Developing a Virtual Reality Wildfire Simulation to Analyze Human Communication and Interaction with a Robotic Swarm During Emergencies

Patricia Chaffey[*], **Ron Artstein**[*] **Kallirroi Georgila**[*], **Kimberly A. Pollard**[†], **Setareh Nasihati Gilani**[*]
David M. Krum[*], **David Nelson**[*], **Kevin Huynh**[*], **Alesia Gainer**[*], **Seyed Hossein Alavi**[*], **Rhys Yahata**[*], **David Traum**[*]

[*] Institute for Creative Technologies, University of Southern California, Los Angeles, CA, 90094
[†]U.S. Army Research Laboratory, Los Angeles, CA, 90094

Abstract

Search and rescue missions involving robots face multiple challenges. The ratio of operators to robots is frequently one to one or higher, operators tasked with robots must contend with cognitive overload for long periods, and the robots themselves may be discomfiting to located survivors. To improve on the current state, we propose a swarm of robots equipped with natural language abilities and guided by a central virtual "spokesperson" able to access "plays". The spokesperson may assist the operator with tasking the robots in their exploration of a zone, which allows the operator to maintain a safe distance. The use of multiple robots enables rescue personnel to cover a larger swath of ground, and the natural language component allows the robots to communicate with survivors located on site. This capability frees the operator to handle situations requiring personal attention, and overall can accelerate the location and assistance of survivors. In order to develop this system, we are creating a virtual reality simulation, in order to conduct a study and analysis of how humans communicate with these swarms of robots. The data collected from this experiment will inform how to best design emergency response swarm robots that are effectively able to communicate with the humans around them.

1. Introduction

In emergency disaster situations, there is extreme time pressure to conduct searches and locate survivors within the first 48 hours, as the mortality rate increases substantially after this time (Murphy, 2004). Emergency responders are limited to whatever manpower can be mustered, as well as what technology is available to them to aid in their tasks. To address this challenge, we propose utilizing a robotic swarm and virtual "spokesperson" equipped with natural language capabilities to assist with both locating survivors in a dangerous area, as well as communicating with them to determine the best course of action.

A robotic swarm can be defined by the coordination between robotic units, which is reliant on shared information and distributed algorithms. Units within the swarm will be able to traverse areas dangerous to human life, mitigating some of the risk that typically befalls emergency responders working in the field. In addition, the swarm will need to be able to communicate effectively with survivors, as well as impart knowledge to emergency personnel responding to or controlling the swarm's movements.

A significant challenge in implementing swarms of robots is controlling the group in a manner that is efficient for one human to oversee (Kolling et al., 2016). Human control and intervention is currently essential for search and rescue operations with swarms. Human involvement increases swarm performance in complex and cluttered environments (Kolling et al., 2013), such as a populated area during a disaster. However, the human can also be a bottleneck to performance if they must control robots individually or approve all decisions (Wang et al., 2009; Velagapudi et al., 2008). To address this, these swarms will need autonomous abilities and new modes of communication that will allow humans to direct, track, and absorb incoming information from multiple entities. Historically, ease of use has been an impediment in using robots during disaster events. In 2001, the World Trade Center rescue response teams had access to various robots, but only utilized a few. Some reasoning for the exclusion included physical constraints, but in one case, the robot was not used because the user interface provided with it was still in its infancy (suitable for developers, but not emergency personnel) (Casper and Murphy, 2003). This highlights the importance of developing an interface that is intuitive to responders with various levels of training with robots, as the rate at which rescuers can be taught to use these tools correlates to the accessibility of training (USFA, 1995). A swarm with these capabilities will have many applications. Here, we discuss the utilization of robotic swarms in a search and rescue context in a wildfire.

In a one-year project, the USC Institute for Creative Technologies is exploring two models of communication: (1) a sports-team inspired "playbook" capability that assigns specific roles to members of the swarm that can be coordinated with other units, and (2) a virtual swarm "spokesperson" who communicates with humans and acts as a human-like front-end interface to the complex swarm and imparts information from the collective knowledge of the swarm. Humans can manage larger numbers of robots if they can issue high level commands (e.g., playbook plays) to a mediating agent (e.g., a "spokesperson") (Sycara and Lewis, 2012). This personified front end allows not only for communication with a swarm operator, but with civilians as well. A natural language interface, e.g., (Traum et al., 2003; Lukin et al., 2018), would allow intuitive interaction between swarm operators and the virtual spokesperson, and between the virtual spokesperson and civilians, obviating the need for extensive training or pre-existing familiarity with robots. Other first steps taken with this project include the creation of a framework for experimentation and data collection, in addition to a demonstration in a simulated environment of how these models of communication can impact human-swarm collaboration.

Due to the difficult nature of collecting data on human responses during emergency situations, this project uses a virtual reality environment depicting an active wildfire encroaching on a small town (shown in Figure 1). A player will assume the role of a swarm operator, tasked with the swift and safe evacuation of a town in imminent danger from an approaching wildfire. At this stage in the project, the player will have to handle where they assign members of the swarm, the majority being unmanned aerial vehicles (UAVs, i.e., drones), as well as handling interpersonal issues with residents in danger from the fire. To reduce cognitive overload, players will at times need to rely on the "spokesperson" to step in and assist with tasks when multiple incidents occur simultaneously. Assisted tasks may include guiding civilians away from danger, deploying drones, and handling larger events. While it is not the current focus, we also intend to investigate interactions between the spokesperson and townspeople and other rescue workers.

Figure 1: A bird's eye view of the town in the simulated environment

2. Related Work

The idea of using robots for search and rescue is well established (Shah and Choset, 2004), with disasters being a driving force for advancement (Casper and Murphy, 2003). Natural catastrophes (e.g., earthquakes) or man-made disasters (e.g., terrorist attacks) can kill thousands of people in a very short amount of time and the critical period (usually estimated to 48-72 hours) after the disaster is when first responders want to utilize all the available assets (Stein, 2004). After this period, victim mortality drastically increases owing to exposure and lack of food, water, and medical treatment.

There are various research groups worldwide who focus on developing robots for Urban Search And Rescue (USAR). The Center for Robot-Assisted Search And Rescue (CRASAR, 2019) is the oldest institution devoted to promoting the use of ground, marine, and aerial unmanned systems for public safety. Their mission is fostering unmanned systems used by formal emergency management agencies through voluntary national and international activities.

In one of the first attempts for designing collaborative systems of mobile robots, (Jennings et al., 1997) introduced an algorithm for a distributed team of autonomous mobile robots to search for an object. Upon finding the object, the robots would gather around it and "rescue" it in a collaborative manner.

(Casper et al., 2000) explored issues that arise when using robots in search and rescue, focusing on a collapsed building scenario. (Penders et al., 2011) developed technology for a swarm of robots assisting fire-fighters. They discussed the swarming algorithms that provide the functionality by which the robots react to and follow humans. Furthermore, (Jacoff et al., 2003; Messina and Jacoff, 2006) have introduced and explored different metrics for performance evaluation of urban search and rescue tasks.

3. The Simulation

The primary objective of this project is to study how humans interact and communicate with swarms of robots in disaster situations, both from the point of view of an operator, as well as a civilian. To this end, we are designing a simulation that allows players to take on these roles, and intend to collect data on how they interact with robotic units during an emergency event. The approach taken in developing this simulation is outlined below.

3.1. Proposed Scenario

For the scope of the project, the simulation is currently limited to that of a wildfire. Future iterations of the project may include other search and rescue scenarios, such as flooding or the aftermath of an earthquake. A wildfire presents challenges in time and resource management unlike other possible scenarios. In this context, the operator must work to remove civilians from the encroaching fire before it reaches their portions of the map. Challenges include stubborn residents who initially refuse to leave their homes, injured civilians in need of medical assistance, and obstruction of vital roads leading out of the town. These challenges will occur in overlapping time blocks, increasing the cognitive load placed on the operator as they navigate the simulation.

One of the overarching difficulties in this simulation will be swarm management. Although a primary goal is exploration of the map and locating civilians who have yet to evacuate, operators will also need to decide how many drones to hold in reserve until critical moments. Operators will face several possible dilemmas, for example multiple incidents may occur at once, requiring the operator to commit several members of the swarm towards addressing these issues. Likewise, an operator may be confronted with the possible loss of a number of drones. If too many members of the swarm are lost or out of commission, the operator's ability to cover ground at a reasonable pace is severely limited. The operator will need to be able to anticipate possible emergency events and try to prepare for them.

To address this difficulty, an operator will be assisted by a virtual "spokesperson". The spokesperson will have basic natural language capabilities and be equipped with certain "plays" to implement at various points in the scenario. An example of this would be of an incident where a previously safe zone becomes dangerous. The spokesperson is tasked by the operator to guide individuals away from a danger to a new, safe location. The spokesperson handles the communication aspect by projecting instructions through the drones to the civilians, and using the drones as a visual guide to the new area. Other areas where the spokesperson can assist include offering a brief description of the current state of the disaster (e.g., explaining to the civilian that the fire is still approaching their current location) and instructing the person on what they should do to evacuate safely. The virtual spokesperson will be initially implemented via Wizard of Oz methodology (Dahlbäck et al., 1993; Green et al., 2004), following a pre-defined set of capabilities and language response protocol. Civilians may be played by additional wizards and/or by research participants, depending upon the aspect of the scenario under study. Future iterations of the simulation can employ progressively more automated natural language and swarm management capabilities, building upon data gathered in earlier iterations, as in (Passonneau et al., 2011; Marge et al., 2017).

The advantages of implementing a spokesperson equipped with these plays include the ability to cover more ground, locate and speak to more civilians, and allow the operator to avoid cognitive overload while handling the aspects of rescue that require personal attention. The events requiring personal attention may be unique or highly contextual in nature, beyond the general issues the spokesperson is equipped to deal with. However, in order to build a system that can communicate effectively with civilians and operators, information on how people communicate with these swarms must be gathered and analyzed.

3.2. Development of Game Principles

Currently, project resources include a developed virtual map of the simulated town the operator will explore and interact with (Figure 1). For the simulation, we created a neighborhood in the Unity game engine (https://unity.com), using the Residential Buildings Pack from Gabro Media[1] with trees from SpeedTree[2]

The incidents that can be discovered by the operator as they explore the zone are balanced to provide an experience that is both realistic and reflective of the cognitive stresses that occur during the initial hours of a search and rescue. Operators who make optimal decisions and wisely allocate their resources will be able to maximize their coverage of the map and increase their chances of evacuating as many civilians as possible.

In order to pilot the simulation, we have used a role-playing game approach, with a *Game Master* (or GM), guiding events and providing information to players, including the operator. When an event involving civilians is discovered, another individual takes on the role of the civilian and interacts with the operator. As events are 'discovered' the operator makes decisions regarding how to deploy resources, which the GM implements within the constraints of the game. Some of the constraints include time-sensitive events and resource limitations. In regard to what is considered time-sensitive, an operator may need to take into account the rate at which the fire spreads, and evacuate civilians closer to it. Additionally, there is a time cost to sending members of the swarm out to specific locations on the map. Resource limitations refer to what drones are available for an operator to use, in addition to other resources such as a fire truck or paramedic team.

An example of an incident is when drones discover a stubborn couple intent on remaining in their home, despite the rapidly encroaching fire. The couple's reasoning for this is that they feel they have taken necessary precautions and will be able to simply ride out the flames as they pass by. As an emergency responder, it is an operator's responsibility to ensure that the couple leaves their home with enough time to escape the wildfire. The operator may initially opt to allow the spokesperson to attempt to persuade the couple into leaving, while they continue their work elsewhere, or they may step in personally right away, e.g., by patching their voice directly through the robot that is nearest to the couple. While this event is occurring, other incidents are being discovered by the swarm of drones moving throughout the map. A simultaneous event might include another family that requires instructions on how to safely evacuate. Once again, the operator will be presented with a choice in how they want to handle the situation. As events of this nature are discovered, the operator will find that they need to make rapid-fire decisions regarding which residents they speak with to persuade or calm as needed, and which can be left to the spokesperson.

With the right decisions in the scope of the simulation, an operator will be able to maintain enough of their resources in order to adequately cover the town while stepping in as needed with the different citizens and their various issues. These decisions, however, will need to be made under time-sensitive conditions and increasing pressure from the fire. How the operator makes their decisions and how they communicate these choices to the spokesperson are items to be explored in this simulation.

4. Data Collection

The first stage of experiments will involve Wizard of Oz testing, where wizards will act out the roles of the spokesperson and the civilians, and assist with drone movement. Following this, future iterations of the simulation will be more autonomous. Participants interacting with the scenario will be recorded as they play. The recordings will be analyzed in order to ascertain what trends emerge when people communicate with robotics systems in emergency contexts. The analysis will inform future designs of the spokesperson, to better enable it to effectively communicate with operators and civilians alike. In addition, the human trials involving wizards will shed light on what participants expect in terms of the spokesperson's ability to communicate.

[1]https://assetstore.unity.com/packages/3d/environments/urban/residential-buildings-pack-39272

[2]https://assetstore.unity.com/packages/3d/vegetation/speedtree/free-speedtrees-package-29170

4.1. The Operator's Perspective

From the operator's point of view, they will need to be able to command the spokesperson at various times to assist, as well as speak through the drones to civilians. Questions that arise include the following: How do operators issue commands and requests to the spokesperson? What tasks to operators feel comfortable delegating and how to they check on progress? How can a spokesperson be designed to elicit trust and reduce cognitive load of the operator?

Answering these questions allows for better design and expectation management in regard to the spokesperson. It also informs what kind of training might be needed for individuals using the system out in the field. Given the time-sensitive nature of disaster response, it is imperative that minimal time be spent clarifying errors in speech and understanding. Ensuring that the spokesperson is suitably equipped to handle general tasks and queries, as well as making certain the operator has a realistic sense of how the spokesperson works, helps to address these issues.

4.2. The Civilian's Perspective

The civilian's perspective is somewhat more challenging to evaluate, as an individual's response to a simulation involving an emergency may not be as high in fidelity to a real disaster situation. However, topics that can be explored include what nature of questions a civilian survivor might ask, in addition to asking participants in the simulation to assess their comfort around robotic members of the swarm.

Many robots in search and rescue are not anthropomorphic, nor can they easily support or be reconfigured to show facial expressions or support affective expression (Bethel and Murphy, 2006). However, in stressful situations, it is particularly important that a robot be able to quickly and smoothly present itself as a reliable entity that is there to help. Given this, identifying how a humanoid virtual spokesperson can use natural language to communicate for the swarm is an important component in designing these swarms.

To expand on this, if the swarm spokesperson can convince a civilian that it is trustworthy and capable, the civilian is far more likely to adhere to advice given by the spokesperson. If the operator is not required to step in personally in each individual situation, this allows the operator's attention to remain free to handle other tasks. An important component to how this system can maximize its effectiveness is to ensure that the human operator is not spending too much time or attention on conversation with civilians that do not require it.

Note that due to limitations in the design and hardware of drones, the drones themselves are not envisioned as having natural language processing capabilities. Each drone will be equipped with a microphone and a speaker. When the civilian speaks to the drone, the drone will just transmit the civilian's speech recorded through the microphone to the spokesperson and the simulation environment. Similarly, when the spokesperson or the operator wants to communicate with the civilian, synthesized or pre-recorded speech will be transmitted to the drone and played through the drone's speaker.

5. Future Steps

Next steps for the current iteration of the project include finalizing the designs of events and subsequently, human trials with the current wildfire module. The data collected regarding how humans interact with the spokesperson and drones will be used to refine how the spokesperson communicates with users. In addition, areas where operators may need to be trained in regard to the system will be identified, and implemented either through tutorial-like activities or added explanation from the spokesperson when a person initially begins using the system.

Furthermore, we will use machine learning in order to improve the autonomous behavior of the spokesperson and the swarm entities, and their ability to communicate and adapt to interaction with humans. Reinforcement Learning (RL) is well suited for this task as it is designed for maximizing long-term goals rather than focusing on immediate results that are not necessarily optimal in the long run. For example, with RL a potential reward to be maximized could be the number of civilians saved. We will use RL to optimize the behavior of the spokesperson. In addition, RL could be used offline for designing new plays to populate our playbook. That is, we can provide the swarm entities with rewards (based on individual and common goals), and let them train against one another in order to come up with the optimal play for tackling a particular problem. Because in this case multiple agents will be learning at the same time, we will have to use multi-agent RL, which unlike standard RL does not assume that the learning environment is stationary (Georgila et al., 2014; Xiao and Georgila, 2018; Bowling and Veloso, 2002).

The simulation may also be expanded to include other search and rescue scenarios or other swarm management scenarios, such as earthquake or flooding response. A swarm of robots able to seek out human life and assist emergency responders can, with the appropriate modifications to behavior, be of service in many different types of situations.

6. Conclusions

The use of swarm robots in search and rescue operations is relatively new, and the best practices for how these units can interact with humans, both from an operational perspective and a civilian perspective, are still being explored. The virtual reality simulation allows for a more realistic setting for participants to interact in, which provides ample opportunity to collect data on how responders and civilians may communicate with these robots.

Although the focus is currently on a wildfire scenario, future iterations of the simulation may include other scenarios. However, these first steps in the wildfire context will be of benefit towards creating an experimental framework. This framework can be used to assess what mode of communication must be supported in order to have a system that is effective and intuitive for users under extreme pressure.

Acknowledgements

This work was supported by the U.S. Army. Any opinion, content or information presented does not necessarily reflect the position or the policy of the United States Government, and no official endorsement should be inferred.

7. References

Bethel, Cindy L. and Robin R. Murphy, 2006. Affective expression in appearance constrained robots. In *ACM SIGCHI/SIGART International Conference on Human-Robot Interaction (HRI)*.

Bowling, Michael and Manuela Veloso, 2002. Multi-agent learning using a variable learning rate. *Artificial Intelligence*, 136(2):215–250.

Casper, Jennifer and Robin R. Murphy, 2003. Human-robot interactions during the robot-assisted urban search and rescue response at the World Trade Center. *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, 33:367–385.

Casper, Jennifer L., Mark Micire, and Robin R. Murphy, 2000. Issues in intelligent robots for search and rescue. In *Unmanned Ground Vehicle Technology II*, volume 4024. International Society for Optics and Photonics.

CRASAR, 2019. The Center for Robot-Assisted Search And Rescue. http://crasar.org/.

Dahlbäck, N., A. Jönsson, and L. Ahrenberg, 1993. Wizard of Oz studies – why and how. *Knowledge-Based Systems*, 6:258–266.

Georgila, Kallirroi, Claire Nelson, and David Traum, 2014. Single-agent vs. multi-agent techniques for concurrent reinforcement learning of negotiation dialogue policies. In *Annual Meeting of the Association for Computational Linguistics (ACL)*.

Green, A., H. Huttenrauch, and K. S. Eklundh, 2004. Applying the Wizard-of-Oz framework to cooperative service discovery and configuration. In *RO-MAN IEEE International Workshop on Robot and Human Interactive Communication*.

Jacoff, Adam, Elena Messina, Brian A. Weiss, Satoshi Tadokoro, and Yuki Nakagawa, 2003. Test arenas and performance metrics for urban search and rescue robots. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)(Cat. No. 03CH37453)*, volume 4.

Jennings, James S., Greg Whelan, and William F. Evans, 1997. Cooperative search and rescue with a team of mobile robots. In *International Conference on Advanced Robotics (ICAR)*.

Kolling, Andreas, Katia Sycara, Steven Nunnally, and Michael Lewis, 2013. Human swarm interaction: An experimental study of two types of interaction with foraging swarms. *Journal of Human-Robot Interaction*, 2:103–129.

Kolling, Andreas, Phillip Walker, Nilanjan Chakraborty, Katia Sycara, and Michael Lewis, 2016. Human interaction with robot swarms: A survey. *IEEE Transactions on Human-Machine Systems*, 46:9–26.

Lukin, Stephanie M., Felix Gervits, Cory J. Hayes, Pooja Moolchandani, Anton Leuski, John G. Rogers III, Carlos Sanchez Amaro, Matthew Marge, Clare R. Voss, and David Traum, 2018. ScoutBot: A dialogue system for collaborative navigation. In *Proceedings of ACL 2018, System Demonstrations*. Melbourne, Australia: Association for Computational Linguistics.

Marge, Matthew, Claire Bonial, Ashley Foots, Cory Hayes, Cassidy Henry, Kimberly A. Pollard, Ron Artstein, Clare R. Voss, and David Traum, 2017. Exploring variation of natural human commands to a robot in a collaborative navigation task. In *Workshop on Language Grounding for Robotics, Association for Computational Linguistics*.

Messina, Elena and Adam Jacoff, 2006. Performance standards for urban search and rescue robots. In *Unmanned Systems Technology VIII*, volume 6230. International Society for Optics and Photonics.

Murphy, Robin R., 2004. Human-robot interaction in rescue robotics. *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, 34:138–153.

Passonneau, Rebecca J., Susan L. Epstein, Tiziana Ligorio, and Joshua Gordon, 2011. Embedded wizardry. In *Annual Meeting of the Special Interest Group on Discourse and Dialogue (SIGDIAL)*.

Penders, Jacques, Lyuba Alboul, Ulf Witkowski, Amir Naghsh, Joan Saez-Pons, Stefan Herbrechtsmeier, and Mohamed El-Habbal, 2011. A robot swarm assisting a human fire-fighter. *Advanced Robotics*, 25(1-2):93–117.

Shah, Binoy and Howie Choset, 2004. Survey on urban search and rescue robots. *Journal of the Robotics Society of Japan*, 22(5):582–586.

Stein, Mark, 2004. The critical period of disasters: Insights from sense-making and psychoanalytic theory. *Human Relations*, 57(10):1243–1261.

Sycara, Katia and Michael Lewis, 2012. Human control strategies for multi-robot teams. In *WSEAS International Conference on Computers*.

Traum, David, Jeff Rickel, Stacy Marsella, and Jonathan Gratch, 2003. Negotiation over tasks in hybrid human-agent teams for simulation-based training. In *Proceedings of AAMAS 2003: Second International Joint Conference on Autonomous Agents and Multi-Agent Systems*.

USFA, 1995. Technical rescue program development manual. Report FA-159, 2/96, United States Fire Administration, Federal Emergency Management Agency.

Velagapudi, Prasanna, Paul Scerri, Katia Sycara, Wang-Hsuan Lee, and J. Wang, 2008. Scaling effects in multi-robot control. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*.

Wang, Huadong, Michael Lewis, Prasanna Velagapudi, Paul Scerri, and Katia Sycara, 2009. How search and its subtasks scale in n robots. In *ACM/IEEE International Conference on Human-Robot Interaction (HRI)*.

Xiao, Gang and Kallirroi Georgila, 2018. A comparison of reinforcement learning methodologies in two-party and three-party negotiation dialogue. In *International FLAIRS Conference*.