DSA Intelligent Transformational Systems
Real-time Individualized Training Vectors For Experiential Learning

Sandia National Laboratories

PM: John Wagner, PI: Elaine Raybourn, Ph.D.

Team: Nathan Fabian, Matthew Glickman, Eilish Tucker, Matthew Willis

Problem

Few approaches to data mining in deployed training games exist, even though game-based training is used by the military around the world. Opportunity exists to collect and analyze these data, to construct a full-history learner model, and to individualize and Improve the effectiveness of the training.

To address this we:

o Establish a human training performance metric on a non-kinetic culturally based training mission in DARWARS Ambush NK

o Establish a baseline performance rating using human coders, after establishing inter-coder reliability.

o Use content analysis to develop an automated coder that predicts human baseline performance evaluation

Adaptive self-paced and customized "anytime-anywhere" training in intercultural human interaction is one of the top five priority U.S. Army Future Force warfighter outcomes.

Approach

	Ratings				
	1	2	3	4	5
Communication	Poor to little communication with Confederate	Poor communication with Confederate	Needs prompting from Confederate	Utilizes Confederate	Engages Confederate in mission
	Poor to little communication with Host National	Poor communication with Host National	Okay communication with Host National	Good Communication with Host National	Great communication with Host National
	Does not attempt to use Arabic greetings		Attempts Arabic Phrases		Uses Arabic
Cultural Awareness	Poor to no Cultural Awareness	Little Cultural Awareness	Attempts Cultural Awareness	Demonstration of Cultural Awareness	Cultural Awareness more prominent across behaviors and communication
Observation	Not aware of surroundings	A little aware of surroundings	Observant, but might not spot crowd or BOLO	Aware of surroundings. Might spot BOLO	Spots BOLO. Aware of crowd. Concerns for safety
Mission	Tries to Arrest Host National			Asks pointed questions	
	Forgets parts of mission	Narrow view of mission		Concern for finishing mission non-kinetically	
	Very Forceful		Forceful		Not forceful

Table 1 - Human coder's player rating guidelines.

oTwo human coders rated each player's performance following rules in Table 1.

o Human coders also considered in-game quantitative ratings of Observer/Evaluator role-players in overall performance ratings

o All communication was transcribed manually to text files. Files were processed in Titan Toolkit using text analysis algorithms (LSA) developed for the Network Grand Challenge LDRD (Figure 1)

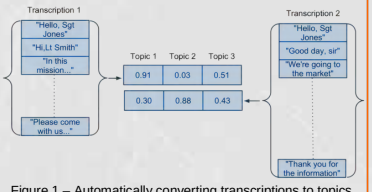


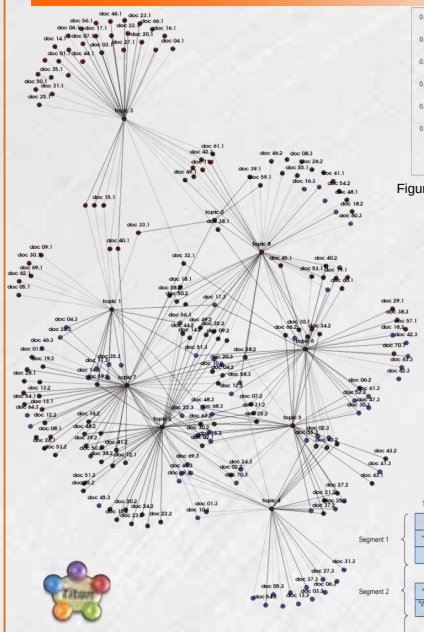
Figure 1 – Automatically converting transcriptions to topics

o Produced a topic space (Table 2) of all in-game communication.

Topic 0	Topic 1	Topic 2	Topic 3	Topic 4	Topic 5	Topic 6	Topic 7	Topic 8
ah	shot	tent	shop	target	uh	uh	uh	uh
uh	hello	shot	inch	medical	bodies	confirmed	bodies	americans
nasir	safe	green	tv	neutralized	lieutenant	injured	village	leave
um	shop	crates	thirty-two	permission	return	rocket	tribes	tribe
sergeant	shooting	injured	greetings	lock	medical	care	leave	strange
okay	rpg	medical	models	clear	touch	bodies	terrorists	weapon
saiful	buy	night	flat	confirmed	happened	appears	members	cousin
hello	greetings	bodies	panel	force	throwing	permission	ambulance	night

Table 2 - Automatically generated topic space.

Approach, contd.



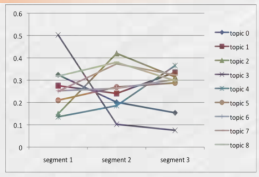


Figure 2 – Topic relevance in temporal segment

o Transcripts were split manually into three temporally distinct sub-transcripts

o Useful temporal patterns appear in the split data (Figures 2 and 3)

o Topic space vectors were concatenated into extended space vectors for each transcription (Figure 4)

Figure 3 – Weighted topic-to-segment graph. First temporal segment rendered in red, second rendered in black, and final rendered in blue

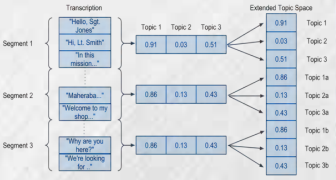


Figure 4 – Splitting documents into temporal segments

Results

o Using Support Vector Machines to provide a mapping from vector to rating, we predicted human coders' scores with a correlation of .62 (with 10-fold cross validation) and mean absolute error of .70.

o The correlation between the two human coders was .79. Figure 5 shows both regression lines and scatterplot data for the predictions vs. human coder.

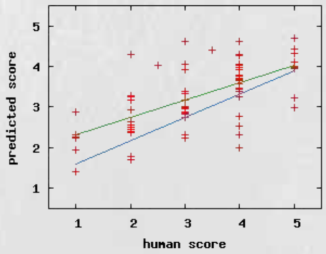


Figure 5 – Comparison of predicted scores to original human scoring. Green line is computer-human regression. Blue line is human-human regression.

Significance

The result implies that, given text transcription or text-only communication, human coders can be effectively replaced by this automated transcript evaluation. Given the potentially very large quantity of game data produced in military training, especially communication data (i.e., social or cultural games), this automated processing provides an analysis path forward to retain the data for historical learner models, and squad or unit-level performance models.

Although not exploited with the size of this data set, the Titan Toolkit is built to take advantage of internet-scale information and to scale its analysis up to the extreme scale possible in military training data.

Finally, this result also provides a path forward to develop a full learner model. Having data split into temporal segments forms a rudimentary trajectory of performance through the course of the serious game. This trajectory can be further studied in the pursuit of real-time learner model construction.

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

SAND 2010-5379P

Sandia National Laboratories