

Measuring the Impact of COVID-19 Restrictions on Mobility: A Real Case Study from Italy

Claudia Cavallaro, Armir Bujari, Luca Foschini, Giuseppe Di Modica, and Paolo Bellavista

Abstract—When COVID-19 first struck the provinces of Northern Italy in early 2020 (especially in Lombardy and in Emilia-Romagna), the conditions there made it a perfect storm. The virus outbreak spread with an unusual violence (in the period from late February to April 2020), with a catastrophic toll in terms of human deaths. Taken by surprise, Italy mandated a complete nation-wide lockdown, successively resorting to ministerial decrees alleviating and postponing the restrictions.

Now more than ever, there is an increased awareness on ICT used to combat the pandemic. In this article, we present a quantitative analysis evidencing the impact of restrictions on mobility. To this end, we rely on a vehicular mobility dataset confined in the downtown area of Bologna, Italy. Pursuing the objective, we propose a modified version of a state-of-the-art data mining algorithm, allowing us to efficiently identify and quantify mobility flows. The proposal, if combined with additional data sources, could allow for a fine-grained and timely decision making, combating the pandemic.

Index Terms—Big data, COVID-19, pattern mining, vehicular mobility.

I. INTRODUCTION

GLOBALLY, as of June 2021, there have been around 175 million confirmed cases of COVID-19 and more than 3.8 million deaths [1]. Up until the time of this writing, Italy has 4.25 million confirmed COVID-19 cases, while the death count amounts to 127 thousand. At the time of the so-called first pandemic wave (March 2020), in order to face the uncontrollable infection rate, the Italian government implemented a three-month nation-wide lockdown that, on a progressive basis, imposed stay-at-home orders (24/7), travel restrictions among provinces, social distancing, suspension of all business activities not connected with food and drugs. Because of the relaxation of restrictions enacted by the government in May 2020, since the end of summer 2020, the infection rate has dramatically increased (second pandemic wave) and hit a serious mark, forcing the authorities to enforce new restrictive measures. To face potential new waves and local outbreaks, the national government did not opt for a new strict nation-wide lockdown, rather, it resorted to prime minister decrees (DPCM) that from time to time, and region

by region, exacerbated or released the restrictions according to the seriousness of the situation. A more precise timeline of the issued decrees is depicted and discussed later on in Section IV.

Because of the ascertained dangers of the COVID-19 virus to transmit by means of respiratory droplets, imposing restriction on people's mobility is one of the main actions governments are undertaking to limit the virus spread. Tracking population mobility, on the one hand, is paramount to assess the efficacy and effectiveness of governments' measures; on the other one, it provides useful information to study and model the cause-effect dynamics that are triggered each time new restrictive measures are taken. Mobile phone records collected and owned by telecommunication providers (carriers) provide exactly the right kind and volume of information necessary to track human movements [2], assess presences and population density [3], devise mobility patterns [4] and eventually forecast future movements. On this hot topic, several collaborations are ongoing among carriers, governments and research communities [5], [6].

In this article, we present a quantitative measurement study, analyzing the impact of restrictions enacted by the Italian government during the first half of 2020. To this aim, we rely on a vehicular mobility dataset confined to the downtown area of the metropolitan city of Bologna, Italy, evidencing the impact that restrictions had on mobility flows. In specific, we start by providing a preliminary analysis quantifying the overall trend of vehicular mobility, characterizing the employed dataset. We then present and discuss a modified version of a state-of-the-art data mining algorithm, *Apriori* [26], later on used to quantify and assess vehicular flow patterns in a fast and reliable manner. The timely identification and measurement of vehicular flows becomes relevant and would allow governmental organizations to apply fine-grained and timely policies in an informed way. This approach, if combined with additional data sources, could provide invaluable insights to organizations.

The article is organized as follows: Section II discusses state-of-the-art techniques on trajectory mining, concluding with a high level overview of the proposed algorithmic approach. Next, Section III presents and characterizes the dataset, while providing a preliminary quantitative analysis, evidencing a first, high level overview, on the impact of restrictions on mobility during the lockdown period in Bologna, Italy. Section IV presents and discusses the algorithmic approach used to quantify vehicular flow evolution in time. Following, in Section V the results of the analysis are discussed according to the timeline of decrees as mandated by the Italian government. Finally, in Section VI the conclusions are drawn.

Manuscript received March 8, 2021; revised June 21, 2021; approved for publication by Christos Verikoukis, Guest Editor, September 16, 2021.

C. Cavallaro is with the National Institute for Nuclear Physics (INFN), CNAF section, Bologna, Italy, email: claudia.cavallaro@cnaif.infn.it

A. Bujari, L. Foschini, G. D. Modica, and P. Bellavista are with the University of Bologna, Department of Computer Science and Engineering, Italy, email: {armir.bujari, luca.foschini, giuseppe.dimodica, paolo.bellavista}@unibo.it

A. Bujari is the corresponding author

Digital Object Identifier: 10.23919/JCN.2021.000034

Creative Commons Attribution-NonCommercial (CC BY-NC).

This is an Open Access article distributed under the terms of Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided that the original work is properly cited.

II. BACKGROUND

In the past twenty years, the widespread of sensing-enabled mobile devices, along with the advancement of technologies for location acquisition, has fuelled a strong interest of the research community around the study of movements of both individuals and vehicles. From the analysis of the trajectories followed by people during their daily activities, useful information may be derived for a number of applications requiring real-time responses or timely decision making such as road traffic and tourism planning, to name a few [7]. In this paper, we apply trajectory mining techniques to a large data set containing traces of vehicle movements in order to derive trajectories followed by vehicles during the first wave of COVID-19 pandemic, and assess whether and how restrictions impacted on the overall vehicle mobility.

The huge, diverse and sometimes uncertain data provided by location-acquisition technologies calls for effective and efficient mining technique to build precise trajectories out of sequences of points characterized with spatio-temporal information. In his systematic survey, Zheng [8] identifies the main stages characterizing the pipeline of activities on which the paradigm of *trajectory data mining* is grounded. Relevant stages identified in the surveys are: (i) Trajectory pre-processing, where procedures are enforced to polish/structure the data in preparation for next stages; (ii) trajectory index and retrieval, devoted to indexing of data in support of efficient querying operations; (iii) trajectory pattern mining, tasked with the identification of the category of patterns that can be discovered from a single trajectory or a group of trajectories; (iv) trajectory classification, that aims to differentiate among trajectories (or its segments) according to different status (e.g., motions, transportation modes, human activities). Herein, we are interested in algorithms devoted to trajectory pattern mining and, more specifically, to sequential patterns detection from mobility data.

In this context, clustering and frequent sequential patterns are two of the main trajectory pattern mining techniques. Trajectory clustering aims at grouping similar trajectories into clusters. Similarity among trajectories is denoted by the “distance” between their respective feature vectors. Several distances and similarity measures can be found in the literature. Notable examples of proposed metrics and distance definitions are dynamic time warping (DTW) [9], longest common subsequence (LCSS) [10], Fréchet distance [11] and Haversine distance [12]. In [13], Wang *et al.* address the problem of robustness of some common measures of trajectories similarity. Mining frequent sequential patterns mainly consists in analyzing multiple trajectories in the aim of finding a certain number of moving objects that travel a common sequence of locations in a similar time interval.

In the following, we discuss state-of-the art techniques relevant to our study along with a critical review of the proposals. Concluding, is a high level overview of the proposed approach and of its merits with respect to the state-of-the-art.

A. Related Work

In [14], Zygoras and Gunopulos propose a technique to detect frequent traffic patterns, referred to as corridors, that could be used in transportation, and more broadly, by municipalities for city-wide planning purposes. In this work, GPS trajectories are discretized using a grid-based approach, applying the latent Dirichlet allocation (LDA) model [15] to extract frequent corridors. Successively, a hierarchical clustering algorithm is applied to each frequent set using a DTW-based approach to compute the distance between two patterns. Finally, the resulting corridor is selected from the candidate one’s by minimizing the principle of minimum description length (MDL). The approach is validated on real datasets collected from taxi trips in the city of Porto and by bus in Dublin.

Bicocchi *et al.* [16] analyze cellular data records to explore urban mobility patterns in Milan and Turin (Italy) metropolitan areas. They detect similar paths in order to propose frequent rides. Common mobility routines are identified through an extension of the LDA model, performed to develop a travel recommendation system for multiple users.

In [17], Crociani *et al.* address the problem of automatic lane detection in heterogeneous pedestrian flows, adopting an unsupervised clustering method. Based on the DBSCAN algorithm, they use a more precise distance function thanks to the angular distance between the vectors and the pedestrian speed considered. Khan *et al.* [18] develop different crowd analysis techniques and segmentation approaches based on the K-means algorithm used to cluster all similar flow vectors. In [19], the authors propose a novel hierarchical clustering technique to detect common trajectories, in which the similarity among tracks is measured by the longest common subsequence (LCSS).

Novel parallel versions of the LCSS algorithm are developed in [20], with a tool on distributed and shared memory for data analytics in bioinformatics. A parallel implementation of flow detection is also presented in [21], in which the Haversine distance is used for comparison, with the aim of extracting meaningful information from large datasets and suggesting places of interest in real time.

Buchin *et al.* [22] consider the problem of detecting commuting patterns by grouping the sub-routes of certain neighboring trajectories using the Fréchet distance. The Fréchet distance is among the most appropriate measures for the distance between continuous curves, in particular in its approximation for polygonal curves called discrete Fréchet distance (DFD). Devogele *et al.* [23] describe a faster variant of DFD, which includes filtering and pruning processes. They also improve DFD accuracy, balancing precision and CPU time reduction.

Rolim *et al.* [24] identify frequent movement patterns of trajectories in Santa Catarina - Brazil. The trajectories are segmented using the MDL principle and then clustering them, using the Fréchet distance. In [25], the authors focus on the problem of corridor detection from large GPS trajectory datasets. By discretizing the dataset, they build three different strategies through the *Apriori* algorithm and refine the obtained result via Radius Neighbors Graph.

B. A Novel Approach

From a thorough analysis of the state-of-the-art, it emerges that most authoritative works rely on well-known and consolidated spatial data mining techniques to devise sequential pattern detection strategies. Some proposals resort to a discretization of the geographical space which allows them to apply methods like trajectory transformation on matrices or on graphs [14], [25]. These class of techniques require a careful study on the discretization step to be adopted, which if not properly set, could result in useful information getting lost, having a negative impact on the precision of the final output. Furthermore, clustering methods adopted in [14], [17], [18], [24] revolve around the manual setting of the problem input parameters (e.g., a distance threshold or a pre-determined number of static clusters) which may strongly affect the goodness of the final outcome. Finally, proposals that rely on the calculus of the distance between pairs of points to assess the similarity of trajectories [22]–[24] are computationally heavy. Should the dataset grow in size - for instance, in the case that a larger geographical area needs to be investigated or a higher precision is requested - these approaches are impractical.

In this article, we propose a different approach that manages to achieve good results, in a fast and efficient way, without having to face the above mentioned issues. In order to identify common frequent trajectories out of a dataset of GPS points, we make use of *Apriori*, a data mining algorithm proposed by Agrawal et al. [26]. *Apriori* was initially proposed in the marketing context, used to determine which set of products (itemsets) are most often bought together by customers. In our work, we borrow the *Apriori* approach and exploit the similarities of *shopping basket-vehicle* and *item-road* to quickly detect which roads have been crossed more frequently than others, by different vehicles, in a considered time slot. Specifically, we developed a bottom-up approach to determine which paths (corridors) are close to each other or overlap in the same road section, with no need to apply a comparison between pairs of trajectories or compute the distances between points.

In addition, quasi-unified sampling strategies are often required to calculate similarity between trajectories, thus introducing errors and information loss. Our approach extracts reliable statistical data, as it is not sensitive to either noise or to the sampling frequency (in particular, it is sufficient that the vehicle registers at least one point for each different road traveled), allowing us to quickly and easily quantify the patterns of travel over the road network. More details on the algorithmic approach and its adoption are discussed in Section IV.

III. VEHICULAR MOBILITY DATASET

The vehicular mobility dataset employed in this study is provided by an Italian nation-wide car insurance company under the project IPODAMO [27] where the authors are involved as scientific advisors. The dataset used herein is geographically scoped to a part of the metropolitan area of Bologna, Italy, depicted in Fig. 1. In the following, we provide

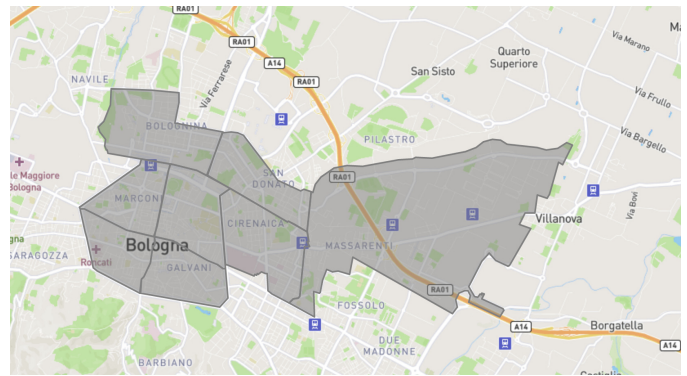


Fig. 1. The gray area denotes the area of interest comprising the city center of the metropolitan area of Bologna, Italy.

a brief overview of the dataset along with a preliminary quantitative evaluation of the vehicular data.

A. Description

The source data extraction and filtering process provides an entire vehicle trip whenever a vehicles' position is found to be inside the area of interest shown in Fig. 1. If this event is verified the trip is reconstructed from a temporal buffer and provided to us in textual format. Note that although the area of interest comprises only parts of Bologna, Italy, it is quite representative as it comprises the downtown of Bologna, a major and vibrant city in Italy, along with an arterial road such as the A14 freeway which is the second largest freeway in Italy.

The historical dataset covers a period from January 2020 to June 2020, including the first lockdown phase announced in Italy (March 2020) and successive ones. Vehicles are equipped with a blackbox, a multi-purpose and autonomous on-board devices with sensing and communication capability, generating data when pre-determined events occur e.g., vehicles engine turned on/off. Monthly datasets comprise daily trips, delimited by a start and stop latitude/longitude along with some additional attributes used to uniquely identify the trip. Below is provided a list with relevant attributes qualifying a trip:

- Trip id: Numerical identifier of the particular trip;
- Device id: Numerical identifier of the on-board blackbox, which changes periodically, each month;
- Start date/time: Time and date when this trip initiated i.e., vehicle engine is turned on;
- End date/time: Time and date when this trip terminated i.e., vehicle engine is turned off;
- Start latitude/longitude: GPS coordinates denoting the place when this trip initiated;
- End latitude/longitude: GPS coordinates denoting the place when this trip terminated;
- Average velocity: Average speed in *km/h* of the trip;
- Vehicle characteristics: A formatted string containing information (if present) regarding the vehicle type and model.

Along with the daily trip header dataset, an additional monthly dataset is provided which contains intermediary trip

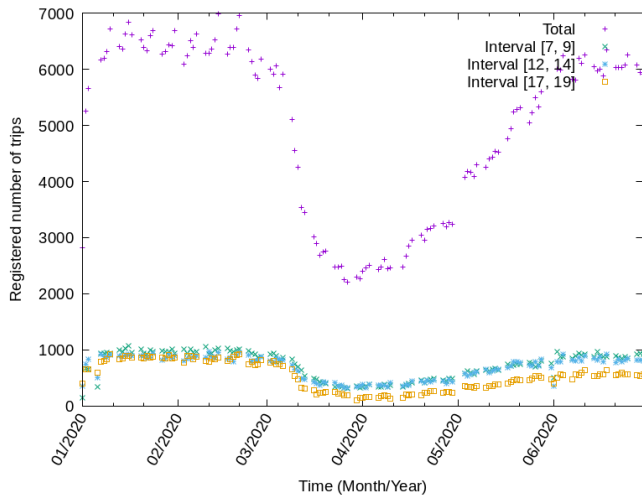


Fig. 2. Daily number of trips evolution in time. The chosen intervals denote common daily rush hours.

samples. An intermediary trip sample is generated whenever a vehicle instantly (de)accelerates, and more generally whenever one of the following conditions is satisfied: either vehicles traverses 1 km or 60 s have gone by since the last position was announced. In specific, a trip detail contains the following information:

- Trip id: Numerical identifier of the particular trip;
- Device id: Numerical identifier used to denote the on-board blackbox. The identifier is changed periodically, each month, for privacy concerns;
- Date/time: Time and date when this sample was taken;
- Latitude/longitude: GPS coordinates denoting the place the sample was generated;
- Type of street: Whether the road is freeway, urban street etc.

As a final note, it is possible to reconstruct the entire vehicle trip information thanks to the trip identifier attribute contained in both traces.

B. A Preliminary View

Fig. 2 quantifies the number of trips evolution in time. To compute this information, we rely on the trip identifier available in the monthly datasets and prior to counting, we exclude the weekend days in order to eliminate any transitory effects due to movements from and to the city. Instead, Fig. 3 shows the normalized number of trips with respect to the busy month of January when the pandemic had not yet invaded the daily lives of the western world. Another complementary view is provided in Fig. 4 showing the empirical cumulative distribution function (ECDF) of the trip frequency distribution for the different months present in the dataset. In particular, to compute this data, we rely on the terminal identifier value which does not vary inside a month timeframe.

In both Fig. 2 and Fig. 3 one can immediately observe the sudden decrease in the volume of traffic starting from March, when the first decree announcing the lockdown measure took effect, reaching its lowest peak in the month of April. The

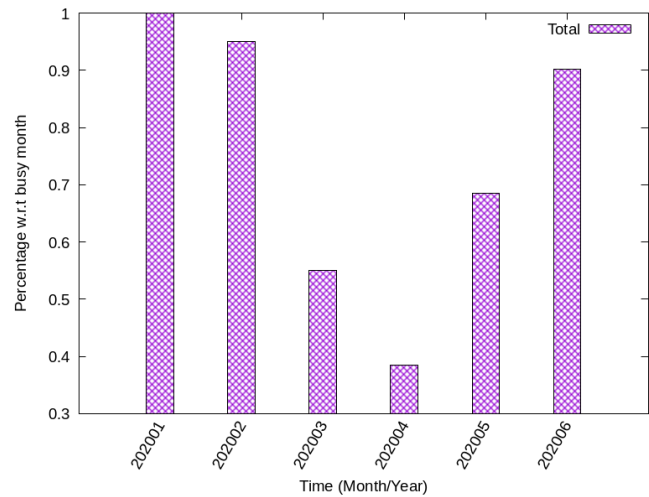


Fig. 3. Histogram showing the evolution in time of the trips contained in the dataset.

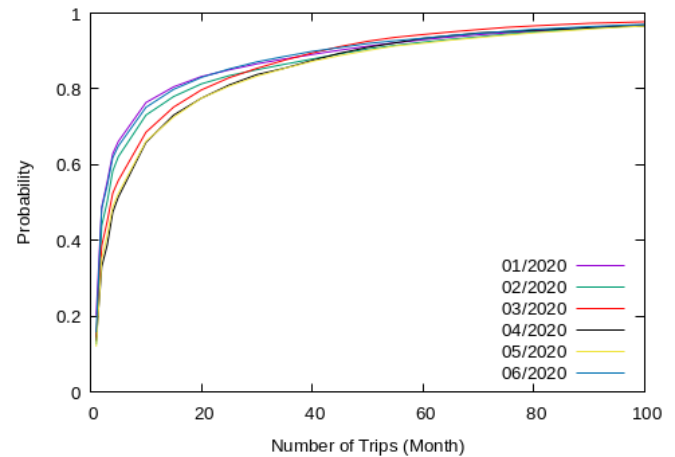


Fig. 4. Empirical cumulative distribution function of trip frequencies per vehicle present in the dataset.

decrease in traffic volumes is noticeable even in typical rush hours (Fig. 2) coinciding with activities such as going to or returning from work/school. From this point onwards, the volume steadily increases, until new and more relaxed measurements come into effect. A similar trend is shown in the trip frequency distribution (Fig. 4) evidencing an impact in normal activity reduction i.e. frequent trips in the interval $[0, 50]$ with April and May having the lowest values, reaching proportionally comparable values in the higher frequencies.

In the following, we propose a new type of analysis, allowing us to efficiently analyze and quantify vehicular mobility patterns over a road network.

IV. MINING VEHICULAR FREQUENT PATTERNS

In this section, we discuss an efficient, modified version of *Apriori* used to extract information on vehicular flow distribution. Following, we present a precise timeline of restrictive measures taken into effect in Italy in the period between

Algorithm: Modified Apriori

```

Input :  $D$ : trajectory database
           $min\_sup$ : minimum vehicular density
          threshold
Output: Set  $E$  containing most frequent corridors
/* compute level-1 information <road,
  # of vehicles> */
 $L_1 = level\_1\_roads\_in(D)$ 
/* iterate until no transaction
  available */
for  $k = 2; L_{k-i} \neq \emptyset; k = k + 1$  do
  /* compute candidate subsets
    filtered by topological order
    and minimum support */
   $C_k = \{c \subseteq subset(L_{k-i}, D) \mid |c| = k \wedge filter(c)\}$ 

  foreach  $vehicleId\ t \in D$  do
     $C_t = subset(C_k, t)$ 
    /* get the subsets of  $t$  that are
      candidates */
    foreach  $candidate\ c \in C_t$  do
      |  $c.cont++$ 
    end
  end
   $L_k = \{c \in C_k \mid c.cont \geq min\_sup\}$ 
end
return  $E = \bigcup_k L_k$ 

```

March-July 2020, later on used to discuss the effects they had on mobility.

A. Modified Apriori

In our context, a vehicular trajectory consists of a sequence of ordered GPS points through which it is possible to reconstruct a path, as a sequence of traversed roads, on an actual road network. To build this synthesized trajectory database, a pre-processing step is involved aimed at reconstructing the reverse geo-coding information from single latitude/longitude samples, which are then grouped to produce the actual path traversed by a vehicle. At first, the grouping is limited to the single road elements.

To obtain the reverse geo-coding information, we exploit the k-nearest neighbor (KNN) algorithm already available in Apache Sedona, a cluster computing framework for spatial data processing [28]. The algorithm relies on the Bologna road network and on the Haversine distance metric to compute the nearest road the sample belongs to. For each road element, referred to as an item in the original version of *Apriori*, it is therefore possible to compute the number of vehicles traversing it in a considered time interval. In *Apriori* terminology, this number is referred to as *support*.

The modified *Apriori* takes as input the above synthesized trajectory dataset which consists of a multiset data structure with entries containing information on: (i) Vehicle trip identifier, (ii) list of roads it traversed during a trip and (iii) a list of

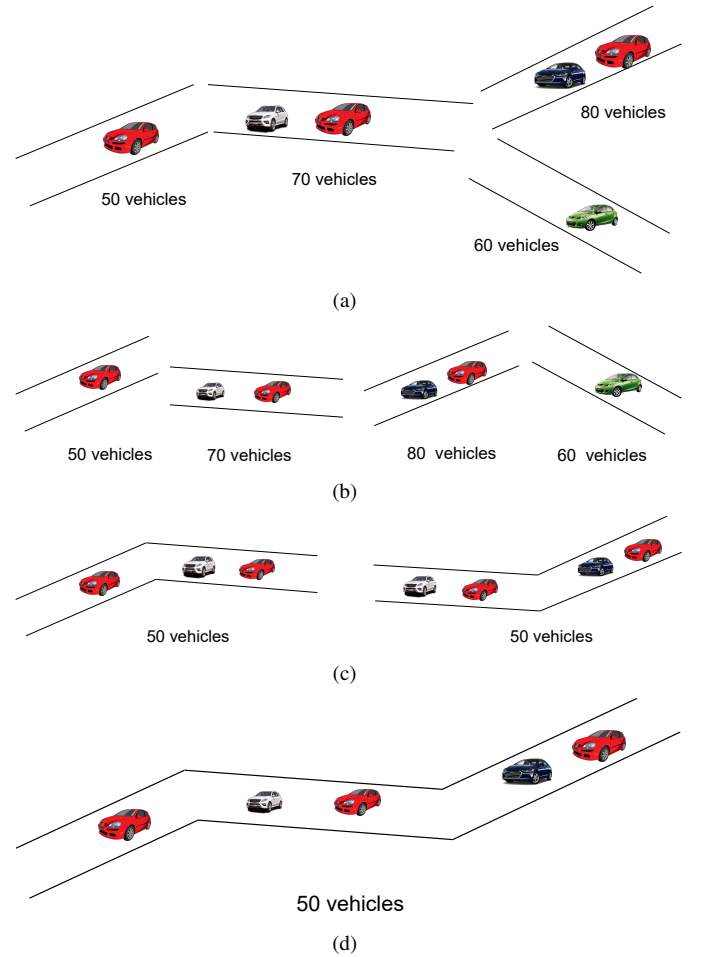


Fig. 5. A minimal running example. Vehicles of different colors represents vehicular flows and their size which might extend to the whole or parts of the road network: (a) Example road topology along with vehicles at time interval t , (b) Level 1 output containing roads satisfying minimum support, (c) Level 2 output containing roads satisfying minimum support, and (d) [Level 3 output containing roads satisfying minimum support.

timestamps, denoting the time lapse of the samples belonging to it. Another important piece of input data is the minimum vehicular density threshold (min_sup), denoting the minimum number of vehicle occurrences that candidate vehicular paths (corridors) must support.

The output consists of a multiset containing ordered sequences of roads or corridors satisfying a minimum support criterion. It is noteworthy to point out, that the time dimension could be omitted from the problem formulation by filtering the input accordingly, simplifying the procedure.

The algorithm (*Modified Apriori*) performs a depth first search of potential frequent subsets of size k , one level at a time, until no further extensions are possible, meaning no corridors of size $k + 1$ with sufficient support (min_sup) can be found. To provide meaningful results, and fruitful next level candidates, the generated subsets C_k at level k are filtered also by considering a topological map order. Indeed, without this criterion the algorithm could due to e.g., errors in the source data, naively combine roads which are far away, qualifying them as corridors. Once suitable candidates have been identified, they are assessed whether the minimum

support condition is satisfied and depending on the outcome are included or excluded from the current level (L_k) identified corridors.

The algorithm is based on the principle of anti-monotonicity, that is, if a set of entities is frequent, then all its subsets are also frequent, but if an itemset is not frequent, then the sets containing it are not frequent either. Through a bottom-up approach we obtain a sequence of frequent roads according to a minimum support criterion.

In Fig. 5 is shown a minimal running example explaining the algorithm's *modus operandi*. In particular, Fig. 5(a) presents a simplified road network along with some vehicular flows depicted by the colored vehicles. From an algorithmic viewpoint, starting at level 1 with a *min_support* of 50 vehicles, all frequent patterns of size 1 with a minimum density of 50 are computed (Fig. 5(b)). At level 2 (Fig. 5(c)) the algorithm discovers all viable pairs of frequent roads (set of candidate corridors of size 2). Among the candidates, those satisfying the *min_sup* criterion are selected and become viable corridors for the next iteration. Note that, at this step, the algorithm pruned the candidate subsets comprising the corridor with the green car. For simplicity, we assume all the shown subsets are viable ones, satisfying the topological order criteria. Similarly, at level 3, all viable subsets of size 3 (corridor of size 3) are built starting from subsets of size 2. A join procedure is applied, generating all the level 3 entries satisfying the criteria.

Since no further candidates are left to assess, the algorithm terminates and returns all identified corridors of size 1 to 3.

B. Restrictions Timeline

In Italy, the epidemiological situation due to Coronavirus disease 2019 exploded in March 2020, although the state of emergency had already been announced on January 31st, 2020, with the occurrence of some COVID-19 outbreaks in specific areas of northern Italy (Codogno in Lombardy region, Vo' Euganeo in Veneto region) in February.

March 9, 2020 is the date of the first national DPCM which transformed Italy into a restricted area, by imposing a temporary suspension of a number of business and public activities which could favor the spread of the COVID-19, such as swimming pools, gyms, school and university classes, museums, cinemas, and recreational centers.

On March 22nd, a new decree was issued jointly by the Minister of Health and the Minister of the Interior which prohibited individuals from moving or travelling by public or private means of transport, except for proven work or health needs, or absolute urgent matters. On April 1st, a new DPCM was adopted which extended the effectiveness of the previous provisions to April 13rd.

With the April 10th DPCM, the suspension of teaching activities was extended up to May 3rd. Shops remained closed, except for groceries, pharmacies, tobacconists, newsagents, and petrol stations.

After almost two months of lockdown, the DPCM of April 27th announces the first relaxation of restrictions for construction companies, manufacturing, mining, automotive, textile, wood, glass and wholesale industries. Among the

provided authorizations, the decree also allowed travels to meet relatives, provided that the prohibition of gathering and distancing is respected, and respiratory protection is used.

On May 18th, Italy restarts. The new decree allowed most business activities to open. Citizens could move within their own region and frequent public places, provided that at least one-meter social distancing is kept, and protective masks are worn.

V. TRAFFIC FLOW ANALYSIS

Herein, we present the results of the quantitative analysis performed exploiting the modified version of the *Apriori* algorithm presented in the prior section.

In Fig. 6, it is shown the vehicular traffic flow evolution over time. In this instance, the data are partitioned and studied in specific time intervals identifying common rush hours e.g., people travelling from/to work. On the x-axis, we report the timeline of restrictions implemented by the Italian government, while the y-axis reports the frequent patterns with a vehicular density greater than or equal to 10. The frequent pattern values denote the number of corridors with a minimum support of 10 as identified by the algorithm.

Up until March 2020 the vehicular flows are subject to periodic fluctuations. A rapid decrease is observed in March 9, 2020, when the lockdown was announced and enacted. Even though the state of emergency was declared in January 31, 2020, i.e., when the first localized outbreaks of Codogno and Vo' Euganeo were discovered, in our dataset no impact on the mobility patterns is noticeable.

A change in traffic patterns, along with a rapid decrease of the traffic flows, is noticeable in the beginning of March 2020, when a nation-wide lockdown was imposed. In this time span, up until May 2020, traffic flows are subject to some fluctuations reaching a minimum after the April 10th decree, which announced the prolongation of the national lockdown to May 4th, 2020. Note that during the entire lockdown period, citizens were not allowed to leave their home, except for strict necessities such as work and health reasons.

On May 4th, restrictions were relaxed and a steady increase of traffic flows can be observed, although its density did not reach the level observed prior to the lockdown. Though people were allowed to move, a large part of the population started working at home (*smart working*).

In Fig. 7 a map-based comparison of frequent roads before and after the lockdown took effect is shown. Herein, we consider the level 1 output of the algorithm covering an entire day. It is evident that the mobility index drastically changed, with little to no activity in the downtown area of the city, and traffic is mostly concentrated in the freeway and city surroundings.

Fig. 8 provides an in-depth analysis on the traffic flow distribution along the restrictions' timeline. To this end, different minimum supports have been adopted, varying the traffic density in the ranges shown in the bottom-left part of the table. The chart groups frequent roads based on density, with darker colors representing the busiest roads.

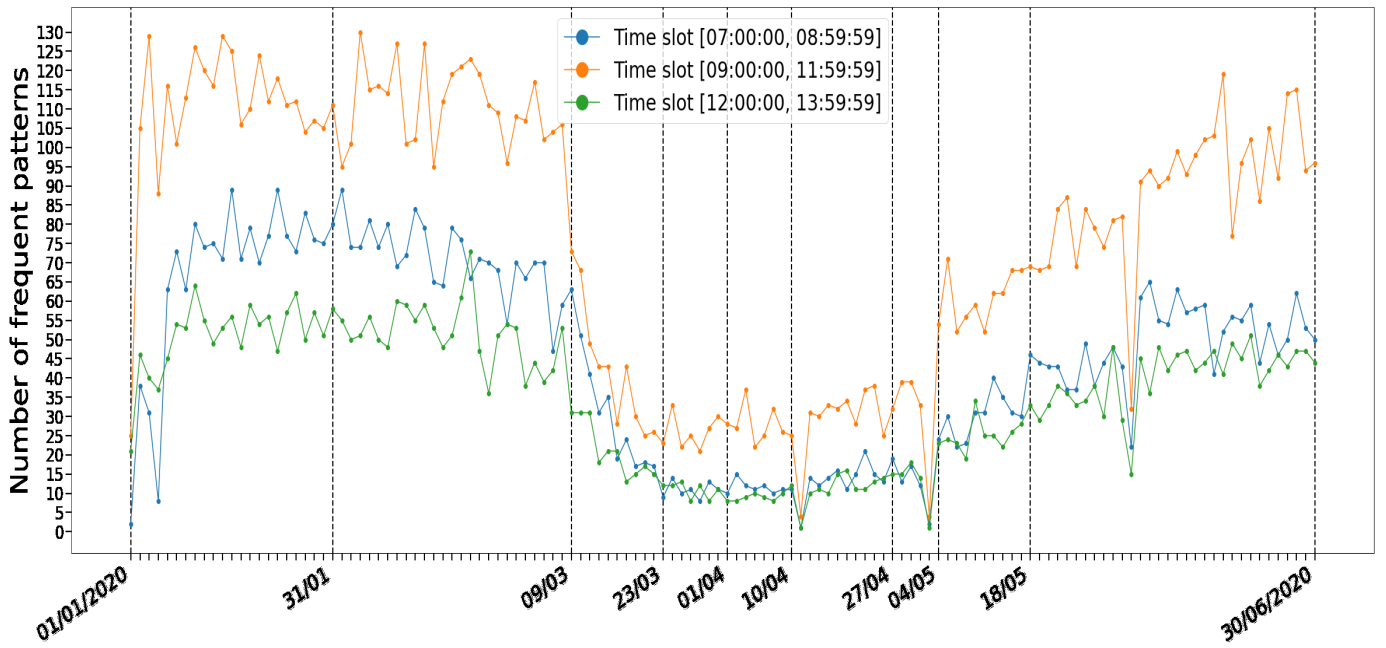


Fig. 6. Traffic flow evolution in time with vehicular density ≥ 10 . On the x-axis the dates where restrictions were announced and took effect.

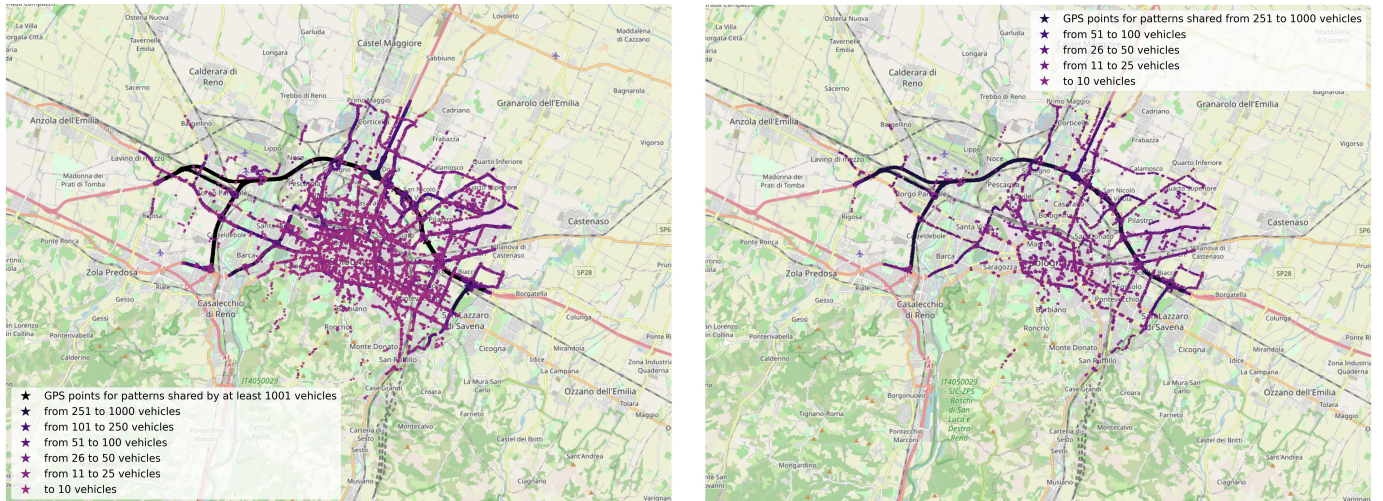


Fig. 7. Comparison of traffic flows between January 31st, 2020 (left) and March 23rd, 2020 (right) for the metropolitan city of Bologna.

The density pattern attribute was discretized into 7 classes according to the number of vehicles crossing the same road on the same date.

At first glance, one can observe the evident change: a decrease of traffic density and a slow (yet steady) increase prior, during and after the lockdown period respectively. Overall, roads subject to higher traffic volumes seem to be less impacted over time. This is reasonable as activity in these roads, although restrictions are in place, can be related to goods transportation from/in Bologna and other essential services.

Low-to-mid frequencies, e.g., roads with traffic levels from 1–100, seem to be the most impacted, reaching their lowest values during the lockdown period followed by a steady increase up to June 2020, where a quasi return to normal is

observed. These roads can be typically associated with urban traffic and everyday movements of people.

For completeness, Table I reports some of the most frequent roads identified by the prior analysis and their traffic evolution along the restriction timeline. The numbers in blue denote the inflection points, i.e., a positive change in the amount of traffic sustained therein.

From what we have shown, it is evident that the pandemic has deeply impacted mobility patterns in the downtown area of Bologna. While now things seem to be returning to a degree of normality, the new coronavirus is still amongst us. This fast has raised the population awareness in understanding the importance of ICT tools for combating the virus in intelligent ways. The timely identification and measurement of vehicular flows becomes relevant and would allow governmental

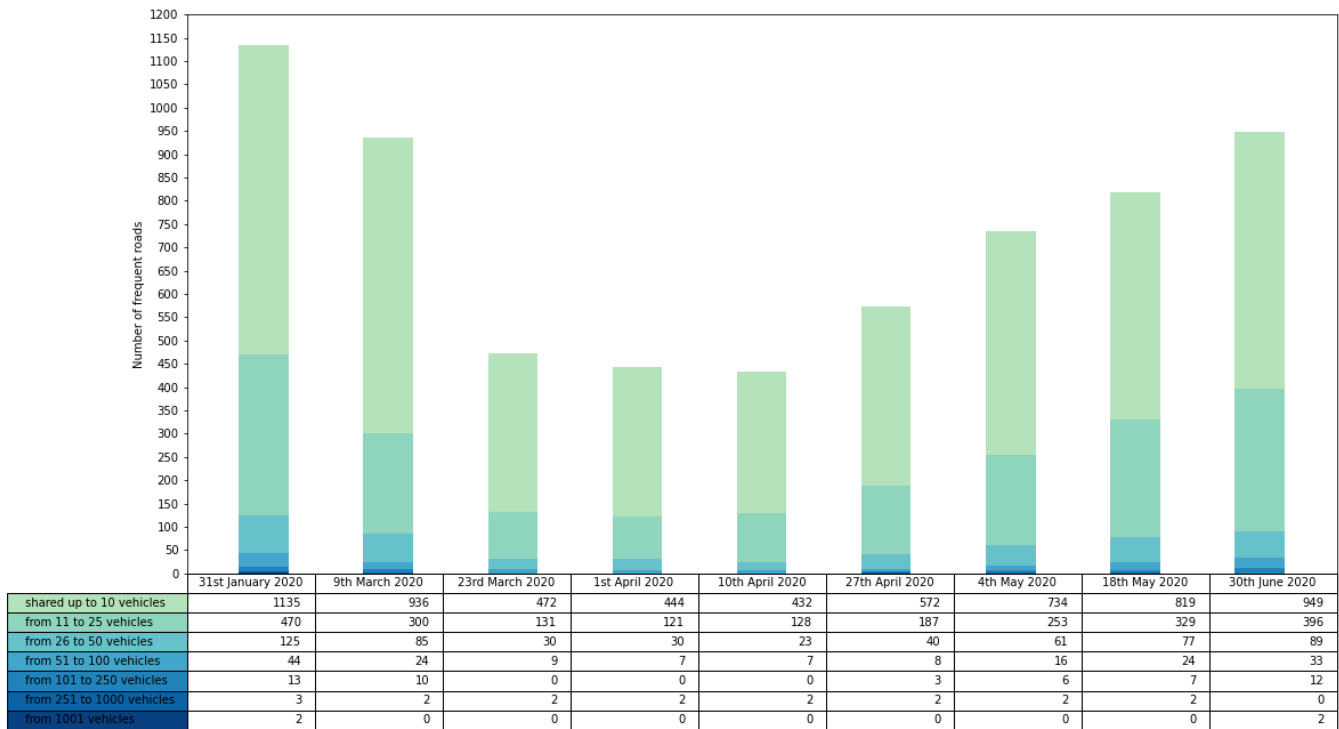


Fig. 8. Distribution of the detected corridor densities. In tabular form are reported the number of roads subject to a particular traffic density on that specific date.

organizations to apply fine-grained and timely policies in an informed way.

VI. CONCLUSION

The pandemic has heavily conditioned our life. More than a year after COVID-19, the world is still struggling and trying to keep the pace. While the first pandemic wave caught most governments by surprise, resorting to desperate initial measures, ICT use plays a crucial role, allowing for a timely and fine-grained decision making.

In this paper, we presented a study measuring the impact restrictions had on the mobility of the metropolitan area of the city of Bologna, Italy. Specifically, the proposed analysis focuses on the evolution of the traffic patterns and density along the restriction timeline. As a further paper contribution, a modified version of the *Apriori* algorithm was presented and discussed. The algorithm helped us quantify and reliably assess vehicular flow patterns.

Currently, we are working on extending this study, by implementing an online, distributed MapReduce-like approach to *Apriori*, which could help improve the flow analysis time. The idea is to distribute the computation, relying on spatial partitioning techniques, among cluster nodes, merging individual outputs to form a coherent global view at e.g., street level. Different levels of aggregations and processing pipelines might be provisioned, depending also on the scope of the analysis.

REFERENCES

- [1] World Health Organization (WHO), "Coronavirus disease (COVID-2019) situation reports." <https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports>, 2020. Last Accessed: June 14, 2021.
- [2] H. Barbosa *et al.*, "Human mobility: Models and applications," *Physics Reports*, vol. 734, no. 6, 2017.
- [3] P. Deville *et al.*, "Dynamic Population Mapping using Mobile Phone Data," *National Academy Sciences*, vol. 111, pp. 15888–15893, Sep. 2014.
- [4] L. Alessandretti, P. Sapiezynski, V. Sekara, S. Lehmann, and A. Baronchelli, "Evidence for a conserved quantity in human mobility," *Nature Human Behaviour*, vol. 2, no. 7, pp. 485–491, June 2018.
- [5] Y. Kang *et al.*, "Multiscale dynamic human mobility flow dataset in the U.S. during the COVID-19 epidemic," *Scientific Data*, vol. 7, no. 390, pp. 1–13, 2020.
- [6] S. Lai *et al.*, "Effect of non-pharmaceutical interventions for containing the COVID-19 outbreak in China," *Nature*, vol. 585, no. 7825, pp. 410–413, 2020.
- [7] S. Jiang, J. Ferreira and M. C. Gonzalez, "Activity-based human mobility patterns inferred from mobile phone data: A case study of Singapore," *IEEE Trans. Big Data*, vol. 3, no. 2, pp. 208–219, 2017.
- [8] Y. Zheng, "Trajectory data mining: An overview," *ACM Trans. Intelligent Syst. Technol.*, vol. 6, no. 29, 2015.
- [9] D. J. Berndt and J. Clifford, "Using dynamic time warping to find patterns in time series," in *Proc. ACM SIGKDD*, 1994.
- [10] M. Vlachos, G. Kollios, and D. Gunopulos, "Discovering similar multidimensional trajectories," in *Proc. IEEE ICDE*, 2002.
- [11] H. Alt and M. Godau, "Compute the Fréchet distance between two polygonal curves," *International J. Comput. Geometry Applications*, vol. 5, no. 1, pp. 75–91, 1995.
- [12] R. W. Sinnott, "Virtues of the Haversine," *Sky Telescope*, vol. 68, p. 158, 1984.
- [13] H. Wang, H. Su, K. Zheng, S. Sadiq, and X. Zhou, "An effectiveness study on trajectory similarity measures," in *Proc. Australian Computer Society ADC*, 2013.
- [14] N. Zygouras and D. Gunopulos, "Corridor learning using individual trajectories," in *Proc. IEEE MDM*, 2018.

	31/01	09/03	23/03	01/04	10/04	27/04	04/05	18/05	30/06
Via dell'Industria	102	94	57	50	48	62	83	90	112
Via Francesco Zanardi	103	78	24	21	18	43	53	75	79
Via Giovanni Gozzadini	122	84	42	39	42	53	58	74	105
Viale Palmiro Togliatti	123	89	31	30	33	48	59	79	94
Viale Mahatma Mohandas Gandhi	125	103	34	30	41	50	72	80	92
Via Nuova Bazzanese	130	105	32	45	29	50	68	83	103
Viale Vittorio Sabena	149	112	52	40	44	63	77	95	125
Via Cristoforo Colombo	181	160	64	62	58	92	131	150	183
Via Stalingrado	205	146	66	52	58	91	95	144	169
Viale Europa	232	162	58	63	66	87	111	133	184
Autostrada A13	244	148	96	89	98	116	118	146	234
Asse Attrezzato Sud-Ovest	256	180	62	59	63	80	142	152	203
Svincolo Bologna Arcoveggio	265	186	89	77	95	114	128	129	229
Viale Roberto Vighi	273	180	85	72	90	115	172	191	202
Autostrada A14	1149	799	398	382	372	464	599	686	1121
Tangenziale Nord RA1	1213	958	433	396	428	525	766	817	1115

TABLE I

MOST FREQUENT ROADS VEHICULAR DENSITY EVOLUTION DURING THE FIRST HALF OF 2020.

- [15] D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent Dirichlet allocation," *J. Machine Learning Research*, vol. 3, pp. 993–1022, Mar. 2003.
- [16] N. Bicchieri and M. Mamei, "Investigating ride sharing opportunities through mobility data analysis," *Pervasive Mobile Computing*, vol. 14, pp. 83–94, 2014.
- [17] L. Crociani, G. Vizzari, A. Gorrini, and S. Bandini, "Identification and characterization of lanes in pedestrian flows through a clustering approach," in *Proc. AI*IA*, 2018.
- [18] S. D. Khan, G. Vizzari, S. Bandini, and S. M. Basalamah, "Detecting dominant motion flows and people counting in high density crowds," *J. WSCG*, vol. 22, no. 1, pp. 21–30, 2014.
- [19] S. D. Khan, S. Bandini, S. M. Basalamah, and G. Vizzari, "Analyzing crowd behavior in naturalistic conditions: Identifying sources and sinks and characterizing main flows," *Neurocomputing*, vol. 177, pp. 543–563, 2016.
- [20] R. Shikder, P. Thulasiraman, P. Irani, and P. Hu, "An OpenMP-based tool for finding longest common subsequence in bioinformatics," *BMC Research Notes*, vol. 12, 2019.
- [21] C. Cavallaro, G. Verga, E. Tramontana, and O. Muscato, "Eliciting cities points of interest from people movements and suggesting effective itineraries," *Intelligenza Artificiale*, vol. 14, no. 1, pp. 75–87, 2020.
- [22] K. Buchin, M. Buchin, J. Gudmundsson, M. Löffler, and J. Luo, "Detecting commuting patterns by clustering subtrajectories," in *Proc. ISAAC*, 2008.
- [23] T. Devoegele, L. Etienne, M. Esnault, and F. Lardy, "Optimized discrete Fréchet distance between trajectories," in *Proc. ACM SIGSPATIAL*, 2017.
- [24] M. Silva, V. Rolim, F. Flamarion, C. Filho, and F. Braz, "A method for identifying patterns of movement of trajectory sets by using the frequency distribution of points," in *Proc. IARIA*, 2017.
- [25] C. Cavallaro and J. Vitrià, "Corridor detection from large GPS trajectories datasets," *Applied Sciences*, vol. 10, no. 14, 2020.
- [26] R. Agrawal and R. Srikant, "Fast algorithms for mining association rules in large databases," in *Proc. VLDB*, 1994.
- [27] "IPPODAMO: Interactive planning platform for city district adaptive maintenance operations," 2020. Online Available: <https://www.ippodamoproject.it>.
- [28] A. Sedona, "A cluster computing system for processing large-scale spatial data," 2020. Online Available: <https://sedona.apache.org/>



Claudia Cavallaro received the Ph.D. in Computer Science in 2021 at the University of Catania. Currently she holds a Postdoc position at Centro Nazionale per la Ricerca e lo Sviluppo nelle Tecnologie Informatiche e Telematiche (CNAF section of National Institute for Nuclear Physics INFN), in Bologna, and she is an Adjunct Professor at the University of Turin, Italy. Her research interests include big data and geo-localized data analysis.



Armir Bujari received the Ph.D. degree in Computer Science from the University of Bologna, Italy, in 2014. He is currently an Assistant Professor of Computer Science and Engineering with the University of Bologna, Italy. His research interests are primarily focused on edge/fog computing applications in the industrial domain, and next generation architectures and networks.



Luca Foschini received the Ph.D. degree in Computer Science Engineering from the University of Bologna, Italy, in 2007. He is currently an Associate Professor of Computer Engineering with the University of Bologna. His interests span from integrated management of distributed systems and services to mobile crowd-sourcing/-sensing, from infrastructures for the deployment of Industry 4.0 solutions to fog/edge cloud systems. Finally, he is serving as secretary of the ComSoc CSIM TC, and as voting member and awards committee chair for the IEEE ComSoc EMEA board.



Giuseppe Di Modica graduated from the University of Catania, Italy. In 2005 he received the Ph.D. in Computer Science and Telecommunication Engineering from the University of Catania, Italy. He is an Assistant Professor with the Department of Computer Science Engineering at the University of Bologna, Italy. He has participated in many regional, national and European R&D projects. His research interests include cloud computing and multi cloud, edge/fog computing, big data, Internet of things, industry 4.0, SOA, microservices, business process

management.



Paolo Bellavista received the M.Sc. and Ph.D. degrees in Computer Science Engineering from the University of Bologna, Italy. He is currently a Full Professor of Distributed and Mobile Systems with the University of Bologna. His research interests include from pervasive wireless computing to online big data processing under quality constraints and from edge cloud computing to middleware for industry 4.0 applications. He serves on several editorial boards, including IEEE Communications Surveys and Tutorials (Associate EiC), ACM CSUR, JNCA

(Elsevier), and PMC (Elsevier). He is the Scientific Coordinator of the H2020 BigData Project IoTwins.