ORIGINAL RESEARCH

Distributed and scalable platform architecture for smart cities complex events data collection: Covid19 pandemic use case

Wadii Basmi[1] · Azedine Boulmakoul[1] · Lamia Karim[2] · Ahmed Lbath[3]

Abstract

For decades, numerous names have been given to boost an urban city: digital city, green city, smart cities and the list goes on. They are all accompanied with ideas and propositions to enrich citizens life quality, by employing latest information technology to improve environment's sustainability, through better energy usage, targeting problems affecting infrastructure costs, automation and efficient human resources distribution. Consequently, cities governors provide plans and conceive laws so society including individuals and organizations collaborate in a cycle of providers and consumers to make steps ahead toward smarti-fication of the city in which they all operate. Hence, hundreds of cities around the world are living example of what a smart cities could be resembling in terms of information technology advancement and everyday usage. Each application or to be general system serve and exist for a specific purpose, using mobile applications and small sensors together to cooperate and deliver a value imposing a huge economical and social value and significant source of data. However, most of this applications are tied to specific domains and solely designed to solve predefined problems. Thus, for a decision maker point of view, decisions' the cost becomes high to correlate multiple data flow in different shapes. As a solution, in this paper we propose a system that is based on abstracting city events of different backgrounds—social, urban and natural, we chose to call them complex space time events. Furthermore, we present its architecture and how it plays with its external actors, and finally, we explain a use case instance made specifically to counter Covid19 pandemic spread and retain public order.

1 Introduction

The internet of things is a giant market and it is continuing to grow in many areas all over the world, and is becoming a solid industry shaping the modern world and everyday's lives. The term smart is applied when a group of microcontroller sensors works together and exchange data throughout a network, to achieve the purpose for which the system they are part of was made (Pradeep and Shashikala 2019). Furthermore, as the data volume is immense, the later is usually part of multiple cycles displayed in visual human readable graphs to keep an eye on the evolution of their subjects.

There are many types of smart services that are aimed to automate and take over repetitive routinised tasks. Also, they are shipped in different shapes depending on the city's population, culture—including citizens awareness level, geography and schedules. Not to mention, on the micro-technical level, they tend to produce large data sets in rapid non-stop frequency, thus, they enter in the big data category, and so,

✉ Azedine Boulmakoul
azedine.boulmakoul@gmail.com

Wadii Basmi
wbasmi@gmail.com

Lamia Karim
lkarim.lkarim@gmail.com

Ahmed Lbath
Ahmed.Lbath@imag.fr

1 Computer Science Department, FSTM, Hassan II University of Casablanca, Casablanca, Morocco

2 LISA Laboratory ENSA Berrechid, Hassan 1st University, Settat, Morocco

3 LIG/MRIM, CNRS, University Grenoble Alpes, Grenoble, France

they can be a likable source of analytical data. Indeed, IoT is truly a blessing and the public tends to call its industrial revolution as "Industrie 4.0" alongside the others (Hui et al. 2017). As a matter of fact, there might be projects producing diverse types of data, yet, for external third parties to access this data, or make use of it, it remains a challenging task and it is usually cost-inefficient or have low return on investment to be made. Consequently, it presents as an obstacle toward an open city whose digital presentation to not necessarily be interconnected.

Our research aims to provide a system that collects all sorts of events that happens in a city. For that purpose, we designed the concept of complex space time events to abstract all kind of events that occur in a city, ranging from basic electronic sensors natural observations to actual changes mutating the states of urban services in a city (Basmi et al. 2020). Additionally, the aim of this abstraction is to construct a system that plays as an exchange node between smart objects so they can enrich their databases, but also, to paint the road forward another system or even a subsystem to take that data and offer generic framework to detect the correlation of anything that happens in a city, at a specific time and a specific geoposition. As a milestone, we took the use case of pandemic Covid19 that ravaged the world in March 2020, we thought of a digital solution that implements a strategy to seize the virus spread, a mobile application that collects complex space-time events related to Covid19 safety measures.

A modern city holds tremendous amount of information that isn't exclusive to electronic devices, in fact, it is an eco-system filled with distinct social, financial and urban phenomenas, therefore, the versatility of events exceeds measurable nature aspect. In this paper, We elaborate these type of events and dissect their importance to impulse the "Smart" effect to cities, by providing a meta model that describes their properties. In addition, we present a system capable of controlling the data flow between the different actors that participate in collecting and seeking the events, over different schemas and scenarios.

2 Related works

In Yadav (2019), they refer to a complex event as a paradigm for processing of data streams in order to provide relevant data and support for timely reactions. Moreover, it is also as shapeless information that can be produced anywhere, anytime and from different sources. In Al-Osta et al. (2019), they are presented as an aggregation of multiple granular events of common characteristics, furthermore, complex events usability is available for consumers as as an input to a complex events processing (CEP) engine (a system to process complex events and produce other meaningful complex

events). In this work, they present complex event as a solution to first: find patterns of appearance of traffic disruptions, then feed it to CEP engine, and discover interesting correlations between cultural events, social media and their influence on traffic. Their approach consists of three phases: collecting data from data sources related to the road, such as types of cultural events and natural conditions, afterward, they classified about 2035 events in eight categories, each category has a set of properties that differs from other categories. Then, they turned these categories into attributes, and applied Aprior algorithm in attempt to discover rules between any of the events and traffic.

In similar fashion, yet, with a different name, Morales and Garcia (2015) defines smart cities as geosmart cities that depends of an infrastructure composed of device and human sensors. They have provided a classification of sensors in three classes:

– *Fixed* They have permanent position and target natural events.
– *Mobile* These sensors are attached to other mobile devices or objects, such as smart phones, vehicles or drones.
– *Human* They are the humans that utilize social networks and third party systems to report data.

These events are represented as g-events and are part of an event-driven geoprocessing framework highlighted in a high level architecture, featuring three layers, where the first layer represents the sensing infrastructure that produce the data stream, then the second layer where all processing happens, and finally the services layer that encloses the consumers of the G-events (Morales and Garcia 2015). Moreover, g-events are subdivided into two types, low level events that represents the atomic finite state of a phenomenon of interest, and high level events which are aggregates of their counterparts. Furthermore, the authors go on giving theoretical representations of core properties of a g-events, which are the phenomenon, space, time interval of the happening, recurrence and derivative.

3 Challenges and contribution

3.1 Smart cities and complex space time events

Smart city defines any city that use internet of things technologies to develop urban services and improve citizens life quality, by giving easier cost efficient access (Kummitha and Crutzen 2019) and enriching the urban living (Hollands 2008). Additionally, a smart city favors technology usage and forces the decision makers to provide facilities that enable smart economy, smart mobility, smart environment

smart living and smart governance (Bruneo et al. 2019). Yet, it is not simply a one sided effort, as citizens have an important role in creating and engaging in IoT business models (Kitchin 2014). In fact, they can demand automated alternative to any traditional routine, calling for companies of private sector to invest and take over the economical wheel, henceforth, they impulse the later to answer the demand to collect economical benefit and speed up the technological revolution. Additionally, internet of things have helped governments to eliminate many job position in favor of automation, thus, easing the load on their budgets in favor of more meaningful services. Nonetheless, IoT's impact also extends to city centers interests profiling. Certainly, sensors have the ability to output grandiose amount of behavioural data, which helps in tracking and predicting their features of interest state changes, namely, air pollution, traffic congestion, unhealthy crowds, or even individuals tracking in countries where citizens privacy is negotiable for a greater cause.

3.2 Complex space time events processing framework data model

In a previous work (Basmi and Boulmakoul 2018), we have discussed complex space time events, which are events related to natural, social or urban phenomenons, and they happen in a specific geolocation and at a specific period of time (Maguerra et al. 2020). We called the system that collects the events Portunus, and in Fig. 1, we confer its data model inspired by Open Geospacial open Consortium (OGC) API; a family of standards being developed to ease development of geospacial data for the web. We chose OGC

Sensors API an open geospatial enabled and unified way to interconnect the Internet of things devices, data and applications across the web (OGC 2017). The data model focus on data streams, which are a collection of events characerised by a value, time, geolocation and the content type sent by sensors; an entity capable of observing a phenomena, capturing its state and sending it to a network, notably, traditional Arduino or Raspberry boards that provides natural readings, and software applications dedicated for monitoring crime declaration or motorcycle accidents in a city area. Additionally, a data stream is a link that helps to determine which actuators should benefit from the collected events in the datastream and which sensors can send its data. Likewise, an actuator is an entity that transform data collected by sensors into a task, for example, a micro-controller sensor or a software application that does use variety of complex space time events to determine the public order in a city area. Finally, a datastream operates on a feature of interest that portrays a geospcial feature, and associated with one phenomena. With this, Portunus is able to host a rich database of heterogeneous events, as a network of abstract things. As a result, this diversity allows consumers to better understand a smart city problems natural social and demographic contexts combined (Basmi et al. 2020).

3.3 Portunus complex space-time events collection framework

3.3.1 System layers

As illustrated in Fig. 2, there are four layers in Portunus:

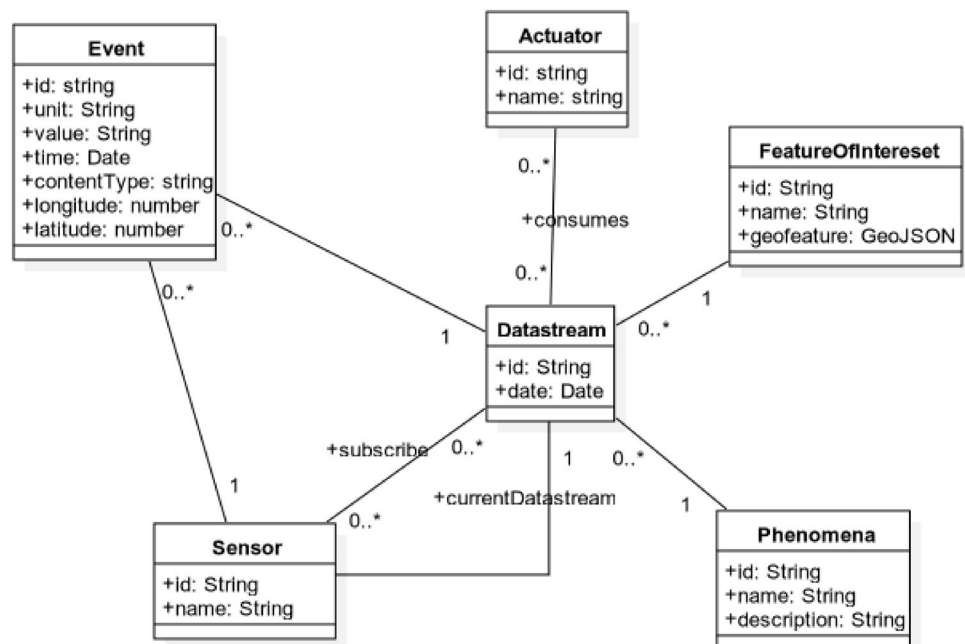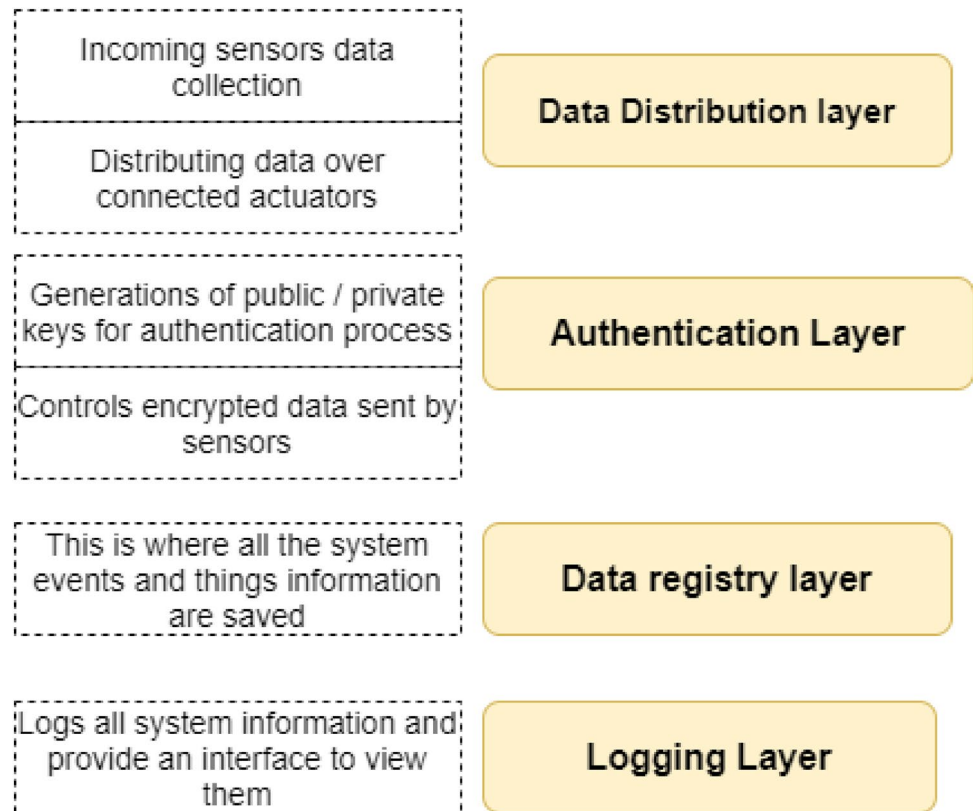**Fig. 1** Portunus entity relationship schema

Fig. 2 Portunus system layers

- *Data distribution layer* Responsible on data circulation between connected objects.
- *Authentication layer* Identifies connected objects and securing their exchange.
- *Data registry layer* It holds the entire system information database in a normalized format.
- *Logging layer* It logs all the system entries and gives an insurance for its health.

Figure 3 gives an internal insight of these layers, and the different components of Portunus architecture. As it can be noted, it is a system based on a hybrid approach when it comes to communication, precisely, asynchronous using a message broker and its messages topic and synchronous through direct communication. Each microservice runs in its own node, and is identified by its own entry in Portunus Domain naming server (DNS) server. With this approach, we isolate each microservice both in domain and as a physical instance, i.e. process. Sensors and actuators alike can communicate with Portunus using either HTTP the web protocol, COAP a lightweight web protocol designed for microcontrollers for better longevity, and MQTT a publish subscriber protocol for sensors too. The microservice responsible on receiving actuators connections and sending them collected events is the *data distributer* microservice, whereas, the *data gatherer* microservice is responsible on collecting events from connected sensors. Both are subscribed to the collected events message topic, and communicate with the security microservice to generate cryptographic keypairs for future communication.

In fact, in order to maintain a fast data flow in this ecosystem, it is important to filter what information is necessary for all the actors, thus, we put a layer that is responsible for registering all information related to concepts and leave only relevant information circulating respectively in channels, to reduce that size of data packets exchanged in the network and not overwhelm the layers with unrelated data. Finally, the last layer is the logging layer, it is in charge of saving all the activities that happens from saving data in the database, objects connections to errors logging, henceforth, it is an essential component to maintain and track the system's health.

3.3.2 Data circulation happy flow

Figures 4 and 5 illustrates data gathering happy flow, that shows how the data gatherer receives data from the sensors, with the help of the authenticator microservice to authenticate and control the access, plus, securing the communication within and outside the system. The scenario starts with the sensor subscribing to a predefined data stream—programmable by its owner after saving it to the portunus database. Using any of the supported protocols, the sensor sends its request to the data gatherer, which in turn takes

Fig. 3 Portunus system architecture
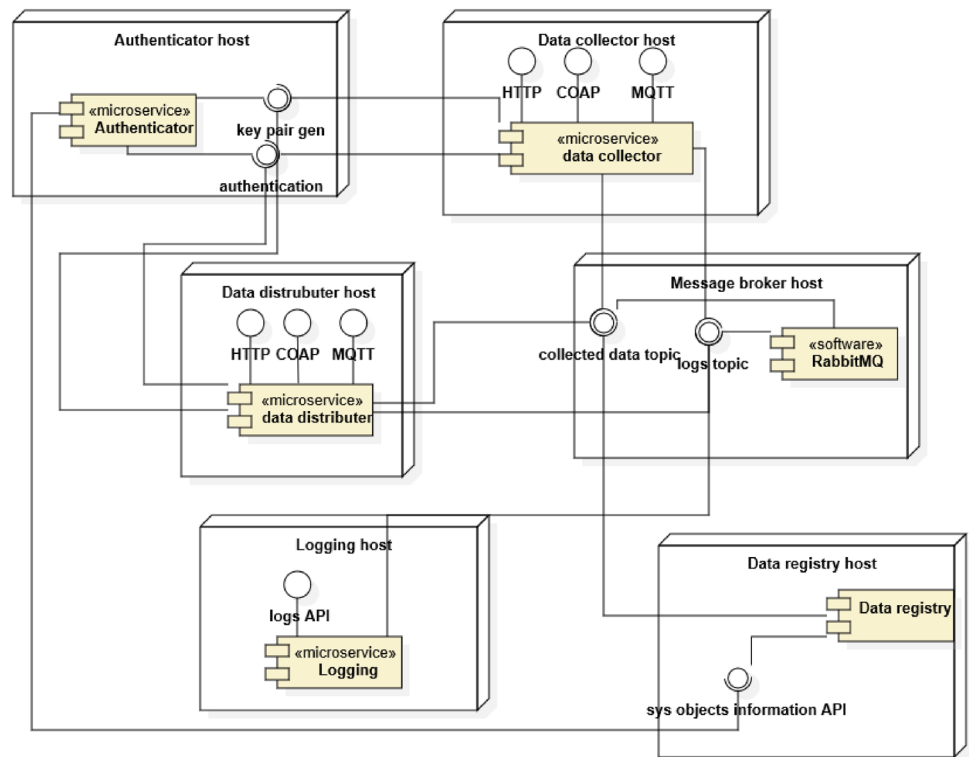
Fig. 4 Data distribution Happy flow (1)
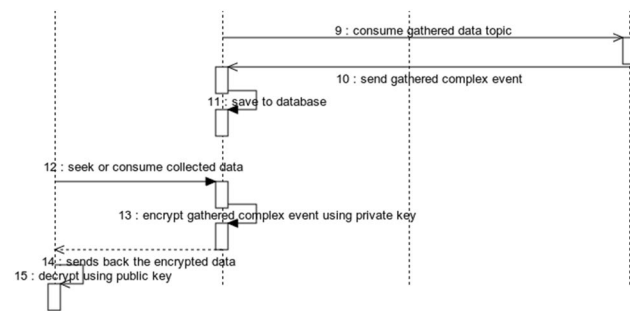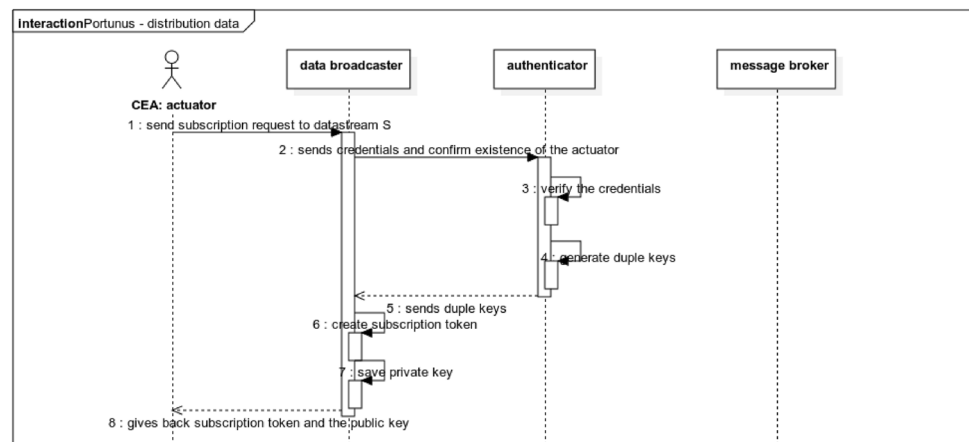
Fig. 5 Data distribution Happy flow (2)

the data stream identifier and the credentials of the sensor, and sends it to the authenticator to check if they are valid using its proper database. Once, the validation succeeds, it creates a duple of private and public keys, and sends it back to the data gatherer, to use it to create The later an expirable session subscription token. Then, it saves the key pair and delivers the public key and the token back to the sensor.

Later, when the authenticated sensor sends an observation, Portunus validates its session token and proceeds with saving *asynchronously* the observation in the events message topic and its own Mongodb database, but at the

same time, it returns the corresponding protocol response, if it awaits any. This is important, because as the number of connected objects increases, regardless of the scalability level, an asynchronous data persistence is much preferable. Moreover, sensors data is open to to loss and uncertainty by nature, i.e., much like in experiments, it is possible that one of the samples drops and fail to be observed either due to sensor deficiency or Portunus'. Nevertheless, if Portunus fails, the fail is reported to the logging microservice, and if failures reach a certain level of recurrence, then, the behaviour becomes abnormal rather than a trait.

3.3.3 Data collection performance evaluation

In the development mode, we have used three machines interconnected in a local network with a DNS server that was installed in a fifth machine where the Mongodb database server was running a replicaset of one primary nodes and two secondary nodes, and the RabbitMQ server too. Each machine had similar hardware and had 4 GB in RAM, with two a quad core. In attempt to test the acquittal of observation process, we performed a load testing that generated around 70 simultaneous request per second. The test was performed for three scenarios, where Portunus microservices were deployed in one machine, then two, then three machines (Fig. 6).

The results show that the difference is not noticeable when the number of connections is low between all the settings, however, as the number grows the acquittal time increases on the first settings earlier than the two other settings. Likewise, the second setting start to deviate from the third, as the number increases at three nodes. However, Portunus focus on delivering and acquitting first, before ensuring the persistence in the message queue, as long, as it has an
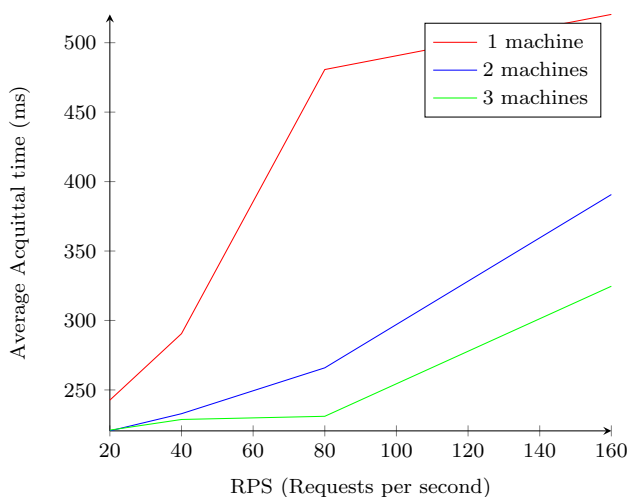

Fig. 6 Load testing complex events collection in different deployment settings

insurance that from the Node.js consumer server that takes care of saving

3.3.4 Data gathering security

Every object send a passcode and its identifier to connect and obtain the necessary session tokens and keys to use its data. Upon failing, Portunus adds it to a blacklist in order to block any further attempts and the connected object in question is marked as failing, and can only be unblocked through the administration panel.

4 Covid19 mobile defender: coronavirus safety measure infringement complex event use case

4.1 Covid19 pandemic and smart cities

Earlier in 2020, the whole world has known the first global pandemic to have touched every reachable area since the Spanish flu in the last century. However, unlike in the past, it took only few months for all the governments to take action thanks to technology, by spreading informative announcements through social and local media, such as taking precautions measures like social distancing and applying lock downs. Moreover, humanity was also able to share and work together throughout the emergency by sharing all their statistics and findings, in order to dissect the behaviour of the new virus and possibly a vaccine. Yet, these efforts didn't stop the countless job losses and heavy economical recession, thus, governments gradually agreed to lift the restrictions and barriers. Though, the cost was far too great in large countries and it has engendered public disorders in many cities, because of misinformation and political intervention to climb on top of the crisis. Equally important, most of security facilities lacked the numbers, strategies and was not able to use information technology to seize control over the spread of the virus, especially, in confinement lifting phase, where noticeable infections were recorded and even rising again in places where it was taught to have been pristine. More importantly, despite the smartness of many cities even in the first world countries, few managed to expect the situation as it was a first in one century, still, highly equipped hospitals where part of wholesome efforts that contributed in providing very accurate statistics in a nearly real time manner, and helped in predicting the causalities based on individuals health factors, such as the age, their genders and their overall medical conditions, giving big credits technology employment in health care industry (Corsi et al. 2020).

It is important to note, that despite the smartness of many cities even in the first world countries, few administrations expected this situation, but still, this is a negative factor

playing in deciding how smart a city is, i.e., how the society put in use technology to keep the safety of its citizens and its economy.

4.2 Covid19 mobile defender concepts

Covid Mobile defender is a mobile application that helps authorities or agents working in an institution responsible on preserving safety measures obedience among citizens in a country. It is based on the concept of complex space time events, where the sensor is the application itself observing well-defined qualitative complex space time events classified in ten classes, notably, infection, fatalities, mask wear, temperature checkpoints, bus, train, taxi, restaurant, shop and coffee shop safety measures infringements. Furthermore, its main functionalities resolve over:

– *Collaborative region monitoring* Users are able to fill the events database with interactive forms based on their geolocation. Each event value holds a flag representing the class, a value to evaluate either the number of people who committed the infringement, or the degree (high or low) depending on the chosen class.
– *Regions zoning* The entire country is shown in the map is subdivided in three categories:

   – *Neutral zones* Zone in which the administrator do not have an idea about the state of safety measures respect.
   – *Green zones* Zones which have low rate of events.
   – *Red zones* Zones which have high rate of events.

– *Events visualisation* The mobile applications provide a map with different filters to customized the displayed events, coverage distance and the zones. On top of that, on the administrator application, it is possible to mark the zones as a progress indicator of virus containment.
– *Reporting to Portunus* With every new event introduced by the application, it is sent to Portunus in a complex space time event format.

4.3 Covid mobile defender application workflow

We have covered the basic concepts of our mobile application and its soft dependency on Portunus. Next, in Figs. 7 and 8, we present the happy flow, its starts by deploying police forces over multiple neutral region, assuming this is the initial state until data says otherwise, even though, it is possible to consume Portunus data related to any event that influence the public order in a region. In fact, the latter can be determined by many observable factors such as crime rates, incidents, poverty, income reports and many other sociological aspects, as long as there are automated services
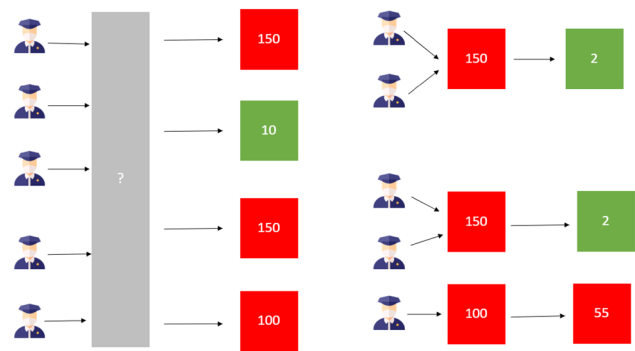

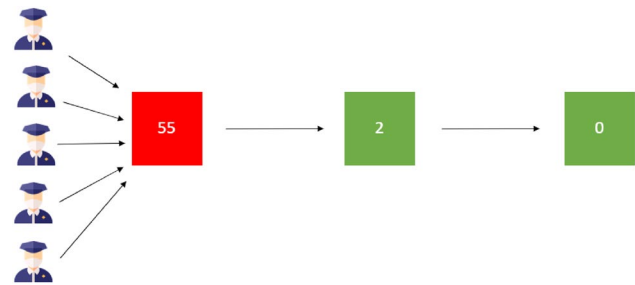**Fig. 7** Mobile defender virus containment 1

Fig. 8 Mobile defender virus containment 2

that can provide Portunus using any of its channels. Hence, the complex space time events platform offers an excellent arsenal of multidimensional data, pertinent to enrich smart cities governors.

Over a period of a week or few days, the administrator leverage the collected events and depending on their distribution on the city's urban areas, they will mark each area with its corresponding color, using a rate per coverage surface expectation metric to decide weather a zone should go red or green. Thereupon, patrols will be much lighter in areas where safety measures are less complied. The process is to repeat, until the red zones disappear or turn into small points, and the rest becomes green. The resources are not limited to police staff, but, based on the number of infections and fatalities, doctors can also be ported to hospitals whom areas are likely to be virus nests. Thus, using this mobile application is effective to reduce the spread of the virus, thanks to the smart allocation of the essential resources that are key to control the virus, in this case, Covid19.

The mobile application runs on Android smart phones, Fig. 9 represents the home page, where the user can view the map with the list of the most recent events, featured by small icons. By clicking on the filter button, the user can control the event types, the coverage radius, zones to display and the maximum period in which the events have been recorded. Moreover, the user can also add new events by selecting the position, then they can feel the form with
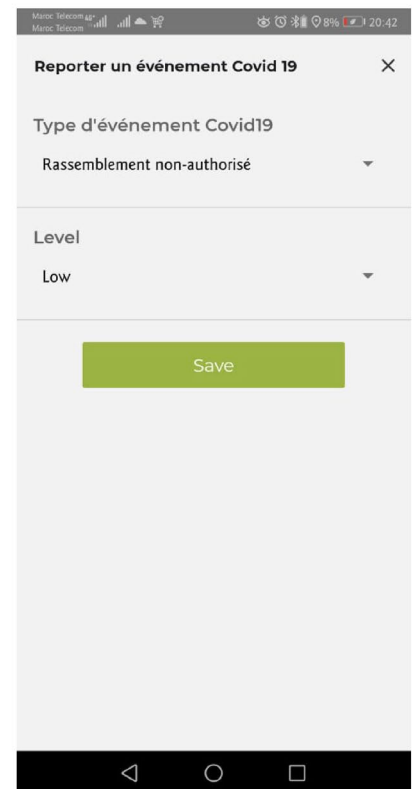
Fig. 9 Covid19 mobile defender
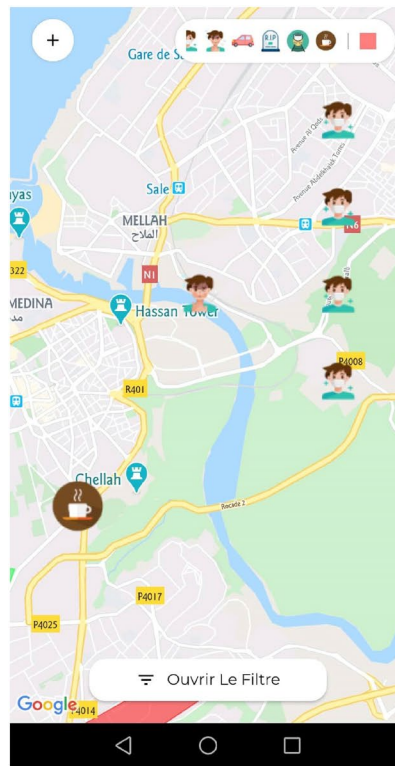
(a) Home page
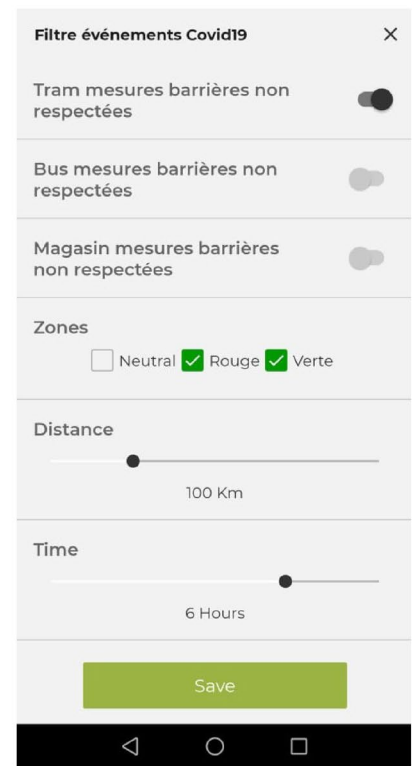
(b) Filter page

(c) Position page

(d) New Covid19 event form

the right type of infringement and its value, then, after submitting, all users can see the new event introduced to the system, and of course, it will be shared with Portunus as a complex space time event.

5 Conclusion

Smart cities are urban areas that utilizes technology and the internet of things to ease access to improve its citizens life quality through faster and smoother access to its urban services. Internet of things represents the smart devices connected to a network to work together in order to accomplish a task. This technology is becoming ubiquitous in the daily life, but also, in cities public and private institutions alike. As a result, companies as an example are able to learn a lot about its customers, and only recently, first world countries are working toward the legislation of many digital practices to protect the common public interest. Complex space time events are concepts that adds an abstract layer on top of anything that happens in a smart city, with the aim to localize it and define the means to collect this information, save it and present it as an opportune source of more accurate decision in other domains, given the presence of a correlation with its own events database. Portunus, a system designed as a microservice ecosystem provides different channels for APIs to communicate with it, in order to either listen on or feed a data stream with complex space time events, in addition, such architecture enforce diversity of technology, separation of concerns and easier scalability. This system is meant to communicate with actuators and sensors, however, they are not restrained to traditional sensors, but can also be any application able to provide Portunus with information with the location and the recording time. Covid19 mobile defender is an example of a sensor that captures the virus safety measures infringement measures, with the intention to help authorities a better control over the spread of the coronavirus and tight resources allocation. As for the perspective, we aim to extend Portunus with an analytics layer to process the information it collects and finds means to classify those events based on anything that can permit to discover links between the collected events and solve a smart city's problems.

References

Al-Osta M, Bali A, Gherbi A (2019) Event driven and semantic based approach for data processing on IOT gateway devices. J Ambient Intell Humaniz Comput 10(12):4663–4678. https://doi.org/10.1007/s12652-018-0843-y

Basmi W, Boulmakoul A (2018) Distributed and scalable framework for smart city real-time complex event processing. In: Conference ASD 2018: big data & applications 12th edition of the conference on advances of decisional systems, 2–3 May 2018, Marrakech Morocco, vol 1. ISBN: 9 789920 356794

Basmi W, Boulmakoul A, Karim L, Lbath A (2020) Modern approach to design a distributed and scalable platform architecture for smart cities complex events data collection. Procedia Comput Sci 170:43–50. https://doi.org/10.1016/j.procs.2020.03.008. http://www.sciencedirect.com/science/article/pii/S1877050920304300 (**the 11th International Conference on Ambient Systems, Networks and Technologies (ANT)/The 3rd International Conference on Emerging Data and Industry 4.0 (EDI40)/Affiliated Workshops**)

Bruneo D, Distefano S, Giacobbe M, Longo Minnolo A, Longo F, Merlino G, Mulfari D, Panarello A, Patanè G, Puliafito A, Puliafito C, Tapas N (2019) An iot service ecosystem for smart cities: the #smartme project. Internet Things 5:12–33. https://doi.org/10.1016/j.iot.2018.11.004. http://www.sciencedirect.com/science/article/pii/S2542660518301008

Corsi A, de Souza FF, Pagani RN, Kovaleski JL (2020) Big data analytics as a tool for fighting pandemics: a systematic review of literature. J Ambient Intell Humaniz Comput. https://doi.org/10.1007/s12652-020-02617-4

Hollands RG (2008) Will the real smart city please stand up? City 12(3):303–320. https://doi.org/10.1080/13604810802479126

Hui TK, Sherratt RS, Sánchez DD (2017) Major requirements for building smart homes in smart cities based on internet of things technologies. Future Gener Comput Syst 76:358–369. https://doi.org/10.1016/j.future.2016.10.026. http://www.sciencedirect.com/science/article/pii/S0167739X16304721

Kitchin R (2014) The real-time city? Big data and smart urbanism. GeoJournal 79(1):1–14. https://doi.org/10.1007/s10708-013-9516-8

Kummitha RKR, Crutzen N (2019) Smart cities and the citizen-driven internet of things: a qualitative inquiry into an emerging smart city. Technol Forecast Soc Change 140:44–53. https://doi.org/10.1016/j.techfore.2018.12.001. http://www.sciencedirect.com/science/article/pii/S0040162518308011

Maguerra S, Boulmakoul A, Karim L, Badir H, Lbath A (2020) Towards a reactive system for managing big trajectory data. J Ambient Intell Humaniz Comput 11(10):3895–3906. https://doi.org/10.1007/s12652-019-01625-3

Morales J, Garcia M (2015) GeoSmart cities: event-driven geoprocessing as enabler of smart cities. In: 2015 IEEE first international smart cities conference (ISC2), Guadalajara, pp 1–6. https://doi.org/10.1109/ISC2.2015.7366207

OGC (2017) Ogc sensorthings API. https://www.opengeospatial.org/standards/sensorthings. Accessed 17 Dec 2019

Pradeep PT, Shashikala KL (2019) Smart city services—challenges and approach. In: 2019 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon), pp 553–558. https://doi.org/10.1109/COMITCon.2019.8862243

Yadav P (2019) High-performance complex event processing framework to detect event patterns over video streams. In: 20th ACM/IFIP international conference MIDDLEWARE 2019. https://doi.org/10.1145/3366624.3368169