Audio Engineering Society

Convention Paper 10540

Presented at the 151st Convention
2021 October, Online

Gunshot Detection Systems: Methods, Challenges, and Can they be Trusted?

John H.L. Hansen[1] and Hynek Bořil[1,2]

[1]*Center for Robust Speech Systems (CRSS), Erik Jonsson School of Engineering and Computer Science, University of Texas at Dallas, USA*

[2]*Electrical and Computer Engineering Department, University of Wisconsin–Platteville, USA*

Correspondence should be addressed to John H. L. Hansen (`John.Hansen@utdallas.edu`)

ABSTRACT

Many communities which are experiencing increased gun violence are turning to acoustic gunshot detection systems (GSDS) with the hope that their deployment would provide increased 24/7 monitoring and the potential for more rapid response by law enforcement to the scene. In addition to real-time monitoring, data collected by gunshot detection systems have been used alongside witness testimonies in criminal prosecutions. Because of their potential benefit, it would be appropriate to ask– how effective are GSDS in both lab/controlled settings vs. deployed real-world city scenarios? How reliable are outputs produced by GSDS? What is system performance trade-off in gunshot detection vs. source localization of the gunshot? Should they be used only for early alerts or can they be relied upon in courtroom settings? What negative consequences are there for directing law enforcement to locations when a false positive event occurs? Are resources spent on GSDS operational costs well utilized or could these resources be better invested to improve community safety? This study does not attempt to address many of these questions including social or economic questions of GSDS, but provides a reflective survey of hardware and algorithmic operations of the technology to better understand its potential as well as limitations. Specifically, challenges are discussed regarding environmental and other mismatch conditions, as well as emphasis on validation procedures used and their expected reliability. Many concepts discussed in this paper are general and will be likely utilized in or have impact on any gunshot detection technology. For this study, we refer to the ShotSpotter system to provide specific examples of system infrastructure and validation procedures.

1 Introduction

Recent years have witnessed an ever increasing presence of surveillance technologies in various aspects of our daily lives, ranging from monitoring of email communication and posts on social media, to location tracking via smartphones, and city deployed CCTV cameras. One surveillance technology that may be less known to the general public and is gaining a prominent position are gunshot detection systems (GSDS). As an example, the ShotSpotter GSDS is currently deployed in about 117 cities in the United States and worldwide [1]. The goal of GSDS is primarily to detect, and second potentially locate, gunshot activity and alert local law enforcement. There are numerous instances where GSDS technology has helped solve violent crimes and, thanks to the quick turnaround between gunshot detecting, raising a system alert, and dispatching a patrol to the location, has contributed to saving victims' lives [2]. However at the same time, city-operations/legal/scientific communities are raising questions regarding GSDS associated annual costs, efficiency, reliability, and social impact factors. Arguably, some of these questions are fueled by limited publicly available, independent conducted validation studies of the technology. While GSDS vendors often emphasize very high system accuracies, it is often unclear how those values were established and how they will fare in real-world deployments. For forensic audio analysis, the domain of automatic speaker recognition has an extensive community of researchers, industry developers, private/government sector users, and very structured publicly available speech corpora and established testing paradigms with well recognized evaluation metrics. Far less, if any, formal testing paradigm(s) exists for GSDS evaluation. A study by the MacArthur Justice Center [3] (Northwestern Univ. Pritzker School of Law's Bluhm Legal Clinic), analyzed ShotSpotter-initiated police deployments from July 1, 2019 to April 14, 2021. The study aimed to determine if ShotSpotter's claimed accuracies would apply in Chicago where the system is used by 12 police districts, and understand GSDS impact on Chicago's marginalized communities. Analysis of records kept by the city's Office of Emergency Management and Communications revealed that over 46,000 system dispatches were initiated in 21.5 months in the Chicago area, of which 10.28 % resulted in a filed incident report of likely involving a gun, and 86 % lead to no report of any crime at all. While dispatching patrols more often than needed could be viewed as precautionary, this comes with its own problems. In an interview in [4], the spokesperson for the MacArthur Justice Center pointed out: "*It sends police racing into communities searching, often in vain, for gunfire. Any resident in the area will be a target of police suspicion or worse. These volatile deployments can go wrong in an instant.*" The numbers found by the study seem to be in sharp contrast to those reported in ShotSpotter advertisement materials. Indeed, this is not the first time system performance was questioned. In 2017, ShotSpotter's employee stated in court testimony [5]: "*Our guarantee was put together by our sales and marketing department, not our engineers... We need to give them [customers] a number. We have to tell them something. . . . It's not perfect. The dot on the map is simply a starting point.*"

The goal of this study is to bring to light possible misconceptions surrounding GSDS technology by reviewing fundamentals of gunshot acoustics and surveying the history and current state in the gunshot detection field. Due to the market coverage, in discussion of typical GSDS architecture and evaluation procedures, we focus on ShotSpotter system. In a 'case study' on evaluations of GSDS systems, we analyze Mazerolle et al. [6] which has been cited by ShotSpotter advertisements at least until 2017. We provide a critical analysis of that study, including discussion of what we believe to be conceptual and 'numerical' issues. Subsequently, we outline several limiting factors and issues that are expected to affect any GSDS technology. Some limitations are inspired by Litch and Orrison [7] which evaluated the SECURES GSDS. Finally, we compare the maturity and state of gunshot forensics versus acoustic speaker recognition forensics. With the latter being arguably more mature in terms of completed scientific research, engineering solutions, and evaluation campaigns, we believe that such a comparison may offer inspiring motivation for the GSDS community in terms of next steps in developing a reliable and repeatable data sets for evaluation and calibration, as well as developing best practices for the field of gunshot forensics.

2 Fundamentals of Gunshot Acoustics and Forensic Acoustics for Gunshots

The physics of how a sound is produced will impact its time-frequency signature response and other defining traits. For gunshot acoustic analysis, firearm design, including length, diameter and construction of the barrel, as well as the type of ammunition all impact the acoustic signature of the firearm discharge [8].

From an acoustics perspective, gunshots have an acoustic structure consisting of the ballistic shockwave and resulting muzzle blast which are dependent on gun type and ammunition used. In general, each gun will have its own signature in terms of acoustics based on barrel size/design, type of weapon, and ammunition employed. Three general broad classes of guns are: (i) Handgun, (ii) Shotgun, and (iii) Rifle, with resulting barrel characteristics summarized as follows [9]. *Handgun*: short barrel with rifling and thick walls to withstand high pressures; Like the rifle, rifling in the handgun puts a spiral spin on a bullet when fired, increasing

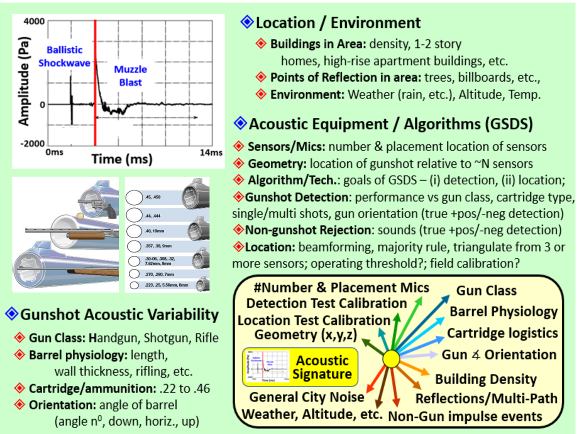

Fig. 1: Sources of acoustic variability & issues that impact system performance for Gunshot Detection Systems (GSDS): (i) Acoustic Variability, (ii) Location / Environment, and (iii) System deployment – acoustic sensor equipment, algorithms, field calibration/testing.

Table 1: Sound level averages in decibel for various handguns, shotguns, and rifles (from a study by Krammer, documented by E.A.R. [13])

CENTERFIRE PISTOL DATA	
.25 ACP	155.0 dB
.32 LONG	152.4 dB
.32 ACP	153.5 dB
.380	157.7 dB
9 mm	159.8 dB
.38 S & W	153.5 dB
.38 Spl	156.3 dB
.357 Magnum	164.3 dB
.41 Magnum	163. 2 dB
.44 Spl	155.9 dB
.45 ACP	157.0 dB
.45 COLT	154.7 dB

SHOTGUN NOISE DATA		
.410 Bore	28" barrel	150 dB
	26" barrel	150.25 dB
	18 1/2 barrel	156.30 dB
20 Gauge	28" barrel	152.50 dB
	22" barrel	154.75 dB
12 Gauge	28" barrel	151.50 dB
	26" barrel	156.10 dB
	18 1/2 barrel	161.50 dB

CENTERFIRE RIFLE DATA	
.223, 55 gr. commercial load 18 1/2" barrel	155.5 dB
.243 in 22" barrel	155.9 dB
.30-30 in 20" barrel	156.0 dB
7mm Magnum in 20" barrel	157.5 dB
.308 in 24" barrel	156.2 dB
.30-06 in 24" barrel	158.5 dB
.30-06 in 18 1/2 barrel	163.2 dB

accuracy and distance; used for firing at stationary targets; barrel bore made for only one specific caliber of ammunition. *Shotgun*: a long barrel, usually with a smooth bore to reduce friction; barrel's walls thinner due to reduced pressures; typically used for shooting at moving targets in the air; barrel bore made for only one specific gauge of ammunition. *Rifle*: long barrel with rifling and thick walls to withstand high pressures; rifling puts spiral spin on a fired bullet, increases accuracy and distance; used for firing at stationary targets; barrel bore made for only one specific caliber of ammunition. Fig. 1(center left) highlights differences in caliber and barrel size after [9]. In addition firearm type, the specific cartridge/ammunition used will also impact the acoustic signature for a gunshot.

Speaker Recognition vs. GSDS: In a manner similar to the rigorous, structured, and systematic work done for voice-based forensic speaker verification due to data mismatch [10], it is possible to establish a similar multi-space set of mismatch issues that impact the acoustic signature and performance of GSDS. It is important to note that while many have explored acoustic analysis of gunshots [8, 11, 12], there has not been an attempt to create and summarize those individual and combination of factors that impact both variability of acoustic gunshot signatures as well as deployment, calibration, and testing mismatch which should occur for GSDS algorithms/systems deployed in the field. Figure 1 is suggested from this study to span the range of issues which should be taken into account in understanding gunshot acoustic signatures and GSDS performance: (i) *Gunshot Acoustic Variability*–gun class, barrel physiology, cartridge/ammunition; (ii) *Location and Environment*–buildings in the area, points of reflection in the area, environment (weather–rain, wind temperature, altitude); (iii) *Acoustic Equipment*–type and placement of sensors, geometry (location and orientation of the gunshot relative to the sensors). Please note that the study by Aguilar [8] does consider a number of these individually but does not suggest a comprehensive space as illustrated in Figure 1.

From an acoustics perspective, there has been extensive work on sound levels from various guns. This has been motivated more for hearing protection for both gun and hunting enthusiasts as well as the military in order to protect an individual against hearing loss. For example, online hearing data [13] shows the different sound levels measured by a sound-level meter in terms of decibels for three classes of guns: handguns, shotguns, and rifles (see Table 1). In this gunfire sound level reference chart, all measured levels vary from 152–163 dB, which far exceed the threshold of pain (120 dB for threshold of pain; conversational speech 60 dB; ([14]). It is clear that the manufacturer of the gun (specific across handguns, etc.), caliber, or length of barrel for shotguns or rifles, will all impact the acoustic sound level response in dB.

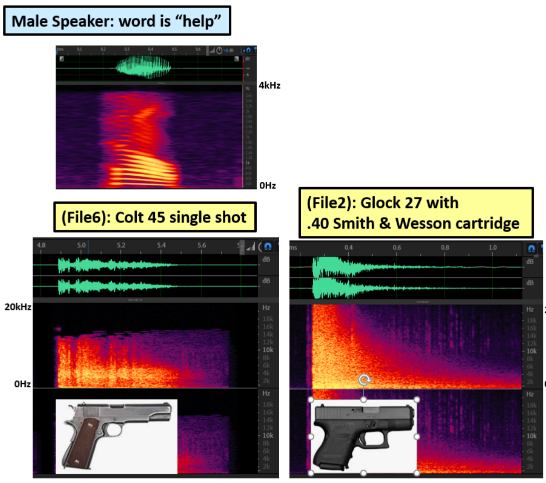

Fig. 2: Time-frequency spectrogram comparison of (i) male speaker producing the word "help" (0–4 kHz); (ii) Colt 45 single shot (0–20 kHz); (iii) Glock 27 single shot (0–20kHz). Time in seconds.

Next, it would be reasonable to briefly note time-frequency structure of the acoustic differences between gunshots. Fig. 2 compares three different time-frequency spectrograms which were analyzed from sample acoustic files: comparison of a (i) male speaker producing the word "help" (0–4 kHz); (ii) Colt 45 single shot (0–20 kHz); (iii) Glock 27 single shot with a .40 caliber Smith & Wesson cartridge (0–20 kHz). While human speech production shows periodic structure (e.g., virtual harmonic lines/bars) seen for phones such as the vowel "e" and semi-vowel "l", the frequency responses for the two handgun gunshots are dramatically different. The acoustic signature decay is reflective of the muzzle blast response. A clear initial acoustic shockwave is present, followed by muzzle blast which decays over time. The upper-left plot in Fig.1 shows that this muzzle blast is approximately 4–5 ms in duration, and travels at the speed of sound ($\sim$340m/s depending on altitude/density of air).

Several recent studies have focused on the acoustics of gunshot analysis and gunshot detection systems for law enforcement. The study by Maher and Routh [11] focused specifically on analysis of gunshot acoustics. The study collected multi-channel acoustical recordings of gunshots under controlled conditions for several firearms. Their approach was to obtain recordings with an elevated platform and spatial array of microphones to explore directional recordings of the muzzle blast. The study only considered waveform acoustic comparisons as well as peak amplitude pressure variations. One important area considered was the consistency and repeatability of gunshot sounds. A study by [15] focused on variations in recorded acoustic gunshot waveforms generated by small firearms under different controlled conditions. A study by Lo and Ferguson [16] focused on localization of small firearms using acoustic measurements of muzzle blast/ballistic shock wave arrivals. Luzi et al. [17] dealt with acoustic firearm discharge detection and classification in an enclosed environment. The study by Aguilar [8] is by far one of the most comprehensive, which covered topics such as: (i) physics of gunfire, (ii) historical development of gunshot detection systems, (iii) current GSDS technologies, and (iv) performance discussion.

Sniper Detection Technology–In early 1990's, there was a need for sniper detection to address sniper fire from Serb snipers in Sarajevo (Bosnian War). NATO intervened and the need existed to localize both gunfire and mortar blasts. A number of systems were developed to address this task, including HALO by Roke Manor Research & BAE Systems (UK), PILAR by Metravib (France), Integrated Sniper Location System by SenTech & Lockheed Martin, Sentinel by SAIC (Science Applications International Corporation), and Bullet Ears by BBN Technologies (see [8] for detailed overview of the systems).

Battlefield Transfer to High Crime Urban Areas: leveraging the DARPA's early advancements in sniper fire detection, there was an effort to transition sniper base detection to urban locations for gunshot detection. The System for Effective Control of URban Environment Security (SECURES) [7], was originally developed by Alliant Technology Systems (contractor Planning Systems, Inc). The first version of SECURES was deployed in Olympic Village in Atlanta, GA, in 1996 [8]. The ShotSpotter system was originally introduced by Trilon Technologies, with its first installation in Redwood City in 1996 [6]. In 2009, ShotSpotter, Inc. acquired the SECURES product line and became a leading company in the area. Currently, ShotSpotter is deployed in nearly 120 US and worldwide cities [1].

3 Gunshot Detection Systems

The following paragraphs detail ShotSpotter's GSDS framework as described in available literature. This description refers to system versions deployed by or prior to 2017, noting that it is likely that further technological advancements have been made since then.

ShotSpotter is a system that uses an array of acoustic sensors installed in an area of interest to detect and locate gunshots, and sometimes other acoustic events such as fireworks or helicopter activity, and alert the customer, typically a law enforcement agency or a security personnel, about those events in real-time [18]. Alerts sent by the system to the customer indicate the location of the incident in the form of a dot on a map [19], latitude and longitude coordinates, and a street address. In the case of detected gunshots, the alerts indicate the number of rounds fired. Before issuing the alert to the customer, the automatically detected events are first reviewed by a human analyst in ShotSpotter's Incident Review Center, who makes the decision whether the alert is justified or if it was just a false alarm. Human analysts may append additional information to the alert indicating the type of gunfire (e.g., a full automatic weapon was fired), presence of multiple shooters, or a shooter being on the move. The typical time interval between the incident and the alert being confirmed by the analyst and issued to the customer is advertised to be less than 45 seconds [18].

The ShotSpotter system consists of four primary elements: (i) *sensors* (includes detection algorithm operating with 1,2, or 4 microphones), (ii) *Location Server*, (requires a certain number of field sensor locations to agree on gunshot detection before location is estimated; uses the time difference of arrival–TDoA), (iii) *Incident Review Center* (human analysts perform a listener assessment to accept or reject automatic detected Alerts), and (iv) *ShotSpotter Alert Console* (provides information on detected gunshot and estimated location information).

ShotSpotter *sensors* are distributed over a geographic area indicated by the customer. Each sensor is a standalone device whose purpose is to detect gunshot pulses (or other events) and send information about the incident to the Location Server. Each sensor typically contains several microphones, a CPU system running the detection algorithm, and a GPS and cellular network modules. All sensors, as well as the Location Server, are synchronized to the GPS satellite time [20]. The task of the *Location server* is to process incident alerts from sensors, calculate geographic locations of the incidents, classify the types of the incidents, store information about the incidents in a database and alert the Incident Review Center [20]. *Incident Review Center* Analysts, also called Flex Operators, assess alerts from the Location Servers and determine whether those alerts will be published to the customer or dismissed. Analysts can change the classification of the incident and add comments to the record [20]. When an alert is issued by the Incident Review Center Analyst, the customer receives the alert on their ShotSpotter *Alert Console*. The Alert Console is a software application running on the customer's (dispatcher's) PC and displays the date, time, and location of the incident. The time refers to the first shot fired in an incident. If multiple shots were fired, only the time of the first shot is given. The location is graphically represented as dot on a map representing the latitude and the longitude of where ShotSpotter located the incident. For the 2017 and earlier versions, ShotSpotter was stating that it would detect at least 80 % of all detectable outdoor gunfire that occurred within the system deployment area and that in those detected cases, the actual shooter location would be within a radius of 25 meters from the location determined by the system [20].

4 A Case Study: ShotSpotter Evaluation in Mazerolle et al. [6]

Mazerolle et al. [6] report presents a controlled field trial of the ShotSpotter System in Redwood Village, a neighborhood located in Redwood City, CA. The field trial was conducted on June 26 and 27, 1997 and the report was completed in November, 1999. At the time of the study, the company operating the ShotSpotter System was referred to as Trilon Technology and the system did not utilize the Incident Review Center, which means that the Location Server was sending automatic alerts directly to the client's Alert Console without a human-based (perceptual) review.

Redwood City Police Dept. gave permission to the Evaluation Team to conduct field tests during two time periods: 10:00 am to 3:00 pm and 7:00 pm to 10:00 pm. These times were decided by the Police Dept. in conjunction with Trilon Technology to avoid heavy traffic hours. The Evaluation Team indicates that "*avoiding heavy traffic hours decreased the possibility of false positive alerts during our field trial as reduced levels of background noises were somewhat artificially restricted (i.e., car backfires and car horns) through this* *process. We acknowledge that, in real life situations, such background noises cannot be ignored.*"

Redwood Village, a neighborhood of approximately one square mile, was chosen for the field test due to its high incidence of celebratory and random gunfire. The test area was covered by 8 sensors mounted on rooftops of residences and other buildings. The sensors were disguised to visually resemble heating vents and bird houses. Based on police calls statistics, the evaluation team randomly selected 32 locations in the test area—27 face block locations and 5 intersections. During the test, blank rounds were discharged from these locations. Three types of weapons were used in the trial-–a .38 Caliber Pistol, a 12 Gauge Shotgun, and an MP5 Assault rifle. Depending on the location, one to four blank rounds were discharged. The trial personnel comprised the Cincinnati Evaluation Team, the Redwood City Police Dept., and Trilon Technology technicians.

A total of 31 test events were employed in the evaluations. The Evaluation Team used a random assignment to determine the number of shots that would be fired in each location-–from a single shot to bursts of two to four shots. The three weapon types were randomly assigned to locations. During the trial, an Evaluation Team member in the field was in contact with an Evaluation Team member in the dispatch center to verify that locations, times, weapon types, and number of rounds fired were all correct.

During the field trial, the Evaluation Team decided to replace the MP5 Assault rifle with either of the other two types weapons on several occasions. The authors state that the MP5 Assault rifle was hardest to detect for the system and this replacement helped ShotSpotter to achieve a higher positive rate: "*Once again, this alteration in the methodology greatly assisted the ability of Shotspotter to achieve a higher true positive rate than what would have been the case if the original design was followed. Nonetheless, we believe that the change in method was warranted since we had so few shots (N = 32) to fire: by repeatedly failing to identify shots from the MP5 assault rifle would not have illuminated additional insights as to the operational accuracy of Shotspotter. We point out, therefore, that reports of the Shotspotter's accuracy as to a system's accuracy needs to take into account the type of weapons that were fired.*" In total, 8 locations used the MP5 assault rifle, 13 locations used the .38 caliber pistol, and 10 locations the 12-gauge shotgun.

The authors summarize the *gunfire detection* performance: "*Overall, the ShotSpotter technology annunciated nearly 80 percent of the test shots (N = 24). Specifically, the technology annunciated shotgun rounds at the highest rate (90 percent) followed by pistol rounds (77 percent) and the MP5 assault rifle (63 percent).*" The *automatic location accuracy* is reported as follows: "*Overall, the system was able to triangulate random gunfire events 84 percent of the time within an average margin of error of 41 feet (see Table 2). In terms of automatic identification, ShotSpotter was able to isolate the location of random gunfire 45 percent of the time with an average margin of error of 26.5 feet.*" The *manual location accuracy* after the intervention of the Trilon Technician: "*With assistance from a Trilon Technician, ShotSpotter was able to locate an additional 39 percent of the gunfire events within 59 feet.*"

4.1 Issues of Mazerolle et al. [6] Study

Alteration of Stimuli During Trial: While the trial was in progress, the Evaluation Team deliberately altered the stimuli (the type of weapon used). The authors openly admit that this alteration was made in observation of one type of weapon, MP5 assault rifle, producing lower detection success rates than the other two weapon types-–a shotgun and a pistol.

Automatic vs. Manual Gunshot Location: The Evaluation Team established three categories for the gunshot location process—*automatic*, *manual*, and *missed*. The missed category refers to instances where the gunfire event was not detected by the system and hence, the system will not be attempting to locate it. The automatic category refers to the system's multi-lateration algorithm calculating the gunfire location without any intervention. Finally, the manual category refers to the case where the Trilon Technician would make adjustments to the system parameters after the automatic location process failed. Mazerolle et al. [6] does not define what are the criteria for the automatic location to be considered failed, neither what types of adjustments the Technician would be making. However, the manual category lacks any meaning in the context of practical applications of the system. In real-world applications, the actual location of the gunfire is unknown and if the system determines the location incorrectly, there is no way to assess it. Unfortunately, the authors of the study assess both automatic and manual location performances and eventually present them cumulative

Table 2: Adopted from Mazerolle – "Table 1. Event by Event Description of Shot Spotter Filed Trial"

Table 1. Event by Event Description of Shot Spotter Field Trial

	Date of Shot	Time of Shot (Military Time)	Shot Location	Type of Location	# of Shots Fired	Weapon Type	System Parameter	Annunciation Yes No	Triangulation Auto Manual Missed	Average Margin of Error (in feet)
*Event 1	06/26/97	1854:57;1855:02,08	1061 Douglas St.	Face block	3	Rifle	21, 14, 8	No	Missed	-
Event 2	06/26/97	1925:35;38	711 3rd Av.	Face block	2	Shotgun	18, 14, 8	Yes	Manual	25
Event 3	06/26/97	1929:16	2424 Spring St.	Face block	1	Rifle	15, 12, 4	No	Missed	-
Event 5	06/26/97	1942:32	644 Stanford Av.	Face block	1	Shotgun	15, 12, 4	Yes	Manual	13
Event 4	06/26/97	1949:32;34	2820 Crocker Av.	Face block	2	Pistol	15, 12, 4	Yes	Auto	45
Event 6	06/26/97	2002:53:54	Warrington/Halsey	Intersection	2	Shotgun	15, 12, 4	Yes	Auto	13
Event 7	06/26/97	2010:10;11;13	888 2nd Av.	Face block	3	Pistol	15, 12, 4	Yes	Manual	154
Event 8	06/26/97	2015:14;15	861 Warrington Av.	Face block	2	Rifle	15, 12, 4	Yes	Manual	162
Event 9	06/26/97	2020:23;25	Charter at Cul de Sac	Intersection	2	Rifle	15, 12, 4	No	Missed	-
Event 10	06/26/97	2027:37;39	2524 Spring St.	Face block	2	Pistol	15, 12, 4	Yes	Manual	20
Event 11	06/26/97	2038:52;39:03	475 Broadway	Face block	2	Rifle	15, 12, 4	No	Missed	-
Event 12	06/26/97	2049:11;12;15	2742 Fair Oaks Av.	Face block	3	Shotgun	15, 12, 4	Yes	Auto	16
Event 13	06/26/97	2055:35;38;45	McArthur/Halsey	Intersection	3	Rifle	15, 12, 4	Yes	Manual	27
Event 14	06/26/97	2106:32;33	Pacific/Middlefield	Intersection	2	Pistol	15, 12, 4	Yes	Auto	27
Event 15	06/26/97	2114:29;30;32;34	473 4th Av.	Face block	4	Rifle	15, 12, 4	Yes	Manual	15
Event 16	06/26/97	2129:11:13	765 Douglas Av.	Face block	2	Shotgun	15, 12, 4	Yes	Auto	22
Event 17	06/27/97	1142:39;40	622 3rd Av.	Face block	2	Pistol	15, 12, 4	Yes	Manual	10
Event 18	06/27/97 06/27/97	1151:56;57 1159:55;56	2205 Middlefield Av. 2205 Middlefield Av.	Face block Face block	2 2	Shotgun Shotgun	15, 10, 4	No	Manual	- 200
Event 19	06/27/97	1211:04;05;05	3117 Hoover St.	Face block	3	Pistol	15, 10, 4	No	Missed	-

as a single metric: "*Overall, the system was able to triangulate random gunfire events 84 percent of the time within an average margin of error of 41 feet (see Table 2).*" Only this cumulative performance is presented in the Executive Summary of the report, without any disclosure of the fact that the real performance of the automatic location algorithm is only **45 %** and that the other **39 %** are coming from manual interventions by the Trion Technician after the fact.

Successful Location of Undetected Gunshots: The report counts one gunshot event that was not detected by the system as successful location (see Table 2, next to the last row–' Shotgun'; 'Annunciation No'; 'Triangulation Manual' instead of 'Missed'). This leads to a curious conclusion by the authors that, while the system was able to detect 9 out of 10 shotgun gunshots (i.e., a 90 % detection accuracy), "*Shotgun events had the highest rate of triangulation at 100 percent (N = 10 of 10 events) with a median margin of error of 23.5 feet.*" This is then propagated to the cumulative results reporting: "Findings from the Shotspotter Field Evaluation indicated that overall, the gunshot location technology was able to **annunciate (detect) nearly 80 percent** of the test shots." followed by, "The gunshot location technology was able to **triangulate (locate) 84 percent** of the test shots (N = 26 of 31 shooting events) within a median margin of error of 25 feet." The authors do not elaborate on how the system could be successfully locating events that it did not detect.

In summary, Mazerolle et al. [6] (i) compromises its experimental protocol by altering the stimuli during the trial to artificially boost the shot detection accuracy; (ii) allows manual interventions by the Trion technician during the location process, results of which are then blended with the automatic location results and reported as a single number in the Executive Summary of the report; and (iii) overstates the location accuracy by counting in as a success a manual location of a gunfire incident that was not detected by the system at all. It is noted that the results from the Executive Summary were relied upon in advertisements of the ShotSpotter performance as recently as in 2016 [20].

5 Limiting Factors and Issues in GSDS Systems

The previous section discussed a ShotSpotter evaluation study. This section first highlights major observations from a similar study conducted on the SECURES system by Litch and Orrison [7]. This is followed by our assessment of limiting factors for GSDS systems.

5.1 SECURES Study by Litch and Orrison [7]

Litch and Orrison [7] is a 2011 study of the SECURES gunshot detection system deployed in Newport News and Hampton. Similar to Mazerolle et al. [6], the executors of the study were not affiliated with the owner/operator of the system. Besides field trials, the authors also analyze law enforcement agencies' data to empirically validate claims and perceptions of the actual usefulness of the system for police officers.

The SECURES system detects and locates gunshots in the coverage area. The system sensors are typically mounted on utility poles and building walls. Sensors triggered by an acoustic event send a signal to a receiver that communicates with a server. Data from the triggered sensors are used to locate the incident in a three-dimensional space. The system produces map coordinates and also information whether the event is on the ground or aerial-based. The system alerts police dispatchers via a law enforcement dispatching system (CAD). While the space limitations prevent a detailed discussion of this study, some of the final observations in Litch and Orrison [7] nicely outline some of the major issues seen in GSDS designs: "*Results attained suggest the complexity of deploying new technologies in urban settings...*"; "*There was a noted tradeoff between true positive errors in the Hampton live fire experiment. If the sensor bar were set higher, false positives would be reduced but so would true positives*"; "*In the field assessment, the problem of unnecessary responses (false positive error) was an issue. In Hampton, 18 % of SECURES-related dispatches had associated 911 calls. This means that 82 % of dispatches would not have occurred but for the gunshot detection system.*"; "*The conclusions of this report suggest that many of the most optimistic notions of how the system would perform and how the system might impact police operations lacked empirical support by the assessment efforts of the investigators.*"

5.2 Limiting Factors for GSDS Systems

Multi-lateration and other techniques relying on TDoA assume that the received signal is arriving directly from the source to be localized (direct sound). *Reflected sound waveforms (echo)* cannot be used towards successful location of sound with these methods. GSDS sensors are prone to confusing echo impulses with direct sound impulses, especially in urban areas where many tall buildings exist.

There are other effects that will impact trajectory/route of the sound waveform arriving to the sensor, including direct sound waveform, such as *wind*, sound *diffraction* and *refraction*. Alterations of the arriving direct sound trajectories may notably reduce accuracy of the location algorithm even if the detected sound is coming from the direct sound waveform.

The composition of the direct sound and reflected sound waveforms arriving all together to the sensors will not depend only on the absolute location of the sound source versus the sensor, but also on the *direction of the gunshot*. For example, the same gunfire location may be producing a nearly pure direct sound arriving to a sensor or a nearly pure reflected sound waveform, depending if the shooter is aiming towards the sensor without any obstructions on the way or aiming towards a wall of a building that is in the opposite direction from the shooter than the sensor.

Mazerolle et al. [6] did not test on *other incidents of interest* (fireworks, helicopter sounds) and controlled *competing impulse noises* (backfires, construction equipment, mufflers); as such there is no indication of what the performance of ShotSpotter-like GSDS would be under such conditions.

In Mazerolle et al. [6], only 45 % of all test gunfires were automatically located by ShotSpotter in 31 trials. The very limited number of trials provides little statistical significance to the already low location accuracy result. Moreover, the trials were conducted outside rush hours, which reduced environmental noises that might have further impacted detection and location accuracy.

6 Summary–Comparison of Gunshot Forensics and Acoustic Speaker Recognition Forensics

In this section, we summarize several major differences between the fields of gunshot detection/location and speech recognition forensics in terms of maturity and research and evaluation frameworks. It is suggested that some of these differences may inspire researchers, engineers, practitioners in the GSDS field for future next steps in developing and solidifying the field.

While studies have considered comparison of acoustic waveform differences between gunshots based on firearm types, ammunition, etc., the field is extremely limited with only a handful of researchers considering the acoustics of gunshots. Maher and Routh [11] notes: "*...acoustical characteristics of gunshots are currently little understood in an objective sense by many law enforcement investigators and acoustical consultants, so there is the possibility of unscientific assumptions, interpretations, and testimony.*"

There have been no efforts to establish formal evaluation procedures to benchmark gunshot detection, gunshot firearm classification, and gunshot location estimation in the field. For comparison, the USA based NIST OSAC–Speaker Recognition Subcommittee [21] has membership from forensic practitioners, researchers/academics, legal/government agencies to establish best practices for forensic speaker recognition in terms of technology use, limitations, testing paradigms, legal/courtroom practices. NIST has established recognized benchmark criteria for evaluating speaker recognition systems. The NIST OSAC–Firearms & Toolmarks Subcommittee only considers analysis of toolmarks on bullets, cartridge cases, firearm function testing, serial number restoration, muzzle-to-object distance determination. There is no specific subcommittee considering forensic acoustics of gunshots within OSAC, suggesting there is no group providing neutral oversight in this field.

The range of sources of acoustic variability & issues that impact system performance for Gunshot Detection Systems (GSDS) shown in 1 have not been comprehensively studied by the community. While GSDS manufacturers understand the issues/dimensions which impact variability/mismatch, their main focus has been to ensure their deployed field solutions can remain operational as they try to claim effective performance. There are limited engineering/scientific evaluation studies vs. those in other forensic areas, including speaker recognition, that validate these commercial systems. The range of causes of acoustic mismatch for gunshots has effectively not yet been addressed in any systematic manner. Therefore, there is serious doubt as to the test/re-test validity of gunshot detection systems.

The speaker recognition community has established well recognized performance acoustic train/development/test sets, a range of algorithm advancements, and formal metrics propagated by NIST– Equal Error Rates (EER), Detection Error Tradeoff (DET) curves, and Detection Cost Functions (DCF/minDCF). In the acoustic gunshot detection domain, open data sets, formalized open competitions, and a community based set of best practices are not yet available. In essence, each company performs a range of experiments/testing, and decide what results to share in their public literature. It is suggested that for the progression of gunshot detection system advancement and acceptance, more formal evaluation and testing procedures are needed. A neutral group which can address and highlight strengths and limitations are needed to help drive this field forward, especially if these systems are to be used beyond simply alerting and directing law enforcement to a location (i.e., detection), versus transitioning these system output data to be used formally used as evidence within legal/courtroom cases. The lack of a formal test corpus where evaluations are conducted by non GSDS-industry personal, transparent field calibration, and periodic re-calibration procedures are needed.

References

[1] SST, "ShotSpotter Precision Policing: Platform Overview," 2021, `https://www.shotspotter.com/platform/` [Accessed: 2021-06-23].

[2] Gecas, A., "Gunfire game changer or big brother's hidden ears?: Fourth Amendment and admissibility quandaries relating to ShotSpotter technology," 2016, pp. 1073–1121, 2016.

[3] MacArthur-Justice-Center, "ShotSpotter Creates Thousands of Dead-End Police Deployments That Find No Evidence of Actual Gunfire," 2021, MacArthur Justice Center, Pritzker School of Law, Northwestern University, Chicago; `https://endpolicesurveillance.com` [Accessed: 2021-06-24].

[4] CST-Editorial-Board, "If ShotSpotter constantly misfires, what's Chicago getting for its $33 million?" 2021, Chicago Sun-Times (Online), May 4, 2021.

[5] Lamb, J. O., "Courtroom Testimony Reveals Accuracy of SF Gunshot Sensors a 'Marketing' Ploy," 2017, San Francisco Examiner (Online), Jul. 11, 2017.

[6] Mazerolle, L. G., Frank, J., Rogan, D., and Watkins, C., "Field Evaluation of the ShotSpotter Gunshot Location System: Final Report on the Redwood City Field Trial," Technical report, Doc. 180112, U.S. DoJ Award 96-MU-MU-0018, 2000.

[7] Litch, M. and Orrison, G., "Draft Technical Report for SECURES: Demonstration in Hampton and Newport News, Virginia," Technical report, Doc. 233342, U.S. DoJ Award 2003-IJ-CX-K029, 2011.

[8] Aguilar, J. R., "Gunshot detection systems in civilian law enforcement," *Journal of the Audio Engineering Society*, 63(4), pp. 280–291, 2015.

[9] Hunter-ed.com, "*Differences Between Rifles, Shotguns, and Handguns*," N/A, https://www.hunter-ed.com/pennsylvania/studyGuide/Differences-Between-Rifles-Shotguns-and-Handguns/20103901_88439/ [Accessed: 2021-06-23].

[10] Hansen, J. H. and Hasan, T., "Speaker Recognition by Machines and Humans: A tutorial review," *IEEE Signal Processing Magazine*, 32(6), pp. 74–99, 2015.

[11] Maher, R. and Routh, T., "Wideband audio recordings of gunshots: waveforms and repeatability," in *Audio Engineering Society (AES) Convention 141*, 2016.

[12] Lilien, R., "Development of Computational Methods for Audio Analysis of Gunshots," Technical report, Doc. 252947, U.S. DoJ Award 2016-DN-BX-0183, 2019.

[13] E.A.R., "Gunfire Noise Level Reference Chart," 2019, https://earinc.com/gunfire-noise-level-reference-chart; https://earinc.com/wp-content/uploads/2019/04/Gunfire-Noise-Level-Chart-EAR-Customized-Hearing-Flyer.pdf [Accessed: 2021-06-23].

[14] Denes, P. and Pinson, E., *The Speech Chain: The Physics and Biology of Spoken Language, Second Edition*, Waveland Press, 2015.

[15] Beck, S. D., Nakasone, H., and Marr, K. W., "Variations in Recorded Acoustic Gunshot Waveforms Generated by Small Firearms," *The Journal of the Acoustical Society of America*, 129(4), pp. 1748–1759, 2011.

[16] Lo, K. W. and Ferguson, B. G., "Localization of Small Arms Fire using Acoustic Measurements of Muzzle Blast and/or Ballistic Shock Wave Arrivals," *The Journal of the Acoustical Society of America*, 132(5), pp. 2997–3017, 2012.

[17] Luzi, L., Gonzalez, E., Bruillard, P., Prowant, M., Skorpik, J., Hughes, M., Child, S., Kist, D., and McCarthy, J. E., "Acoustic Firearm Discharge Detection and Classification in an Enclosed Environment," *The Journal of the Acoustical Society of America*, 139(5), pp. 2723–2731, 2016.

[18] SST, "ShotSpotter Frequently Asked Questions (FAQ)," 2018, https://www.shotspotter.com/system/content-uploads/SST_FAQ_January_2018.pdf [Accessed: 2021-06-23].

[19] SST, "ShotSpotter Flex Alert & Analysis Service: Best Practices, Strategies, and Recommendations," 2013, https://www.seattleprivacy.org/wp-content/uploads/2014/05/SST-Best-Practices-March-20131.pdf [Accessed: 2021-06-23].

[20] SST, "SST Expert Testimony Common Questions and Answers," 2016, https://hobbydocbox.com/72285000-Radio/Sst-expert-testimony-common-questions-and-answers.html [Accessed: 2021-06-23].

[21] OSAC, U. S., "OSAC: Speaker Recognition Subcommittee," 2021, https://www.nist.gov/osac/speaker-recognition-subcommittee [Accessed: 2021-06-23].