2012

A Voice-Based Automated System for PTSD Screening and Monitoring

Roger Xu
Intelligent Automation, Inc.

Gang Mei
Intelligent Automation, Inc.

Guangfan Zhang
Intelligent Automation, Inc.

Pan Gao
Intelligent Automation, Inc.

Timothy Judkins
Intelligent Automation, Inc.

See next page for additional authors

Original Publication Citation

Xu, R., Mei, G., Zhang, G., Gao, P., Judkins, T., Cannizzaro, M., & Li, J. (2012). A voice-based automated system for PTSD screening and monitoring. In J. D. Westwood, S. W. Westwood, L. Felländer-Tsai, R. S. Haluck, R. A. Robb, S. Senger, & K. G. Vosburgh (Eds.), *Medicine meets virtual reality 19* (pp. 552-558). IOS Press. https://doi.org/10.3233/978-1-61499-022-2-552

Authors
Roger Xu, Gang Mei, Guangfan Zhang, Pan Gao, Timothy Judkins, Michael Cannizzaro, Jiang Li, James D. Westwood (Ed.), Susan W. Westwood (Ed.), Li Felländer-Tsai (Ed.), Randy S. Haluck (Ed.), Richard A. Robb (Ed.), Steven Senger (Ed.), and Kirby G. Vosburgh (Ed.)

A Voice-Based Automated System for PTSD Screening and Monitoring

Roger XU [a,1] Gang MEI [a], Guangfan ZHANG[a], Pan GAO [a], Timothy JUDKINS [a],
Michael CANNIZZARO [b], and Jiang LI [c]

[a] *Intelligent Automation, Inc. 15400 Calhoun Dr., Suite 400, Rockville, MD 20855, USA*
[b] *University of Vermont*
[c] *Old Dominion University*

Abstract. Comprehensive evaluation of PTSD includes diagnostic interviews, self-report testing, and physiological reactivity measures. It is often difficult and costly to diagnose PTSD due to patient access and the variability in symptoms presented. Additionally, potential patients are often reluctant to seek help due to the stigma associated with the disorder. A voice-based automated system that is able to remotely screen individuals at high risk for PTSD and monitor their symptoms during treatment has the potential to make great strides in alleviating the barriers to cost effective PTSD assessment and progress monitoring. In this paper we present a voice-based automated Tele-PTSD Monitor (TPM) system currently in development, designed to remotely screen, and provide assistance to clinicians in diagnosing PTSD. The TPM system can be accessed via a Public Switched Telephone Network (PSTN) or the Internet. The acquired voice data is then sent to a secure server to invoke the PTSD Scoring Engine (PTSD-SE) where a PTSD mental health score is computed. If the score exceeds a predefined threshold, the system will notify clinicians (via email or short message service) for confirmation and/or an appropriate follow-up assessment and intervention. The TPM system requires only voice input and performs computer-based automated PTSD scoring, resulting in low cost and easy field-deployment. The concept of the TPM system was supported using a limited dataset with an average detection accuracy of up to 95.88%.

Keywords. PTSD, committee-machine, Gaussian Mixture Model

1. Introduction

Post-Traumatic Stress Disorder (PTSD) [1] is an anxiety disorder that can develop after exposure to one or more traumatic events that threatened or caused great physical harm. It has been reported that nearly 10 percent of Iraq veterans have screened positive for PTSD [2]. It is often difficult and costly to diagnose PTSD due to its varieties of symptoms and potential patients are often reluctant to seek help owing to the stigma associated with PTSD. Previous research has shown that several prominent features of speaking behavior and voice sound characteristics are closely related to the severity of patients' mental illness as well as the depression recovery time course [3][4]. Therefore, a low cost and non-intrusive voiced based PTSD monitoring system has the potential to address the aforementioned need.

[1] Corresponding Author: Roger Xu. Intelligent Automation, Inc. Email: hgxu@i-a-i.com

In this paper, we present a voice-based Tele-PTSD Monitor (TPM), which can remotely screen, monitor, and provide assistance to clinicians in identifying and diagnosing a patient at high risk for PTSD. Section 2 describes the conceptual architecture of the voice-based TPM system. Section 3 introduces the PTSD scoring engine. Section 4 presents the TPM prototype. Preliminary experimental results using a limited dataset are described in Section 5. The final section concludes the potential benefits of using such a system.

2. System Architecture

The architecture of the TPM system is shown in Figure 1. The system consists of voice-based TPM software and a Commercial-Off-The-Shelf (COTS) Interactive Voice Response (IVR) System. The TPM software consists of a PTSD Scoring Engine (PTSD-SE), a PTSD scoring web service, a Graphical User Interface (GUI) and a database.

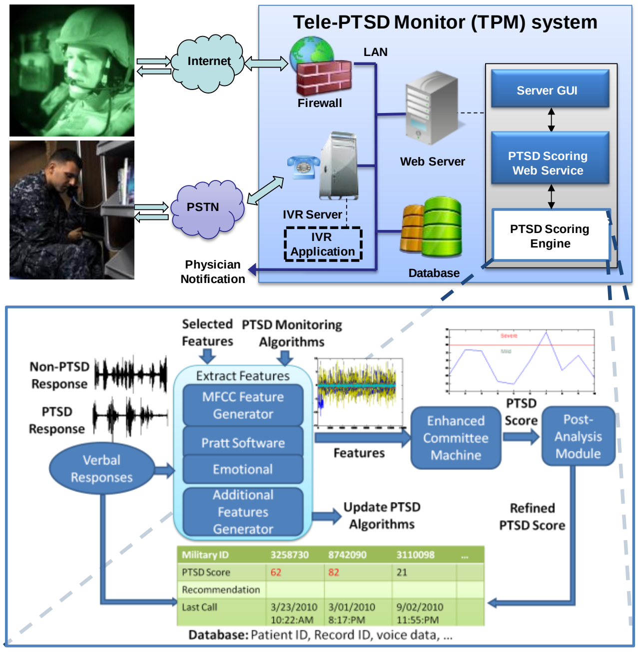

Figure 1. A functional diagram of the voice-based Tele-PTSD Monitor (TPM) system

A soldier (or other service men/women likely to suffer from PTSD) can access the TPM system via a Public Switched Telephone Network (PSTN) as well as the Internet. With PSTN, a soldier only needs to call a phone number to engage with an Interactive Voice Response (IVR) server; while with Internet, the soldier can access the TPM system via a web browser. The PTSD monitoring application on the web server greets the soldier and asks him/her questions that are related to PTSD symptoms and their impact on social and occupational functioning. The soldier's speech contentindependent responses are stored in a database for future analysis by a clinician. At the same time, the voice data is sent to the web server connected to the IVR server. The PTSD scoring web service deployed on the web server invokes the PTSD-SE to compute a PTSD score of the soldier. The scores are also saved in the database and accessible by clinicians. If the score indicates that the user is at risk of demonstrating identifiable PTSD symptoms, the system notifies clinicians to initiate a follow-up consultation via email or text messaging. The server side Graphical User Interface (GUI) software is used to monitor the status of the whole TPM system. The TPM system is low cost and can be accessed through Internet or a phone without additional hardware. It is also field deployable thanks to a common server-client structure.

The voice-based TPM system provides three different capabilities. First, it can be used as a screening tool for initial screening of the patients who may present with symptoms of PTSD. Second, if a patient is diagnosed with PTSD, the TPM system can be used for continuous remote PTSD monitoring and tracking without the needs of costly and frequent clinician visits. Third, although the TPM cannot act as an independent diagnostic system, it can assist psychologists in making better diagnosis by providing objective PTSD scores, free-up valuable time of psychologists, and mine the voice data for further PTSD diagnosis and monitoring.

It is also worth noting that the voice-based PTSD screening and monitoring is done remotely with a confidential connection to protect patients' privacy and confidential information. As such, the stigma of patients can be alleviated without frequent clinician visits.

3. PTSD Scoring Engine (PTSD-SE)

The PTSD-SE is the "brain" of the voice-based TPM system. It processes the voice data by extracting numerous salient acoustic features (e.g., intensity and frequency-based raw or computed values), applies feature selection algorithm to select the most relevant features, computes PTSD scores based on an enhanced committee machine algorithm [5], and performs post analysis for score refinement.

More specifically, we have identified several features that can reflect the acoustic parameters of a PTSD patient. These features include commonly used Mel-Frequency Cepstral Coefficients (MFCC) features, standard deviation of pause and standard deviation of fundamental frequency, among others. Given the input feature sets, a committee machine method can relate these features to mental health states. The basic idea of committee machine [5] is to aggregate the outputs from several PTSD models (committee members) specified by users. Each of the committee members can be trained based on the selected features corresponding to the patient's mental health states (PTSD/non-PTSD). Different algorithms can function as committee members, such as Neural Networks (NNs) [6], GMMS [7], and Support Vector Machines (SVMs) [8]. Finally, a post processing module can be utilized to adjust the PTSD scores based on PTSD score trending information.

4. TPM Prototype

Figure 2 shows the current implementation of the TPM system. It was based on a TCP/IP network which can be directly deployed on the internet or a Local Area Network (LAN). Main functionalities of the TPM system were implemented as web services (C#/.Net) on a Microsoft IIS web server. The way the TPM software works is as follows. The user uses the client GUI software to initiate the process, records the user's speech, and sends the recorded speech to the secure server via Simple Object Access Protocol (SOAP), in the format of a web service call. Upon receiving the call, the PTSD scoring web service on the web server will in turn call the PTSD Scoring Engine (PTSD-SE) to compute the user's PTSD score. Suggested measures based on the PTSD score will be sent back to the client GUI for display. The patient's status and information will also be retrieved from and updated to a database. The server side GUI software is used for the operator to monitor the status of the whole system.

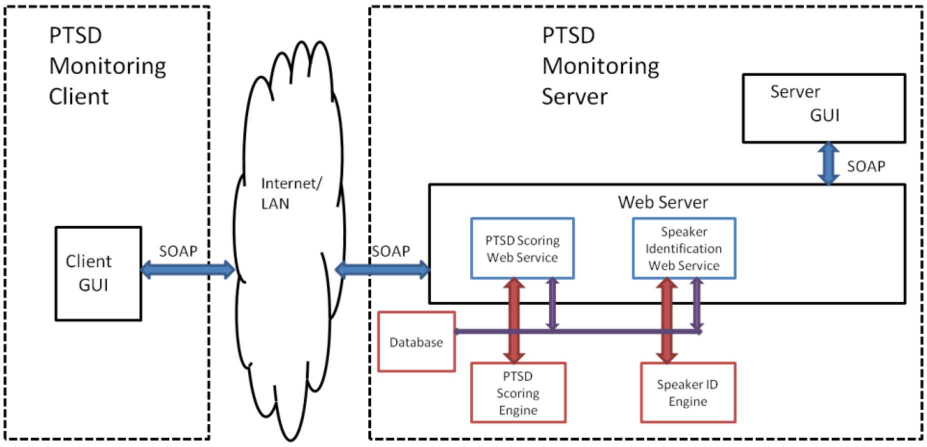

Figure 2. Remote PTSD monitoring system on TCP/IP

The hardware of the PTSD monitoring client is flexible. It can be a PC or WIFI enabled cell phone running the client side software. Depending on the computing power of client hardware, we can relocate some processing tasks to the client side for minimizing network traffic and latency. In this research, we have implemented the TPM client on both a PC and an android cell phone. As an example, the android phone based client is shown in Figure 3.

With the Android-based client, audio is sampled (16-bit raw PCM data) from the microphone and is streamed at 16 kHz to the PTSD monitoring Server. The user can configure the IP address, control port and audio port of the server on the cell phone. On the server, after receiving data from the client, it will first utilize a Voice Activity Detection algorithm (VAD) to remove silence segments. The remaining pure voice data will be used to invoke the PTSD-SE. The PTSD score generated by the PTSD-SE is displayed on the GUI for the algorithm verification purpose. And if the score is higher than a predefined threshold, the voice data is flagged as PTSD data for clinician follow-up. If this is the case, the color of the score bar on the GUI will then be changed from green to red for easy identification. In the GUI, the history of past PTSD scores and the current waveform of voice data can also be displayed. As an example,

Figure 4 shows a score snapshot when a non-PTSD audio clip (left) and a PTSD audio clip (right) are tested.

Figure 3. Android phone-based TPM

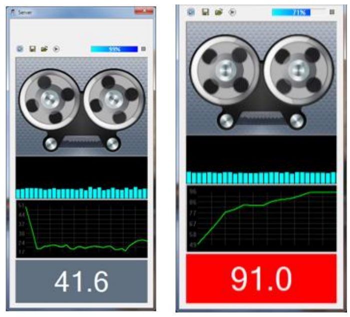

Figure 4. PTSD monitoring: non-PTSD (left); PTSD (right)

5. Performance Evaluation

To evaluate the performance of the PTSD-SE, we used a criterion-based search strategy to acquire publicly available video/audio of soldiers discussing their PTSD. We initially identified 15 public available videos, which meet certain inclusion criteria (e.g., US soldiers who are English speakers without noticeable accents or dysarthria). Among the acquired video clips, we further down selected 5 clips based on video recording quality. Similarly, we have selected 5 public available video clips containing speeches by 5 matched non-PTSD individuals. Furthermore, voice segments were stripped from these video clips using Soundflower (intermediary software for audio extraction between software programs) [9]. Preconditioning was performed to enhance the usability of the audio data (such as removal of interviewer's questions and other non-speech and isolation of background noises).

Figure 5 shows an example of a single extracted salient voice feature, which is a plot of the standard deviation of fundamental frequency of PTSD voices (bottom, samples 1-5) and non-PTSD voices (top, samples 6-10). A smaller standard deviation is found for the PTSD group, and the statistical significance level of the difference between these two groups is 0.0245 (smaller than 0.05). With numerous features extracted and selected, an enhanced committee machine-based regression model has been trained to differentiate PTSD and non-PTSD speeches.

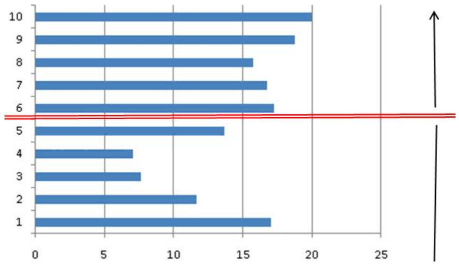

Figure 5. Exemplar feature plot: standard deviation of fundamental frequency.

To evaluate the enhanced committee machine-based PTSD monitoring model, a leave-one-out test scheme was used. For instance, if PTSD datasets 1-4 and non-PTSD datasets 1-4 are used for training of a PTSD GMM model and a non-PTSD GMM model, respectively, the PTSD dataset 5 and the non-PTSD dataset 5 will then be used for testing the detection accuracy. The results for training with only 1, 2, and 4 seconds of data are shown in Table 1. Two approaches are evaluated, GMM and GMM-based committee machine (GMM-CM).

We first evaluated the performance of using a GMM classifier. With only 4 seconds of voice data, we can achieve an average detection accuracy of 94.4% after score normalization. If using the whole audio clip from each test subject (about 2 minutes), the accuracy can reach 100%. When we combined multiple GMM classifiers using the enhanced committee machine method, we can improve the detection accuracy to 95.9% using 4 seconds of voice data.

Table 1. PTSD detection accuracy of GMM and GMM-CM classifiers

Data length	Accuracy for PTSD data		Accuracy for Non-PTSD data		Overall Accuracy	
Algorithm	GMM	GMM-CM	GMM	GMM-CM	GMM	GMM-CM
1 second	84.18%	83.21%	89.48%	94.26%	86.83%	88.74%
2 seconds	89.80%	89.17%	93.34%	96.81%	91.57%	92.99%
4 seconds	92.86%	93.11%	95.99%	98.64%	94.42%	95.88%

6. Conclusions & Discussion

We are currently developing a voice-based automated Tele-PTSD Monitor (TPM), which can remotely screen, monitor and provide assistance to clinicians in diagnosing a PTSD patient's mental healthiness and readiness in the field. Potential users include soldiers returning from warzone or other service man/women who are at risk or may suffer from PTSD. TPM system is accessible via a Public Switched Telephone Network (PSTN) or the Internet. It can be used as a screening tool to identify individuals with high PTSD risk and subsequently notify clinicians for follow-up contact. It can also serve as a monitoring tool to track the "mental healthiness" of persons already diagnosed with PTSD. Preliminary experimental results on a limited dataset consisting of both persons with PTSD and without PTSD show that the TPM system is able to achieve high detection accuracy using only a few seconds of data. Finally, demonstration software was developed based on PC/cell phone platforms and extensive tests have been performed to illustrate the robustness of the TPM system. The future directions of our work include: 1) improve the TPM server capabilities, such as multi-user access and integration with a COTS Interactive Voice Response (IVR) server; 2) enhance the PTSD monitoring performance by integrating additional features (such as emotional features) and improving the PTSD monitoring model and post-processing technique for a more accurate and robust PTSD assessment; and 3) evaluate TPM performance via a large scale dataset.

Acknowledgments

The development of TPM is being supported by the OSD/Army under contract number W81XWH-10-C-0204. We thank Ms. Ashley Fisher, Dr. Eva Lai, and Dr. Jay Shore for their comments and suggestions as we performed this research.

Disclaimer

The views, opinions and/or findings contained in this research/presentation/ publication/Press Release are those of the author(s)/company and do not necessarily reflect the views of the Department of Defense and should not be construed as an official DoD/Army position, policy or decision unless so designated by other documentation. No official endorsement should be made.

Reference herein to any specific commercial products, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the U.S. Government.

References

[1] American Psychiatric Association (1994). Diagnostic and statistical manual of mental disorders: DSM-IV. Washington, DC: American Psychiatric Association. ISBN 0890420610.

[2] Aaron Levin, "Nearly 10 Percent of Iraq Vets Screen Positive for PTSD", Psychiatr News, Volume 41, Number 7, page 1, April 7, 2006.

[3] James C. Mundt, Peter J. Snyder, Michael S. Cannizzaro, Kara Chappie and Dayna S. Geralts, "Voice acoustic measures of depression severity and treatment response collected via interactive voice response (IVR) technology", Journal of Neurolinguistics, Vol. 20, Issue 1, Pages 50-64, January 2007.

[4] A. Nilsonne, "Acoustic analysis of speech variables during depression and after improvement", Acta Psychiatrica Scandinavica, Vol. 76, Issue 3, Pages 235 - 245, Sep 1987.

[5] Guangfan Zhang, Roger Xu, Wei Wang, Jiang Li, Tom Schnell, Mike Keller, Individualized Cognitive Modeling for Closed-Loop Task Mitigation, ModSim World Conference, 2009, Virginia Beach, VA

[6] Jiang Li, J. Yao, N. Petrick, R. M. Summers, and A. K. Hara, "Hybrid Committee Classifiers for a Computerized Colonic Polyp Detection System," Medical Imaging 2006: Image Processing, vol. 6144, pp. 1701-1709, 2006.

[7] D.A. Reynolds and R.C. Rose (1995). "Robust text-independent speaker identification using Gaussian mixture speaker models". IEEE Transactions on Speech and Audio Processing

[8] Corinna Cortes and V. Vapnik, "Support-Vector Networks", Machine Learning, 20, 1995

[9] http://cycling74.com/products/soundflower/