US 20190251138A1

(19) United States

(12) Patent Application Publication (10) Pub. No.: US 2019/0251138 A1

Patton et al. (43) Pub. Date: Aug. 15, 2019

(54) **DETECTING EVENTS FROM FEATURES DERIVED FROM MULTIPLE INGESTED SIGNALS**

(71) Applicant: **Banjo, Inc.**, Park City, UT (US)

(72) Inventors: **Damien Patton**, Park City, UT (US); **Rish Mehta**, Redwood City, CA (US); **Tilmann Bruckhaus**, Redwood City, CA (US)

(21) Appl. No.: **16/029,481**

(22) Filed: **Jul. 6, 2018**

Related U.S. Application Data

(60) Provisional application No. 62/628,866, filed on Feb. 9, 2018, provisional application No. 62/654,274, filed on Apr. 6, 2018, provisional application No. 62/654,277, filed on Apr. 6, 2018, provisional application No. 62/664,001, filed on Apr. 27, 2018, provisional application No. 62/682,176, filed on Jun. 8, 2018, provisional application No. 62/682,177, filed on Jun. 8, 2018.

Publication Classification

(51) **Int. Cl.**

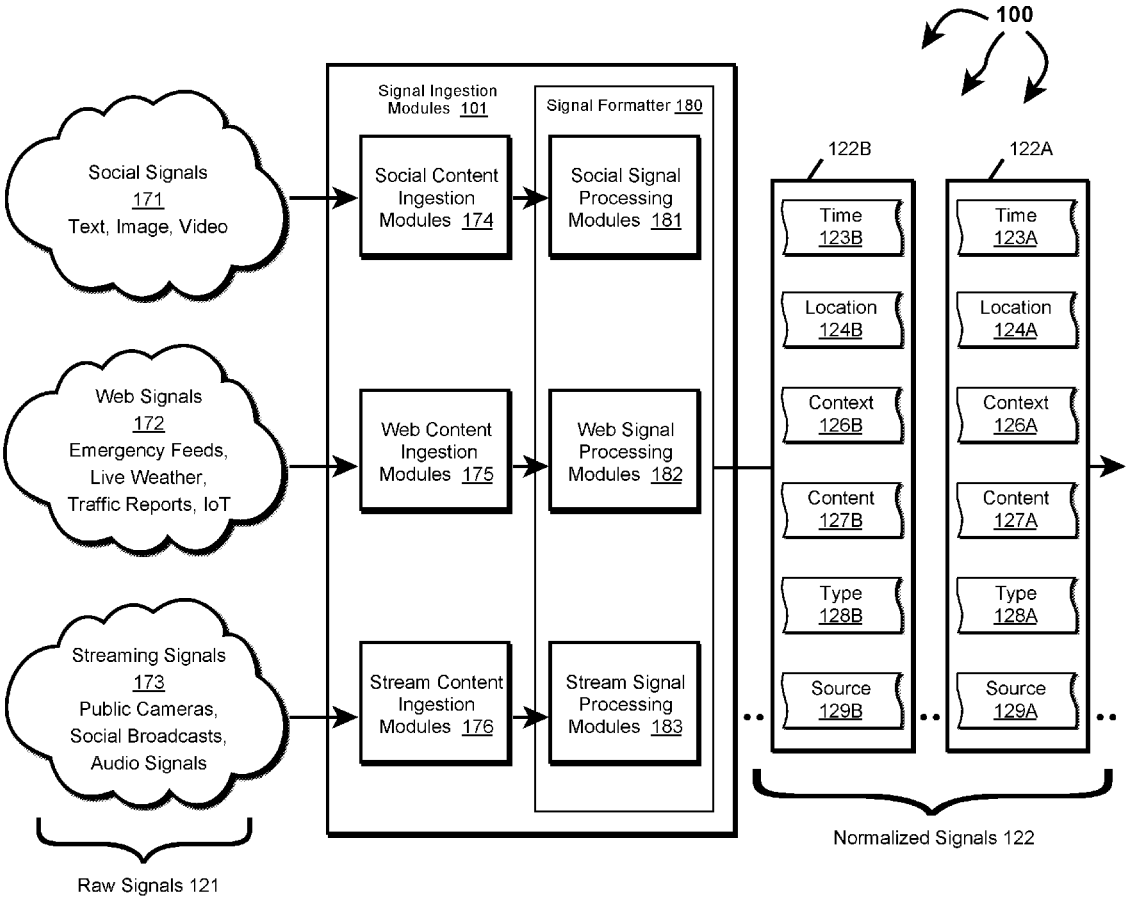
G06F 17/18	(2006.01)
G06F 17/30	(2006.01)
G06F 15/18	(2006.01)

(52) **U.S. Cl.**

CPC ***G06F 17/18*** (2013.01); ***G06F 17/30241*** (2013.01); *H04W 4/90* (2018.02); ***G06F 15/18*** (2013.01); ***G06F 17/30598*** (2013.01)

(57) **ABSTRACT**

The present invention extends to methods, systems, and computer program products for detecting events from features derived from multiple signals. In one aspect, an event detection infrastructure determines that characteristics of multiple signals, when considered collectively, indicate an event of interest to one or more parties. In another aspect, an evaluation module determines that characteristics of one or more signals indicate a possible event of interest to one or more parties. A validator then determines that characteristics of one or more other signals validate the possible event as an actual event of interest to the one or more parties. Signal features can be used to compute probabilities of events occurring.



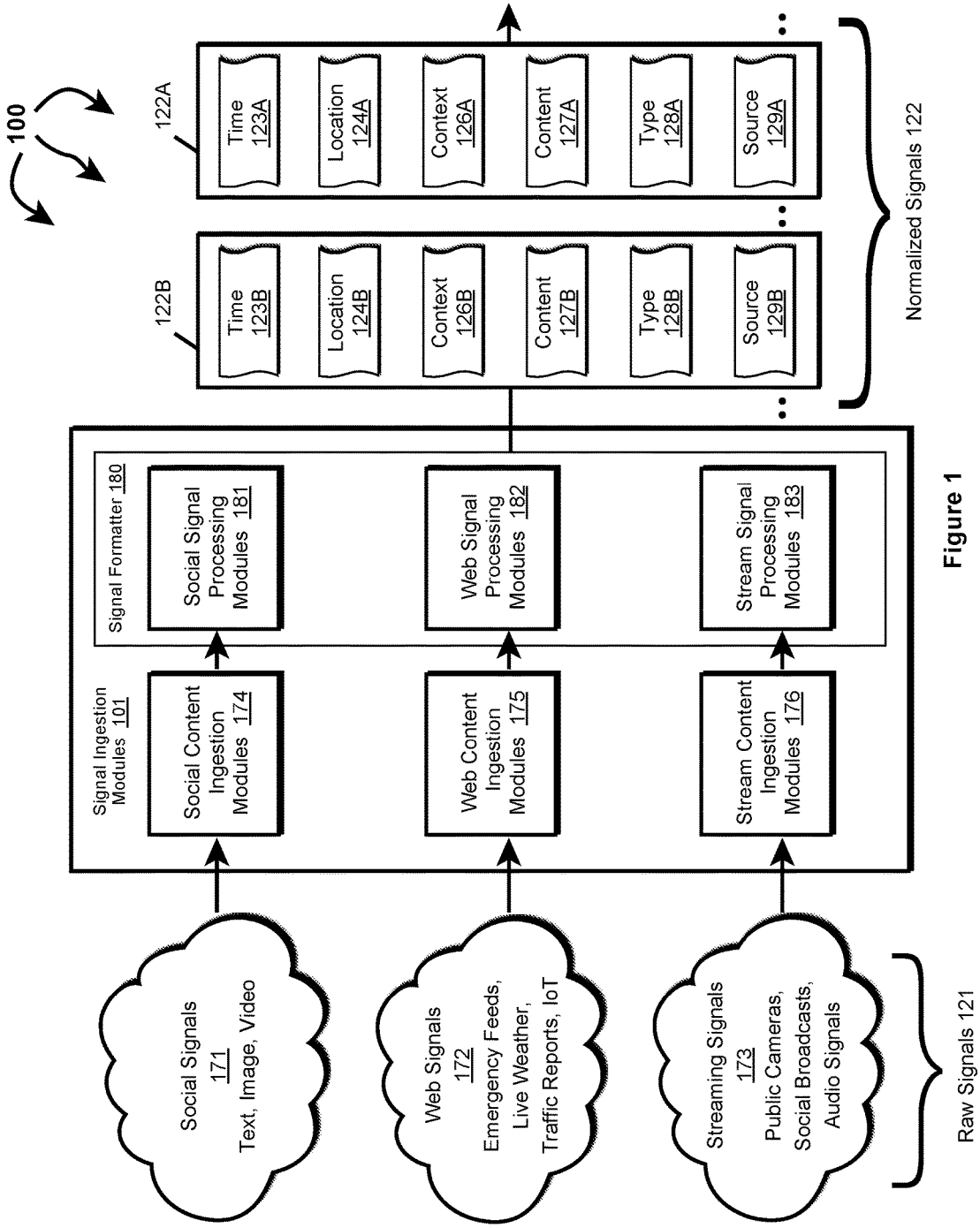

100
Social Signals 171 Text, Image, Video
Web Signals 172 Emergency Feeds, Live Weather, Traffic Reports, IoT
Streaming Signals 173 Public Cameras, Social Broadcasts, Audio Signals
Raw Signals 121
Signal Ingestion Modules 101
Social Content Ingestion Modules 174
Web Content Ingestion Modules 175
Stream Content Ingestion Modules 176
Signal Formatter 180
Social Signal Processing Modules 181
Web Signal Processing Modules 182
Stream Signal Processing Modules 183
122B
Time 123B
Location 124B
Context 126B
Content 127B
Type 128B
Source 129B
122A
Time 123A
Location 124A
Context 126A
Content 127A
Type 128A
Source 129A
Normalized Signals 122

Figure 1

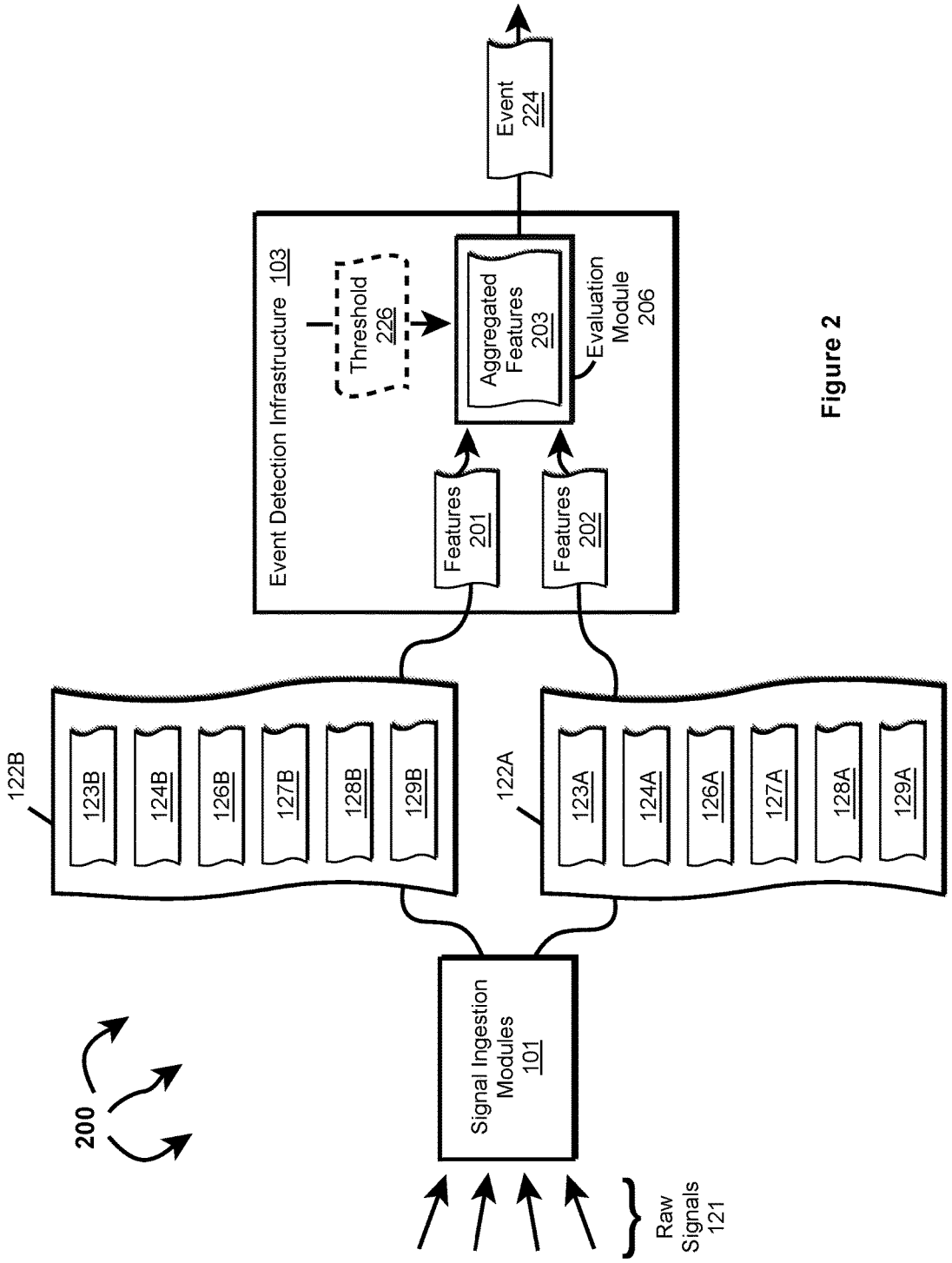

200
Raw Signals 121
Signal Ingestion Modules 101
122B
123B
124B
126B
127B
128B
129B
122A
123A
124A
126A
127A
128A
129A
Features 201
Features 202
Event Detection Infrastructure 103
Threshold 226
Aggregated Features 203
Evaluation Module 206
Event 224

Figure 2

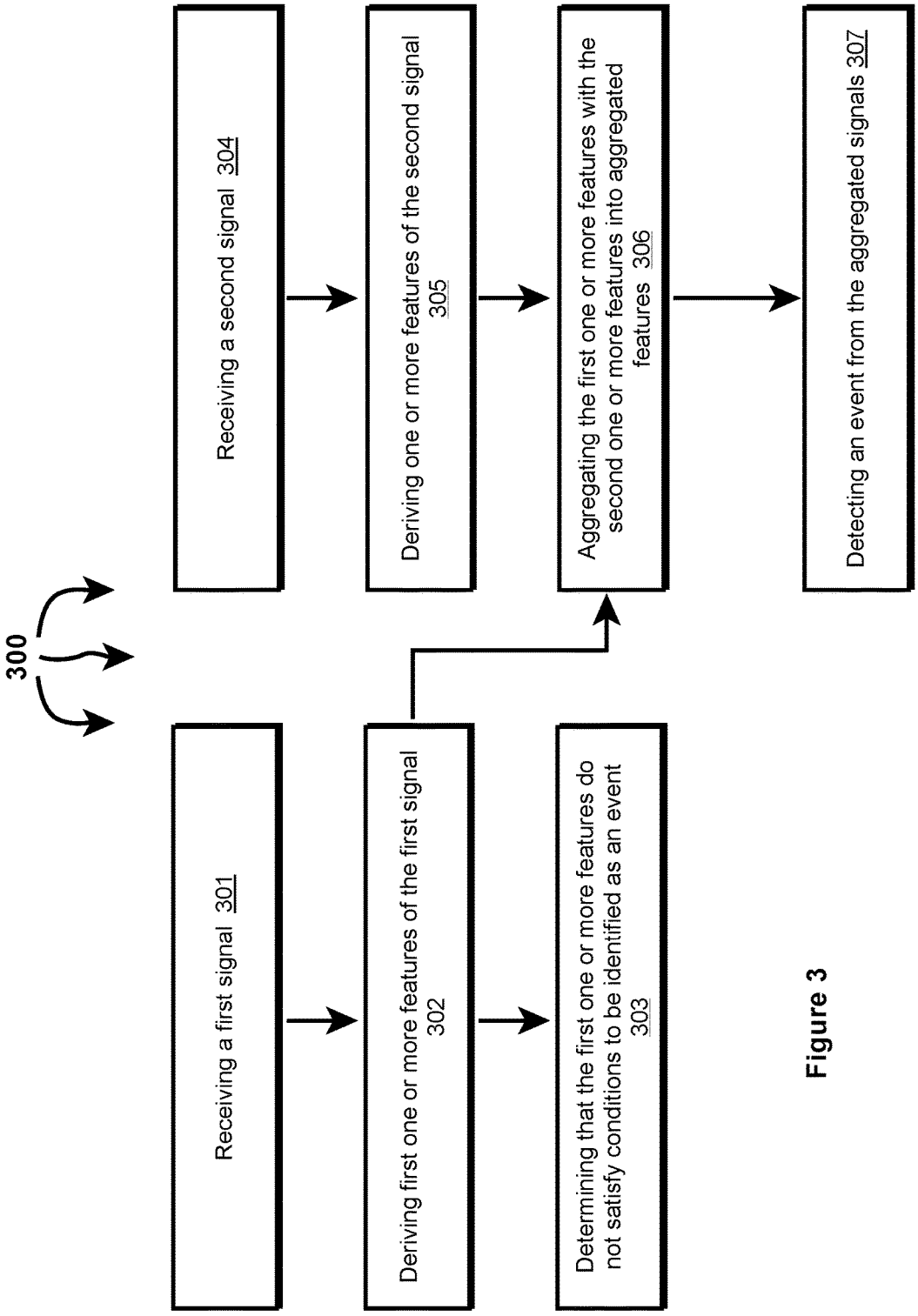

300
Receiving a first signal 301
Deriving first one or more features of the first signal 302
Determining that the first one or more features do not satisfy conditions to be identified as an event 303
Receiving a second signal 304
Deriving one or more features of the second signal 305
Aggregating the first one or more features with the second one or more features into aggregated features 306
Detecting an event from the aggregated signals 307

Figure 3

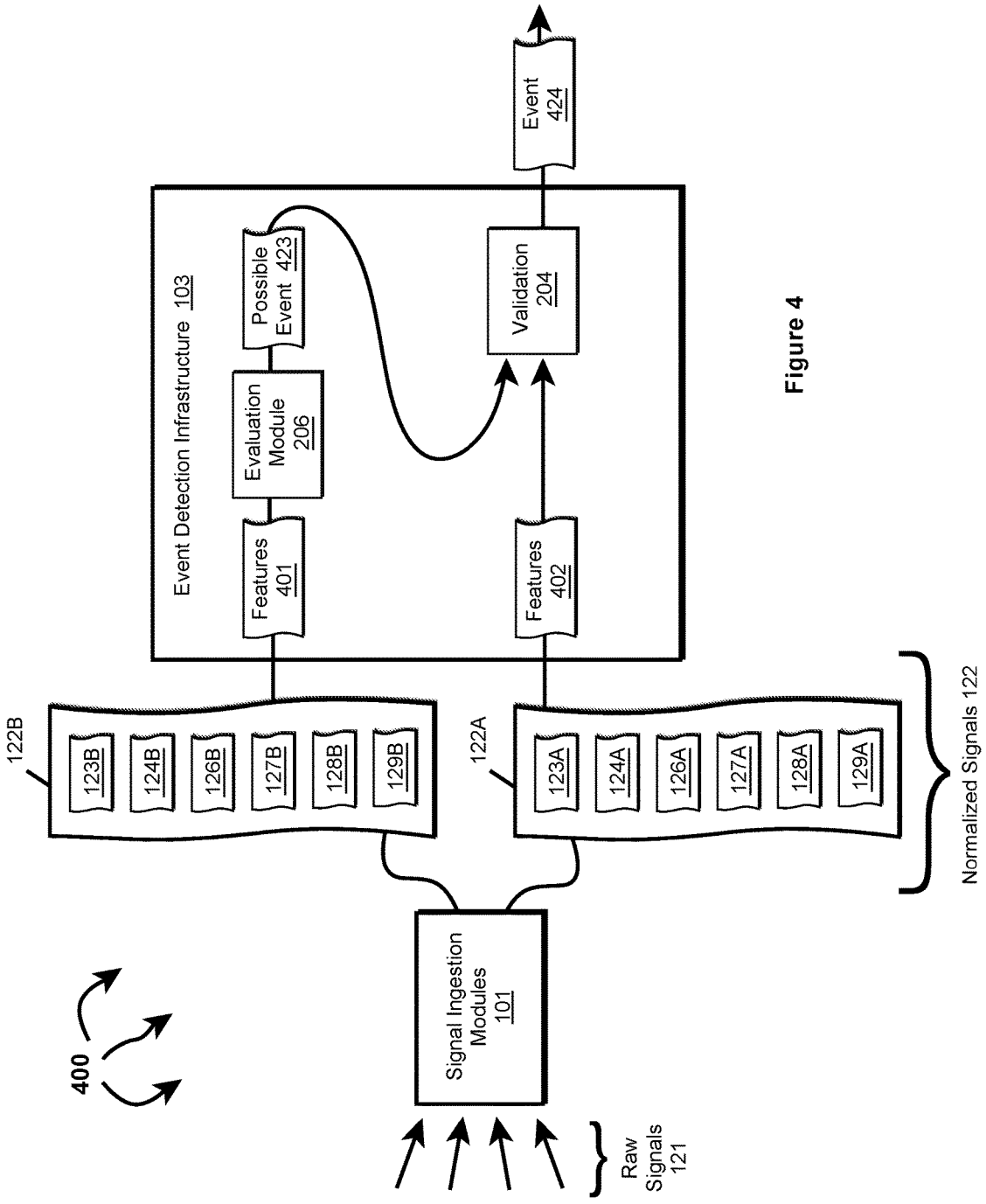

400
Raw Signals 121
Signal Ingestion Modules 101
122B
123B
124B
126B
127B
128B
129B
122A
123A
124A
126A
127A
128A
129A
Normalized Signals 122
Event Detection Infrastructure 103
Features 401
Evaluation Module 206
Possible Event 423
Features 402
Validation 204
Event 424

Figure 4

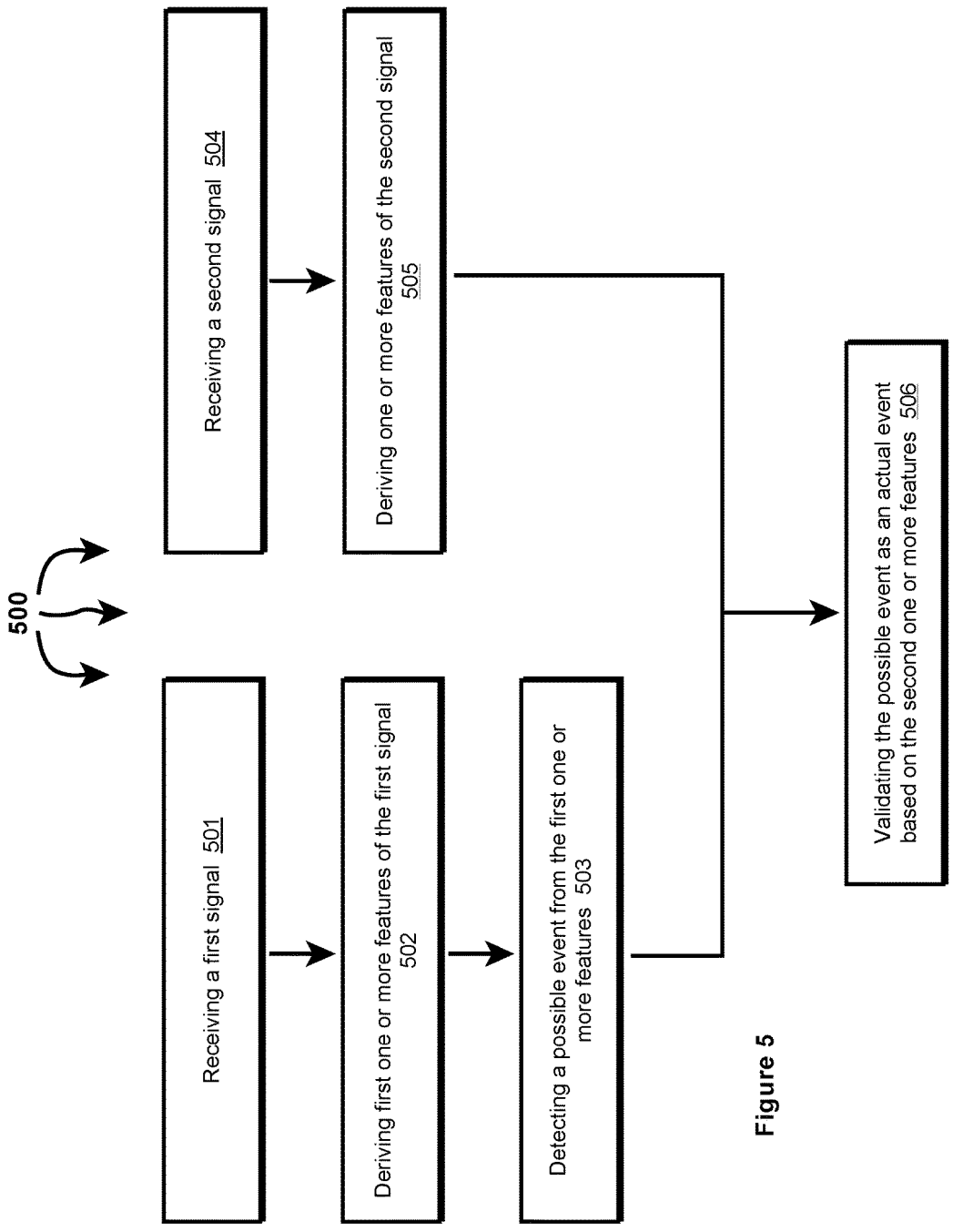

500
Receiving a first signal 501
Deriving first one or more features of the first signal 502
Detecting a possible event from the first one or more features 503
Receiving a second signal 504
Deriving one or more features of the second signal 505
Validating the possible event as an actual event based on the second one or more features 506

Figure 5

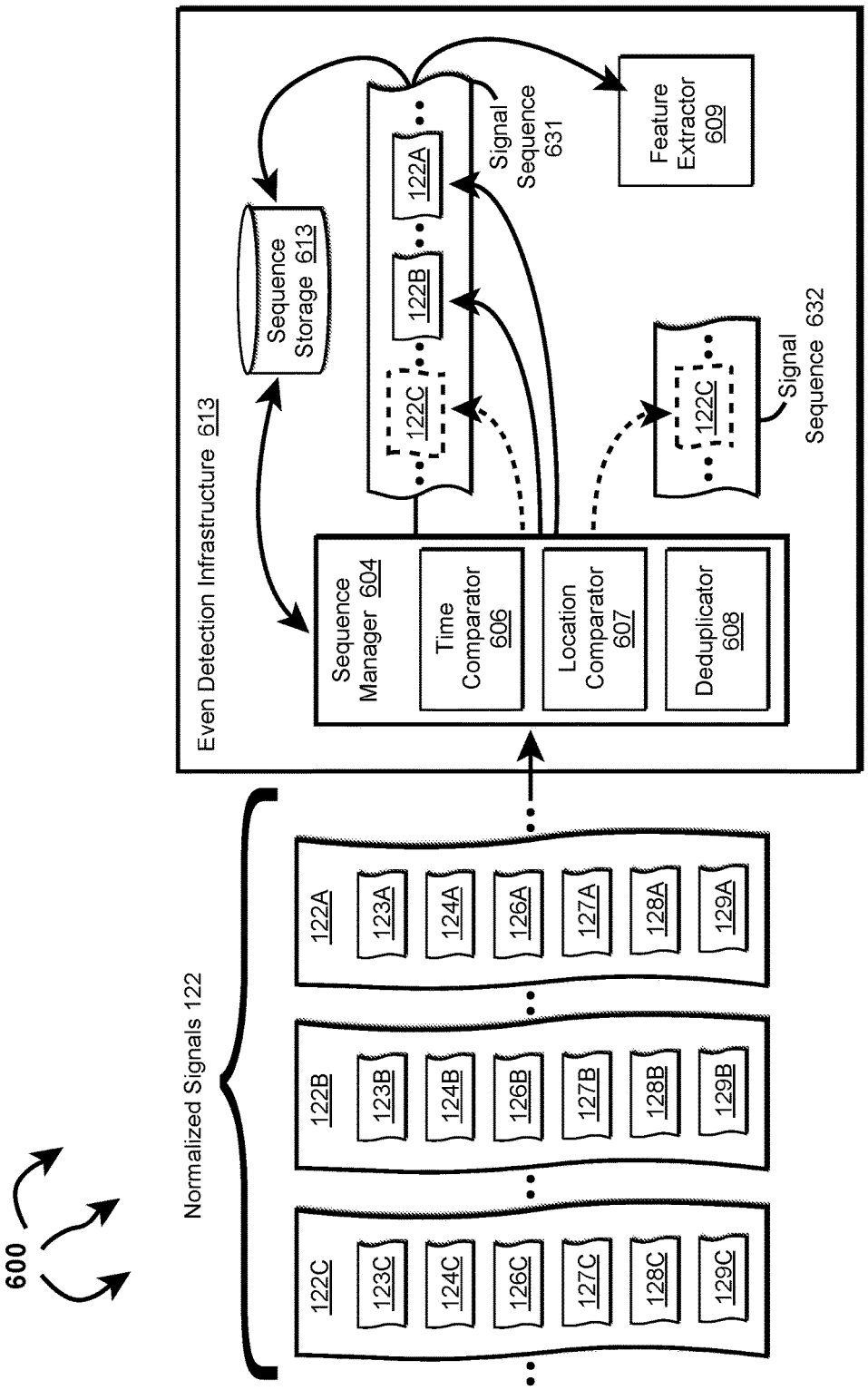

600
Normalized Signals 122
122C
123C
124C
126C
127C
128C
129C
122B
123B
124B
126B
127B
128B
129B
122A
123A
124A
126A
127A
128A
129A
Even Detection Infrastructure 613
Sequence Manager 604
Time Comparator 606
Location Comparator 607
Deduplicator 608
Sequence Storage 613
122C
122B
122A
Signal Sequence 631
Feature Extractor 609
122C
Signal Sequence 632

Figure 6A

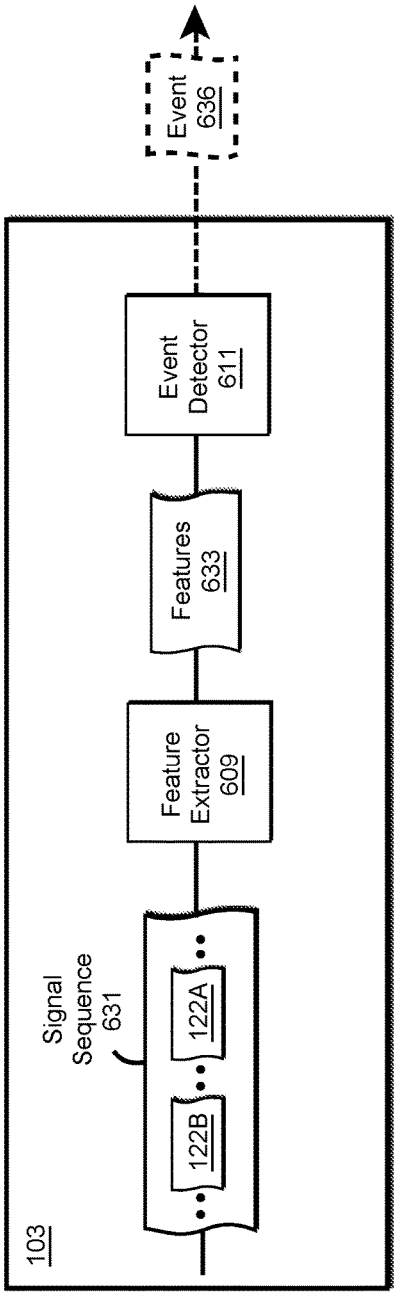


Figure 6B

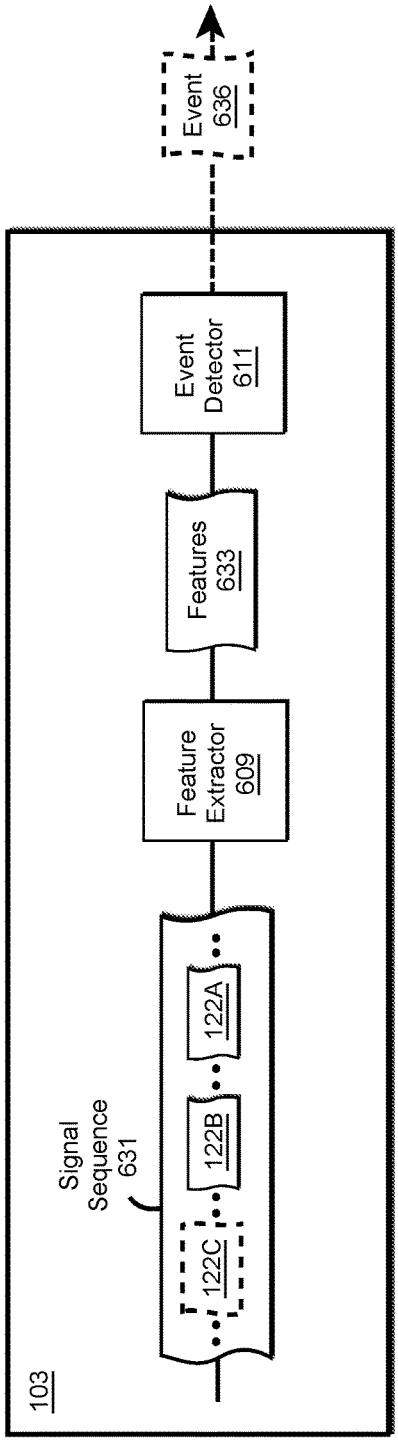


Figure 6C

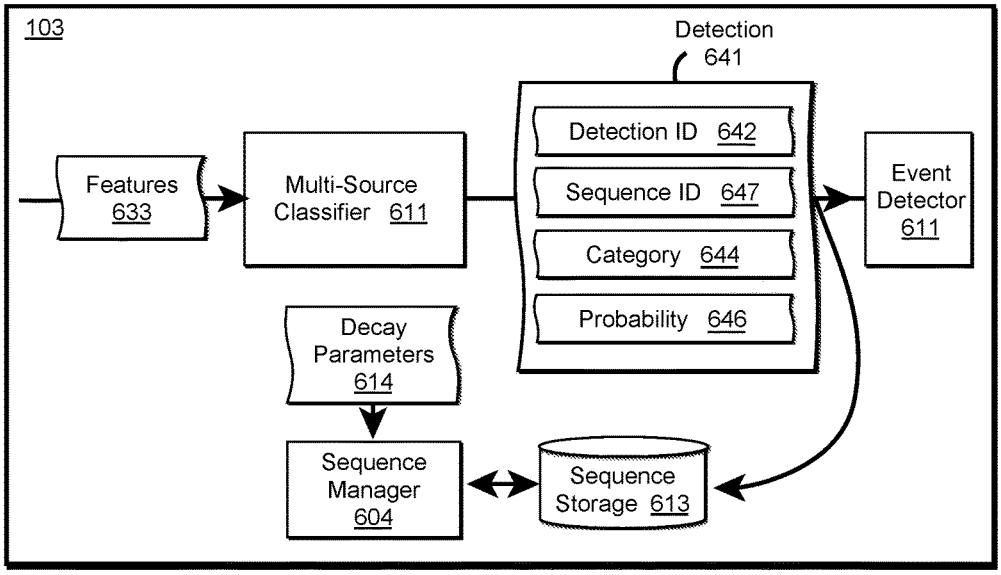


Figure 6D

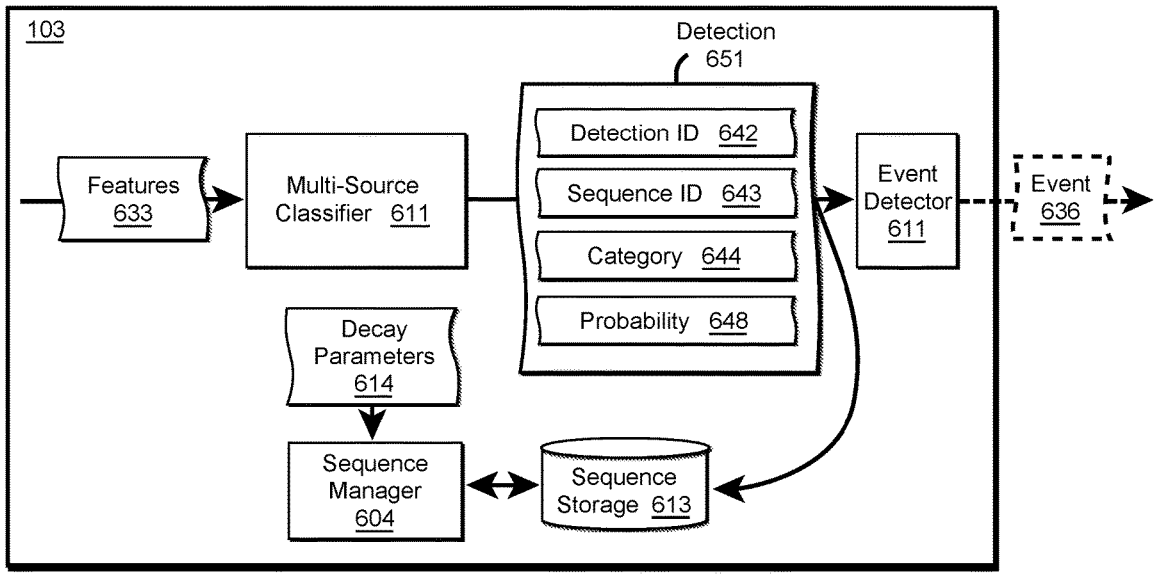


Figure 6E

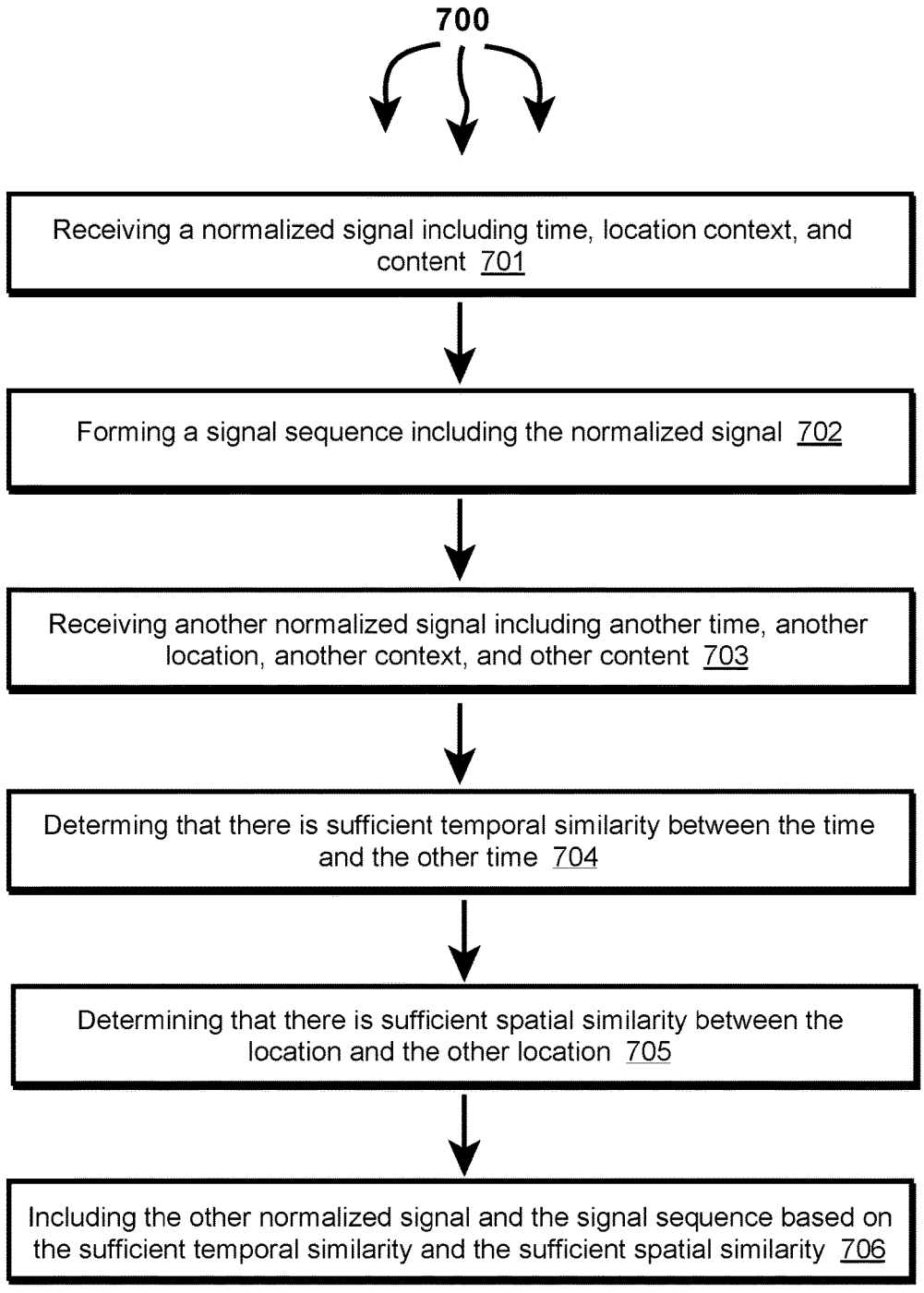

700
Receiving a normalized signal including time, location context, and content 701
Forming a signal sequence including the normalized signal 702
Receiving another normalized signal including another time, another location, another context, and other content 703
Determing that there is sufficient temporal similarity between the time and the other time 704
Determining that there is sufficient spatial similarity between the location and the other location 705
Including the other normalized signal and the signal sequence based on the sufficient temporal similarity and the sufficient spatial similarity 706

Figure 7

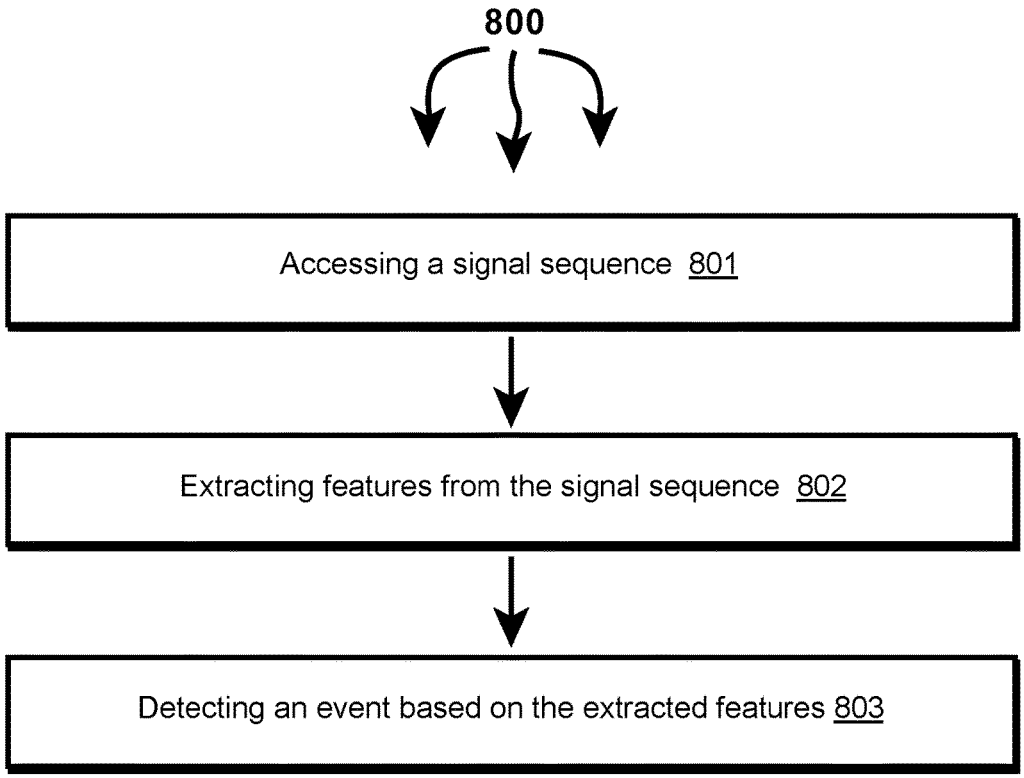

800
Accessing a signal sequence 801
Extracting features from the signal sequence 802
Detecting an event based on the extracted features 803

Figure 8

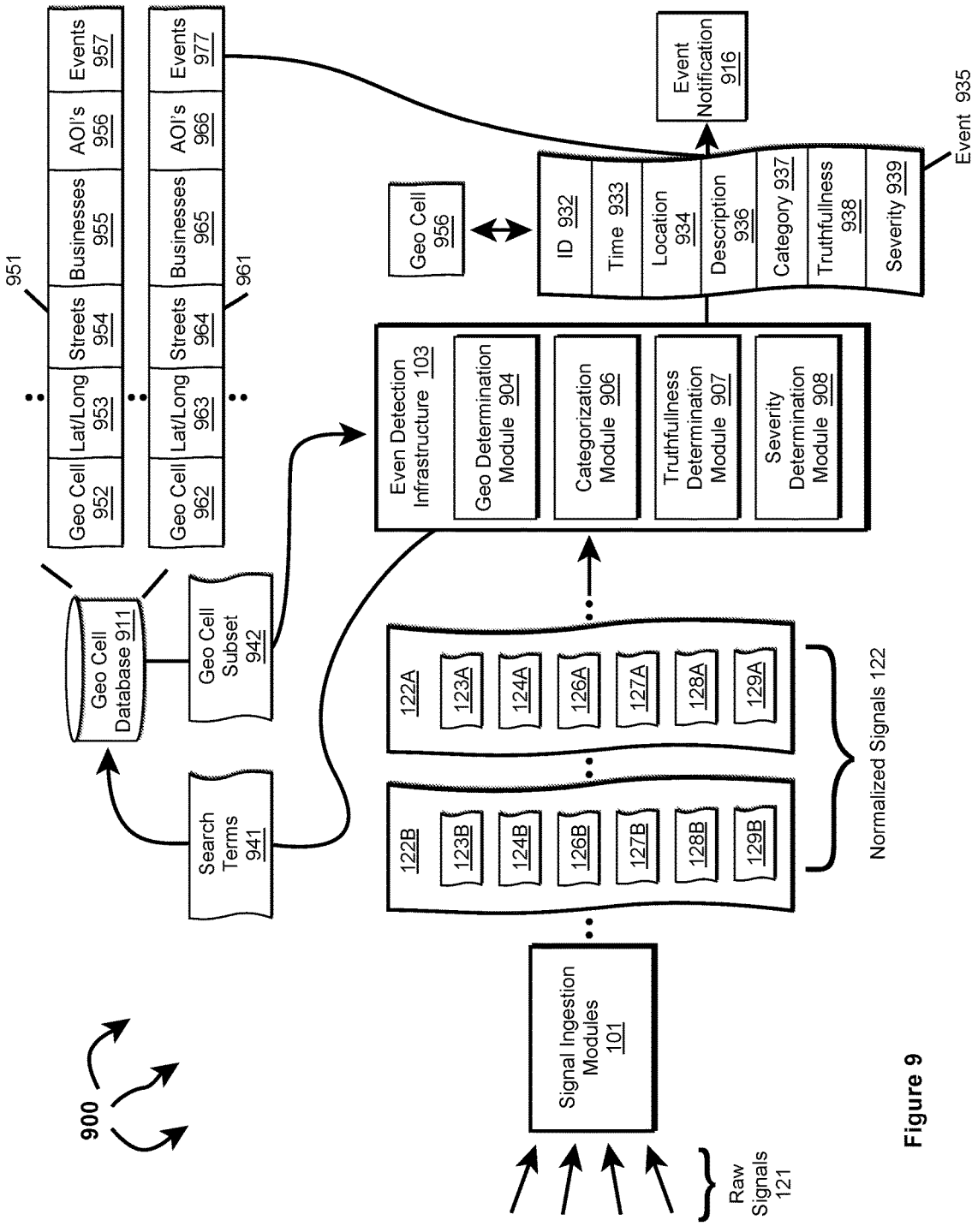

900
Raw Signals 121
Signal Ingestion Modules 101
Normalized Signals 122
122A
123A
124A
126A
127A
128A
129A
122B
123B
124B
126B
127B
128B
129B
Even Detection Infrastructure 103
Geo Determination Module 904
Categorization Module 906
Truthfullness Determination Module 907
Severity Determination Module 908
Search Terms 941
Geo Cell Database 911
Geo Cell Subset 942
951
961
Geo Cell 952
Lat/Long 953
Streets 954
Businesses 955
AOI's 956
Events 957
Geo Cell 962
Lat/Long 963
Streets 964
Businesses 965
AOI's 966
Events 977
Geo Cell 956
ID 932
Time 933
Location 934
Description 936
Category 937
Truthfullness 938
Severity 939
Event 935
Event Notification 916

Figure 9

DETECTING EVENTS FROM FEATURES DERIVED FROM MULTIPLE INGESTED SIGNALS

CROSS-REFERENCE TO RELATED APPLICATIONS

[0001] This application claims the benefit of U.S. Provisional Patent Application Ser. No. 62/628,866, entitled "Multi Source Validation", filed Feb. 9, 2018 which is incorporated herein in its entirety. This application claims the benefit of U.S. Provisional Patent Application Ser. No. 62/654,274, entitled "Detecting Events From Multiple Signals", filed Apr. 6, 2018 which is incorporated herein in its entirety. This application claims the benefit of U.S. Provisional Patent Application Ser. No. 62/654,277 entitled, "Validating Possible Events With Additional Signals", filed Apr. 6, 2018 which is incorporated herein in its entirety. This application claims the benefit of U.S. Provisional Patent Application Ser. No. 62/664,001, entitled, "Normalizing Different Types Of Ingested Signals Into A Common Format", filed Apr. 27, 2018. This application claims the benefit of U.S. Provisional Patent Application Ser. No. 62/682,176 entitled "Detecting An Event From Multiple Sources", filed Jun. 8, 2018 which is incorporated herein in its entirety. This application claims the benefit of U.S. Provisional Patent Application Ser. No. 62/682,177 entitled "Detecting An Event From Multi-Source Event Probability", filed Jun. 8, 2018 which is incorporated herein in its entirety.

BACKGROUND

1. Background and Relevant Art

[0002] Entities (e.g., parents, guardians, friends, relatives, teachers, social workers, first responders, hospitals, delivery services, media outlets, government entities, etc.) may desire to be made aware of relevant events (e.g., fires, accidents, police presence, shootings, etc.) as close as possible to the events' occurrence. However, entities typically are not made aware of an event until after a person observes the event (or the event aftermath) and calls authorities.

[0003] In general, techniques that attempt to automate event detection are unreliable. Some techniques have attempted to mine social media data to detect the planning of events and forecast when events might occur. However, events can occur without prior planning and/or may not be detectable using social media data. Further, these techniques are not capable of meaningfully processing available data nor are these techniques capable of differentiating false data (e.g., hoax social media posts)

[0004] Other techniques use textual comparisons to compare textual content (e.g., keywords) in a data stream to event templates in a database. If text in a data stream matches keywords in an event template, the data stream is labeled as indicating an event.

[0005] Additional techniques use event specific sensors to detect specified types of event. For example, earthquake detectors can be used to detect earthquakes.

BRIEF SUMMARY

[0006] Examples extend to methods, systems, and computer program products for detecting events from features derived from multiple signals.

[0007] In general, signal ingestion modules ingest different types of raw structured and unstructured signals on an ongoing basis. The signal ingestion modules normalize raw signals into normalized signals having a Time, Location, Context (or "TLC") format. Time can be a time of origin or "event time" of a signal. Location can be anywhere across a geographic area, such as, a country (e.g., the United States), a State, a defined area, an impacted area, an area defined by a geo cell, an address, etc.

[0008] Context indicates circumstances surrounding formation/origination of a raw signal in terms that facilitate understanding and assessment of the raw signal. The context of a raw signal can be derived from express as well as inferred signal features of the raw signal.

[0009] Signal ingestion modules can include one or more single source classifiers. A single source classifier can compute a single source probability for a raw signal from (inferred and/or express) signal features of the raw signal. A single source probability can reflect a mathematical probability or approximation of a mathematical probability of an event (e.g., fire, accident, weather, police presence, etc.) actually occurring. A single source classifier can be configured to compute a single source probability for a single event type or to compute a single source probability for each of a plurality of different event types.

[0010] As such, single source probabilities and corresponding probability details can represent Context. Probability details can indicate (e.g., can include a hash field indicating) a probability version and (express and/or inferred) signal features considered in a signal source probability calculation.

[0011] Concurrently with signal ingestion, an event detection infrastructure considers features of different combinations of normalized signals to attempt to identify events of interest to various parties. For example, the event detection infrastructure can determine that features of multiple different signals collectively indicate an event of interest to one or more parties. Alternately, the event detection infrastructure can determine that features of one or more signals indicate a possible event of interest to one or more parties. The event detection infrastructure then determines that features of one or more other signals validate the possible event as an actual event of interest to the one or more parties. Signal features can include: signal type, signal source, signal content, signal time (T), signal location (L), signal context (C), other circumstances of signal creation, etc.

[0012] The event detection infrastructure can group signals having sufficient temporal similarity and sufficient spatial similarity to one another in a signal sequence. In one aspect, any signal having sufficient temporal and spatial similarity to another signal can be added to a signal sequence.

[0013] In another aspect, a single source probability for a signal is computed from features of the signal. The single source probability can reflect a mathematical probability or approximation of a mathematical probability of an event actually occurring. A signal having a signal source probability above a threshold can be indicated as an "elevated" signal. Elevated signals can be used to initiate and/or can be added to a signal sequence. On the other hand, non-elevated signals may not be added to a signal sequence.

[0014] A multi-source probability can be computed from features of multiple normalized signals, including normalized signals in a signal sequence. Features used to compute

a multi-source probability can include multiple single source probabilities as well as other features derived from multiple signals. The multi-source probability can reflect a mathematical probability or approximation of a mathematical probability of an event actually occurring based on multiple normalized signals (e.g., a signal sequence). A multi-source probability can change over time as normalized signals age or when a new normalized signal is received (e.g., added to a signal sequence).

[0015] This summary is provided to introduce a selection of concepts in a simplified form that are further described below in the Detailed Description. This Summary is not intended to identify key features or essential features of the claimed subject matter, nor is it intended to be used as an aid in determining the scope of the claimed subject matter.

[0016] Additional features and advantages will be set forth in the description which follows, and in part will be obvious from the description, or may be learned by practice. The features and advantages may be realized and obtained by means of the instruments and combinations particularly pointed out in the appended claims. These and other features and advantages will become more fully apparent from the following description and appended claims, or may be learned by practice as set forth hereinafter.

BRIEF DESCRIPTION OF THE DRAWINGS

[0017] In order to describe the manner in which the above-recited and other advantages and features can be obtained, a more particular description will be rendered by reference to specific implementations thereof which are illustrated in the appended drawings. Understanding that these drawings depict only some implementations and are not therefore to be considered to be limiting of its scope, implementations will be described and explained with additional specificity and detail through the use of the accompanying drawings in which:

[0018] FIG. **1** illustrates an example computer architecture that facilitates ingesting signals.

[0019] FIG. **2** illustrates an example computer architecture that facilitates detecting an event from features derived from multiple signals.

[0020] FIG. **3** illustrates a flow chart of an example method for detecting an event from features derived from multiple signals.

[0021] FIG. **4** illustrates an example computer architecture that facilitates detecting an event from features derived from multiple signals.

[0022] FIG. **5** illustrates a flow chart of an example method for detecting an event from features derived from multiple signals

[0023] FIG. **6**A illustrates an example computer architecture that facilitates forming a signal sequence.

[0024] FIG. **6**B illustrates an example computer architecture that facilitates detecting an event from features of a signal sequence.

[0025] FIG. **6**C illustrates an example computer architecture that facilitates detecting an event from features of a signal sequence.

[0026] FIG. **6**D illustrates an example computer architecture that facilitates detecting an event from a multisource probability.

[0027] FIG. **6**E illustrates an example computer architecture that facilitates detecting an event from a multisource probability.

[0028] FIG. **7** illustrates a flow chart of an example method for forming a signal sequence.

[0029] FIG. **8** illustrates a flow of an example method for detecting an event from a signal sequence.

[0030] FIG. **9** illustrates an example computer architecture that facilitates detecting events.

DETAILED DESCRIPTION

[0031] Examples extend to methods, systems, and computer program products for detecting events from features derived from multiple signals.

[0032] Aspects of the invention normalize raw signals into a common format that includes Time, Location, and Context (or "TLC") format. Per signal type, signal ingestion modules identify and/or infer a time, a location, and a context associated with a signal. Different ingestion modules can be utilized/tailored to identify time, location, and context for different signal types. Time (T) can be a time of origin or "event time" of a signal. Location (L) can be anywhere across a geographic area, such as, a country (e.g., the United States), a State, a defined area, an impacted area, an area defined by a geo cell, an address, etc.

[0033] Context (C) indicates circumstances surrounding formation/origination of a raw signal in terms that facilitate understanding and assessment of the raw signal. The context of a raw signal can be derived from express as well as inferred signal features of the raw signal.

[0034] Signal ingestion modules can include one or more single source classifiers. A single source classifier can compute a single source probability for a raw signal from features of the raw signal. A single source probability can reflect a mathematical probability or approximation of a mathematical probability (e.g., a percentage between 0%-100%) of an event actually occurring. A single source classifier can be configured to compute a single source probability for a single event type or to compute a single source probability for each of a plurality of different event types. A single source classifier can compute a single source probability using artificial intelligence, machine learning, neural networks, logic, heuristics, etc.

[0035] As such, single source probabilities and corresponding probability details can represent Context. Probability details can indicate (e.g., can include a hash field indicating) a probability version and (express and/or inferred) signal features considered in a signal source probability calculation.

[0036] Concurrently with signal ingestion, the event detection infrastructure considers features of different combinations of normalized signals to attempt to identify events of interest to various parties. Features can be derived from an individual signal and/or from a group of signals.

[0037] For example, the event detection infrastructure can derive first features of a first normalized signal and can derive second features of a second normalized signal. Individual signal features can include: signal type, signal source, signal content, signal time (T), signal location (L), signal context (C), other circumstances of signal creation, etc. The event detection infrastructure can detect an event of interest to one or more parties from the first features and the second features collectively.

[0038] Alternately, the event detection infrastructure can derive first features of each normalized signal included in a first one or more normalized individual signals. The event detection infrastructure can detect a possible event of inter-

est to one or more parties from the first features. The event detection infrastructure can derive second features of each normalized signal included in a second one or more individual signals. The event detection infrastructure can validate the possible event of interest as an actual event of interest to the one or more parties from the second features.

[0039] More specifically, the event detection infrastructure can use single source probabilities to detect and/or validate events. For example, the event detection infrastructure can detect an event of interest to one or more parties based on a single source probability of a first signal and a single source probability of second signal collectively. Alternately, the event detection infrastructure can detect a possible event of interest to one or more parties based on single source probabilities of a first one or more signals. The event detection infrastructure can validate the possible event as an actual event of interest to one or more parties based on single source probabilities of a second one or more signals.

[0040] The event detection infrastructure can group normalized signals having sufficient temporal similarity and/or sufficient spatial similarity to one another in a signal sequence. Temporal similarity of normalized signals can be determined by comparing Time (T) of the normalized signals. In one aspect, temporal similarity of a normalized signal and another normalized signal is sufficient when the Time (T) of the normalized signal is within a specified time of the Time (T) of the other normalized signal. A specified time can be virtually any time value, such as, for example, ten seconds, 30 seconds, one minute, two minutes, five minutes, ten minutes, 30 minutes, one hour, two hours, four hours, etc. A specified time can vary by detection type. For example, some event types (e.g., a fire) inherently last longer than other types of events (e.g., a shooting). Specified times can be tailored per detection type.

[0041] Spatial similarity of normalized signals can be determined by comparing Location (L) of the normalized signals. In one aspect, spatial similarity of a normalized signal and another normalized signal is sufficient when the Location (L) of the normalized signal is within a specified distance of the Location (L) of the other normalized signal. A specified distance can be virtually any distance value, such as, for example, a linear distance or radius (a number of feet, meters, miles, kilometers, etc.), within a specified number of geo cells of specified precision, etc.

[0042] In one aspect, any normalized signal having sufficient temporal and spatial similarity to another normalized signal can be added to a signal sequence.

[0043] In another aspect, a single source probability for a signal is computed from features of the signal. The single source probability can reflect a mathematical probability or approximation of a mathematical probability of an event actually occurring. A normalized signal having a signal source probability above a threshold (e.g., greater than 4%) is indicated as an "elevated" signal. Elevated signals can be used to initiate and/or can be added to a signal sequence. On the other hand, non-elevated signals may not be added to a signal sequence.

[0044] In one aspect, a first threshold is considered for signal sequence initiation and a second threshold is considered for adding additional signals to an existing signal sequence. A normalized signal having a single source probability above the first threshold can be used to initiate a signal sequence. After a signal sequence is initiated, any normalized signal having a single source probability above the second threshold can be added to the signal sequence.

[0045] The first threshold can be greater than the second threshold. For example, the first threshold can be 4% or 5% and the second threshold can be 2% or 3%. Thus, signals that are not necessarily reliable enough to initiate a signal sequence for an event can be considered for validating a possible event.

[0046] The event detection infrastructure can derive features of a signal grouping, such as, a signal sequence. Features of a signal sequence can include features of signals in the signal sequence, including single source probabilities. Features of a signal sequence can also include percentages, histograms, counts, durations, etc. derived from features of the signals included in the signal sequence. The event detection infrastructure can detect an event of interest to one or more parties from signal sequence features.

[0047] The event detection infrastructure can include one or more multi-source classifiers. A multi-source classifier can compute a multi-source probability for a signal sequence from features of the signal sequence. The multi-source probability can reflect a mathematical probability or approximation of a mathematical probability of an event (e.g., fire, accident, weather, police presence, etc.) actually occurring based on multiple normalized signals (e.g., the signal sequence). The multi-source probability can be assigned as an additional signal sequence feature. A multi-source classifier can be configured to compute a multi-source probability for a single event type or to compute a multi-source probability for each of a plurality of different event types. A multi-source classifier can compute a multi-source probability using artificial intelligence, machine learning, neural networks, etc.

[0048] A multi-source probability can change over time as a signal sequence ages or when a new signal is added to a signal sequence. For example, a multi-source probability for a signal sequence can decay over time. A multi-source probability for a signal sequence can also be recomputed when a new normalized signal is added to the signal sequence.

[0049] Multi-source probability decay can start after a specified period of time (e.g., 3 minutes) and decay can occur in accordance with a defined decay equation. In one aspect, a decay equation defines exponential decay of multi-source probabilities. Different decay rates can be used for different classes. Decay can be similar to radioactive decay, with different tau (i.e., mean lifetime) values used to calculate the "half life" of multi-source probability for different event types.

[0050] Implementations can comprise or utilize a special purpose or general-purpose computer including computer hardware, such as, for example, one or more computer and/or hardware processors (including any of Central Processing Units (CPUs), and/or Graphical Processing Units (GPUs), general-purpose GPUs (GPGPUs), Field Programmable Gate Arrays (FPGAs), application specific integrated circuits (ASICs), Tensor Processing Units (TPUs)) and system memory, as discussed in greater detail below. Implementations also include physical and other computer-readable media for carrying or storing computer-executable instructions and/or data structures. Such computer-readable media can be any available media that can be accessed by a general purpose or special purpose computer system. Computer-readable media that store computer-executable

instructions are computer storage media (devices). Computer-readable media that carry computer-executable instructions are transmission media. Thus, by way of example, and not limitation, implementations can comprise at least two distinctly different kinds of computer-readable media: computer storage media (devices) and transmission media.

[0051] Computer storage media (devices) includes RAM, ROM, EEPROM, CD-ROM, Solid State Drives ("SSDs") (e.g., RAM-based or Flash-based), Shingled Magnetic Recording ("SMR") devices, Flash memory, phase-change memory ("PCM"), other types of memory, other optical disk storage, magnetic disk storage or other magnetic storage devices, or any other medium which can be used to store desired program code means in the form of computer-executable instructions or data structures and which can be accessed by a general purpose or special purpose computer.

[0052] In one aspect, one or more processors are configured to execute instructions (e.g., computer-readable instructions, computer-executable instructions, etc.) to perform any of a plurality of described operations. The one or more processors can access information from system memory and/or store information in system memory. The one or more processors can (e.g., automatically) transform information between different formats, such as, for example, between any of: raw signals, normalized signals, signal features, single source probabilities, possible events, events, signal sequences, signal sequence features, multisource probabilities, etc.

[0053] System memory can be coupled to the one or more processors and can store instructions (e.g., computer-readable instructions, computer-executable instructions, etc.) executed by the one or more processors. The system memory can also be configured to store any of a plurality of other types of data generated and/or transformed by the described components, such as, for example, raw signals, normalized signals, signal features, single source probabilities, possible events, events, signal sequences, signal sequence features, multisource probabilities, etc.

[0054] A "network" is defined as one or more data links that enable the transport of electronic data between computer systems and/or modules and/or other electronic devices. When information is transferred or provided over a network or another communications connection (either hardwired, wireless, or a combination of hardwired or wireless) to a computer, the computer properly views the connection as a transmission medium. Transmissions media can include a network and/or data links which can be used to carry desired program code means in the form of computer-executable instructions or data structures and which can be accessed by a general purpose or special purpose computer. Combinations of the above should also be included within the scope of computer-readable media.

[0055] Further, upon reaching various computer system components, program code means in the form of computer-executable instructions or data structures can be transferred automatically from transmission media to computer storage media (devices) (or vice versa). For example, computer-executable instructions or data structures received over a network or data link can be buffered in RAM within a network interface module (e.g., a "NIC"), and then eventually transferred to computer system RAM and/or to less volatile computer storage media (devices) at a computer system. Thus, it should be understood that computer storage media (devices) can be included in computer system components that also (or even primarily) utilize transmission media.

[0056] Computer-executable instructions comprise, for example, instructions and data which, in response to execution at a processor, cause a general purpose computer, special purpose computer, or special purpose processing device to perform a certain function or group of functions. The computer executable instructions may be, for example, binaries, intermediate format instructions such as assembly language, or even source code. Although the subject matter has been described in language specific to structural features and/or methodological acts, it is to be understood that the subject matter defined in the appended claims is not necessarily limited to the described features or acts described above. Rather, the described features and acts are disclosed as example forms of implementing the claims.

[0057] Those skilled in the art will appreciate that the described aspects may be practiced in network computing environments with many types of computer system configurations, including, personal computers, desktop computers, laptop computers, message processors, hand-held devices, wearable devices, multicore processor systems, multi-processor systems, microprocessor-based or programmable consumer electronics, network PCs, minicomputers, mainframe computers, mobile telephones, PDAs, tablets, routers, switches, and the like. The described aspects may also be practiced in distributed system environments where local and remote computer systems, which are linked (either by hardwired data links, wireless data links, or by a combination of hardwired and wireless data links) through a network, both perform tasks. In a distributed system environment, program modules may be located in both local and remote memory storage devices.

[0058] Further, where appropriate, functions described herein can be performed in one or more of: hardware, software, firmware, digital components, or analog components. For example, one or more Field Programmable Gate Arrays (FPGAs) and/or one or more application specific integrated circuits (ASICs) and/or one or more Tensor Processing Units (TPUs) can be programmed to carry out one or more of the systems and procedures described herein. Hardware, software, firmware, digital components, or analog components can be specifically tailor-designed for a higher speed detection or artificial intelligence that can enable signal processing. In another example, computer code is configured for execution in one or more processors, and may include hardware logic/electrical circuitry controlled by the computer code. These example devices are provided herein purposes of illustration, and are not intended to be limiting. Embodiments of the present disclosure may be implemented in further types of devices.

[0059] The described aspects can also be implemented in cloud computing environments. In this description and the following claims, "cloud computing" is defined as a model for enabling on-demand network access to a shared pool of configurable computing resources. For example, cloud computing can be employed in the marketplace to offer ubiquitous and convenient on-demand access to the shared pool of configurable computing resources (e.g., compute resources, networking resources, and storage resources). The shared pool of configurable computing resources can be provisioned via virtualization and released with low effort or service provider interaction, and then scaled accordingly.

[0060] A cloud computing model can be composed of various characteristics such as, for example, on-demand self-service, broad network access, resource pooling, rapid elasticity, measured service, and so forth. A cloud computing model can also expose various service models, such as, for example, Software as a Service ("SaaS"), Platform as a Service ("PaaS"), and Infrastructure as a Service ("IaaS"). A cloud computing model can also be deployed using different deployment models such as private cloud, community cloud, public cloud, hybrid cloud, and so forth. In this description and in the following claims, a "cloud computing environment" is an environment in which cloud computing is employed.

[0061] In this description and the following claims, a "geo cell" is defined as a piece of "cell" in a grid in any form. In one aspect, geo cells are arranged in a hierarchical structure. Cells of different geometries can be used.

[0062] A "geohash" is an example of a "geo cell".

[0063] In this description and the following claims, "geohash" is defined as a geocoding system which encodes a geographic location into a short string of letters and digits. Geohash is a hierarchical spatial data structure which subdivides space into buckets of grid shape (e.g., a square). Geohashes offer properties like arbitrary precision and the possibility of gradually removing characters from the end of the code to reduce its size (and gradually lose precision). As a consequence of the gradual precision degradation, nearby places will often (but not always) present similar prefixes. The longer a shared prefix is, the closer the two places are. geo cells can be used as a unique identifier and to represent point data (e.g., in databases).

[0064] In one aspect, a "geohash" is used to refer to a string encoding of an area or point on the Earth. The area or point on the Earth may be represented (among other possible coordinate systems) as a latitude/longitude or Easting/Northing—the choice of which is dependent on the coordinate system chosen to represent an area or point on the Earth. geo cell can refer to an encoding of this area or point, where the geo cell may be a binary string comprised of 0 s and 1 s corresponding to the area or point, or a string comprised of 0 s, 1 s, and a ternary character (such as X)—which is used to refer to a don't care character (0 or 1). A geo cell can also be represented as a string encoding of the area or point, for example, one possible encoding is base-32, where every 5 binary characters are encoded as an ASCII character.

[0065] Depending on latitude, the size of an area defined at a specified geo cell precision can vary. In one aspect, the areas defined at various geo cell precisions are approximately:

GeoHash Length/Precision	Width × Height
1	5,009.4 km × 4,992.6 km
2	1,252.3 km × 624.1 km
3	156.5 km × 156 km
4	39.1 km × 19.5 km
5	4.9 km × 4.9 km
6	1.2 km × 609.4 m
7	152.9 m × 152.4 m
8	38.2 m × 19 m
9	4.8 m × 4.8 m
10	1.2 m × 59.5 cm
11	14.9 cm × 14.9 cm
12	3.7 cm × 1.9 cm

[0066] Other geo cell geometries, such as, hexagonal tiling, triangular tiling, etc. are also possible. For example, the H3 geospatial indexing system is a multi-precision hexagonal tiling of a sphere (such as the Earth) indexed with hierarchical linear indexes.

[0067] In another aspect, geo cells are a hierarchical decomposition of a sphere (such as the Earth) into representations of regions or points based a Hilbert curve (e.g., the S2 hierarchy or other hierarchies). Regions/points of the sphere can be projected into a cube and each face of the cube includes a quad-tree where the sphere point is projected into. After that, transformations can be applied and the space discretized. The geo cells are then enumerated on a Hilbert Curve (a space-filling curve that converts multiple dimensions into one dimension and preserves the locality).

[0068] Due to the hierarchical nature of geo cells, any signal, event, entity, etc., associated with a geo cell of a specified precision is by default associated with any less precise geo cells that contain the geo cell. For example, if a signal is associated with a geo cell of precision 9, the signal is by default also associated with corresponding geo cells of precisions 1, 2, 3, 4, 5, 6, 7, and 8. Similar mechanisms are applicable to other tiling and geo cell arrangements. For example, S2 has a cell level hierarchy ranging from level zero (85,011,012 km^2) to level 30 (between 0.48 cm^2 to 0.96 cm^2).

[0069] Signal Ingestion and Normalization

[0070] Signal ingestion modules ingest a variety of raw structured and/or unstructured signals on an on going basis and in essentially real-time. Raw signals can include social posts, live broadcasts, traffic camera feeds, other camera feeds (e.g., from other public cameras or from CCTV cameras), listening device feeds, 911 calls, weather data, planned events, IoT device data, crowd sourced traffic and road information, satellite data, air quality sensor data, smart city sensor data, public radio communication (e.g., among first responders and/or dispatchers, between air traffic controllers and pilots), etc. The content of raw signals can include images, video, audio, text, etc. Generally, the signal ingestion modules normalize raw signals into normalized signals, for example, having a Time, Location, Context (or "TLC") format.

[0071] Different types of ingested signals (e.g., social media signals, web signals, and streaming signals) can be used to identify events. Different types of signals can include different data types and different data formats. Data types can include audio, video, image, and text. Different formats can include text in XML, text in JavaScript Object Notation (JSON), text in RSS feed, plain text, video stream in Dynamic Adaptive Streaming over HTTP (DASH), video stream in HTTP Live Streaming (HLS), video stream in Real-Time Messaging Protocol (RTMP), etc.

[0072] Time (T) can be a time of origin or "event time" of a signal. In one aspect, a raw signal includes a time stamp and the time stamp is used to calculate Time (T). Location (L) can be anywhere across a geographic area, such as, a country (e.g., the United States), a State, a defined area, an impacted area, an area defined by a geo cell, an address, etc.

[0073] Context indicates circumstances surrounding formation/origination of a raw signal in terms that facilitate understanding and assessment of the raw signal. The context of a raw signal can be derived from express as well as inferred signal features of the raw signal.

[0074] Signal ingestion modules can include one or more single source classifiers. A single source classifier can compute a single source probability for a raw signal from features of the raw signal. A single source probability can reflect a mathematical probability or approximation of a mathematical probability (e.g., a percentage between 0%-100%) of an event (e.g., fire, accident, weather, police presence, shooting, etc.) actually occurring. A single source classifier can be configured to compute a single source probability for a single event type or to compute a single source probability for each of a plurality of different event types. A single source classifier can compute a single source probability using artificial intelligence, machine learning, neural networks, logic, heuristics, etc.

[0075] As such, single source probabilities and corresponding probability details can represent Context (C). Probability details can indicate (e.g., can include a hash field indicating) a probability version and (express and/or inferred) signal features considered in a signal source probability calculation.

[0076] Per signal type and signal content, different normalization modules can be used to extract, derive, infer, etc. time, location, and context from/for a raw signal. For example, one set of normalization modules can be configured to extract/derive/infer time, location and context from/for social signals. Another set of normalization modules can be configured to extract/derive/infer time, location and context from/for Web signals. A further set of normalization modules can be configured to extract/derive/infer time, location and context from/for streaming signals.

[0077] Normalization modules for extracting/deriving/inferring time, location, and context can include text processing modules, NLP modules, image processing modules, video processing modules, etc. The modules can be used to extract/derive/infer data representative of time, location, and context for a signal. Time, Location, and Context for a signal can be extracted/derived/inferred from metadata and/or content of the signal. For example, NLP modules can analyze metadata and content of a sound clip to identify a time, location, and keywords (e.g., fire, shooter, etc.). An acoustic listener can also interpret the meaning of sounds in a sound clip (e.g., a gunshot, vehicle collision, etc.) and convert to relevant context. Live acoustic listeners can determine the distance and direction of a sound. Similarly, image processing modules can analyze metadata and pixels in an image to identify a time, location and keywords (e.g., fire, shooter, etc.). Image processing modules can also interpret the meaning of parts of an image (e.g., a person holding a gun, flames, a store logo, etc.) and convert to relevant context. Other modules can perform similar operations for other types of content including text and video.

[0078] Per signal type, each set of normalization modules can differ but may include at least some similar modules or may share some common modules. For example, similar (or the same) image analysis modules can be used to extract named entities from social signal images and public camera feeds. Likewise, similar (or the same) NLP modules can be used to extract named entities from social signal text and web text.

[0079] In some aspects, an ingested signal includes expressly defined Time, Location, and Context upon ingestion. In other aspects, an ingested signal lacks an expressly defined Location and/or an expressly defined Context upon ingestion. In these other aspects, Location and/or Context can be inferred from features of an ingested signal and/or through reference to other data sources.

[0080] In further aspects, Time may not be included, or an included time may not be given with high precision and is inferred. For example, a user may post an image to a social network which had been taken some indeterminate time earlier.

[0081] Normalization modules can use named entity recognition and reference to a geo cell database to infer location. Named entities can be recognized in text, images, video, audio, or sensor data. The recognized named entities can be compared to named entities in geo cell entries. Matches indicate possible signal origination in a geographic area defined by a geo cell.

[0082] As such, a normalized signal can include a Time, a Location, a Context (e.g., single source probabilities and probability details), a signal type, a signal source, and content.

[0083] In one aspect, frequentist inference technique is used to determine a single source probability. A database maintains mappings between different combinations of signal properties and ratios of signals turning into events (a probability) for that combination of signal properties. The database is queried with the combination of signal properties. The database returns a ratio of signals having the signal properties turning into events. The ratio is assigned to the signal. A combination of signal properties can include: (1) event class (e.g., fire, accident, weather, etc.), (2) media type (e.g., text, image, audio, etc.), (3) source (e.g., twitter, traffic camera, first responder radio traffic, etc.), and (4) geo type (e.g., geo cell, region, or non-geo).

[0084] In another aspect, a single source probability is calculated by single source classifiers (e.g., machine learning models, artificial intelligence, neural networks, etc.) that consider hundreds, thousands, or even more signal features of a signal. Single source classifiers can be based on binary models and/or multi-class models.

[0085] Output from a single source classifier can be adjusted to more accurately represent a probability that a signal is a "true positive". For example, 1,000 signals with classifier output of 0.9 may include 80% as true positives. Thus, single source probability can be adjusted to 0.8 to more accurately reflect probability of the signal being a True event. "Calibration" can be done in such a way that for any "calibrated score" the score reflects the true probability of a true positive outcome.

[0086] FIG. **1** depicts computer architecture **100** that facilitates ingesting and normalizing signals. As depicted, computer architecture **100** includes signal ingestion modules **101**, social signals **171**, Web signals **172**, and streaming signals **173**. Signal ingestion modules **101**, social signals **171**, Web signals **172**, and streaming signals **173** can be connected to (or be part of) a network, such as, for example, a system bus, a Local Area Network ("LAN"), a Wide Area Network ("WAN"), and even the Internet. Accordingly, signal ingestion modules **101**, social signals **171**, Web signals **172**, and streaming signals **173** as well as any other connected computer systems and their components can create and exchange message related data (e.g., Internet Protocol ("IP") datagrams and other higher layer protocols that utilize IP datagrams, such as, Transmission Control Protocol ("TCP"), Hypertext Transfer Protocol ("HTTP"),

Simple Mail Transfer Protocol ("SMTP"), Simple Object Access Protocol (SOAP), etc. or using other non-datagram protocols) over the network.

[0087] Signal ingestion module(s) **101** can ingest raw signals **121**, including social signals **171**, web signals **172**, and streaming signals **173** (e.g., social posts, traffic camera feeds, other camera feeds, listening device feeds, 911 calls, weather data, planned events, IoT device data, crowd sourced traffic and road information, satellite data, air quality sensor data, smart city sensor data, public radio communication, etc.) on going basis and in essentially real-time. Signal ingestion module(s) **101** include social content ingestion modules **174**, web content ingestion modules **176**, stream content ingestion modules **177**, and signal formatter **180**. Signal formatter **180** further includes social signal processing module **181**, web signal processing module **182**, and stream signal processing modules **183**.

[0088] For each type of signal, a corresponding ingestion module and signal processing module can interoperate to normalize the signal into a Time, Location, Context (TLC) format. For example, social content ingestion modules **174** and social signal processing module **181** can interoperate to normalize social signals **171** into the TLC format. Similarly, web content ingestion modules **176** and web signal processing module **182** can interoperate to normalize web signals **172** into the TLC format. Likewise, stream content ingestion modules **177** and stream signal processing modules **183** can interoperate to normalize streaming signals **173** into the TLC format.

[0089] In one aspect, signal content exceeding specified size requirements (e.g., audio or video) is cached upon ingestion. Signal ingestion modules **101** include a URL or other identifier to the cached content within the context for the signal.

[0090] Signal formatter **180** can include one or more single signal classifiers classifying ingested signals. The one or more single signal classifiers can assign one or more signal source probabilities (e.g., between 0%-100%) to each ingested signal. Each single source probability is a probability of the ingested signal being a particular category of event (e.g., fire, weather, medical, accident, police presence, etc.). Ingested signals with a sufficient single source probability (e.g., >=to 4%) are considered "elevated" signals.

[0091] In one aspect, signal formatter **180** includes modules for determining a single source probability as a ratio of signals turning into events based on the following signal properties: (1) event class (e.g., fire, accident, weather, etc.), (2) media type (e.g., text, image, audio, etc.), (3) source (e.g., twitter, traffic camera, first responder radio traffic, etc.), and (4) geo type (e.g., geo cell, region, or non-geo). Probabilities can be stored in a lookup table for different combinations of the signal properties. Features of a signal can be derived and used to query the lookup table. For example, the lookup table can be queried with terms ("accident", "image", "twitter", "region"). The corresponding ratio (probability) can be returned from the table.

[0092] In another aspect, signal formatter **180** includes a plurality of single source classifiers (e.g., artificial intelligence, machine learning modules, neural networks, etc.). Each single source classifier can consider hundreds, thousands, or even more signal features of a signal. Signal features of a signal can be derived and submitted to a signal source classifier. The single source classifier can return a probability that a signal indicates a type of event. Single source classifiers can be binary classifiers or multi-source classifiers.

[0093] Raw classifier output can be adjusted to more accurately represent a probability that a signal is a "true positive". For example, 1,000 signals whose raw classifier output is 0.9 may include 80% as true positives. Thus, probability can be adjusted to 0.8 to reflect true probability of the signal being a true positive. "Calibration" can be done in such a way that for any "calibrated score" this score reflects the true probability of a true positive outcome.

[0094] Signal ingestion modules **101** can include one or more single source probabilities and corresponding probability details in the context of a normalized signal. Probability details can indicate a probability version and features used to calculate the probability. In one aspect, a probability version and signal feature are contained in a hash field.

[0095] Thus in general, any of the received raw signals can be normalized into normalized signals including Time, Location, Context, signal source, signal type, and content. Signal ingestion modules **101** can send normalized signals **122** to event detection infrastructure **103**. For example, signal ingestion modules **101** can send normalized signal **122**A, including time **123**A, location **124**A, context **126**A, content **127**A, type **128**A, and source **129**A to event detection infrastructure **103**. Similarly, signal ingestion modules **101** can send normalized signal **122**B, including time **123**B, location **124**B, context **126**B, content **127**B, type **128**B, and source **129**B to event detection infrastructure **103**. Signal ingestion modules **101** can also send normalized signal **122**C (depicted in FIG. **6**), including time **123**C, location **124**C, context **126**C, content **127**C , type **128**C, and source **129**C to event detection infrastructure **103**.

[0096] Multi-Signal Detection

[0097] FIG. **2** illustrates an example computer architecture **200** that facilitates detecting an event from features derived from multiple signals. As depicted, computer architecture **200** further includes event detection infrastructure **103**. Event infrastructure **103** can be connected to (or be part of) a network with signal ingestion modules **101**. As such, signal ingestion modules **101** and event detection infrastructure **103** can create and exchange message related data over the network.

[0098] As depicted, event detection infrastructure **103** further includes evaluation module **206**. Evaluation module **206** is configured to determine if features of a plurality of normalized signals collectively indicate an event. Evaluation module **206** can detect (or not detect) an event based on one or more features of one normalized signal in combination with one or more features of another normalized signal.

[0099] FIG. **3** illustrates a flow chart of an example method **300** for detecting an event from features derived from multiple signals. Method **300** will be described with respect to the components and data in computer architecture **200**.

[0100] Method **300** includes receiving a first signal (**301**). For example, event detection infrastructure **103** can receive normalized signal **122**B. Method **300** includes deriving first one or more features of the first signal (**302**). For example, event detection infrastructure **103** can derive features **201** of normalized signal **122**B. Features **201** can include and/or be derived from time **123**B, location **124**B, context **126**B, content **127**B, type **128**B, and source **129**B. Event detection

infrastructure **103** can also derive features **201** from one or more single source probabilities assigned to normalized signal **122**B.

[0101] Method **300** includes determining that the first one or more features do not satisfy conditions to be identified as an event (**303**). For example, evaluation module **206** can determine that features **201** do not satisfy conditions to be identified as an event. That is, the one or more features of normalized signal **122**B do not alone provide sufficient evidence of an event. In one aspect, one or more single source probabilities assigned to normalized signal **122**B do not satisfy probability thresholds in thresholds **226**.

[0102] Method **300** includes receiving a second signal (**304**). For example, event detection infrastructure **103** can receive normalized signal **122**A. Method **300** includes deriving second one or more features of the second signal (**305**). For example, event detection infrastructure **103** can derive features **202** of normalized signal **122**A. Features **202** can include and/or be derived from time **123**A, location **124**A, context **126**A, content **127**A, type **128**A, and source **129**A. Event detection infrastructure **103** can also derive features **202** from one or more single source probabilities assigned to normalized signal **122**A.

[0103] Method **300** includes aggregating the first one or more features with the second one or more features into aggregated features (**306**). For example, evaluation module **206** can aggregate features **201** with features **202** into aggregated features **203**. Evaluation module **206** can include an algorithm that defines and aggregates individual contributions of different signal features into aggregated features. Aggregating features **201** and **202** can include aggregating a single source probability assigned to normalized signal **122**B for an event type with a signal source probability assigned to normalized signal **122**A for the event type into a multisource probability for the event type.

[0104] Method **300** includes detecting an event from the aggregated features (**307**). For example, evaluation module **206** can determine that aggregated features **203** satisfy conditions to be detected as an event. Evaluation module **206** can detect event **224**, such as, for example, a fire, an accident, a shooting, a protest, etc. based on satisfaction of the conditions.

[0105] In one aspect, conditions for event identification can be included in thresholds **226**. Conditions can include threshold probabilities per event type. When a probability exceeds a threshold probability, evaluation module **106** can detect an event. A probability can be a single signal probability or a multisource (aggregated) probability. As such, evaluation module **206** can detect an event based on a multisource probability exceeding a probability threshold in thresholds **226**.

[0106] FIG. **4** illustrates an example computer architecture **400** that facilitates detecting an event from features derived from multiple signals. As depicted, event detection infrastructure **103** further includes evaluation module **206** and validator **204**. Evaluation module **206** is configured to determine if features of a plurality of normalized signals indicate a possible event. Evaluation module **206** can detect (or not detect) a possible event based on one or more features of a normalized signal. Validator **204** is configured to validate (or not validate) a possible event as an actual event based on one or more features of another normalized signal.

[0107] FIG. **5** illustrates a flow chart of an example method **500** for detecting an event from features derived from multiple signals. Method **500** will be described with respect to the components and data in computer architecture **400**.

[0108] Method **500** includes receiving a first signal (**501**). For example, event detection infrastructure **103** can receive normalized signal **122**B. Method **500** includes deriving first one or more features of the first signal (**502**). For example, event detection infrastructure **103** can derive features **401** of normalized signal **122**B. Features **401** can include and/or be derived from time **123**B, location **124**B, context **126**B, content **127**B, type **128**B, and source **129**B. Event detection infrastructure **103** can also derive features **401** from one or more single source probabilities assigned to normalized signal **122**B.

[0109] Method **500** includes detecting a possible event from the first one or more features (**503**). For example, evaluation module **206** can detect possible event **423** from features **401**. Based on features **401**, event detection infrastructure **103** can determine that the evidence in features **401** is not confirming of an event but is sufficient to warrant further investigation of an event type. In one aspect, a single source probability assigned to normalized signal **122**B for an event type does not satisfy a probability threshold for full event detection but does satisfy a probability threshold for further investigation.

[0110] Method **500** includes receiving a second signal (**504**). For example, event detection infrastructure **103** can receive normalized signal **122**A. Method **500** includes deriving second one or more features of the second signal (**505**). For example, event detection infrastructure **103** can derive features **402** of normalized signal **122**A. Features **402** can include and/or be derived from time **123**A, location **124**A, context **126**A, content **127**A, type **128**A, and source **129**A. Event detection infrastructure **103** can also derive features **402** from one or more single source probabilities assigned to normalized signal **122**A.

[0111] Method **500** includes validating the possible event as an actual event based on the second one or more features (**506**). For example, validator **204** can determine that possible event **423** in combination with features **402** provide sufficient evidence of an actual event. Validator **204** can validate possible event **423** as event **424** based on features **402**. In one aspect, validator **204** considers a single source probability assigned to normalized signal **122**B in view of a single source probability assigned to normalized signal **122**B. Validator **204** determines that the signal source probabilities, when considered collectively satisfy a probability threshold for detecting an event.

[0112] Forming and Detecting Events from Signal Groupings

[0113] In general, a plurality of normalized (e.g., TLC) signals can be grouped together in a signal group based on spatial similarity and/or temporal similarity among the plurality of normalized signals and/or corresponding raw (non-normalized) signals. A feature extractor can derive features (e.g., percentages, counts, durations, histograms, etc.) of the signal group from the plurality of normalized signals. An event detector can attempt to detect events from signal group features.

[0114] In one aspect, a plurality of normalized (e.g., TLC) signals are included in a signal sequence. Turning to FIG. **6**A, event detection infrastructure **103** can include sequence

manager **604**, feature extractor **609**, and sequence storage **613**. Sequence manager **604** further includes time comparator **606**, location comparator **607**, and deduplicator **608**.

[0115] Time comparator **606** is configured to determine temporal similarity between a normalized signal and a signal sequence. Time comparator **606** can compare a signal time of a received normalized signal to a time associated with existing signal sequences (e.g., the time of the first signal in the signal sequence). Temporal similarity can be defined by a specified time period, such as, for example, 5 minutes, 10 minutes, 20 minutes, 30 minutes, etc. When a normalized signal is received within the specified time period of a time associated with a signal sequence, the normalized signal can be considered temporally similar to signal sequence.

[0116] Likewise, location comparator **607** is configured to determine spatial similarity between a normalized signal and a signal sequence. Location comparator **607** can compare a signal location of a received normalized signal to a location associated with existing signal sequences (e.g., the location of the first signal in the signal sequence). Spatial similarity can be defined by a geographic area, such as, for example, a distance radius (e.g., meters, miles, etc.), a number of geo cells of a specified precision, an Area of Interest (AoI), etc. When a normalized signal is received within the geographic area associated with a signal sequence, the normalized signal can be considered spatially similar to signal sequence.

[0117] Deduplicator **608** is configured to determine if a signal is a duplicate of a previously received signal. Deduplicator **608** can detect a duplicate when a normalized signal includes content (e.g., text, image, etc.) that is essentially identical to previously received content (previously received text, a previously received image, etc.). Deduplicator **608** can also detect a duplicate when a normalized signal is a repost or rebroadcast of a previously received normalized signal. Sequence manager **604** can ignore duplicate normalized signals.

[0118] Sequence manager **604** can include a signal having sufficient temporal and spatial similarity to a signal sequence (and that is not a duplicate) in that signal sequence. Sequence manager **604** can include a signal that lacks sufficient temporal and/or spatial similarity to any signal sequence (and that is not a duplicate) in a new signal sequence. A signal can be encoded into a signal sequence as a vector using any of a variety of algorithms including recurrent neural networks (RNN) (Long Short Term Memory (LSTM) networks and Gated Recurrent Units (GRUs)), convolutional neural networks, or other algorithms.

[0119] Feature extractor **609** is configured to derive features of a signal sequence from signal data contained in the signal sequence. Derived features can include a percentage of normalized signals per geohash, a count of signals per time of day (hours:minutes), a signal gap histogram indicating a history of signal gap lengths (e.g., with bins for 1 s, 5 s, 10 s, 1 m, 5 m, 10 m, 30 m), a count of signals per signal source, model output histograms indicating model scores, a sequent duration, count of signals per signal type, a number of unique users that posted social content, etc. However, feature extractor **609** can derive a variety of other features as well. Additionally, the described features can be of different shapes to include more or less information, such as, for example, gap lengths, provider signal counts, histogram bins, sequence durations, category counts, etc.

[0120] FIG. **7** illustrates a flow chart of an example method **700** for forming a signal sequence. Method **700** will be described with respect to the components and data in computer architecture **600**.

[0121] Method **700** includes receiving a normalized signal including time, location, context, and content (**701**). For example, sequence manager **604** can receive normalized signal **622**A. Method **700** includes forming a signal sequence including the normalized signal (**702**). For example, time comparator **606** can compare time **623**A to times associated with existing signal sequences. Similarly, location comparator **607** can compare location **124**A to locations associated with existing signal sequences. Time comparator **606** and/or location comparator **607** can determine that normalized signal **122**A lacks sufficient temporal similarity and/or lacks sufficient spatial similarity respectively to existing signal sequences. Deduplicator **608** can determine that normalized signal **122**A is not a duplicate normalized signal. As such, sequence manager **604** can form signal sequence **631**, include normalized signal **122**A in signal sequence **631**, and store signal sequence **631** in sequence storage **613**.

[0122] Method **700** includes receiving another normalized signal including another time, another location, another context, and other content (**703**). For example, sequence manager **604** can receive normalized signal **622**B.

[0123] Method **700** includes determining that there is sufficient temporal similarity between the time and the other time (**704**). For example, time comparator **606** can compare time **123**B to time **123**A. Time comparator **606** can determine that time **123**B is sufficiently similar to time **123**A. Method **700** includes determining that there is sufficient spatial similarity between the location and the other location (**705**). For example, location comparator **607** can compare location **124**B to location **124**A. Location comparator **607** can determine that location **124**B has sufficient similarity to location **124**A.

[0124] Method **700** includes including the other normalized signal in the signal sequence based on the sufficient temporal similarity and the sufficient spatial similarity (**706**). For example, sequence manager **604** can include normalized signal **124**B in signal sequence **631** and update signal sequence **631** in sequence storage **613**.

[0125] Subsequently, sequence manager **604** can receive normalized signal **122**C. Time comparator **606** can compare time **123**C to time **123**A and location comparator **607** can compare location **124**C to location **124**A. If there is sufficient temporal and spatial similarity between normalized signal **122**C and normalized signal **122**A, sequence manager **604** can include normalized signal **122**C in signal sequence **631**. On the other hand, if there is insufficient temporal similarity and/or insufficient spatial similarity between normalized signal **122**C and normalized signal **122**A, sequence manager **604** can form signal sequence **632**. Sequence manager **604** can include normalized signal **122**C in signal sequence **632** and store signal sequence **631** in sequence storage **613**.

[0126] Turning to FIG. **6**B, event detection infrastructure **103** further includes event detector **611**. Event detector **611** is configured to determine if features extracted from a signal sequence are indicative of an event.

[0127] FIG. **8** illustrates a flow chart of an example method **800** for detecting an event. Method **800** will be described with respect to the components and data in computer architecture **600**.

[0128] Method **800** includes accessing a signal sequence (**801**). For example, feature extractor **609** can access signal sequence **631**. Method **800** includes extracting features from the signal sequence (**802**). For example, feature extractor **609** can extract features **633** from signal sequence **631**. Method **800** includes detecting an event based on the extracted features (**803**). For example, event detector **611** can attempt to detect an event from features **633**. In one aspect, event detector **611** detects event **636** from features **633**. In another aspect, event detector **611** does not detect an event from features **633**.

[0129] Turning to FIG. **6**C, sequence manager **604** can subsequently add normalized signal **122**C to signal sequence **631** changing the signal data contained in signal sequence **631**. Feature extractor **609** can again access signal sequence **631**. Feature extractor **609** can derive features **634** (which differ from features **633** at least due to inclusion of normalized signal **122**C) from signal sequence **631**. Event detector **611** can attempt to detect an event from features **634**. In one aspect, event detector **611** detects event **636** from features **634**. In another aspect, event detector **611** does not detect an event from features **634**.

[0130] In a more specific aspect, event detector **611** does not detect an event from features **633**. Subsequently, event detector **611** detects event **636** from features **634**.

[0131] An event detection can include one or more of a detection identifier, a sequence identifier, and an event type (e.g., accident, hazard, fire, traffic, weather, etc.).

[0132] A detection identifier can include a description and features. The description can be a hash of the signal with the earliest timestamp in a signal sequence. Features can include features of the signal sequence. Including features provides understanding of how a multisource detection evolves over time as normalized signals are added. A detection identifier can be shared by multiple detections derived from the same signal sequence.

[0133] A sequence identifier can include a description and features. The description can be a hash of all the signals included in the signal sequence. Features can include features of the signal sequence. Including features permits multisource detections to be linked to human event curations. A sequence identifier can be unique to a group of signals included in a signal sequence. When signals in a signal sequence change (e.g., when a new normalized signal is added), the sequence identifier is changed.

[0134] In one aspect, event detection infrastructure **103** also includes one or more multisource classifiers. Feature extractor **609** can send extracted features to the one or more multisource classifiers. Per event type, the one or more multisource classifiers compute a probability (e.g., using artificial intelligence, machine learning, neural networks, etc.) that the extracted features indicate the type of event. Event detector **611** can detect (or not detect) an event from the computed probabilities.

[0135] For example, turning to FIG. **6**D, multi-source classifier **612** is configured to assign a probability that a signal sequence is a type of event. Multi-source classifier **612** formulate a detection from signal sequence features. Multi-source classifier **612** can implement any of a variety of algorithms including: logistic regression, random forest (RF), support vector machines (SVM), gradient boosting (GBDT), linear, regression, etc.

[0136] For example, multi-source classifier **612** (e.g., using machine learning, artificial intelligence, neural networks, etc.) can formulate detection **641** from features **633**. As depicted, detection **641** includes detection ID **642**, sequence ID **643**, category **644**, and probability **646**. Detection **641** can be forwarded to event detector **611**. Event detector **611** can determine that probability **646** does not satisfy a detection threshold for category **644** to be indicated as an event. Detection **641** can also be stored in sequence storage **613**.

[0137] Subsequently, turning to FIG. **6**E, multi-source classifier **612** (e.g., using machine learning, artificial intelligence, neural networks, etc.) can formulate detection **651** from features **634**. As depicted, detection **651** includes detection ID **642**, sequence ID **647**, category **644**, and probability **648**. Detection **651** can be forwarded to event detector **611**. Event detector **611** can determine that probability **648** does satisfy a detection threshold for category **644** to be indicated as an event. Detection **641** can also be stored in sequence storage **613**. Event detector **611** can output event **636**.

[0138] As detections age and are not determined to be accurate (i.e., are not True Positives), the probability declines that signals are "True Positive" detections of actual events. As such, a multi-source probability for a signal sequence, up to the last available signal, can be decayed over time. When a new signal comes in, the signal sequence can be extended by the new signal. The multi-source probability is recalculated for the new, extended signal sequence, and decay begins again.

[0139] In general, decay can also be calculated "ahead of time" when a detection is created and a probability assigned. By pre-calculating decay for future points in time, downstream systems do not have to perform calculations to update decayed probabilities. Further, different event classes can decay at different rates. For example, a fire detection can decay more slowly than a crash detection because these types of events tend to resolve at different speeds. If a new signal is added to update a sequence, the pre-calculated decay values may be discarded. A multi-source probability can be re-calculated for the updated sequence and new pre-calculated decay values can be assigned.

[0140] Multi-source probability decay can start after a specified period of time (e.g., 3 minutes) and decay can occur in accordance with a defined decay equation. Thus, modeling multi-source probability decay can include an initial static phase, a decay phase, and a final static phase. In one aspect, decay is initially more pronounced and then weakens. Thus, as a newer detection begins to age (e.g., by one minute) it is more indicative of a possible "false positive" relative to an older event that ages by an additional minute.

[0141] In one aspect, a decay equation defines exponential decay of multi-source probabilities. Different decay rates can be used for different classes. Decay can be similar to radioactive decay, with different tau values used to calculate the "half life" of multi-source probability for a class. Tau values can vary by event type.

[0142] In FIGS. **6**D and **6**E, decay for signal sequence **631** can be defined in decay parameters **114**. Sequence manager **104** can decay multisource probabilities computed for signal sequence **631** in accordance with decay parameters **614**.

[0143] The components and data depicted in FIGS. **1-8** can be integrated with and/or can interoperate with one another to detect events. For example, evaluation module **206** and/or validator **204** can include and/or interoperate with one or more of: a sequence manager, a feature extractor, multi-source classifiers, or an event detector.

[0144] FIG. **9** illustrates an example computer architecture **900** that facilitates detecting events. The components and data described with respect to FIGS. **1-8** can also be integrated with and/or can interoperate with the data and components of computer architecture **900** to detect events.

[0145] As depicted, computer architecture **900** includes geo cell database **911** and even notification **916**. Geo cell database **911** and even notification **916** can be connected to (or be part of) a network with signal ingestion modules **101** and event detection infrastructure **103**. As such, geo cell database **911** and even notification **916** can create and exchange message related data over the network.

[0146] As descried, in general, on an ongoing basis, concurrently with signal ingestion (and also essentially in real-time), event detection infrastructure **103** detects different categories of (planned and unplanned) events (e.g., fire, police response, mass shooting, traffic accident, natural disaster, storm, active shooter, concerts, protests, etc.) in different locations (e.g., anywhere across a geographic area, such as, the United States, a State, a defined area, an impacted area, an area defined by a geo cell, an address, etc.), at different times from time, location, and context included in normalized signals.

[0147] Event detection infrastructure **103** can also determine an event truthfulness, event severity, and an associated geo cell. In one aspect, context information in a normalized signal increases the efficiency of determining truthfulness, severity, and an associated geo cell.

[0148] Generally, an event truthfulness indicates how likely a detected event is actually an event (vs. a hoax, fake, misinterpreted, etc.). Truthfulness can range from less likely to be true to more likely to be true. In one aspect, truthfulness is represented as a numerical value, such as, for example, from 1 (less truthful) to 10 (more truthful) or as percentage value in a percentage range, such as, for example, from 0% (less truthful) to 100% (more truthful). Other truthfulness representations are also possible.

[0149] Generally, an event severity indicates how severe an event is (e.g., what degree of badness, what degree of damage, etc. is associated with the event). Severity can range from less severe (e.g., a single vehicle accident without injuries) to more severe (e.g., multi vehicle accident with multiple injuries and a possible fatality). As another example, a shooting event can also range from less severe (e.g., one victim without life threatening injuries) to more severe (e.g., multiple injuries and multiple fatalities). In one aspect, severity is represented as a numerical value, such as, for example, from 1 (less severe) to 5 (more severe). Other severity representations are also possible.

[0150] In general, event detection infrastructure **103** can include a geo determination module including modules for processing different kinds of content including location, time, context, text, images, audio, and video into search terms. The geo determination module can query a geo cell database with search terms formulated from normalized signal content. The geo cell database can return any geo cells having matching supplemental information. For example, if a search term includes a street name, a subset of one or more geo cells including the street name in supplemental information can be returned to the event detection infrastructure.

[0151] Event detection infrastructure **103** can use the subset of geo cells to determine a geo cell associated with an event location. Events associated with a geo cell can be stored back into an entry for the geo cell in the geo cell database. Thus, over time an historical progression of events within a geo cell can be accumulated.

[0152] As such, event detection infrastructure **103** can assign an event ID, an event time, an event location, an event category, an event description, an event truthfulness, and an event severity to each detected event. Detected events can be sent to relevant entities, including to mobile devices, to computer systems, to APIs, to data storage, etc.

[0153] As depicted in computer architecture **900**, event detection infrastructure **103** detects events from information contained in normalized signals **122**. Event detection infrastructure **103** can detect an event from a single normalized signal **122** or from multiple normalized signals **122**. In one aspect, event detection infrastructure **103** detects an event based on information contained in one or more normalized signals **122**. In another aspect, event detection infrastructure **103** detects a possible event based on information contained in one or more normalized signals **122**. Event detection infrastructure **103** then validates the potential event as an event based on information contained in one or more other normalized signals **122**.

[0154] As depicted, event detection infrastructure **103** includes geo determination module **904**, categorization module **906**, truthfulness determination module **907**, and severity determination module **908**.

[0155] Geo determination module **904** can include NLP modules, image analysis modules, etc. for identifying location information from a normalized signal. Geo determination module **904** can formulate (e.g., location) search terms **941** by using NLP modules to process audio, using image analysis modules to process images, etc. Search terms can include street addresses, building names, landmark names, location names, school names, image fingerprints, etc. Event detection infrastructure **103** can use a URL or identifier to access cached content when appropriate.

[0156] Categorization module **906** can categorize a detected event into one of a plurality of different categories (e.g., fire, police response, mass shooting, traffic accident, natural disaster, storm, active shooter, concerts, protests, etc.) based on the content of normalized signals used to detect and/or otherwise related to an event.

[0157] Truthfulness determination module **907** can determine the truthfulness of a detected event based on one or more of: source, type, age, and content of normalized signals used to detect and/or otherwise related to the event. Some signal types may be inherently more reliable than other signal types. For example, video from a live traffic camera feed may be more reliable than text in a social media post. Some signal sources may be inherently more reliable than others. For example, a social media account of a government agency may be more reliable than a social media account of an individual. The reliability of a signal can decay over time.

[0158] Severity determination module **908** can determine the severity of a detected event based on or more of: location, content (e.g., dispatch codes, keywords, etc.), and volume of normalized signals used to detect and/or otherwise related to an event. Events at some locations may be inherently more severe than events at other locations. For

example, an event at a hospital is potentially more severe than the same event at an abandoned warehouse. Event category can also be considered when determining severity. For example, an event categorized as a "Shooting" may be inherently more severe than an event categorized as "Police Presence" since a shooting implies that someone has been injured.

[0159] Geo cell database **911** includes a plurality of geo cell entries. Each geo cell entry includes a geo cell defining an area and corresponding supplemental information about things included in the defined area. The corresponding supplemental information can include latitude/longitude, street names in the area defined by the geo cell, businesses in the area defined by the geo cell, other Areas of Interest (AOIs) (e.g., event venues, such as, arenas, stadiums, theaters, concert halls, etc.) in the area defined by the geo cell, image fingerprints derived from images captured in the area defined by the geo cell, and prior events that have occurred in the area defined by the geo cell. For example, geo cell entry **951** includes geo cell **952**, lat/lon **953**, streets **954**, businesses **955**, AOIs **956**, and prior events **957**. Each event in prior events **957** can include a location (e.g., a street address), a time (event occurrence time), an event category, an event truthfulness, an event severity, and an event description. Similarly, geo cell entry **961** includes geo cell **962**, lat/lon **963**, streets **964**, businesses **965**, AOIs **966**, and prior events **967**. Each event in prior events **967** can include a location (e.g., a street address), a time (event occurrence time), an event category, an event truthfulness, an event severity, and an event description.

[0160] Other geo cell entries can include the same or different (more or less) supplemental information, for example, depending on infrastructure density in an area. For example, a geo cell entry for an urban area can contain more diverse supplemental information than a geo cell entry for an agricultural area (e.g., in an empty field).

[0161] Geo cell database **911** can store geo cell entries in a hierarchical arrangement based on geo cell precision. As such, geo cell information of more precise geo cells is included in the geo cell information for any less precise geo cells that include the more precise geo cell.

[0162] Geo determination module **904** can query geo cell database **911** with search terms **941**. Geo cell database **911** can identify any geo cells having supplemental information that matches search terms **941**. For example, if search terms **141** include a street address and a business name, geo cell database **911** can identify geo cells having the street name and business name in the area defined by the geo cell. Geo cell database **911** can return any identified geo cells to geo determination module **904** in geo cell subset **942**.

[0163] Geo determination module can use geo cell subset **942** to determine the location of event **935** and/or a geo cell associated with event **935**. As depicted, event **935** includes event ID **932**, time **933**, location **934**, description **936**, category **937**, truthfulness **938**, and severity **939**.

[0164] Event detection infrastructure **103** can also determine that event **935** occurred in an area defined by geo cell **962** (e.g., a geohash having precision of level 7 or level 9). For example, event detection infrastructure **103** can determine that location **934** is in the area defined by geo cell **962**. As such, event detection infrastructure **903** can store event **935** in events **967** (i.e., historical events that have occurred in the area defined by geo cell **962**).

[0165] Event detection infrastructure **103** can also send event **935** to event notification module **916**. Event notification module **916** can notify one or more entities about event **134**.

[0166] The present described aspects may be implemented in other specific forms without departing from its spirit or essential characteristics. The described aspects are to be considered in all respects only as illustrative and not restrictive. The scope is, therefore, indicated by the appended claims rather than by the foregoing description. All changes which come within the meaning and range of equivalency of the claims are to be embraced within their scope.

What is claimed:

1. A method comprising:

receiving a first normalized signal;

deriving first one or more features of the first normalized signal;

determining that the first one or more features do not satisfy conditions to be identified as an event;

receiving a second normalized signal;

deriving second one or more features of the second signal;

aggregating the first one or more features with the second one or more features into aggregated features; and

detecting an event from the aggregated features.

2. The method of claim **1**, wherein aggregating the first one or more features with the second one or more features into aggregated features comprises:

detecting a possible event from the first one or more features;

validating the possible event as an actual event based on the second one or more features.

3. The method of claim **1**, further comprising including the first normalized signal in a signal sequence;

determining that the second normalized signal has sufficient temporal similarity to the first normalized signal;

determining that the second normalized signal has sufficient spatial similarity to the first normalized signal; and

including the second normalized signal in a signal sequence that contains the first normalized signal.

4. The method of claim **3**, wherein aggregating the first one or more features with the second one or more features into aggregated features comprises deriving features of the signal sequence from the first one or more features and the second one or more features.

5. The method of claim **4**, wherein deriving features of the signal sequence comprises deriving one or more of: a percentage, a count, a histogram, or a duration.

6. The method of claim **1**, wherein the first normalized signal is one of: a social post with geographic content, a social post without geographic content, an image from a camera feed, a 911 call, weather data, IoT device data, satellite data, satellite imagery, a sound clip from a listening device, data from air quality sensors, a sound clip from radio communication, crowd sourced traffic information, or crowd sourced road information.

7. The method of claim **6**, wherein the second normalized signal is a different one of: a social post with geographic content, a social post without geographic content, an image from a traffic camera feed, a 911 call, weather data, IoT device data, satellite data, satellite imagery, a sound clip from a listening device, data from air quality sensors, a sound clip from radio communication, crowd sourced traffic information, or crowd sourced road information.

8. The method of clam **1**, wherein deriving first one or more features of the first normalized signal comprises deriving the first one or more features from a first single source probability assigned to the first normalized signal;

wherein deriving second one or more features of the second normalized signal comprises deriving the first one or more features from a second single source probability assigned to the second normalized signal;

wherein aggregating the first one or more features with the second one or more features into aggregated features comprises aggregating the first single source probability and the second single source probability into a multisource probability;

wherein detecting an event from the aggregated features comprises detecting an event from the multisource probability.

9. A method, the method comprising:

receiving a normalized signal including time, location, context, and content;

forming a signal sequence including the normalized signal;

receiving another normalized signal including another time, another location, another context, and other content;

determining that there is sufficient temporal similarity between normalized signal and the other normalized signal;

determining that there is sufficient spatial similarity between the normalized and the other normalized signal; and

including the other normalized signal in the signal sequence based on the sufficient temporal similarity and the sufficient spatial similarity

10. The method of claim **9**, wherein determining that there is sufficient temporal similarity between the normalized signal and the other normalized signal comprises determining that the time and the other time are within a specified time of one another.

11. The method of claim **9**, wherein determining that there is sufficient spatial similarity between the normalized and the other normalized signal comprises determining that the location and the other location are within a specified distance of one another.

12. The method of claim **9**, wherein determining that there is sufficient spatial similarity between the normalized and the other normalized signal comprises determining that the location and the other location are within a specified number of geo cells of one another.

13. The method of claim **9**, further comprising determining that the other normalized signal is not a duplicate of the normalized signal prior to including the other normalized signal in the signal sequence.

14. The method of claim **9**, further comprising:

deriving one or more features of the signal sequence based on the normalized signal and the other normalized signal;

detecting an event from the derived one or more features.

15. The method of claim of claim **14**, wherein deriving one or more features of the signal sequence comprise deriving a multisource probability for the signal sequence.

16. The method of claim **15**, wherein deriving a multisource probability indicating the probability of the normalized signals in the signal sequence indicate a specified type of event.

17. A method, the method comprising:

accessing a signal sequence of normalized signals, normalized signals included in the signal sequence having a sufficient temporal similarity to one another and having a sufficient spatial similarity to one another;

extracting features from the signal sequence; and

detecting an event based on the extracted features.

18. The method of claim **17**, further comprising prior to detecting the event:

detecting that the extracted features do not indicate the event;

adding an additional normalized signal to the signal sequence;

extracting further features from the signal sequence based on the additional normalized signal; and

wherein detecting an event based on the extracted features comprises detecting an event based on the further extracted features.

19. The method of claim **17**, further comprising deriving a multisource probability from the extracted features; and

wherein detecting an event based on the extracted features comprises detecting an event based on the multisource probability.

20. The method of claim **17**, wherein extracting features of the signal sequence comprises deriving one or more of: a percentage, a count, a histogram, or a duration.

* * * * *