bayes networks:

1. A graphical component composed of a directed acyclic graph (DAG) where vertices represent events and edges are relations between events.

2. A numerical component consisting in a quantification of different links in the DAG by a conditional probability distribution of each node in the context of its parents.

Naive Bayes are very simple Bayes networks which are composed of DAGs with only one root node (called parent), representing the unobserved node, and several children, corresponding to observed nodes, with the strong assumption of independence among child nodes in the context of their parent [117].

To ensure the classification, it is critical to consider the parent node to be a hidden variable stating to which class each object in the testing set should belong. And child nodes represent different attributes specifying this object. Therefore, only the conditional probabilities should be computed when training the data set because of the unique structure. Once the network is quantified, any new object giving values of attributes can be classified by using the Bayes rule.

For example, let the set of attributes that are used for prediction be termed $X_1, X_2, \ldots, X_n$, and the outcome be A. If the new instance with attribute values $\underline{X} = \underline{x}$, then the posterior probability estimate for each possible outcome value a_i is $P(A = a_i \mid \underline{X} = \underline{x})$.

However, this simplest way to estimate probability could lead to enormous training data requirement when larger sets of attributes are considered, or attributes are allowed to have multiple values. Naive bayes classification allows a frequent assumption to solve this issue which is called conditional independence. This assumption means every attribute (X_1,...X_n) in the set are independent with others given outcome A=a_i.

Expressed by:

$$
\begin{aligned}
P(A = a_i \mid \underline{X} = \underline{x}) &= \frac{P(\underline{X} = \underline{x} \mid A = a_i) \cdot P(A = a_i)}{P(\underline{X} = \underline{x})} \\
&= \frac{P(A = a_i) \cdot \prod_i P(X_i = x_i \mid A = a_i)}{P(\underline{X} = \underline{x})}
\end{aligned}
\tag{3.2.1}
$$

where A is a possible value in the session class and $\underline{\mathrm{X}}$ is the total evidence on attributes nodes.

In the past decades, naive bayes classifier have been applied in various domains. There are many studies [118–120] have been done to implement naive bayes classifier in text data mining. Moreover, [121] has proposed a rapid assignment of rRNA sequences into the new bacterial taxonomy based on naive bayes classifier. And [122] developed an intelligent system using naive bayes classifier to predict heart disease.

In this study, we chose naive bayes classifier over other machine learning algorithms such as Support Vector Machine (SVM) or decision tree for two reasons. First, naive bayes classifier has very high computation efficient, which can provide real-time response to the system. Second, because of the conditional independence assumption, naive bayes classifier can remain high accuracy with small training data set.

3.3 Context-Aware Support System

Through the proposed context-aware supportive system, an easily accessible virtual caregiver/mentor which can assist their daily activities is provided to young adults with Autism Spectrum Disorder (ASD). They can improve their living skills progressively and learn how to live independently by using this proposed system. Moreover, this system can work anywhere, anytime, and gives personalized prompts for the special needs of each user. Besides, caregivers and mentors who take responsibility for taking care of individuals with ASD can also get benefits from the system. The application provides activity records and reduces workload, stress, and financial burden.

3.3.1 System Architecture

In general, this system follows the traditional client and server (C-S) network architecture, as shown in Figure 3.2. There is one dedicated central server which takes charge of data logging, context analysis, and decision-making. By analyzing the data collected from context sensors, the server can advise users to finish activities by sending guidance and suggestions.

The client in the proposed system refers to the individuals with ASD who are using the mobile application. With help from the application, users can perform general activities and living tasks. During this process, user operation data will be

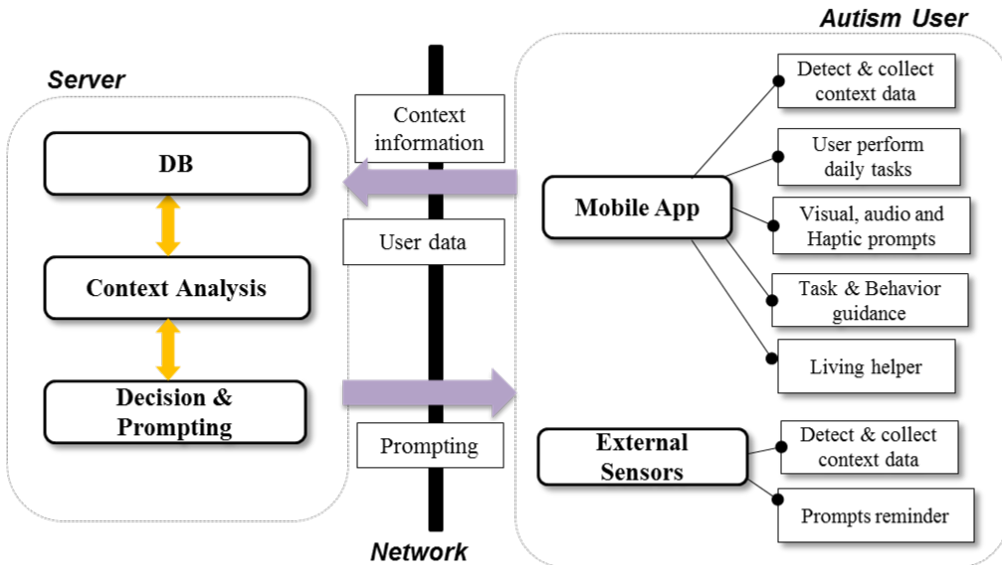

Figure 3.2: System architecture

captured and transfered to the server for situation awareness analysis. In addition, external sensors for location and event detecting are included to provide precise and detailed context information. Similarly, data from external sensors are also sent to the server automatically. Furthermore, application for caregivers and mentors is another client in the proposed system. They will be notified and get involved whenever users with ASD get stuck in a certain procedure or get lost in navigation.

The overall system comprises of five basic components, as shown in Figure 3.3. The task module represents user activities such as cooking, getting dressed and so on. This module highly relies on the database in the server. The user module is used to store user profile information, such as schedule, preferences, and functionality. This module can help the system make informed decisions for every specific user.

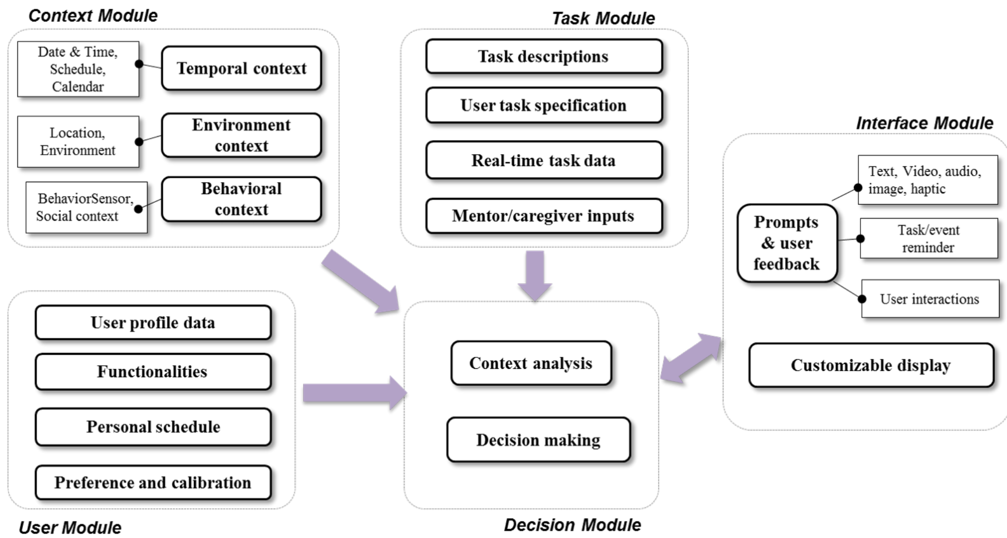

Figure 3.3: System modules

There are three critical context information types in the context module, which are temporal, environmental, and behavioral. These three kinds of input help the proposed system collect and utilize the context information to infer the situations of users. Interface modules in the proposed system provide prompts and instructions displays. Multiple output options through text, video, audio and image are also included in this module. The last module is decision module, incorporates inputs from the other modules and performs informed decisions to provide accurate and timely cues and appropriate prompts. The interaction amongst these modules enables and contributes towards the efficiency and precision of the proposed solution.

3.3.2 Mobile Application

We believe a smartphone based solution can address our challenges better, as it is easily accessible, cost effective and has many useful built-in sensors. Therefore, a mobile application, working as an intermediary between supportive system and end users, is included in the system. It serves as the user's agent and is in charge of detecting and collecting context information from different sensors. For example, when a user with a mobile device enters the kitchen, after communicating with the sensor at the kitchen's door the mobile application can detect this event and send the corresponding context data to the server. In addition, the application on a smartphone is the platform where users can follow instructions to perform actual daily activities. Figure 3.4 is a screenshot of mobile application on an Android smartphone and the task shown is cooking.

Figure 3.4: Mobile Application User Interface

3.3.3 Context Sensors

Location sensors are the key context sensors for independent living assistance such as a cooking assistant. Based on the location of the user and the data from other modules, decisions as to whether to send a prompt and what prompt should be sent can be made. Two kinds of sensors have been adapted in the system, namely infrared sensors and Bluetooth wireless sensors.

As one of the most popular infrared sensors, the Sharp GP2D12 [22] has been employed for location measuring in this system. The sensor uses an LED drive circuit with integrated signal processing and then outputs analog voltage. The effective range is 10 to 80cm while output is a voltage signal from 0.4V to 2.4V. In order to process and transfer these analog voltage signals, an Arduino board with a 32-bit ARM processor and six analog input channels is used. Moreover, with the help of a CC3000 Wi-Fi shield manufactured by Adafruit, the Arduino board could host a micro-server in wireless Ethernet and transfer real-time distance data via the telnet method. Figure 3.5 shows the infrared sensor for capturing location contexts.

In addition, Estimote beacon sticker [21], a small wireless sensor that broadcasts radio signals to communicate with smartphones, has also been used to get the location information. It includes a powerful 32-bit ARM processor, 256kB memory, and 2.4GHz Bluetooth Smart module. The broadcasting range could be as little as 2 inches and as far away as 230 feet. This kind of sensor follows the iBeacon protocol

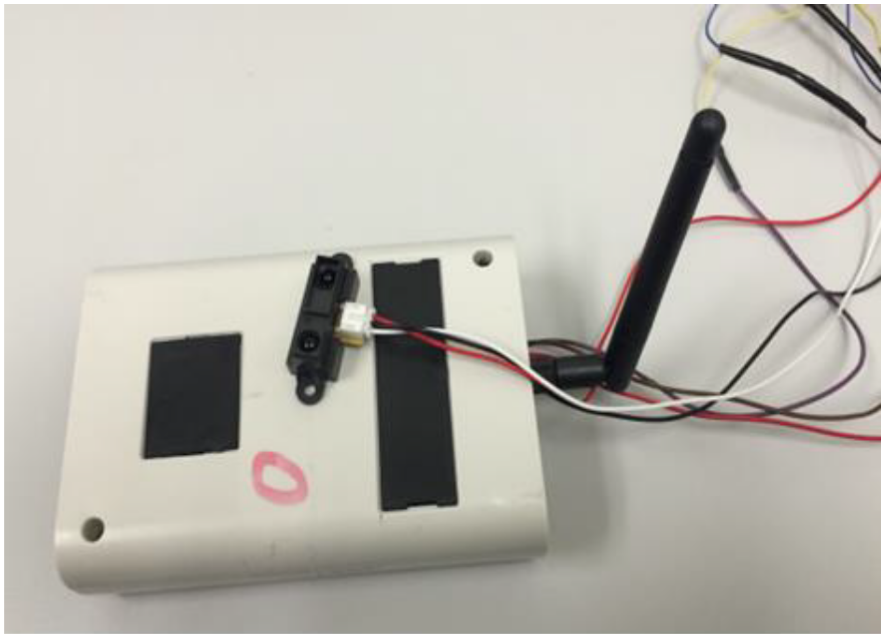

Figure 3.5: Infrared sensor system

and uses Bluetooth for data transmission. When a mobile device enters range of a beacon, it will receive various kinds of information including distance, which is broadcasted by the beacon. We put the beacon sticker, which has a unique ID, at major locations such as on the refrigerator, on the microwave, in front of a sink, on an entry door, and so on. When the user with the smartphone gets close to any sticker, his/her exact location can be tracked immediately. Figure 3.6 demonstrates a user with a smart phone that communicates with the beacon sensors.

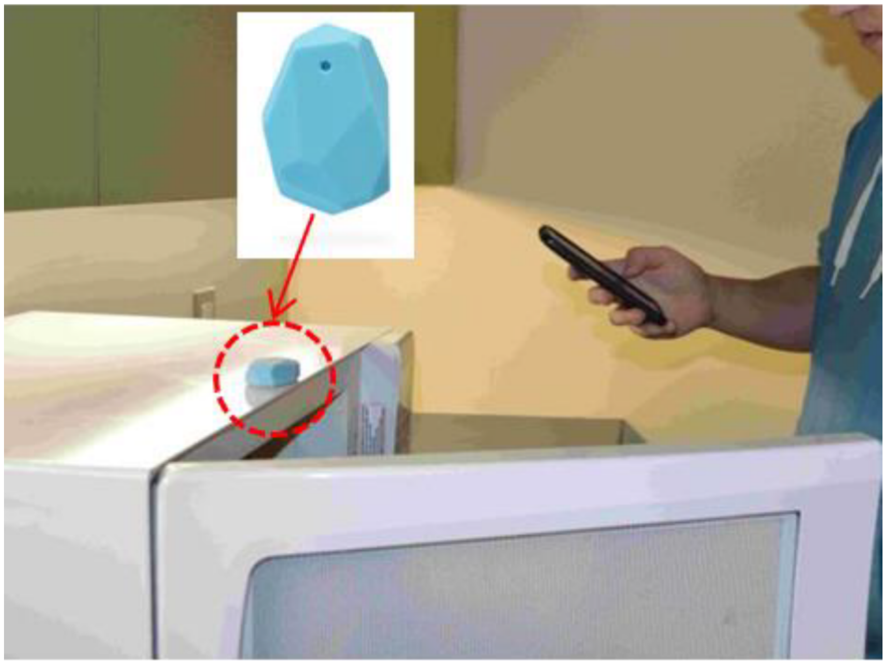

Figure 3.6: Bluetooth sensor system

Infrared sensors are much cheaper than Bluetooth sensors, but its range of coverage and data precision are not as good as the Bluetooth sensors. Infrared sensors can work independently without interaction with mobile devices, but the Bluetooth sensors should communicate with mobile devices to get the location data. So, choosing which location sensors is highly dependent on the application objectives.

3.4 Experiment

Two phases have been used in the user study: the first phase is to identify user needs and challenges for project design and development, and the second phase is a field study for user evaluation and feedback collection. Ten young adults with

Autism Spectrum Disorder (ASD) and their caregivers/mentors have participated in cooking tasks in an independent living setting and the user study is in two-sessions. The functionalities of these participants are at various levels and most of them do not have any cooking experience. Since the first session focuses on the user-guided learning functionalities and the second session requires the user with ASD to complete the same cooking task independently. During the second phase, a caregiver observes user and send correct prompting information when they are stuck in one procedure or get lost in navigation pages. Data is collected in the second session is used for decision module training.

3.4.1 Environment Setting

The proposed system server is based on PHP and implemented on a PC with the Debian Operating System. MySQL is employed for building the database to react to application requests and record experimental data. The system allows the user to access by both HTML and Android mobile APP. Two Samsung Galaxy S3 smartphones with Android 4.2.2 are provided to the participants to conduct cooking tasks.

As the main objective of the proposed system is to support independent living for users with ASD, experiments are conducted in a daily living setting which is a Life Skills Lab allocated in a special Autism Center at a public university. The Lab is a fully functional apartment with kitchen, living room, bathroom, dining room,

Figure 3.7: The Life Skills Lab for user study

and laundry room. Figure 3.7 displays the Life Skills Lab.

3.4.2 Dataset Attributes

From the experiments of ten participants with ASD, a dataset is collected in the MySQL database. There are 976 instances recorded in the dataset and each of them includes timestamp, user id, and page information. By performing preprocessing, four attributes are identified as parameters of training data:

- Gender (G): among ten participants, six of them are male and generate 583 instances out of 976. The other four are female users with ASD who generate 393 instances. Since some previous studies [123, 124] have shown that patients with ASD in different sex and gender have various symptoms and behaviours. Therefore, setting gender as one of the attributes for prediction is significative.

- Functionality (F): represents the living abilities and skills of a participant and the level of each user with ASD as defined by HPs. The functional levels of ten participants cover a wide range. Two of them are highly functional and can manage various daily routines independently. One participant has very low functionality, and needs constant assistance and reminders from mentor. The other seven participants are identified as medium functionality by HPs.

- Page type (Pt): means what kind of step of current instance. Three possible values are identified by HPs: browse, short, and long. Type browse means the user navigated to these pages and took a quick look. The staying time on these pages is normally less than 2 seconds. The second type is short step which defines some specific quick jobs, like taking out eggs from the refrigerator. Commonly these jobs take no more than 50 seconds but more than 3 seconds. The other pages, namely long, are time consuming, which take more than 50 seconds and even up to a couple minutes; for example, participants were waiting for eggs to boil.

- Page difficulty (Pd): reveals how possible it is that a user with ASD would be stuck in the current page and need help from HPs. This attribute is based on a statistic in regards to how many prompts of the current page were needed among all users. If a page was prompted less than 3 times in its history, it could be easy for most users and they are less likely to need help, namely low difficulty. However, when a page was prompted more than 5 times, it is extremely possible that the user needs assistance from HPs namely high difficulty. In the training dataset, 861 instances are low difficulty while 26 instances are high difficulty. The other 89 instance are identified as medium difficulty.

During the experiment, there is always a caregiver who is qualified as HP observing the participants finish cooking task. When a user gets stuck in a certain procedure or gets lost in the page navigation, the caregiver will send a prompt to help the user. The most critical evidence of judging a user need help is staying time on the current page. For example, if a user stays in a page which is getting eggs from the refrigerator whereas he/she keeps reading the description without any action, the user is likely to need a prompt. Therefore, our situation awareness (SA) model needs to calculate a reasonable staying time for a certain page. When users stay longer than the reasonable staying time, the system will send a prompt which simulates the action of a caregiver.

It is apparent that the calculation should consider the attributes identified by HPs above, which means the system will output a reasonable duration (D) of current page based on four input variables. In order to train the Naive Bayes classifier, we have done preprocessing to extract true values of duration. These values are acquired by calculating the number of standard divisions from the mean, where standard division and mean is based on values of staying time for the current page over all history data. To keep the calculation simple, we discrete the duration in two possible values: short and long; and this is the objective of classification. The threshold of discretization is set by HPs and it is adjustable for different user groups.

3.4.3 Situation Awareness Model

Based on the equation 3.2.1, then we can build the SA model and determine the probabilities associated with the different outcome states. To illustrate, we can consider a situation: a male user in high functionality is finishing a cooking task. The current step is a short step and very few people got stuck in this step. Therefore, we can use the model to calculate the possibility of both high duration and low duration given that the gender is male, functionality is high, page type is short, and page difficulty is low. This can be expressed by:

$$
\begin{aligned}
P(D = high \mid G = male \& F = high \& Pt = short \& Pd = low) \\
= \frac{1}{K} \cdot P(G = male \mid D = high) \\
\cdot P(F = high \mid D = high) \\
\cdot P(Pt = short \mid D = high) \\
\cdot P(Pd = low \mid D = high) \\
\cdot P(D = high),
\end{aligned}
\tag{3.4.1}
$$

and

$$
\begin{aligned}
P(D = low \mid G = male \& F = high \& Pt = short \& Pd = low) \\
= \frac{1}{K} \cdot P(G = male \mid D = low) \\
\cdot P(F = high \mid D = low) \\
\cdot P(Pt = short \mid D = low) \\
\cdot P(Pd = low \mid D = low) \\
\cdot P(D = low),
\end{aligned}
\tag{3.4.2}
$$

where

$$
\begin{aligned}
K = & P(G = male \mid D = high) \cdot P(F = high \mid D = high) \\
& \cdot P(Pt = short \mid D = high) \cdot P(Pd = low \mid D = high) \\
& \cdot P(D = high) \\
& + P(G = male \mid D = low) \cdot P(F = high \mid D = low) \\
& \cdot P(Pt = short \mid D = low) \cdot P(Pd = low \mid D = low) \\
& \cdot P(D = low).
\end{aligned}
\tag{3.4.3}
$$

All terms in the above expressions can be estimated with historical data.

3.4.4 Result and Discussion

In this study, Weka 3 [125] is selected to train the SA model based on data which is collected in the experiments. Weka provides a collection of machine learning algorithms for data mining tasks, which includes the Naive Bayes classifier. Another attractive feature provided by Weka is called cost/benefit analysis. This can help to achieve a certain threshold to minimize overall costs. Figure 3.8 is a screenshot of the Weka preprocess window which shows that a total of 976 instances are used for training the SA model. In order to validate the accuracy of our SA model, we employee 10-folds cross validation to calculate the confusion matrix. It should be noted that this matrix is based on a default threshold which is 50%. That means if the possibility of duration equals long given evidences is larger than 50%, this instance will be classified to the long group.

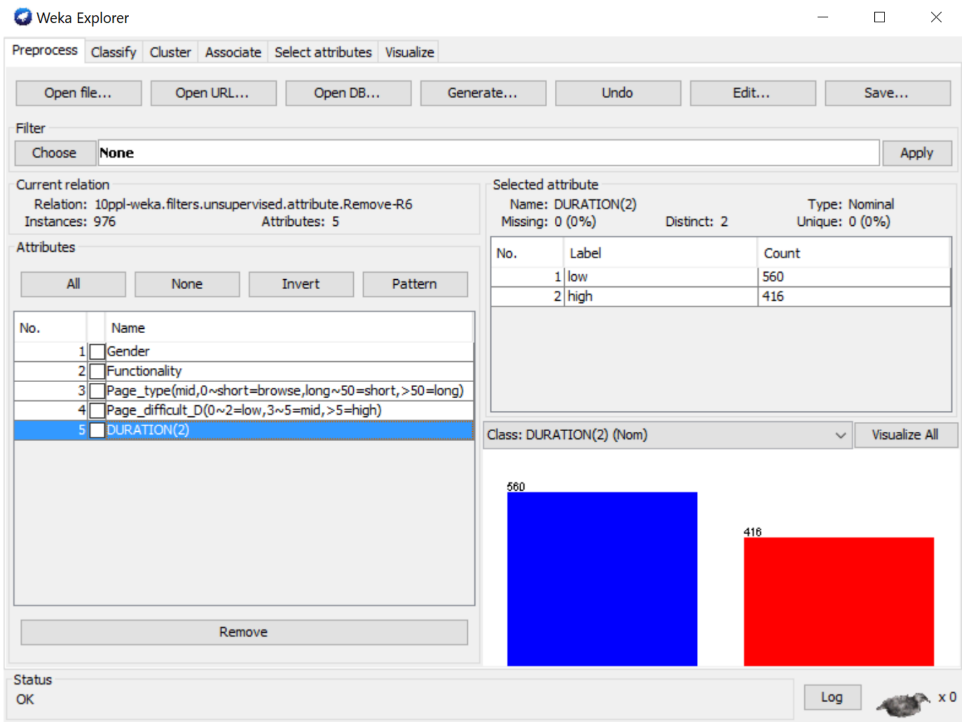

Figure 3.8: Weka 3 Preprocess Window

Figure 3.9 presents the detailed accuracy report of cross validation which shows that 57.68% of total instances are correctly classified. Thereinto, 448 instances belonging to low duration are classified correctly , which is 80%. While the other 20% instances belonging to low duration are classified to the incorrect class which is high duration. For true value of instances equal to high duration, only 115 instances (27.6%) are assigned to the correct class while the other 301 instances (72.4%) have the incorrect classification result.

Then we take advantage of cost/benefit analysis to tune the threshold so that we can achieve minimized overall costs. Figure 3.10 is a screenshot of the cost/benefit

```
=== Detailed Accuracy By Class ===

               TP Rate   FP Rate   Precision   Recall  F-Measure   ROC Area  Class
                 0.8       0.724      0.598     0.8       0.684      0.555    low
                 0.276     0.2        0.507     0.276     0.358      0.555    high
Weighted Avg.    0.577     0.5        0.559     0.577     0.545      0.555

=== Confusion Matrix ===

   a   b   <-- classified as
 448 112 |   a = low
 301 115 |   b = high
```

Figure 3.9: Detailed Accuracy Report

analysis result. It shows that when the threshold is set to 53.84%, global classification accuracy is 59.32% which is improved a little. It has the similar result with default threshold: 425 instances (75.89%) belonging to low duration are assigned to group low duration correctly while the other 135 instances (24.11%) are classified to the wrong class; 154 instances (37.02%) belonging to high duration are assigned to the correct class while the other 262 instances (62.98%) have incorrectly classification results.

There are two possible reasons behind this unsatisfied result. The first reason could be lacking data, since there is only one participant identified in low functionality by HPs and most instances generated by this user belong to high duration. However, when the system tries to calculate the possibility given evidences, there is little history data matched. As a result, lots of instances belonging to the high duration class are assigned to false group thus decreasing the accuracy of the system. To solve this problem, we will involve more participants into our user study and capture more data. These data will help to train a Naive Bayes model with higher accuracy.

The other possible reason is that the attributes identified for prediction are not conditional independent, which violates the assumption of Naive Bayes classifier and leads to unsatisfied result. For example, gender and functionality maybe have internal relationships which decrease the accuracy of our model. Therefore, some existing approaches that combine relevant attributes for possibility calculation could be involved in our system. It is noted that these approaches also require more data from a user study to achieve high accuracy model.

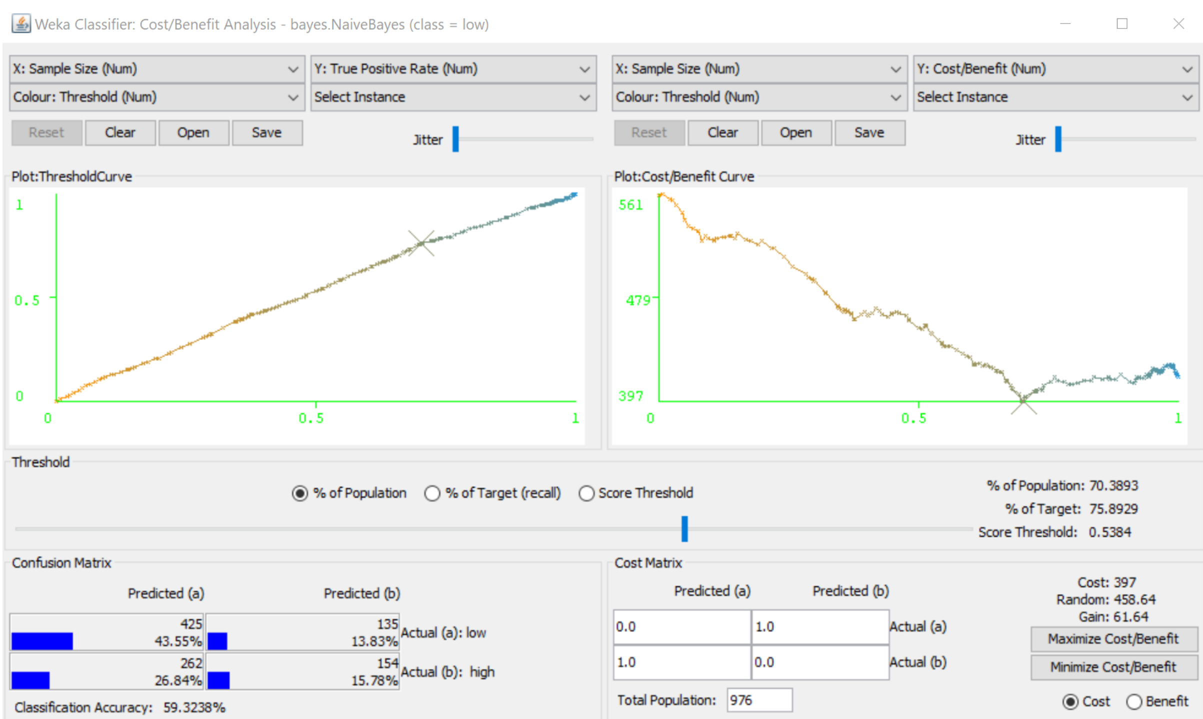

Figure 3.10: Cost/Benefit Analysis

3.5 Conclusion

This chapter has presented an ambient intelligence based context aware assistive system for young adults with Autism Spectrum Disorder (ASD). The proposed system relies on client-server architecture and provides a smartphone application for users to interact. Two kinds of low-cost context sensors are included in the system to collect indoor location data. Besides, we have involved ten participants belonging to various functionalities to finish cooking tasks in a public laboratory. During the experiments, a bunch of data are collected and HPs identified four critical attributes to predict reasonable staying time. With these training data, we build up a Naive Bayes classifier based situation awareness model to help HPs decrease their workload and working stress.

Since our current experiments can only provide very limited data to improve the accuracy of situation awareness model, the next step of the research is to extend our user study and involve more users with ASD.

Chapter 4

Exploring Assistive Healthcare System for Young Children with Spinal Muscular Atrophy

4.1 Introduction

Spinal Muscular Atrophy (SMA) is an autosomal recessive neurodegenerative disease characterized by the degeneration of spinal cord motor neurons, atrophy of skeletal muscles, and generalized weakness [126]. Most of the nerve cells that control muscles are located in the spinal cord, which accounts for the word spinal in the name of the disease. SMA is muscular because its primary effect is on muscles, which do not receive signals from these nerve cells. Atrophy is the medical term for getting smaller, which is what generally happens to muscles when they are not active [127]. The incidence of SMA is about one in every 6000 to 10,000 livebirths with a carrier frequency of one in 50 [128, 129]. Nowadays, SMA affects more than 750,000 patients worldwide and it is the number 1 genetic killer of infants and toddlers [130]. Unfortunately, there was no effective medical treatment until Nusinersen was approved by the FDA in December 2016 [131]. However, the cost of this treatment is humongous: the first year of treatment will cost $750,000 and $375,000 per year after that [132]. Moreover, previous research has shown that the median total expenditures for patients with SMA were $83,652 over an average of 6.9 years and the median cost for patients with early diagnosis was $167,921 [133].

Individuals with SMA and their families have significant medical expenditures.

Previous research [134] has shown that approximately 80 percent of individuals with SMA fall into type I, which is the most severe category. Type I SMA is also called Werdnig-Hoffmann disease; it leads infants to severe muscle weakness before 6 months of age. These young children never achieve the ability to sit independently when placed on their own; therefore they are completely dependent on others for all aspects of their care and participation. Furthermore, most of them have extremely limited muscle functionality, patients may only use several fingers with very little pressure. As a result, they cannot use regular interactive devices such as a mouse or a keyboard. Since most patients with type I SMA cannot walk or talk, they may face significant challenges from the following three aspects. First, they have difficulty efficiently communicating with others, so it is hard to present their opinions and ask for help. Second, individuals with SMA cannot control the environment around them, having problems running on the light in living room or turning off the television. The third challenge is that children with type I SMA have limited resources in regards to learning more about the world, they cannot interact with computer systems effectively. Therefore, solving these challenges is very critical and helpful. Since young children desire to interact and engage with the world, it is a crucial time for intelligence development. How to provide communication assistance becomes a fairly vital research question for this special user group.

With the rapid development of information technology and various sensors, research on living assistance systems and communication systems have became very active in the past few decades. [135] has proposed an assistive system to improve independence for people with autism spectrum disorder and [101, 102] have done some research on improving social, academic or organization skills. Furthermore, there is also research for other populations such as deaf children [106] and older adults [107]. Although different methods and technologies have been proposed, none of them can be applied for our targeted users since the targeted users of previous research have strong muscular functionality. Young children diagnosed with type I SMA require special switches and sensors that they can interact with due to their disabilities. And this study can be extended to help other populations such as people with regular muscular atrophy and amyotrophic lateral sclerosis.

In this study, we propose an assistive healthcare system to improve targeted users' quality of life. We present an multi-modal interaction approach which employs multiple interactive sensors to improve the interactive ability of children with SMA. Moreover, we design the architecture of the assistive system and implement partial functions in a web-based application. Our goal is to utilize the system to achieve highly efficient communication, environment control, and decrease users' living dependence.

The rest of this chapter is organized as follow: section 2 introduces the background and related work in the field of assistive systems for users with disability; section 3 presents our multiple-modal interaction system. A brief pilot study is also involved in this section. Section 4 introduces the architecture of the system and a system prototype. Section 5 concludes this chapter.

4.2 Literature Review

Existing research on young children with Spinal Muscular Atrophy (SMA) is quite limited. No directly related work was found from the assistive technology and accessible computing field that examines how to use information technology to help these children communicate with other people or control their environment. Previous research on SMA [136] has shown that type I SMA will cause the most serious weakness on movement and language ability; however, individuals have intact attentiveness and intellect. There are some studies that have examined multiple systems for users with variable disabilities which are similar to SMA.

Switch-based devices are effective and easy to use for users to input with. Typically, switches are mechanically actuated, may come in a variety of sizes and materials, and may be placed in various locations including near the hands, elbows, shoulders, feet, or head [137]. Most researchers use the simplest switches that support only one function and can be mapped to a single keyboard key when connected to a PC. [138] introduces a one-key system with a simple switch and [139] presents a customized system with wireless embedded switches for teenagers with Duchenne Muscular Dystrophy. Another input category is accessible touch and gesture. [140] explores touch screen use and tapping performance for people with motor impairments. The usability of multi-touch gestures on a smartphone touch screen is evaluated in [141]. Besides these two input groups, pressure-sensitive input can also be used. [142] applies pressure input pad to power wheelchair armrests, which enables multiple inputs and force feedback joystick control for power wheelchairs is studied by [143].

With the rapid development of information technology, some research turns to non-contact inputs, like eye-tracking and brainwave measurement. Eye-tracking systems can measure the point of a gaze and the motion of the eye ball relative to the head; brainwave measurements monitors the electrical activities of the brain and predicts the intention of users. [144] introduces eye tracking methodology and [145] presents eye tracking in human-computer interaction and usability research.

However, this previous research cannot be applied to our targeted users because type I SMA patients possessed significantly low physical capabilities. Most switches are too strong for them to push down and they cannot even keep their hands on switches. A gesture touchpad requires precise control of hand movements, which is impossible for our targeted users. And there is no directly related research to assess the usability of pressure-sensitive sensors for type I SMA patients. Moreover, since eye tracking and brainwave measurement techniques require the users to be in a specific position, for example in front of an eye tracker, which means the applicability is extremely limited.

4.3 Multi-Modal Interaction Approach

The focus of our research is to explore how to provide an accessible and responsive interactive way to allow young children to easily communicate and control their environment. In addition, considering the technology background of our user population, our solution should be customizable and usable, which means it should not take more than a few minutes to set up and it can also easily be understood and modified by families who are not technically trained. In the following sections, we explain our approaches and findings in detail.

4.3.1 User Requirements

The targeted user population of our project is very young children, approximately 18 months-5 years, with type I Spinal Muscular Atrophy (SMA). Real-time feedback is the first priority requirement since children of this age can be impatient and easily frustrated. Delayed response from the system may cause emotional reactions and decreased usability. Well-designed user interface is the second requirement.

Young children are easily attracted by their favorite character, such as super heroes and adorable animals. Moreover, subdued color can relax children and they will adapt to the system quickly. The third requirement is that it must be low-cost and customizable. To be able to widely support our targeted users, we require the proposed system employs inexpensive sensors and application platforms to achieve our economic goal. And every single user is a special individual, which means user interface should be customizable and multiple design solutions should be provided.

4.3.2 Wearable Interactive Device

As explained before, we aim to develop low-cost, customizable, wearable interactive devices which can be used by individuals affected by neurodegenerative conditions like SMA to control their communication devices and their environments. Considering the physical conditions of our users and the needs for daily usage, we believe the wearable interactive device should have the following unique characteristics:

- Customizable (i.e., can be modified/adapted for others with SMA and similar disorders)
- Affordable/accessible to people in the SMA community
- Usable (does not take more than 5 minutes to set up and can easily be understood and modified by families who are not technically trained)
- Durable

- Responsive/sensitive
- Accurate
- Allows for maximal functional movement patterns (i.e., palms of hands are not obstructed so they can still be used for play)
- Position independent. Rather than having to craft additional custom positioning solutions which bring the user to the input device, we bring the device to the user in a position which is comfortable and safe for long term regular use
- Improves quality of life and meets the needs of very young children (approximately 18 months-5 years), for whom very few therapeutic products currently exist

In order to manage and connect sensors and switches, a Feather 32u4 single board microcontroller is employed, which is manufactured by Adafruit. This board is Arduino-compatible and integrated with Bluetooth Low Energy (BLE). Arduino is an open-source electronics platform based on both its hardware and software. We can use the Arduino board and software library to read various analog or digital inputs from switches and sensors. BLE is a new low-power, 2.4GHz spectrum wireless protocol and can be compatible with iOS, Android, Mac OS X and Windows 8+. The core chip of this board is ATmega32u4 clocking at 8MHz and at 3.3V logic. With 32K of flash and 2K of RAM, it can act like a mouse or keyboard easily and effectively. Moreover, micro USB and battery charging ports are both provided in this board, and it is extremely convenient for portable projects. In this system, a 3.7V Lithium polymer battery is used as the power supply.

To connect sensors and switches with the microcontroller, a 1/4 size breadboard and multiple 10K ohm pull-down resistors are used for the circuit. There are two channels designed in the breadboard for various switches and sensors so they can work synchronously.

The sensors we consider to adapt to use in the interactive device need to be not only sensitive enough to capture tiny finger/hand movements, but also comfortable, light-weight and portable, as it is designed for long term regular use. After careful research, we propose to include the following four types of sensor/switch.

4.3.2.1 Micro Light Switch

This kind of switch allows the user to trigger it with a feather light touch. For example, the one manufactured by AbleNet only requires 0.4 oz (10 g) of force, which meets the requirement for our targeted users. With the feedback of auditory click and tactile, users can receive instant interaction results. Furthermore, the material of this switch is ABS, which is skin-friendly and provides easy mounting. Figure 4.1 is an example of a micro light switch used in our system. The box shown in the figure is an adaptor which converts an analog signal to a digital signal and connects to a computer via a USB. The switches can be mounted on wheelchair, standing frame, and even bed. These kinds of switches are highly effective for users who still maintain some functions on fingers. Moreover, these micro light switches can be connected to microcontrollers and paired to any devices that support Bluetooth.

Figure 4.1: Micro Light Switches

4.3.2.2 Flex Sensor

This sensor is also manufactured by Adafruit, which can detect flexing or bending in one direction. They were popularized by being used in the Nintendo PowerGlove as a gaming interface. Basically, this sensor is a resistor that changes value based on how much it is flexed. The resistance is about 25K ohm when it is unflexed, while the number will rise to about 100K ohm when flexed all the way. Both directional bendings can be traced and will produce electric signals. Figure 4.2 shows the sensor which is wired with a microcontroller. Two different sizes of sensors can be used to adapt to users of different ages and in various joints. These kinds of sensors can be used for users who have functionality on their wrist, ankle, or elbow.

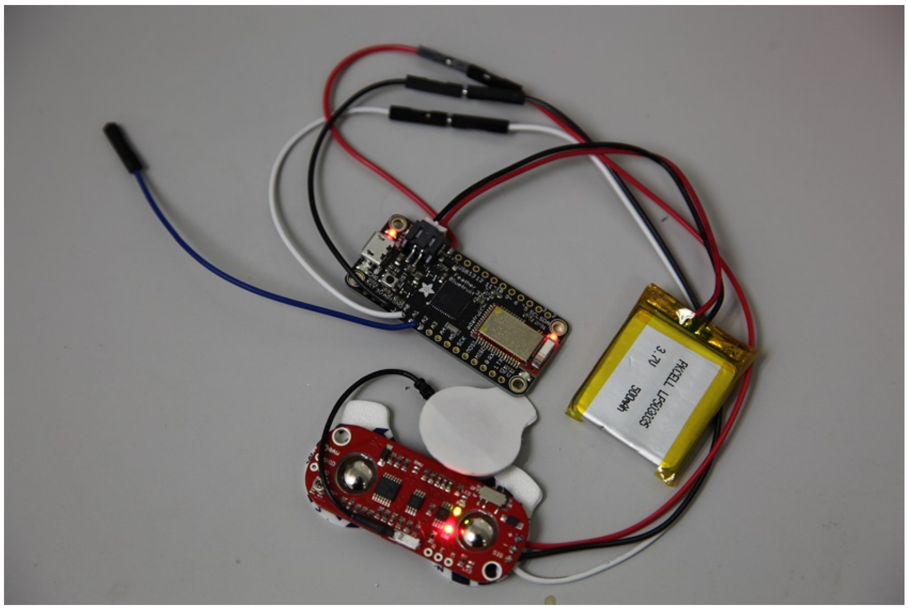

Figure 4.2: Flex Sensor

4.3.2.3 EMG-Based Muscle Sensor

MyoWare Muscle Sensor is used with Muscle Sensor Surface EMG Electrodes, which can track electromyograph and reasoning muscle contraction. Electromyography is an electrodiagnostic medical technique used for evaluating and recording the electrical activity produced by skeletal muscles. With this sensor, users can use muscle contractions as inputs to control the proposed system. If a user with SMA can contract muscles on thigh or sup brows, an EMG sensor can capture these tiny movement. Figure 4.3 shows the sensor connected with a microcontroller and battery.

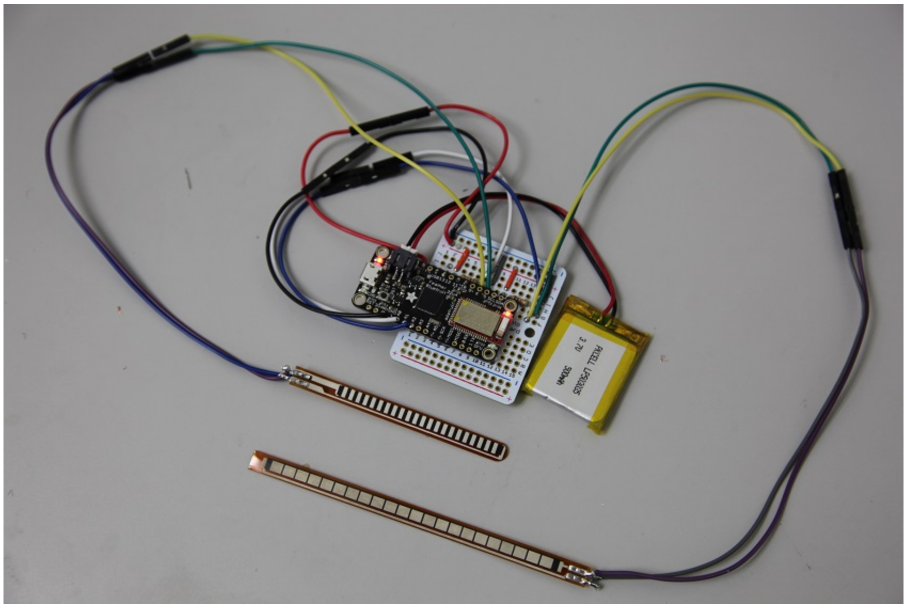

Figure 4.3: EMG Based Muscle Sensor

4.3.2.4 Stretch Sensor

This sensor is based on conductive yarn material which can conduct electricity. The conductive yarn spun from a stainless-steel polyester fiber blends and provides extremely pliable and soft touch. It also performs like an alternative resistor and the resistance will change when the yarn is stretched. In this study, conductive yarn is woven as a finger glove and the user can bend fingers to produce electric signals. Figure 4.4 shows the sensor.

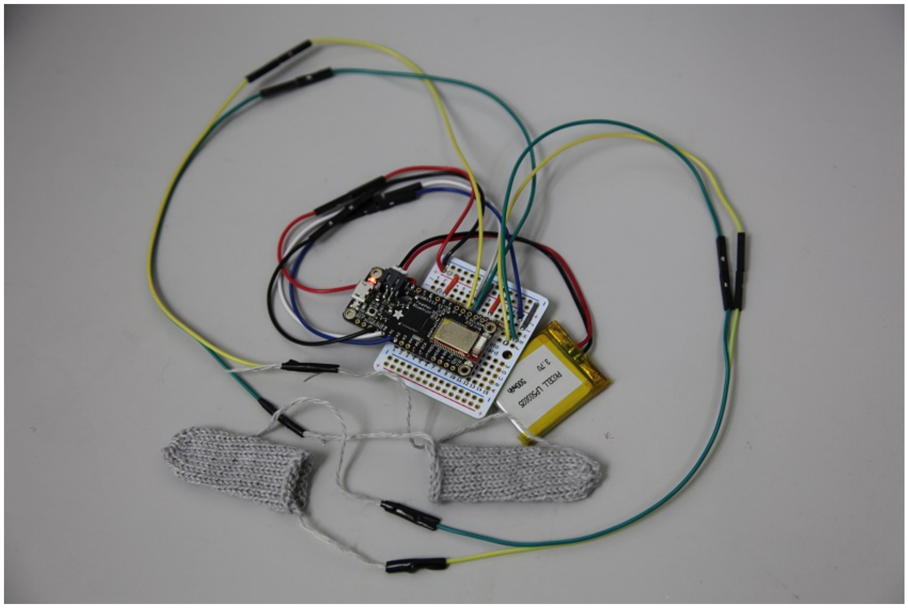

Figure 4.4: Stretch Sensor

4.3.3 Interactive Game

In order to evaluate the proposed interactive systems and test them from different aspects including effectiveness and user satisfaction, we have designed a selection game. It should be noticed that our targeted users are young kids with SMA, our game has simple graphic interfaces, instant audio feedback, and is very easy to understand. The game has multiple objects and asks users to select the Minion icon from bear icons. There are four different levels, and the selections are from three to nine. Users need to use two input devices, like switches or sensors. One input is for switching the cursor and the other input performs as the select function. Users can only access the next level by selecting the correct minion icon. The switches and sensors used for input are binding to a key on a regular keyboard. Figure 4.5

is a screenshot of the interactive selection game.

Figure 4.5: Game Screenshot

When a user plays the interactive selection game, every action from the user is recorded and delivered to Healthcare Professionals (HPs). Figure 4.6 is a screenshot that shows the operational data from a user. With the help from HPs, the following attributes are identified and logged when users play the interactive game:

- Total time of current level: It represents the staying time of a user in current level. Long duration normally means the user cannot use the current sensor very well.

- Number of false selections: When a user selects a bear icon which is incorrect answer, a counter will self-increase until the user completes current level. It reveals how possible the user trigger sensors unconscious.

- Total clicks of current level: It is a counter that records the number of actions taken by a user. This represents the efficiency of designed user interface layout.

- Time taken to select answer: This attribute records the time taken between a user moves cursor to correct answer and triggers selecting operation.

Therefore, we can utilize the proposed interactive game to collect user data and assess the usability of sensors for a certain user. In the following section we designed a pilot study for targeted users who are with type I SMA.

4.3.4 Pilot Study

The designed pilot study focuses on a participant who is a three year old boy with SMA Type I. Like most children with the same disease, he is fully intact cognitively but unable to function independently in many domains. He uses a tracheotomy tube and ventilator to breathe, and requires assistance to sit and to hold his head up. Due to extreme muscle weakness, our participant wears an exoskeleton and other assistive/mobility devices to help him utilize his minimal movements. Furthermore, he requires 24-hour supervision to monitor his vital signs. The participant does not

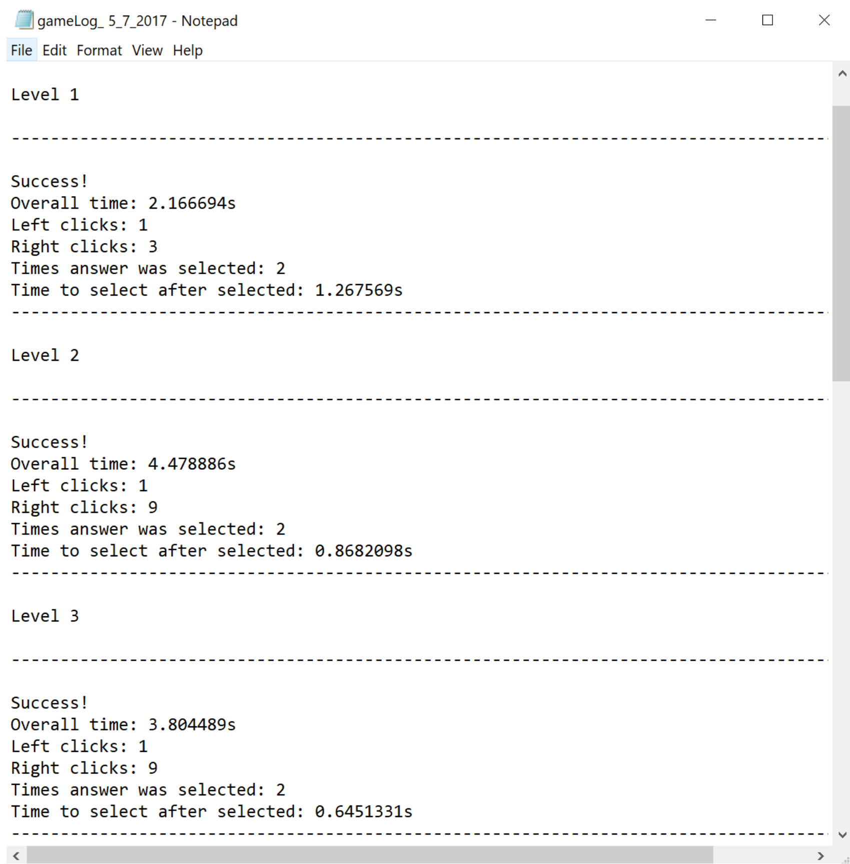

Figure 4.6: Operational Data

finish the pilot study and he is currently in the third phase, so we cannot present his experimental result in this dissertation.

Three phases are included in this user study: the first phase is to identify muscle functionality of the user and to interview parents; the second phase is used to implement input sensors and help users understand the test system; the last phase is for user evaluation and feedback collection.

4.3.4.1 First Phase User Study

During this phase, the major task is observing participants, since every individual of the targeted user group is unique and in different functionalities. For example, some children may use their left hands better while others are good with their left ankles. Therefore, we need to observe participants play with toys, watch videos, and communicate with parents in this study, then identifying the strongest muscles will be simple and straight forward. Moreover, interviewing parents can also increase the accuracy of result because they are very familiar with the preferences and habits of participants. For example, some children with SMA have both strong muscles on their legs and wrists; however, they are so active that they cannot keep their wrists on a stable switch. At this point, sensors for leg movement could perform better.

4.3.4.2 Second Phase User Study

When we are fully aware of the functionality of participants and their preferences, implementation and introduction of the game based system are the next phase. First, we need to deploy one or multiple sensors and switches for the user and guide him/her to trigger them based on information collected from first phase. Several adjustments are required during this phase to make participants comfortable and convenience to use sensor inputs. Furthermore, the position and mounting method should be considered in safety aspect since our targeted users are young children.

After sensor implementation, the next task in this phase involves user-guided learning. We need to introduce the interactive game and help users understand the objective of this game. Based on preference information gathered in the first phase, we are able to set the favorite cartoon character as a goal in each level such as Minions, and then adjust the game difficulty based on user functionality. After this task, the second phase of the user study is finished.

4.3.4.3 Third Phase User Study

During this phase, the user keeps practicing the interactive game with implemented sensors for two weeks. Besides the participant practicing the interactive game, parents will also take part in answering a questionnaire at the end. Their feedbacks is extremely critical and helpful since they understand their children very well and can provide an evaluation of the interaction experiences. Moreover, parents will report potential weaknesses of the system and provide valuable advise. All these feedbacks will be helpful to improve the user study and collecting valuable data in the future.

4.4 Assistive System for Young Children with Spinal Muscular Atrophy

In the previous section, we have designed a multi-modal interaction approach to enable young children with Spinal Muscular Atrophy (SMA) to interact with computer systems. Therefore, an assistive healthcare system is designed and implemented to improve living independence of the targeted users. The objective of our study is to increase the targeted users' quality of life.

4.4.1 System Architecture

Due to severe muscle weakness, most children with type I SMA have minimum movements and cannot speak. In addition, they are usually using medical devices that continuously track vital health data such as heart rate, blood pressure, and blood oxygen level. These data should be involved in the system, which can help to examine Situation Awareness (SA) of the user. The architecture of proposed system is described in Figure 4.7. The system input data can be divided into passive and active categories and the output is used for providing three services, namely environment control, mobility control and communication. The following sections will introduce the functions in detail.

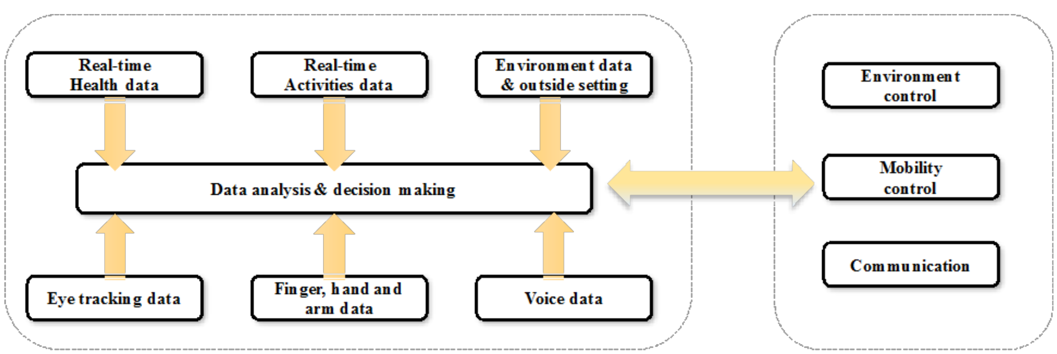

Figure 4.7: System Architecture of Assistive System for Users with SMA

4.4.1.1 Data Input and Output

- **Passive Data Input:** Passive data, including health and activities information, can be used to decide users' current status. That can be acquired from various devices such as traditional heart rate monitor or wear-able devices. Due to the special physical conditions, the vital health data of children are monitored continuously. In our system, we would like to explore commonly monitored health da-ta such as blood pressure and heart rate.

  Moreover, environmental data and external settings will be included as passive data input too. For example, the current room/location (e.g. bedroom, playroom, or gym), current time and pre-defined personal schedule will be recorded automatically in the background to help analyzing users' needs. In addition, we will also track wheelchair moving speed, obstacles around users and other similar data to better understand user status.

- **Active Data Input:** Even though the children with type I SMA have very limited physical capabilities, their voluntary movements are the most important and valuable interaction channel. Considering the physical conditions of the children, we will focus on collecting three types of user data, namely eye movements, finger/arm movements, and vocalization.

- **Data Output:** Given the passive and the active data inputs, data analysis and decision making module will use machine learning technology to determine the user status and provide services from the following three aspects.

4.4.1.2 Environment control

The proposed system will allow users to have some control over their environment. In the current stage, we would like to focus on the basic controls such as turning on or off the lights or TV. Some automatic controls will be considered. For example, when the current time has passed the typical wake-up time of the child and the system detects that the child is awake, the display screen can be turned on automatically and ask the child whether he would like to turn on the light or pull up the curtain. Environmental control will also allow users to change their location. For example, the child can choose to go to the gym room after staying in the living room for some time.

4.4.1.3 Mobility Control

With the limited physical capability, people with SMA largely depend on their wheelchairs to achieve mobility. Therefore, mobility control is a very important service we would like to support through the system. Even though there are many advanced wheelchairs available, it is still very challeng-ing for children with type I SMA to use them because of safety concerns and the precise positioning required to operate the wheelchairs. In this project, we will explore the possibilities of using data from multiple sources (e.g., EEG, voice, eye and finger movements) to improve the accuracy of the control. Similar to environmental control, we will consider some automatic control in this service as well. For example, if the system detects the wheelchair speed is too fast, it will automatically slow it down.

4.4.1.4 Communication

As the most important service component, assisting communication of children with SMA is the key module of the proposed system. We will utilize both passive and active data from multiple sources to understand user intentions. In the first stage of the study, we will focus on three main categories, namely self-care, entertainment, and environmental control. For example, if a user wants to play with puzzle, the system should be able to allow the user to easily specify 'puzzle' from the 'play' category. Alternatively, the user may need to be repositioned or wish to request suctioning of his airway, so he can select these options from the 'self-care' menu. We will use a Visual Scene Display (VSD) to embed the language concepts within the Augmentative and Alternative Com-munication (AAC) system. A VSD is a visual scene such as a photograph or picture that depicts or represents a situation, place, or experience for a child. For young children, grid layouts impose more metalinguistic demands that are beyond their skill levels [6]. When language concepts are represented within a context, it allows the child to generalize new vocabulary and find target vo-cabulary faster and easier, while also reducing metalinguistic demands on the child [7]. Literature supports the fact that VSDs are beneficial for young children and can be modified to fit the child's specific needs. Hot spots and specific vocabulary concepts can be easily adapted and are simple to learn in relation to a photograph or picture that the child is familiar with. We propose to first address the areas of self-care, environmental control, and play for the system, and will then expand to other functional areas as he learns and his needs change.

4.4.2 System Prototype

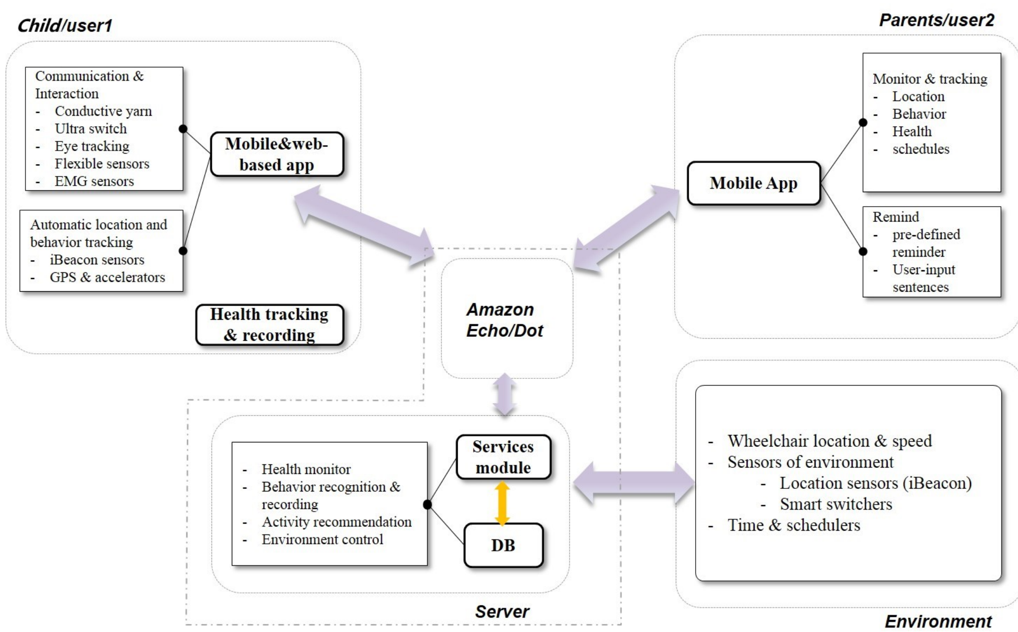

Figure 4.8: Assistive System Design

At the current stage of research, we are following the system architecture presented in the last section and have designed an assistive healthcare system which is shown in Figure 4.8. Partial functions are implemented and the system is host on a public web server. Thus users could access our system in any devices connected with Internet, which means our system is cross-platform. With the multi-modal interaction approach presented in the last section, users could utilize our system online easily.

4.4.2.1 Web-Based Application

Figure 4.9 shows some screenshots of the web-based assistive healthcare application. Based on the system architecture, three categories are identified by Healthcare Professionals (HPs), which are environment control, entertainment, and self-care. Figure 4.9 (a) shows the homepage of our web-based application that contains three icons. Figure 4.9 (b) and (c) are detailed available selections in self-care and entertainment pages. For some pages, for example, computer games in the entertainment page contains a lower level selections which provide specific games. Figure 4.9 (d) is a screenshot of specific computer games in the lowest level.

Since our targeted users are very young and they have limited ability to read text, every icon in our application is image based. This user-friendly design can help the targeted users understand the application easily and increase usability. The application allows a user interacting with various inputs, such as regular keyboard, regular mouse, and any sensors proposed in multi-modal interaction approach. Two operational functions in the application are very similar to that in the interactive

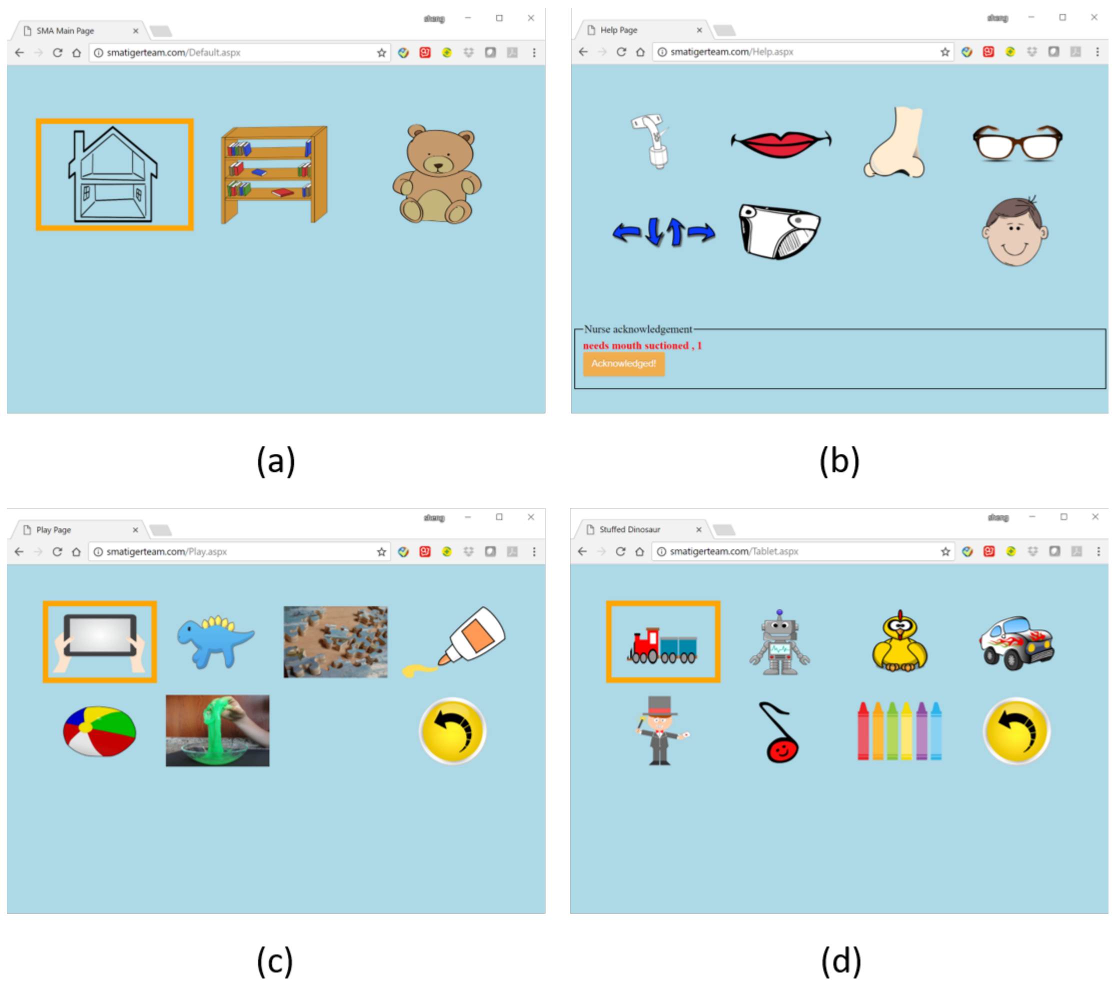

Figure 4.9: Web-Based Assistive Healthcare Application

game, which are switching and selecting. Users can use switching function to move cursor to the desired icon and enter the next page by selecting function. In order to decrease misunderstanding, there is a voice notification when the cursor move to a alternative icon. For example, when the cursor is moved to the bear icon describing self-care page, the system will say ’help please’.

4.4.2.2 Environment Control Module

In order to implement the environment control module, we involve two popular commercial products on the market: Amazon Echo and Samsung SmartThings. These two products compatible with each other and users could use voice command to control smart home devices. Figure 4.10 illustrates the design of environment control module. Considering most of targeted users have limited speaking ability, our web-based application can generate standard Amazon Echo voice command for them. For example, if a user want to turn on the TV, he/she can use the environment control module in the application and then select TV. The application will generate the voice command and speak out'Alexa, turn on the TV'. When this voice command is received by Amazon Echo, it will then control connected Samsung SmartThings to switch on the TV.

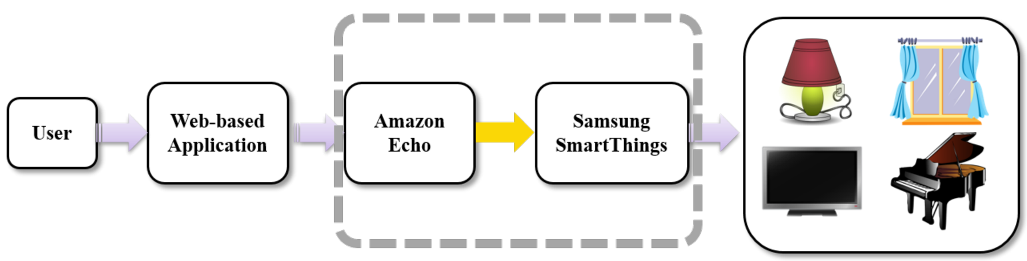

Figure 4.10: Environment Control Module

4.5 Conclusion

This chapter introduces an assistive healthcare system for young adults with Spinal Muscular Atrophy (SMA). Since the targeted users have limited muscular functionality, a multi-modal interaction approach is presented in the study. Some reliable and low-cost switches and sensors are selected to increase interactivities for young children. Furthermore, we design a pilot user study for a participant who has type I SMA. The participant is currently conducting the pilot study. Three phases of the user study are designed and they will be finished in the future work. Based on the approach, then we have designed the architecture of the system and implemented a web-based application with partial functions.

In the future, we will include more participants in different ages and functionality in this project and perform the pilot study. Another future work is implementing full functional application and involving participants to the user study.

Chapter 5

Conclusion

5.1 Conclusion and Discussion

In this dissertation we have studied some research questions regarding Situation Awareness (SA) applications. Today SA techniques are widely used in various and complex systems which belong to both commercial and military domains. The objective of our work is to research how to utilize SA techniques to help Subject Matter Experts (SMEs) and Healthcare Professionals (HPs) make decisions in military and healthcare contexts. We have also identified that SME/HP based human computation is necessary during the SA process since this can improve the performance and accuracy of the applications.

In the military context, we have studied how to calibrate the Value of Information (VoI) system for a specific battlefield so that it can provide more valuable intelligence to support military operations. Furthermore, we have extended our system into social media and have provided a detailed explanation on the affection of social media data on SA. In the healthcare domain, we have developed assistive healtcare systems for users under special health conditions. Our focus is set on two groups of users who are affected by Autism Spectrum Disorder (ASD) and Spinal Muscular Atrophy (SMA). The designed system can assist their daily activities and improve their quality of life.

Today military operations utilize information from a myriad of sources that provide overwhelming amounts of data. A primary challenge of decision makers at all levels is to identify the most important information with respect to the mission at hand, and often do so within a limited amount of time. Recently, a fuzzy logic based VoI calculation system has been developed by [29, 31], which becomes the foundation of our research on SA in the military context. The previous research has only provided standard triangular shape membership function which cannot represent complicated and complex military contexts. We extend the study and present trapezoidal shape membership functions in the current system.

Moreover, an adjustable VoI determination system is developed in chapter 2 which allows SMEs to tune customized membership functions for different battlefields. With the adjustable method, our system can achieve a higher accuracy of VoI calculation and lead to better SA.

Another research topic in chapter 2 focuses on combining multiple information culled from a dissimilar array of sources that describe the same event. The existing research only focuses on the value of a single piece of information and we present an investigation of the effect of integrating subsequent complementary and/or contradictory information into the VoI process. Since the integration requires a reverse Fuzzy Associative Memory (FAM) matrix to update Information Content, three different approaches are explored to generate empty rules. Regretfully, the initial results show that the proposed approach cannot work well with the integration.

Moreover, we have also presented the utilization of human processing for military SA based on social media data. Social media data are considered as a data input for the SA architecture and we identify the features and challenges that come with it. Besides, a novel approach to integrating human computation to the complicated process of interweaving social media to a military environment is presented in chapter 2. This chapter acquires various data and extracts high value information to compute SA, which answers the Research Question (RQ) 2 and RQ 3. We have also answered RQ 4 and RQ 5 by studying the integration of complementary and contradictory information and identifying that SME based human computation is necessary

Besides the military context, we have also applied SA techniques to the healthcare domain. With the fast development of information technology and sensor technology recently, more information related to healthcare becomes available. These data can be utilized to help individuals with special health conditions such as ASD and SMA, thus improving their quality of life.

In chapter 3, we focus on an SA based assistive healthcare system for young adults with ASD. An ambient intelligence based context aware assistive system is developed to improve their living skills and to help them perform daily activities like cooking and getting dressed. HPs can also use the system to help users with

ASD when they get stuck in a procedure of tasks. We have done some experiments, focusing on cooking assistance, and the HPs use the system to send prompts when users need help. During the experiments, user data are collected for decision model training. A Naive Bayes classifier is selected and cross validation results show that the automated prompting function can certainly perform like HPs. This makes a contribution to the independent living efforts of users with ASD. Chapter 3 acquires user information and situations which answer RQ 1. Since multiple ambient and user data is collected for SA, RQ 2 is answered. We have also involved HP based human computation to identify valued attributes for modeling which answers RQ 5.

In chapter 4 young children with SMA become our targeted users and we have designed an assistive healthcare system for them. Since most SMA patients have extremely limited muscular functionality so that they cannot use regular input devices to interact with computer systems, we have developed a multi-modal interaction approach. Some special and low-cost sensors are explored for targeted users and an interactive selection game is designed to verify usability of sensors. We have also designed a pilot study including three phases that collect feedbacks from both users and parents. Moreover, based on the communication approach, a comprehensive assistive system is designed. It provides multiple functions that lead to highly efficient communications, environment control, and which decrease living dependency.

In chapter 4, we acquire user information to compute SA, which is user inputs, this answers RQ 1. RQ 2 is answered by utilizing multiple sensors to compute SA and RQ 5 is answered by involving HPs, who identify muscular functionality and user requirements.

We have made several publications through this dissertation, and the list includes:

1. S. Miao, R.J. Hammell II, T. Hanratty, and Z. Tang, "Comparison of Fuzzy Membership Functions for Value of Information Determination," in MAICS, 2014, pp. 53-60.

2. S. Miao, R.J. Hammell II, Z. Tang, T. Hannatty, J. Dumer, and J. Richardson, "Integrating Complementary/Contradictory Information into Fuzzy-based VoI Determinations," in CISDA, 2015, IEEE Symposium on pp. 1-7.

3. R.J. Hammell II, T. Hanratty, and S. Miao, "Empirical Study on Combining Complementary and Contradictory Information in a Fuzzy-based System," in Computational Intelligence (SSCI), 2016, IEEE Symposium on pp. 1-8.

4. S. Miao and Z. Tang, "Utilizing Human Processing for Fuzzy-Based Military Situation Awareness Based on Social Media," in FUZZ-IEEE, 2017.

5. Z. Tang, J. Guo, S. Miao, S. Acharya, and J.H. Feng, "Ambient Intelligence Based Context-Aware Assistive System to Improve Independence for People

with Autism Spectrum Disorder," in System Science (HICSS), 2016, pp. 3339-3348.

6. S. Miao, Z. Tang, J.H. Feng, A. Jozkowski, and M. Lichtenwalner "An Exploratory Case Study to Support Young Children with Spinal Muscular Atrophy (SMA)" in Proceedings of ACM International Conference on Accessible Computing, 2017, Baltimore.

5.2 Limitation and Future Work

Even though we have developed an adjustable method for determining the value of single piece of information and combining complementary/contradictory information, all the analyses are based on value distribution. So it is necessary to vet the Value of Information (VoI) approaches with SMEs to provide direct feedback on applicability and to exercise the VoI construct within a task network model to assess the potential impact. Moreover, conducting human computation experiments to measure how cognitively aligned interfaces improve task performance will be another future task. Since the approaches presented in chapter 2 still cannot solve the empty rules problem, additional knowledge elicitation should be favored. Aside from the current system, some other fuzzy system architecture must be considered. Furthermore, we are interested in implementing the SA system based on social media data and then conducting some experiments after consulting with SMEs. We will also study how to improve computing efficiency so that we can achieve quick and accurate SA.

The assistive healthcare system for people with autism spectrum disorder in chapter 3 only utilizes limited data to guide living activities. In the future, more data will be collected such as daily schedule, indoor temperature, accelerated speed, and will be added to the assistive healthcare system. These will help to improve the accuracy of the SA process and help targeted users efficiently. Moreover, the cross-validation result shows that a naive bayes classification model cannot yet get very good results and provides limited help in decreasing the workload for HPs. That is because our experiment data are limited and we need to conduct more user studies. With more data from future experiments, we can train a better possibility model and achieve better accuracy.

At the current research stage, we only have one participant in the research examined in chapter 4 and the pilot study is not finished yet. More participants should be included in our study to assess the usability of the interaction approach so that we can collect more valuable user data. These data can be used for training a possibility model which provides automated calibration when users test various sensors for the first time. In addition, some new devices will be tested in the future, such as eye-tracking devices and vocal control devices. Another future task is implementing fully functional application and involving various participants into user study. Furthermore, designing customizable graphic user interface for the system is necessary, as our targeted users could have different preferences.

References

[1] Anonymous, “Us army field manual (fm) 2-22.3,” in *Human Intelligence Collector Operations*. US Army, 2006.

[2] J. Bleiholder and F. Naumann, “Conflict handling strategies in an integrated information system,” 2006.

[3] S. Tzu, *The art of war*. e-artnow, 2012.

[4] C. Billings, “Situation awareness measurement and analysis: A commentary,” in *Proceedings of the International Conference on Experimental Analysis and Measurement of Situation Awareness*, vol. 1. Daytona Beach, FL: Embry-Riddle Aeronautical University Press, 1995.

[5] M. R. Endsley, “Measurement of situation awareness in dynamic systems,” *Human factors*, vol. 37, no. 1, pp. 65–84, 1995.

[6] R. Taylor, “Situational awareness rating technique(sart): The development of a tool for aircrew systems design,” *AGARD, Situational Awareness in Aerospace Operations 17 p(SEE N 90-28972 23-53)*, 1990.

[7] M. Vidulich, C. Dominguez, E. Vogel, and G. McMillan, “Situation awareness: Papers and annotated bibliography,” DTIC Document, Tech. Rep., 1994.

[8] M. R. Endsley and M. W. Smolensky, “Situation awareness in air traffic control: The picture.” 1998.

[9] F. T. Durso, C. A. Hackworth, T. R. Truitt, J. Crutchfield, D. Nikolic, and C. A. Manning, “Situation awareness as a predictor of performance for en route air traffic controllers,” *Air Traffic Control Quarterly*, vol. 6, no. 1, pp. 1–20, 1998.

[10] D. N. Hogg, K. FOLLES, F. Strand-Volden, and B. Torralba, “Development of a situation awareness measure to evaluate advanced alarm systems in nuclear power plant control rooms,” *Ergonomics*, vol. 38, no. 11, pp. 2394–2413, 1995.

[11] J. Wirstad, “On knowledge structures for process operators,” Ergonomiraad AB, Tech. Rep., 1981.

[12] “Ushahidi, year=2017, howpublished="https://www.ushahidi.com/".”

[13] “Crisiscommons,” https://crisiscommons.org/, 2017.

[14] Anonymous, “Us army training and doctrine command (tradoc) pamphlet 525-66,” in *Military Operations Force Operating Capabilities*. US Army, March 2008.

[15] B. C. Dostal, “Enhancing situational understanding through the employment of unmanned aerial vehicles,” *Interim Brigade Combat Team Newsletter*, no. 01-18, 2007.

[16] M. T. Flynn, M. F. Pottinger, and P. D. Batchelor, “Fixing intel: A blueprint for making intelligence relevant in afghanistan,” DTIC Document, Tech. Rep., 2010.

[17] T. P. Hanratty, R. J. Hammell II, B. A. Bodt, E. G. Heilman, and J. C. Dumer, “Enhancing battlefield situational awareness through fuzzy-based value of information,” in *System Sciences (HICSS), 2013 46th Hawaii International Conference on*. IEEE, 2013, pp. 1402–1411.

[18] D. A. Shea and S. A. Lister, “The biowatch program: detection of bioterrorism.” Congressional Research Service [Library of Congress], 2003.

[19] “Biosense platform,” https://www.cdc.gov/nssp/biosense/, 2017.

[20] M. Keller *et al.*, “Use of unstructured event-based reports for global infectious disease surveillance-volume 15, number 5—may 2009-emerging infectious disease journal-cdc,” 2009.

[21] “Estimote beacon,” http://estimote.com/.

[22] “Sharp gp2d12,” http://zuff.info/SharpGP2D12_E.html.

[23] M. R. Endsley, "Toward a theory of situation awareness in dynamic systems," *Human Factors: The Journal of the Human Factors and Ergonomics Society*, vol. 37, no. 1, pp. 32–64, 1995.

[24] D. S. Alberts, J. J. Garstka, R. E. Hayes, and D. A. Signori, "Understanding information age warfare," DTIC Document, Tech. Rep., 2001.

[25] Y. Liang, "An approximate reasoning model for situation and threat assessment," in *Fuzzy Systems and Knowledge Discovery, 2007. FSKD 2007. Fourth International Conference on*, vol. 4. IEEE, 2007, pp. 246–250.

[26] M. Magnani and D. Montesi, "A survey on uncertainty management in data integration," *Journal of Data and Information Quality (JDIQ)*, vol. 2, no. 1, p. 5, 2010.

[27] J. Yen and R. Langari, *Fuzzy logic: intelligence, control, and information*. Prentice Hall Upper Saddle River, 1999, vol. 1.

[28] L. A. Zadeh, "Fuzzy sets," *Information and control*, vol. 8, no. 3, pp. 338–353, 1965.

[29] T. Hanratty, R. J. Hammell II, and E. Heilman, "A fuzzy-based approach to the value of information in complex military environments," in *International Conference on Scalable Uncertainty Management*. Springer, 2011, pp. 539–546.

[30] R. J. Hammell, T. Hanratty, and E. Heilman, “Capturing the value of information in complex military environments: A fuzzy-based approach,” in *Fuzzy Systems (FUZZ-IEEE), 2012 IEEE International Conference on*. IEEE, 2012, pp. 1–7.

[31] T. Hanratty, E. Heilman, J. Dumer, and R. J. Hammell II, “Knowledge elicitation to prototype the value of information.” in *MAICS*, 2012, pp. 173–179.

[32] E. A. Newcomb and R. J. Hammell II, “Examining the effects of the value of information on intelligence analyst performance,” in *Proceedings of the Conference on Information Systems Applied Research ISSN*, vol. 2167, 2012, p. 1508.

[33] ——, “Validating a fuzzy-based mechanism for improved decision support,” in *Proceedings of the 14th IEEE/ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD 2013)*, 2013, pp. 143–148.

[34] C. Bisdikian, L. M. Kaplan, M. B. Srivastava, D. J. Thornley, D. Verma, and R. I. Young, “Building principles for a quality of information specification for sensor information,” in *Information Fusion, 2009. FUSION'09. 12th International Conference on*. IEEE, 2009, pp. 1370–1377.

[35] D. E. Wilkins, T. J. Lee, and P. Berry, "Interactive execution monitoring of agent teams," *J. Artif. Intell. Res.(JAIR)*, vol. 18, pp. 217–261, 2003.

[36] L. A. Zadeh, "The concept of a linguistic variable and its application to approximate reasoning—i," *Information sciences*, vol. 8, no. 3, pp. 199–249, 1975.

[37] F. Wayland, *The limitations of human responsibility*. D. Appleton, 1838.

[38] A. M. Turing, "Computing machinery and intelligence," *Mind*, vol. 59, no. 236, pp. 433–460, 1950.

[39] L. von Ahn, "Human computation," *Doctoral Thesis*, 2005.

[40] E. Law and L. Von Ahn, "Input-agreement: a new mechanism for collecting data using human computation games," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2009, pp. 1197–1206.

[41] M.-C. Yuen, L.-J. Chen, and I. King, "A survey of human computation systems," in *Computational Science and Engineering, 2009. CSE'09. International Conference on*, vol. 4. IEEE, 2009, pp. 723–728.

[42] A. J. Quinn and B. B. Bederson, "Human computation: a survey and taxonomy of a growing field," in *Proceedings of the SIGCHI conference on human factors in computing systems*. ACM, 2011, pp. 1403–1412.

[43] L. Von Ahn and L. Dabbish, "Labeling images with a computer game," in *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM, 2004, pp. 319–326.

[44] T. Chklovski and Y. Gil, "Towards managing knowledge collection from volunteer contributors." in *AAAI Spring Symposium: Knowledge Collection from Volunteer Contributors*, 2005, pp. 21–27.

[45] J. Wolfers and E. Zitzewitz, "Prediction markets," *The Journal of Economic Perspectives*, vol. 18, no. 2, pp. 107–126, 2004.

[46] "Chacha search engine," http://www.chacha.com/, 2017.

[47] J. Bian, Y. Liu, E. Agichtein, and H. Zha, "A few bad votes too many?: towards robust ranking in social media," in *Proceedings of the 4th international workshop on Adversarial information retrieval on the web*. ACM, 2008, pp. 53–60.

[48] I. King, J. Li, and K. T. Chan, "A brief survey of computational approaches in social computing," in *Neural Networks, 2009. IJCNN 2009. International Joint Conference on*. IEEE, 2009, pp. 1625–1632.

[49] D. J. Watts, *Six degrees: The science of a connected age*. WW Norton & Company, 2004.

[50] A.-L. Barabási, "Linked: The new science of networks," 2003.

[51] B. He, C. Macdonald, and I. Ounis, "Ranking opinionated blog posts using opinionfinder," in *Proceedings of the 31st annual international ACM SIGIR conference on Research and development in information retrieval*. ACM, 2008, pp. 727–728.

[52] H. Deng, M. R. Lyu, and I. King, "Effective latent space graph-based re-ranking model with global consistency," in *Proceedings of the Second ACM International Conference on Web Search and Data Mining*. ACM, 2009, pp. 212–221.

[53] A. Agarwal and S. Chakrabarti, "Learning random walks to rank nodes in graphs," in *Proceedings of the 24th international conference on Machine learning*. ACM, 2007, pp. 9–16.

[54] R. Saga and H. Tsuji, "Ranking method for mediators in social network," in *International Conference on Online Communities and Social Computing*. Springer, 2007, pp. 151–159.

[55] J. Abernethy, O. Chapelle, and C. Castillo, "Web spam identification through content and hyperlinks," in *Proceedings of the 4th international workshop on Adversarial information retrieval on the web*. ACM, 2008, pp. 41–44.

[56] F. Benevenuto, T. Rodrigues, V. Almeida, J. Almeida, C. Zhang, and K. Ross, “Identifying video spammers in online social networks,” in *Proceedings of the 4th international workshop on Adversarial information retrieval on the web*. ACM, 2008, pp. 45–52.

[57] B. Krause, C. Schmitz, A. Hotho, and G. Stumme, “The anti-social tagger: detecting spam in social bookmarking systems,” in *Proceedings of the 4th international workshop on Adversarial information retrieval on the web*. ACM, 2008, pp. 61–68.

[58] T. Sudkamp and R. J. Hammell, “Interpolation, completion, and learning fuzzy rules,” *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 24, no. 2, pp. 332–342, 1994.

[59] H. Zimmermann, *Fuzzy Set Theory and Its Applications*. Allied Publishers, 1996. [Online]. Available: https://books.google.com/books?id=oJYSpp_0M6MC

[60] T. Hanratty, J. Dumer, R. Hammell, S. Miao, and Z. Tang, “Tuning fuzzy membership functions to improve value of information calculations,” in *Norbert Wiener in the 21st Century (21CW), 2014 IEEE Conference on*. IEEE, 2014, pp. 1–7.

[61] J. Bleiholder and F. Naumann, "Data fusion," *ACM Computing Surveys (CSUR)*, vol. 41, no. 1, p. 1, 2009.

[62] R. Ahmed, P. DeSmedt, W. Du, W. Kent, M. A. Ketabchi, W. A. Litwin, A. Rafii, and M.-C. Shan, "The pegasus heterogeneous multidatabase system," *Computer*, vol. 24, no. 12, pp. 19–27, 1991.

[63] D. Draper, A. Y. Halevy, and D. S. Weld, "The nimble integration engine," *ACM SIGMOD Record*, vol. 30, no. 2, pp. 567–568, 2001.

[64] J. Hammer, J. McHugh, and H. Garcia-Molina, "Semistructured data: The tsimmis experience." 1997.

[65] T. Eiter, M. Fink, G. Greco, and D. Lembo, "Efficient evaluation of logic programs for querying data integration systems," in *International Conference on Logic Programming*. Springer, 2003, pp. 163–177.

[66] J. Chomicki, J. Marcinkowski, and S. Staworko, "Computing consistent query answers using conflict hypergraphs," in *Proceedings of the thirteenth ACM international conference on Information and knowledge management*. ACM, 2004, pp. 417–426.

[67] V. d. Subrahmanian, S. Adali, A. Brink, R. Emery, J. J. Lu, A. Rajput, T. J. Rogers, R. Ross, and C. Ward, "Hermes: A heterogeneous reasoning and mediator system," 1995.

[68] A. Motro, J. Berlin, and P. Anokhin, "Multiplex, fusionplex and autoplex: three generations of information integration," *ACM SIGMOD Record*, vol. 33, no. 4, pp. 51–57, 2004.

[69] A. Bilke, J. Bleiholder, F. Naumann, C. Böhm, K. Draba, and M. Weis, "Automatic data fusion with hummer," in *Proceedings of the 31st international conference on Very large data bases*. VLDB Endowment, 2005, pp. 1251–1254.

[70] F. Naumann, A. Bilke, J. Bleiholder, and M. Weis, "Data fusion in three steps: Resolving schema, tuple, and value inconsistencies." *IEEE Data Eng. Bull.*, vol. 29, no. 2, pp. 21–31, 2006.

[71] X. L. Dong, L. Berti-Equille, and D. Srivastava, "Integrating conflicting data: the role of source dependence," *Proceedings of the VLDB Endowment*, vol. 2, no. 1, pp. 550–561, 2009.

[72] B. Zhao, B. I. Rubinstein, J. Gemmell, and J. Han, "A bayesian approach to discovering truth from conflicting sources for data integration," *Proceedings of the VLDB Endowment*, vol. 5, no. 6, pp. 550–561, 2012.

[73] M. Wu and A. Marian, "Corroborating answers from multiple web sources." in *WebDB*, 2007.

[74] X. Yin, J. Han, and S. Y. Philip, "Truth discovery with multiple conflicting information providers on the web," *IEEE Transactions on Knowledge and Data Engineering*, vol. 20, no. 6, pp. 796–808, 2008.

[75] L. Blanco, V. Crescenzi, P. Merialdo, and P. Papotti, "Probabilistic models to reconcile complex data from inaccurate data sources," in *International Conference on Advanced Information Systems Engineering*. Springer, 2010, pp. 83–97.

[76] X. Dong, L. Berti-Equille, and D. Srivastava, "Truth discovery and copying detection from source update history," Technical report, AT&T Labs–Research, Florham Park, NJ, Tech. Rep., 2009.

[77] L. Berti-Equille, A. D. Sarma, A. Marian, D. Srivastava *et al.*, "Sailing the information ocean with awareness of currents: Discovery and application of source dependence," *arXiv preprint arXiv:0909.1776*, 2009.

[78] R. J. Hammell, T. Hanratty, and S. Miao, "Empirical study on combining complementary and contradictory information in a fuzzy-based system," in *Computational Intelligence (SSCI), 2016 IEEE Symposium Series on*. IEEE, 2016, pp. 1–8.

[79] "Netpop research," https://netpop.com/.

[80] "Alexa: The web information company," http://www.alexa.com/.

[81] "Using social media for enhanced situational awareness and decision support," https://www.dhs.gov/publication/using-social-media-enhanced-situational-awareness-decision-support, 2017.

[82] "Lessons learned: Social media and hurricane sandy," https://www.dhs.gov/publication/lessons-learned-social-media-hurricane-sandy, 2017.

[83] S. Brooks, "Emergency mgmt 2.0: How# socialmedia & new tech are transforming preparedness, response, & recovery# disasters# part2# govt/ngos," *Washington, DC: US House of Representatives*, 2013.

[84] S. D'Amico, "Earthquake research and analysis," *Tech Press*, 2012.

[85] Y. Qu, P. F. Wu, and X. Wang, "Online community response to major disaster: A study of tianya forum in the 2008 sichuan earthquake," in *System Sciences, 2009. HICSS'09. 42nd Hawaii International Conference on*. IEEE, 2009, pp. 1–11.

[86] D. H. Sonnenwald and L. G. Pierce, "Information behavior in dynamic group work contexts: interwoven situational awareness, dense social networks and contested collaboration in command and control," *Information Processing & Management*, vol. 36, no. 3, pp. 461–479, 2000.

[87] L. Zedeh, "Outline of a new approach to the analysis of complex systems and decision processes," *IEEE Transactions on Systems, Man, and Cybernetic*, vol. 3, pp. 28–44, 1973.

[88] J. Yen and R. Langari, *Fuzzy logic: intelligence, control, and information*. Prentice-Hall, Inc., 1998.

[89] L. Bing and K. C. Chan, "A fuzzy logic approach for opinion mining on large scale twitter data," in *Utility and Cloud Computing (UCC), 2014 IEEE/ACM 7th International Conference on*. IEEE, 2014, pp. 652–657.

[90] Z. Wang, V. Joo, C. Tong, and D. Chan, "Issues of social data analytics with a new method for sentiment analysis of social media data," in *Cloud Computing Technology and Science (CloudCom), 2014 IEEE 6th International Conference on*. IEEE, 2014, pp. 899–904.

[91] L. Yang, X. Geng, and H. Liao, "A web sentiment analysis method on fuzzy clustering for mobile social media users," *EURASIP Journal on Wireless Communications and Networking*, vol. 2016, no. 1, p. 128, 2016.

[92] V. Chitraa and A. S. Thanamani, "Web log data analysis by enhanced fuzzy c means clustering," *arXiv preprint arXiv:1405.5509*, 2014.

[93] T. P. Hanratty, E. A. Newcomb, R. J. Hammell II, J. T. Richardson, and M. R. Mittrick, "A fuzzy-based approach to support decision making in complex military environments," *International Journal of Intelligent Information Technologies (IJIIT)*, vol. 12, no. 1, pp. 1–30, 2016.

[94] A. Speaks, "What is autism?" https://www.autismspeaks.org/what-autism, 2017.

[95] ——, "Facts about autism," https://www.autismspeaks.org/what-autism/facts-about-autism, 2017.

[96] V. K. Jensen and S. C. Spannagel, "The spectrum of autism spectrum disorder: A spectrum of needs, services, and challenges," *Journal of Contemporary Psychotherapy*, vol. 41, no. 1, pp. 1–9, 2011.

[97] A. Palmen and R. Didden, "Task engagement in young adults with high-functioning autism spectrum disorders: Generalization effects of behavioral skills training," *Research in Autism Spectrum Disorders*, vol. 6, no. 4, pp. 1377–1388, 2012.

[98] R. J. Landa and M. C. Goldberg, "Language, social, and executive functions in high functioning autism: A continuum of performance," *Journal of autism and developmental disorders*, vol. 35, no. 5, pp. 557–573, 2005.

[99] D. Hendricks, "Employment and adults with autism spectrum disorders: Challenges and strategies for success," *Journal of Vocational Rehabilitation*, vol. 32, no. 2, pp. 125–134, 2010.

[100] M.-H. Tseng, C.-P. Fu, S. A. Cermak, L. Lu, and J.-Y. Shieh, "Emotional and behavioral problems in preschool children with autism: Relationship with sensory processing dysfunction," *Research in Autism Spectrum Disorders*, vol. 5, no. 4, pp. 1441–1450, 2011.

[101] H. Sampath, R. Agarwal, and B. Indurkhya, "Assistive technology for children with autism-lessons for interaction design," in *Proceedings of the 11th Asia Pacific Conference on Computer Human Interaction*. ACM, 2013, pp. 325–333.

[102] C. Putnam and L. Chong, "Software and technologies designed for people with autism: what do users want?" in *Proceedings of the 10th international ACM SIGACCESS conference on Computers and accessibility*. ACM, 2008, pp. 3–10.

[103] M. Cramer, S. H. Hirano, M. Tentori, M. T. Yeganyan, and G. R. Hayes, "Classroom-based assistive technology: collective use of interactive visual schedules by students with autism." in *CHI*, 2011, pp. 1–10.

[104] L. Escobedo, D. H. Nguyen, L. Boyd, S. Hirano, A. Rangel, D. Garcia-Rosas, M. Tentori, and G. Hayes, "Mosoco: a mobile assistive tool to support children with autism practicing social skills in real-life situations," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2012, pp. 2589–2598.

[105] C. Fage, L. Pommereau, C. Consel, É. Balland, and H. Sauzéon, "Tablet-based activity schedule for children with autism in mainstream environment," in *Proceedings of the 16th international ACM SIGACCESS conference on Computers & accessibility*. ACM, 2014, pp. 145–152.

[106] C. Guimaraes, M. H. Pereira, and S. Fernandes, "A framework to inform design of learning objects for teaching written portuguese (2nd language) to deaf children via sign language (1st language)," in *System Sciences (HICSS), 2015 48th Hawaii International Conference on*. IEEE, 2015, pp. 2–10.

[107] A. Mihailidis and G. R. Fernie, "Context-aware assistive devices for older adults with dementia," *Gerontechnology*, vol. 2, no. 2, pp. 173–188, 2002.

[108] R. O. Duda and P. E. Hart, "Pattern recognition and scene analysis," 1973.

[109] P. Langley and S. Sage, "Induction of selective bayesian classifiers," in *Proceedings of the Tenth international conference on Uncertainty in artificial intelligence*. Morgan Kaufmann Publishers Inc., 1994, pp. 399–406.

[110] J. Nehmer, M. Becker, A. Karshmer, and R. Lamm, "Living assistance systems: an ambient intelligence approach," in *Proceedings of the 28th international conference on Software engineering*. ACM, 2006, pp. 43–50.

[111] A. Mileo, D. Merico, and R. Bisiani, "Wireless sensor networks supporting context-aware reasoning in assisted living," in *Proceedings of the 1st international conference on PErvasive Technologies Related to Assistive Environments*. ACM, 2008, p. 54.

[112] Y.-J. Chang, W. C. Chang, and T.-Y. Wang, "Context-aware prompting to transition autonomously through vocational tasks for individuals with cognitive impairments," in *Proceedings of the 11th international ACM SIGACCESS conference on Computers and accessibility*. ACM, 2009, pp. 19–26.

[113] C. Bodine, G. McGrew, M. Melonis, P. Heyn, T. Struemph, S. Elfner, A. Goldberg, G. Philips, and M. Lanning, "Context-aware prompting system (caps) for persons with cognitive disabilities: A feasibility pilot study," in *Poster presentation at the 11th Annual Coleman Institute Conference*, 2011.

[114] M. Melonis, A. Mihailidis, R. Keyfitz, M. Grzes, J. Hoey, and C. Bodine, "Empowering adults with a cognitive disability through inclusion of non-linear context aware prompting technology (n-caps)," in *RESNA Annual Conference. Arlington (VA): RESNA*, 2012, pp. 59–108.

[115] F. V. Jensen, *An introduction to Bayesian networks*. UCL press London, 1996, vol. 210.

[116] J. Pearl, *Probabilistic reasoning in intelligent systems: networks of plausible inference*. Morgan Kaufmann, 2014.

[117] N. B. Amor, S. Benferhat, and Z. Elouedi, "Naive bayes vs decision trees in intrusion detection systems," in *Proceedings of the 2004 ACM symposium on Applied computing*. ACM, 2004, pp. 420–424.

[118] C. C. Aggarwal and C. Zhai, *Mining text data*. Springer Science & Business Media, 2012.

[119] A. McCallum, K. Nigam *et al.*, "A comparison of event models for naive bayes text classification," in *AAAI-98 workshop on learning for text categorization*, vol. 752. Madison, WI, 1998, pp. 41–48.

[120] S. Ting, W. Ip, and A. H. Tsang, "Is naive bayes a good classifier for document classification," *International Journal of Software Engineering and Its Applications*, vol. 5, no. 3, pp. 37–46, 2011.

[121] Q. Wang, G. M. Garrity, J. M. Tiedje, and J. R. Cole, "Naive bayesian classifier for rapid assignment of rrna sequences into the new bacterial taxonomy," *Applied and environmental microbiology*, vol. 73, no. 16, pp. 5261–5267, 2007.

[122] S. A. Pattekari and A. Parveen, "Prediction system for heart disease using naïve bayes," *International Journal of Advanced Computer and Mathematical Sciences*, vol. 3, no. 3, pp. 290–294, 2012.

[123] A. K. Halladay, S. Bishop, J. N. Constantino, A. M. Daniels, K. Koenig, K. Palmer, D. Messinger, K. Pelphrey, S. J. Sanders, A. T. Singer *et al.*, "Sex and gender differences in autism spectrum disorder: summarizing evidence gaps and identifying emerging areas of priority," *Molecular autism*, vol. 6, no. 1, p. 36, 2015.

[124] D. M. Werling and D. H. Geschwind, "Sex differences in autism spectrum disorders." *Current opinion in neurology*, vol. 26, no. 2, pp. 146–153, 2013.

[125] "Weka3," http://www.cs.waikato.ac.nz/ml/weka/.

[126] M. R. Lunn and C. H. Wang, "Spinal muscular atrophy," *The Lancet*, vol. 371, no. 9630, pp. 2120–2133, 2008.

[127] "What is spinal muscular atrophy?" https://www.mda.org/disease/spinal-muscular-atrophy, 2017.

[128] J. Pearn, "Classification of spinal muscular atrophies," *The Lancet*, vol. 315, no. 8174, pp. 919–922, 1980.

[129] S. Ogino, D. G. Leonard, H. Rennert, W. J. Ewens, and R. B. Wilson, "Genetic risk assessment in carrier testing for spinal muscular atrophy," *American journal of medical genetics*, vol. 110, no. 4, pp. 301–307, 2002.

[130] B. Hillebrands and K. Ounadjela, *Spin dynamics in confined magnetic structures I*. Springer Science & Business Media, 2003, vol. 83.

[131] E. W. Ottesen, "Iss-n1 makes the first fda-approved drug for spinal muscular atrophy," *Translational Neuroscience*, vol. 8, no. 1, pp. 1–6, 2017.

[132] "The cost of biogen's new drug: $750,000 per patient," http://www.cbsnews.com/news/the-cost-of-biogens-new-drug-spinraza-750000-per-patient/, 2017.

[133] E. P. Armstrong, D. C. Malone, W.-S. Yeh, G. J. Dahl, R. L. Lee, and N. Sicignano, "The economic burden of spinal muscular atrophy," *Journal of medical economics*, vol. 19, no. 8, pp. 822–826, 2016.

[134] "Werdnig-hoffmann disease," https://rarediseases.org/rare-diseases/werdnig-hoffmann-disease/, 2016.

[135] Z. Tang, J. Guo, S. Miao, S. Acharya, and J. H. Feng, "Ambient intelligence based context-aware assistive system to improve independence for people with autism spectrum disorder," in *System Sciences (HICSS), 2016 49th Hawaii International Conference on*. IEEE, 2016, pp. 3339–3348.

[136] A. Von Gontard, K. Zerres, M. Backes, C. Laufersweiler-Plass, C. Wendland, P. Melchers, G. Lehmkuhl, and S. Rudnik-Schöneborn, "Intelligence and cognitive function in children and adolescents with spinal muscular atrophy," *Neuromuscular Disorders*, vol. 12, no. 2, pp. 130–136, 2002.

[137] A. M. Cook and J. M. Polgar, *Assistive technologies: Principles and practice*. Elsevier Health Sciences, 2014.

[138] I. S. Mackenzie and T. Felzer, "Sak: Scanning ambiguous keyboard for efficient one-key text entry," *ACM Transactions on Computer-Human Interaction (TOCHI)*, vol. 17, no. 3, p. 11, 2010.

[139] M. Zimmerling, F. Ferrari, R. Lim, O. Saukh, F. Sutton, R. Da Forno, R. S. Schmidt, and M. A. Wyss, "A reliable wireless nurse call system: Overview and pilot results from a summer camp for teenagers with duchenne muscular dystrophy," in *Proceedings of the 11th ACM Conference on Embedded Networked Sensor Systems*. ACM, 2013, p. 35.

[140] T. Guerreiro, H. Nicolau, J. Jorge, and D. Gonçalves, "Towards accessible touch interfaces," in *Proceedings of the 12th international ACM SIGACCESS conference on Computers and accessibility*. ACM, 2010, pp. 19–26.

[141] S. Trewin, C. Swart, and D. Pettick, “Physical accessibility of touchscreen smartphones,” in *Proceedings of the 15th International ACM SIGACCESS Conference on Computers and Accessibility*. ACM, 2013, p. 19.

[142] P. Carrington, A. Hurst, and S. K. Kane, “The gest-rest: a pressure-sensitive chairable input pad for power wheelchair armrests,” in *Proceedings of the 16th international ACM SIGACCESS conference on Computers & accessibility*. ACM, 2014, pp. 201–208.

[143] A. Fattouh, M. Sahnoun, and G. Bourhis, “Force feedback joystick control of a powered wheelchair: Preliminary study,” in *Systems, Man and Cybernetics, 2004 IEEE International Conference on*, vol. 3. IEEE, 2004, pp. 2640–2645.

[144] A. Duchowski, *Eye tracking methodology: Theory and practice*. Springer Science & Business Media, 2007, vol. 373.

[145] R. Jacob and K. S. Karn, “Eye tracking in human-computer interaction and usability research: Ready to deliver the promises,” *Mind*, vol. 2, no. 3, p. 4, 2003.

Appendices

Appendix A

Approval for Human Subjects Research – *Continuation of Protocol #1609004866*

1. Co-Principal Investigators: Dr. Amanda C. Jozkowski, Dr. Jinjuan (Heidi) Feng, & Dr. Ziying (Katherine) Tang.

Investigator's Affiliation with Towson University (Please check one):
Faculty _X_ Staff ___ Student ___ Outside Principle Investigator ___

Principal Investigator Signature:________________________ Date: 2/17/2017

Title of Research: Assistive Technology for Young Children with SMA and other Special Healthcare Needs ("Project yCAT")

Period of Research (start and end dates) Start: August 2016 End: August 2018

Institution & Department: Towson University Dept. of Occupational Therapy & Occupational Science and Dept. of Computer and Information Sciences

Address to which approval should be sent: Towson University Enrollment Services, Room 245

Applicant's Day Time Phone: (410) 704-2272 Applicant's E-mail Address: ajozkowski@towson.edu

Co-Investigator(s): Sheng Miao (doctoral student in CIS), Molly Lichtenwalner (masters student in SLP).

2. If you are a student please provide the following:

Faculty Sponsor Name: ________________________ Phone: ______________

3. Has this research project been previously considered by the IRB?

Yes X* No Last approval date: 09/06/2016

(*Note: current submission re-titled to reflect continuation of existing approved project with new focus on other diagnoses in addition to children with SMA)

4. If the research is funded, indicate the source:

Towson University Fisher College Endowment fund
CIO Forum Scholarship Award

5. Check if the following is true: (Be sure that you check all appropriate responses)

Does the research include:

X minors ___ prisoners ___ pregnant women

X use of audio, video, digital recordings, or participant photos to collect subject data

X use of protected health information

___ information that could affect the participant's employability, financial standing or reputation

___ information which deals with sensitive aspects of the participant's own behavior, such as illegal conduct, drug use, sexual behavior, or use of alcohol.

___ information which would place the participant at risk of criminal or civil liability if it became known outside the research

___ the use of educational tests (cognitive, diagnostic, aptitude, or achievement)

___ survey or interview instrument

X the participants being fully informed of the research project

X voluntary participation by all participants

___ interviewing or surveying only elected or appointed public officials or candidates for public office

___ observation of public behavior

X the collection or study of existing data, documents, records or specimens

___ procedures in which the anonymity** of the participant will be insured

6. Designate the category you believe describes your research:

    This is an addendum to an existing approved protocol that went through full review (#1609004866)

7. What is the objective of the study? (Be clear and concise. Do not use jargon)

Due to muscle weakness or fluctuations in neuromuscular tone, some children with special healthcare needs such as Spinal Muscular Atrophy (SMA) have limited movements in their limbs or heads, and sometimes cannot speak. Therefore, they are not able to communicate with other people or the environment through the typical approaches (e.g., speaking, traditional switches or remote controls, traditional input devices such as the mouse, keyboard, touchscreen, etc.). However, these children often have normal voluntary eye movements that allow them to interact with eye-gazed based applications, or limited but voluntary finger and arm movements that may allow them to use special ultra-sensitive switches.

Over the past year, the team has worked to develop and test a wearable system/device prototype (called "yCAT" for Young Children's Assistive Technology) that allows one child with special healthcare needs to:

(1) more easily communicate

(2) control his environment including light, temperature, and entertainment/play options
(3) navigate while sitting and control his power wheelchair

The most important and unique characteristics of yCAT are that it is:

A. Customizable (i.e., can be modified/adapted for others with SMA and similar disorders)
B. Affordable/accessible to people in the SMA community
C. Usable (does not take more than 5 minutes to set up and can easily be understood and modified by families who are not technically trained)
D. Durable
E. Responsive/sensitive
F. Accurate
G. Allows for maximal functional movement patterns (i.e., palms of hands are not obstructed so they can still be used for play)
H. Improves quality of life and meets the needs of very young children (approximately 18 months-5 years), for whom very few therapeutic products currently exist

The key research question is: "Can wearable technology be used to increase communication and provide environmental and mobility device control capabilities young children with special healthcare needs?"

Now that the team has worked with one family to develop a prototype system for yCAT (First Stage User Study - see ***Figure 1***), we are interested in collecting qualitative and quantitative feedback on the device's utility, acceptability, and design features. This will involve two new arms of the study: 1) Family use at home with children with SMA, and 2) user testing in the community for the speech-language components.

8. What is the research design and what will be required of each participant? (Attach extra page if needed)

<u>Arm 1 Second Stage User Study - Fall 2017-Spring 2018</u>

Up to 5 other families of young children (age 18 months – 5 years) with neuromuscular disorders (SMA) in the community will be targeted. We propose three direct contacts with each family:

1) Meet the family and child and obtain informed consent, take measurements of the child's arms and photos of their home environment and mobility devices in order to prepare the system for the individual child's needs.
2) Introduce the system to each family, train them in its use and assist with customization of the associated software, and direct them to use it in the home for 4-6 weeks.
3) Return to the home and collect feedback from the child and their family.

Each session is anticipated to last approximately 2 hours, and will be scheduled at the family's convenience. Additional communication with the team via email or video conferencing will require approximately 1 hour per week of the family's time, for a total of 10-12 hours for the study.

<u>Arm 2 Initial Testing for Generalization to Other Health Conditions – Spring 2018-Summer 2018</u>

In addition, we seek to test the AAC (speech-language) components of the device with a group of young children participating in a community-based therapy program, to assess generalizability to other special healthcare needs populations beyond SMA, including those with mild cognitive deficits and speech difficulties not related to movement disorders.

For this arm of the study, we will recruit up to 10 children from the Infant and Toddler (age 18 months to 5 years) therapy programs at the Towson University Institute for Well-being. We propose three contacts with each child:

1) Meet the child and review therapy file in order to customize the device for the child.
2) Approximately 2 weeks later, orient the child and therapist to the device and instruct on its use, making modifications as needed to the set-up, images, etc.
3) The following week, provide the device to the child to use during a therapy session and observe their use while remaining "hands-off," recording clinical observations and collecting objective data on accuracy through the system software.

Each session will last approximately 20 minutes, for a total of 1 hour required of the participants over the three sessions.

The entire project cycle is listed in Figure 1 below. We are now moving into year 2 of the project and are on track with our benchmarks.

Year 1	Year 2	Year 3
• Understand user population • Review literatures • Explore new research approaches • 1st stage system development • 1st stage user study	• 2nd stage system development • Web-based service development • 2nd stage user study • Deliver results to SMA community	• System finalization and evaluation • Refine and generalize results to other health conditions

***Figure 1*. Project Cycle**

9. How will the participants be selected? If you intend to recruit volunteers, please attach all advertisements and flyers.

Arm 1 of this project will involve up to 5 children with SMA and their families. We will recruit participants through the "SMA Support System" Facebook group, of which the research team has been a member for over a year. The SMA community is small but very open and a number of families have already expressed interest in partnering with researchers on future projects without any solicitation on our end. We will also send an announcement through the local chapters of the Cure SMA association, which has been supportive of our project in the past and allowed us to distribute informational brochures about our work to date at a recent event (see ***Appendices A*** and ***B***).

Participants in Arm 2 of this project will be recruited through the Towson University Occupational Therapy and Speech-Language Pathology centers at the Institute for Well-being. Because the study will take place within the therapeutic programs, the families have already provided consent to participate in student learning opportunities including University research, as part of the program intake process. Information letters will be sent home with all children enrolled in a therapeutic program who meet study criteria (18 months – 5 years with documented speech-language delay). The letters will be sent home as a courtesy and will outline the purpose and details of study participation. Parents will be given contact information for the study team in case they have questions, and the ability to opt out of participation will be provided, if desired (see ***Appendix C***). Therefore, formal consent is not required, but a short

Assent procedure will be used with the children (see ***Appendix D***).

10. What are the risks to the human participant?

    There are no known risks or discomforts associated with participation in this study, other than the slight risk that the individuals outside the study may obtain personal information about the family beyond that which is expected. Due to the desire of the families for the results of the study to be shared widely within the SMA community and beyond, they are very comfortable with allowing us to work with their children, and have indicated a willingness to have their names, images, and data (videorecordings, examples of user response output from the system) posted on the project website for all to see and use. The team will consult with the families on any shared content, and will respect their wishes to edit/modify as needed to protect privacy as desired.

11. How will confidentiality of the participants be maintained? (Is the study anonymous? Who will know the identity of the participants? If pre- and post-test, how will participants be identified?)

    Please see response to question #10, above. The participant families will have the right to view photographs, video-recordings and all response data for up to one year after conclusion of the study. All data will be securely stored in a locked filing cabinet or fingerprint- and password-protected computer system in the research office to prevent access by unauthorized personnel, in accordance with IRB policy.

12. Is there any information with regard to protocol or intention that will not be disclosed to the participant on the informed consent form? If so, what is it, and why will it not be disclosed?

    No.

13. What debriefing information will be given to the participants following their participation? If any information was withheld from the participants, it must be disclosed at the debriefing.

    The study team will provide a final research report no later than 6 months after the study conclusion including key study findings and an open-access, community-focused report via the project website.

14. Specify the participant characteristics required (age, gender, etc.) and the number of participants. (Be specific)

    <u>Arm 1</u>
    Each of the participant families will include one young child with SMA type 1 or 2 (age 18 months-5 years at the start of the data collection period), and at least one parent or guardian willing to assist with set-up, use of the device, and providing feedback.

    <u>Arm 2</u>
    Each child participant will be currently enrolled in one of the therapeutic programs offered through the Towson University Institute for Well-being, be 18 months – 5 years old at the start of the study, and have a documented speech-language delay (confirmed by chart review).

15. How will the data be recorded and stored? (Be specific). PLEASE NOTE: All original data must be kept for a minimum of three years. Data of student researchers must be kept in a secure place in the faculty sponsor's office.

    All data will be stored securely in a Co-PI's office for at least three years in a locked filing cabinet or password-protected dedicated drive accessible only to the research team. Video-recorded data and digitally scanned paper field note files will be securely stored on a dedicated drive accessible only to the research team (paper will be destroyed securely).

16. Financial consequences:

    Intellectual property and patent/copyright/trademark protections (under the three PIs) will be sought with advisement from Towson University for any technologies developed, once device and protocols are refined. The participant families and funding agencies have no financial claims to any of the materials developed as a result of this project. The family will agree to not seek intellectual property or invention ownership, as their participation is considered a service to the SMA community.

<u>***Please see attached:</u>
Appendix A. Consent Form (Arm 1)
Appendix B. Recruitment flyer for Facebook and email announcement (Arm 1)
Appendix C. Information letter (Arm 2)
Appendix D. Assent form (Arm 2)

Appendix A.

INFORMED CONSENT

STUDY TITLE: Assistive Technology for Young Children with SMA and other Special Healthcare Needs ("yCAT" study)

PRINCIPAL INVESTIGATORS:
Dr. Amanda C. Jozkowski, Dr. Jinjuan (Heidi) Feng, & Dr. Katherine (Ziying) Tang

DEPARTMENT:
Department of Occupational Therapy and Occupational Science / Department of Computer and Information Sciences

TELEPHONE NUMBER: (410) 704-2272

We invite you to take part in a research study. Please take as much time as you need to read the consent form. You may want to discuss it with your family, friends, or your personal doctor. You may find some of the language difficult to understand. If so, please ask questions. If you decide to participate, you will be asked to sign this form.

<u>WHY IS THIS STUDY BEING DONE?</u>

Drs. Jozkowski, Feng, and Tang at Towson University have formed a team to address the challenges your child has in communicating, and navigating in and controlling their environment. This group has been working with another local SMA family to develop solutions to make your child more independent in their daily life. We are now seeking other families to trial the prototype system and provide feedback on further refinements needed to the design and setup. The following is a brief outline of the goals of this project, which were developed in collaboration with other families in the global SMA community:

The team hopes to develop and test a wearable system/device that would allow your child to:
(1) more easily communicate
(2) control their environment including light, temperature, and entertainment/play options
(3) navigate while sitting and control their power wheelchair

The most important and unique characteristics of this device are that it is:
A. Customizable (i.e., can be modified/adapted for others with SMA and similar disorders)
B. Affordable/accessible to people in the SMA community
C. Usable (does not take more than 5 minutes to set up and can easily be understood and modified by families who are not technically trained)
D. Durable
E. Responsive/sensitive
F. Accurate
G. Allows for maximal functional movement patterns (i.e., palms of hands are not obstructed so they can still be used for play)
H. Improves quality of life and meets the needs of very young children (approximately 18 months-5 years), for whom very few therapeutic products currently exist

WHAT IS INVOLVED IN THE STUDY?

If you decide to take part, this is what will happen:

You are invited to participate in a family case study using a “participatory action research” model. This includes working with the team on co-generation of problem statements, collaboration on design and mounting of the system to meet your child’s needs and interests, and feedback on technological challenges and system functionalities. This approach includes frequent communication with you by email or video conferencing, as well as 3 intermittent home visits to observe, set up the system, and measure response to intervention. Responses will be gathered through video-recording, still photography, written summary of correspondence, and field notes, as well as system-generated user data (such as number of times each command is accessed or length of time to select items on the screen). We anticipate visiting your family in person 3 times for 2 hours each session over the course of the study, with approximately 6 weeks between visits. Other communication will require approximately 1 hour per week of your time.

In some cases, the researchers may wish to use a video recording or photograph of your child for academic presentations or publications. You will be asked to indicate your agreement to allow videos/photos of your child to be shown, and you have the option to request that your child’s face be obscured. Should you decide that you’d rather not have your child’s video/photograph used for academic purposes, you may decline and still participate in the study.

WHAT ARE THE POSSIBLE RISKS AND DISCOMFORTS?

There are no known risks or discomforts associated with participation in this study, other than the slight risk that individuals outside the study may obtain personal information about your family beyond that which is expected.

Due to the desire of the team and your family for the results of the study to be shared widely within the SMA community and beyond, we appreciate your willingness to have your names, images, and data (video-recordings, examples of user response output from the system) posted on the project website for all to see and use. The team will consult with your family on any shared content, and will respect your wishes to edit/modify as needed to protect privacy as desired.

There may be other risks/discomforts associated with participating in this study that we do not know at this time.

WILL YOUR INFORMATION BE KEPT PRIVATE?

We may publish the information from this study in journals or present it at meetings. If we do, we will not use your name, your child’s name, or any information that would identify you or your child unless explicitly approved by you in writing.

We will remove your name and other information that may identify you or your child from our research data. We will store the data in a locked filing cabinet or fingerprint- and password-protected computer system in our research office to prevent access by unauthorized personnel.

You have the right to view video-recordings and other data generated from the study for up to one year after conclusion of the study. All data will be destroyed or securely stored in a locked filing cabinet or fingerprint- and password-protected computer system in our research office to prevent access by unauthorized personnel according to IRB guidelines.

WHAT ARE THE POSSIBLE BENEFITS OF TAKING PART IN THIS STUDY?

The possible benefits for taking part in this study may include:

You are contributing to research that will advance our understanding of how to help children with severe disabilities (Type I SMA in particular) gain more independence through information technologies. The proposed system will provide a better match between the demands of the tasks and the abilities of the user, and accordingly, can lead to increased accessibility and more effective solutions. The main benefits of the proposed solution include:

- The system will allow your child and other children with similar conditions to communicate with people and control their own environment, which is crucial for independent living;
- This system bears much lower cost compared to existing commercial solutions that only support one fragment of the proposed activities (e.g., commercial eye-gaze applications cost several thousand dollars);
- This system can be available to assist the user 24 hours a day, 7 days a week;
- This system is highly flexible and customizable. Caregivers and therapists could update related data such as schedule or specific settings whenever needed.

WHAT OTHER OPTIONS ARE THERE?

An alternative would be to not participate in this study.

ARE THERE ANY PAYMENTS TO YOU FOR TAKING PART IN THE STUDY?

In appreciation of your time and participation, you and your child will be provided with a custom wearable adaptive device for use in your home and the community. In addition, we will also give you a thank-you letter in acknowledgment of you and your child's participation in the project.

WHAT ARE YOUR RIGHTS AS A PARTICIPANT, AND WHAT WILL HAPPEN IF YOU DECIDE NOT TO PARTICIPATE?

Your participation in this study is voluntary. Your decision whether or not to take part will not affect your current or future relationship with Towson University. You are not giving up any legal claims or rights. If you do decide to take part in this study, you are free to change your mind and stop being in the study at any time.

As the parent/guardian of a young child, you are asked to provide consent on their behalf, and assent will be waived. However, all efforts will be made to keep the child comfortable and aware of expectations. You may ask that we hold or discontinue study procedures at any time.

When agreeing to participate, you acknowledge that you have no financial claims to any of the materials developed as a result of this project. You will not seek intellectual property or invention ownership, as your participation is considered a service to the SMA and greater scientific communities.

CAN YOU BE REMOVED FROM THE STUDY?

You may be removed from this study without your consent if you or your child appear uncomfortable or become too ill to participate in the study.

WHOM DO YOU CALL IF YOU HAVE QUESTIONS OR CONCERNS?

You may contact Dr. Amanda Jozkowski at (410) 704-2272 with any questions, concerns, or complaints about the research or your participation in this study prior to or after enrollment in the study. If you are unable to contact the research team, please contact Dr. Elizabeth Katz, Chair of the Towson University Institutional Review Board at irb@towson.edu or (410) 704-3072. You will get a copy of this consent form.

AGREEMENT:

I have read (or someone has read to me) the information provided above. I have been given a chance to ask questions by contacting the researchers in person and by phone if needed. All my questions have been answered. By signing this form, I confirm that I am at least 18 years old, and I am agreeing to take part in this study, and to allow my child to participate.

Name of participant 1 (parent/guardian)	Signature	Date and time signed

Name of participant 2 *optional (parent/guardian)	Signature	Date and time signed

Name of participant 3
(child)

In some cases, the researchers may wish to use a video-recording or photograph of your child for academic presentations or publications (please check and initial one):

- ❑ I agree to allow my child's full face to be shown in academic presentations or publications as necessary.____**(initials)**
- ❑ I prefer that a portion of my child's face be obscured in all academic presentations or publications so that he/she cannot be identified.____**(initials)**
- ❑ I do not want photographs of my child included in any academic presentations or publications.____(initials)

- -

RESEARCHER USE ONLY:

I have personally explained the research to the research participant by phone or in person, and answered all questions. I believe that the potential participant understands the information described in this informed consent and freely consents to participate.

Name of Person Obtaining Informed Consent	Signature	Date and time signed

Appendix B.

Families of Young Children with SMA Type 1 or 2

Young Children's Assistive Technology

We invite you to participate in an exciting research project with Towson University to develop and test a new wearable assistive technology device to enhance communication and functional independence.

Is your child:

- Age 18 months – 5 years old?
- Unable to communicate verbally (or has difficulty)?
- Desiring more independent control of his/her environment (turning on lights, music, TV, etc)?

If so, please contact us for more information about Project yCAT!

TOWSON UNIVERSITY™

Details:

- Researchers will come to your home on three occasions to work with you and your child and customize a device compatible with our newly developed software using switches, eyegaze, or other Assistive Technology solutions.
- A total of 10-12 hours of your time is requested over the course of 6-8 weeks.
- At the conclusion of the study, the device is yours to keep!

Contact:

Amanda Jozkowski, PhD, OTR/L
(410) 704-2272
ajozkowski@towson.edu

Appendix C.

TOWSON UNIVERSITY

Project yCAT

Dear parent/guardian:

Your child has been identified as a potential participant in an exciting collaboration with the __________(therapeutic) Program and Towson University researchers.

We have develop a new assistive technology device ("yCAT") to enhance communication for very young children (age 18 months – 5 years), for whom few products are currently available.

During your child's time in the ________(therapeutic) Program this semester, student and faculty researchers will introduce our new device to your child and take observations on how he/she is able to use it. The researchers will be present for approximately 20 minutes over three sessions, and will coordinate with the Program supervisor and therapists to avoid interruption to the Program goals.

We thank you for your continued support of Towson student learning and innovation as clients of the Children's Therapy Center. If you have questions about this project or wish to opt out, you may do so without penalty or explanation required. Please contact Dr. Jozkowski for more information.

yCAT

Young Children's Assistive Technology

Contact:
Amanda Jozkowski, PhD, OTR/L
(410) 704-2272
ajozkowski@towson.edu

Appendix D. **Child Participant Assent Script**

Young Children's Assistive Technology

Session 1:

Hello, my name is ____________________(researcher).

I'm going to get to know you a bit today. Is it ok if I take a picture of you and watch you play for a while with your friends here?

YES NO (circle child's verbal or nonverbal response)

It is ok if you don't want me to stay. If you want me to stop, please tell me STOP, or come tap my shoulder.

Thank you, this will be fun! I'm excited to meet you.

Session 2:

Hi, __________________(name of child)! It's your friend ___________________(name of researcher) again.

Did you know I helped make a computer toy called "yCAT" that can help kids talk?

Is it ok if I show you how to play with the "yCAT" so you can talk to your friends here today? I will help you.

YES NO (circle child's verbal or nonverbal response)

It is ok if you don't want to play with yCAT. If you want me to stop, please tell me STOP, or come tap my shoulder.

Thank you, this will be fun! I'm excited to show you yCAT.

Session 3:

Hi, __________________(name of child)! It's your friend ___________________(name of researcher) again.

Do you remember when we played with yCAT to help you talk to your friends here?

I want to see how fun and easy it is for you to use without my help. Is it ok if I watch you play with the "yCAT" today?

YES NO (circle child's verbal or nonverbal response)

It is ok if you don't want to play with yCAT. If you want me to stop, please tell me STOP, or come tap my shoulder.

Thank you, this will be fun! I'm excited to meet see you play with yCAT!

Curriculum Vitae

Sheng Miao

[redacted]

SKILLS AND QUALIFICATIONS

- Programming ability in C++, Java, Python, Matlab, Android
- Developing ability in Arduino
- Data fusion and Data mining
- Communication engineering
- Signal processing and image processing
- Networking

EDUCATION

Doctor of Science in Information Technology July, 2017

Towson University, Towson, MD

Master of Science in Applied Information Technology May, 2012

Towson University, Towson, MD

Bachelor of Science in Communication Engineering June, 2011

Qingdao University, Qingdao, Shandong, China

PROFESSIONAL EXPERIENCE

Research Assistantship for Dr. Robert J. Hammell II in Towson University.
April 2013~now
Participated in designing, constructing and implementing Fuzzy logic based Value of Information assessment system for US Army Research Laboratory

- Research on Value of Information, Adjustable Fuzzy models, Complementary and Contradictory information fusion.
- Delivery an adjustable fuzzy model based application to ARL subject matter experts

System developer for Dr. Jinjuan Feng, Dr. Ziying Tang, and Dr. Subrata Acharya in Towson University

- Sept. 2013-June. 2015

Participated in developing assistive system for people with Autism. Build a classification model based on Naïve Bayes method for automotive prompting.

System developer for Dr. Jinjuan Feng, Dr. Ziying Tang, and Dr. Amanda C. Jozkowski in Towson University

- Nov. 2015-June, 2017

Participated in developing assistive system for children with Spinal Muscular Atrophy.

Technical engineer for Qingdao Sinetec Technology Corporation.

September 2008~May 2010

Programming for Professor Qun Miao in Qingdao Technological University

March. 2010

Publication

S. Miao, R.J. Hammell II, T. Hanratty, and Z. Tang, "Comparison of Fuzzy Membership Functions for Value of Information Determination," in MAICS, 2014, pp. 53-60.

S. Miao, R.J. Hammell II, Z. Tang, T. Hannatty, J. Dumer, and J. Richardson, "Integrating Complementary/Contradictory Information into Fuzzy-based VoI Determinations," in CISDA, 2015, IEEE Symposium on pp. 1-7.

R.J. Hammell II, T. Hanratty, and S. Miao, "Empirical Study on Combining Complementary and Contradictory Information in a Fuzzy-based System," in Computational Intelligence (SSCI), 2016, IEEE Symposium on pp. 1-8.

S. Miao and Z. Tang, "Utilizing Human Processing for Fuzzy-Based Military Situation Awareness Based on Social Media," in FUZZ-IEEE, 2017.

Z. Tang, J. Guo, S. Miao, S. Acharya, and J.H. Feng, "Ambient Intelligence Based Context-Aware Assistive System to Improve Independence for People with Autism Spectrum Disorder," in System Science (HICSS), 2016, pp. 3339-3348.

S. Miao, Z. Tang, J.H. Feng, A. Jozkowski, and M. Lichtenwalner "An Exploratory Case Study to Support Young Children with Spinal Muscular Atrophy (SMA)" in Proceedings of ACM International Conference on Accessible Computing, 2017, Baltimore.