CAPTURING CRITICAL INFORMATION USING THE WEARABLE INTELLIGENT REPORTING ENVIRONMENT (WIRE)*

Michael Orr, Helen Hastie, Daniel Miksch, Joshua Flanders, and Celeste Corrado

Lockheed Martin Advanced Technology Laboratories
3 Executive Campus
Cherry Hill, New Jersey 08002
[morr, hhastie, dmiksch, jflander, ccorrado]@atl.lmco.com

ABSTRACT

First responders and patrolling soldiers equally share the problem of reporting observations from the field. Current methods are often tedious and require a significant amount of attention to complete, distracting the mobile user from attending to the current situation. This paper presents the Wearable Intelligent Reporting Environment (WIRE), a capability allowing deployed personnel to record observations from the field in an easy intuitive spoken language interface that is natural to use. In addition to supporting an interface to collect observations, WIRE enables transmission from the field, dramatically increasing the dissemination and analysis of critical intelligence information.

INTRODUCTION

Deployed emergency and military personnel have the difficult task of performing their duties and making observations of the situation. Making these observations and filing them upon return from the field is burdensome. As a result, only a tiny amount of current information is obtainable, thus decisions are often made based on outdated information. This inability to disseminate crucial information about a situation significantly decreases safety of deployed personnel still out in the field.

To address this issue, Lockheed Martin Advanced Technology Laboratories (LM ATL) is developing the Wearable Intelligent Reporting Environment (WIRE). WIRE makes information access and dissemination an easy, painless process. Mobile users equipped with WIRE can create digital reports in the field, record notes, and access critical intelligence information. Allowing responders to create digital reports in the field is not a new approach, but most systems require data input by keyboard, mouse, or pen, all which rely on displays and hands. WIRE is hands-free and incorporates the latest in Spoken Language Understanding and lightweight wearable technologies. It requires no display and operates by natural speech. This allows the user to remain focused on the tasks at hand yet able to create thorough, standardized reports. Report distribution can be automatic and wirelessly from the field or manually by downloading directly from the wearable system at a later time.

---

* This system was previously referred to as the Emergency Responder Intelligent Interaction Environment (ERI^{2}E)

DOMAIN PROBLEM

Many organizations deploy personnel to potentially hostile or dangerous environments to perform duties as well as collect information. The United States Army relies on the human intelligence gathering of small, forward scouting groups to increase battlefield situational awareness. Hazardous materials technicians must investigate and direct clean up of toxic and dangerous sites. Emergency responder personnel must deal with unknown situations, injured people, and disasters on a regular basis.

In all of these cases, the personnel deployed are sent into situations with little or no advanced knowledge. The primary tasks of the personnel may be different depending on the type of event, but in each case they must gather information about the situation, whether it is patient information, enemy troop movements, structural damage assessments, or a potential biological agent outbreak. Current procedures for reporting this information vary in format and depth but all possess two major drawbacks.

First, current reporting procedures usually involve personnel taking notes at the scene and submitting a report upon return. This is burdensome to the people in the field who must take time away from their duties to make notes, fill out a report in the proper format, and submit it. Second, the current procedures discourage quality data gathering by making it difficult. The process of handwriting and submitting reports slows the turn-around time for the information that does make it back from the field. This hampers analysis and slows down situational awareness.

Army Lieutenant General Alexander, stated that as of November 2003, 400,000 Post-Conflict patrols have been conducted in Iraq, and only 6,000 digitized reports have been created and pushed to the Battalion or higher level. Another example is the pre-hospital patient care reports filed by EMS workers. These are often submitted in a haphazard and barely understandable format with shorthand notations unique to individual ambulance groups. The Oregon Department of Human Services [1] has spent eight years trying to develop a common format and electronic submission capability because it takes too long to collect and transcribe these time sensitive local reports. The federal government is also looking into a national standard for this reason.

THE WIRE SOLUTION

WIRE attempts to solve these problems by allowing the responder to gather more information faster and easier, increasing the speed information can be analyzed and disseminated, and increasing safety and capability of personnel in the field through increased intelligence. Deployed personnel can converse with WIRE as naturally as giving a report to a human. The system constantly listens so that the personnel do not have to take their hands off of equipment to turn on the microphone. The user would proceed and speak the report or record notes, which would be customized to the user's domain, with the system prompting the user if key information is missing. Additionally, WIRE collects sensor data from the wearable sensors, including GPS, heading, and image data, and automatically integrates these

data into the report, saving the user from manual entry of that data. When finished the user would say “Report Complete” and the system would revert back to listening mode. Should the user be interrupted during the generation of a report, they can interrupt their input to the report with a command and continue with the system when ready.

EXAMPLE DOMAIN — SALUTE REPORTING

Soldiers conducting patrols in post-conflict areas acquire a wealth of knowledge and intuition about the region. These soldiers collect large quantities of human intelligence data which do not get exploited. Each patrol is required to report on their patrol. The key useable report is a collection of Size Activity Location Unit Time and Equipment (SALUTE) reports per patrol. There are almost 10 times as many patrols as there are reports filed because of the burden on the soldiers of creating the digital reports upon return to base. Figure 1 contains a sample SALUTE report [2].

TO: G2, V CORPS
FROM: 1-96 FA, 23 AD DTG: 230900Z AUG 98
REPORT NO: 07-035

1. SIZE: N/A

2. ACTIVITY: Captured Ammunition

3. LOCATION: West bank of Fulda River, south of Bebra, six-digit grid NB 553476

4. UNIT: 1-96 FA, 23 AD (capturing unit)

5. TIME: Ammunition captured at 230230Z Aug 98

6. EQUIPMENT: N/A

7. REMARKS/OTHER INFORMATION: Response to priority intelligence requirements (IPR) 23-0016-93. Ammo site secured, awaiting disposition instructions.

Figure 1. Sample SALUTE Report

WIRE is configured to generate semantically rich SALUTE reports, will assist the soldier in dynamically generating SALUTE reports based on soldier input and automated sensor suite readings for location, time, and heading. Table 1 contains a sample dialogue for using WIRE in the SALUTE reporting domain that demonstrates its robust, mixed initiative capabilities.

In this environment, the completed reports are transmitted wirelessly to the command post for immediate processing. If a network link is not available when complete, the reports are queued for a future transmission. In the worst case, the reports will be transmitted when the soldier returns to base.

Table 1. Sample WIRE Dialogue

Responder: New report	**System: New SALUTE, what is Size?**	*WIRE assists with filling in the report*
Responder: There are 15 enemy soldiers traveling south	**System: Roger. What is Location?**	*Dialogue is mixed initiative, system accepts any combination of information*
Responder: 2 miles south of my current location	**System: Location is latitude 57 longitude 99. What is Unit?**	*WIRE utilizes current location information from GPS sensors*
Responder: They are dressed in fatigues with blue patches and arabic writing.	**System: Roger that. What is Time?**	*WIRE can handle complex linguistics such as anaphoric references.*
Responder: uhh about 10 minutes ago, and they are carrying machine guns and grenade launchers	**System: Roger. Time is 0800.**	*WIRE easily handles disfluencies*
Responder: End report	**System: Report saved.**	*Quick, easy interactions make the interface more useable*

SYSTEM ARCHITECTURE

The architecture for WIRE has been designed for a mobile user. Depending on the type of tasks performed by the user, WIRE can be installed on either a wearable or handheld computer. In most cases, it is preferable to run WIRE on a wearable computer so both hands can be used to perform other tasks. With WIRE, no assumptions are made on network availability; therefore all processing is performed locally on the portable device. The only exception for network availability falls with transmitting completed reports. In this case, if a network link is not available at the time the report is completed, it is queued for transmission at a later time, such as when a soldier returns to base or when a responder comes within range of a wireless network.

This section describes the wearable hardware configuration for running WIRE as well as the software components that enable the spoken language-understanding interface. A functional view of the system is depicted in Figure 2.

Hardware

Considering the noise levels the responder may experience while in the field, the audio will be captured by a throat microphone. Throat microphones are activated only by skin vibration from the vocal cord and are capable of canceling background noise by 90%. In addition, they enable the responder to whisper when necessary. The hardware for audio output consists of a speaker or ear piece connected to the audio output of the WIRE system's sound card. The earpiece, either a behind-the-ear speaker or bone conduction headset, will be used for system generated output audio, such as tones or speech.

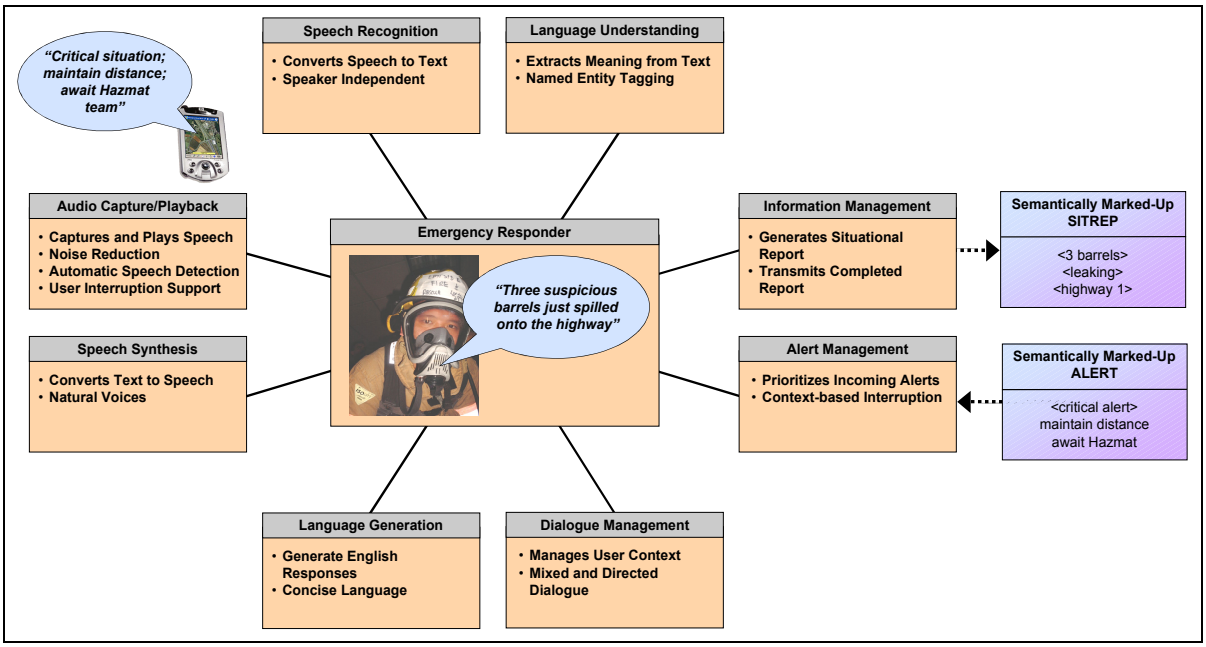

Figure 2. WIRE Functional Architecture

The WIRE system will have an integrated sensor suite to assist the responder in capturing critical information while operating in various patrol modes. The integrated sensor suite will include a digital capture camera, GPS, and a digital compass.

Finally, each WIRE system will include a PC/104 based wearable CPU running Linux. PC/104 is a standard for PC-compatible circuit boards that can be stacked together to create an embedded computer system. PC/104 boards are very powerful for their size—approximately 4-in. by 4-in. They are designed for minimal power consumption, small footprint, modularity, expandability, and ruggedness. The PC/104 system will use a Transmeta Crusoe BGA processor, running at 800 MHz enclosed in a small, lightweight housing to ensure optimal performance in the operational environment. By using hot-swappable batteries, the WIRE system will be able to support extended operations, without losing data or functionality.

Software

Speech is a robust, hands and eyes-free mode of interacting with WIRE. Since the interface will understand a wide range of events and paraphrases, it will be easy to use and require little training. This has been shown to be the case in our Shipboard interface and our Marine Logistics interface that were well accepted into every day use by personnel [3]. The system will be trained on numerous accents of English in a variety of noise levels, making it speaker independent. As we have previously demonstrated, it will function in high noise environments up to 103 db. In addition, the understanding or Language Processor module is advanced enough to remember dialogue context and make inferences based on this, thus reducing the need to repeat inferable information such as anaphoric references.

The speech is collected by the audio capture component, and passed to the Speech Recognizer server that is responsible for converting the audio waveform to text. The output

of the Speech Recognizer is a list of the n most likely possible utterances (known as the n-best list) that map to the corresponding speech. Language Understanding process analyzes the n-best list produced by the Speech Recognizer to extract the underlying semantics of the responder's utterance to generate a semantic frame. The semantic frame is passed to the Dialogue Manager that maintains the context of the most recent dialogue with the user and interfaces with other components of WIRE. The Dialogue Manager checks the information contained in the semantic frame to determine which data elements are present, which are missing, and whether the information is consistent with domain knowledge. The Dialogue Manager is configured for each domain with knowledge and the appropriate reporting representation model. The Dialogue Manager may need to interact with the user in which case a new semantic frame, or meaning of the response to send back to the user, is generated and passed to the Language Generation component where it is converted to a concise English prompt. The text prompt is converted to audio via a speech synthesis engine and played to the user.

The Information Management component within WIRE is responsible for generating the semantically rich report. The report is continually updated as additional information is provided by the user. When complete, either when all fields are provided or as determined by the user, the report is transmitted to an operations center or command post via radio or wireless link for immediate processing and action. These processing centers may consider reports from WIRE as human intelligence (HUMINT) and fuse with information collected from sensors, and intelligence data sources (FBI, ONI, etc.) to develop an understanding and model of the current situation with the ability to make predictions on future activities [4]. At any time, alerts can be issued to a WIRE user. Some notional system, such as one located in the operations center, can issue a semantically marked-up alert to the WIRE system. These alerts can be issued depending on a mobile user's proximity to an incident, types of observations reported, etc. The alert is received by the Alert Management component, prioritized, and then passed to the user at an appropriate time.

CONCLUSION

LM ATL's extensive experience fielding speech understanding [3] systems provides us with a unique insight into crafting usable systems, low impact interfaces. WIRE's capabilities will be thoroughly tested in a MOUT environment. WIRE avoids the serious obstacles of recognition, acceptance, data transmission, and power consumption. The throat microphone coupled with previously fielded speech recognition technology achieves high recognition in noisy field environments. Problems with data transmission are overcome by providing local storage options within the system itself, and the system will be designed for low power usage and have the ability to hot swap batteries for longer usage.

WIRE will allow deployed personnel to record data and create reports in the field in an easy, intuitive manner that is less burdensome than current methods. As a result, agencies will gather more critical information faster, analyze that information, and disseminate it more efficiently. WIRE provides these capabilities through its intuitive speech interface and ability to send information wirelessly from the field.

REFERENCES

[1] *The Quick Connect: The Newsletter of the Oregon Volunteer Firefighters Association*, 2003. Retrieved April 20, 2005, from http://www.ovfa.org/quick_connect_december.pdf

[2] *Chapter 12: Captured Enemy Ammunition* (n.d.) Retrieved April 20, 2005, from http://www.globalsecurity.org/military/library/policy/army/fm/4-30-13/ch12.htm

[3] Daniels, J. and Hastie, H., "The Pragmatics of Taking a Spoken Language System Out of the Laboratory," *Proc. of Human Language Technology (HLT) North American Chapter of the Association for Computational Linguistics (NAACL) 2003*, Edmonton, Canada, May 27-June 1, 2003.

[4] Gigli, Sergio, *Situation Understanding Technologies for Homeland Security*, R&D Partnerships in Homeland Security Proceedings, Boston, MA, April 27-28 2005.