Collaborative Information Sensemaking for Multi-Robot Search and Rescue

Vladimir Zadorozhny
School of Information Sciences
University of Pittsburgh
Pittsburgh, PA
vladimir@sis.pitt.edu

Pei-Ju Lee
School of Information Sciences
University of Pittsburgh
Pittsburgh, PA
pel30@pitt.edu

Michael Lewis
School of Information Sciences
University of Pittsburgh
Pittsburgh, PA
mlewis @sis.pitt.edu

ABSTRACT

In this paper, we consider novel information sensemaking methods for search and rescue operations that combine principles of information fusion and collective intelligence in scalable solutions. We will elaborate on several approaches that originated in different areas of information integration, sensor data management, and multi-robot urban search and rescue missions.

Keywords

Mobile robots, information fusion, crowdsoursing, sensemaking, search and rescue mission.

1. INTRODUCTION

Efficient utilization of large robot teams for urban search and rescue (USAR) operations is a grand challenge in the design of advanced mobile cyber-physical systems. A basic approach assumes that human operators navigate robots through the environment and gather information about locations of immobilized, or moving victims. The operators observe video feeds from the robots to detect the victims and to direct the rescue mission. One operator may control multiple robots in a round-robin style. Robots can explore different areas and concurrently produce video streams.

The basic approach does not perform well for large multi-robot systems with numerous victims spread over large areas. As the scale of the search and rescue mission increases, the level of the operator's load and number of detection errors become significant. After observing a victim in a video frame, the operator estimates proximity of the victim to the robot and marks the coordinates on a map. This sequence of actions may distract the operator from noticing other victims in the queue of images. As a result, some victims may be overlooked while other victims are double-counted, which may mislead the rescue team.

In this paper we propose to approach this task of large-scale data utilization and information sensemaking via collaborative efforts within a large network of mobile robots and human observers. Our major contribution is a systematic exploration of the principles of crowdsourcing and collective intelligence to support efficient information fusion for large-scale multi-robot search and rescue missions. The related technologies have evolved independently over a long time. In prior work we investigated several methods to exploit their synergy in the context of the USAR missions [ZL3a,ZL13b]. In this paper, we extend and systematically consider the collaborative information fusion and sensemaking to estimate victim presence probability, as well as number of victims in particular location at specific time interval.

The concept of information sensemaking and fusion has been applied in various domains. While there are various interpretations of the term "sensemaking", conceptually it refers to how people understand complex phenomena, often overcoming "gaps" in reality via ideas, emotions, thoughts, etc., making sense of situations and enacting this sense into the world they explore [PR10, De03,We95]. Many studies of sensemaking have been conducted in the context of individual information seeking tasks, while some of the efforts are focused on sensemaking in collaborative environments [PR10]. Collaborative sensemaking involves multiple people. [Mu08] reports on applicability of sensemaking methods in crisis situations. Related concept of information fusion refers to multi-sensor data fusion [Ha04], information fusion for data integration [ZH13, BN08], and more recently, human-centered information fusion methods [HJ10]. The common data fusion algorithms perform some kind of data aggregation (e.g., averaging, or weighted averaging in extension) [Ha90]. More sophisticated probabilistic and statistical techniques are required as the number of sensors grows, which may also cause severe data conflicts and data inconsistencies. Our proposed collaborative framework is extendable and can utilize various invormation fusion techniques.

A crowdsourcing process commonly include the following components: (1) dividing the tasks into *microtask*s , (2) motivating users to contribute in solving the microtasks and (3) combining user solutions (responses) into *consensus* solution [Ba11]. In this paper we propose to define (1) so as to perform (3) automatically (via automatic information fusion). This is in contrast to related works, where generating consensus require notable efforts from domain experts. For example, [Ba11] explores feasibility of crowdsourcing solutions to utilize large aerial and satellite image datasets of a disaster area. An online user community (crowd) can annotate parts of images to identify, classify, and prioritize damaged regions. The crowd can include both experts and general public.

We are not aware of any related works that would conduct a systematic study of sensemaking methods in the context of automatic information fusion based on crowdsourcing. The paper is organized as follows. In the next section we explain our collaborative information sensemaking framework for multi-robot search and rescue missions. Section 3 elaborate on the information fusion methods to make sense of annotated reports. We present some experimental results in Section 4. Section 5 concludes.

2. GENERAL FRAMEWORK FOR COLLABORATIVE INFORMATION SENSEMAKING

We propose a general framework for collaborative information sensemaking that utilize collective intelligence of mobile robots and human operators/observers in the process of victim detection. The framework is shown in Figure 1. The stream of images from numerous mobile robots is presented to a group of human observers/operators. The operators acknowledge presence of the victim in an image and estimate number of victims providing corresponding image annotation. The images together with human annotations represent *reports* from individual robots at different locations and time intervals. The task of finding victim location and estimating their distribution is performed via automatic fusion of the robot reports. This fusion is conducted continuously; as robots collectively explore larger areas, it is expected that the estimates converge to actual victim numbers in specific locations and within different time intervals. We assume that robots can explore only accessible areas. This approach can be combined with the traditional victim detection process involving an active operator who marks the victim locations on the map.

In the rest of the paper, we assume that each mobile robot is equipped with a scanning laser range finder (SLRF), - a high-resolution environmental sensor [Ca05]. The SLRF probes the search area with pulses of laser light and measures the round-trip time for each pulse. Using the round-trip time T, robots can estimate the distance to an obstacle along the laser beam as $(C \times T)/2$, where C is the speed of light. The accuracy of the SLRF depends on scanning frequency (how often the full range is swept by the scanning sensor, e.g. 10Hz), scanning resolution (angular distance between two consecutive laser beams, e.g., 1 degree), and round-time measurement precision.

SLRFs have become very popular in mobile robotics applications. Meanwhile, the high frequency and high-resolution data generated by these sensors are still underutilized for automatic information fusion that could significantly facilitate and optimize multi-robot search and rescue missions. Design and development of such information fusion methods is a major contribution of this paper. We propose to integrate numerous laser scans from multiple mobile robots to help human operatosr with estimating victim locations. Our approach utilizes the concept of an occupancy grid representing a map of the environment as an evenly spaced field of random variables [El89,Ko94]. Each variable corresponds to a cell of the occupancy grid - a map quadrant; it reflects the presence of a victim in that quadrant. Our information fusion methods utilize laser scans from mobile robots to compute posterior estimates for these random variables.

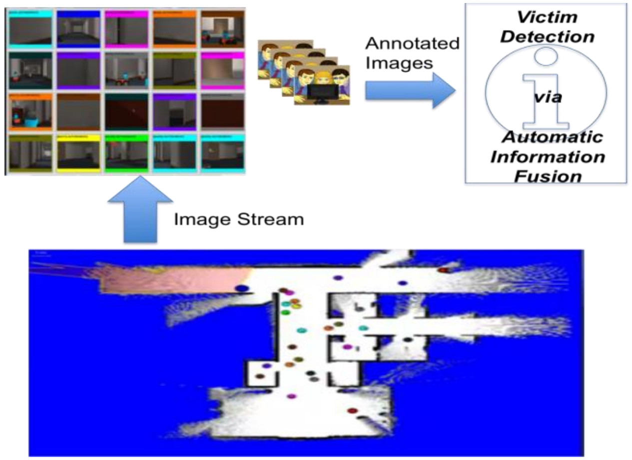

Figure 1: Collaborative Information Fusion Framework

Figure 2 shows one laser scan of a round area within an environment split into cells of an occupancy grid. The reflecting beams of the laser scan reveal the presence and locations of two circular objects in the scanned area. The operator would annotate images obtained from both robots with a victim presence tag, as well as with number of observed victims.

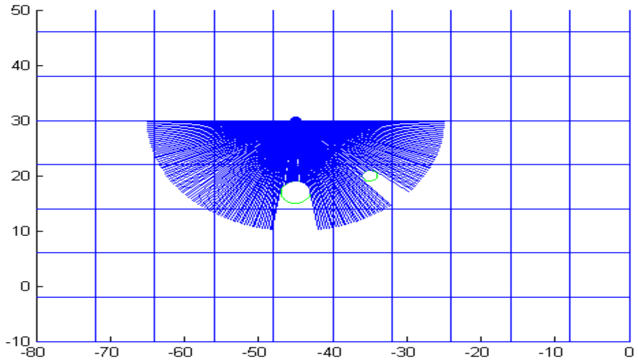

Figure 2: Laser scan of a round area with two circular objects

3. MAKING SENCE OF ANNOTATED REPORTS

3.1 Estimating Victim Presence Probabilities

Our collaborative framework continuously collects and processes reports from individual robots. The reports include robot observations and human annotations about victim presense in particular image, as well as number of victims. First, we explain how this information can be utilized to esimate victim presence probability in a specific cell of occupancy grid. In the next subsection we will elaborate on esimating number of victims in each cell.

In our previous work [ZL13a], we introduced a basic information fusion approach to estimate victim distribution over the cells of the occupancy grid. This approach considers any grid cell overlapping with robot scan lines as a potential victim location. We call it a *potential victim cell (PVC)*. The *victim presence probability (VPP)* in a *PVC* can be estimated as a ratio of number of victim scans overlapping with the cell *VS* to the total number *N* of victim scans, i.e., $VPP = \frac{VS}{N}$. For the example, $N = 14$ and the cell where the victim is actually located has the highest estimated victim presence probability of *2/14*. For the rest of the *PVCs* this probability is *1/14*. We demonstrated that as the number of scans grows, the estimated probability distribution converges to the actual distribution of victims over the occupancy grid.

In [ZL13b] we considered more advanced and more accurate information fusion methods. Taking into account the information from robots that currently do not observe any victims can increase the accuracy of the estimates for victim presence probabilities. Their scans are empty, i.e., an operator does not annotate them with a victim presence tag. If a cell overlaps with both victim scans and empty scans, we call it a controversial cell (*CC*). We applied various probabilistic sensor models [Ko94] to estimate the victim presence probabilities. In particular, we use Bayes' rule to estimate the *VPP* as a conditional probability $P(V|VS)$, where V is a property reflecting victim presence in a cell, VS is a condition that the cell overlaps with a victim scan. This probability can be estimated as follows:

$$VPP = P(V\,|\,VS) = \frac{P(VS|V)P(V)}{P(VS|V)P(V) + P(VS|noV)P(noV)},$$

where $P(V)$, $P(noV)$ are the prior probabilities of victim presence and victim absence in the cell, $P(V) + P(noV) = 1$. The prior probabilities can also be assigned based on such factors as strength of the disaster hit, time and characteristics of the disaster area.

3.2 Estimating Number of Victims

A major challenge in estimating number of victims form multiple annotated reports is handling redundant and, possibly, inconsistent information obtained from robots and human observers. The infromation related to the number of oserved victims presented in the annotated reports may have both temporal and spatial redundancy as we explain below.

Temporal Redundancy. It is possible to have multiple concurrent reports about the number of victims in the same occupancy grid cell within *overlapping time intervals*. Figure 3 shows an example of concurrent reports from two robots including observed number of victims in a grid cell *C1* (*Report_1* and *Report_2*) within overlapping time intervals *10:00-10:30* and *10:15-10:45*. The number of victims in two reports may differ due to obstacles preventing robots from observing all victims in the cell, or due to moving victims. We cannot simply add the numbers of victims to find the total number of victims in the cell *C1* from *10:00* to *10:45*. There is a temporal redundancy between *Report_1* and *Report_2.*

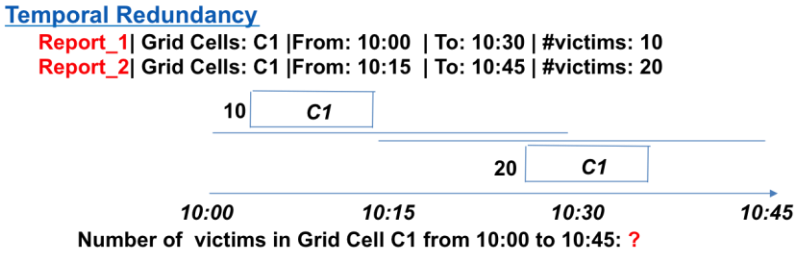

Figure 3: Example of Temporal Redundancy

Spatial Redundancy. We may also have multiple robots reporting on number of victims in *overlapping locations*. Figure 4 shows an example of reports from two robots for the number of victims observed in cells *C1*, *C2* (*Report_3*) and *C2*, *C3* (*Report_4*) within the same time interval from 11:00 to 11:30. We cannot simply add up their corresponding numbers to obtain the total number of victims in cells *C1*,*C2,* and *C3*. There is a spatial redundancy between *Report_3* and *Report_4*.

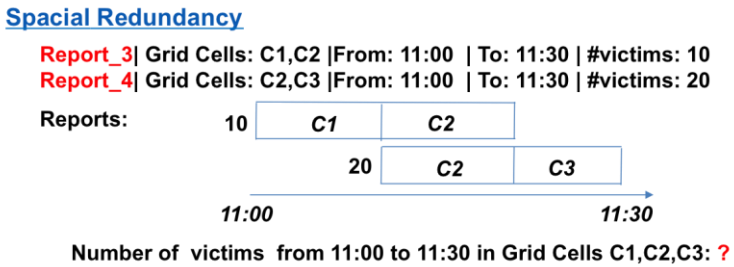

Figure 4: Example of Spatial Redundancy

Spatial and Temporal Redundancy. Finally, we may have multiple robots reporting on number of victims in *overlapping locations and for overlapping time intervals*. Figure 5 shows an example of reports from four robots for the number of victims observed in cells *C1, C2* and *C2, C3* within overlapping time interval from 10:00 to 10:30 and from 10:15 to 10:45. We cannot simply add up their corresponding numbers to obtain the total number of victims in cells *C1*,*C2*, and *C3* from *10:00* to *10:45*. There is a spatial redundancy between *Report_3* and *Report_4*.

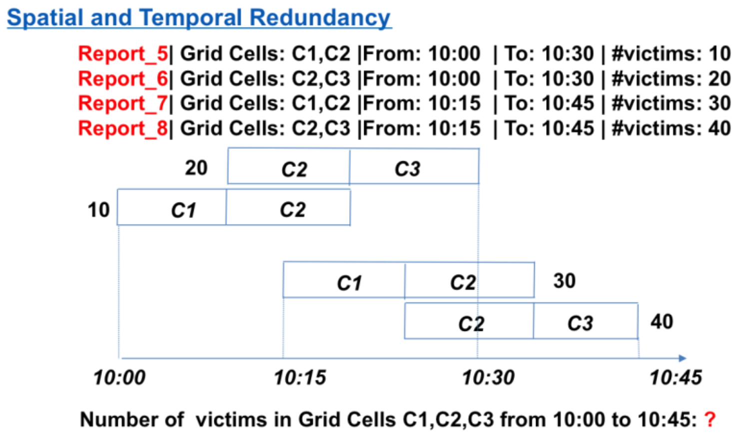

Figure 5: Example of Spatial and Temporal Redundancy

We can perform fusion of redundant annotated reports along temporal or spatial dimensions (one-dimensional information fusion), or along both temporal and spatial dimensions (two-dimensional information fusion). We will elaborate on it in the next two subsections.

3.3 One-Dimensional Information Fusion

First we consider a one-dimentional information fusiton utilizing either spatial or temporal redundancy. We propose to constructs an underdetermined linear system corresponding to redundant reports from multiple robots (*characterisitic linear system*). The goal is to estimate a number of victims per cell of an occupancy grid and per selected time interval as an approximate solution of the characteristin linear system.

Consider a simple example of merging reports from four robots (Table I). The four reports reflect detection of victims in an occupancy grid of size *3x2*. Here *Ri* represents a report on total number of victims *Vi* in occupancy grid cells *covered by Ri*. The overlapping reports cover the whole occupancy grid as showed in Figure 6.

Table I. Example of spatially overlapping reports

Report ID R_i	Covered Space Unit	Report Value
R_1	C_1, C_2, C_3, C_4	V_1 (700)
R_2	C_2, C_3, C_4	V_2 (500)
R_3	C_3, C_4, C_5	V_3 (600)
R_4	C_4, C_5, C_6	V_4 (700)

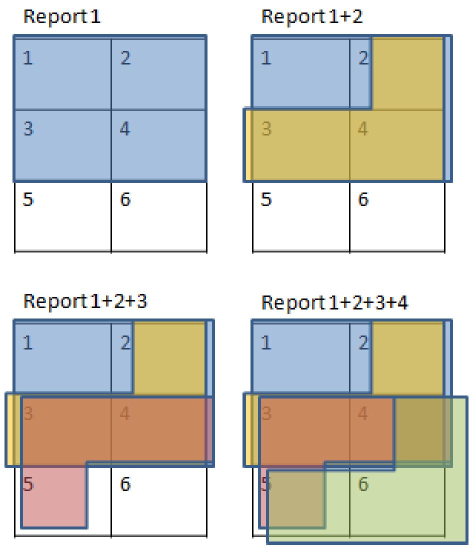

Figure 6. The overlapping reports cover the whole occupancy grid

Finding the th number of victims over space intervals can be represented as the following linear optimization problem.

$$\begin{aligned} &\textit{Max.} && x_1+x_2+x_3+x_4+x_5+x_6 \\ &\textit{Subject to} && x_1+x_2+x_3+x_4 = 700 \\ &&& x_2+x_3+x_4 = 500 \\ &&& x_3+x_4+x_5 = 600 \\ &&& x_4+x_5+x_6 = 700 \\ &&& x_1, x_2, x_3, x_4, x_5, x_6 \geq 0, \end{aligned}$$

where *xi* is an estimated number of victims in an occupancy grid cell *Ci*. The underdetermined linear system can be represented in matrix form ***AX=b***:

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{bmatrix} = \begin{bmatrix} 700 \\ 500 \\ 600 \\ 700 \end{bmatrix}.$$

An optimal solution for linear programming system can be found using different methods such as the basis exchange method, the branch and cut method, etc. In this paper, we apply non-negative least-square method to find a solution of our characteristic linear system.

Solving linear equations ***AX* = *b*** using non-negative least-square method estimates ***X*** as $\boldsymbol{X' = A^T(AA^T)^{-1}b}$. For our example the solution is as follows:

$$\boldsymbol{A^T}(\boldsymbol{AA^T})^{-1}\boldsymbol{b'} = \boldsymbol{X'} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{bmatrix} = \begin{bmatrix} 200 \\ 0 \\ 0 \\ 500 \\ 100 \\ 100 \end{bmatrix}.$$

After that we substitute estimated solution ***X'*** in the original equation to obtain estimated reported values ***b''***. The matrix ***b''*** generated by the solution set X' for our example is as follows:

$$\boldsymbol{AX'} = \boldsymbol{b''} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{bmatrix} = \begin{bmatrix} 700 \\ 500 \\ 600 \\ 700 \end{bmatrix}.$$

Next we estimate a difference between estimated values ***b''*** and actual values ***b***. For our example:

$$\boldsymbol{\delta} = \begin{bmatrix} \delta_1 \\ \delta_2 \\ \delta_3 \\ \delta_4 \end{bmatrix} = \begin{bmatrix} b_1'' - b_1 \\ b_2'' - b_2 \\ b_3'' - b_3 \\ b_4'' - b_4 \end{bmatrix} = \begin{bmatrix} 700 - 700 \\ 500 - 500 \\ 600 - 600 \\ 700 - 700 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

If reported values are inconsistent, the above δ will be non-zero. We call this process of checking the difference between estimated value and actual values the *Reverse Substitution Method (RS)*. Using the RS method we monitor the number and distribution of victims in a specific area. Note, that for some cases we could do quick estimates before solving the full linear system. For example, from reports R_1 and R_2 in Table 1 we can quickly conclude that number of victims in C_4 is 200.

Our approach can be used for continuos information fusion in collaborative enviroments. Figure 7 consider a series of 10 time units (*TU*); the number of spatial reports is accumulated as the number of time unit increases. In other word, there is only one spatial report at *TU1*, two reports at *TU2* since moving robots have explored more areas, and so on. We expect that the victim detection accuracy will increase with time, since the number of reports is increasing. However, the detection delay will also increase.

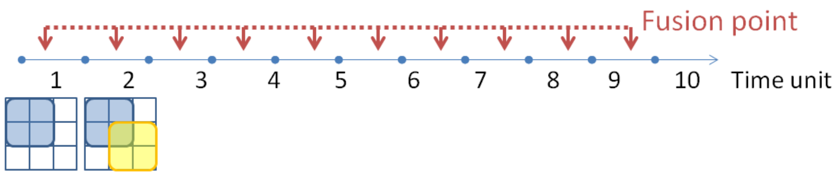

Figure 7. Continuous information fusion

3.4 Two-Dimensional Information Fusion

In the previous sub-section we introduced one-dimensional information fusion for redundant reports and illustrated this approach on one-dimensional spatial information fusion. Similar we could perform one-dimensional temporal information fusion with overlapping reports with temporal redundancy. In this subsection we outline the idea of two-dimensional information fusion combining spatial and temporal fusion to provide better accuracy in inconsistency detection, and to provide more efficient victim detection in each location for a specific time interval.

The idea of the two-dimensional information fusion studies is to use temporal fusion to estimate the number of targets per time interval for a giving location of reports (according to the laser data of robots, location can be defined by a group of points) and to use the spatial fusion to estimate the number of targets per location (group of points) for a given time interval. This would require to maintain and two characteristic linear systems – one for temporal fusion that estimates total number of victims per time intervals, and another one for spatial fusion that estimates number of victims in a specific location for specific time interval. With the estimated values generated by multiple linear systems for each time and location, we can refine positions of immobilized victims and describe the moving trajectories of mobile victims/targets.

To sum up, we propose the following cases for the collaborative information fusion: (1) static victims, static robots; (2) static victims, moving robots; (3) moving victims, static robots; (4) moving victims and moving robots. While two-dimensional information fusion would work for all of them, most efficiently it would be utilized for the most challenging case of moving victims and moving robots.

4. SOME RESULTS

We used the USARSim framework [Ca06] to explore the impact of different information fusion strategies on the convergence of estimated probability distribution to the actual distribution of victims as the number of laser scans increases. USARSim is a high fidelity simulation framework developed at the Carnegie Mellon University Robotic Institute. USARSim is now maintained by NIST personnel on SourceForge, where it has been downloaded more than 70,000 times. Numerous validation studies show close agreement in behavior between USARSim models and real robots. We set up a scenario with 24 robots exploring an urban environment of 80x60 unit cells with 32 victims. The total number of laser scans collected for 15 minutes of a real-time simulation exceeded 21,000. We performed large-scale annotation of victim images based on visual ranges of robot laser scans. I.e., each time a laser scan hits a victim in a visual range of the robot camera the image is annotated with a victim tag.

Figure 7 shows estimated distribution convergence to actual victim distribution for different granularities of the occupancy grid. We used Jensen-Shannon

Divergence (JSD) [Li91] to measure similarity between estimated and actual victim distributions at different stages of the search mission (the smaller JSD reflects higher distribution similarity). Each point represents a state of information fusion after approximately 200 visual scans have been collected. We observe better convergence for a coarser granularity of the occupancy grid, where we estimate VPP in larger map quadrants. It is possible to conclude that the naïve distribution estimation can be used to roughly outline the victim locations and select large areas with the higher VPP to optimize the search and rescue. Meanwhile, the naïve approach does not converge well for finer granularity occupancy grids, which requires higher victim detection accuracy.

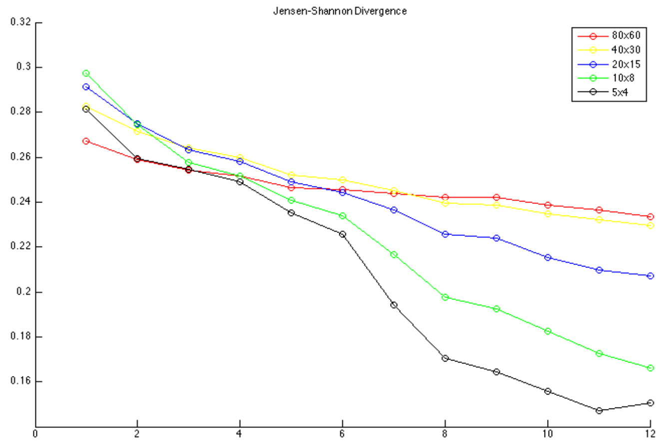

Figure 8: Convergence of estimated victim distribution

We also obtained preliminary performance estimations of our one-dimentional spatial fusion method (reffered to as RS in Figure 9, for *Reverse Substitution* method explained in Section 3.3). For this study we set up three different levels of sparsity. Sparsity refers to how many spatial cells have zero victim. We assumed up to 20 robots exploring the spatial environment of 36 space units of an occupancy grid within 20 time units. Figure 9 compares the relative detection error (Relative Distance, RD) of the RS method under different sparsities. The RD value reflects the relative difference between the estimated and the actual number of victim in each space unit. Figure 9 shows that the RDs in these three scenarios are close. The medium sparsity corresponds to the highest RD followed by low and high sparsity.

In both cases (Figure 8 and Figure 9) we observe performance improvements for continuously accumulated reports in our collaborative environment. We conclude that our collaborative approach represents a feasible solution for information sensemaking in large-scale search and rescue missions.

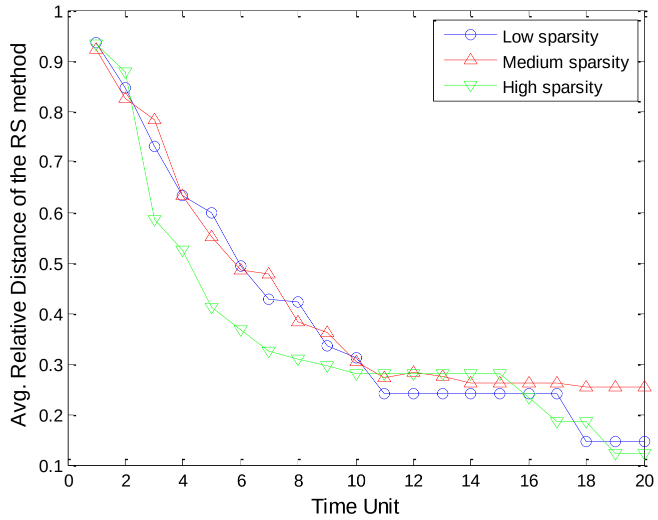

Figure 9. Relative distance of the RS method

CONCLUSION

We introduced a novel collaborative framework for scalable information sensemaking in search and rescues operations. Our framework utilizes collaborative efforts of a large network of mobile robots and human observers. We explained and demonstrated how our proposed framework can be efficiently used for the task of estimating probability of victim presence, as well as estimating number of victims in particular location. Our framework is extendable; it can utilize new information fusion techniques, as well as combine existing methods for better information sensemaking strategies.

REFERENCES

[Ba11] Barrington, L., Ghosh, S., Greene, M., Har-Noy, S., Berger, J., Gill, S., Lin, A., Huyck, C. Crowdsourcing earthquake damage assessment using remote sensing imagery. Annals of Geophysics, 54(6), 2011.

[BN08] Bleiholder, J., Naumann, F. Data Fusion. ACM Computing Surveys 41(1), 2008.

[Ca06] Carpin, S., Stoyanov, T., Nevatia, Lewis, M., & Wang, J. Quantitative assessments of USARSim accuracy, *Proceedings of PerMIS*, 2006.

[Ca05] Carpin, S., Wang, J., & Lewis, M. High fidelity tools for rescue robotics: Results and perspectives, Bredenfeld (Ed.) *Robot Soccer World Cup IX*, Berlin-Heidelberg: Springer LANI, 2005.

[De03] Dervin, B., L. Foreman-Wernet, and E. Lauterbach, eds. Sense-Making METHODOLOGY Reader: Selected Writings of Brenda Dervin. Hampton Press Inc., Cresskill, NJ, USA, 2003.

[El89] Elfes, A. Using occupancy grids for mobile robot perception and navigation. *Computer,22(6),*1989.

[Ha04] Hall,D. Mathematical techniques in multi-sensor data fusion. Artech House, 2004.

[Ha90] Hackett, J. and Shah,M. Multi-sensor fusion: a perspective, in Robotics and Automation, IEEE, 1990.

[HJ10] Hall,D, Jordan,J. Human-centered information fusion. Artech House, 2010.

[Ko94] Konolige, K. Improved occupancy grids for map building. *Autonomous Robots. 4(4)*, 1994.

[Li91] Lin, J. Divergence Measures Based on the Shannon Entropy. IEEE Transactions on Information Theory, 37(1), 1991.

[Mu08] Muhren. W., Van Den Eede, G., Van de Walle, B. Sensemaking as a Methodology for ISCRAM Research: Information Processing in an Ongoing Crisis. Proceedings of ISCRAM, 2008.

[PR10] Paul, S., Reddy, M. Understanding Together: Sensemaking in Collaborative Information Seeking. Proceedings of CSCW, 2010.

[We95] Weick, K.E. Sensemaking in Organizations. Sage Publications, Thousand Oaks, CA, USA, 1995.

[ZL13a] Zadorozhny, V., Lewis, M. Information Fusion based on Collective Intelligence for Multi-Robot Search and Rescue Missions. Proceedings of the 14th International Conference on Mobile Data Management (MDM'13), 2013.

[ZL13b] Zadorozhny, V., Lewis, M. Information Fusion for USAR Operations based on Crowdsourcing. Proceedings of the 16th International Conference on Information Fusion (FUSION'13), 2013.

[ZH11] Zadorozhny, V. Hsu, Y.-F. Conflict-aware Fusion of Historical Data. Proc. of the International Conference on Scalable Uncertainty Managemnt (SUM'11), 2011.