Crowdsourcing for the spatialization and signaling of Covid-19 transmission predictors: an approach based on risk perception

Murilo Guerreiro Arouca[1], Carlos Daniel S. Cruz[1], Marcos Ennes Barreto[1,2], Isa Beatriz da C. Neves[1], Federico Costa[1], Hussein Khalil[3], Ricardo Lustosa Brito[1]

[1]Federal University of Bahia (UFBA)

[2]London School of Economics and Political Science (LSE)

[3]Swedish University of Agricultural Sciences (SLU)

{murilo.guerreiro, cdansantana, marcosb, isabeatriz}@ufba.br

{hussein.khalil88, fcosta2001, lustosaricardo}@gmail.com

Abstract. *Popular participation in public health actions is essential for fighting Covid-19, especially in vulnerable urban communities where the lack of geographical data at fine resolution scale hinders appropriate spatial responses. This work proposes a crowdsourcing-based solution that captures georeferenced data regarding the population's perception of risk in relation to transmission predictors of Coronavirus. The proposed solution allows for mapping and sending real-time alerts regarding the presence of such transmission predictors. A validation study involving 20 people from a community in the city of Salvador revealed that the proposed solution is highly acceptable as user-centred alert tool, especially among young people.*

1. Introduction

Progressively increasing in terms of innovation in collaboration, crowdsourcing is a process of co-production of ideas, content and artifacts that usually involves strategies for collecting information or other material or immaterial resources from an indefinite number of individuals who aim to achieve specific objectives or implement a particular project [de Freitas and Ewerton 2019]. It is notable the potential that mobile devices have to amplify this model of problem solving, so that through these devices it is possible to make the contribution easier and ubiquitous.

Existing studies in the literature suggest that data collected trough crowdsourcing approach are significantly relevant for coping, responding to, and mitigating damage from epidemics, and more recently, the Covid-19 pandemic [Leung and Leung 2020]. In some circumstances, these data can be used to complement official data.

However, the task of informing the population about the risks and prevention of Covid-19 can be a quite challenging task. Therefore, working in the context of risk perception can be an effective alternative for the communication of health policy, since the high media exposure of the disease may have contributed to the construction of different perspectives regarding the risk of Covid-19 to people's health, which may be associated with political, economic and geographic aspects [Huynh et al. 2020].

In Brasil, Covid-19's spatial data usually cover the visualization of data on an inter-municipal or higher geographically scale, showing the lack of fine resolution geographical scale spatial data to support responses during health emergencies. This is particular worse in in vulnerable urban communities. This problem is not observed only in Brazil but also in other Western countries, where privacy policies hinder the use of contact tracking devices and applications. This contributes negatively to this lack of fine-grain information.

In this scenario, crowdsourcing associated with the population's perception of risk is an opportune alternative for confronting Covid-19, mainly in vulnerable urban communities. The success with previous studies resulting from the use of crowdsourcing, participatory mapping, and gamification for the mapping of zoonoses and their predictors in vulnerable urban communities [Arouca et al. 2019], indicates that the use of crowdsourcing can be significantly effective for the spatialization and signaling of transmission predictors of Covid-19.

2. Background and related work

With the advent of the pandemic context that came with the spread of the Coronavirus, several technologically supported approaches started to be used in several countries to support the confrontation of Covid-19, as well as to mitigate the local and global impacts caused by the disease [Ferreira and Pena 2020]. Among these approaches, participatory surveillance stands out as one of the main approaches.

Often used as a form of social monitoring during the Covid-19 pandemic, the participatory surveillance approach is explored in different ways in several countries, being marked by the use of applications for mobile devices. Due to the pandemic, and even with scarce resources and a deficient health system, India inaugurated the country's first participatory disease surveillance initiative, implemented through the app Aarogya Setu [Garg et al. 2020b].

Aarogya Setu aims to track possible carriers of Covid-19. Using device geolocation data and bluetooth to detect the approximation of infected people, the application also has a chat bot with updates from the country's health ministry about the disease. The use of the application by the population is optional.

The tracking is performed with the help of Bluetooth technology and social graphs generated by location or GPS, which shows the user's interaction with anyone who has been tested positive for coronavirus and notifies them. It detects and tracks the user's movement with the help of GPS and Bluetooth sensors. It is also able to identify other nearby smartphones that also have the application installed and can send a notification if it contacts infected people using its database and algorithms.

Like India, China, Russia, South Korea and Singapore also stood out for using the participatory surveillance approach. However, unlike India, these countries used this approach as a form of social monitoring. China, for example, demanded that its citizens install an application on their smartphones, so that, based on personal information, the system generates a QRcode with a color referring to the individual's quarantine status. In this way, the color of the code determines the user's mobility restriction, and whether he can attend public places, such as subways and shopping malls [Courtney 2020].

When observing the contact tracking methodologies elucidated in the aforementioned works, it can be seen that initiatives that involve contact tracking are useful for coping with infectious diseases such as Covid-19. However, due to the privacy policies adopted by Western countries, this methodology is not widespread in these countries [Garg et al. 2020a], unlike Eastern countries, since in the Eastern world there are several totalitarian government regimes and the use of these technologies in some cases is mandatory.

Therefore, when considering the privacy policies in force in Western countries, there is a gap regarding the use of Contact Tracing methodologies for spatialization and signaling of data related to Covid-19. In this way, the present work proposes a solution that uses crowdsourcing as an alternative to the use of contact tracing methodologies, and that is feasible with regard to the privacy policies of western countries.

3. Proposed solution: +Lugar platform (Covid-19 version)

+Lugar is a platform that aims to map zoonoses and their predictors in vulnerable urban communities using collaborative mapping, crowdsourcing and gamification. The first version of the platform had an institutional section aimed at covering other projects that aimed to carry out collaborative mapping in public health. Due to the impacts caused by the advent of the new Coronavirus, as well as the difficulty in obtaining social engagement to face Covid-19, a new version of the platform was developed for, among other objectives, collecting data for the spatialization of Covid-19 transmission predictors on an intra-neighborhood scale, as well as alerting users of the application about the proximity to areas indicated as at risk of transmission of Covid-19 in real time.

The Covid-19 version of the +Lugar platform aims to use its intuitive and interactive interface that uses a theory-oriented gamification design model implemented on the platform as support to crowdsourcing and participatory mapping approaches in practices aimed at coping with Covid-19. Figure 1 depicts the main section of the application, comprised by resources developed in the Covid-19 version of the platform (left side), including a section where some of the application's gamification resources are located.

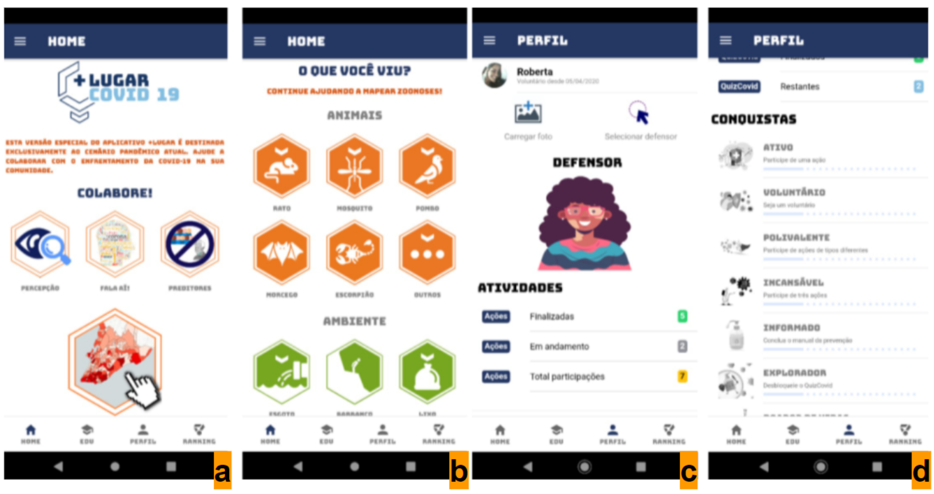

Figure 1. a) Covid-19 data collection and visualization environment, b) environment for collecting data on zoonoses, c) interface customization environment, d) gamification and engagement resources.

Due to the need for a collective effort to mitigate damage and cope with the Covid-

19 pandemic [Wolf et al. 2020], some functions have been implemented on the +Lugar platform in order to promoting the mapping of predictors, such as agglomeration points and the absence of the use of Personal Protective Equipment (PPE) in public places, as well as to create a proximity alert mechanism of these areas with the user.

Although public health policies have widely publicized the main predictors related to the transmission of the new Coronavirus, as well as the main conducts adopted to cope with it, it is still not possible to observe in the literature a validated method for measuring data about compliance with sanitary measures, mainly that involves the spatialization of these data on a fine resolution geographical scale.

Thus, this work defined as itś bases the use of crowdsourcing and participatory mapping approaches associated with data based on the population's risk perception. Although this work is limited to the use of georeferenced data regarding to risk perception, in the application there are questionnaires that are not georeferenced but that involve risk perception data.

Increasingly present in the context of epidemiology, studies that use the risk perception of individuals are frequently carried out, since they are able to signal subjective indicators that are generally not available in other research approaches. This is because risk perception studies consider, in a unique way, factors of cultural and social origin that can interfere with the perception of individuals [Giulio et al. 2015].

The proximity alert mechanism works by communicating the application with the server, so that, based on the georeferenced data stored on the server, the application notifies the user in real time if he enters a radius of 30 meters from an area indicated by the presence of a transmission predictor. It is important to highlight that this mechanism is addressed by the gamification strategies present on the platform.

+Lugar platform is structured in two parts, so that one side has a server with an API and the other has two applications, a hybrid application for mobile devices and a Web application corresponding to the client interface. For the development and implementation of this work, only the Android version of the mobile application was used.

In the current version, the mobile application of the +Lugar platform was updated to use the Ionic Framework version 4, as well as the platform API was rewritten in the Python language using the Flask framework, and integration with SQLAlchemy, a open source object-relational mapper (ORM) library developed for the Python programming language. The server application is running in a virtual machine on the Google Cloud platform. The machine runs a Linux distribution, configured to host all the technologies described before.

This work demanded a specific form of communication between the application and the server. For this, Socket.IO was used, which is a JavaScript library allowing for two-way communication between client and server in real time. According to the documentation, it can run on the majority of platforms and browsers. The communication between the server and their clients is established with a WebSocket connection whenever possible, and will fallback to use HTTP long-polling method when necessary.

In order to work with geographic data, such as user's position, some technologies were used. On the mobile application, GPS data was collected from the device's sensors, using Ionic's built-in features to access the sensor. Regarding user's visualization and interaction with the map features, LeafLet was used. Leaflet is a library developed for the projection of interactive maps with an intuitive design, being relatively light compared to other libraries in the segment, for example, the Google Maps API. This library is essential to obtain greater coverage of devices, as it can be used on on devices with a slow connection, low processing capacity and memory. Figure 2 shows a Leaflet map of the +Lugar platform.

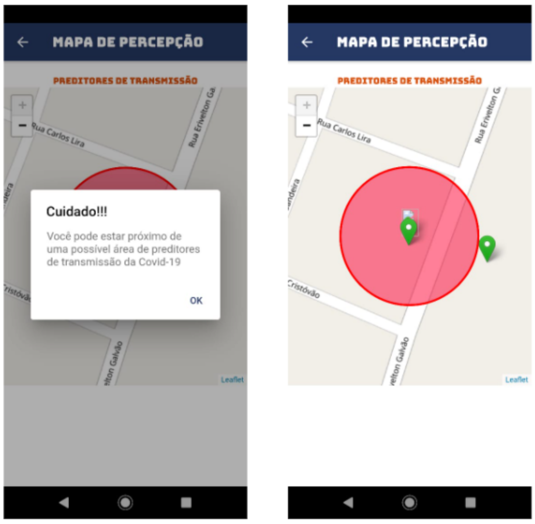

Figure 2. Perception map where all transmission predictors signaled by users will be spatialized. It also allows visual identification of the user's position in relation to signaled points.

In order to handle the geodata received on the server application from the mobile app, the GeoPy module was choosen together with the HaverSine module. These modules allow to execute some functions, such as distance calculation based on latitude and longitude, converting these positions to text addresses, among others. These modules are important for calculating the distance between the user and the Covid-19 transmission predictors previously signaled on the platform.

4. Acceptance and use of technology

To assess the acceptance and use of the proposed solution, we have used UTAUT (Unified Theory of Acceptance and Use of Technology). It is an acceptance model that uses constructs with four moderating conditions (Gender, Age, Experience, and Voluntariness of use) and four determining conditions (Performance expectation, Effort expectation, Social influence, and Facilitating conditions). Thus, the implication of moderating factors in relation to determining factors is observed, aiming at behavioral intention and usage behavior [Venkatesh et al. 2003].

The means used to collect the data was the application of an electronic form made available to research participants through an internet link. The questionnaire was prepared on Google Forms and included questions about the constructs of the theory developed by [Venkatesh et al. 2003], using the Likert scale (Table 1). As shown in Table 2, to obtain each determining factor, three questions were listed in the questionnaire.

Table 1. Likert scale used.

1	2	3	4	5	6	7
I completely disagree	Largely disagree	I partially disagree	Neutral	Partially agree	I largely agree	I completely agree

Table 2. Questions about determining factors.

Questões	
Performance expectation	Q1 - I believe the platform is/will be useful for the Covid-19's confrontation.
	Q2 - Using the platform allowed/will allow me to contribute to the fight against Covid-19.
	Q3 - Using the platform, I have/will have more ways to contribute to facing Covid-19.
Effort expectation	Q1 - My interaction with the platform is/will be clear and easy to understand.
	Q2 - It was/will be easy for me to become a skilled user of the platform.
	Q3 - Learning to use the platform was/will be easy for me.
Social influence	Q1 - People who influence my behavior think I should use the platform.
	Q2 - People who matter to me think I should use the platform.
	Q3 - In general, the Collectives and Community Associations in my neighborhood have encouraged the use of the platform.
Facilitating conditions	Q1 - I have/will have the necessary resources (hardware, mobile device, internet) to use the platform.
	Q2 - I have/will have the necessary knowledge to use the platform.
	Q3 - A specific person or group will be available to help when I have difficulties with the platform.

4.1. Characterization of participants

This study was carried out with young people from the community of Marechal Rondon, located in the rail suburb of the city of Salvador. Twenty young people aged between 13 and 23 years participated in this study. Demographic analysis of the sample indicated that 65% of respondents were men and 35% were women. Only 15% of participants had no prior experience of using and contributing to collaborative platforms. Furthermore, 70% of respondents were between 15 and 20 years of age at the time of filling out this questionnaire.

4.2. Preliminary results

To assess the reliability of the questionnaire applied in the research, Chrobach's alpha coefficient was used, a remarkable coefficient with regard to checking the internal consistency of a scale for a set of construct indicators through a correlation [Bland and Altman 1997]. The result of the reliability analysis obtained through the Chrobach's Alpha coefficient was 0.86, as well as the standardized alpha coefficient was 0.86. Thus, this analysis demonstrates a satisfactory internal consistency.

Still in progress, the analyzes performed so far showed satisfactory results. In Figure 3, it is possible to observe the distribution of responses related to the determining factors of the UTAUT model in this preliminary verification. The Performance expectation boxplot shows that the survey respondents believe that the methodology presented in this version of the platform is relevant for confronting Covid-19.

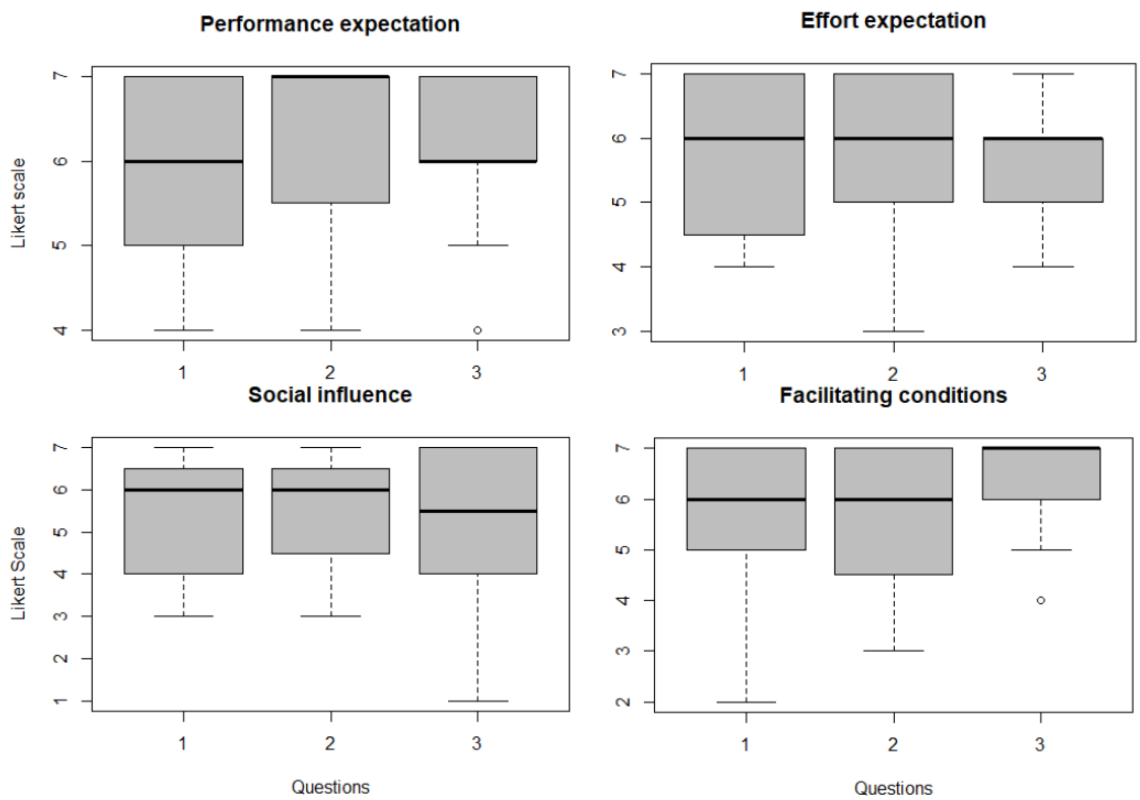

Figure 3. Distribution of responses related to the determining factors.

As can be seen in the Effort expectation boxplot, the young people did not indicate problems in terms of understanding the platform and its functions. Regarding Social influence, it is observed that community associations and other people have an influence on the use of the platform by young people. In the boxplot about the facilitating conditions, it is evident that most of the young participants have the necessary resources to use the platform. In the future, to perform multivariate analyses, this questionnaire will be applied in a new study with more participants from Marechal Rondon and other communities in the suburb of Salvador.

5. Conclusion and future work

This paper presented an althernative solution for the spatialization and signaling of Covid-19 transmission predictors using the crowdsourcing. This solution is very interesting in view of the ethical implications that permeate the use of the participatory surveillance and contact tracing approaches. Through the approach proposed in this work, it is possible to obtain data on Covid-19 transmission predictors without the need for social monitoring. Although this approach does not make use of data with the same reliability as data from official sources, it manages to provide greater precision regarding the spatialization of the collected data. Tests with synthetic data have shown excellent metrics, as well as the assessment of the acceptance and use of the technologies showed promising results.

Through the presented solution, it is expected to obtain spatially distributed indications of the presence of predictors of SARS-CoV-2 transmission in a participatory way, as well as real-time notification regarding the approximation of users of the application in relation to the areas previously signaled. Through the proposed risk perception approach, it is also expected to generate georeferenced data on an intra-neighborhood scale, this can be done due to the characteristics of this form of collaboration.

In addition to the spatial understanding of transmission predictors being important for the development of prevention and decision-making strategies, in the future this information may also be associated with other data on the disease. It is expected that this solution can be explored in other possible epidemiological scenarios.

References

Arouca, M. G., Neves, I. B. C., and Brito, R. L. (2019). +Lugar: Um aplicativo gamificado destinado à saúde coletiva. *Anais do Seminário Tecnologias Aplicadas a Educação e Saúde*.

Bland, J. M. and Altman, D. G. (1997). Statistics notes: Cronbach's alpha. *Bmj*, 314(7080):572.

Courtney, C. (2020). Covid-19 and China's Health Code System. *Somatosphere, April*, 5.

de Freitas, C. S. and Ewerton, I. N. (2019). Networks for Cyberactivism and their Implications for Policymaking in Brazil. In *Crowdsourcing: Concepts, Methodologies, Tools, and Applications*, pages 1363–1378. IGI Global.

Ferreira, C. A. A. and Pena, F. G. (2020). O uso da tecnologia no combate ao Covid-19: uma pesquisa documental/the use of technology in the combat of Covid-19: a documentary research. *Brazilian Journal of Development*, 6(5):27315–27326.

Garg, L., Chukwu, E., Nasser, N., Chakraborty, C., and Garg, G. (2020a). Anonymity preserving iot-based Covid-19 and other infectious disease contact tracing model. *IEEE Access*, 8:159402–159414.

Garg, S., Bhatnagar, N., and Gangadharan, N. (2020b). A Case for Participatory Disease Surveillance of the Covid-19 pandemic in India. *JMIR Public Health Surveill*, 6(2):e18795.

Giulio, G. M. D., Vasconcellos, M. d. P., Günther, W. M. R., Ribeiro, H., and Assunção, J. V. d. (2015). Percepção de risco: um campo de interesse para a interface ambiente, saúde e sustentabilidade. *Saúde e Sociedade*, 24:1217–1231.

Huynh, T. L. et al. (2020). The Covid-19 risk perception: A survey on socioeconomics and media attention. *Econ. Bull*, 40(1):758–764.

Leung, G. M. and Leung, K. (2020). Crowdsourcing data to mitigate epidemics. *The Lancet Digital Health*, 2(4):e156–e157.

Venkatesh, V., Morris, M. G., Davis, G. B., and Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS quarterly*, pages 425–478.

Wolf, L. J., Haddock, G., Manstead, A. S., and Maio, G. R. (2020). The importance of (shared) human values for containing the Covid-19 pandemic. *British Journal of Social Psychology*, 59(3):618–627.