Social-Context Middleware for At-Risk Veterans

Nadiyah Frances Johnson
Marquette University

SOCIAL-CONTEXT MIDDLEWARE FOR AT-RISK VETERANS

By
Nadiyah Johnson

A Thesis submitted to the Faculty of the Graduate School,
Marquette University.
In Partial Fulfillment of the Requirements for
The Degree of Master of Science

Milwaukee, WI

May 2016

ABSTRACT

SOCIAL-CONTEXT MIDDLEWARE FOR AT-RISK VETERANS

Nadiyah Johnson

Marquette University, 2016

Many veterans undergo challenges when reintegrating into civilian society. These challenges include readapting to their communities and families. During the reintegration process veterans have difficulties finding employment, education or resources that aid veteran health. Research suggests that these challenges often result in veterans encountering serious mental illness. Post-Traumatic Stress Disorder (PTSD) is a common mental disease that veterans often develop. This disease impacts between 15-20% of veterans.

PTSD increases the likelihood of veterans engaging in high risk behaviors which may consist of impulsivity, substance abuse, and angry outbursts. These behaviors raise the veterans' risk of becoming violent and lashing out at others around them. In more recent studies the VA has started to define PTSD by its association to specific high risk behaviors rather than defining PTSD based on a combination of psychiatric symptoms. Some researchers have suggested that high risk behaviors -- extreme anger (i.e., rage or angry outbursts) is particularly problematic within the context of military PTSD. Comparatively little research has been done linking sensor based systems to identify these angry episodes in the daily lives of military veterans or others with similar issues.

This thesis presents a middleware solution for systems that work to detect, and with additional work possibly prevent, angry outbursts (also described in psychological literature as "rage") using physiological sensor data and context-aware technology. This paper will cover a range of topics from methods for collecting system requirements for a subject group to the development of a social-context aware middleware. In doing such, the goal is to present a system that can be constructed and used in an in lab environment to further the research of building real-world systems that predict crisis events, setting the state for early intervention methods based on this approach.

ACKNOWLEDGMENTS

Nadiyah Johnson

First and above all, I praise God for granting me the capability to proceed successfully. I warmly thank and appreciate my mother and father, Cynthia and Lyndon, for being a constant source of motivation. I would also like to thank my brother, sister, friends and entire family for their continuous support throughout my academic career. I would like to express my sincere gratitude to my uncle Tracey for providing assistance in numerous ways. Additionally, I owe my deepest gratitude to my grandparents for paving the way to my academic success.
.

I am extremely grateful to my advisor Dr. Sheikh Iqbal Ahamed for his support and guidance throughout this process. I would also like to thank my committee members Dr.Zeno and Dr.Praveen for providing me with valuable ideas and insights as I prepared my thesis.

Finally, I would like to acknowledge the veterans for their help with this project. It has been a great honor to work with them.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

Civilian reintegration is a difficult process for many veterans. Military culture is much different from civilian life. Prior to transitioning into civilian status veterans are accustomed to a structured and militant lifestyle. According to a poll from the Washington Post approximately 50% of veterans suffer from readjustment issues[1]. Over time if these issues are not properly dealt with, veterans may become more susceptible to mental illness [2].

Post-traumatic stress disorder (PTSD) is one of the main problems associated with civilian reintegration. It is a condition of consistent emotional and mental stress resulting from trauma. Recent studies from the United States Veteran Affairs (VA) have started to define PTSD based on high risk behaviors rather than a cluster of psychiatric symptoms. This diagnosis increases the likelihood of veterans engaging in at risk behavior. At-risk behavior is a lifestyle or activity that places a person at increased risk of violent, dangerous, or unhealthy outcomes. This type of behavior involves heavy consumption of alcohol, substance abuse, impulsive activities, angry outburst (AOB) etc. AOB is particularly problematic in the context of military based PTSD. Continuous engagement in these at-risk behaviors such as AOB can result in homelessness, social isolation, and even suicidal thoughts [1]. Veteran's likelihood of engaging in these activities is heavily correlated with their social environment [2].

In this thesis, guidelines are presented for developing a middleware used to detect and predict psychological crisis events. For example, one crisis event of particular interest is angry outbursts (rage). The middleware will use physiological sensor data and social context technology to infer that rage is occurring within a social interaction that the veteran is party to. This system will output a value which represents whether or not the user's current social state indicates that they are likely to engage in an at-risk activity. To make this inference about the social context, it will take into account three types of underlying information, physiological data, text analysis, and GPS data. This middleware will be particularly useful in the development of an angry outburst detection and prediction system.

The paper is organized as follows: first the motivation behind this research is presented, followed by the related works section which is based on past context-aware systems, next the solution for this project is discussed; Finally, the broad impact of this approach and a discussion of its weaknesses are presented.

CHAPTER 2: BACKGROUND

In the past there have been several attempts to develop systems that help veterans to combat their PTSD. Mobile devices are ideal for mobile health (mHealth) interventions. Their built in sensor technology makes it possible to obtain relevant and useful data. Several studies suggest the use of mobile devices for developing systems that aid or facilitate mental health care. The VA has developed several apps that have a focus on PTSD. PTSD Coach, PE Coach, PTSD Support, PTSD Eraser are a few of many mobile apps that have been developed to target issues associated with PTSD. These apps either work to inform veterans about PTSD, help them to self-assess, or provide management techniques to help them to cope with the disorder [3]. There have been a few projects which work on the social aspect of civil reintegration. POS REP connects veterans within a perimeter range. Veterans are able to use this app to reach out to other vets when they are in need of support [4]. There has yet to be a mobile app dedicated to the detection and prevention of AOB of veterans.

Social computing is an interdisciplinary research and application field based on theoretical computational and social sciences [5]. Social computing is a focal point of many information and communication technology (ICT) systems. Social context is essential to developing an efficient model of the social system surrounding a particular individual or other social unit of interest. In order to understand a social interaction it is important to recognize the social context. Successful social interactions construct common knowledge which is often an unspoken shared understanding that enforces social norms [6]. In order for technology to successfully detect or predict behavior it must be able to accurately describe social context, typically in quantitative ways. To this end, the current project proposes social context middleware that can be used in a variety of systems to increase the system's understanding of the social state of veterans.

CHAPTER 3: MOTIVATION

Members of the armed forces experience an abundance of lifestyle changes as they transition from uniformed duty to civilian status. These changes can result in some veterans having to contend with challenges. Many veterans return to civilian society with post-traumatic stress (PTSD), along with other physical and emotional ailments. Due to the lack of veteran assistive programs a large number of veterans face their problem alone which eventually increases the severity of their issues. PTSD is a common mental illness attributed to military life. This illness along with other emotional disorders can lead to violent behavior, homelessness and in severe cases suicide.

The following scenarios highlight the targeted research problem. The goal of these scenarios is to expose the common issues associated with anger related PTSD. In these scenarios the mobile phone and wearable device are constantly on the user’s person allowing the device to collect the necessary sensor data for the middleware to capture and analyze. These scenarios emphasize the need for a social context middleware.

3.1 Scenario 1- Nightmares

Michael has returned from Iraq. He has been involved with the Dryhootch peer mentoring program for several weeks. His peer mentor made a note in the peer report that Michael has been experiencing stress lately. The answers from Michael’s self-report indicates that he hasn't been getting much sleep. It is learned from the peer and self-reported data that Michael has recently been suffering from symptoms of PTSD. It is not clear, however, that his PTSD is manifesting as vivid nightmares. Over the past couple of weeks Michael has been waking up from these nightmares. The sensor data from the wearable device and mobile phone reads in GPS, and physiological data which indicate that he paces around until he falls back to sleep. Recently Michael's PTSD has worsened and he is now suffering from more intense nightmares of his time in Iraq causing him to relive his past experience (i.e. re-experiencing symptoms, commonly referred to as "flashbacks"). As he wakes up his flashbacks overlap with reality and he begins swinging his arms to punch whatever is near him. When Michael heads to bed, aware of his serious condition, he locks his bedroom door, turns off the lights and lies down.

A system developed to detect angry outburst or PTSD related symptoms can greatly benefit from social context middleware. The middleware would first reads in Michael's environmental data using his mobile phone and wearable devices. The system would use the middle wear to recognize the inactivity and deduces that Michael is sleeping. Again, michael wakes up in the middle of the night in a cold sweat, heart racing, and swinging his arms in a boxing-like motion. He is mentally unable to differentiate reality from his nightmare. His phone is sitting on the nightstand but he is wearing a wearable device. This device reads in social context data (his fast pulse, high temperature etc). The middleware produces an output which indicates that there is a high possibility Michael is engaging or will engage in further at risk behavior. The middleware enables the angry outburst detection system to conclude that Michael is waking up in the middle of the night with symptoms of PTSD. This means Michael is more likely to experience at-risk behavior. The angry outburst detection system will be able to use the information provided by the middleware to prevent, stop, or later predict an angry outburst.

3.2 Scenario 2: At-risk Driving

Craig has returned from war. We learned from the peer and self-report that he has recently been suffering from symptoms of PTSD that is manifesting as stress induced uncontrollable anger. Craig has been having a hard time finding a job and the bills are beginning to pile up. Craig comes home after going through what he considers a bad job interview. He starts reading the mail and sees that his house is going into foreclosure. Enraged, Craig grabs his keys jumps in the car and searches for a familiar escape. An angry outburst detection system may utilize social context middleware to read in and analyze context data. The system will be able to use the data retrieved from the middleware in an algorithm to understand the user's current behaviors (speeding, erratic driving). In this particular scenario the phone is laying on the passenger seat. Michael is wearing his wearable device. He is violently gesturing to other drivers on the road. The wearable device is able to read in context data (fast heart rate and location). The angry outburst detection system, with the help of the social context middleware, can conclude that he is experiencing at-risk behavior- road rage or angry outburst.

3.3 Scenario 3: At-risk Communication

Jordan has been feeling stressed lately because his wife just lost her job making his job the main source of income. As a result of this stress he has become less motivated, more depressed and more hostile towards others. One day Jordan begins arguing with his wife via text message. He starts using violent language. His text messages are filled with swear words and threatening phrases. Later that day he posts on Facebook one sentence that says “Today my demons won, farewell”. Our social-context middleware would be used to detect his unusual and suspicious communication with his loved ones. An angry-outburst/ PTSD assistive app would be able to use a middleware which provides social-context data allowing it to successfully detect dangerous behavior.

3.4 Characteristics of the Middleware

The goal of this project is to help people in similar situations as those in the above scenarios. The three scenarios are common situations that at-risk veterans face on a regular basis. These common scenarios have very high costs. When a veteran experiences an AOB they put themselves at an increased likelihood of a dangerous, violent, or unhealthy outcome. AOB can lead to an individual isolating themselves from their family. It can lead to a person losing their job. It can also cause others around them distress by placing them in an unsafe environment.

It is apparent that new technology is necessary to help combat AOB. The middleware is designed to help veterans in situations similar to the ones described above. In the future the middleware may be extended to help people suffering from other crisis events.

The goal is to create a middleware that helps systems/mobile apps:

- Detect AOB
- Predict AOB
- Gather pertinent user data to expand knowledge of AOB

CHAPTER 4: RELATED WORK

4.1 Context-Aware Definitions

In 1995 Schilit and Theimer were one of the firsts to introduce the concept of context -awareness. They defined context as: where the subject is physically located, the people that they are interacting with the subject, and objects or resources nearby [7]. Their research suggests that context-aware systems should be able to react to changes in the environment based off of the information it encompasses. Schilit and Theimer break context aware-computing down into three applications (1) Proximate selection involves strategically placing an object on a user interface to increase accessibility;(2) Automatic Contextual Reconfiguration, which is the process of adding or removing a network connection based on the user's environment; (3)Contextual information and commands, which is the concept of a system accounting for the subjects situation particularly their location and adapts to system commands as a result. These applications were prototyped on a PARCTAB, a small palm size computer. This study shows the range of context-aware applications but is limited to proximate-environmental context.

In 2006 Andid Dey defined context awareness technology as a system that provides relevant services or information to a user based on the user’s tasks [8]. This research generalizes context-aware computing by categorizing context- computing features into three categories, (1) Presentation of information and services, (2)Automatic execution of services for a user and (3)tagging of context to information to support later retrieval. These types of features were combined with abstractions which were then implemented in a context tool-kit that is specific for building developers. The toolkit makes it easy for software to develop into a context-aware system. Their system uses widgets, interpreters and aggregators. The widgets use conventional poll and subscribe methods, the interpreters fuse data to derive context and the aggregators gather the entire context about one particular entity. The tool-kit approach is useful because it generalizes context-awareness in such a way where it can have many applications however it is limited in its ability to derive the situational context as it relates to relevant information between the user and target outcome of the software. Situational context is one of the most significant aspects of context-aware computing. It determines the relevancy of data.

Jason I. Hong and James A. Landay conducted a research study in 2001 to describe the benefits of a service infrastructure for context awareness [9]. They claimed the network based service infrastructure approach will benefit the developers and administrators and end-users. One benefit is that this approach will allow for compatibility with a wider range of devices due to its independence of hardware, operating system and programming languages. Another benefit of this approach is that the sensors can be changed independently and dynamically. This means other sensors or services may continue running as the context-aware service is being changed or updated etc. The third highlighted benefit in this research was that the context aware devices and applications will be easier to develop and deploy because all of the processing power, storage of data and sensors will be shared. This is beneficial because individual devices and apps will not need to handle every sensor in order to collect relevant context data. Instead, the burden can be placed on the infrastructure to find suitable nearby sensors. A context-aware network based middleware service approach is beneficial as it provides external sensors and processing power that is not application dependent but it lacks the ability to provide specific context-data for a particular need.

4.2 Physiological Context Data

Physiological sensor data can be a useful indicator of a patient's health status. As technology advances it is becoming easier to obtain physiological data via mobile devices. Many devices such as mobile phones and wearable devices are used to collect such data because they have built in sensors such as accelerometers and gyroscopes. The increased use of mobile devices makes them an ideal source for physiological data. For example, Won-Jae Yi et al uses an android mobile phone to collect and fuse sensor data to determine whether the user is at risk of falling [10]. This study points to how activity sensor data can be used to achieve accurate health assessments.

While most mobile devices have built in sensors many lack sensors that can obtain extensive physiological sensor data. Often research studies use external sensors to obtain accurate physiological data. A study by Rajan, D. et al uses a small wireless low sensor platform called SHIMMER which stands for Sensing Health with Intelligence, Modularity, Mobility and Experimental Reusability [11]. This platform combines a micro controller with daughter boards (bio sensors) making it perfect for medical sensing applications. In this study an app was developed using this platform pointing to how affective physiological sensors can be used as an effective way to extract relevant features from sensor measurements such as oxygen saturation and respiratory intake.

Physiological sensor data can be additional context-aware data. A study on relationships between behavioral and physiological sensor data involves using an analytic tool to mitigate the challenges of interfacing multiple software programs and implementing the data analysis functionality within and across these software programs [12]. In this case the physiological data was used to identify its correlation between behavioral data. This over all helps the subjects to comprehend the situation or context in which the behaviors and events transpire. This can result in the advancement of effective therapies.

4.3 Emotion Based Context-awareness

There have been many attempts to advance context-aware technology and increase its application use in systems however many approaches fail to address specific needs for context-aware data. Berthelon, F. et al focuses on a particular sphere of context-awareness which relates to emotion. They incorporated ontology, a well-known interoperable reasoning tool into the process of emotion detection. Their ontological model is broken into two parts, categorization and contextualization. Categorization maps the emotion while the contextualization part works to utilize emotion expression as a function of a person's contextual knowledge. The focus of this was limited to the study of philia and phobia but offered information for future research of context-awareness in respect to emotion detection.

Context- aware technology can provide impactful information in emotion detection research. Emotion detection is largely dependent on the situation at hand. Text is a growing mode of communication due to the rapid production and use of mobile devices. Text has become a tool for deriving context data. Douiji, Y. et al has developed a system that utilizes text as context data in order to recognize emotions [13]. In this research study text data is collected from the subject's mobile devices when they engaged in instant messaging and social media posts. Once collected the text data is sent to the server for text processing and emotion extraction. The text data is also combined with other context data such as location and upcoming calendar event data to improve the overall accuracy of emotion detection. This system takes into account three points of data (1) Presence of affective keywords in text - this context data is examined as historical data; (2) General knowledge context data of the subject which involves the collection of religion, social information from the subject prior to using the app; (3) How the subject emotionally perceives their surroundings - this involves using a case-based methodology. This study combines a text and context-data to interpret emotion. It considers situational data and user's current perception and historical data to reach a concluded emotion.

Analyzing text is a common way to infer emotion however auditory communication also consists of meaningful emotional data. Na Yang and Arjmand Samuel developed a system which recognizes emotion based on the way people speak [14]. Their system, implemented on a Windows Mobile phone, uses signal processing methods to extract speech features. Logistic Regression is the Machine learning algorithm used in the emotion prediction aspect of this research. Data is trained using the Prosody Database. In this study the speech data was combined with physiological and environmental context data such as an accelerometer and location in order to improve the accuracy of the system. Its overall accuracy is 71%.

The research work above demonstrates the usefulness of context-aware applications as it relates to emotion. The presented social middleware not only focuses on extracting data which points to a specific emotion it also focuses on extracting data which points to a specific social behavior or activity.

4.4 Social Context Awareness

Social context is information that sheds light on a person's social state. It is determined by how a person's physically and emotionally interacts with another person. In past research emotion recognition has been approached by analyzing how people emotionally react to others. Rana El Kaliouby et.al researched the effect of temporal facial-expression context on emotion recognition [15]. Their experiment involved investigating the effect of facial-expression cues on the recognition accuracy of basic and complex emotions. It aimed to understand the extent of how temporal context impacts the recognition of emotions from facial expressions. It was discovered that small amounts of temporal facial expression context data had an obvious effect on recognition accuracy of complex emotions but not for basic emotions. In this study it's clear how a facial expression, a facet, of social context can be indicative of emotions.

Varol Akman of Bikent University wrote a paper titled Rethinking context as a social construct [16]. This article highlights the effectiveness of social sciences stance as it relates to a better understanding of context. He used examples from literary theory to show that interpretation is possible only within shared contexts. One particular model highlighted in Akmans paper was Dell Hymes SPEAKING Model. This model takes on the cultural approach of communication. The acronym stands for :Setting and Scene, Participants, Ends, Act, Keys, Instrumentalities, Norms, and Genre. Setting and Scene refers to the time and place of the social interaction. Participants represent the individuals involved. Ends, refers to the purpose of the social interaction. Act, is the order of the events within the interaction. Keys are the cues which establish the tone. Instrumentalities are the style of speech. Norms are the social rules. Lastly, genre represents the type of categories the social interaction falls under (ie, prayer, lecture). While this model targets literary discourse it is reflective of the social interactions that occur in everyday life. The SPEAKING model highlights the necessary elements to produce a full understanding of a social interaction. Varol Akman suggests these elements are the basis of social context.

A. Pentland, built a socially aware system which predicted social outcomes [6]. This system measures a set of nonlinguistic social signals, like engagement, mirroring, activity level, and stress, using tone of voice over sixty second periods. The goal was to quantify the subject's attitude rather than their internal state. In this system, social context is considered to be the identity of participants in the subject's immediate presence. There are several methods to determine social context, including Bluetooth-based proximity detection, infrared (IR) or radio-frequency (RF) tags, and vocal analysis. This system yielded a 90% accuracy for determining behavioral outcomes such as which couples would exchange phone numbers at a bar; who would exchange business cards etc. It is apparent from this research that social context is an effective tool to determining behavioral outcomes.

Below is a taxonomy chart which compares the different types of context applications from previous works. This project extends on user context to make systems or applications more socially intelligent by providing social context as it relates to AOB.

Table 1 Context Taxonomy Chart

Type of Context	Examples	Use
Computing Context	• network connectivity • communication cost, communication bandwidth „ • nearby resource	Improving business applications
Physical/environmental Context	• Lighting • Sound • Temperature	Improving accessibility of rooms, traffic
Time Context	• Day • Week • Month	Creates historical context.
User Context	• User Profile • User Location • Social Situation	Improves social intelligence of systems. *-AOB social context*

CHAPTER 5: SOLUTION

In order for emotion detection/prediction systems to have accurate results it must consider situational context. One aspect of situational context is social context. This thesis presents the design of a middleware which outputs a value that represents a person's current social state. This middleware involves 3 information elements, physiological data, GPS data, and text analysis. Each of the elements will uses specific services and or machine learning models to produce a final value representative of the user's current social state.

5.1 Gathering Requirements

It is necessary to strategically collect requirements to get an understanding of the type of data needed to produce the intended output. The goal of the middleware is to collect social context data associated with AOB and output a value which represents the user's social state. Veterans are the target user for this middleware. For this project requirements were derived from informal focus groups. These sessions involved discussion between an interdisciplinary team of psychologists, computer scientists, and veterans. The veterans were asked open ended questions which helped to shed light on what an AOB looks like, when it occurs, why it occurs and who it occurs with. Below in figure 1 is a picture of notes taken from veterans as they discussed the AOB phenomenon. In this particular discussion veterans were asked what emotions they feel prior to the angry outburst. Their response provided us with information which allowed us to deduce the type of data needed to accurately assess their social state as it relates to

AOB. It was learned from these informal discussions that text, GPS, and physiological data are significant factors in determining a veteran's social state.

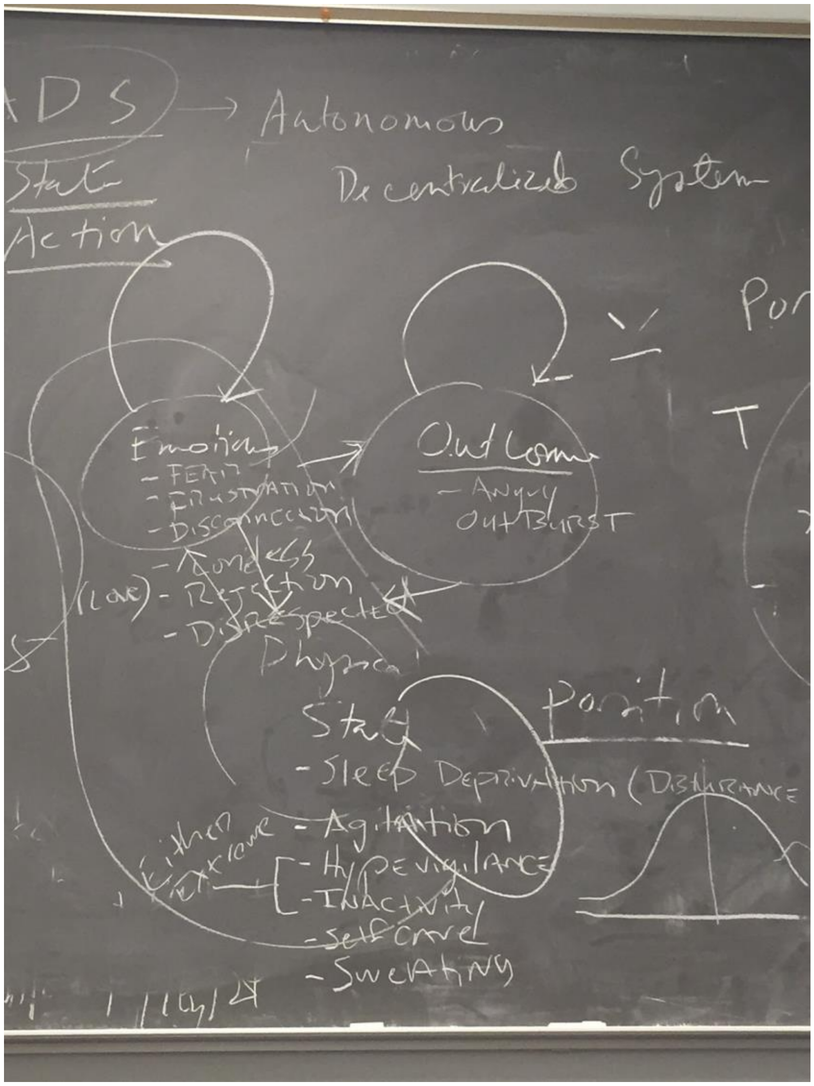

Figure 1. Notes from informal focus group

5.2 Components

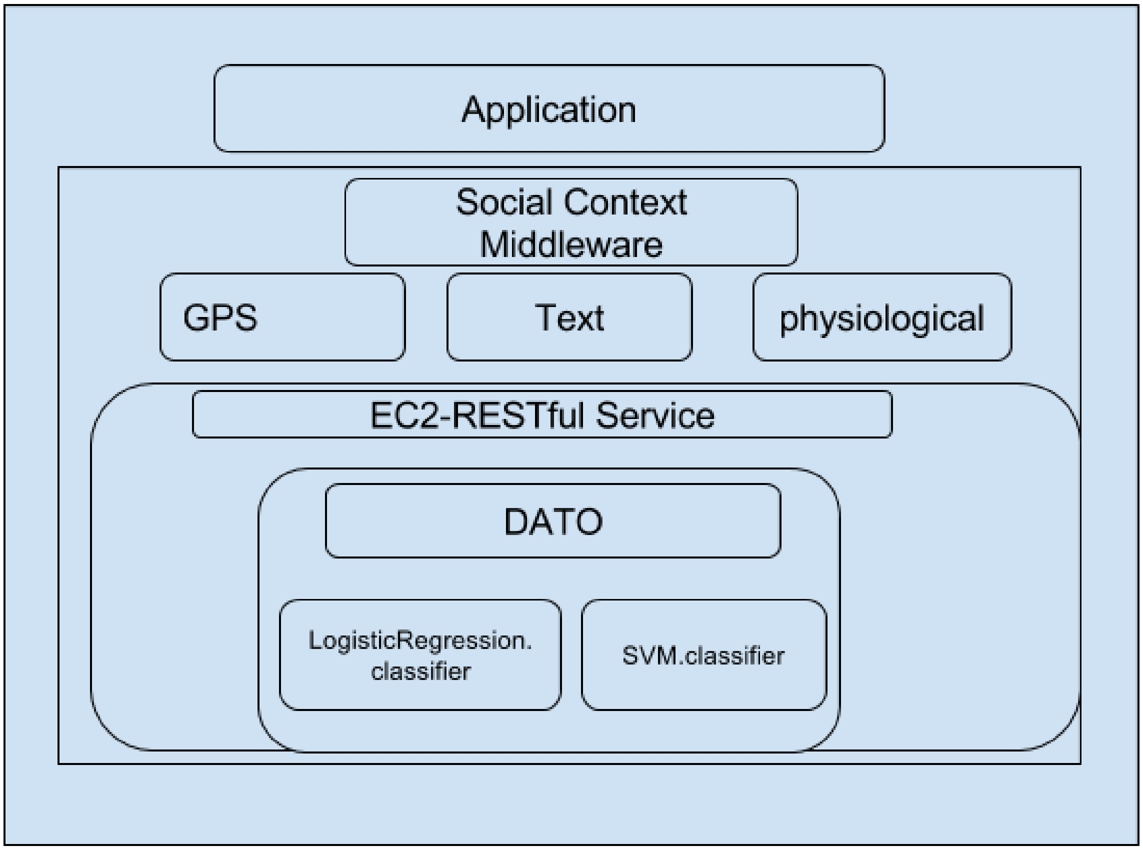

Figure 2 Middleware framework

The physiological data of a person is an appropriate indicator of a person's social environment. A person's pulse, blood pressure, etc. can help to distinguish whether the user is in a fun, relaxing, exciting, or stressful social environment. The middleware will use the physiological data to calculate a value which will be used in the final calculation of the person's social state.

GPS data is critical when determining a person's immediate social state. For example, users who have history with drug and alcohol abuse should not be near neighborhoods that have a high number of drug related crime reports. Literature points to the fact that communities characterized by poverty, inequality and socioeconomic disadvantages have an increased risk of having negative outcomes including recidivism [17]. This middleware design will use GPS data to produce a value representing the geographical element of the social context of the user.

Text-analysis is another component of the middleware. This component will analyze the text of the user’s text messages, emails and social media updates to determine the person's social state. When people are frustrated, angry or depressed it often shows up in the way they communicate with others. The middleware will use an algorithm to produce a value that represents the communication element of the user’s social context.

CHAPTER 6: IMPLEMENTATION

6.1 Study 1 – GPS Component

This section discusses how high risk environments impact a person's social state. The methods of using the Android Location Manager and the calculation of the social context of this component are also discussed in this section.

6.1.1 High Risk Environments

GPS is the first entity of our social context- middleware. Location is an important element of social context. It assists in the characterization of a situation [18]. A user's location determines how they interact with a person, place, or object. Location information brings awareness to the user's social environment. When developing a social context-middleware for a specific group, such as at-risk veterans, it is critical to factor in social settings that are likely to cause a target outcome. In our case we are developing a social-context middleware for at risk veterans. It is important to consider locations that we consider risky behavior triggers [17]. There are two major high risk settings that our system will consider ;(1) General high risk settings,(2) personal high risk settings.

6.1.1.1 General High Risk settings

General high risk settings are social environments which are known to generally increase the likelihood of someone engaging in at-risk behavior. The goal of our social-context middleware is to generate a value that informs an angry outbursts/ crisis detection system of the user's current social state. Veterans that find themselves in crisis situations often suffer from PTSD. Veterans, particularly those who have suffer from PTSD, are more likely to engage in risky behaviors. Locations such as neighborhoods with high crime reports, liquor stores, casinos all fall under high risk areas [17]. Heavy alcohol consumption, illegal drugs, and gambling are common outlets for at-risk veterans.

6.1.1.2 Personal High Risk environments

Personal high risk environments are user specific social environments that increase the likelihood of the user engaging in at-risk behaviors. These social settings can range anywhere from the users mother's house to a specific bank in their neighborhood. For example a user may not have a good relationship with their mother. Every time the user has a conversation with their mother they feel small or unaccomplished. As a result they get into a heated argument making this location a high risk setting. These locations are optionally set by the AOB or crisis/detection systems as a form of preprocessing data. Once these personal high risk settings are specified our social context middleware will use the information when evaluating the user's social context.

6.1.2 Methods

The middleware uses Android GPS location Manager API to track the user's location. This API provides the service of updating software on the device's geographical location. It also sends an application-specified Intent when the device enters the proximity of a particular geographical location [19]. Creating a functioning Proximity feature involved referencing Android Location Manager Class in the code. This allowed the use of services such as requestLocationUpdates which notifies the middleware whenever the user changes location. This service allows the calculation of the distance between the user's current location and the area of interest. The Proximity Alert service is used in the middleware to set specific longitude and latitude coordinates along with a radius of the areas of interest. The combined use of these services allows for acute tracking of the user for the purposes of gaining social context.

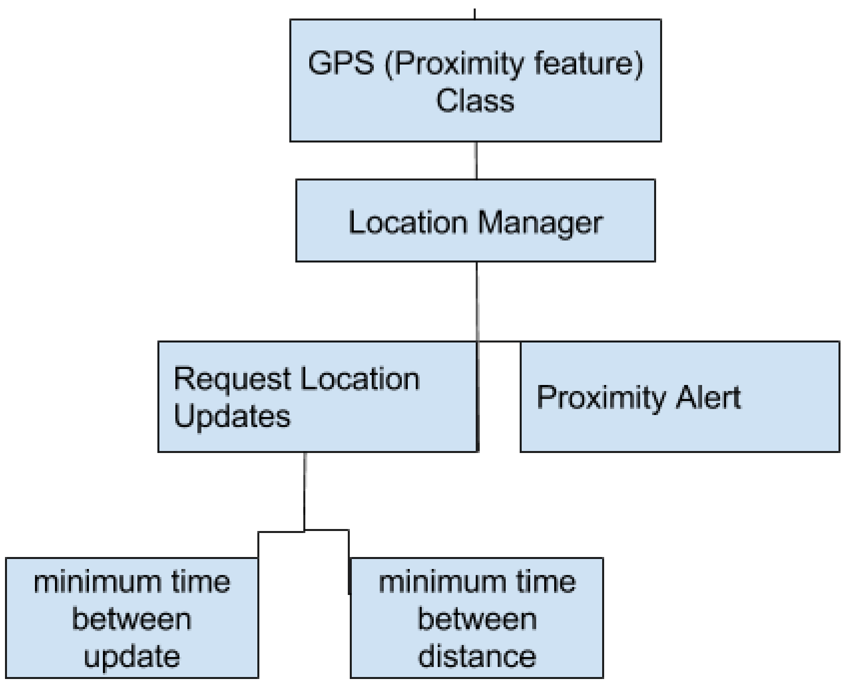

Figure 3 Android Location Services Framework

6.1.2.1 Setting Parameters

When developing middleware which utilizes the android GPS service there are a few parameters that need to be set based on the target user. In our case we are tracking location of an at-risk veteran using proximity service. Time and distance are parameters that have to be sensibly set so that the frequency of location updates are at a minimum to conserve battery life. Each location update requires power from GPS, WIFI, Cell and other radios which can be a strain on the battery. Our system updates the user location every 5 minutes or whenever a user has traveled 503 meters.

6.1.2.2 Calculating Social Context Value

Developing a social-context middleware for at-risk veterans involves considering environments that may trigger at-risk behavior. After many round table informal discussions with veterans it has been discovered that alcohol abuse and drug use are common for those struggling with civil reintegration. In one discussion a veteran revealed that when he was using drugs he began to burglarize homes to pay for drugs. Neighborhoods with high crime reports are high-risk environments because they may compel an at-risk person to engage in drugs, theft, or other criminal activity [17]. Our social-context middleware focuses on at-risk veterans in the city of Milwaukee WI. Based on our findings on neighborhoodscout.com the most crime ridden areas in Milwaukee are listed in the chart below. We have set the proximity points in our code to the locations in the chart.

Table 2 Crime ridden neighborhoods in Milwaukee

neighborhood	longitude	latitude
Atkinson avenue/14th street	43.089244	-87.928317
N Teutonia Ave / W Chambers S	43.073318	-87.931036
E Keefe Ave / N Holton St	43.082255	-87.905021
W Center St / N 16th St	43.067988	-87.932557
Burleigh St / W Roosevelt Dr	43.075621	-87.987238
N Milwaukee St / E Juneau Ave	43.045988	-87.907259

The social context value for the GPS component of the middleware ranges on a scale of 0 to 6. 0 indicates that the user is in a neutral environment, they are not near an high risk location. The closer a person gets to the high risk area the more the social context value will increase. If the user is within 1 miles or closer of the high-risk neighborhood or location the social context value will return a 6. This number indicates that the user is near a location where they are more likely to engage in at risk behavior. Figure 4 is a representation of the calculation of the social context value.

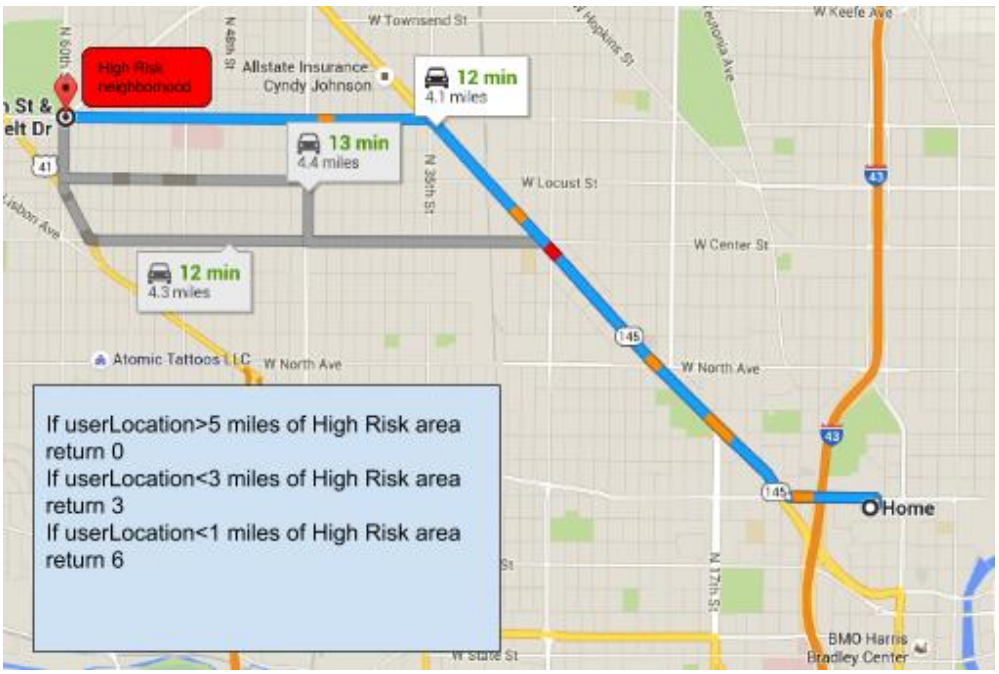

Figure 4 Representation of social state algorithm for GPS

6.2 Study 2 – Text Analysis Component

This section discusses the pros and cons of selecting linear and logistic machine learning classifiers. It also reviews the process of learning weights of specific words, and determining the accuracy and error of the model. The specific methods of the implementation of this component are then discussed toward the end of this section. Lastly the results of the developed model and calculation of social context value for this component are presented.

6.2.1 Choosing a Classifier – Linear Classifier

A linear classifier can be used as a simple threshold classifier. In the case of a text analysis system the positive words and negative words would be considered in the sentence. One approach to a simple linear classifier is to use the word count of text. If the number of positive words is greater than the number of negative words then we can conclude that we are observing a positive sentence. For example: Today was great, the game was awesome, but the weather was terrible. Here there are two positives and one negative. (great, awesome, and terrible). The sentence could be labeled as having a positive sentiment. There are some things to consider when taking this approach. (1)Where does the list of positive and negative words come from? (2)What are the degrees of sentiment or weight of the words? How are the words weighted? For example: Is the word amazing better than the word great?(3) Single words are not enough. For example: Today was good. Today was not good. The first two issues are addressed by using a machine learning classifier. Number 3 is addressed by more elaborate features.

A simple linear classifier takes a list of words and adds weights to them based on the algorithm used in the training dataset. Table 3 includes words with a hypothetical corresponding weight.

This shows how some of the weights gradually decline as the sentiment become more negative. Some words get a weight zero because they are not sentiment indicators. Given these weights a linear classifier would calculate a score. In the case that the input sentence is “Today was great, the game was awesome, but the weather was terrible”. The score of the input sentence would be X = 1.2+1.7+-2.1 = 0.8. Since the score is greater than 0 it is a positive sentence. If the score was less than 0 it would be considered a negative sentence. A Linear classifier such as this one considers the weight of the sum of the input.

If the sentences are plotted based on the score it would yield a graph similar to the graph in figure 5. Below the threshold are sentences that scored positively and above the threshold are sentences that scored negatively. The red line is the linear decision boundary. This separates the positive sentences from the negative.

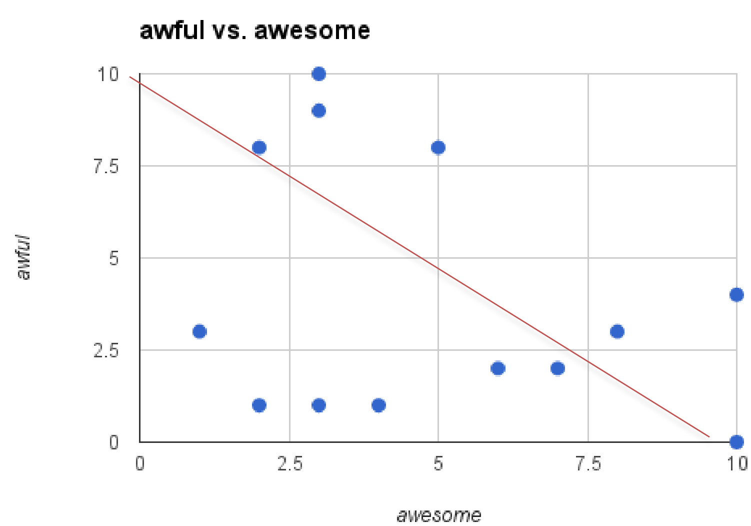

Figure 5 Example of Graph of sentiment text-analysis

Simple linear classifier
Score(x) = sum of weighted words in a sentence
If Score(x) >0 : $\hat{y}$ is a positive sentiment
else:
$\hat{y}$ is a negative sentiment

Figure 6 Linear classifier algorithm

6.2.2 Choosing a Classifier - Logistic Classifier:

Logistic Regression classifiers can predict the probability of an outcome that only has two values (ie dichotomy). In our text analysis feature we distinguish between negative and positive sentiment {0,1}. The words in the sentences are also given weights similar to Table 3. Instead of having a threshold of zero like the above linear classifier, this classifier uses a probability of the text being positive. The probability of the number being 1(positive) is far more reflective of the sentiment of the text than a threshold score.

Logistic regression

$$h_\theta(x) = g(\theta^T x)$$

$$g(z) = \frac{1}{1+e^{-z}}$$

Equation 1 Logistic Regression

Word	Weight
Good	1.0
great	1.5
awesome	2.7
bad	-1
terrible	-2.1
awful	-3.3
restaurant, the, we ,where,etc	0.0

Table 3 Hypothetical weights for example sentiment text analysis

6.2.3 Logistic Regression Classifier vs Linear Classifier

6.2.3.1 Skewed Decision Boundary Line

Linear regression classifiers are not always ideal or appropriate classifiers. A linear classifier is a simple threshold classifier and thus may not always hold up well when applied to an enormous amount of data. Linear regression uses the best fit line as a threshold for determining the output. Given an outlier data point the best line may be skewed in such a way that causes a worse hypothesis or prediction.

6.2.3.2 Output Values Greater or Less than Target Range

Linear classifier gives an output of exact numbers. These numbers may be outside of the target range. In our case we will be working with binary classification {0,1}, negative or positive respectively. The linear regression model can output values much larger than one or less than zero even when the training values are 0 and 1.

6.2.4 Learning the Weights

In order to learn the weights of the words in a sentence the classifier needs to be trained. To do this a dataset which involves positives and negatives sentences are used as input data .The data is then split into a training set and a test set. Next, the training set is fed into the classifier and the logistic algorithm learns the weights of words.

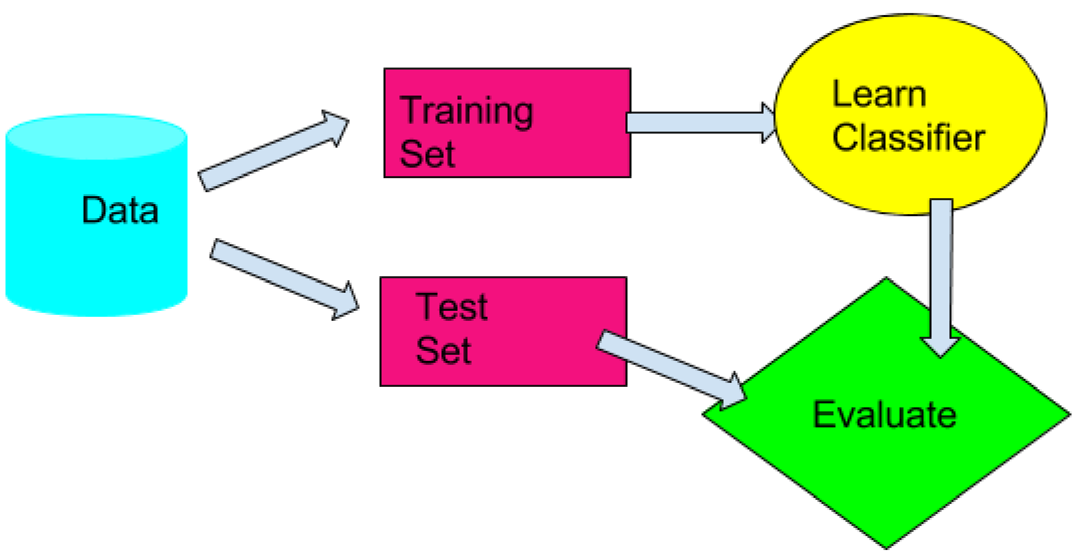

Figure 7 work flow of classification process

6.2.5 Determining the Accuracy and Error

Calculating the error involves feeding the classifier the test data. The classifier then outputs correct or incorrect results. For example the input data will be a sentence and have a sentiment ie. (Sentence=Today was great, Sentiment= positive).The sentiment is hidden from the classifier. The classifier is only able to see the sentence. Let's say in this case the classifier correctly classified the input data. In another case the input data might be (today was ok, negative). If the classifier incorrectly classifies the data as positive we will then have 1 correct and 1 incorrect. This process would continue for every sentence in the data set. To measure the error the numberof mistakes are divided by the total # of test sentences. To calculate accuracy the # of correct predictions are divided by the total # of sentences. The best possible value for error is 0 and the best possible value for accuracy is 1.

6.2.5.1 Determining Good Accuracy

It is important to consider what good accuracy is for any classifier. In this case, to determine the accuracy of the text-analysis classifier a base case is used. The simple scenario of guessing whether or not the sentence is positive or negative is examined. Because this case deals with a binary classification our guess has the a .50 percent probability of being correct. For k number of classes we would use 1/k as the equation to calculate the accuracy of our guess. The classifier should be substantially better than the accuracy of our random guessing or the classifier is considered pointless. The text-analysis feature is veteran specific. The methodology used to determine the sentiment of the text involves counting the number of curse words. The use of profanity is a part of veteran culture. In the case that a veteran uses profanity in 90% of their sentences if our classifier classifies every sentence as negative our system may still yield 90% accuracy. This is due to class imbalance which means one class is more common than others. In this scenario our 90% accuracy is not meaningful. This is something that is addressed later on. The goal is to develop a classifier that is able to beat simple base line approaches such as random guessing and majority class cases.

6.2.6 Methods

The methods of the implementation of the text analysis tool are presented below. The methods of developing this component involve selecting a machine learning tool, developing a classifier, and finally training the model with selected features.

6.2.6.1 Machine Learning Tool

In the implementation of the text analysis feature the python programming language and the IPython Notebook is used. Ipython notebook is an easy interactive environment for programming in Python.

A powerful python package called Graphlab Create is also used [20]. This is a scalable machine learning library for Python. It also includes the SFrame, which is great for data manipulation [18]. SFrames allows for the manipulation of large data sets because it is not limited by memory issues.

6.2.6.2 Developing the Classifier

The data set used for the development of the model encompasses positive and negative sentiment text to create and test the sentiment classifier. The texts in the dataset are simulated messages. The texts with positive sentiment were developed based on a common understanding of what is considered positive. The texts with negative sentiment were created based off of texts found on the internet that are associated with acts of suicide or other crisis events. In the beginning there were two columns in the original data set labeled message and rating. The first column labeled message contained text. The second column, rating, was the positive or negative rating of the text. Once the csv file was imported into graphlab the first step was to develop a word count vector for each message using graphlab text analytics function. This created a column in the data set which contained the word count of all the words in each message. The next step was to develop the sentiment classifier. This process involved data engineering. The 'rating' column was used to define a sentence as positive or negative. The ratings ranged from 1-5. It was decided to ignore all ratings of three because it lands in the middle of the positive and negative ratings. In order to do this the data was filtered so that the only observable data were messages that did not have a rating of 3. This step was necessary because the rating of three did not offer meaningful data. To define sentences as positive or negative another column in the data set called sentiment was created. If the message rating was greater than or equal to 4 the sentiment column was populated with a 1 and if the product rating was less than 4 it was populated with a 0.

6.2.6.3 Training the Data

Once the data had been defined as positive or negative sentiment, the next step was to train the data. Training the data involved taking the dataset and randomly splitting the data, 80% for training and 20% for testing. Curse words are often indicators of negative or angry sentiment while words like 'great' and 'awesome' have more positive connotations. A function was created to count the number of times a curse word appeared in the message. This function was used to populate a column of word counts for the following words {fuck, life, love, never, sucks, demons, damn, hell , kill, shit, great}. These words were selected because of their negative and positive connotations.

message	rating	sentiment	word_count	predicted_sentiment	fuck	life	love
Today was a great day	5	1	{'a': 1L, 'great': 1L, 'was': 1L, 'day': 1L, ...	0.998436171592	0	0	0
I love you	5	1	{'i': 1L, 'you': 1L, 'love': 1L} ...	0.995032856454	0	0	1
you are the best	5	1	{'the': 1L, 'are': 1L, 'best': 1L, 'you': 1L} ...	0.997038919129	0	0	0
I want to kill myself, Life isn't worth it ...	1	0	{'life': 1L, 'anymore.': 1L, 'everyone.': 1L, ...	0.000753499972446	1	1	0
I want to die	1	0	{'i': 1L, 'die': 1L, 'want': 1L, 'to': 1L} ...	0.00314737704611	0	0	0
Life sucks	2	0	{'life': 1L, 'sucks': 1L}	0.00267932227117	0	1	0
Fuck Life	1	0	{'life': 1L, 'fuck': 1L}	0.00515561844538	1	1	0
I love life	5	1	{'i': 1L, 'life': 1L, 'love': 1L} ...	0.992458503138	0	1	1
You are a shitty person	2	0	{'a': 1L, 'person': 1L, 'you': 1L, 'are': 1L, ...	0.00300879895814	0	0	0
My demons won today	1	0	{'won': 1L, 'my': 1L, 'demons': 1L, 'today': ...	0.00151547863159	0	0	0

Figure 8 Spread sheet containing text analysis data

The goal was to use the training data to create a model which classified the messages by analyzing the sentiment column of 1's and 0's and their corresponding selected word features. The graphlab logistic classifier, training data and my selected word features were used as parameters to build the sentiment model. It is part of veteran's culture to use curse words. This cause the majority class case discussed in section 6.1.5.1. To combat this problem I added more features.

6.2.6.3.1 All Caps and Punctuation

Selecting features for text analysis should not only be based on a specific outcome, in this case Angry outburst, but they should also be selected based on the target user(veterans). This text analysis feature is to be implemented in a middleware that aims to detect angry outburst in veterans. Cursing is a common part of veteran culture. Thus foul words may not always be indicative of anger or other emotions. Hence it is critical to add more features as a consideration to this point. Words that are in all caps often represent that a person is yelling or emphasizing a point. Words in all caps are frequently coupled with hostility or frustration. Punctuation is also very indicative of emotion. Multiple periods, exclamation and question marks can hint toward frustration or anger from the user. The number of words in all caps and punctuation count were also added as features to the text analysis model.

6.2.6.4 Results

In the first iteration of training the model the selected words feature, which contained words from the input data that had positive and negative connotations, was the only feature used. Using only the selected words as features yielded a result of 83% accuracy. This was not bad but given a larger data set this accuracy might not hold. After adding new features such as number of punctuation and number of words with all capital letters the model yielded a result of 87% accuracy. The accuracy of the model increased by 4% . Using a priori modelling of how humans behave has created meaningful, incremental increases in the accuracy of this model.

6.2.6.5 Calculating the Social Context Value

Once the model has reached the goal accuracy it will be used in the middleware to assist with the calculation of the users overall social state. If the user's text data yields an output from our model between 0 and .50 then a value of 6 will be returned. If the model outputs a value between .50 and .80 a value of 3 will be returned. Lastly if the model outputs a predicted value over .83 a value of 0 will be returned. This value will be combined with the value from the physiological feature and GPS feature to arrive at a final social context value.

6.3 Study – 3 Physiological Component

This section discusses the development of the physiological component. The sensors used in the Empatica E4 are discussed in the beginning of this section. The specific methods of collecting data, choosing a stimuli, choosing a classifier and developing a classier are then discussed toward the end of this section. Subsequently the results of the developed model and calculation of social context value for this component are presented.

6.3.1 Developing the Physiological Component

This social context middleware uses the Empatica E4 [21]. The objective is to fuse the physiological sensor data from this device to find distinguishing output that allows us to determine the social state of the user. When choosing physiological features it is best to extract features which have a strong correlation to the target outcome of the middleware. The middleware functions to detect social-context as it relates to AOB. This section describes the physiological data that correlates to anger. The development of the physiological element of the middleware involved four steps - data collection, feature selection, and classification.

6.3.1.1 Electrodermal Activity Sensor

The physiological sensor data is extracted using the Empatica E4 wristband. This wrist band uses an Electrodermal Activity Sensor(EAS). This is used to measure the sympathetic nervous system arousal which extracts features related to stress, engagement, and excitement. Hypertension is heavily associated with the sympathetic activation in the human body [22]. Sympathetic activation increases when one experiences excitement or anticipation. It also increases as a result of physical, emotional or cognitive stress making the EAS a good source for estimating potential AOB from the user.

Empatica uses the skin to measure sympathetic activation. Skin is the only organ that is completely innervated by the sympathetic nervous system (and not affected by parasympathetic activation).The sympathetic activation can be monitored by subtle electrical changes across the surface of the skin. Electrodermal activity(EDA) is typically measured in microSiemens. The EDA is a combination of two components, the skin conductance level (SCL) and the skin conductance response(SCR). The E4 records the SCL while keeping a sufficient sensitivity level to distinguish the SCR under any condition. This high resolution data is used to measure the sympathetic activation and autonomic stress of the user. The E4 device can measure conductance in the [0.01, 100] uS range . It has a default sampling rate of 4Hz and the digital resolution is 1 digit per 900 Pico Siemens [23]. The high resolution data will provide us with three levels of zoom (5 hours, 30 minutes, 4 minutes) . The high dynamic range and sensitivity of the sensor are used in the analysis process to cardiovascular activity with AOB. Real time user data from this sensor will be streamed from the Empatica to our middleware to be analyzed by our model.

6.3.1.2 Photoplethysmography Sensor

Heart Rate can also be an indicator of anger [24]. Light can be absorbed by different biological substances such as pigments in the skin, bone, and arterial and venous blood. Changes in blood flow occur mainly in the arteries and arterioles (but not in the veins). Arteries typically contain more blood volume during the systolic phase of the cardiac cycle where the ventricles contract and pump blood to the arteries than during the diastolic phase where the heart fills with blood. PPG sensors optically detect changes in the blood flow volume by detecting the changes in the light intensity through the reflection from or transmission through the tissue [25].

The E4 wristband uses a PPG sensor to measure blood volume pulse (BVP). This gives critical information about the heart rate, heart rate variability (HRV), and other cardiovascular features.

6.3.1.3 Linking to Anger

The EDA and PPG sensors are used to measure nervous system arousal associated with stress engagement, excitement and cardiac information. The goal is to fuse the data extracted from these features to discover a correlation with anger. A machine learning algorithm is implemented to train a classifier. In order to classify data a subject used the E4 event marker feature. This provided us with a dataset that was useful in the data training process. This is discussed later on in the chapter.

6.3.2 Empatica platform

6.3.2.1 Bluetooth

The Empatica E4 is able to collect real-time data using a Bluetooth 4.0 (Bluetooth Low Energy – BLE) interface and recording mode using its internal flash memory. This device allows for collection of quality continuous data representing the two main branches of the autonomic nervous system stress response outside of the lab which is ideal for research. The bluetooth connection is established using the function deviceManager.connectDevice(device). Once a connection is established the middleware collects data from the PPG and the EDA sensors. The data will be saved and stored as a csv file where it will be sent to a cloud server to be analyzed by a DATO support vector machine based model.

6.3.3 Methods

Creating a machine learning based model requires collecting a set of data and splitting it up to train and test to improve its accuracy. Below describes the methods of collecting physiological data for sentiment analysis.

6.3.3.1 Collecting data

Mohammad Soleymani did a study on EEG-based emotion recognition. Four minute movie clips were used as stimuli to arouse human emotion [23].In this experiment the primary focus was on positive and negative emotions. This is due to the fact that emotions are very complex. Often time's people cannot accurately describe or distinguish their emotions [26]. For this reason each of the movie clips were classified into positive or negative sentiment.

6.3.3.2 Choosing a clip

Choosing a movie clip for data collection is critical. The videos must be chosen based on the desired output. The goal of the middleware in this project is to detect angry out bursts in veterans. Experimental studies show that anger is associated to increased cardiovascular activities [27]. Fear and Anger evoke similar fight or flight physiological reactions. Thus fear evoking movie clips were selected as our stimuli. Comical clips were selected to evoke happy emotions in our subjects.

The physiological feature in this project uses a support vector based model to determine whether a person is experiencing negative or positive emotions. During the data collection process a subject was observed watching several scary and funny movie clips. The length of each clip ranged between 3-5 minutes. The subject wore the Empatica E4 device as they watched each clip. Once the physiological data was collected the data engineering process began.

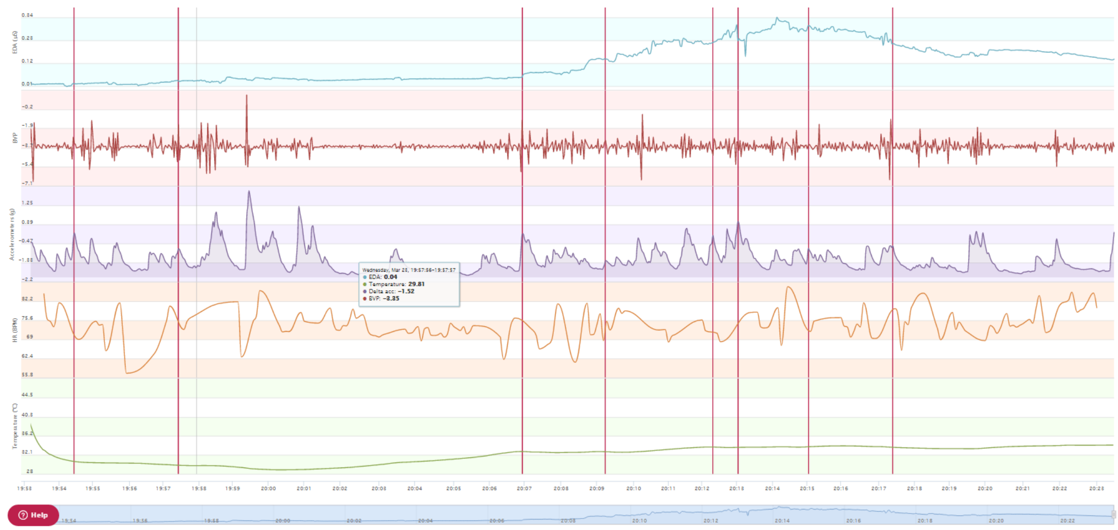

Figure 9 physiological data of subject

6.3.3.3 Choosing a classifier

Support Vector Machines (SVM) finds the optimal separating hyperplane which separates different label sets. In this case it is the goal to distinguish positive and negative sentiment. The physiological sensor data feature of the middleware implements SVM because it is a robust classifier. It can be used to classify large and small amounts of samples. This classifier can also be used for both simple and highly complex classification models. Lastly it can be used to implement sophisticated mathematical principles to avoid overfitting [28].

In our SVM model the data is represented as points in multidimensional space and mapped so that the data points from the two classes are divided by linear separator.

Emotions are complex and not always easy to distinguish. Developing a physiological sentiment classifier requires a soft support vector machine algorithm because the data extracted from the sensors may not always be linearly separable. This form of SVM implements a loss function. The loss function penalizes misclassification which means the margin between the positive and negative sentiments are smaller. This is not ideal. The SVM model used from Dato minimizes the loss function below:

fi(θ)=max(1−θTx,0)fi(θ)=max(1−θTx,0)

Figure 10. Loss Function

The goal is to minimize the trade-off between maximum margin and classification error. Delta is our penalty parameter which determines our tradeoff. An intercept term is added by appending a column of 1’s to the features. Regularization is necessary to prevent overfitting by penalizing models with extreme parameter values. The blended equation is shown below in figure 11.

$$\min_{\theta} \sum_{i=1}^{n} f_i(\theta) + \lambda \|\theta\|_2^2$$

Figure 11. SVM Equation

6.3.3.4 Developing a classifier

The Empatica E4 extracts physiological data from the subject which can be exported to individual csv files. Before training the data the csv files for Bvp, Temperature, Heart rate and EDA for both funny and scary observations were consolidated into one sheet. The sentiment column was populated with 1 or 0 indicating positive and negative respectively. These labels were placed by their corresponding data.

6.3.3.5 Results

Once the data was consolidated into one file the data was split into training data and test data. Out of 2499 samples 80% was designated to training data and 20% to test data. Dato's support vector machine classifier was used to construct the physiological sentiment model. The model used temperature BVP, Heart rate and EDA as features. This model yielded 100% accuracy.

Table 4 Evaluation of the model

```
{'accuracy': 1.0, 'confusion_matrix': Columns:
        target_label    int
        predicted_label int
        count   int

 Rows: 2

 Data:
 +--------------+-----------------+-------+
 | target_label | predicted_label | count |
 +--------------+-----------------+-------+
 |      1       |        1        |  140  |
 |      0       |        0        |  134  |
 +--------------+-----------------+-------+
```

6.3.3.6 Calculating Social Context Value

This model will be used in the middleware to assist with the calculation of the final social state. If the user's physiological data yields an output from our model which represents negative sentiment a value of 6 will be returned. This value will be combined with the value from the text analysis and GPS feature to arrive at a final social context value.

6.4 Using the trained Physiological and Text Models

Models for the Text and Physiological elements of the middleware used the predictive services from DATO. Once the models have been developed in graphlab they must be deployed to the EC2 instances so that they can be used within the middleware. In order to deploy the model first an EC2 instance must be configured using the code below.

```
import graphlab as gl

# make sure to replace the following with your own information
ps_state_path = 's3://<your-bucket-name>/predictive_service/ps'

# Create an EC2 config
# You can either specify your AWS credentials using environment variables, or
# set them as arguments to this object's constructor
ec2_config = gl.deploy.Ec2Config(
  aws_access_key_id='<your access key>',
  aws_secret_access_key='<your secret key>')
```

Next the predictive service must be implemented. This process requires consideration of setting several optional parameters it is advised to have at least 3 nodes for cache utility and high availability.

```
# use the EC2 config to launch a new Predictive Service
# num_hosts specifies how many hosts the Predictive Service cluster has. You can scale up and down later after initial creation.
ps = gl.deploy.predictive_service.create(
  name='sklearn-predictive-service',
  ec2_config=ec2_config,
  state_path=ps_state_path,
  num_hosts=1)
```

Once the Ec2 instance and predictive services have been configured the next step is to add the models to expose the trained text and physiological model as a REST endpoint in the newly configured predictive service. This step requires wrapping the model in a python function then adding it to the predictive service code below. This will enable the middleware to use the model to produce an output reflective of the user's social state.

```
def classifyText(x):
  prediction = textModel.predict(x)

  # convert into a json serializable value
  return list(prediction)

# add your predictive function that wraps scikit-learn model
ps.add('classify', classifyText)
```

To use the model in the middleware a client config file must be created using the code below. Once the file is created the model can be queried by either using REST or a python client package.

```
ps.save_client_config(file_path='/tmp/ps_client.conf', predictive_service_cname =
https://models.ubicomp.com).
```

The models developed in this project will be accessed in the middleware to either attain an output which is reflective on the users' social state as it relates to AOB or to update the model based on specific user data.

6.5 Calculating Final Social Context Value

The GPS component produces a social context value between 1 and 6. This value is based on the user's current location. Given the user is greater than 5 miles of a high risk area the component will return an output of 0. If the user is between 1 and 3 miles of a high risk neighborhood, the component will return a social context value of 3. Whenever the user is within a mile of a high-risk neighborhood or near an area they have specified to be particularly stressful for them, the middleware will output a value of 6. This value indicates the user is at a higher risk of experiencing an angry outburst.

Once the physiological data is extracted from the Empatica E4 device it is collected and analyzed using the machine learning model trained in study 3. A value of 0 or 6 will be outputted from this component. If the model returns a value of 1, indicating the user is experiencing a positive sentiment, the social context value of 0 will be returned. This indicates that their physiological data is not reflective of someone who is at-risk of an angry outburst. If the model yields a value of 0, which indicates the user is experiencing a negative sentiment, then the component will out put a social context value of 6. This value indicates that the user is more likely to experience an angry outburst.

The Text analysis component uses a machine learning model trained in study 2 to compute the output of a social context value between 1 and 6. Higher values produced from the model indicate that the user is more at risk of an angry outburst. Given the model returns a value between .80 and 1 the text analysis component will return a value of 0. If the model returns a value between .50 and .80 the component will return a 3. Lastly if the model produces a value between 0 and .50 the text analysis component will return a value of 6. This indicates that the user is at a higher likely hood of experiencing an angry outburst.

The values from the GPS, and text analysis and physiological components will be summed together. The sum is the final value used to indicate the current social state of the user. This value is the output of the middleware. Applications will be able to use this middleware to determine the current social state of the user in order to detect or predict an emotion.

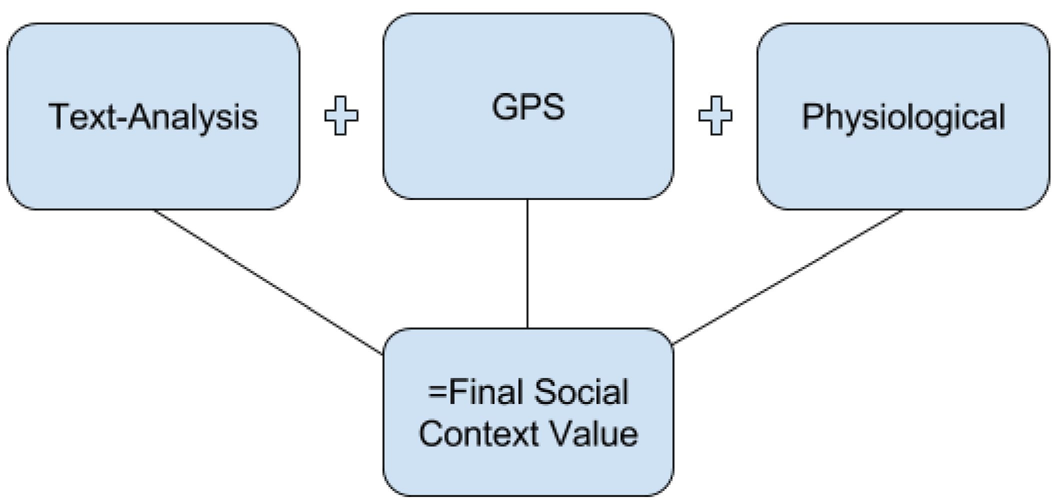

Figure 12 Final Social Context Equation

CHAPTER 7: CONCLUSION & FUTURE WORK

In this project I have defined several guidelines for developing a social context middleware. These guidelines involve methods of attaining requirements for a social context system, methods of developing a GPS proximity feature for high-risk location, and methods of distinguishing between positive and negative sentiment of a subject using physiological and text data. The process of developing machine learning models for the text and physiological features are explained in detail. The text analysis model yielded an accuracy of 87% and the physiological model was 100% accurate.

Our social context middleware has a lot of potential. While the goal is to produce a value reflective of a veteran's social state as it relates to AOB, future versions of the middleware can be applied to a broad range of users who are at-risk of various crisis events.

Currently the middleware is designed to only take in physiological data from the Empatica E4. In the future the goal is to further develop the middleware to handle data from wearable devices that are widely available to the public. Other goals for the future are to improve the overall accuracy of the models by minimizing the false positive/false negative within the middleware models. Currently this project faces the 'cold start' problem which means we are unable to draw any inferences for a specific user because we have not gathered enough information. This can prevent us from reaching a high level of accuracy for a specific user. We can achieve better accuracy by using DATO's update functions which allows us to use data gathered from the user and update the model to be user specific. This is necessary in order to effectively assists AOB/ other crisis event detection systems. Incorporating this middleware into applications will allow for an in depth observations of social context as it relates to anger. In the future there will be enough data gathered to inform psychological science, providing a deeper understanding of the specific behaviors associated with this AOB.

BIBLIOGRAPHY

[1] B. H. M. N. H. a. M. V. M. S. Kaplan, "Suicide risk and precipitating circumstances among young, middle aged, and older male veterans,," *American Journal of Public Health vol. 102 ,* pp. pp. S131–7, 201.

[2] A. M. M. ,. N. C. M. ,. J. C. B. ,. M. A. ,. J. R. &. D. M. B. Christine Timko, " Treatments for Recidivism RiskAmong Justice-Involved Veterans," *ournal of Offender Rehabilitation,,* pp. 620-640, 2014.

[3] "Department of Veterans Affairs, O. National Resource Directory," Ebenefits.va.gov, 2015. [Online]. Available: https://www.ebenefits.va.gov/ebenefits/nrd.

[4] "Pos-rep.com," POS REP., 2015. [Online]. Available: http://pos-rep.com/. [Accessed 25 Feburary 2016].

[5] K. M. C. D. Z. a. W. M. F. Y. Wang, "Social Computing: From Social Informatics to Social Intelligence," *IEEE Intelligent Systems,* Vols. vol. 22, no. 2, pp. 79-83, 2007.

[6] A. Pentland, "Socially aware, computation and communication," *Computer,* vol. vol. 38, no. 3, pp. 33-40, 2005.

[7] N. A. a. R. W. B. Schilit, "Context-Aware Computing Applications," *ieee,* 1994.

[8] G. A. a. D. S. A. Dey, "A Conceptual Framework and a Toolkit for Supporting the Rapid Prototyping of Context-Aware Applications," *Human-Comp. Interaction,* vol. vol.16, no. 2, pp. pp. 97-166, 2001.

[9] J. A. L. Jason Hong, "An infrastructure approac to context-aware computing," *Human-Computer Interaction,* vol. 16, no. 2, pp. 287-303, 2001.

[10] O. S. S. M. a. J. S. W. Yi, "Wearable sensor data fusion for remote health assessment and fall detection," *IEEE International Conference on Electro/Information Technology,* 2014.

[11] A. S. S. R. M. B. a. P. S. D. Rajan, "Health monitoring laboratories by interfacing physiological sensors to mobile android devices," in *2013 IEEE Frontiers in Education Conference (FIE)*, 2013.

[12] M. S. M. P. a. K. K. J. Kim, "Visual Analysis of Relationships between Behavioral and Physiological Sensor Data," in *2015 International Conference on Healthcare Informatics*, 2015.

[13] Y.] Douiji and H. ".-C. I. C. A. s. f. R. E. f. t. Mousanif, "I-CARE: Intelligent Context Aware system for Recognizing Emotions from text," in *Intelligent Systems: Theories and Applications (SITA) 2015 10th International Conference*, 2015.

[14] A. S. Na Yang, "Context-rich Detection of User's Emotions using A Smartphone," Microsoft Research, 2011.

[15] P. R. S. K. E. D. C. Rana El Kaliouby, "Temporal Context and the Recognition of Emotion from Facial Expression," in *Proceedings of the HCI International Conference*, American Psychological Association, 2003, pp. 631-635.

[16] V. Akman, "Rethinking context as a social construct," *Journal of Pragmatics,* pp. 743-759 , 2000.

[17] P. Thad Q. Strom, P. Jennie Leskela, P. Lisa M. James and P. Paul D. Thuras, "An Exploratory Examination of Risk-Taking Behavior and PTSD," *MILITARY MEDICINE,* vol. 177, pp. 390-396, 2012.

[18] G. Biamino, "Modeling social contexts for pervasive computing environments," in *Pervasive Computing and Communications Workshops (PERCOM Workshops), 2011 IEEE International Conference*, Seattle, 2011.

[19] "Location and Maps," Android, [Online]. Available: http://developer.android.com/guide/topics/location/index.html. [Accessed 5 January 2016].

[20] "DATO," 2016. [Online]. Available: https://dato.com/products/create/. [Accessed 10 January 2016].

[21] M. L. D. B. R. W. P. a. S. T. M. Garbarino, "Empatica E3 — A wearable wireless multi-sensor device for real-time computerized biofeedback and data acquisition," Athens, 2014.

[22] G. G. C. G. G. S. Giuseppe Mancia, "Sympathetic Activation in the Pathogenesis of Hypertension and Progression of Organ Damage," *Scientific Contributions,* pp. 724-728, 1999 .

[23] X. W. W. L. C. S. a. B. L. L. D. Nie, "EEG-based emotion recognition during watching movies," Cancun, 2011.

[24] ,. ,. M. G. G. R.-A. R. E. A. S. Neus Herreroa, "What happens when we get angry? Hormonal, cardiovascular and asymmetrical brain responses," *Hormones and Behavior,* vol. 57, no. 3, p. 276–283, 2010.

[25] Y. M. 2. M. S. 1. a. M. Y. 1. Toshiyo Tamura 1, "Wearable Photoplethysmographic Sensors—Past and Present," *Wearable Electronics,* pp. 282-302, 2014.

[26] M. P. G. I. R. a. T. S. H. Z. Zeng, "A survey of affect recognition methods: audio, visual, and spontaneous expressions," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 31, no. 1, 2009.

[27] S. C. S. J. L. K. A. M. T. W. K. A. L. M. Pamela J. Feldman, "Negative emotions and acute physiological responses to stress," *Annals of Behavioral Medicine,* pp. 216-222, 1999.

[28] D. H. Alexander Statnikov*, "A Gentle Introduction to SVM," 14 November 2009. [Online]. [Accessed 15 Febuary 2016].

[29] C. R. K. D. T. T. L. a. F. G. J. o. N. a. M. D. T. L. Hartl, "Predicting high risk behaviors in veterans with posttraumatic stress disorder," *Journal of Nervous and Mental Disease,* vol. 193, p. 464472, 2005.