UNDERSTANDING SPEAKER ROLES WITHIN NATURALISTIC TEAM MISSION COMMUNICATIONS: FEARLESS STEPS APOLLO-11

M.M. Chandra-Shekar, J.H.L. Hansen

{meena.chandrashekar, john.hansen}@utdallas.edu
CRSS-RSTL: Center for Robust Speech Systems, Robust Speech Technologies Lab (RSTL),
The University of Texas at Dallas (UTD), Richardson, Texas 75080.

INTRODUCTION: Apollo-11 (A-11) was the first manned space mission to successfully bring astronauts to the moon and return them safely. Effective team based communications is required for mission specialists to work collaboratively to learn, engage, and solve complex problems. As part of NASA's goal in assessing team and mission success, all vital speech communications between these personnel were recorded using the multi-track SoundScriber system onto analog tapes, preserving their contribution in the success of one of the greatest achievements in human history. More than +400 personnel served as mission specialists/support who communicated across 30 audio loops, resulting in +9k hours of data for A-11. To ensure success of this mission, it was necessary for teams to communicate, learn, and address problems in a timely manner. Previous research has found that compatibility of individual personalities within teams is important for effective team collaboration of those individuals. Hence, it is essential to identify each speaker's role during an Apollo mission and analyze group communications for knowledge exchange and problem solving to achieve a common goal. Assessing and analyzing speaker roles during the mission can allow for exploring engagement analysis for multi-party speaker situations.

METHOD: The UTDallas Fearless steps Apollo data is comprised of 19,000 hours (A-11,A-13,A-1) possessing unique and multiple challenges as it is characterized by severe noise and degradation as well as overlap instances over the 30 channels. For our study, we have selected a subset of 100 hours manually transcribed by professional annotators for speaker labels. The 100 hours are obtained from three mission critical events: 1. Lift-Off (25 hours) 2. Lunar-Landing (50 hours) 3. Lunar-Walking (25 hours). Five channels of interest, out of 30 channels were selected with the most speech activity, the primary speakers operating these five channels are command/owners of these channels. For our analysis, we select five speaker roles: Flight Director (FD), Capsule Communicator (CAPCOM), Guidance, Navigation and, Control (GNC), Electrical, environmental, and consumables manager (EECOM), and Network (NTWK). To track and tag individual speakers across our Fearless Steps audio dataset, we use the concept of 'where's Waldo' to identify all instances of our speakers-of-interest across a cluster of other speakers. Also, to understand speaker roles of our speaker-of-interests, we use speaker duration of primary speaker vs secondary speaker and speaker turns as our metrics to determine the role of the speaker and to understand their responsibility during the three critical phases of the mission. This enables a content linking capability as well as provide a pathway to analyzing group engagement, group dynamics of people working together in an enclosed space, psychological effects, and cognitive analysis in such individuals.

IMPACT: NASA's Apollo Program stands as one of the most significant contributions to humankind. This collection opens new research options for recognizing team communication, group dynamics, and human engagement/psychology for future deep space missions. Analyzing team communications to achieve such goals would allow for the formulation of educational and training technologies for assessment of STEM knowledge, task learning, and educational feedback. Also, identifying these personnel can help pay tribute and yield personal recognition to the hundreds of notable engineers and scientist who made this feat possible.

ILLUSTRATION: In this work, we propose to illustrate how a pre-trained speech/language network can be used to obtain powerful speaker embeddings needed for speaker diarization. This framework is used to build these learned embeddings to label unique speakers over sustained audio streams. To train and test our system, we will make use of Fearless Steps Apollo corpus, allowing us to effectively leverage a limited label information resource (100 hours of labeled data out of +9000 hours). Furthermore, we use the concept of 'Finding Waldo' to identify key speakers of interest (SOI) throughout the Apollo-11 mission audio across multiple channel audio streams.

REFERENCES: ***[1]*** A. Joglekar, S.O. Sadjadi, M. Chandra-Shekar, C. Cieri, J.H.L. Hansen, "Fearless Steps Challenge Phase-3 (FSC P3): Advancing SLT for Unseen Channel and Mission Data across NASA Apollo Audio," ISCA INTERSPEECH-2021, pp. 986-990, Sept. 2021; ***[2]*** M. C. Shekar, Master's Thesis, CRSS-UTDallas, p. 48; ***[3]*** M. Chandra Shekar, J.H.L. Hansen, ""Historical Audio Search and Preservation: "Finding Waldo" within the Fearless Steps Apollo-11 Naturalistic Audio Corpus," IEEE Signal Proc. Mag. (accepted Sept. 2022); ***[4]*** "Apollo flight controller 101: Every console explained," https://arstechnica.com/science/2019/12/apollo-flight-controller-101-every-console-explained/?comments=1, accessed: 2022-03-28