Gunshot Detection System for JTRS Radios

Michael Smith, Scott Buscemi, and Dexiang John Xu, Ph.D.
NED T&E; Software Integration and Advanced Technology Lab
SPAWAR SYSTEM CENTER at Charleston SC

Abstract— This paper presents a novel solution for detecting and locating gunshots using minimal equipment on a JTRS radio is proposed. Each radio acts as a sensor node to determine and share muzzle blast time of arrival information in order to determine a shooter location. The feasibility of such a system was investigated by developing detection and location algorithms and testing each algorithm's effectiveness in both simulation and with real data. The detection algorithm is used to accurately pinpoint the arrival of the muzzle blast of a gunshot at a single microphone and was realized using a rake-correlation filter loop. The location algorithm used the position and time of arrival information gathered from multiple sources to determine the shooter location and was realized by an extended time-invariant Kalman filter. The Kalman filter is very adaptable and can incorporate additional measurements or sensors such as terrain information to eliminate large vertical errors; or Angle of Arrival information to incorporate additional gunshot systems into the network. Results showed that in realistic scenarios, with only time of arrival information, gunshot location accuracy is dictated by radio positioning.

Keywords- gun shot; JTRS; Kalman; detection; correlation filter, Cramer Rao

I. INTRODUCTION

JTRS radios provide soldiers with the ability to form an ad-hoc network, capable of reporting a soldier's position with military GPS units [1]. This project investigated the feasibility of using this network to create a low cost, light weight system to detect gunshots and locate shooters. The gunshot detection system proposed would offer several advantages over similar systems:

- Integration into the JTR set will enable immediate and automatic communication of the gunshot analysis information in a battlefield environment.
- The war fighter would be required to carry little extra equipment.
- Any existing gunshot monitoring device, such as Boomerang, could be incorporated in this system to improve accuracy of both systems and communicate locations.
- Integration into the battlefield communications infrastructure will obviate the need for additional dedicated frequencies.

Currently, there exist several different shooter location systems. One that is fielded for the military is Boomerang [2]. It uses a fixed array of microphones mounted on a HMMWV (Humvee) to locate a shooter's position. It does not currently have the ability to take advantage of networked radios. A system that uses networked sensors is the Shot Spotter [3]. It is a civilian application popular with police agencies that can quickly alert and direct law enforcement to shots being fired. This system requires fixed detector positions and an array of

microphones at each sensor. Each sensor operates autonomously and shooter position is reported through the network.

The gunshot location system in this project was developed on the basis of using a wireless ad-hoc network with single microphones on each node and no fixed node locations. This setup makes detection harder. Mobile nodes require the use of inaccurate GPS data and the absence of microphone arrays forces the system to work with less information. Each sensor alone is unable to determine Angle of Arrival (AoA) information and therefore cannot operate autonomously.

The algorithms can be run on the radios by utilizing available processor resources. Alternatively, a small dedicated processor could be included in the microphone hardware to process the algorithms and interface with the radio's data port.

Finding a shooter location is divided into two major categories. The first is detecting the gunshot and estimating its time of arrival. There are two components to most gunshots: a shockwave and a muzzle blast. The focus in this project is on detecting the muzzle blast. The muzzle blast is the easier of the two to detect and it does not require as high quality audio recording as it does to detect the shockwave. Shockwaves have a much lower duration in time and fewer distinguishing features than muzzle blasts. Shockwave detection may be included in future iterations of the software. The next step is to use the time of arrival information from each radio node to determine the shooter position. This problem is an inverse problem which cannot be solved for directly. The algorithm chosen to solve this problem was an extended time-invariant Kalman filter, due to its quick convergence and adaptability. The Kalman filter can easily be expanded to take in different types of measurements and has the ability to improve current technologies such as the Boomerang system through networking.

II. ALGORITHMS

A. Detection

The sound of a gunshot often depends more on the environment in which the gun was fired than the gun itself. The sound of a gunshot can vary greatly between different locations and microphone placements [4]. In this project it was assumed the gunshots are in an outdoor environment.

The criteria for judging the effectiveness of a detection algorithm is the False Positive Rate (FPR) and False Negative Rate (FNR). FPR is defined as the number of false positives divided by the total number of negative instances. FNR is defined as the total number of false negatives divided by the total number of positive instances. Other criteria include

algorithm complexity, runtime, and timing accuracy [5]. It is important to reduce complexity and runtime so the algorithm can be implemented on low-cost platforms and operate real-time. The accuracy of an algorithm detecting the Time of Arrival (ToA) of a gunshot is important for locating the shooter [6].

The optimal solution was found to be a combination of a correlation filter and rake receiver [7]. The correlation filter helps resolve the signal from uncorrelated surrounding noise. The rake receiver helps eliminate much of the multipath present in the signal. The implementation is low cost requiring three filters of 60 taps each and had a 10% FNR and a 3% FPR. The algorithm works on overlapping windows of 48K samples, or one second of data. The windows are overlapped 50% so that each sample is processed twice, as shown in Figure 1.

Figure 1. Gunshot Detection Process Sequence

Each window is passed into the correlation filter. The correlation filter is obtained through the following steps:

1. Obtain several high quality recordings of gunshots on high quality microphones.
2. Average the gunshots together. The gunshots can be lined up in the time domain by finding the maximum power within the signal.
3. Pass the averaged gunshot through the microphone that the soldiers will use. The filter will start from where the gunshot first rises above and lastly falls below three standard deviations from the noise floor.

The signal out of the correlation filter is tested for peaks. If no peaks are detected above three standard deviations from the noise floor the signal is considered to have no gunshots. If peaks are found the signal is processed further. The signal is tested for up to 8 peaks; each peak having a minimum separation of 64 samples.

The peaks are then used to create a rake filter. The taps of the filter are delayed by the same number of samples the peaks are separated and the weights are set using the maximum ratio-combining principle. The rake filter is then applied to the original signal.

The process can then be repeated, where the signal that passed through the rake filter is now the input signal. In

practice the optimum result was obtained by repeating the process three times.

At the final iteration the signal through the rake receiver is passed through the correlation filter. The peak value of this output is used to determine if a gunshot is present.

Once the signal is processed, the absolute value of the signal is taken. This output signal is then divided by its mean. This process normalizes the signal without using multiplication except in calculating the mean. If the output of this process passes a certain threshold, a gunshot is detected, and the time of the gunshot is reported. A simple diagram is presented in Figure 2 showing the flow of data through the detection algorithm.

Figure 2. Gunshot Detection Data Flow Diagram

B. Kalman Filter

The Kalman Filter used in this application is an extended and time-invariant Kalman filter. It is time-invariant in that the observations do not change from iteration to iteration, and it is extended because the equations that describe shooter location are non-linear. Currently, the algorithm runs on one set of observations that all occur from the same gunshot event. The filter is used to find a shooter position in space (x, y, and z) and time (time the gun was fired) for which the set of observations agree in the Minimum Mean Squared Error (MMSE) sense. Error in this case is defined by the difference between a measured Time of Arrival (ToA) and a calculated ToA. The calculated ToA is determined based on an estimation of the shooters position and the radio's location. There are sources of error in both the measured ToA and the calculated ToA. The measured ToA for each radio can only be as accurate as the detection algorithm allows it to be, and the calculated ToA has error added from the GPS measurements of the radio location.

The first step is to define the problem. The algorithm is looking for the shooter position in space and time, which minimizes the error seen by multiple observers. The estimate is:

$$\tilde{x} = [\tilde{x}_x \quad \tilde{x}_y \quad \tilde{x}_z \quad \tilde{t}_0]$$

Where $\tilde{x}$ is the estimated shooter position in the x, y, and z dimensions and $\tilde{t}_0$ is the estimated time the gunshot went off.

This estimated shooter location will be modified with each new iteration. During an iteration, the error in the shooter location estimate (y) is calculated for each radio according to the difference between the measured ToA and the calculated ToA.

$$y(r_i) = t_i - \left(\frac{1}{c} \|p_i - \tilde{x}\|_2 + t_0 \right)$$

Where y is a vector of length N (N = number of radios) and is therefore the residual for each radio. The position and time measured for the i^{th} radio is r_i and t_i respectively. The term

$\|r_i - \hat{r}_i\|_2$ is the distance between the i^{th} radio position and the estimated shooter location.

After the error is calculated, if the sum of squared errors is below a certain threshold, the shooter position is determined to be optimal in the MMSE sense. Otherwise, the iteration continues to determine which direction to step in order to reduce the error. To do this, the gradient field is calculated by taking the partial derivatives of the error with respect to shooter position in time and space.

$$H(i, j) = \frac{\partial y(r_i)}{\partial x(j)}$$

The result is an N by 4 matrix where H (i, j) is the derivative of the error of the i^{th} radio with respect to the j^{th} shooter dimension. The first three shooter dimensions are the x, y, and z location and the fourth is the time the gun was fired. Here are those derivatives:

$$H(i, j) | j \in [1, 2, 3] = \frac{z(j) - r_i(j)}{c * \|r_i - \hat{z}\|_2}$$

$$H(i, 4) = -1$$

The last thing that needs to be defined prior to the final calculations is the covariance matrix R. This is the covariance of measurement dimensions. The error in each dimension was considered to be independent. This makes the H matrix diagonal, with each non-zero entry being the variance of each dimension. From here, the Kalman filter equations can be followed to determine the next iteration's guess at the shooter location.

$$\begin{aligned} P_{k+1} &= (I - K * H) * P_k \\ S &= H P_k H^T + R \\ K &= P_k H^T S^{-1} \\ \hat{x}_{k+1} &= \hat{x}_k + g K_{k+1} y_k \end{aligned}$$

This algorithm's equations are set up such that each step should reduce the residual. However, since this is a non-linear Kalman Filter that is not guaranteed, if the step is sufficiently large to result in moving out of the zone of the local minima (the area that the gradient points towards the local minima), the algorithm will not converge. To combat this, a gain 'g' is introduced on the step size the above equations. This was initially set to 1/2 and was decreased if extensive oscillations were seen indicating an under-damped system. This gain helped to ensure an over-damped system such that no step would leave the zone of the local minima. In practice, it significantly helped the convergence of the algorithm in many cases; however, it did increase the number of iterations required to converge.

Setting the initial condition, $\hat{x}_k$ for the first iteration, is an important aspect of the algorithm. A good point to start is near the first radio that heard the gunshot as this radio must be the closest one to the shooter. The only problem is that the shooter cannot be placed in the exact same position as the radio. To solve this, the shooter position is started just off this radio, however this introduces a new problem, as the direction of the shooter is unknown. This is solved by running the Kalman filter multiple times each from a starting point one meter off the radio. The result with the smallest residual is used as the final shooter location.

Another addition was built into the algorithm that can be turned on and off. If the assumption is made that the shooter is on the ground, a map of the terrain elevation can be used as another measurement that can help find the location of the shooter. Terrain information would not be as useful in mountainous or urban regions. An X, Y, and Z matrix was included as an optional input to the algorithm. These define the terrain around the shooter's location. The beauty of the Kalman filter is that this is simply an addition to the current algorithm. To do this, the length of the h and H matrices are increased by one to account for this new measurement. First, a bilinear interpolation is performed on the current estimate of the shooter location and saved as z_s . This is used to calculate h:

$$h(N+1) = \frac{(x(3) - z_s)^2}{c}$$

The H matrix is the derivative of the residual with respect to the unknown variables. This means that the H matrix is the same as before, with an additional row for the new terrain measurement.

$$\begin{aligned} H(N+1, j \in [1, 2]) &= 2/c * (x(3) - z_s) * \frac{\partial z_s}{\partial x(j)} \\ H(N+1, 3) &= 2/c * (x(3) - z_s) * \left(1 - \frac{\partial z_s}{\partial x(j)}\right) = 2/c * (x(3) - z_s) \\ H(N+1, 4) &= 0 \end{aligned}$$

The partial derivative of z_s , with respect to position, is simply the gradient of the terrain which can be calculated ahead of time. Terrain is two dimensional; there is one measurement for each x and y pair, therefore, the derivative of z_s in the z direction is zero leaving a slightly different equation when the partial derivative in the z direction is taken. In addition, there is no dependence on the time the shot was fired; therefore, this partial derivative is zero. The remainder of the algorithm can be run as described before using the new h and H matrices.

In much the same way as terrain information is added, Angle of Arrival (AoA) information that could be incorporated as well. Additional rows would be added to the h matrix for each AoA measurement and additional rows and columns would be added to the H matrix. This is how it is possible to incorporate other systems with this system. By simply extending the Kalman filter we can include any measurement that could increase the accuracy of locating the shooter beyond both systems working independently.

C. The Cramer Rao Bound

The Cramer Rao Bound (CRB) is a statistical tool that defines a lower bound on the variance of an estimator. In other words, the CRB defines how accurate the solution to an algorithm can be, given a set of measurements. In this case, the estimator is the shooter location and the measurements are the radio positions and ToAs. The application of this bound to the problem, will give a "best case scenario" to compare the Kalman Filter Algorithm against. The key point is that no matter how good an algorithm is, it is theoretically impossible for it to be more accurate than the Cramer Rao Bound.

For this CRB case we are assuming a Multivariate Normal Distribution. Fortunately, this case has been well studied and the solution to the CRB has been defined. X is a random

variable defined as a multivariate normal distribution with mean: $\mu(\theta)$ and covariance $C(\theta)$. The variable θ is the unknown parameter; in this case, shooter position and time the shot was fired.

$$\begin{aligned} x &\sim N_d(\mu(\theta), C(\theta)) \\ \theta &= [s_x \quad s_y \quad s_z \quad t_s] \\ \mu_i(\theta) &= \frac{1}{c} \|r_i - s\|_2 \end{aligned}$$

The mean is a vector of length N (N =number of radios), where each entry is the time it should take a shot to travel from the shooter to the radio. Each measurement is independent; therefore, the covariance matrix will be diagonal with values only on the diagonal. $C(i,i)$ will have the total variance of all the measurement errors arising from radio i as reflected in the time measurement. Because of this, it is necessary to represent the error in radio position as an error in the TOA. The result is the following covariance matrix:

$$C[i, j] = \begin{cases} \left(3 * \frac{pos_sd}{c}\right)^2 + time_sd^2 & \text{for } i = j \\ 0 & \text{otherwise} \end{cases}$$

The pos_sd is the standard deviation in each of the x, y, and z dimensions, c is the speed of sound, and the $time_sd$ is the standard deviation measured in the TOA.

The general method of determining any CRB is to find the Fisher Information Matrix (FIM) and invert it. This case is a multivariate normally distributed case with independent measurements. As such, the FIM is calculated as follows:

$$I_{\mu, \mu} = \frac{\partial \mu^T}{\partial \theta} C^{-1} \frac{\partial \mu}{\partial \theta} \quad \text{where:} \quad \frac{\partial \mu}{\partial \theta_k} = \frac{r_k - s_k}{c * \|r_i - s\|_2}$$

This makes the FIM matrix, I , a 4 by 4 matrix with each element as defined above. Next, it is necessary to take the inverse of I to get the CRB.

Now that the CRB is obtained, the team must make sense of the four dimensional covariance matrix. It is not possible to visualize in four dimensions so an Eigen-analysis is done on the CRB matrix. The result is 4 orthogonal unit vectors in time and space, the Eigenvectors, and a scalar for each, the Eigen values. In all cases encountered, one of the Eigenvectors had a very large time component and a negligible spatial component. This vector was ignored and the time components of the other vectors were also ignored so that the error could be visualized. The Eigenvectors are plotted in their positive and negative direction and scaled by twice the square root of corresponding Eigen values (to get two standard deviations) to obtain an ellipse of error. Inside this ellipse is where 95% of the solutions will fall if the Kalman Filter is ideal, ignoring error in time. In addition, this defines the maximum accuracy of any algorithm attempting to locate a shooter.

There are a few caveats to this application of the CRB. First, ignoring the time error makes the shape of the error that the CRB predicts not completely correct. The actual accuracy will have a curve associated with the ellipse, which will make it look more like a banana in many cases. Visualize grabbing two ends of the ellipse and pulling away from the radios to

make the correct shape. The ellipse is a good estimate of the error and further analysis would have been past the point of diminishing returns.

The second caveat is that terrain information was not used in the determination of the CRB. This is because the ellipse was still a good estimation, with the exception of the z-dimension which was now locked to the terrain. The error in that dimension can simply be ignored when running the Kalman solution with the terrain.

D. Monte Carlo Simulations

In order to test and verify the Kalman filter, Monte Carlo Simulations (MCS) were run for different scenarios. A scenario consists of a set of radio positions, a shooter location, and, optionally, terrain information. From each location, the exact time of arrival of the sound of a gunshot is calculated for each radio. The Kalman filter is run multiple times, and with each iteration, a different amount of error is added to both the exact radio positions and the exact TOA measurements. These “noisy” measurements are given to the Kalman filter and a solution is found for the shooter location given the noisy measurements. Because the amount of error added with each iteration is different, the resulting shooter location is different. After several iterations (usually 100 to 1000) are completed; all of the calculated shooter positions can be plotted. The standard deviation of these locations can be determined and this can be compared to the CRB to verify the quality of the Kalman filter.

The amount of error to add during each iteration is very important because it is desirable to mimic the real world measurements as accurately as possible. A normal distribution was assumed for both the radio locations and the ToA measurements. For the TOA measurements, the team assumed that the ToA could be guessed within 4 milliseconds (ms), and so a standard deviation of 2ms was used. For the radio position, an assumption was made that there would be military accurate GPS. This accuracy is classified information and our assumption will be left out for security reasons. These educated guesses on standard deviation dramatically affect the calculated shooter positions in almost all scenarios.

III. FIELD TEST SETUP

The field test setup used the following equipment: Multimix8 Alesis 8-channel mixer, Nady U-81 Wireless 8-channel receiver, Windows pc (base station), 10 msi U100 laptops, and 10 GPS units.

The tests were conducted at a local shooting range. Eight wireless microphones and laptops were placed at various locations throughout the shooting range. Each laptop had a GPS unit attached. The laptops recorded all of the data collected by the GPS units. Two laptops were placed along the line of shooters to record the GPS data at those locations. The 8-channel mixer was connected to the base station so that each channel could be recorded simultaneously. The recordings could be started and stopped at the base station. The data collected was then processed by the gunshot detection algorithm to get the times of gunshots. The GPS data was averaged to approximate the military GPS accuracy.

IV. RESULTS

A. Gunshot Detection Results

The gunshot detection algorithm detected 90% of shots present. Ten percent of shots detected were false positive. The data revealed that the requirements that were originally placed for an impulse to be considered a gunshot had to be dramatically lowered. In testing with recordings found in online databases, the detection rate was 96% and there were no false positives. These recordings were done with very high quality microphones with very flat frequency responses up to 20 kHz. The microphones used in this test were only capable of recording frequencies up to 12 kHz [10].

In the case where the shooter is close to a microphone, the shock wave and the muzzle blast fall on top of each other making it harder to discern the muzzle blast. An example of a recording from our closest microphone is presented below with its output from the detection algorithm.

Figure 7. Gunshot Recorded with a Close Microphone

It is apparent that the detection algorithm is having difficulty separating the two. It is almost impossible to tell by looking at the recording where the muzzle blasts begins and the shockwave ends. In comparison, a recording of the same gunshot at a farther distance is shown in figure 8 as the shock wave is clearly in front of the muzzle blast. The detection algorithm also peaks much higher, at around 35, which is well above the previous shot that peaks below 15 at the output of the detection algorithm.

To increase the detection rate a tradeoff had to be made. The threshold was lowered for what was considered to be a gunshot. This increased the number of false positives. Figure 9 is an example of a false positive.

Figure 8. Gunshot Recorded with a Distant Microphone

Figure 9. False Positive Gunshot

The peak through the filter in this image is roughly 25. The signal is clearly an impulse. If the closer shots were removed, the threshold could be set higher and this false positive could have been removed. One solution for this problem is to have two detection algorithms; one for close shots and one for far shots. This would require more processing power, and would likely need an adjustment to the time of arrival for one filter versus the other. Additionally in this case the low amplitude of the false positive hints that it is not a gun shot. The peak in this case is on the order of 10 times lower than that of the previous signal as recorded by the microphone

B. Kalman Results-Monte Carlo Simulations

The results of the simulations varied dramatically on radio orientation. In general, the better separation of radios in a single direction, the more accurate the shooter location will be in that direction. Figures 10 and 11 illustrate this point.

Figure 11 is an example of a fairly good estimate. The solutions are accurate within 15 meters. Figure 10 shows how inaccurate the solution can be. Here the solution is accurate within 15 meters in one direction, but only accurate to 150 meters in the orthogonal direction. In the graphs below, the radio positions are eliminated and error bars are included for the CRB.

Figure 10. Monte Carlo Simulation: Large Variance

Figure 11. Monte Carlo Simulation: Small Variance

Figures 12 (Left) and 13 (Right): CRB of Fig. 10 and 11

Inspecting Figure 12, note that the CRB does not extend as far to the right as the Kalman filter MCS does. This is due to the caveat noted in CRB materials and methods section. Error in the fourth dimension (time) results in stretching the ellipse away from the radios. There are an infinite number of radio and shooter scenarios possible and the team analyzed many. These are typical of the majority of the results.

C. Field Test Results

Issues with obtaining access to a gunshot range proved a difficult problem to overcome. The team was able to get one day in which the equipment worked properly and data was recorded. The figure below shows the field test results.

Figure 14. Gunshot Detection in Rifle Shooting Range

Figure 14 shows clustering of the shots demonstrating that the experiment is repeatable, however, nearly all the results are

well forward of the shooting line. The GPS locations are accurate and the wireless microphone system recorded the sounds of the gunshots well. After examining the data, the arrival times at some of the microphones are not what would be expected which caused the systematic error on display.

V. CONCLUSION

The team successfully demonstrated the ability for a multiple radio system to identify the location of a gun shooter and the ability of such a system to incorporate stand alone systems to improve the accuracy over both systems independently. The team demonstrated that the simulations are valid in a real scenario by demonstrating a gunshot location system in the field. It is clear that this could be a fielded technology when SRW radio's OE is available to government organizations and JTRS radios are deployed. In that stage, the team can integrate the gunshot detection algorithm into the SRW radio and the gunshot location algorithm in the WNW radios.

Acknowledgements: We would like express our appreciation to Dr. Robert Maher from Montana State University to provide quality gunshot recording database for this project.

REFERENCES

- [1] Duckworth, G.L., et. al. "Fixed and wearable acoustic counter sniper systems for law enforcement." BBN Technologies White Papers. http://www.bbn.com/resources/pdf/sniper_spic98.pdf. 27 March 2005.
- [2] Showen, Robert "An Operational Gunshot Location System." Proc SPIE, Vol. 2935, 130-139 (1997)
- [3] "Waveform Design Specification for WNW", Boeing corporation document for JTRS program.
- [4] David Crane, "Anti-Sniper/Sniper Detection/Gunfire Detection Systems at a Glance", www.defensereview.com
- [5] Geoffrey Peeters, "A large Set of Audio Features for Sound Description (similarity and classification) in the CUIDADO project", April 23, 2004
- [6] Ken S. Fisher. "Systems and Methods Related to Identifying and/or Locating Weapon Fire Incidents", USPTO Application 20080219100. US, Sept. 11, 2008.
- [7] R. C. Maher, "Modeling and Signal Processing of Acoustic Gunshot Recordings," in *Proc. IEEE Signal Processing Society 12th DSP Workshop*, September 2006, p.p. 257-261.
- [8] J. C. Lahr and F. G. Fischer "Location of Acoustic Sources Using Seismological Techniques and Software" Open-File Report 93-221. U.S. Geological Survey, Menlo Park, California, 1993.
- [9] Kevin C. Baxter, Ken Fisher "Gunshot detection sensor with display", USPTO Application #: 20080192574, Aug. 31, 2007.
- [10] Fred Holmes, Ken Fisher, and Kevin Baxter. "System and method for precision acoustic event detection", USPTO patent 7420878, Sept. 2, 2008