This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3135054, IEEE Access

IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2017.DOI

Effective Response to Natural Disasters: a Data Science Approach

GHULAM MUDASSIR [1], EVANS ETRUE HOWARD[1], LORENZA PASQUINI[1], CLAUDIO ARBIB[1], ELISEO CLEMENTINI [2], ANTINISCA DI MARCO [1], AND GIOVANNI STILO [1]

[1]Department of Information Engineering, Computer Science and Mathematics, University of L'Aquila, Italy (e-mail: name.surname@univaq.it)
[2]Department of Industrial and Information Engineering and Economics, University of L'Aquila, Italy (e-mail: name.surname@univaq.it)

Corresponding author: Ghulam Mudassir (e-mail: ghulam.mudassir@graduate.univaq.it).

This work was partly supported by *Territori Aperti* (a project funded by Fondo Territori, Lavoro e Conoscenza CGIL, CISL and UIL) and by *SoBigData-PlusPlus* H2020-INFRAIA-2019-1 EU project, contract number 871042. The open data used in the evaluation comes from *opendata.regione.abruzzo.it*.

ABSTRACT Natural disasters can cause widespread damage to buildings and infrastructures and kill thousands of living beings. These events are difficult to be overcome both by the populations and by government authorities. Two challenging issues require in particular to be addressed: find an effective way to evacuate people first, and later to rebuild houses and other infrastructures. An adequate recovery strategy to evacuate people and start reconstructing damaged areas on a priority basis can then be a game changer allowing to overcome effectively those terrible circumstances. In this perspective, we here present DiReCT, an approach based on *i)* a dynamic optimization model designed to timely formulate an evacuation plan of an area struck by an earthquake, and *ii)* a decision support system, based on double deep Q Network, able to guide efficiently the reconstruction the affected areas. The latter works by considering both the resources available and the needs of the various stakeholders involved (e.g., residents social benefits and political priorities). The ground on which both the above solutions stand was a dedicated geographical data extraction algorithm, called "GisToGraph", especially developed for this purpose. To check applicability of the whole approach, we dovetailed it on the real use-case of the historical city center of L'Aquila (Italy) using detailed GIS data and information on urban land structure and buildings vulnerability. Several simulations were run on the underlining network generated. First, we ran experiments to safely evacuate in the shortest possible time as many people as possible from an endangered area towards a set of safe places. Then, using DDQN, we generated different reconstruction plans and selected the best ones considering both social benefits and political priorities of the building units. The described approaches are comprised in a more general data science framework delved to produce an effective response to natural disasters.

INDEX TERMS Data Science, City Reconstruction Planning, Decision-support System, Deep Reinforcement Learning, Evacuation Planning, Flow Model, Geographical Information Systems, Natural Disaster, Network, Optimization Model.

I. INTRODUCTION

Almost all of us have sometimes found ourselves in a situation where a large number of people gathered together in a particular place: spectators at concerts or sports events, students in school premises, commuters in railway and metro stations, employees in large office buildings. To ensure people safety and comfort, not only a careful design of pedestrian facilities is required, but also effective strategies for crowd management. In particular, strategies for an efficient facility evacuation are of paramount importance in case of emergency, be it natural (earthquakes, tsunamis, hurricanes) or caused by human activities (fire, gas leak, bomb threat etc.). Recent examples are the Australian bushfires [1], [39] killing at least 34 people between June 2019 and May 2020; Hurricane Eta, November 2020, killing at least 150 people in Central America [4], [5]; flash floods killing more than 150 people in Afghanistan in August 2020 [3], [22]; earthquake and tsunami killing a total of 117 people in Greece and Turkey in October 2020 [2] (not to mention animals). These and many more examples identify the need for

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3135054, IEEE Access

developing good crowd management and *emergency evacuation procedures*. The study of pedestrian and evacuation dynamics is very complex, especially due to the nonlinear interactions among the (large) number of people involved. Those interactions also include, in general, hardly predictable psychological factors influencing human behavior, and also such external factors as the layout of pedestrian passageways is to be taken into account.

Managing the situation after a disaster is also quite complex due to the damages suffered by public/private buildings and such infrastructures as bridges and roads [44]. Recovering damaged infrastructures is lengthy and costly [20]. For example, after Hurricane Katrina in 2005, the recovery phase took five years and needed one billion dollars for completion; for the Hyogoken-Nanbu earthquake in Japan, just the highway reconstruction process took 20 months. In post-disaster situations, authorities are forced to make quick and efficient decisions to overcome the problems. Effective management of resources, funding and time are then required to reduce shocks on economy, but decision-makers and public authorities often lack of comprehensive *post-disaster recovery plans* to restore normal life at a fast pace [27]. In fact, full recovery from a disaster always requires long time, and this very badly affects local communities from both a social and an economical viewpoint. This challenging phase was defined by Contreras et al. [14] as a *complex multidimensional, long-term process involving planning, financing, decision-making, and reconstruction*.

Among the major difficulties generally faced by public decision-makers, one finds the problem of defining a *balanced* recovery plan, that is, one that considers all formal and informal requirements and makes sure to reconstruct damaged areas in reasonable time [34]. To mention one critical aspect, the plan should, on the one hand, consider immediate needs concerning the vulnerability of buildings; on the other hand, it should take into account resilience factors to tackle disasters in the future [33] and everything should be seen in the perspective of the relevant budget-to-time issues. Another aspect often not tackled by public decision-makers deals with the social benefits that local communities can obtain from the recovery plan implementation. In fact, societal impact and benefits are of course different for different plans, so that the plan itself can play a vital role in all post-disaster phases. A traditional reconstruction approach is not suitable for this purpose, and we need a reconstruction framework that encompasses social benefits for the communities affected by the disaster [50].

Analyzing and observing data science patterns is not only a beneficial practice for business, but can also improve the efficiency and effectiveness of emergency and disaster management organizations. The availability of such mobile devices as smartphones [51], as well as of social media, helps monitor disasters with real-time information and can allow a rapid and accurate response. Whether through crisis mapping or event simulation, data science is pioneering new emergency management methods. It can for instance analyse the information gathered in real-time from a community in order to assist those in need or those looking for friends and relatives lost in a disaster. By accessing to real-time information on the most involved areas, emergency personnel can sensibly reduce search time and intensify recovery efforts. Working along with professional insight and satellite imagery, data science indicated trends and started practices that saved actual lives and proved to be truly effective for emergency management [31]. For example, during the 2010 earthquake in Haiti the US Marine Corps used an open-source software created by Ushahidi for gathering information useful to rescue the victims [18].

Meanwhile, we observe that data science [48] can play a vital role during the preparation of evacuation plans that guide people in safe places and reconstruction/recovery plans oriented to social benefits. This role is a direct consequence of combining multiple fields and sources considered by data scientists in their mission of interpreting data for the advantage of decision making. In urban applications, the availability of data from different sources (like municipalities, civil protection, environmental agencies, volunteered geographic information, social networks, and so on) represents a big opportunity, only marginally explored by now, to build an integrated knowledge base useful for evacuation and recovery plans. Heterogeneous sources and big data make it difficult to reach this integration without resorting to specialized tools. It is not sufficient that municipalities make available specific data-sets on their institutional websites, as it is often the case. Information should in fact be processed and transformed into services that can be of immediate use to citizens and emergency operators.

Considering the above discussion and scenarios, we recognised the lack of a methodology to integrate heterogeneous data and deploy effective services for emergency management. For this reason, we here propose DiReCT (Disaster Reconstruction City), a data science-based methodology able to cope with different post-disaster emergencies.

To do so, our approach merges the available information into semantically enriched graph data structures. We note that both the initial evacuation and reconstruction plans can be dealt with by the proposed methodology. Specifically, we will focus (i) on ways, methods and tools to help citizens reach pre-identified safe areas even in highly dynamic scenarios, and (ii) on the definition of viable plans to reconstruct buildings within appropriate time spans and cost budgets.

These plans fulfill such major constraints as those relevant to completion time, financing, political priorities and physical dependencies, and also include the consideration of social benefits for the affected community.

Hereafter we highlight the main contributions and novelties proposed by this paper:

- the work presents the first data science-based methodology (DiReCT) able to cope with different post-disaster emergencies;

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3135054, IEEE Access

- we propose an improved [1] version of a crowd evacuation optimization model to appropriately manage the congestion that may occur during evacuation;
- we show the feasibility and applicability of the DiReCT methodology and of its methods to complex scenarios based on real-world data;
- we experimentally validate the proposed methodology by measuring its performance on both evacuation and reconstruction use cases;
- we finally show the approach at work on the city centre of L'Aquila (Abruzzo, Italy), gathering real data from publicly available official sources of information[2]

The paper is organised as follows: in Section II we present and discuss related work. The proposed methodology is described in Section III, and its application to a real case study is outlined in Section IV. Experimental results are discussed in Section V. In SectionVI, staring from the presentation of approach limitations, we present main future work. Conclusions are finally drawn in Section VII.

II. RELATED WORK

The presentation of related work is organized in three parts: in the first we discuss data extraction; in the second, evacuation planning in natural disasters; the third part is finally focused on the discussion of reconstruction planning in post-disaster management.

Data extraction: Nowadays, one of the most important application areas of Geographical Information Systems (GIS) is specialized in the analysis of transport networks and related problems: those systems are identified as GIS-T. The first work one can highlight is a book [19] by Manfred M. Fischer that discusses improvements to the traditional transportation data model, which are needed to support advanced network analysis in a ground transportation context. Deelesh Mandloi et al. [30] present an object-oriented data model to represent a multi-modal, indoor/outdoor transportation network in an urban area; this model is particularly suitable for network analyses regarding route planning and navigation in the network. As reported by [6], [35], GIS applications in disaster management are progressively becoming very useful tools that help process the related emergency activities and reduce the extremely critical duration of emergency management operations.

Evacuation planning in natural disaster: In the evacuation planning literature one can identify three main, distinct yet interrelated, research streams. The first one focuses on the empirical study of pedestrian behavior and crowd dynamics. In contrast, a second stream is concerned with the development of mathematical models to describe movement and interactions of pedestrians as realistically as possible [46]. Finally, a third stream of research uses optimization-based methodologies to develop models that determine optimal evacuation plans or design solutions [8]. Most of the research falls under the first two categories. Several review articles discuss the empirical research and modeling of pedestrian and evacuation dynamics. Schadschneider et al. [42] provide a summary of the empirical studies and theoretical models developed that far, and give two examples of possible application of such a research. Helbing and Johansson [24] give a similar overview and, additionally, discuss research issues on panic and critical crowd conditions. Schadschneider and Seyfried [43] investigate the quantitative data on pedestrian dynamics for the calibration of evacuation models. They considered the implications for cellular automata models. Papadimitriou et al. [37] assess two different topics of research, namely: route choice models and crossing behavior models, the latter studying how pedestrians cross the street under different traffic conditions. Gwynne et al. [23] classify evacuation models based on the nature of the model application, the enclosure representation, the population perspective, and the behavioral perspective. Zheng et al. [52] distinguish seven methodological approaches: cellular automata, lattice-gas, social-force, fluid dynamics, agent-based, game-theoretic models, and experiments with animals. They also look at the possibility of modeling heterogeneous individuals, the scale of representation, whether time and space are discrete or continuous, whether a normal situation or an emergency is assumed, and the typical phenomena that the model can represent. In addition, Duives et al. [16] identify eight motion base cases and six self-organizing crowd phenomena that a simulation model should be able to reproduce. Furthermore, they look at ten other model characteristics, such as the ability to simulate pressure in crowds and the computational requirements of the model to assess model applicability. Their classification distinguishes between cellular automata, social-force, activity-choice, velocity-based, continuum, hybrid, behavioral, and network models. Kalakou and Moura [26] present a general overview of models from different research areas to analyze the design of pedestrian facilities, while Lee et al. [29] focus on models for the evacuation of ships. Finally, Bellomo et al. [11] focus on the mathematical properties of models for pedestrian behavior. The third category of research has received less attention in the literature. Our work falls under the third category. Using an optimization-based methodology, specifically a dynamic maximum flow model, we also model possible congestion that can occur along the arcs of the underlying network.

Post-disaster reconstruction planning: The reconstruction planning literature is characterized by works that use mathematical models and/or machine-learning algorithms. Opricovic et al. [36] develop a multi-criteria decision model (MCDM) for the analysis of post-disaster planning. The MCDM selects the best reconstruction plan on the basis of previously defined parameters. The methodology makes also use of fuzzy multi-criteria optimization and encompasses qualitative as well as non-quantifiable variables. Also Goujon et al. [21] propose a multi-criteria decision support model based on the Myriad tool. Their work defines the priorities and evaluation of reconstruction projects also consid-

[1] With respect to previous publications such as [25]
[2] Open data comes from opendata.regione.abruzzo.it

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3135054, IEEE Access

ering social population needs. Tavakkol et al. [45] propose an entropy-based framework for the post-disaster decision-making process that helps manage collapsed buildings and road closures. Using the data collected, the Authors define a prioritization strategy concerning damaged bridges and buildings. Post-disaster data were collected with the help of CCTV, cellphones, and cameras, while prioritization strategies are evaluated by Pareto analysis. The whole methodology was then evaluated by extensive experiments using data related to the Nisqually earthquake.

Ghannad et al. [20] propose a post-disaster recovery model that produces and defines a reconstruction plan based on the priority of damaged project/facilities. Priorities are assigned considering socioeconomic factors of the affected communities, as well as time and cost. The method is based on an Analytical Hierarchy Process (AHP) for decision making and on an optimization model for resource allocation. The evaluation of the model is conducted using data from counties (called parishes). Similarly, Rodriguez in [40] proposes a model for reconstructing roads and removing debris in post-disaster situations. They use a two-stage Steiner Tree Model (to check roads to be reconstructed first) and scheduling algorithms for restoring roads and crews assignments. For verification, they implemented the proposed method using data from the 1994 Northridge California earthquake.

In [17], Eid et al. developed an innovative decision framework by adopting an agent-based approach. They set short-term and long-term redevelopment goals by considering three-dimensional vulnerabilities of communities like social, economic, and environmental ones. For this purpose, they use residential agents, economic agents, and state disaster recovery agents (SDRC). The main purpose of SDRCs is to evaluate the recovery plan and prioritize objectives, via aggregated equations, after every simulation.

The main difference between our model and the literature commented so far, is that ours considers, in reconstruction planning, the social benefits of local communities, includes dependencies among reconstruction units, cost, time, damage level, and also prioritizes every unit according to the intentions of political authorities.

III. METHODOLOGY

This section presents our data science methodology to effectively respond to natural disasters, focusing mainly on evacuation and reconstruction plans. We can distinguish three main parts: i) the construction of a network (or graph) based on the GIS data (GisToGraph algorithm) (subsection III-A), ii) the dynamic flow modeling, and the solution to evacuation planning (subsection III-B), and iii) the modeling of the reconstruction planning and its solution built by employing double deep Q-learning network (DDQN) approach (subsection III-C). 'DiReCT' loop is used for feedback to GisToGraph algorithm after evacuation and reconstruction. The proposed methodology is overall illustrated in Fig. 1, where the rounded rectangles represent the processes, i.e, the set of operations for the different algorithms, and the rectangles represent the outputs of the various operations.

As one can understand from Fig. 1, graph generation is the preliminary step for evacuation and reconstruction phases. We extract data from several sources, and so obtain an enriched graph that associates the area or city considered with all the information necessary to evacuate citizens (e.g., street lengths and widths, population of each building, etc.) and to reconstruct the city (e.g., budget and time for reconstruction, building priority as per political assessment, etc.).

Section III-A details the graph generation, Section III-B reports on the evacuation planning approach, and Section III-C details the reconstruction planning.

A. GISTOGRAPH ALGORITHM

The available information in the form of GIS data relating to a damaged city needs to be transformed into a network structure that incorporates useful information for subsequent phases of evacuation and reconstruction. The resulting network, called *Enriched Undirected Graph* (EUG), helps the effective manipulation of information in all such phases.

The EUG of the damaged area is built in two steps (see also [25]): (i) collection of the required information of the city that is needed for evacuation and reconstruction in disaster management (such data could come from shapefiles or other city's data repositories); (ii) by using the *GisToGraph* algorithm we transform these input data into the nodes, edges, and attributes of the EUG.

The generated EUG allows an efficient search of the optimal routes in case of evacuation during the relief phase and represents the basis of a decision support system during the recovery and reconstruction phase. The choice of the optimal path does not consider only the distance between the nodes, but it also takes into account other factors like the risk of buildings, street capacities, and the number of people who need to be evacuated. Additionally, the EUG is enriched with other information like the damage level of buildings and roads that are useful for reconstruction planning.

Geographical data describing a city are mostly available in some specific spatial data format (e.g. shapefiles). The latter describe geographical features as the base geometric elements (Points, Linestrings, and Polygons), flanked by textual attributes. The object-oriented modeling approach allows us to organize collections of simple features: the crossroads as points, the streets as lines, the sets of points of interest as points or polygons, and the sets of census areas as polygons, respectively. The points of interest (e.g., buildings or safe areas) need to be located at starting or ending points in a route. Census areas are not part of the network but they are useful to estimate the number of people in each building or the number of people in a particular area of the city.

Those are the classes used by the GisToGraph algorithm, to perform data extraction:

- class *City*: is the main class which is used to store all data related to the area of interest, actually, it could be considered the whole city or just a region of it; the limit of this area is represented by defined boundaries.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3135054, IEEE Access

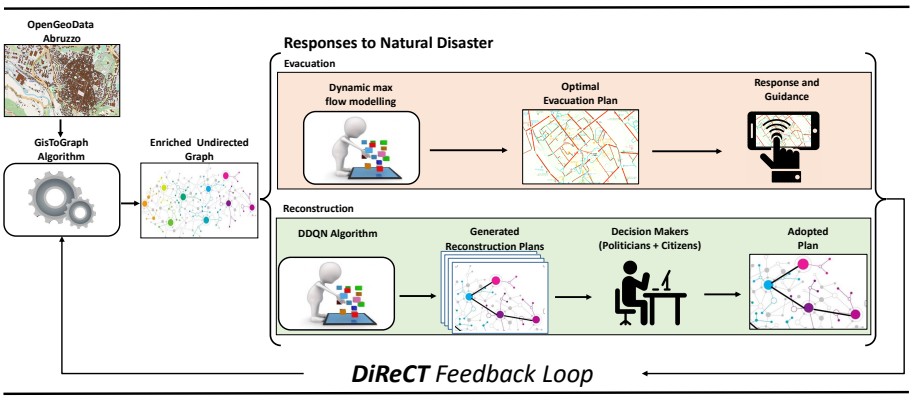

FIGURE 1: Overview of the proposed methodology (the rounded rectangles represent automatic processes, the rectangles show data, while icons without rectangles indicate user interaction phases)

In particular, its attributes are an identification name, a geometry that represents the boundaries of the area of interest, a list of all the crossroads, streets, points of interest, and census areas that compose it, and a Spatial Reference System (SRS) used for all spatial data. It also contains all the methods to load its components from shapefiles.

- class *Crossroad*: represents the intersection of two or more roads. It has only two attributes: an ID to identify it and its position expressed in geometric coordinates.
- class *Street*: represents every road that goes from one crossroad to another without passing from another crossroad. It has a unique ID, an optional name that is by default the name of the original way. This class contains its geometry represented as a line string, a dictionary to collect all its other possible properties, and a list of the points of interest that can be accessed from that street. The dictionary is populated accordingly to the use case (e.g. its width, its danger level, and so on).
- class *Point of interest*: represents the specific location in the city that someone may find useful or interesting. It could be a building, a bus stop, or a gas station, depending on the study case. It has an ID, a category (e.g., "School," "Safe Area," "Strategic Building"), a geometry that could be either a polygon describing the real boundary of the point of interest or a point that approximates its original position, a dictionary to add all its other possible properties depending on the case of study (for example the number of students for a school or the number of people a safe area can host), the ID of the associated street and the position of its entrance from the street expressed in geometric coordinates. The two latter associations are not always explicitly available from datasets since the position of the point of interest is not directly in the road network. For this reason in most cases, we adopt a straight method to find the association by calculating the street situated at the shortest distance from the point of interest and determining the nearest point on the street network that represents to access the building. More sophisticated methods can be devised if more precise information is available, such as building street numbers.
- class *Census Area*: is a geographic region defined to take a census. Every city could have one or more census areas that intersect with its geometry. Each census area has an ID, a geometry represented by a polygon, an estimate of the number of people in that area, and a list of the points of interest in that area.

As anticipated, besides this main data, it is necessary to add other textual information which are required for disaster management to start evacuation and reconstruction planning in an emergency situation, in particular:

- road length and width: are useful to understanding how many people a road can contain at a given time;
- people in buildings: to estimate the number of people in any particular building during disaster;
- time of reconstruction: time to reconstruct any building after disaster occurred;
- building state: state of a building under reconstruction;
- state of reconstruction;
- the required amount for reconstructing the building and the amount actually granted;
- practicability of the building;
- type of the buildings: useful to give priority to the buildings in the reconstruction phase.

The *EUG* is generated considering all the aforementioned information. Note that the EUG is *undirected* because we assume that pedestrians move in both directions, and it is a *multigraph* because the same crossroads could be the endpoints of different streets. The EUG can be used for both evacuation planning and reconstruction planning to find out damage buildings and roads. Table 1 reports all attributes we used in EUG, evacuation planning and reconstruction planning.

To improve the computational efficiency, we adopted a hierarchical approach: thus we defined two graphs, which are called the generic G_A and the detailed one G_D. In the de-

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3135054, IEEE Access

TABLE 1: Parameters Notations

Notation	Definition
V	Set of vertices v that represent reconstruction units (e.g, building, hospital, bridge).
E	Set of edges e=(v_1, v_2) that represent connections between two reconstruction units, namely v_1 and v_2.
s_v	$s_v \in \{0, 1\}$ represents the status of v, if $s_v = 1$ then v is not damaged or its reconstruction is completed.
T_v	the time needed to reconstruct v. If $s_v = 1$, T_v is zero.
c_v	Established cost for the reconstruction of v. $s_v = 1$, c_v is zero.
p_v	political priority of v: $p_v \in \{1, 2 \cdots 10\}$ where 1 represents *Low Priority* and 10 represents *Highest Priority*.
b_v	Number of people that take direct advantage of v, $b_v \in N$. If v is a residence building, b_v represents the number of people having the residence in v.
A_v	Represents all possible actions A of agent on unit v within an environment
$S(v)$	Number of people that take directly and indirectly advantage after the reconstruction of v, $S(v) \in N$.
$d(v_1, v_2)$	Function returning the distance between v_1 and v_2 reconstruction units, v_1 and $v_2 \in V$xV. Such distance is calculated considering the minimum path connecting v_1 and v_2.
P	P represents the generated plan specified as an ordered list of reconstruction units $[v_1, v_2, ..., V_n]$
S_p	Social benefits of a plan P
T_e	Ending time of the reconstruction plan
G'	G'=(V', E') is the physical dependencies graph where $V' \subseteq V$ and E' is a set of edges $e' = (v'_1, v'_2)$ representing that v'_1 cannot be reconstructed before v'_2.
T	set of unit time slots where $T = \{0, 1, 2, ..., \tau\}$
y_i^t	state of cell $i \in V$ at time $t \in T$, that is, the number of persons that occupy i *at* t: this number is a known model parameter for $t = 0$ (in particular, $y_0^0 = 0$) and a decision variable for $t > 0$
n_i	capacity of cell i: it is a measure of the maximum nominal amount of people that i can host at any time (in particular, $n_0 \geq \sum_i y_i^0$); this amount depends on cell shape and size; if cells can be assumed uniform one can set $n_i = n, \forall i \in V, i \neq 0$
x_{ij}^t	how many people move from cell i to an adjacent cell j in $(t, t + 1]$: this gives the average speed at which the flow proceeds from *i to j*
c_{ij}	capacity of the passage between cell i and cell j: this is the maximum number of people that, independently on how many persons are in cell j, can traverse the passage in the time unit (independence on cell occupancy means neglecting system congestion: we will consider this issue later).

tailed one $G_D = (V, E)$, the set of nodes $V = (V_C \cup V_{POI})$ is given by the union of:

- V_C: nodes that represent all the crossroads in the city, they have a position in coordinates and all the other information saved in the first phase.
- V_{POI}: nodes thst represent all the points of interest (buildings) in the city, their position is not the original one but it is the point on the street, where they can be accessed. Also, these nodes contain all the information saved in the collection data phase.

Then, the set of edges $E = (E_S \cup E_{HS})$ is given by the union of streets E_S and half-streets E_{HS}, where the set of edges,

- E_S: represents all the streets in the city that connect the crossroad nodes V_C. This edge set contains all the necessary information about the streets in the city/area under consideration.
- E_{HS}: called *half-street*, represents the parts of streets that connect points of interests to each other or to a crossroad. They are found by splitting original streets E_S with respect to access points to a point of interest (for example the entrance of a building). Every half street is saved with respect to the original street by inheriting its information.

Then, the generic graph is a more abstract one $G_A = (V_C, E_S)$, which can be seen as a subgraph of G_D, where the set of nodes are made of only the crossroads, V_C, and the set of edges are made of streets, E_S. Obviously, the detailed graph contains more nodes than the generic graph: computing a path between points of interest in G_D can be faster than exploiting G_A. In order to find a route between two points of interest, we can distinguish three steps: *departure, search and arrival*. In the *departure* step, the detailed graph is explored in order to move from the starting point of interest (the origin) to the nearest crossroad that brings us closer to the destination. In the *search* step, the algorithm uses G_A to find a route from the crossroad found in the *departure* step to a crossroad that is as close as possible to the destination (i.e., one of the endpoints of the street that contains it). Finally, in the *arrival* step one moves from crossroads to the destination, hence it is necessary to use G_D. Using this method, the CPU time of the routines used in the elaboration of the evacuation plan can be substantially reduced.

Similarly for reconstruction planning, we need time to reconstruct the unit, cost of rebuilding that unit, the damage level of buildings/units, number of people inhabiting the unit, political priority, and physical dependencies (damage roads/bridges). All the required information needed for the reconstruction planning model are not directly available as external data. For this reason, a second step completes the input data with the following information:

- reconstruction *time* of a node v is determined by following the approach described by M. Polese et al. in [38];
- *politically priority* is integrated into the dataset according to type of the buildings [15];
- on the basis of the obtained units' status, we are able to generate the physical dependencies graph G'.

In the output graph, we represent damaged buildings with red color nodes and fine/reconstructed ones with green color nodes. Similarly, damaged roads and bridges are represented by red dotted lines and fine/reconstructed ones by black lines, respectively.

B. EVACUATION PLANNING

After the generation of the EUG, we discuss our strategy for evacuation planning. The emergency evacuation plan is developed by consecutive steps: simulation and revision carried out following domain-specific rules. Basic steps require the definition of safe areas at nearby locations, safe routes to reach those areas, and a correct partition of the evacuees proceeding to safe areas. Modeling the plan is a

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3135054, IEEE Access

formal process that starts from all the relevant information (population and census, maps, regulations, etc.) output by the GisToGraph algorithm and ends up with structured data and software tools to access and use them.

Designing an evacuation plan encompasses the following actions:

- Determination of safe areas. There are two ways to find out safe areas in a city. One way is to take advantage of the ones built in the GIS files of land use. Another way is to use satellite imagery. The former method is more efficient for a planned city than the latter because land use (mapped by organizations like transportation, municipalities) is already known.
- Determination of paths. An important design action in the definition of an evacuation plan is finding out the safest (not necessarily the shortest paths) for providing humanitarian supply, closest facility, and deploying the search and rescue team.
- Ramification of people. Disaster management authority, police, a paramilitary force, army, NDRF during calamity emergencies do these works better than any other. They have good coordination and synchronization according to the situation. Before evacuation, we gather all the information about the capacity of the crowd, census data of inhabitants, and exit points of the incident area so that easy ramification of people could take place. Disaster managers command all the organizations working in search and rescue operations and make coordination among them. Soldiers and volunteers take the instructions from disaster managers and help evacuees by indicating the safe areas and helping them evacuate. There is no certain rule or algorithm applicable in efficient and effective distribution of people because when disaster triggers, it creates havoc at that place. So ramification of people should be done by accurate analysis and inspection of the situation.

To generate the emergency evacuation plan that minimizes the total evacuation time while maximizing the number of evacuees evacuated to safety, we adapted the linear optimization model originally developed by [10] for building evacuation. The model was then calibrated with respect to several parameters and re-scaled to the network of several orders of magnitude.

1) Dynamic maximum flow modeling

With reference to the earlier works of [13], [10] and [25] where a discrete-time network stock-and-flow model is devised to find, at increasing time slots τ, the maximum amount of people that can be evacuated within τ to a given set of safe areas. The model starts with a static oriented network $G = (N, A)$ obtained from the graph G_D generated by the GisToGraph algorithm by suitably embedding the city streets into a set N of elementary cells; the arcs in A connect geometrically adjacent cells in both directions. Cells may in general have different shapes or sizes: for the purpose of this work, it is important that every cell can approximately be traversed in a single time unit. Depending on size, the i-th cell has a capacity n_i equal to the maximum number of people it can host and, at any given time t, contains some number $y_i^t \leq n_i$ of people. Moreover, depending on street size, a limited amount $x_{ij}^t \leq c_{ij}$ of people can move in the unit interval $[t, t+1]$ from cell i to an adjacent cell j. Finally, depending on scenarios, the network G may consist of a number of maximal connected components: in each component, safe places collectively correspond to a single super-sink 0 with a capacity large enough to host all evacuees.

From each component of the static network G, a dynamic network $G_T = (N_T, A_T)$ is then constructed as the time expansion of G over a time horizon $T = \{0, 1, \ldots, \tau\}$, with:

$N_T = \{(i,t) | i \in N,\ t \in T\}$;

$A_T = A_M \cup A_H$, where A_M, the movement arcs, link (i,t) to $(j, t+1)$ for $(i,j) \in A$, and A_H, the holdover arcs, link (i,t) to $(i, t+1)$ for $i \in N$.

Using then x_{ij}^t and y_i^t as decision variables, the model assigns the initial cell occupancy, expresses flow conservation, and enforces the appropriate capacities (possibly considering congestion phenomena). Distinct models are formulated for different τ, with the objective of maximizing the total in-flow y_0^τ in the super-sink at time τ. One then seeks the least τ^* within which the totality of people can be evacuated from the endangered area: to reduce CPU time, τ^* is computed by logarithmic search. In this way, the method provides the decision maker with the Pareto-frontier of the conflicting objectives $\min\{\tau\}$, $\max\{y_0^\tau\}$.

Due to the linear structure of the model, large number of variables can be allowed in solution. Adding these variables can help improve model granularity by reducing space and time units. More importantly, it can also help approximate the non-linearities along the arc capacities. When c_{ij} is constant, we fail to model arc-congestion, which is a situation where the speed at which the system empties is a decreasing function of room occupancy y_i^t. A more accurate model of congestion requires arc capacity to be a concave decreasing function of room occupancy, see Figure 2. Hence the dynamic maximum flow model can be written as:

$$max \quad \{y_0\}_{t=1}^{\tau} \tag{1}$$

$$y_j - y_j^{t-1} - \sum_{i:ij \in A} x_{ij}^{t-1} + \sum_{i:ji \in A} x_{ji}^{t-1} = 0 \qquad j \in V, t \in T, t > 0 \tag{2}$$

$$0 \leq y_i^t \leq n_i \qquad t \in T, i \in V \tag{3}$$

$$y_i^{t-1} = u_i^{t-1} + v_i^{t-1} + w_i^{t-1}, \quad x_{ij}^t = \phi_{ij}^t + \chi_{ij}^t + \psi_{ij}^t \tag{4}$$

with u_i^{t-1}, v_i^{t-1}, w_i^{t-1} non negative and subject to upper bounds

$$u_i^{t-1} \leq n_i', \quad v_i^{t-1} \leq n_i'' - n_i', \quad w_i^{t-1} \leq n_i - n_i'' \tag{5}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3135054, IEEE Access

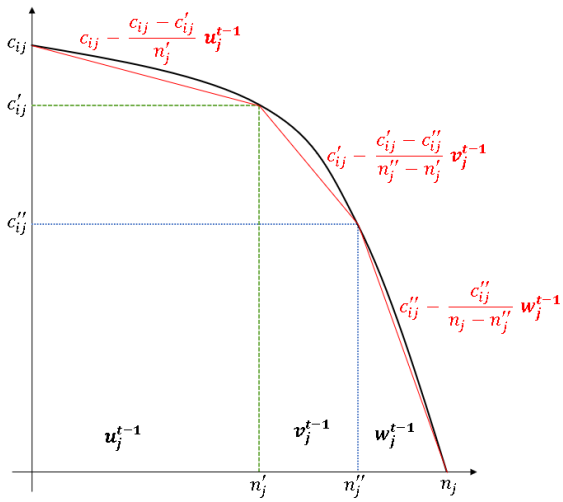

FIGURE 2: Linearization of a congestion curve.

$$
\begin{aligned}
0 \le \phi_{ij}^t &\le c_{ij} - \frac{c_{ij} - c'_{ij}}{n'_j} u_j^{t-1} \\
0 \le \chi_{ij}^t &\le c'_{ij} - \frac{c'_{ij} - c''_{ij}}{n''_j - n'_j} v_j^{t-1} \\
0 \le \psi_{ij}^t &\le c''_{ij} - \frac{c''_{ij}}{n_j - n''_j} w_j^{t-1}
\end{aligned} \tag{6}
$$

The consistency of the ϕ, χ and ψ variables with the x flow variables requires $\chi = 0 (\phi = 0)$ if ϕ (if χ) does not saturate its capacity. This is ensured, at optimality, by the properties of basic solutions. After rephrasing 6

$$
\begin{aligned}
0 \le \phi_{ij}^t, &\qquad \phi_{ij}^t + a_{ij} u_j^{t-1} \le c_{ij} \\
0 \le \chi_{ij}^t, &\qquad \chi_{ij}^t + a'_{ij} v_j^{t-1} \le c'_{ij} \\
0 \le \psi_{ij}^t, &\qquad \psi_{ij}^t + a''_{ij} w_j^{t-1} \le c''_{ij}
\end{aligned} \tag{7}
$$

where

$$
a_{ij} = \frac{c_{ij} - c'_{ij}}{n'_j}, \qquad a'_{ij} = \frac{c'_{ij} - c''_{ij}}{n''_j - n'_j}, \qquad a''_{ij} = \frac{c''_{ij}}{n_j - n''_j}
$$

and $a_{ij} < a'_{ij} < a''_{ij}$, we observe the following fact (that can be generalized to any piece-wise linear approximation of the congestion curve)

Equation 2 is the flow conservation law, which expresses the occupancy of cell j at time t as the number y_j^{t-1} of persons present at time $t-1$, augmented of those that during interval $(t-1, t]$ move to j from another cell $i \neq j$, less those that in the same interval leave cell j for another room $i \neq j$. Box constraints (3) and (6) reflect the limited hosting capability of the elements of G.

Variables ϕ, χ and ψ are introduced to linearize a non-linear capacity constraint that simulates link congestion. We in fact assume that the capacity c_{ij} of link ij is not constant, but decreases with the flow in the link. From Figure 2, in order to linearize the congestions on the arcs or passages between adjacent cells, we introduce constant parameters c, c', c'' and n, n', n'' in inequalities (5) - (7). The meaning of these parameters is shown in Figure 2: in particular, c_{ij} is the link capacity when no flow traverses the link. As the flow of people increases in the interval $[0, n'_j]$, the capacity linearly goes down to a value c'_{ij}, with gradient $\frac{c_{ij} - c'_{ij}}{n'_j}$, corresponding to the green-dotted lines. A similar linearization is defined for flows in $[n'_j, n''_j]$ shown with the blue-dotted lines with capacity c''_{ij} with slope $\frac{c'_{ij} - c''_{ij}}{n''_j - n'_j}$, and $[n''_j, n_j]$: in the latter case, the capacity is reduced to 0 (in fact, the link is blocked when n people try to cross it).

2) Parameter Setting

The model complexity increases with both τ and the size of G. The more people evacuate, the larger the τ^*, so the former parameter, in turn, increases with the number of evacuees. As the area considered has a scale much larger than a single building in terms of both network and people involved, model size increases accordingly. Particular care must be taken in parameter setting and other implementation choices to reflect the real numbers. We next survey the main model parameters: model granularity, walking velocity, cell capacity, and street capacity.

Model Granularity Model granularity touches both spatial and temporal units. It affects the shape and size of the unit cells in which the network is decomposed and the slots that form the evacuation time horizon. As we assumed above, we embed the crossroad network into a grid whose cells are assumed to be isometric: they can be crossed in any direction in the same amount of time. This amount helps define the time slot duration, and cells are regarded as virtual unit open spaces that communicate to one another via virtual doors. The virtual door capacities are assumed to be the width of the streets. The geometry of the grid can vary and, due to the structure of the streets, we used a rectangular grid - where each street is split into an integer number of cells.

Walking Velocity The basis on which the length of each unit time slot was established is the *free flow walking velocity*, that is, the speed at which humans prefer to walk in non-congested and non-hampered conditions. This parameter is important to perceive the distance an individual can walk during a specific period of time. Through its evaluation, one can define the cells in which an area is to be divided for the best approximation of traveling time. In literature there are different evaluations of pedestrian free flow velocity, including those depending on their age ([47] [28]). Not having this information we assumed a free flow walking speed for a flat surface of 1.00 m/s ([9], [7], [8]).

Cell Capacity The pedestrian density, which is the number of persons per square meter monitored at any time, is vital information for crowd safety and evacuation performance, as movements are dramatically reduced in highly dense areas. According to UK fire safety regulations, the maximum allowed density corresponds to 0.3 m^2 per standing person, which increases to 0.5 m^2 for public houses, to 1.0 m^2 for

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3135054, IEEE Access

dining places, to 2.0 m^2 for sports areas, and to 6.0 m^2 for office areas. In our case study, the maximum capacity of each cell is calculated by assuming 0.5 m^2 per evacuee.

Street Capacity We considered "virtual doors" as the width of the streets. We assume a constant door capacity of 1.8 persons per second per 1-meter door width (p/m/s), meaning that a maximum number of 9 persons can pass through a 1-meter street width per time slot (5 seconds). Also, capacities are assumed to be proportional to street width (i.e., streets are categorized based on their width being either 3 m or 5 m or 7 m).

C. POST-DISASTER RECONSTRUCTION PLANNING

According to the methodology diagram, we have proposed reconstruction planning with the help of EUG generated by the GisToGraph algorithm, which contains details of the area to be reconstructed as well physical dependencies (damage roads and bridges). Starting from the EUG, alternative reconstruction plans have been generated by using deep reinforcement learning (DDQN algorithm). All these generated plans are later handed over to decision-makers to select the best plan and start reconstruction.

The post-disaster reconstruction planning model considers the social benefits of the affected community during rehabilitation and the political benefits of politicians because they also play a vital role in the post-reconstruction plan. Additionally, physical dependencies (damage roads/bridges) are also taken into account during reconstruction. All these concepts are explained below.

1) Social benefits:

Social benefit is a metric that quantifies the benefits an affected community gets after reconstruction in terms of the average number of people staying in a particular building at a particular time. For example, when we rebuild a hospital, the majority of people will get both direct (including doctors, nurses, staff, and patients) and indirect (medical and pharmaceutical stores that open in the immediate surroundings since all these entities are interconnected) benefits.

We have defined social benefit S_P of a reconstruction plan P in the following way:

$$S_P = \sum_{v \in P} S(v).(T_e - T_v) \tag{8}$$

2) Political Priority:

Political priority is associated with every reconstruction unit, it describes political importance for a specific strategy, and it is modeled as an integer in the interval $[1, 10]$. If the maximum integer is associated with any building, it represents that the building is the most important concerning political strategies. For example, decision-makers accordingly with policy actors can decide that public buildings will first be constructed, private residences, and commercial services. This strategy is implemented by assigning priority numbers to reconstruction units according to their importance. Additionally, another major constraint defined for political priority is called *threshold*. According to this concept, every reconstruction plan should satisfy a political priority of $\geq 80\%$.

3) Physical Dependencies:

Physical dependencies specify the reconstruction order among damaged units like buildings, roads, and bridges. For example, in a post-disaster situation where a bridge is damaged and is the only possible path to access other destroyed buildings, it is compulsory to reconstruct it first [12]. Such a bridge is considered a physical dependency and during reconstruction it will be treated as a unit. We have used an undirected graph to describe these kinds of situations and model physical dependencies.

4) Constraints:

Apart from the above concepts, we have several constraints that must be satisfied in every reconstruction plan.

- **Cost:** it describes the *Budget* limit, which is defined by the governments in annual financial statements for the reconstruction of the territories affected by natural disasters. According to this constraint, the total cost of the plan (or a set of plans) must be less than or equal to the defined *Budget*. Such constraint is expressed as follows:

$$\sum_{v \in P} C_v \leq Budget \tag{9}$$

- **Plan Duration:** it is needed to accelerate the reconstruction process. The government sets constraints on the plan regarding specific time to complete it. This constraint is represented as:

$$\sum_{v \in P} T_v \leq T_e \tag{10}$$

- **Political Strategy:** as we have previously discussed, our approach considers political strategies. According to this aim, each unit v is associated with political priority p_v and each plan must satisfy the following constraint, otherwise it will not be acceptable:

$$\frac{\sum_{v \in P} P_v}{|P|} \geq Th_p \tag{11}$$

  where Th_p is a threshold forcing the desired political strategy. This guarantees that the 80% political strategy must be satisfied in every plan.
- **Physical dependencies constraint:** if any street has width less than 3 meters and there are many buildings to be reconstructed, in this case all the damaged buildings are supposed to be constructed in a specific order that will be decided based on the actual condition of the road. We cannot start from the building in the middle of the street because the next damaged buildings will not be accessible. Additionally, political priority of all buildings in that street will be same. According to this,

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3135054, IEEE Access

every plan must respect the dependencies represented by the graph $G' = (V', E')$. This constraint is modelled as:

$$\forall\, v \in P,\ \nexists\, e \in E' \ such\ that \\ e = (v, \overline{v}),\ s_{\overline{v}=0\ and\ \overline{v} \notin P} \tag{12}$$

5) Reconstruction planning by double deep Q-Learning Network (DDQN)

The double deep Q-Learning algorithm is used in our approach for defining alternative reconstruction plans. In DDQN, convolutional neural networks (CNNs) are used to approximate action-value non-linear functions called Q-functions [41]. In our approach, state, action and reward are defined as follows:

State: is a tuple depicted as *(current location, remaining budget, remaining time)*;

Action: represents all possible agent moves in the action space, which is composed by *reconstruction units ID*;

Reward: is the social benefit as defined in Eq. 8.

Agent keeps learning (one node to another) from action reward function by iteratively updating Q-value with the help of equation (13) which is fundamentally know as Bellman equation:

$$Q(s, a; \theta) = S_r(v) + \gamma \max_{a' \in A_v} Q'(s', a'; \theta_i^-) \tag{13}$$

where $a' \in \mathrm{A}_v$ and a' represents the agent action to next node v which has maximum Q value.

$Q(s, a; \theta)$ defines Q value of state s, action a is *(units/roads ID)*, θ is neural network parameter, $S_r(v) = S(v)$ (details are in section IV-B) is the immediate social reward achieved by optimal action a on behalf of current state s, γ is the discount factor that trades off the importance of immediate and later rewards, s' is followed by s after taking action a, and θ_i^- is the network parameter used to compute the target network.

According to generic approach of reinforcement learning [32], network is trained by minimizing a sequence of loss function ($L_i(\theta_i)$) that changes at each iteration i. Loss function in DQN is the squared difference between Q-target and Q-network.

$$L_i(\theta_i) = E[(\overbrace{S_r(v) + \gamma \max_{a' \in A_v} Q'(s', a'; \theta_i^-)}^{\text{Q-target}} - \overbrace{Q(s, a; \theta_i)}^{\text{Q-network}})]^2 \tag{14}$$

Here, θ_i is used to compute Q-network and θ_i^- is used for Q-target computation.

IV. REAL CASE STUDY IMPLEMENTATION

In this section, we discuss a real case study to explain the feasibility and applicability of our data science approach. We have applied our methodology to a portion of the historic city centre of L'Aquila, in the Abruzzo region (Italy) (see Fig. 3), which was severely affected by the 2009 earthquake. For our case study, we considered a small area extending towards the northwest of the crossing of three main streets: Corso Vittorio Emanuele (north-south axis of the city), Corso Principe Umberto, and Via San Bernardino, for a total land size of 246,684.28 m^2 (see Fig. 3). The aim is to provide more details on the implementation and check the results of the optimization model described in Section III-B and the reconstruction planning model described in Section III-C.

In the considered area, we can find 133 buildings and 297 streets with 216 crossroads. Applying the GisToGraph algorithm, we generated the EUG shown in Fig. 4. Such a graph contains 349 nodes (216 nodes for crossroads and 133 nodes for buildings' entrances) and 533 edges (297 are the streets while the remaining 236 are the half-streets and the connections between a building and another node, crossroad or building, reachable from the same street).

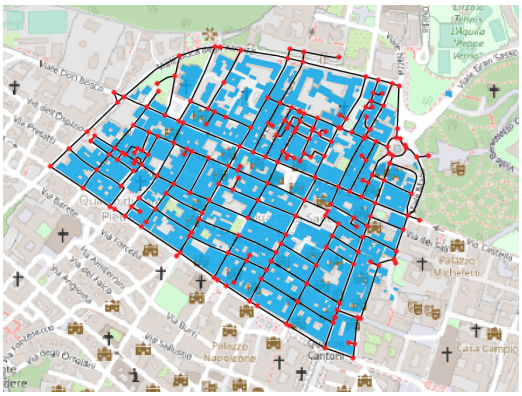

FIGURE 3: Shapefiles of the considered area

All the collected data was gathered from the open data portals of the Abruzzo Region OpenGeoData Abruzzo[3], the Istat[4] database and the project USRA (Special Office for the Reconstruction of L'Aquila)[5]. The portal OpenGeoData Abruzzo is the official provider of the public accessible regional land database containing several layers at a resolution of 1:5000 meters. As shown in Fig. 3, we extracted the street layer (represented by black lines), the crossroad layer (represented by red dots), and the building layer (represented by blue polygons). In addition to the mandatory data of each component of the city, we extracted from the shapefiles two more attributes of the streets: the length expressed in meters and the width categorized as "less than 3.5 meters", "between 3.5 and 7.0 meters", "greater than 7.0 meters". We extracted from the Istat database the shapefiles of the census areas used to estimate the number of people in every building by dividing the total amount of people in each census area by the number of buildings. Thus, we added a new attribute namely "the number of people" of each building. Finally, we enriched the information of the buildings with the help of data provided from the USRA project, which deals with the reconstruction of buildings in L'Aquila. The information extracted from the latter source of information are the following:

[3]http://opendata.regione.abruzzo.it/content/dbtr-regione-abruzzo-scala-15000-edizione-2007-formato-shp

[4]https://www.istat.it/it/archivio/104317

[5]https://bde.comuneaq.usra.it/bdeTrasparente/openData/openDataSet/

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3135054, IEEE Access

- the state of reconstruction (buildings/roads/bridges) that could be "in reconstruction" or "reconstructed";
- the usability of the building that is classified "A" for usable buildings, "B" for buildings temporarily unusable but usable with emergency measures, "C" for partially unusable buildings, "D" for unusable buildings subject to medium reconstruction, "E" for unusable buildings subject to heavy reconstruction, and "F" for unusable buildings due to external risk;
- the required amount (cost/budget) for reconstructing the building in Euros;
- the amount actually granted in Euros;
- cadastral category, which indicates the intended use of the building;
- number of sub-units (e.g., apartments), useful to understand the value of the building.

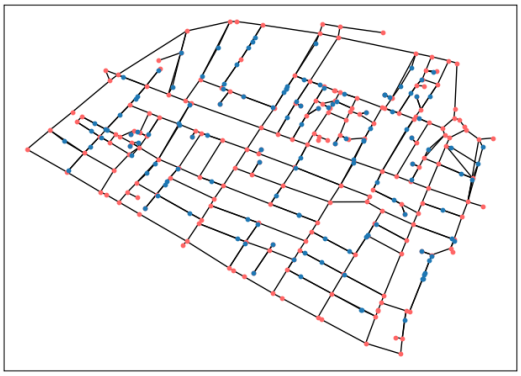

FIGURE 4: Generated graph of the considered area

- **Scenario A:** The crossroad graph G_A was embedded into a grid of cells with a length of 3 meters and width set as the street's width. After the splitting, we obtained a new graph of 3271 nodes corresponding to the cells of the crossroads, including the supernode 0 as a safe place, and 6704 arcs linking adjacent cells that allow people to traverse cells.
- **Scenario B:** Similarly, graph G_A was embedded into a grid of cell length of 5 meters and width equal to the street's width. The splitting resulted in a graph of 1933 nodes corresponding to the cells and 4020 arcs linking adjacent cells that allow people to traverse cells.
- **Scenario C:** Finally, graph G_A was embedded into a grid of cell length of 7 m and width equal to the street's width. This embedding resulted in a graph of 1364 nodes corresponding to the cells and 2882 arcs linking adjacent cells that allow people to traverse cells.

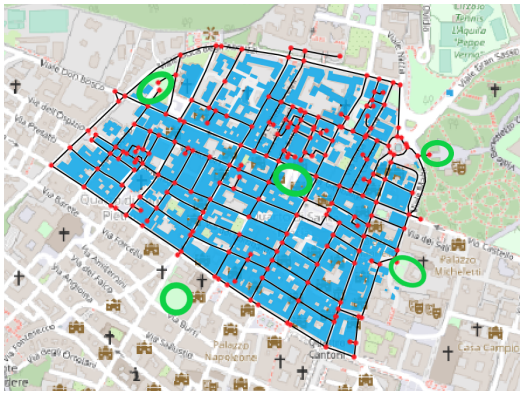

FIGURE 5: List of considered safe locations

A. EVACUATION PLANNING

According to subsection III-B, the dynamic maximum flow model is used for simulating the safe evacuation of people in the historic center of L'Aquila. We start considering the crossroad graph G_A, which consists of 216 nodes (crossroads/junctions) and 297 interconnecting streets with varying widths and lengths. The streets were split into elementary cells, each acting as a (virtual) quasi-rectangular grid that can be traversed in a unit time slot. Three different scenarios with different granularity were considered in this experiment. Hence, the crossroad graph G_A was embedded into three different grid sizes. Without loss of generality, in each scenario, the supernode 0 collectively comprises the set of all safe places. All arcs are assumed to be bidirectional except those towards the safe place. A time slot is the time required for crossing one cell: considering the average free flow walking velocity from subsection III-B2 and the cell size, we determined as 3, 5, and 7 seconds, respectively, as possible time slots for the three different scenarios considered. Note that we considered a street capacity of 9 persons through a 1-meter street width per time slot, meaning that no more than 9 persons can pass through a 1-meter width "virtual" door per monitoring frequency:

In the following section, we report the simulation runs of an emergency evacuation planning that safely evacuates 5000 people for all the three aforementioned scenarios. The people are randomly distributed over the cells. In all three scenarios, we computed the minimum time required to complete the evacuation from the cells into a set of 5 safe locations randomly selected. Moreover, we runned an extra experiment (the controlled one) where the safe locations are fixed. The selected safe locations are highlighted by green circles in figure 5. We like to note that in this case we have chosen to place three of the five safe locations outside the boundaries of the area on the contrary of the random selection where those places are inside the boundary (by chance) of the selected area.

The code for simulation was written in python using the Gurobi API and solved using Gurobi Optimizer version 9.1.1. All the experiments were run on a Core i7-3rd generation 2.9GHz computer with 16Gb of RAM memory under Windows 10 pro 64-bits.

B. RECONSTRUCTION PLANNING

As mentioned in subsection III-C, we used deep reinforcement learning (DDQN) to solve the post-disaster reconstruc-

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3135054, IEEE Access

tion planning and datasets of the city center of L'Aquila is used for experimental evaluation. For this purpose, we are going to present at first the immediate social reward function, then constraints implementation, and lastly the training of the DDQN agent:

- *Immediate Social Reward Function* $(S_r(v))$*:* is computed using Equation 15: given a unit v, the social benefits $S(v)$ is the number of people who take benefit directly and indirectly from its reconstruction. For that reason, the number of direct beneficiaries is combined with the number of indirect ones. Direct beneficiaries are those who are living/working in the direct neighborhood of the unit v. Once a damaged unit v is get reconstructed (now such a reconstructed unit v will become unit u), the agent considers reconstructing the next damaged unit v, in the neighborhood of u, which has the highest social benefits normalized by the distance. The α and β are constants that are used for parameter tuning during reward calculation (in this experiment α= 0.5 and β= 0.8, value range is always between 0 and 1) and allow to determine the weights of the two social benefit components (direct and indirect).

$$S_r(v) = \left[\alpha.b_v + \beta\left(\sum_{u \in V|s_u=1} \frac{S(u)}{d(u,v)}\right)\right] \quad (15)$$

where:

$$\alpha, \beta \in [0,1],\ \alpha + \beta = 1$$

- *Constraints Check:* Constraints verifications are performed when the agent selects a new node v and generates a reconstruction plan. Constraints functions validate input parameters (*'Budget'* and 'T_e') in every cycle[6], which defines duration of the plan. If the conditions are satisfied, then the time and cost required to construct specific damage units are deducted from the total input parameters. Agent will keep reconstructing damage units until the total *budget* or *time* bounds are not reached. The *political priority* is another constraint that the agent needs to meet in every reconstruction. The proposed plan should meet the defined threshold ($P_v \geq 8$) [15]. According to this, at least 80% of set political strategies (Th_p) should be covered in every cycle (see Table 4). The agent will tackle this aspect according to the following formula:

$$Th_p = (Maxp_v - TrainingCycle + 1) * 0.8 \quad (16)$$

Where $Maxp_v$ is the set of maximum political priority of 'v' and $TrainingCycle$ is the round of agent training. The overall threshold of political priority is decreased (see Table 4) in a way that less beneficial buildings are considered in later cycles.

Physical dependency constraint is used to check if there is any damaged road/bridge that prevents accessing any other damaged unit. In that case, that road/bridge will be included in the plan because cost and time will be required for its reconstruction.

- *Post-disaster Rebuilding Planning implementation:* The planning is solved by using the double deep Q-learning network (DDQN) technique. Thus, the training of the deep Q-networks starts by initializing the replay memory and the neural network. Then the learning proceeds by randomly select any of the damaged units 'v' which satisfies all defined constraints. On every action, Q-value (see Eq. 13) gets updated based on immediate social reward $(S_r(v))$ and comprised into the deep neural networks (DNNs), a phase that is called *Training*. We used the ϵ-greedy approach for the exploration part. When the action is chosen then its social reward $S_r(v)$ is calculated, and the agent moves to the next state s'. Typically, the approximation of Q-value, calculated by non-linear function, is very unstable [32], thus the DNNs can be easily overfitted. To solve this problem, we used the experience replay technique, which stores all the experience set $< s, a, S_r(v), s' >$ during forward pass. The random mini-batches taken from the replay store are used to update parameter θ by minimizing the loss function (for imbalance data) [49] (see Eq. 14). The training process will go on until a defined number of iterations are executed.
Subsequently, the training verification of trained agent is done through random agent by running equal number of iterations of trained agent. Once training is completed we get different alternative reconstruction plans (see Eq. 8). Furthermore, multiple plans can be derived from the generated ones since several units can be reconstructed in parallel (see Table 7a, 7b). All the successful reconstructed units in the plan change their status from '0' to '1' and an updated data set will be created and saved in .XLSX and .CSV format. The agent will be trained on the updated dataset for the remaining damaged units, and this process will go on until all the units and roads/dependencies are reconstructed.

The training and verification comparison of our approach with the random agent is showed by Figures 6 and 7, respectively. In both figures, the X-axis represents the number of episodes that we run for the agent during the training and the Y-axis reports the social reward. In Fig. 6, the light blue line represents the actual reward, while the dark blue line represents the mean reward value. The curve is getting less steep after 20000 episodes, which shows agent is fully trained. Similarly, in Fig. 7, the orange line represents random agents and the blue line shows the learning level of the trained agent. The plotting shows the trained agent is way more expert and accurate than a random agent because the

[6]The trained agent returns alternative reconstruction plans. On successful implementation of the selected plan, the data set will be updated and the agent will be trained again to determine new reconstruction plans for the remaining damaged units. A cycle is composed of the agent training and successful implementation of the reconstruction plan.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3135054, IEEE Access

social reward of the random agent remains around 50 and the trained agent reward is more than 150 cycle-1.

We implemented our solution in Python using the TensorFlow framework. The training was conducted on a MacBook Pro with a quad-core intel core i5 processor and Intel Iris Plus Graphics 645 graphics. The deep neural network is composed of three fully connected hidden layers that use the rectified linear activation function. The hidden layers contain 8,64 and 128 neurons respectively. We used the Adam optimizer with a learning rate of 0.001. We used the mean square error as a loss function. The reconstruction planning agent is fully trained on 15000 episodes and we compare the results produced by our trained agent with the results produced by a random agent. Table 2 reports all the hyperparameters specific details used by our approach.

TABLE 2: Fixed Parameters

Fixed Parameters	Value
Optimizer	Adam optimizer, learning rate = 0.001
Loss function	Mean squared error, Eq. 14
Q-Learning function	Q(s,a;θ) = $S_r(v) + \gamma \max_{a' \in A_v} Q'(s', a'; \theta_i^-)$
Batch size	32
Steps before training	20000
Maximum memory size	2000
Political Priority	Minimum=0 , Maximum =10
Exploration strategy	Epsilon greedy policy (Epsilon $\in 10^{-7}$, 1 and self.epsilon_ decay=0.0003.)
Reward discount factor	self.discount_factor = 0.95
Input Parameters	'Budget' and 'Time' (T_e)

V. RESULTS AND DISCUSSION

In this section, we discuss the results of the experiments made on both approaches: evacuation planning and reconstruction planning.

A. EVACUATION PLANNING:

In evacuation planning, we solved problems (1) - (7) for $\tau = 1, 2, 3, ..$ until a solution of value $N = 5000$ evacuees is found. Fig. 8 reports the number of evacuees and the time slot τ required for them to reach a safe place. It is possible to notice in Fig. 8 that, in terms of evacuation time, anyone has reached a safe location in $\tau = 110$, which is 330 seconds ($\approx$ 6 minutes) for scenario A, while in scenario B everyone is safely evacuated in $\tau = 140$ corresponding to 700 seconds ($\approx$ 12 minutes), and finally it takes $\tau = 188$, which is 1316 seconds ($\approx$ 22 minutes), to put everyone to safety in Scenario 3. From Fig. 8, it can be seen that there were some bottlenecks between time slots [0, 80] and [0, 109] in scenarios B and C, respectively. This is due to the fact that the model takes into consideration congestion that could occur on arcs along the passages between adjacent cells. The congestion of some arcs results in bottlenecks, thereby reducing the amount and speed flow along those arcs. After these bottlenecks, the evacuees' pattern follows a certain steady-state distribution.

We report in Fig. 9 the total time required to evacuate different fractions of the entire number of evacuees in all three scenarios. For instance, evacuating 50% of the evacuees to safety took $\tau = 72, 105, 150$ time slots in case of scenarios A, B, C, respectively. As expected, the tail of evacuated people increases with initial occupancy as time progresses.

Lastly, we report the differences among the experiments where the safe places were randomly chosen (original) and the one with the handcrafted safe places (controlled one). In all three scenarios, additional time-slots were needed to safely evacuate the entire population of 5000 people, this is due to the fact that in the original experiment all the safe locations were inside the boundary of the area, thus closer to the citizens. On the opposite, the controlled experiment has three of the five safe places outside of the considered area. The details of this comparison are reported in Fig.10 where:

- **Scenario A**: the evacuees in the original experiment were evacuated in 110 time-slots, whereas in the controlled one they were in safety in 121 time-slots (i.e., an additional 11 time-slots were needed to evacuate everybody);
- **Scenario B**: 140 time-slots in the original experiment and 160 time-slots in the controlled one;
- **Scenario C**: 188 time-slots in the original experiment and 200 time-slots in the controlled one.

B. RECONSTRUCTION PLANNING:

The results of the evaluation of post-disaster reconstruction planning are based on the aforementioned dataset and information of the damaged area. In the graph representation (see Fig.11), red nodes represent damaged units and green nodes show reconstructed/usable units. Similarly, red dotted edges show damaged roads/bridges and dark edges show reconstructed/usable roads and bridges. In the selected area, there are 37 damaged units/buildings out of 133 and 20 damaged roads/bridges out of 150 roads.

Every damaged unit/building and road/bridge has a specific political priority (i.e., from 1 to 10). This priorities are reported in Table 3 according to [15]. Every reconstruction plan have specific thresholds (Eq. 16) as reported in Table 4. These thresholds decrease every cycle in a way that in later cycles less beneficial buildings can be considered.

TABLE 3: Buildings Priority

Buildings	Priority
Hospitals	10
Colleges/School	9
Residential Area	9
Public Points	8
Public Buildings	7
Business Centers	6
Gym Centers	5
Banquet Halls	5
Private Buildings	4
Museums	3
Bars/Cinemas	2
Other Places	1

TABLE 4: Cycles Th_p

Cycles	Th_p (Eq. 16)
1	>= 8.0
2	>= 7.2
3	>= 6.4
4	>= 5.6
5	>= 4.8
6	>= 4.0
7	>= 3.2
8	>= 2.4
9	>= 1.6
10	>= 0.8

When we run the proposed reconstruction planning algorithm on the L'Aquila dataset, the alternative reconstruction

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3135054, IEEE Access

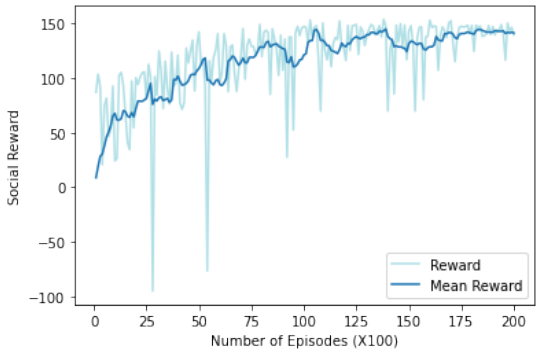

FIGURE 6: Cycle-1 Training

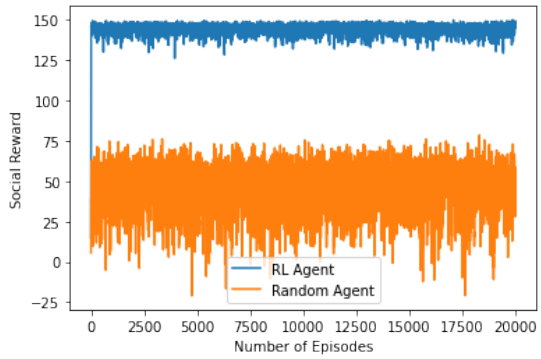

FIGURE 7: Cycle-1 Training Verification

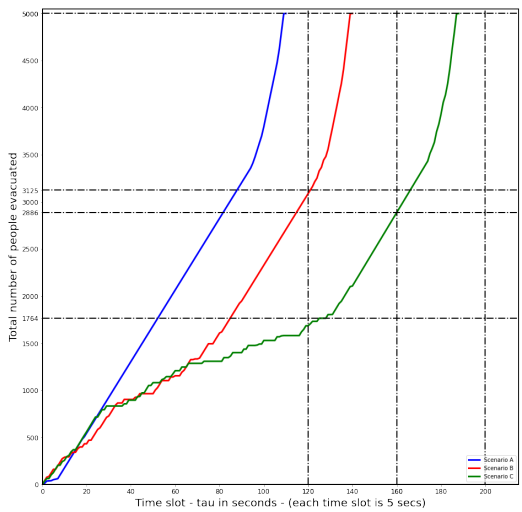

FIGURE 8: Number of people safely evacuated in each scenario at every time slot τ .

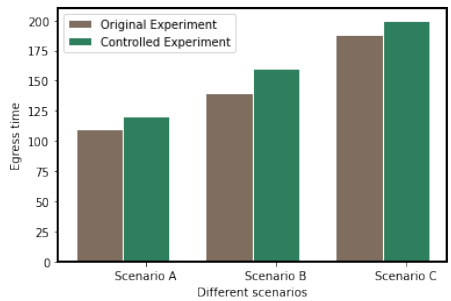

FIGURE 10: Comparison of the total time taken to complete the evacuation process in every scenario

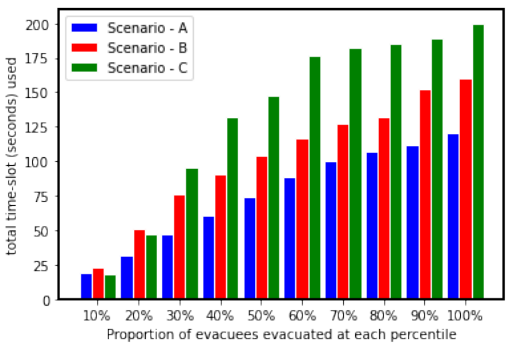

FIGURE 9: Total time taken to evacuate various fractions of total evacuees for each scenario.

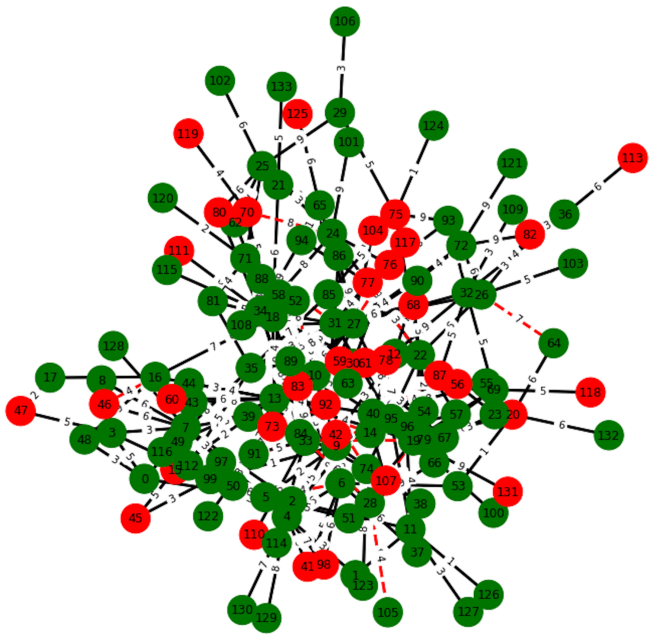

FIGURE 11: Damaged buildings and roads

plans, computed once the agent is fully trained, are reported in Table 5 and Table 6. We used as input parameters $100,000 for the budget and 24 months for the time. In Cycle-1 Plan 1, we count 11 damaged buildings with 4 physical dependencies (PD), a political priority of 9.4, and a total social benefit S_p of 3132. Similarly, in Cycle-1 Plan 2, we find 11 damaged buildings with 5 PDs, a political priority of 9.3, and a S_p of 3120.

It is possible to notice a slight difference regarding the physical dependencies, the political priority, and total social benefits in the two plans since the agent randomly selects

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3135054, IEEE Access

TABLE 5: Cycle-1 Plan 1

Budget: $100000 Time:24 Months						
Sr. No	**Units ID**	**Type**	**Buildings**	**PD**	**PP**	**S_P**
1	104	Civil building	11	4	9.4	3132
2	87	Civil building				
3	87-9	P.Dependency				
4	22-77	P.Dependency				
5	77	Civil building				
6	82	Civil building				
7	86-65	P.Dependency				
8	125	Civil building				
9	15	Civil building				
10	41	P.Dependency				
11	20	Civil building				
12	79-33	P.Dependency				
13	80	Civil building				
14	70	Civil building				
15	83	Civil building				

TABLE 6: Cycle-1 Plan 2

Budget: $100000 Time:24 Months						
Sr. No	**Units ID**	**Type**	**Buildings**	**PD**	**PP**	**S_P**
1	77	Civil building	11	5	9.3	3120
2	77-22	P.Dependency				
3	9-87	P.Dependency				
4	87	Civil building				
5	82	Civil building				
6	86-65	P.Dependency				
7	125	Civil building				
8	15	Civil building				
9	41	Civil building				
10	20	Civil building				
11	79-33	P.Dependency				
12	80	Civil building				
13	70	Civil building				
14	83	Civil building				
15	16-46	P.Dependency				
16	46	Civil building				

the starting damaged building. After the definition of the reconstruction plans, we can highlight those units that can be constructed in parallel. The parallelization of the reconstruction will save time making more efficient use of resources. For this purpose, Table 7a and Table 7b show the lists of the units which can be parallelized in plan 1 and plan 2, respectively.

TABLE 7: Parallel Units Construction

Sr.No	Parallel Units
1	[104,87,87-9]
2	[22-77]
3	[77,82,86-65]
4	[125,15,41,20,79-33]
5	[80,70,83]

(a) Plan 1

Sr.No	Parallel Units
1	[77,77-22]
2	[9-87]
3	[87,82,86-65]
4	[125,15,41,20,79-33]
5	[80,70,83,16-46]
6	[46]

(b) Plan 2

Due to the available budget and time, the reconstruction process completes in different cycles. Table 8 shows the summary of remaining cycles. In cycle 2, a total of 19 units (12 buildings and 7 physical dependencies) are considered and cycle 3 has 21 units (14 buildings and 8 physical dependencies). Additionally, political priority of cycle 2 is 7.9 with $S_P = 2871$ whereas cycle 3 has an overall political priority 6.8 and $S_p = 2252$.

TABLE 8: Summary Table (Budget: $100,000 Time: 24 Months)

Sr.No	Cycles	Units	Buildings	PD/Roads	PP	S_P
1	Cycle 2	19	12	7	7.9	2871
2	Cycle 3	21	14	8	6.8	2252

VI. APPROACH LIMITATIONS AND FUTURE WORK

In this section we examine the assumptions, among those which the DiReCT framework is based on, that might be seen as a cause of limitation of our work. We begin with reviewing those assumption, then we sketch possible directions of future research.

First, this paper addresses the problem of defining or improving the evacuation plans of an urban areas. Evacuation planning can be either done *in a pre-disaster phase* to determine the robustness of the emergency procedures defined, and possibly improve them using a scenario-based analysis; or *during the disaster*, providing citizens (perhaps by real-time digital services) with run-time guides to evacuate indoor places and reach safe areas.

Both phases leverage on risk estimations given by urban experts and on data already available provided by various actors (such as municipalities, rescue operators and/or the civil protection department). In case of emergency – in order to produce an updated picture of the city and redefine the evacuation guidance for those city areas whose predictions were not accurate – such data should be integrated with updated/live data coming, for instance, from citizens who use specific apps, or produced by a smart city infrastructure.

The approach to evacuation planning proposed by the DiReCT framework can be used both in a pre-disaster phase and during an emergency phase, provided that is fed with adequate data. To deal with data inaccuracy, one can approach the problem using the DiReCT evacuation planning algorithm in different scenarios. For instance, one can consider the worst case in which every node has reached its maximum occupancy, i.e, $y_i = n_i \forall i \in V$, and run the algorithm to determine the time needed to move the whole population from dangerous to specific safe location, taking arc bottlenecks into account. However, this highly conservative scenario may not represent real-life situations, where not all the nodes are actually filled at the maximum capacity. So, instead of assigning nodes the maximum nominal occupancy, one can gradually increment this value within ranges based on certain attributes and characteristics of streets and buildings.

While this approach is adequate for a pre-disaster phase, for an emergency phase it can present limitations due to a potential lack of updated information, which introduces uncertainty. Future research is required to remove such limitations. Possible directions are:

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3135054, IEEE Access

- Considering past disasters and the specificity of the area, an accurate data prediction function and scenarios generation can be devised to understand how the uncertainty of data feeding the evacuation plan model can influence the planning prediction.
- Software services and connected mobile apps can be devised, to involve citizens in sending updated data during the emergency.
- A robust and resilient smart city infrastructure can be designed and realized: for a first prototype we suggest to work on controlled small areas, such as a university or high school campus.

Note that during a disaster and some days later, the damaged area can experience problems in the telco infrastructure which, even if not damaged, can be requisitioned by the civil protection department for urgent communication during rescue activities.

Differently from evacuation plan, reconstruction planning is a post-disaster activity that leverages on building practicability checks made by civil protection or professionals during the emergency phase. In L'Aquila, experts needed more than six months to determine the status of all buildings, and the information collected was used to determine reconstruction procedures. In general, during the reconstruction phase the information required is entirely available with high accuracy. Hence, data uncertainty is very limited and should not influence the effectiveness and efficiency of the proposed approach.

Let us now depict further possible extensions that we are considering for this work.

For evacuation planning, we are studying to incorporate additional risk factors into the model, like those associated with each node/building and/or each arc/street. Other aspects we are willing to explore in the future are further optimization models, exact or approximate, to be employed in order to reduce the computational effort presently required by simulations.

For reconstruction planning, we will proceed along two main directions: *i*) define more complex social benefit functions based on more specific attributes than the mere number of residents living in an area; *ii*) analyse and compare machine learning approaches to determine the best algorithm to apply. Since reconstruction is not a real-time problem, we will focus more on results accuracy than on computational efficiency.

Ongoing work, instead, is focused on the integration of further data sources with additional information associated with spatial data. A first additional source of information that we are presently integrating is OpenStreetMap (OSM). OSM database provides geodata of the entire world: this will allows us to extend and test both the evacuation model and the reconstruction plans not only into the Italian region of Abruzzo, but potentially to any place on earth, as long as census data are available.

Another ongoing activity is related to quality assessment of the geo-data employed. To quote an example, OSM provides information about home addresses, pedestrian pathways (including private ones), fences, and other obstacles between streets and buildings. This information opens to improved estimates of the pedestrian paths that must be travelled from the way out of a building to the street and, eventually, to the safe places.

VII. CONCLUSION

In this paper, we have proposed a comprehensive data science framework called DiReCT for evacuation and reconstruction planning in case of natural disasters. The contributions of this paper are:

- an integrated framework that, based on data science, can help decision makers to face natural disasters. As first realization, we embed automatic support to evacuation and reconstruction planning.
- the definition of the *GisToGraph* algorithm to generate an enriched underlining network of any location, specifically tailored to include useful information for disaster management, especially in the *preparedness, response and reconstruction* phases.
- the adaptation and validation of the optimization model developed by [10] to a real outdoor case study, i.e, the historical city centre of L'Aquila in Italy, for emergency evacuation purposes.
- *Reconstruction planning* of damaged buildings and physical dependencies during post-disaster situation in damaged area by using *double deep Q-network (DDQN)* learning algorithm.
- the feasibility and applicability of our framework on a real case study, the historical center of L'Aquila city.

The network generated by the *GisToGraph* algorithm represents the city map in terms of buildings, crossroads and streets. Different from other similar algorithms, we are able to manage additional information, needed for evacuation planning and reconstruction, added as attributes to network nodes and arcs. Concerning the *evacuation planning* model, we adapted the linear optimization model originally developed by Arbib et al. [10] for the evacuation of the interior of a building. The model had to be customized with respect to several parameters, and re-scaled to the network of several orders of magnitude. In *reconstruction planning* we have considered all key attributes like cost, time, physical dependencies, social benefits of affected community and political priority to consider politicians' input. At the end, we have successfully validated our framework on one of the four quarters of the "L'Aquila" city center.

REFERENCES

[1] Australia's fires 'killed or harmed three billion animals'. https://www.bbc.com/news/world-australia-53549936, Jul 2020. Accessed: 2021-05-30.

[2] Earthquake hits greece and turkey, bringing deaths and floods. https://www.bbc.com/news/world-europe-54749509, Oct 2020. Accessed: 2021-05-30.

[3] Flash floods kill more than 70 in afghanistan. https://www.reuters.com/article/us-afghanistan-floods-idUSKBN25M148, Aug 2020. Accessed: 2021-05-30.

[4] Former hurricane eta leaves 150 dead or missing in guatemala, president says. https://www.cbsnews.com/news/hurricane-eta-150-dead-missing-guatemala/, Nov 2020. Accessed: 2021-05-30.

[5] Storm eta hits cuba after devastating central america. https://www.bbc.com/news/world-latin-america-54864963, Nov 2020. Accessed: 2021-05-30.

[6] Rifaat Abdalla. Evaluation of spatial analysis application for urban emergency management. SpringerPlus, 5(2081), 2016.

[7] Ahmed Abdelghany, Khaled Abdelghany, Hani Mahmassani, Hasan Al-Ahmadi, and Wael Alhalabi. Modeling the evacuation of large-scale crowded pedestrian facilities. Transportation research record, 2198(1):152–160, 2010.

[8] Ahmed Abdelghany, Khaled Abdelghany, Hani Mahmassani, and Wael Alhalabi. Modeling framework for optimal evacuation of large-scale crowded pedestrian facilities. European Journal of Operational Research, 237(3):1105–1118, 2014.

[9] Ahmed Abdelghany, Khaled Abdelghany, Hani S Mahmassani, and Saad A Al-Gadhi. Microsimulation assignment model for multidirectional pedestrian movement in congested facilities. Transportation research record, 1939(1):123–132, 2005.

[10] Claudio Arbib, Mahyar T Moghaddam, and Henry Muccini. Iot flows: a network flow model application to building evacuation. In A View of Operations Research Applications in Italy, 2018, pages 115–131. Springer, 2019.

[11] Nicola Bellomo, Benedetto Piccoli, and Andrea Tosin. Modeling crowd dynamics from a complex system viewpoint. Mathematical models and methods in applied sciences, 22(supp02):1230004, 2012.

[12] J. D. Brooks, K. Kar, and D. Mendonça. Dynamic allocation of entities in closed queueing networks: An application to debris removal. In 2013 IEEE International Conference on Technologies for Homeland Security (HST), pages 504–510, 2013.

[13] Wonjoon Choi, Horst W Hamacher, and Suleyman Tufekci. Modeling of building evacuation problems by network flows with side constraints. European Journal of Operational Research, 35(1):98–110, 1988.

[14] Diana Contreras, Thomas Blaschke, Stefan Kienberger, and Peter Zeil. Myths and realities about the recovery of l×³aquila after the earthquake. International Journal of Disaster Risk Reduction, 8:125 – 142, 2014.

[15] Mauro Dolce and Agostino Goretti. Building damage assessment after the 2009 abruzzi earthquake. Bulletin of Earthquake Engineering, 13(8):2241–2264, 2015.

[16] Dorine C Duives, Winnie Daamen, and Serge P Hoogendoorn. State-of-the-art crowd motion simulation models. Transportation research part C: emerging technologies, 37:193–209, 2013.

[17] M.S. Eid and I.H. El-Adaway. Decision-making framework for holistic sustainable disaster recovery: Agent-based approach for decreasing vulnerabilities of the associated communities. Journal of Infrastructure Systems, 24(3), 2018. cited By 4.

[18] BE Eze and OO Olaiya. Understanding the role of big data management during crisis. 2020.

[19] Manfred M. Fischer. GIS and Network Analysis, pages 43–60. 2006.

[20] Pedram Ghannad, Yong-Cheol Lee, Carol J Friedland, Jin Ouk Choi, and Eunhwa Yang. Multiobjective optimization of postdisaster reconstruction processes for ensuring long-term socioeconomic benefits. Journal of Management in Engineering, 36(4):04020038, 2020.

[21] B. Goujon and C. Labreuche. Use of a multi-criteria decision support tool to prioritize reconstruction projects in a post-disaster phase. In 2015 2nd International Conference on Information and Communication Technologies for Disaster Management (ICT-DM), pages 200–206, Nov 2015.

[22] Rahmat Gul and Tameem Akhgar. Death toll from heavy afghanistan flooding surpasses 150. https://www.pbs.org/newshour/world/death-toll-from-heavy-afghanistan-flooding-surpasses-150, Aug 2020. Accessed: 2021-05-30.

[23] Steve Gwynne, Edward R Galea, M Owen, Peter J Lawrence, and Lazaros Filippidis. A review of the methodologies used in the computer simulation of evacuation from the built environment. Building and environment, 34(6):741–749, 1999.

[24] Dirk Helbing and Anders Johansson. Pedestrian, crowd, and evacuation dynamics. arXiv preprint arXiv:1309.1609, 2013.

[25] Evans Howard, Lorenza Pasquini, Claudio Arbib, Antinisca Di Marco, and Eliseo Clementini. Definition of an enriched GIS network for evacuation planning. In Proceedings of the 7th International Conference on Geographical Information Systems Theory, Applications and Management - GISTAM,, pages 241–252. INSTICC, SciTePress, 2021.

[26] Sofia Kalakou and Filipe Moura. Bridging the gap in planning indoor pedestrian facilities. Transport reviews, 34(4):474–500, 2014.

[27] Matthew G Karlaftis, Konstantinos L Kepaptsoglou, and Sergios Lambropoulos. Fund allocation for transportation network recovery following natural disasters. Journal of Urban Planning and Development, 133(1):82–89, 2007.

[28] Richard L Knoblauch, Martin T Pietrucha, and Marsha Nitzburg. Field studies of pedestrian walking speed and start-up time. Transportation research record, 1538(1):27–38, 1996.

[29] Dongkon Lee, Hongtae Kim, Jin-Hyoung Park, and Beom-Jin Park. The current status and future issues in human evacuation from ships. Safety Science, 41(10):861–876, 2003.

[30] Deelesh Mandloi and Jean-Claude Thill. Object-Oriented Data Modeling of an Indoor/Outdoor Urban Transportation Network and Route Planning Analysis, volume 99, pages 197–220. 03 2010.

[31] Muhammed Miah. Use of data mining in emergency evacuation planning. International Journal, 1(1), 2012.

[32] Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Alex Graves, Ioannis Antonoglou, Daan Wierstra, and Martin Riedmiller. Playing atari with deep reinforcement learning, 2013.

[33] Ghulam Mudassir. Social-based physical reconstruction planning in case of natural disaster: A machine learning approach. In Fabiano Dalpiaz, Jelena Zdravkovic, and Pericles Loucopoulos, editors, Research Challenges in Information Science, pages 604–612, Cham, 2020. Springer International Publishing.

[34] Ghulam Mudassir and Antinisca Di Marco. Social-based city reconstruction planning in case of natural disasters: a reinforcement learning approach. In 2021 IEEE 45th Annual Computers, Software, and Applications Conference (COMPSAC), pages 493–503, 2021.

[35] Matthias Müller, Lars Bernard, and Rico Vogel. Geographic Information and Cartography for Risk and Crisis Management, chapter Multi-criteria Evaluation for Emergency Management in Spatial Data Infrastructures, pages 273–286. Springer-Verlag Berlin Heidelberg, 01 2010.

[36] Serafim Opricovic and Gwo-Hshiung Tzeng. Multicriteria planning of post-earthquake sustainable reconstruction. 2002.

[37] Eleonora Papadimitriou, George Yannis, and John Golias. A critical assessment of pedestrian behaviour models. Transportation research part F: traffic psychology and behaviour, 12(3):242–255, 2009.

[38] M. Polese, M. Di Ludovico, and A. Prota. Post-earthquake reconstruction: A study on the factors influencing demolition decisions after 2009 l,Äôaquila earthquake. Soil Dynamics and Earthquake Engineering, 105:139 – 149, 2018.

[39] Kevin Tolhurst Hon. Assoc. Prof. It's 12 months since the last bushfire season began, but don't expect the same this year. https://theconversation.com/its-12-months-since-the-last-bushfire-season-began-but-dont-expect-the-same-this-year-139757, Apr 2021. Accessed: 2021-05-30.

[40] Diana Mercedes Rodriguez Coca. Methodology for planning reconstruction activities after a disaster considering interdependencies and priorities. PhD thesis, 2020.

[41] H. Sasaki, T. Horiuchi, and S. Kato. A study on vision-based mobile robot learning by deep q-network. In 2017 56th Annual Conference of the Society of Instrument and Control Engineers of Japan (SICE), pages 799–804, 2017.

[42] Andreas Schadschneider, Wolfram Klingsch, Hubert Klüpfel, Tobias Kretz, Christian Rogsch, and Armin Seyfried. Evacuation dynamics: Empirical results, modeling and applications. arXiv preprint arXiv:0802.1620, 2008.

[43] Andreas Schadschneider and Armin Seyfried. Empirical results for pedestrian dynamics and their implications for cellular automata models. In Pedestrian Behavior. Emerald Group Publishing Limited, 2009.

[44] Sasan Tavakkol, Hien To, Seon Ho Kim, Patrick Lynett, and Cyrus Shahabi. An entropy-based framework for efficient post-disaster assessment based on crowdsourced data. In Proceedings of the Second ACM SIGSPATIALInternational Workshop on the Use of GIS in Emergency Management, pages 1–8, 2016.

[45] Sasan Tavakkol, Hien To, Seon Ho Kim, Patrick Lynett, and Cyrus Shahabi. An entropy-based framework for efficient post-disaster assessment based on crowdsourced data. In Proceedings of the Second ACM SIGSPATIALInternational Workshop on the Use of GIS in Emergency Management, EM-GIS '16, pages 13:1–13:8, New York, NY, USA, 2016. ACM.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3135054, IEEE Access

[46] Kardi Teknomo. Microscopic pedestrian flow characteristics: Development of an image processing data collection and simulation model. arXiv preprint arXiv:1610.00029, 2016.
[47] TranSafety. Study compares older and younger pedestrian walking speeds. TranSafety, Inc, Road Engineering Journal, 1997.
[48] Dimiter Velev and Plamena Zlateva. An analysis of the relation between natural disasters and big data. International Journal of Data Science, 1(4):370–381, 2016.
[49] Shoujin Wang, Wei Liu, Jia Wu, Longbing Cao, Qinxue Meng, and Paul J. Kennedy. Training deep neural networks on imbalanced data sets. In 2016 International Joint Conference on Neural Networks (IJCNN), pages 4368–4374, 2016.
[50] Xiaoming Zhang, Bingyu Sun, Tao Mei, and Rujing Wang. Post-disaster restoration based on fuzzy preference relation and bean optimization algorithm. In 2010 IEEE Youth Conference on Information, Computing and Telecommunications, pages 271–274, 2010.
[51] Ling Yin, Jie Chen, Hao Zhang, Zhile Yang, Qiao Wan, Li Ning, Jinxing Hu, and Qi Yu. Improving emergency evacuation planning with mobile phone location data. Environment and Planning B: Urban Analytics and City Science, 47(6):964–980, 2020.
[52] Xiaoping Zheng, Tingkuan Zhong, and Mengting Liu. Modeling crowd evacuation of a building based on seven methodological approaches. Building and Environment, 44(3):437–445, 2009.

GHULAM MUDASSIR received the B.S degree in computing from University of Bradford, Bradford, UK, in 2012, and the M.S. degree (Gold Medalist) in computer science from the University of Lahore, Lahore, Pakistan, in 2014. Now he is currently pursuing the Ph.D. degree with the Department of Information Engineering, Computer Science and Mathematics, University of L'Aquila, Italy. His research interests include applied data sciences and machine learning specifically in reinforcement learning.

EVANS ETRUE HOWARD received the B.S degree in mathematics from Kwame Nkrumah University of Science and Technology, Kumasi, Ghana, in 2011, and his M.S. degree in mathematical engineering (mathmods) in University of L'Aquila, Italy, in 2016, where he is currently pursuing the Ph.D. degree. His research interests include emergency evacuation planning and operations research methods and applications specifically in network and discrete (integer and combinatorial) optimization.

LORENZA PASQUINI received the bachelor's degree in Information and Communication Technologies (ICT) Engineering from the University of L'Aquila (Italy) in 2018. Now she is a master degree student in Computer and Systems Engineering and in the 2020 she was awarded a scholarship for research activity on geographic information systems and spatial data analysis by the Department of Information Engineering, Computer Science and Mathematics of the University of L'Aquila.

CLAUDIO ARBIB is full professor of Operations Research at the Dept. of Information Engineering, Computer Science and Mathematics (DISIM), University of L'Aquila, Italy. He received a PhD in Computer and System Science from the University of Rome "Sapienza" in 1987 and the same year begun his academic career at the University of Rome "Tor Vergata". He directed the DISIM from 2008 to 2011, and from 2014 to 2019 was research dean of the University of l'Aquila by appointment of the Rector. Presently, he is vice-president of the centre for excellence DEWS at the University of L'Aquila. In 2013 he was visiting professor at Bilkent University, Dept. of Industrial Engineering (Ankara, Turkey). He was team leader of relevant national/international research projects funded by the EU, the Italian National Research Council (CNR) and the Italian Ministry of University and Research on telecommunications, transportation, integrated logistics, cutting & packing problems, and graph algorithms. He acts as referee for relevant international OR journals, as well as of academic and industrial research projects upon appointment of public authorities. His main research interests are in combinatorial optimization as applied to telecommunications, transportation, logistics, cutting, packing and scheduling problems.

ELISEO CLEMENTINI is an associate professor of computer science at the Department of Industrial and Information Engineering and Economics of the University of L'Aquila (Italy). He received a M.Eng. in Electronics Engineering from University of L'Aquila in 1990 and a Ph.D. in Computer Science from University of Lyon (France) in 2009. He has been a visiting professor at the National Center for Geographic Information and Analysis of the University of Maine, at the Department of Geography and Geomatics of the University of Glasgow, at the Department of Geography of the University of Liege, and at the Department of Geodesy and Geoinformation of Vienna University of Technology. His research interests are mainly in the fields of spatial databases and geographical information science. One of his well-known contributions is the DE+9IM model for topological relations, which is part of Open Geospatial Consortium recommendations and ISO/TC 211 standard. He has been an invited speaker in various international symposia. He is in the editorial board of the ISPRS International Journal of Geo-Information and is a member of the program committee of various conferences such as the Conference on Spatial Information Theory (COSIT), the ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems (ACM GIS), and the International Conference on Geographic Information Science (GIScience).

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3135054, IEEE Access

ANTINISCA DI MARCO is associate professor at the University of L'Aquila where she held her Ph.D. in Computer Science in June 2005. In 2014, she funded SMARTLY s.r.l., an academic SPIN OFF of the University of L'Aquila, and Off Site Art no-profit association established in L'Aquila that promotes culture initiative. Previously she was Research Fellow at the University College London, UK. Her main research interests include software engineering, software performance engineering, context modeling and awareness, data science and Bioinformatics. She published more than 70 journals and conference papers on such topics. She has served as organizing comitee member and program committee member for several international conferences and workshops, and as reviewer for many journals on her research topics. She has been member and coordinator of several national and international research projects and, currently, she is in involved in the coordination of Territori Aperti and as leader of University of L'Aquila team in SoBigData++ EU project. She is one of the creator and coordinator of PINKAMP project, an intensive summer school for girls in STEM topics.

GIOVANNI STILO is an Assistant Professor in the Department of Information Engineering, Computer Science and Mathematics at the University of L'Aquila. He received his PhD. in Computer Science in 2013, and in 2014 he was a visiting researcher at Yahoo! Labs in Barcelona. Between 2015 and 2018, he was a researcher in the Computer Science Department at La Sapienza University, in Rome. His research interests are in the areas of machine learning and data mining, and specifically reinforcement learning, anomaly detection, temporal mining, social network analysis, network medicine, and semantics-aware recommender systems. He has organized several international workshops, held in conjunction with top-tier conferences (ICDM, CIKM, and ECIR), and he is involved as editor and reviewer of top-tier journals, such as TITS, TKDE, DMKD, AI, KAIS, and AIIM.