August 2023

SAFETY PATROL: MECHANISMS FOR IMPROVING EMERGENCY RESPONSE TIMES

Tracy Caivano

Paul B Giralt

David C White, Jr.

Jason F Haase

SAFETY PATROL: MECHANISMS FOR IMPROVING EMERGENCY RESPONSE TIMES

AUTHORS:
Tracy Caivano
Paul B Giralt
David C White, Jr.
Jason F Haase

ABSTRACT

To equip first responders with critical, time-sensitive information and accelerate emergency services response times, various solutions are provided herein through several techniques. Under a first technique, after an emergency event such as a gunshot is either automatically detected by a camera or manually initiated by a user, or when a dangerous object such as a gun is detected by a camera, a network may react by associating the source of the dangerous event or object with a person based on proximity data; identifying the physical characteristics of the person (such as height, hair color, clothing, visible tattoos, etc.); attaching such characteristics as textual metadata; and then transmitting that metadata to first responders. A second technique automatically develops a radio frequency (RF) signature profile of a person of interest (from RF signals emitted by devices carried by the person), associates that profile to the person, and leverages that profile to track the person as they move throughout a building or campus, allowing a user to look back in time (to, for example, identify where a person came from and how they entered a building) by tracking the RF profile over time. The above-described data is extremely important during any ongoing emergency and equips first responders with critical information which only a network can provide.

DETAILED DESCRIPTION

During an emergency situation such as a mass shooting, every second counts. Time is of the essence when it comes to notifying law enforcement personnel and other first responders of the unfolding situation. Further, providing accurate, real-time information is critical to the ability of such individuals to successfully assess the situation and quickly respond, thereby saving valuable time and lives.

During such a situation, the questions that first responders will be eager to resolve include are gunshots still being fired; if so, where are those shots coming from; who the shooter is; what the description of the shooter is; and if the shooter is moving, to where they are moving.

The military knows well that information is the key to winning a battle. And network equipment vendors empower schools and businesses with such information through their networks as well as their Internet Protocol (IP) cameras and sensors. As a result, those vendor’s products could form part of a solution that provides first responders with key information to help them assess and respond to attacks.

To support the development and the dissemination of the type of information that was described above, various solutions are provided herein through several techniques. Each of the techniques will be briefly introduced here and then described in detail later in the instant narrative. Under a first technique, when an emergency event (such as a gunshot, the presence of a gun, etc.) is either automatically detected or manually asserted, a system according to the technique may quickly begin to connect the technology dots to provide critical, situational information. A second technique leverages the reality that people today almost always carry with them some form of device that emits RF signals and supports the automatic development of an RF profile (based on those signals) and the use of such a profile to track a person as they move about a space.

Turning to the first technique, as referenced above, aspects of this technique include, among other things, three key features. A first feature encompasses the tasks of detecting and associating, a second feature encompasses the task of describing, and a third feature encompasses the task of acting.

Under the first key feature (detecting and associating), computer vision methods may, using known approaches, be employed to identify people and objects (such as weapons), detect any person that appears to be holding a weapon (or other suspicious object), and highlight that person on a floorplan (using existing approaches that facilitate the mapping of a person that is seen on a camera to a location on a floorplan).

Additionally, a first responder may manually tag a person of interest from a video feed even if that person had not automatically been detected as a dangerous person (such as someone carrying a weapon or other suspicious object). Further, the first responder or a public safety individual can indicate to a system (according to the first technique) more than one known appearance of a suspect. Still further, such a system can also automatically correlate multiple angles of a person on its own without requiring human intervention.

Under the second key feature (describing), after the system has acquired one or more sample snapshots of a potential suspect, a profile of that individual (including textual descriptions) may be developed based on image(s) in the video, along with relevant situational and demographic information, for use by the authorities. A computer vision method may analyze one or more snapshots or video snippets of the potential suspect and translate different characteristics of that person into text. Such visual context information may describe facial features, hair color, hair style, estimated age, height, clothing, a weapon, etc.

Traditionally, the recognition of facial features is often challenged due to the quality of the original video image or the fact that a person of interest may not be looking straight into a camera. Such issues impact the ability of facial recognition software (and humans) to be able to accurately identify a suspect. The second feature improves upon existing methods because once someone has been identified, cameras may track that person and save the best (e.g., 10) images that most clearly show the person's face. If the person is not looking directly at the camera, the system may continue to run a still frame buffer and build on the profile each time a 'better' or clearer image is captured (meaning that a greater percentage of the face was captured in the frame, both eyes were directed to the camera, etc.). Such an approach dramatically saves time for individuals attempting to review past footage looking for a clear image of a suspect. Instead, the system may capture the best images for review and display and then continue monitoring to improve on the set of captured images. Artificial intelligence (AI) methods may also be used to gather multiple snapshot inputs and construct a single photograph of what the suspect looks like, by aggregating the characteristics of the person, while simultaneously improving the text-based description of the person.

By leveraging a network and its devices, the approach described above provides for the fast, accurate identification of a suspect and the means to provide the developed information to the appropriate authorities.

Under the third key feature (acting), the developed profile information (as described above) may be provided to local authorities either through a portal or directly to an emergency response team. Additionally, such information may also be provided through a text-to-911 facility. Alternatively, a system according to the first technique may interact with a text-to-speech facility to relay the information over radio frequency (RF) channels, or even over a site's intercom system, to convey to everyone in a building the description of the attacker.

Through the first technique (as described above), first responders are able to arrive on a scene with accurate, descriptive information (including a description of a suspect) already at hand, thus reducing the time that would otherwise be spent collecting situational information and reducing misinformation.

Figure 1, below, presents elements of an exemplary solution that is possible according to aspects of the first technique and which is reflective of the above discussion.

Figure 1: Exemplary Solution

Importantly, aspects of the first technique may also be utilized in similar events (such as a building fire, etc.) that require situational awareness on the part of first responders.

As described and illustrated above, a central element of the first technique encompasses the activities that take place following the automatic or manual triggering of an event.

One of those activities encompasses the creation of a description of a target individual, capturing key identification characteristics (including, but not limited to, height, gender, hair color, eye color, weight, clothing, etc.) and location information (such as, for example, “third floor, room 302”) for that individual.

Regarding the capture of identification characteristics, under a system according to the first technique image recognition may continue to track a person and store a buffer of (e.g., five) images, constantly looking for better images (since ceiling-mounted cameras may not always capture crisp and clear images) as well as using multiple similar images to improve the quality of an image and a description (since a layering of multiple images allows for the emergence of a clearer picture). A constant comparison with the top or best (e.g., 5) clearest images allows the system to know if it has captured a ‘better’ shot, specifically considering clarity, and also determine whether a candidate shot captures an individual’s full facial features. During such a comparison, an ideal image may be one where the suspect is looking directly at the camera, or which contains distinguishing features. As a result of the above-described process, a developed description represents a composite of different attributes that are determined from a variety of images of the person of interest.

Another of the activities encompasses acting on the developed information by disseminating the same to authorities, so when those individuals enter a scene, they know exactly what a suspect looks like. Optionally, this information (which is now in textual form) may also be automatically sent to different staff through a text or other messaging channel or even announced over an intercom. In the case of the latter, the announcement may be tailored to limit the amount of information (e.g., “Person wearing white sweatshirt and blue jeans detected with a gun on the 3rd floor, room 302. Please evacuate the building.”).

Turning to the second technique, as referenced above, aspects of this technique combine video object tracking and RF-based location tracking capabilities in a novel way to enable the seamless following of a subject of interest in an emergency or other situation that requires the ability to track a person’s movement throughout a space that is equipped with both video surveillance and a wireless network that is capable of detecting devices (even if those devices are not actively connected to the wireless network).

People today almost always carry with them some form of device that emits RF signals. Such devices include mobile phones, smart watches, activity trackers, tracking devices, laptops, cameras (with Wi-Fi or Bluetooth capabilities), etc. Such devices emit RF signals that are based on the specific device and its protocol, but commonly the devices employ a cellular protocol (such as long-term evolution (LTE) technology, 3rd Generation Partnership Project (3GPP) fourth generation (4G) technology, 3GPP third generation (3G) technology, etc.), Wi-Fi (such as Wi-Fi 6 or the Institute of Electrical and Electronics Engineers (IEEE) standard 802.11ax, Wi-Fi 5 or the IEEE standard 802.11ac, etc.) and/or Bluetooth or Bluetooth Low Energy (BLE), all of which may be used to detect and track a device.

Aspects of the second technique leverage such pervasive devices to automatically create one or more RF profiles and then associate the same with the corresponding device-carrying person. This may be accomplished by identifying the RF devices that a network 'sees,' confirming their location on a floorplan or map, and then tracking their movement throughout an area using existing approaches to triangulate and store location data.

For devices which are connected to the network, operating system (OS) details and other metadata may be used to confirm whether a device is static (i.e., it is stationary) or dynamic (i.e., it moves). RF devices which are static and do not move may be automatically tagged as such and then excluded from a human profile. RF devices that move over time may be further evaluated to identify those devices which always move together, which may then be included in the same profile.

At the end of the above-described process, a system according to the second technique knows what devices are moving around together, where they are, and where they have been at particular times. Additionally, the system may indicate whether a profile has been designated as belonging to a human but for which the human's identity is currently unknown (a state which will be resolved later in the instant process, as described below).

The above-described process may operate constantly, attempting to group devices that are moving together, as captured in the illustrative process flow that is presented in Figure 2, below.

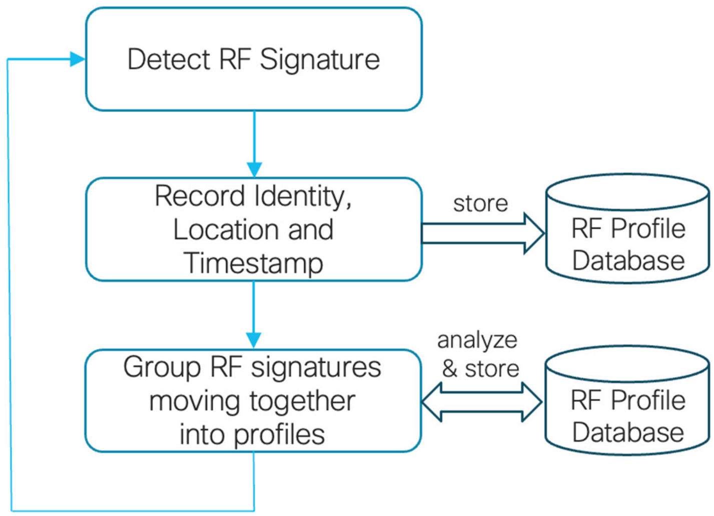

Figure 2: Illustrative Process Flow

Considering the above-described background information, the second technique leverages two existing technologies – first, the use of a camera to detect a person or object and second, the ability to map a visual detection of a person or object to a location on a floorplan – that may be optionally triggered by an emergency event (such as a gunshot) and then identifies and links the RF profile of a person to allow for RF tracking by correlating the location of the person of interest with the RF profile that is found in that location at the same time. Multiple visual identifications of the person of interest at different locations along with an association of the same RF profile may be used to further increase a confidence in an accurate correlation (e.g., the same person may be identified at two different times by two different cameras, that reside in two separate locations, and at those two times the same RF profile may be detected in those locations).

Aspects of the second technique may be further understood through an illustrative example. Under that example, at 8:42 a.m. a person walks into the north entrance of a building and, at 8:43 a.m., that person brandishes and fires a gun. Existing object detection and sound detection facilities may identify this activity as an emergency event. Alternatively, security personnel may manually identify a person of interest (from, for example, their monitoring of the footage from a security system) and indicate to a system according to the second technique that a specific person seen at a specific time on a specific camera is the person that they wish to track. Figure 3, below, depicts elements of the above-described activities.

Figure 3: Detection of Object and Association with a Person

At that point, the security personal could look back at the footage from 8:42 am and optionally further identify where the suspect was located at a previous time, giving the system more than one known location of the subject to improve correlation accuracy. Alternatively, existing methods for the visual tracking of a subject across cameras may provide the system with additional known locations of the suspect at different times leading up to the instant event.

The above-described trigger (of an emergency event or the identification of a person of interest) may prompt the network to identify the RF profile that is associated with the person of interest (in this case the person that was identified as having a gun). Existing methods may be employed to map a person who is seen in a camera feed with a physical location on a floorplan or other geographic map and existing products may map an RF device to a location on a floorplan. The second technique automatically correlates these two pieces of location data to be able to match an RF signature to a person that is seen on a video camera capture.

Using the above-described suspect location data, the system knows where the individual is and may examine all of the RF profiles that are located at that spot at that time. The system can then associate an RF profile to the suspect and subsequently track that suspect throughout a building or campus based on their unique RF profile, as depicted in Figure 4, below.

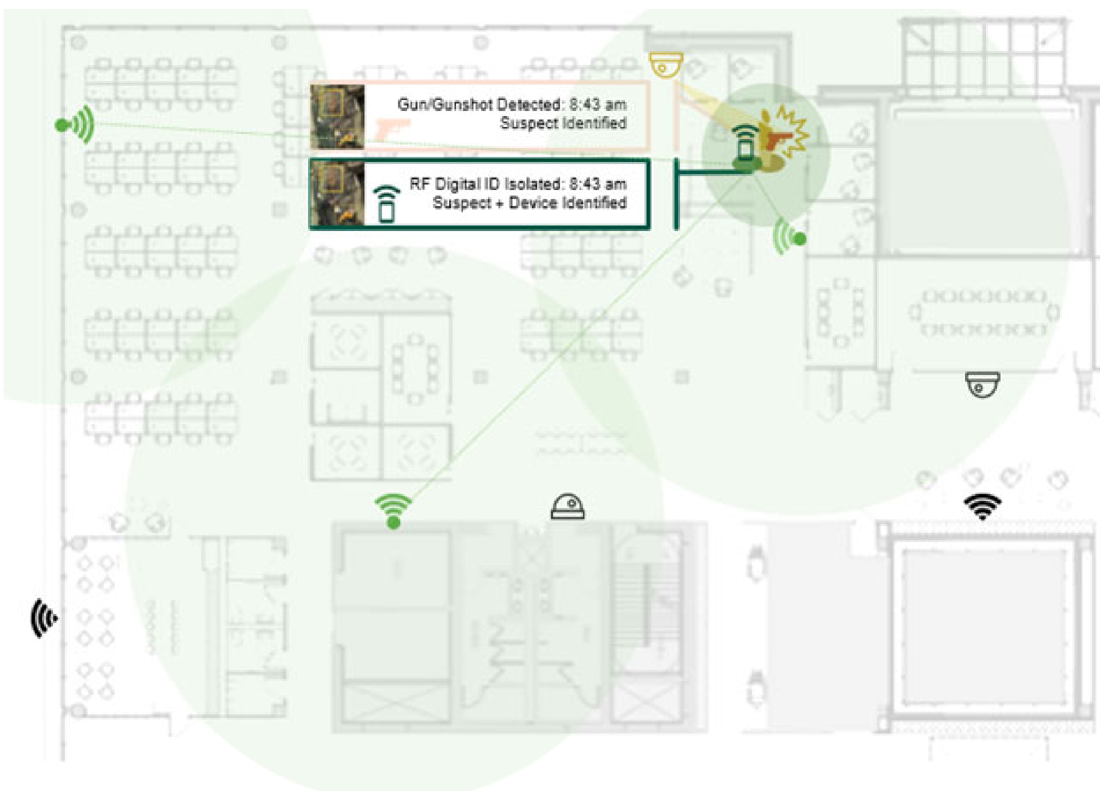

Figure 4: RF Signature Isolation and Correlation with Suspect

It is important to note that RF emissions may be intermittent and the devices that are being carried by the suspect may not have transmitted any radio energy at the exact moment that a video capture was taken. However, a system according to the second technique would have tracked the different devices in a given area before and after the event and can therefore interpolate where a detected device could have been at the location of the camera footage at the time of the capture. The system may therefore estimate the positions of the different RF devices at the time in question based on measurements before and after the event in question.

It is possible that multiple RF profiles may be found in the vicinity of the location where the suspect was visually located. In such a case, additional identifications may be necessary to discover the specific profile that is associated with the person of interest. For example, if at 8:42 a.m. the suspect was seen by a camera near the north entrance and the system detected that two different RF profiles (e.g., Profile A and Profile B) were identified as being in the vicinity of that location at that time; and then at 8:43 a.m., in a different part of the building, the suspect was identified by a camera and the system detected that Profile A, Profile C, and Profile D were located in the vicinity of this second location; the system can conclude that Profile A must be that of the suspect. Such additional location confirmations can help to increase a confidence level.

As indicated previously, aspects of the second technique support a suspect tracking capability. Continuing with above example, as a suspect moves throughout a building or campus, they will pass in and out of the view of different cameras, but they will never leave the RF range of the various access points. Thus, the network is able to track the suspect's RF profile and constantly show the suspect's location on a floorplan, even when the suspect is out of camera view. Additionally, since the network has been alerted to track the person, it may not only provide real-time location information but also provide historical tracking of where the person came from (based on backtracking their RF profile from the moment it entered the network). Figure 5, below, depicts elements of the above-described activities.

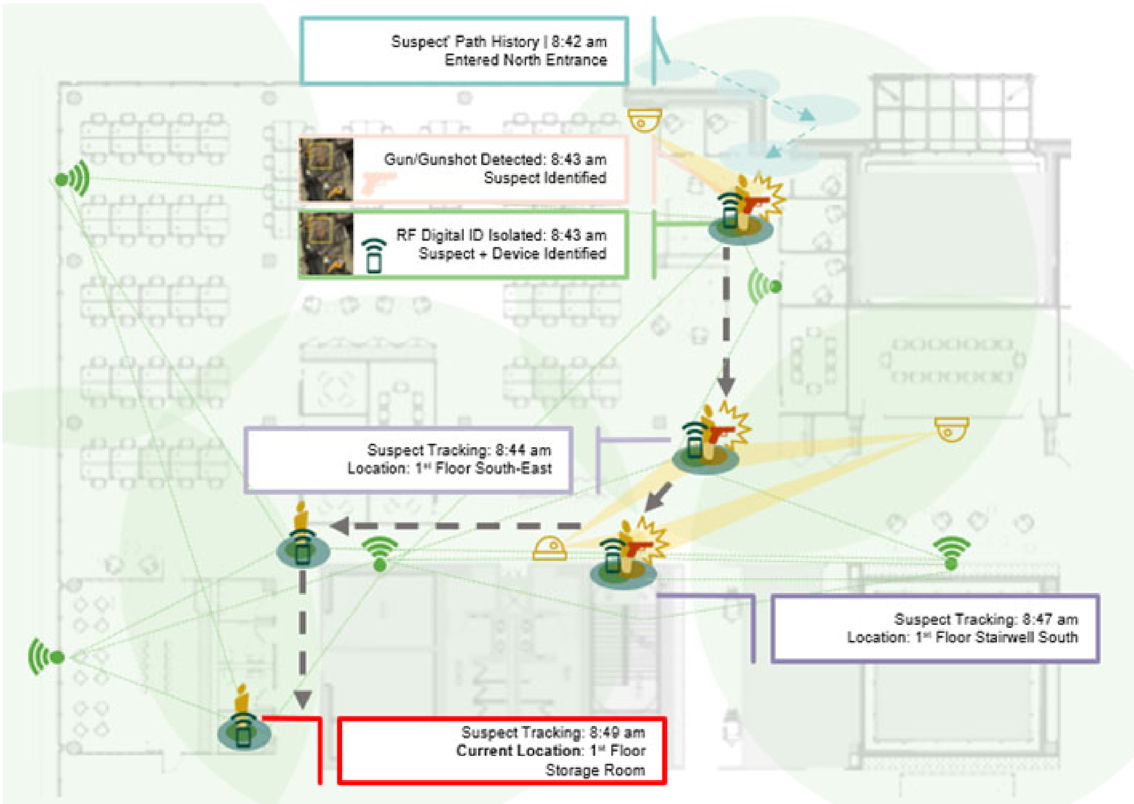

Figure 5: Movement Tracking of Visual and RF Data

As can be seen from Figure 5, above, the suspect is currently hiding out of the view of any cameras, but the RF profile indicates that they are in a storage room on the first floor of the building.

The flowchart that is presented in Figure 6, below, outlines the above-described process, from initial detection through ongoing tracking, as supported by aspects of the second technique.

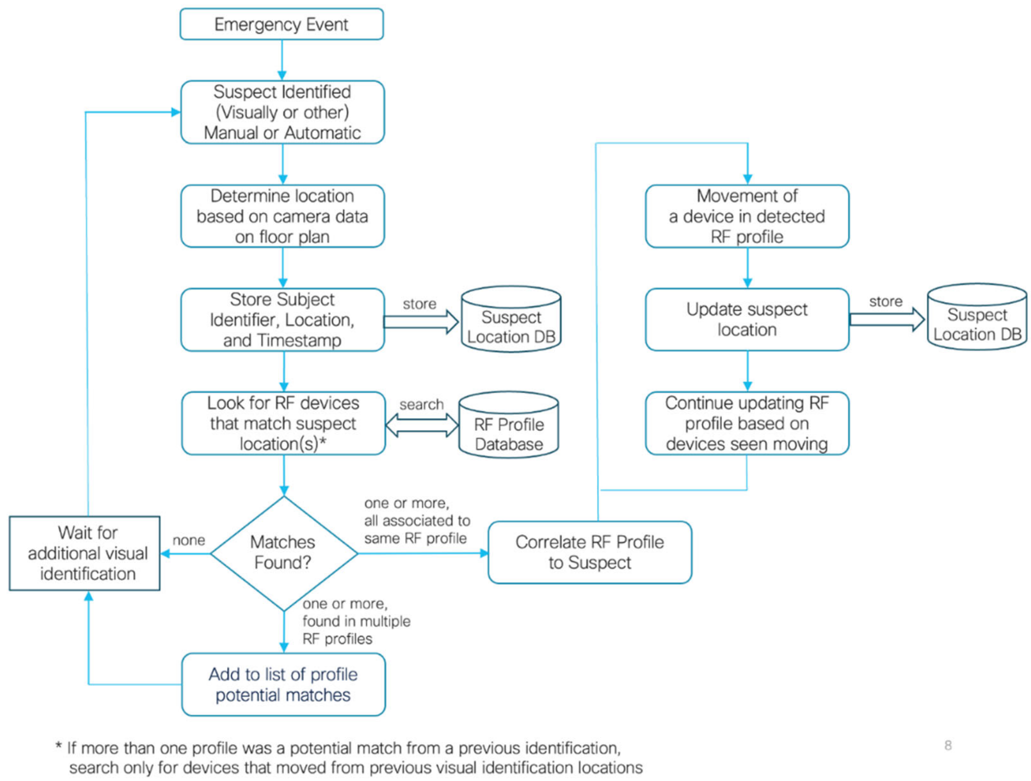

Figure 6: Process Flow

In the event that multiple profiles are found as potential matches, the system may track and display the movements of multiple profiles simultaneously. As those profiles diverge (e.g., as they separate in latitude and longitude) the system can then pinpoint which of the profiles is associated with the suspect. If the identified user is outside of a camera range, then multiple profiles may be tracked independently until various of the profiles come into camera range at which point one of the profiles may be verified as the profile that is associated to the suspect.

The above-described location information is critically important to first responders in emergency situations, since knowing a suspect's location is vital. However, knowing the suspect's location is only possible through the second technique, which extends the visibility of someone by associating an RF profile to a person and then tracking that RF profile as it moves throughout a building or campus. In connection with the process that was described above, standard RF triangulation techniques may be employed to identify the precise location of an RF emission and then overlay that location on the floorplan of a building.

Under the second technique, a person's device does not need to 'join' or 'associate' to a network for a profile to be created or for that profile to be tracked. As everyday consumer devices emit RF 'chatter' just by being powered on, it is possible to leverage that chatter to track the device (and, by extension, the person carrying the device). Additionally, existing access points provide Wi-Fi as well as Bluetooth and even LTE services, all of which allow for device tracking and triangulation.

According the second technique, the data that may be pushed into a suspect location database can then be used by a variety of systems to provide real-time location information to first responders in the form of maps, floorplans, audio recordings, or augmented reality applications that visually direct first responders to the suspect.

As described and illustrated above, the second technique encompasses a number of novel elements. First, the technique supports the automatic and continuous creation of an RF profile for each of the (possibly multiple) devices that are carried by a given person and the grouping of devices that are seen to be moving together over time. Second, the technique associates an RF profile to a person that has been identified on a camera or that has been selected through another method. Third, the technique tracks an RF profile as an overlay (with timestamps) on a floorplan map, from where and when the profile was first seen to where the profile is currently located and provides the suspect's location data to other systems for rendering in other ways. Fourth, the technique provides alerts on a suspect's current location through various communication means including a building's intercom system, broadcast messages (such as pages, email messages, or any other notification mechanism), etc. Fifth, the technique provides a portal through which an individual may be manually selected after which a selected person may be tracked (as, for example, they navigate a building or campus) through their RF profile signature.

In summary, to equip first responders with critical, time-sensitive information and accelerate emergency services response times, various solutions have been provided herein through several techniques. Under a first technique, after an emergency event such as a gunshot is either automatically detected by a camera or manually initiated by a user, or when a dangerous object such as a gun is detected by a camera, a network may react by associating the source of the dangerous event or object with a person based on proximity data; identifying the physical characteristics of the person (such as height, hair color, clothing, visible tattoos, etc.); attaching such characteristics as textual metadata; and then transmitting that metadata to first responders. A second technique automatically develops a RF signature profile of a person of interest (from RF signals emitted by devices carried by the person), associates that profile to the person, and leverages that profile to track the person as they move throughout a building or campus, allowing a user to look back in time (to, for example, identify where a person came from and how they entered a building) by tracking the RF profile over time. The above-described data is extremely important during any ongoing emergency and equips first responders with critical information which only a network can provide.