White Paper

Version 1.01

Balint Seeber

Bastille Tracking Number	43
DHS ICS-CERT Advisory Number	ICSA-18-100-01
Public Release	April 10, 2018

Bastille

Table of Contents

Version History

1.01	Fixed typo, formatting.
1.0	Initial version

Level of Technical Detail

Full technical details are currently being withheld as Emergency Warning Systems are considered public safety critical infrastructure, and it is unclear how many deployments are vulnerable to the exploit. The vendor has indicated it has worked on a patch that increases the level of security of the protocol. System owners are encouraged to contact the vendor.

Overview

What is SirenJack?

SirenJack is a vulnerability found in ATI Systems' emergency warning systems that can be exploited via Radio Frequencies (RF) to activate sirens and trigger false alarms.

How is the vulnerability exploited?

The radio protocol used to control the sirens is not secure (activation commands are sent 'in the clear' - no encryption is used). A bad actor can find the radio frequency assigned to a deployment, craft malicious activation messages, and transmit them from their own radio to set off the system. All that is required is a $30 handheld radio and a computer.

How many ATI systems are affected?

In news reports, ATI claims more than 5,000 systems in and around cities, military installations, universities and industrial sites (including oil and nuclear) across North America and around the globe. The SirenJack vulnerability was found in San Francisco and confirmed in two other locations, but it is not known how many other ATI systems are subject to SirenJack. We urge all ATI customers to work with ATI to understand if their system is impacted and employ a remediation immediately.

What are the potential effects of SirenJack?

The public relies on emergency warning systems to be activated only for legitimate threats, often weather or security related. False alarms cause widespread concern and increasing distrust in these systems, particularly as seen in 2017 after the Dallas siren incident that set off over 150 tornado warning sirens citywide for more than 90 minutes.

More information, overview and proof of concept videos can be found at:

https://sirenjack.com/

Basic Configuration

An emergency siren system, one type of Emergency Warning System (EWS), consists of one or more active sound-producing elements to warn populations of various natural or man-made threats. Historically, these have also been known as air raid sirens. Modern systems tend to use discrete electronics and speakers (‘horns’ and ‘drivers’) to produce very loud audible alerts. These alerts can be a variety of tones (constant frequency, ‘wailing’, ‘whooping’, etc), spoken messages, or a combination of the two.

Deployments typically consist of a number of sirens that cover a wide geographical area when used to alert the population of, for example, a city, county, industrial site, military installation or educational institution. Due to the spatial distribution of such systems, it makes sense to implement them with radio frequency (RF) controls, as opposed to a more expensive wired approach.

Each individual ‘siren’ node is usually a telegraph pole, on top of which an array of speakers is placed to produce audio in multiple directions. Power is supplied by the electricity grid, batteries and/or solar cells. A cabinet (‘enclosure’) attached to the pole houses electronics, such as an embedded computer, amplifier boards, and radio communications equipment) used to control the siren. Siren nodes can also be placed on the roofs of buildings, or exist as mobile units.

The siren system as a whole is commanded from one or more control locations. For example there might be one central control point (‘controller’), and another mobile, redundant control point. A controller can be used to activate the siren system (i.e. send out an alert tone and/or spoken message), configure the siren nodes, and perform diagnostics (e.g. check the operating status of each siren node, especially after a test).

A siren system controller will utilise a radio link to send and receive messages to and from the siren system. This can be considered as a one-to-many network, where a controller will either broadcast messages to the entire network (e.g. during siren activation, for simultaneous triggering of all nodes), or communicate directly with a specific node (e.g. for diagnostics).

Radio Links

Radio links provide the medium through which messages are conveyed by using RF modulation. Two of the simplest forms of modulation are basic FM and AM used by broadcast radio stations. These are analog modulation schemes, which are intended to convey the ‘analog’ information of human speech or music. Digital modulation schemes are used to convey binary data (1s and 0s) in a more efficient manner. There are a myriad of digital modulation schemes in use today, each with their own features. Examples are the many schemes used in cell phones (GSM, CDMA, LTE), aviation (ACARS, ADS-B), digital public safety networks used by first responders (APCO Project 25, or P25), satellite data links, and control and telemetry in industrial control systems.

In smaller deployments, direct RF communications (‘simplex mode’) on one frequency can be sufficient, where the radio attached to a controller has sufficient transmit (TX) power to be heard by each siren, and each siren’s radio has enough transmit power to be heard by the controller.

In larger deployments, such as a city, a radio repeater can be used effectively to boost coverage (the service area). Repeaters are installed at geographically advantageous locations, such as on a tall building, or hill, along with a tall radio mast. When relaying through a repeater, there are typically two frequencies in use (the input and output frequencies). A radio wishing to have its transmission heard over a wider area will transmit on the input frequency. The repeater will listen for any transmissions on the input frequency, and will then (in real time) rebroadcast an incoming transmission on the output frequency at a much higher power than the incoming transmission used. In this way, a smaller-powered node (or mobile unit), with the assistance of a repeater, can broadcast over a much larger area.

Many deployed radio repeaters are of the analog FM flavour: they are designed specifically to listen for incoming narrowband FM transmissions, and rebroadcast the same analog signal. Repeaters exist in the digital flavour too, in which they are designed to receive a specific mode of digital signal. For example, in a P25 public safety network, a P25-compliant repeater will receive P25 digital transmissions, demodulate the incoming data, perform ‘bit regeneration’ (e.g. error correction), and then transmit the re-modulated, re-generated data.

The Dallas Siren Hack: DTMF Replay Attack

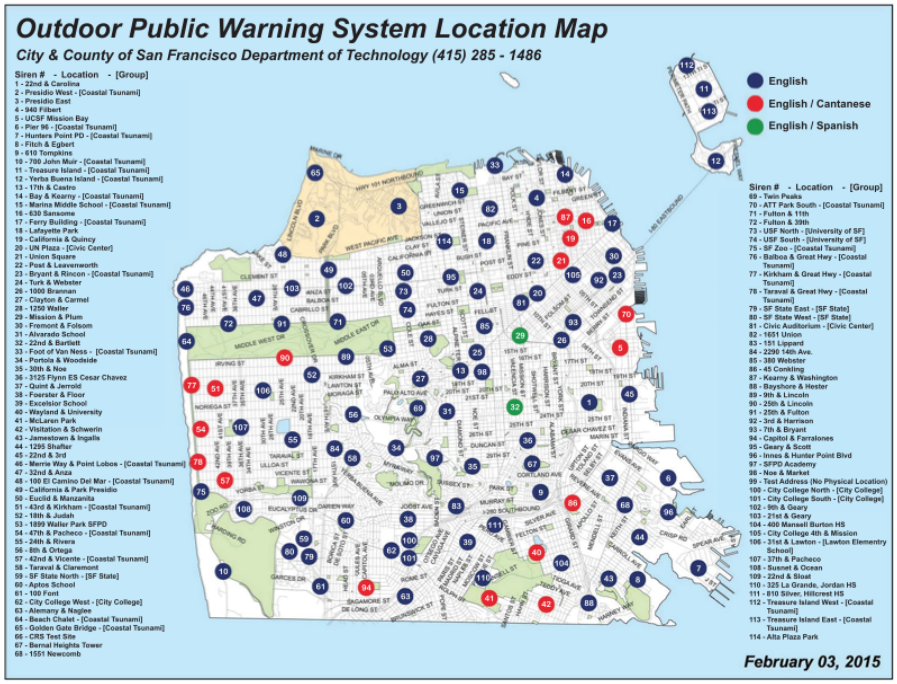
The SirenJack vulnerability is distinct from the replay attack that struck the Federal Signal-manufactured Dallas tornado warning system on April 7th, 2017. The older Dallas system used Dual Tone Multi Frequency (DTMF) tones to activate the system over an analog radio link. It is trivial to record the audio of those tones (e.g. on a laptop or tape recorder), and then replay them on the same frequency while transmitting. The activation ‘code’ usually is fixed, and therefore can be accepted multiple times.

San Francisco's Outdoor Public Warning System

The City of San Francisco deployed a siren network throughout the city and refers to it as the Outdoor Public Warning System (OPWS). The system was manufactured by Acoustic Technology, Inc ('ATI Systems', or 'ATI'). The OPWS is maintained by the San Francisco Department of Technology, and is one tool used by the San Francisco Department of Emergency Management to alert residents.

The San Francisco deployment currently consists of 144 siren poles. ATI models HPSS16 and HPSS32 are in public view around the city. A repeater is used to relay radio transmissions between one or more controllers (ATI model CCU) and the rest of the network.

Each Tuesday at midday, the siren network is tested ("the Tuesday Noon test"). A loud 'wail' tone is heard through the city, followed by a male voice announcing that it is "only a test".



A publicly-available map of siren locations from 2015 [OPWSMap].

An HPSS16 siren node.

Research Process

The findings that led to the disclosure were based on passive observation of activity on the public safety frequencies used by the San Francisco OPWS. There was no direct interaction with the system: no active transmissions were made, nor was there access to equipment, software or firmware.

There were two phases to this research:

1) Observing the radio spectrum for activity (signal transmissions) that correlated with the weekly Tuesday noon test, and then identifying the system's frequency

2) Analysing transmissions on the system's frequency to determine the signal modulation scheme, thereby revealing whether the system was secure, or vulnerable to attack.

Phase One: Signal Search

As nothing was known about the implementation of the radio link in the San Francisco OPWS, it was necessary to search the entire radio spectrum for candidate signals that might be used to activate the siren system for the Tuesday noon test. The search was based on the assumption that there should be radio transmissions correlated in time with the siren system activation. Those transmissions would contain the activations commands for the network.

There are many hundreds of active frequencies and potentially thousands of transmissions in an observation window. However it would only be possible to perform any observations in a finite time and frequency window (time as there is only one test per week, frequency because it is not tractable to capture the entire radio spectrum for a sufficient period of time with a moderate amount of consumer radio equipment). In order to expedite the search, a number of Ettus Research Universal Software Radio Peripheral (USRP) Software Defined Radio (SDR) receivers were used to capture select parts of the radio spectrum, and then those captures were reviewed to check for any correlated activity.

To help prioritise the frequency ranges to look at, additional intelligence was collected by noting radio-related features of siren nodes, such as the type, dimensions and direction of the pole-mounted antennas.

After identifying some unique and more readily searchable clues, such as the model of Yagi antenna used at one location, attention was placed on a particular band. After one week's test, it was obvious there was a good correlation of transmissions with the test. This observation was repeated over the following weeks, where the same activity was detected on the same frequency. This frequency became the first candidate to demodulate and analyse further.

It also became evident that a repeater was in use, because all transmissions on the identified frequency appeared to have the same signal power, yet listening to the transmissions themselves revealed differences that would be due to degradation in signal quality, likely because of varying distances between a node and the repeater. It can be expected that the signal from the furthest nodes might experience some interference.

The use of a repeater was confirmed when the same transmissions were found on second frequency (with a fixed offset from the first). Those transmissions varied in signal strength, which is consistent with it being the repeater input frequency, and with nodes being at various distances from the observing receiver. Transmissions that appeared to be from the master of the network (the controller) always appeared to have the same signal strength, which was also expected.

Phase Two: Signal Analysis

Once a candidate signal was identified, it was necessary to characterize the modulation scheme using standard signal characterization processes. This involved turning the raw radio signal into binary data (0s and 1s).

The modulation of this signal was quickly identified as using a common scheme. Blind signal analysis techniques were used to discover the signal modulation parameters, such as symbol rate, symbol-to-bit mapping and framing.

Once a compatible demodulator was implemented, it was necessary to build a corpus of the raw data contained in the transmissions (raw packets). This specific frequency was monitored by both recording the raw RF and decoding transmissions later, as well as decoding transmissions in real time.

The key to analysis of a digital protocol, or any other data source with an unknown structure, is being able to recognise patterns in the data. After framing the data correctly, initial analysis showed that the raw data had low entropy (i.e. did not appear random), which is often a sign that there is structure that can be interpreted. After collecting several weeks worth of individual transmissions' raw data, it became apparent that:

1) no encryption was being used

2) certain packets were activation commands, based on their timing relative to when the sirens were heard

3) there were elements in the packets that remained static

4) of the elements that were changing, patterns were recognisable, and therefore could be extrapolated

Additional traffic was observed following weekly tests, which was found to be the individual siren status check process. After the weekly siren test, the controller attempts to contact each siren in the system and request its status, presumably to check that all system components are in working order and that the test was successful. It was also found that multiple nodes would not respond to status requests, either because they had experienced a malfunction, or had been intentionally removed from the network but not from the controller's configuration. Assuming the former, an observer can build up a picture of the overall state of the network.

Protocol Security

There are some important definitions, and distinctions, that apply to data security generally, and, in this context, radio protocol security.

A communications system is generally considered secure when it implements both strong encryption **and** authentication mechanisms.

The following descriptions should be considered as introductions to these concepts. There are many nuances that can have a significant impact on the security of a protocol. Readers are encouraged to deepen their understanding from more detailed sources.

Shared Medium

The RF spectrum is a shared medium (akin to a shared communications bus). Any transmission on a frequency can be received by anyone receiving on that same frequency (or in the case of using an SDR, having the transmitter's frequency fall in the usable bandwidth of the tuned receiver). Therefore transmissions will propagate through an inherently insecure medium. This necessitates the implementation of security features in radio protocols, such as encryption and authentication, where a radio network should only be accessible to authorised parties.

The primary consideration of an external observer is to be in a position to receive a strong signal. This is influenced by one's distance from the transmitter, gain of the receiver antenna (e.g. a directional antenna will perform better than an omnidirectional one, but assumes it is known where to point it), and the performance of the receiver radio (this can be improved by using RF filters to suppress interference caused by strong out-of-band signals).

As an external observer can exploit the shared medium by observing regular communications of the system, every time a system transmits, it can potentially leak some information that can be used to better understand its protocol and the information being transmitted through it.

Error Detection

When data is transmitted through an unreliable medium, such as the radio spectrum, portions of the data may be corrupted due to interference. For a receiver to verify the integrity of the message (check that it has not been modified or corrupted), additional data can be added to the

original message that represents a compact summary of the message. This is often known as a checksum. The method to compute the checksum is built into the communications protocol, so both sender and receiver are aware of the algorithm. Since the sender transmits the original data and a checksum, the recipient can receive the original data and transmitted checksum, and then compute its own checksum of the received data. If the receiver's computed checksum matches the received checksum, then the message is likely free of errors. If they do not match, either the original data or the transmitted checksum was corrupted. Either way it is not possible to verify the integrity of the message, and so it is commonly discarded. A system will often be tolerant of such packet loss and employ a retransmission scheme. The simplest is using acknowledgement messages (ACKs) on the reverse path, where a receiver will transmit an ACK to the original sender of the message. If the sender does not receive an ACK within a timeout, it assumes that the recipient did not receive the message, and transmits it again.

The simplest form of error detection is the parity bit, used in serial communications. Another involves applying binary mathematical operations to the raw data to calculate a basic checksum. More advanced error detection algorithms include Longitudinal Redundancy Check (LRC) and Circular Redundancy Check (CRC) algorithms. CRCs are one of the most common forms of error detection algorithm. It is easy to confuse CRCs with other error detection algorithms and use the term generically, but CRCs are a particular class of algorithm with well-defined properties.

Although error detection algorithms allow a receiver to verify the integrity of a message, they should not be considered as adding security to a protocol. Most checksum algorithms are well understood, documented and optimised for certain applications. There are mathematical techniques to recover the parameters of various checksum algorithms when those parameters are unknown in advance, and a sufficiently large data set has been collected. There also exist programs, such as 'reveng', that can recover these parameters by brute force from such a data set. Moreover the parameters of a checksum, by definition, are fixed from one message to the next.

Error Correction

Error correction (Forward Error Correction, FEC) algorithms build on error detection by adding a minimal amount of redundant data to the original data, allowing sophisticated algorithms on the receiver to recover portions of the received data that were found to be corrupt.

The simplest FEC is a repetition code, where the same data is sent multiple times in one message - it is also the most inefficient. A receiver can then compare multiple copies and recover corrupt data by using 'majority rule' with the repetitions.

All modern communications systems employ some form of error detection or correction. The goal of more advanced FEC codes is to minimise the amount of redundant data added, while maintaining a minimal error rate and being bound by any computational complexity limits in the receiver. Some common types are: convolutional codes (used satellite communications), turbo codes (used in LTE), BCH codes (used in pagers) and Reed-Solomon codes (used in P25).

Unencrypted Data

When data is unencrypted, the original data ('plain text' in cryptography) is being sent 'in the clear'. There is no attempt to modify the data before transmission to make it more difficult to read by an external party. It is important to note that even though raw data may at first glance look random or unintelligible, that does not make it encrypted. Even when one, or even a handful of packets, look random, patterns can emerge when analysing large data sets. Once patterns are established, the underlying format of these packets is revealed.

Some communications need to be unencrypted by design so that they can be easily interpreted by anyone. This is particularly true in the amateur radio bands.

Obfuscation

Obfuscation is a general technique whereby unencrypted data is changed to make it more difficult to detect or understand, and ultimately make it harder to perform the sort of analysis that is easier on plain text. Data can be obfuscated in a number of ways, including scrambling and encryption. An authorised receiver would have the required knowledge to be able to un-obfuscate the inbound data and restore the original message.

Scrambling

Scrambling is a data 'pre-formatting' process used particularly in radio communications to increase the entropy of raw data prior to being transmitted over the air, which helps a receiver synchronise to the incoming data. This is because a receiver's timing recovery algorithm will be able to establish and maintain a better lock on the incoming signal when that signal contains many state transitions. In most communications systems, it is not possible to guarantee that raw data will cause a sufficient amount of state transitions when modulated, so scrambling all raw data overcomes this issue.

In the context of this document, scrambling is not regarded as secure encryption, even though it might make the raw data transmitted over the air look random. Analysis can be attempted by searching for repetition of scrambled sequences and making assumptions about the regularity of the underlying raw data.

Scrambling and descrambling can be achieved by using additive or multiplicative scramblers, and are commonly implemented using Linear Feedback Shift Registers. They are found in some form or another in all modern communications systems.

Encryption

Encryption is the first requirement to secure a communications system. It is a technique that modifies plain text data to secure communications between multiple parties where each party shares some, or all, knowledge about how that modification took place, such as an encryption key. Ideally it becomes impossible for an external party to be able to communicate with that group without the additional knowledge. However encryption alone does not prevent someone from retransmitting (replaying) a previously encrypted message, and having that illegitimate message being considered valid by the original group.

Common implementations use either:

1) symmetric key encryption (each party has a shared key, which must be kept secure),

or

2) asymmetric key encryption (each party has part of the knowledge required to communicate securely).

Asymmetric encryption requires greater complexity in its implementation, computational resources, and management, but also has increased benefits, such as key security.

Even though asymmetric algorithms are superior, symmetric algorithms are still very common in large-scale communication systems. One, or a small number of symmetric keys might be used on all devices in a system. If any one device in the system is compromised and a key is recovered (for example through known plaintext analysis, or key dumping with physical access), then the entire system is compromised. An example of this can be found in [Glass11] where a known plaintext attack can be launched against a large-scale digital P25 network to recover the key. This attack exploits the fact that known plaintext ‘silence frames’ are transmitted at the end of each transmission (‘over’). The key can be recovered with a bruteforce key space search because the encryption algorithm uses keys of insufficient length, making such a search tractable with modern FPGAs and GPUs. Therefore it is important to use cryptographically secure ciphers with sufficiently long keys to make such attacks intractable.

For devices that have low power usage requirements, such as embedded devices, asymmetric algorithms may be too expensive for the processor and power budget. Therefore symmetric algorithms are used instead (if at all, or with shorter keys).

Authentication

Authentication is the other requirement for a secure communication system. It is used to verify the authenticity of a message, and the sender. Confirming that the data in a transmission is genuine, and is not being replayed, prevents external parties from masquerading as legitimate, authorised parties (spoofing). This is also known as being able to check the 'freshness' of a message. Note that secure authentication has stronger guarantees than simply computing and comparing the checksum of a message.

The simplest form of authentication with 'freshness' can be achieved by adding some element to the raw data that changes each time a new message is sent, such as sequence counter. Communicating parties maintain knowledge of the last-seen sequence number, and transmit with the next consecutive sequence number. Recipients can then authenticate the message by checking if the received sequence number is the next expected consecutive value. If messages are lost (are not able to be received by the recipient), one technique is to accept a small range ahead of the last-seen sequence number. This is similar to vehicle alarm rolling codes.

Authentication functions are often more complex, cryptographically-secure hash functions that produce a digest of the message. A checksum can also be considered like a hash, but hashes are typically much longer to reduce the change of there being a hash collision. A collision occurs when different input data produce the same hash value. Longer hashes make it much more difficult to analyse the behaviour of a hash function to understand how it works. In this sense, they are considered 'one-way' functions.

Authentication alone does not make for a secure communications system. In the sequence counter example above, it is trivial for an external observer to note sequence numbers in previous messages between legitimate parties, and then forge their own message with the next expected sequence number. Therefore authentication must be combined with encryption.

A more secure protocol employing sequence numbers would first encrypt the sequence number, then the raw data (so that it cannot be read by an external observer), and also append a hash of the original message (for increased security a secure keyed-hash should be used). This guarantees that:

1) the original message is not sent in the clear

2) the same raw data sent multiple times will appear different ‘over the air’ in each transmission

3) the integrity of the raw data can be verified in a more secure manner

Sophisticated implementations of authentication using sequence counters exist in many protocols, including the authentication layer of LTE and 802.15.4/ZigBee Networking.

There are some well-known general-purpose algorithms that provide secure authentication, such as the Hash-based Message Authentication Code (HMAC), which yields a secure hash based on the original message, a cryptographic hash function and a key, making it more immune to certain types of cryptographic attacks.

Security Through Obscurity

The trap that many protocol designers have fallen into is thinking that their system will be secure if they design their own proprietary protocol. The term ‘security through obscurity’ is usually applied retroactively to describe such designs after they have been revealed to be insecure.

The assumption is that, since those implementing such systems are the only ones that know the details, no one else will be able to figure them out. This is not an acceptable security approach, especially with the abundance of tools and techniques, both hardware and software, that can be used to analyse such protocols.

Although potentially counterintuitive, the most secure protocols are those that have been exposed, and survived, public scrutiny. This is because if the underlying mathematics can be shown to be secure, then its security largely depends on a solid implementation and good key management. It is possible for a system to suffer from vulnerabilities even when a secure algorithm is chosen because the software or hardware implementation of that algorithm introduced a bug and/or leaked information that could be used in a cryptographic attack.

Secure Protocols

Some examples of widespread secure wireless protocols used in different domains include: BlueTooth (consumer electronics), encrypted 802.15.4/ZigBee Networking (Low Power Wide Area Networks, LPWANs) and encrypted P25 (public safety networks). The strength of their security relies on them being implemented and configured correctly with strong ciphers, long keys and good key management.

The ATI Protocol

The proprietary digital radio protocol used by ATI to control the San Francisco OPWS was found to have no encryption. As messages were sent in the clear, the patterns of changing elements became easy to interpret. These patterns could be extrapolated to craft malicious messages that conform to the protocol's format and therefore look legitimate, such as activation commands to trigger false alarms. In a deployment where regular testing takes place, knowledge gained by passive observation of test activation commands can be used to trigger the siren system in that deployment at will. That same knowledge may be used to trigger sirens in a different, vulnerable deployment. Different deployments may have different configurations and be running different versions of the software, but still be vulnerable. A motivated bad actor could visit multiple deployments to check if there are any differences in the packet format/command structure, and then begin to generalise the attack.

The protocol does not draw on any truly secure practices to prevent analysis of the relevant fields, and thwart potential interference with the system. It is therefore vulnerable due to its reliance on security through obscurity.

Bastille began its Responsible Disclosure process with ATI on Jan 8th 2018. By the time of public disclosure on April 10th 2018, ATI had created a patch that it had provided to at least one of its customers. See the Remediation section below.

A San Francisco siren node being tested during the upgrade

Proof-of-Concept

A Proof-of-Concept (PoC) was demonstrated on an ATI siren node with a single horn at a low volume at an isolated location. A modulator and transmitter were created using GNU Radio and a USRP B200mini SDR. Knowledge of the protocol gained by passive observation of two active deployments (San Francisco, CA and Sedgwick County, KS) provided sufficient information to enable the crafting of legitimate activation commands for this node, the configuration for which was unknown. The node was standalone and was not part of a siren network during the PoC. Some additional, limited protocol fuzzing was required to make the node believe it was part of an active deployment. Neither this fuzzing, nor the knowledge gained from multiple deployments, is required if a malicious actor wished to target a real deployment assuming a previous activation has been observed at that deployment. The end result of using the knowledge gained from observing these two deployments could have also been achieved by further, limited fuzzing.

The ‘listening’ radio inside the siren cabinet was pre-programmed to operate on a licensed frequency so it would have violated FCC rules to send it a signal over the air. Therefore a cable was run from the transmitter antenna port of the SDR to the receiver antenna port on the siren’s radio (via an RF attenuator). This works exactly as it would in a broadcast situation, other than using copper as a stand-in for air. This meant that no RF energy was radiated from any antennas. An alternative approach commonly used by security researchers is to place the devices under examination in Faraday cage, however this would have been impractical because the size of the equipment would have required a very large cage.

The PoC demonstrated sending a command packet to activate the siren and place it into live Public Address mode (also found with very limited fuzzing, which a malicious actor could easily perform). Once the digital command for this mode is sent, any subsequent analog narrowband FM signals received by the radio are rebroadcast through the siren horns. Some canned messages were played back in this way.

The PoC video can be found here:

https://www.youtube.com/watch?v=YdnTBOBGjiA

PoC setup

Remediation

ATI has stated they have worked on increasing the level of security of their radio protocol, and this fix has now been reported to be rolled out across San Francisco's OPWS. During the weeks leading up the public disclosure, the OPWS frequency in San Francisco was active with an increasing number of packets that displayed higher entropy (appeared random), and activation commands in San Francisco have no longer been seen in the clear since public disclosure. No cryptanalysis has been performed to determine the efficacy of the fix.

Details of remediation steps have not been made available publicly. As such, we urge all customers to contact ATI to determine if their system is vulnerable, and if it is, confirm what remediation steps they should take. We encourage other vendors to check if their systems are vulnerable to this type of attack.

Suggested Remediations

The standard method to implement a minimum layer of security is twofold: require encryption and authentication. Encrypting messages makes it harder to analyse the underlying message content, but alone does not prevent a replay attack. Adding a layer of authentication with 'freshness' provides resilience against replay attacks by helping a recipient verify the authenticity of message (accept only new, legitimate messages, not old ones).

A suggested approach would be to add symmetric key encryption, with a sufficiently long key to prevent brute-force attacks to each packet's payload. This would require a change to the radio protocol, controller software and embedded firmware. These components would then be aware of encrypted messages, could recover the original message and act on them as they would ordinarily. Adding 'freshness' or randomisation, traditionally done with sequence numbers or Initialization Vectors (IVs), to each message would also mean that the same encrypted payload would appear different over-the-air when transmitted multiple times.

The other option available is to procure the P25 digital radio network upgrade from ATI, meaning that the siren system's communications will ride upon a separate, fully digital and standardized transmission scheme. The P25 protocol can then have strong encryption, such as AES, applied on top of it (so the siren system gets encrypted 'by proxy'). Encrypted P25 networks are used all over the globe, for example, in cities for public safety radio networks used by first responders (police, ambulance, etc). It does require, though, that the customer also install or upgrade their existing radio infrastructure (separate from the siren system) to be P25-compliant, and then upgrade it further to support encryption. That is, P25-enabled sirens will likely still need a separate P25 digital radio network for any sizable deployment, such as a

city.

Contact

To contact ATI Systems, please visit their website:

atisystem.com

More Information

Please visit:

sirenjack.com

References

[OPWSMap] http://sfdem.org/sites/default/files/FileCenter/Documents/2851-OPWS%20150205%20letter.pdf

[Glass11] S. Glass, V. Muthukkumarasamy, M. Portmann and M. Robert: “Insecurity in Public-Safety Communications: APCO Project 25”

Appendix A: Public Disclosure Document

This is the public-facing document that describes the vulnerability. Key details have been removed so as to avoid providing any information that could be used to attack unpatched systems.

This document can be found in plain-text here:

https://sirenjack.com/s/bastille-43-ati-insecure-protocol-public.txt

And as a formatted web page here:

https://www.sirenjack.com/bastille-advisory-43-sirenjack

```
Bastille Tracking Number 43



Overview

The deployment of ATI's mass notification system in San Francisco, CA
(known by the City and County of San Francisco as the city's Outdoor
Public Warning System) is vulnerable to false alarms. It can be
triggered by specially-crafted malicious radio transmissions, without
requiring physical access to any part of the system.



Affected Devices

In the basic configuration used by the SF OPWS:

HPSS16
HPSS32
MHPSS
ALERT4000*
```

Proof-of-Concept Details

These findings apply to the current SF OPWS configuration:

1. Discovering input/output frequencies
2. Air interface analysis
3. Malicious packet construction and transmission

The SF OPWS is exercised every Tuesday at noon by the SF Department of Emergency Management (DEM) in a test mode where all nodes are simultaneously triggered, and then each node is queried for its state in series (to which the node responds with its status).

The configuration uses a standard dedicated analog repeater approach. A central console (at the SF DEM) is used to trigger and monitor the nodes. The console, and the nodes, connect to analog radios that transmit and receive through the repeater to enable coverage across San Francisco.

A custom digital packet-based protocol is used to transmit commands and telemetry between the console and nodes.

The frames to activate the nodes are broadcast during the SF midday test, and contain the same payload each week. As they are transmitted in the clear, it is possible to construct similarly valid sequence of commands and transmit them on the repeater’s input frequency, thereby triggering all nodes. It is trivial to construct a different sequence of payloads, and/or attempt fuzzing, to see what other functions can be activated (for instance, live audio re-broadcast via radio).

* Learnt/guessed from public research only.

Mitigation

An attacker must find the dedicated input/output frequencies in use, reverse engineer the air interface, be able to craft malicious packets, and transmit those packets.

Suggested Solutions

Use standard approaches for message encryption and authentication.

Consideration of anti-jamming strategies is also advised.

Test Environment

San Francisco Outdoor Public Warning System (OPWS)

Note: No active testing was performed (i.e. no signals were transmitted). All analysis has been conducted passively.

Credits

Balint Seeber, Bastille

Appendix B: ICS-CERT Advisory

The advisory number is:

ICSA-18-100-01

The official web page can be found here:

https://ics-cert.us-cert.gov/advisories/ICSA-18-100-01

The following is a copy-and-paste of the original document. Please check the link above for the latest version.

Advisory (ICSA-18-100-01)

ATI Systems Emergency Mass Notification Systems

Original release date: April 10, 2018

Legal Notice

---

1. EXECUTIVE SUMMARY

- **CVSS v3 5.3**
- **ATTENTION**: Exploitable remotely.
- **Vendor**: Acoustic Technology, Inc. (ATI Systems)
- **Equipment**: ATI Emergency Mass Notification Systems
- **Vulnerabilities**: Improper Authentication, Missing Encryption of Sensitive Data.

2. RISK EVALUATION

Successful exploitation of these vulnerabilities could trigger false alarms.

3. TECHNICAL DETAILS

3.1 AFFECTED PRODUCTS

The following ATI's Emergency Mass Notification Systems devices are affected:

- HPSS16
- HPSS32
- MHPSS, and
- ALERT4000.

3.2 VULNERABILITY OVERVIEW

3.2.1 IMPROPER AUTHENTICATION CWE-287

Improper authentication vulnerability caused by specially crafted malicious radio transmissions may allow an attacker to remotely trigger false alarms.

CVE-2018-8862 has been assigned to this vulnerability. A CVSS v3 base score of 5.3 has been calculated; the CVSS vector string is (AV:A/AC:H/PR:N/UI:N/S:U/C:N/I:H/A:N).

3.2.2 MISSING ENCRYPTION OF SENSITIVE DATA CWE-311

Missing encryption of sensitive data vulnerability caused by specially crafted malicious radio transmissions may allow an attacker to remotely trigger false alarms.

CVE-2018-8864 has been assigned to this vulnerability. A CVSS v3 base score of 5.3 has been calculated; the CVSS vector string is (AV:A/AC:H/PR:N/UI:N/S:U/C:N/I:H/A:N).

3.3 BACKGROUND

- **Critical Infrastructure Sectors:** Commercial Facilities, Defense Industrial Base, Emergency Services, Government Facilities, Nuclear Reactors, Materials, and Waste
- **Countries/Areas Deployed:** Worldwide
- **Company Headquarters Location:** Massachusetts

3.4 RESEARCHER

Balint Seeber of Bastille reported these vulnerabilities to NCCIC.

4. MITIGATIONS

ATI has created a patch which adds additional security features to the command packets sent over the radio. ATI is testing this patch, and it will be available upon request. Many systems are engineered to meet specific user needs and users need to make sure any upgrades are appropriate for their systems.

ATI recommends that, where feasible, simple voice radios be replaced with digital P-25 (APCO) radios, which provide highly secure encrypted links.

NCCIC reminds organizations to perform proper impact analysis and risk assessment prior to deploying defensive measures.

NCCIC also provides a section for control systems security recommended practices on the ICS-CERT web page. Several recommended practices are available for reading and download, including Improving Industrial Control Systems Cybersecurity with Defense-in-Depth Strategies.

Additional mitigation guidance and recommended practices are publicly available in the ICS‑CERT Technical Information Paper, ICS-TIP-12-146-01B--Targeted Cyber Intrusion Detection and Mitigation Strategies, that is available for download from the ICS-CERT website.

Organizations observing any suspected malicious activity should follow their established internal procedures and report their findings to NCCIC for tracking and correlation against other incidents.

No known public exploits specifically target these vulnerabilities. High skill level is needed to exploit.