Natural hazard risk of complex systems – the whole is more than the sum of its parts: I. A holistic modelling approach based on Graph Theory

Marcello Arosio[1], Mario L.V. Martina[1], and Rui Figueiredo[1]

[1]Scuola Universitaria Superiore IUSS Pavia, Palazzo del Broletto - Piazza della Vittoria 15, 27100 Pavia (Italy)

Correspondence to: Marcello Arosio (marcello.arosio@iusspavia.it)

Abstract: Assessing the risk of complex systems to natural hazards is an important and challenging problem. In today's intricate socio-technological world, characterized by strong urbanization and technological trends, the connections, interdependencies and interactions between exposed elements are crucial. These complex relations call for a paradigm shift in collective risk assessments, from a reductionist approach to a holistic one. Most commonly, the risk of a system is estimated through a reductionist approach, based on the sum of the risk of its elements individually. In contrast, a holistic approach considers the whole system as a unique entity of interconnected elements, where those connections are taken into account in order to more thoroughly assess risk. To support this paradigm shift, this paper proposes a new holistic approach to assess the risk in complex systems based on Graph Theory. The paper is organized in two parts: part I describes the proposed approach, and part II presents an application to a pilot study in Mexico City. In part I, we demonstrate that by representing a complex system such as an urban settlement by means of a network (i.e. a graph), it is possible to take advantage of the techniques made available by the branch of mathematics called Graph Theory to analyse its properties. Moreover, it is possible to establish analogies between certain graph metrics (e.g. authority, degree, hub values) and risk variables (exposure, vulnerability and resilience). Leveraging these analogies, one can not only obtain a deeper knowledge of the system (structure, weaknesses, etc.), but also understand its risk mechanisms (how the impacts of a single or multiple natural hazards are propagated, where they are exacerbated), and therefore assess the disaster risk of the system as a whole, including second-order impacts and cascade effects.

1. Introduction

We live in an increasingly complex world. Today's societies are interconnected in complex and dynamic social-technological networks, and have become more dependent on the services provided by critical facilities. Population and assets in hazard-prone areas are increasing, which translates into higher economic losses (Bouwer et al., 2009). In coming

years, climate change is expected to exacerbate these trends (Alfieri et al., 2017). In this context, disaster risk is a worldwide challenge that institutions and private individuals need to face at both global and local scales. Today, there is growing attention to the management and reduction of disaster risk, as illustrated for example by the wide adoption of the

30   Sendai Framework for Disaster Risk Reduction (SFDRR, 2015).

The effective implementation of strategies to manage collective risk, i.e. the risk assembled by the collection of risk assets/elements, requires support from Risk Assessment (RA) studies that quantify the impacts that hazardous events may have on the built environment, economy and society (Grossi and Kunreuther, 2005). As such, RA is a fundamental step in the Disaster Risk Management cycle, particularly within a risk management framework where Disaster Risk Reduction

35   (DRR) is directly integrated in national development strategies. The research community concerned with DRR, particularly in the fields of physical and environmental science, has generally agreed on a common approach for the calculation of risk (R) as a function of hazard (H), exposure (E), and vulnerability (V): $R = f (H, E, V)$ (e.g. Balbi *et al.*, 2010). Hazard defines the potentially damaging events and their probabilities of occurrence, exposure represents the population or assets located in hazard zones that are therefore subject to potential loss, and vulnerability links the intensity

40   of a hazard to potential losses to exposed elements. This framework has been in use by researchers and practitioners in the field of seismic risk assessment for some time (Bazzurro and Luco, 2005; Crowley and Bommer, 2006), and has more recently also become standard practice for other types of hazards, such as floods (Arrighi et al., 2013; Falter et al., 2015). Despite consensus on the conceptual definition of risk, different stakeholders tend to have their own specific perspectives. For example, while insurance and reinsurance companies may focus on physical vulnerability and potential economic

45   losses, international institutions and national governments may be more interested in the social behaviour of society or individuals in coping with or adapting to hazardous events (Balbi et al., 2010). As such, even though this risk formulation can be a powerful tool for RA, it has its limits. For instance, it does not consider social conditions, community adaptation or resilience (i.e. a system's capacity to cope with stress and failures and to return to its previous state). In fact, resilience is still being debated and there is not an common and consolidated approach to assess it (Bosetti et al., 2016; Bruneau et

50   al., 2004; Cutter et al., 2008, 2010). To overcome these limits, different approaches have been proposed in recent research. For example, (Carreño et al., 2007a, 2007b, 2012) have proposed to include an aggravating coefficient in the risk equation in order to reflect socio-economic and resilience features. Another example can be found in the Global Earthquake Model, which aims to assess so-called integrated risk by combining hazard (seismic), exposure and vulnerability of structures with metrics of socio-economic vulnerability and resilience to seismic risk (Burton and Silva, 2015). These new integrated

55   approaches are seen with increasing interest, particularly with regard to Climate Change Adaptation (CCA). In fact, CCA-related research activities are strongly focusing on how local communities are able to adapt to and cope with a disaster (Balbi et al., 2010). In addition to integrated risk, there is also an increasing international interest in multi-risk assessment

as a combination of multi-hazard (natural and manmade) with multi-vulnerability (physical and social) assessments (Gallina et al., 2016; Karagiorgos et al., 2016; Liu et al., 2016). Some research has explored the potential of an integrated

60   approach to risk and multi-risk assessment of natural disasters in various fields. However, quantitative collective RA (both scenario- and event-based) still requires further development to consider the connections and interactions between exposed elements. Although holistic approaches are in strong demand (Cardona, 2003; Carreño et al., 2007b; IPCC, 2012), the majority of methods and especially models developed so far are based on a reductionist paradigm, which estimates the collective risk of an area as the combination of the risk of its exposed elements individually, neglecting the

65   links between them.

In a changing society which increasingly relies on interconnections, the links between elements are crucial, especially considering the urbanization and technological trend which modern-day society is strongly promoting. Urban population growth means that people are depending more and more on critical facilities, and there is an increasing interdependency between infrastructures. Complex socio-technological networks, which increase the impact of local events on broader

70   crises, characterize the modern technology of present-day urban society (Pescaroli and Alexander, 2016). This partly supports the perception that the risk assessment paradigm, where the total risk is the combination of the risk of elements individually, requires a more comprehensive approach than the traditional reductionist one, as it involves "whole systems" and "whole life" thinking (Albano et al., 2014). The reductionist approach contrasts with the complex nature of disasters. These tend to be strongly non-linear i.e. the ultimate outcomes (losses) are not proportional to the initial event (hazard

75   intensity and extensions) and are expressed by emergent behaviour (i.e. macroscopic properties of the complex system) that appear when the number of single entities (agents) operate in an environment, giving rise to more complex behaviours as a collective (Bergström, Uhr and Frykmer, 2016). In the last decade, many disasters have shown high levels of complexity and the presence of nonlinear paths and emergent behaviour that have led towards secondary events. Examples of such large-scale extreme events are the eruption of the Eyjafjallajokull volcano in Iceland in 2010, which affected

80   Europe's entire aviation system, the flooding in Thailand in 2011, which caused a worldwide shortage of computer components, and the energy distribution crisis triggered by hurricane Sandy in New York in 2012.

Secondary events (or indirect losses) due to dependency and interdependency have been thoroughly analysed in the field of critical infrastructures: telecommunications, electric power systems, natural gas and oil, banking and finance, transportation, water supply systems, government services and emergency services. Rinaldi et al. (2001), in one of the

85   most quoted papers on this topic, proposed a comprehensive framework to identify, understand and analyse the challenges and complexities of interdependency. Since then, numerous works have tried to study the increase in systemic vulnerability (Menoni et al., 2002) due to the increase in interdependencies in our modern society: *"systemic risk is the risk of having not just statistically independent failures, but interdependent, so-called 'cascading' failures in a network*

of N interconnected system components. That is, systemic risks result from connections between risks ('networked risks').
90   *In such cases, a localized initial failure ('perturbation') could have disastrous effects and cause, in principle, unbounded*
damage as N goes to infinity (Menoni, 2001). Ouyang (2014) reviews the existing modelling approaches and categorizes
them into six groups: empirical, agent-based, system dynamics-based, economic theory-based, network-based, and others.
This wide range of models reflects the different levels of analysis of critical infrastructures (physical, functional or socio-
economic). Trucco et al. (2012) propose a functional model aimed at i) propagating impacts, within and between
95   infrastructures, in terms of disservice due to a wide set of threats and ii) applying it to a pilot study in the metropolitan
area of Milan. However, this well-developed branch of research is mostly focused on the analysis of a single infrastructure
typology, and the aim is usually to assess the efficiency of the infrastructure itself rather than the impact that its failure
may have on society.

A full research branch analyses the complex socio-physical-technological relationships of our society considering a
100   System-of-System (SoS) perspective, whereby systems are merged into one interdependent system of systems. In a SoS,
people belong to and interact within many groups, such as households, schools, workplaces, transport, health care systems,
corporations and governments. In a SoS, the dependencies are therefore distinguished between links within (intra) the
same system, or between (inter) different systems (Alexoudi et al., 2011). The relation between different systems are
modelled in literature using qualitative graphs or flow diagrams (Kakderi et al., 2011) and by matrices (Abele and Dunn,
105   2006). Tsuruta and Kataoka (2008) use matrices for determining damage propagation due to interdependency based on
earthquake data and expert judgment considering infrastructure networks (e.g. electric power, waterworks, sewerage,
telecommunication, road, and social functions like finance, medical treatment and administration). Menoni (2001)
proposes a framework showing major systems interacting in a metropolitan environment based on observations on the
Kobe earthquake. Lane and Valerdi (2010) provide a comparison of various SoS definitions and concepts, while Kakderi
110   et al. (2011) have delivered a comprehensive literature review of methodologies to assess the vulnerability of a SoS.

Although the aspects of complexity and interdependency have been investigated by various models of critical
infrastructure as a single system, or as systems of systems, we would like to further explore the complexity of the exposed
elements and their interconnections, and propose an approach to develop more holistic collective risk assessments. In
particular, it is necessary to better analyse the interaction between elements at risk and their influences on indirect impact
115   assessment. Given this context, this paper proposes an insight into collective risk assessment from an innovative holistic
perspective.

The aims of this paper are:

- to present a new perspective to promote a paradigm shift from a reduction to holistic approach;

- to propose a new approach to analyse the risk of complex systems based on Graph Theory, the branch of mathematics
120   for the treatment of networks;
- to link traditional risk variables (exposure, vulnerability, and resilience) to certain properties of a graph;
- to introduce a debate on the new perspective and approach, and directions for future developments.

2.   A Graph Theory approach to modelling complex systems

2.1.   Background and relevant graph properties

125   As discussed in detail above, a network could allow to portray the complexity of a risk system. In the scientific community, the mathematical properties of a network are studied using Graph Theory, which could provide a better angle also to assess risk from a holistic and systemic viewpoint. The following paragraphs review the main aspects of Graph Theory, on which our approach is based.

Over recent decades, studies of network concepts, connections and relationships have strongly accelerated in every area
130   of knowledge and research (from physics to information technology, from genetics to mathematics, to building and urban design), showing the image of a strongly interconnected world in which relationships between individual objects are often more important than the objects themselves (Mingers and White, 2009).

Graph Theory is the branch of mathematics that studies the properties of networks. Networks can comprise physical elements in the Euclidean space (e.g. electric power grids, the Internet, highways, neural networks) or entities defined in
135   an intangible space (e.g. collaborations between individuals).

Since its inception in the eighth century, Graph Theory has provided answers to questions in different sectors, such as pipe networks, roads, and the spread of epidemics. During the last decade, there has been an increase in interest in the study of complex networks (e.g. irregular structures, dynamically evolving in time), paying renewed attention to the dynamic properties of networks (Börner et al., 2007; Newman, 2003).

140   Formally, a complex network can be represented by a graph G which consists of a finite set of elements $V(G)$ called vertices (or nodes), and a set $E(G)$ of pairs of elements of $V(G)$ called edges (or links) (Boccaletti et al., 2006). The graph can be undirected or directed (Figure 1a and b). In an undirected graph, each of the links is defined by a pair of nodes i and j, and is denoted as l_{ij}. The link is said to be incident in nodes i and j, or to join the two nodes; the two nodes i and j are referred to as the end-nodes of link l_{ij}. In a directed graph, the order of the two nodes is important: l_{ij} stands for a link
145   from i to j, node i points to node j, and $l_{ij} \neq l_{ji}$. Two nodes joined by a link are referred to as adjacent.

A graph could have edges of different capacity or intensity. In this case, a real number representing the weight (i.e. the relative importance/capacity/intensity) of the link is associated to it, and the graph is said to be weighted. The weighted

graph can also be directed or undirected. A graph that shows variability in the capacity and intensity values of edges is called weighted, and a real number representing the strength or value of the connection is associated to each link of the

150 network (Figure 1c) (Börner et al., 2007).

A short list of the most common set of node, edge and graph measurements used in Graph Theory is presented here. There are measurements that analyse the properties of nodes or edges, local measurements that describe the neighbourhood of a node (single part of the system), and global measurements that analyse the entire network (whole system). From a holistic point of view, it is important to note that since some node/edge measurements require the examination of the

155 complete network, this allows looking the studied area as a unique entity that results from the connections and interactions between its parts and characterizing the whole system.

The degree (or connectivity, k) of a node is the number of edges incident with the node. If the graph is directed, the degree of the node has two components: the number of outgoing links (referred to as the degree-out of the node), and the number of ingoing links (referred to as the degree-in of the node). The distribution of the degree of a graph is its most basic

160 topological characterization, while the node degree is a local measurement that does not take into account the global properties of the network. On the contrary, path lengths, closeness and betweenness centrality are properties that consider the complete network. The path length is the geodesic length from node i to node j: in a given graph, the maximum value of all path lengths is called diameter and the average shortest path length is named characteristic path length. Closeness computes the distance (number of links) of a node to all others, and betweenness is defined as the number of shortest

165 paths between pairs of nodes that pass through a given node.

Other relevant characteristics that are commonly analysed in directed graphs to find the relative importance of a node, in terms of the global structure of the network, are the hub and authority properties. A node with high hub value points to many other nodes, while a vertex with high authority value is linked by many different hubs. Mathematically, the authority value of a node is proportional to the sum of the node hubs pointing to it and the hub value of a node is proportional to

170 the sum of authority of nodes pointing to it. In the World Wide Web, for example, websites (nodes) with higher authorities contain the relevant information on a given topic (e.g. wikipedia.com) while websites with higher hubs point to such information (e.g. google.com).

The mathematical properties presented above are useful metrics to analyse the structural (i.e. network topology, arrangement of a network) and functional (i.e. network dynamics, how the network status changes after perturbation)

175 properties of large networks. Depending on the statistical properties of the degree distributions, there are two broad classes of network: homogeneous, and heterogeneous. Homogeneous networks show a distribution of the degree with a typically exponential and fast decaying tail, such as Poissonian distribution, while heterogeneous networks have a heavy-tailed distribution of the degree, well-approximated by a power-law distribution. Many real-world complex networks show

power-law distribution of the degree and these are also known as scale-free networks, because power-laws have the same
functional form on all scales (Boccaletti et al., 2006). Networks with highly heterogeneous degree distribution have few
nodes linked to many other nodes (i.e. few hubs), and a large number of poorly connected elements.

The properties of the static network structure are not always appropriate to fully characterize real-world networks that
also display dynamic aspects. There are examples of networks that evolve with time or according to external environment
perturbations (e.g. removal of nodes/links). Two important properties to explore the dynamic response to a perturbation
are percolation thresholds and fragmentation modes.

Percolation was born as the model of a porous medium but soon became a paradigm model of statistical physics. Water
can percolate in a medium if an infinite number of links exists (i.e. the presence of links means the possibility of water
flowing through the medium), and this depends largely on the fraction of links that are maintained. When the graph has
most of the links, there is a higher probability that connection between two nodes may exist and, in this case, the system
percolates. Vice versa, if most links are removed, the connected components tend to be small and the network becomes
fragmented (Van Der Hofstad, 2009). The percolation threshold is an important network feature resulting from
percolation: obtained by removing vertices or edges from a graph. When a perturbation is simulated as a removal of
nodes/links, the fraction of nodes removed is defined as $f = \frac{Nodes_{removed}}{Nodes_{Total}}$, and the probability of nodes/links present in

a percolation problem is $p = 1 - f = \frac{Nodes_{remainig}}{Nodes_{Total}}$. Consequently, it is possible to define the percolation threshold (p_c)
as the minimum value of p that leads to the connectivity phase of the graph (Gao et al., 2015), i.e. the threshold that
discriminates between the connected and fragmented phases of the network. In a random network, for example, $p_c = 1/\bar{k}$,
where $\bar{k}$ is the mean of degree k (Bunde and Havlin, 1991).

The second property that investigates dynamic evolution is the fragmentation property (i.e. number and size of the portions
of the network that become disconnected). The number and the size of the sub-networks obtained after removing the
vertices/edges provide useful information. In the case of a so-called giant component fragmentation, the network retains
a high level of global connectivity even after a large amount of nodes have been removed, while in the case of total
fragmentation, the network collapses into small isolated portions. For this reason, "*keeping track of the fragmentation
evolution permits the determination of critical fractions of removed components (i.e., fraction of component deletion at
which the network becomes disconnected), as well as the determination of the effect that each removed component has
on network response*" (Dueñas-Osorio et al., 2004).

2.2. Proposed workflow

Within the framework of Graph Theory, we propose an approach based on the following two major phases:

- Network conceptualization;
- Graph construction.

210   The workflow is presented in Figure 2.

2.2.1.   Network conceptualization

According to the objects of each specific context, the network conceptualization phase defines the hypothesis of the analysis and the system boundaries. In particular, it establishes the two main objects of the network: nodes (vertices) and links (edges) and their characteristics.

215   The nodes can theoretically represent all the entities that the analysis wants to consider: physical elements like a single building, bridge and electric tower, suppliers of services such as schools, hospitals and fire brigades, or beneficiaries such as population, students or specific vulnerable groups such as elderly people. Due to the very wide variety of elements that can be chosen, it is necessary to select the category of nodes most relevant to the specific context of analysis. It is also necessary to define, for each node, the operational state that can be characterized, from the simplistic Boolean

220   (functional/non-functional), to discreet states (30/60/100% of service/functionality), or even a complete continuous function (similarly to vulnerability functions). In a graph, the states of each node depend both on the states of the adjacent nodes and on the hazard. In this paper, we use the term *node* to refer to its graph characteristics and term *element* to refer to the entity that it represents in the real world.

The links between the nodes that create the graph can range from physical, geographical, cyber or logical connections

225   (Rinaldi et al., 2001). Two nodes are physically connected if one node is dependent on the material output(s) of the other (e.g. a power grid system provides electricity to an industrial area). Two nodes are geographically connected if their close spatial proximity is able to influence the reciprocal state (e.g. building close to a nuclear power plant). If a node depends on the information provided by another node, these two nodes are cyber connected (e.g. supervisory control and data acquisition (SCADA) systems, industrial productivity). When the nodes are connected but the mechanisms are not

230   physical, geographical and cyber, then the connection is logical (e.g. the increasing of petroleum price has an influence on the traffic congestion). Depending on the selected category for the node, it is necessary to define the type of connection between elements, which are essential for the analysis, and its characteristics: direction and weight. The graph will be directed when the direction of the connection between elements is relevant, otherwise it will be undirected (e.g. the world wide web vertices represent webpages and directed edges represent hyperlinks; network based on friendship is an

235   undirected graph). Furthermore, it is possible to assign a value to each link and to weight the different connections in the graph.

In defining the topology, it is crucial to define the level of analysis details coherently with the scope and scale, both for the selection of elements and for the relationship between elements that need to be considered. In the case of a very high detail for example, a node of the graph could represent a single person within a population, and in the case of lower resolution, it could represent a large group of people with a specific common characteristic, such as living in the same block or having the same hobby. In the case of analyses at a coarser level, an entire network (e.g. electric power system) can be modelled as a single node of another larger network (e.g. national power system).

2.2.2. Graph construction

Once the network is conceptually defined, in order to actually build the graph, it is then necessary to establish the connection between all the selected elements. The conceptual network determines the existence of connections between categories of elements, but it does not define how a single node of one category is linked to a node of another category. Therefore, it is necessary to define rules that establish the connections between each single node. For the sake of clarity, an example could be the following: in the conceptual network, the relationship is defined between students and school ("students go to school"); subsequently, it is necessary to make the link between each student and a school in the area, applying the following rule - "students go to the closest school". This is an example of geographical connection with nodes that are linked by their spatial proximity.

The connections between the single elements can be represented either by a list of pairs of nodes or, more frequently, by the adjacency matrix. Any graph G with N nodes can be represented in fact by its adjacency matrix $A(G)$ with $N \times N$ elements A_{ij}, whose value is $A_{ij} = A_{ij} = 1$ if nodes i and j are connected, and 0 otherwise. If the graph is weighted, $A_{ij}=A_{ji}$ can have a value between 0 and 1 expressing the weight of the connection between the nodes. The properties of the nodes are represented in both cases by another matrix, with a column for each property associated with the node (e.g. name, category, type).

In practical terms, the list of all connections or the adjacency can be automatically obtained via GIS analysis, in the case of geographical connections, or by database analysis, in the case of other categories of connections. The list of nodes, together with either the list of links or the adjacency matrix, are the inputs to build the mathematical graph using dedicated tools. For example, the igraph package (http://igraph.org/r/), the specific library and package for network analysis of the R environment, provides a set of data types and functions for the implementation of graph algorithms, and is able to handle large graphs with millions of vertices and edges.

3. Graph concepts in the context of collective risk assessment

265 Once a network has been setup and a graph has been constructed, it is then possible to compute and analyse its properties by means of Graph Theory. Recall that the topologic properties (i.e. degree, path length, degree distribution, hub, authority) represent the structural characteristics of the network, while dynamic properties (i.e. percolation threshold and fragmentation) provide information about the response of the network to a perturbation or disruption. These analytical tools can be very useful for understanding risk mechanisms: How is it generated? How is it propagated? Which are the

270 weaknesses of the system? Information such as this is key to perform more thorough risk assessments, where second-order and cascade effects are considered, and to support the implementation of more effective risk mitigation actions.

The traditional conceptual skeleton to describe risk as a function of hazard, exposure, vulnerability and resilience can still be adopted within the framework of the proposed graph-based approach. In fact, the properties calculated from a graph consist in a new layer of information for some of those risk variables that go beyond their traditional interpretations within

275 the reductionist paradigm. In particular, they provide a more comprehensive characterization of the single nodes (deriving from their relationships with other nodes), as well as of the system as a whole. In the following we present the analogies that can be established between the graph properties described in Section 2.1 and certain risk variables, and then provide a simple illustrative example.

3.1. Analogy between graph properties and risk variables

280 The proposed graph properties can be used to more thoroughly characterize systems of exposed elements. As such, from the four risk variables shown above, the hazard preserves its traditional definition as an event that can impact such systems, or part(s) of it, with certain intensities and associated probabilities of occurrence. For the three other variables, namely, exposure, vulnerability, and resilience, below we provide a brief discussion on their analogies with the graph properties presented previously. It is worth noting that the concept of impact that occurs as a result of the above factors

285 can also be expanded within the graph-based approach. For example, the indirect impact due to a hazardous event suffered by a certain node may be defined as a function of the direct damage sustained by one or more of its parent nodes (i.e. traditional impact), and of the type of service the latter provide to the former. This could be given by a vulnerability function defining the consequences of such a cascade effect. The integration of indirect impact quantification within the graph-based framework will be addressed in future research.

3.1.1. Exposure

290

In analogy with the traditional approach but at the same time extending its concept, the value of each exposed node can be assessed as the relative importance that is given to it by the graph, which is measured by the network itself by means

of the connections that point to each node. In Graph Theory, this relative importance between elements, based on standardized values, can be investigated through the authority analysis. The distribution of authority shows which elements might have a lower or higher value within the system. In particular, a node (element) has high authority value if in the whole system there are many nodes (or otherwise some hubs) that provide services (i.e. providers) to it in order to make it functional; in this case, higher authority values imply that the system privileges some receiver nodes compared with others according to their connections with the provider nodes. A factory settled in an industrial district receives more services (e.g electric power, special road and logistic systems) compared with a factory historically developed in the old quarter of the city. This simple example shows that the factory in the district is privileged by the system connections compared to the industry settled in the old quarter.

3.1.2. Vulnerability

In the reductionist approach, vulnerability is referred to the propensity of an asset to be damaged because of a hazardous event. In the graph perspective, the vulnerability can be defined both for the whole system and the single node.

In the first case, the vulnerability of a system can be defined as the propensity of the network to be split into isolated parts after a hazardous event. In that condition, an isolated part is unable to provide and receive services, which can translate into indirect losses. The system vulnerability can be expressed by the following graph properties: hubs, betweenness and degree out distribution. The presence of nodes with high values of hub values (namely hubs) indicates a propensity of the network to be indirectly affected more extensively by a hazard event, since a large number of nodes are connected with the hubs. Betweenness manifests the tendency to create isolated sub-networks. As an example, in a road graph, the bridge node has a higher value of betweenness because all the nodes of a sub-graph (e.g. one side of the river) need to pass through the bridge node in order to connect to the nodes of the other sub-graph (the other side of the river). In the case of a bridge failure, the two sides of the river are isolated and the original road graph splits into two sub-graphs. A network that has nodes with high betweenness values has a higher tendency to be fragmented, because it has a strong aptitude to generate isolated sub-networks. Finally, the degree distribution, which expresses network connectivity of the whole system (i.e. the existence of paths leading to pairs of vertices), has a strong influence on network vulnerability after a perturbation. The shape of the degree distribution determines the class of a network: heterogeneous graphs (power-law distribution and scale-free network) are more resistant to random failure, but they are also more vulnerable to intentional attack (Schwarte et al., 2002). As emphasised above, scale-free networks have few nodes linked to many nodes (i.e. few hubs), and a large number of poorly connected elements. In the case of random failure, there is a low probability of removing a hub, but if an intentional attack hits the hub, the consequences for the network could be catastrophic.

In the second case, the single node vulnerability depends on the relationship that one node has with the others. In particular, the closeness represents the likelihood of a node to be affected indirectly by a hazard event due to the lack of services provided by other nodes. A lower value of closeness, i.e. a shorter path between a node and the network, means
325   a higher probability of a node to remain isolated because of a hazard event. On the other hand, high value of closeness, i.e. a longer path between a node and the network, means a low probability that the node i will be isolated.

3.1.3. Resilience

Resilience differentiates from vulnerability in terms of dynamic features. The properties and functions used to model vulnerability are static characteristics that do not consider any time evolution; instead, the definition of resilience implies
330   a time evolution of the characteristics of the whole system. This difference can be expressed by a cinematography analogy: vulnerability is a single frame of the resilience video.

In this context, the study of the percolation threshold (p_c) can be used to explain the resilience of the network after a perturbation. The p_c value distinguishes between the connectivity phase (above p_c) and the fragmented phase (below p_c). In the connectivity phase, the network can lose nodes without losing the capacity to cope with the perturbation as a
335   network, while in the fragmented phase, the network does not actually exist anymore and the remaining nodes are unable to cope with the disruption alone.

This critical behaviour is a common feature also observed in natural disasters. In some cases, the exposed elements withstand some damage and loss, but the overall system maintains its structure. However, there are events in which the amount of loss (affected nodes) is so relevant that the system loses the overall network structure. In the first case, the
340   system has the capacity to cope independently and tackle the event, while in the second case, the system is unable to cope. In the case of an earthquake, for example, if a large number of exposed elements are damaged, it is common for affected regions to be unable to cope with the situation and require help from outside their borders (i.e. the graph needs new nodes). Asking for support from outside would increase the extension of the area (graph) and offers new resources (nodes), therefore decreasing the percolation threshold value increasing connectivity. In this perspective, the percolation threshold
345   should define when a specific area is not self-sufficient to overcome that level of loss and requires help from outside.

The dynamic responses are characterized by the network fragmentation property, which describes the performance of a network when its components are removed (Dueñas-Osorio and Vemuru, 2009). For instance, the so-called giant component fragmentation (the largest connected sub-network) and the total fragmentation describe network connectivity and determine the failure mechanism (Dueñas-Osorio et al., 2004). Keeping track of fragmentation evolution makes it
350   possible to determine both the critical fraction of components removed (i.e. the smallest component deletion that disconnects the network), and the effect that each component removed has on the network response.

For these reasons, we consider percolation threshold and network fragmentation a good indicators of resilience, also because it is able to show the emergent behaviour of the whole system beyond just considering the single parts of the network (e.g. node).

3.2. Illustrative example

In order to illustrate the application of Graph Theory in the characterization of a system exposed to natural hazards, in Figure 3 we present an example of a hypothetical city comprising various elements of different types which provide services among them. In specific, our example includes 20 elements: 8 Blocks of residential buildings, 1 Hospital, 2 Fire Stations, 3 Schools, 3 Fuel Stations and 2 Bridges. We assume that these elements are located in a flood-prone area. Blocks are intended to represent the population, which receives services from the other nodes. Bridges provide a transportation service, Fire Stations provide a recovery service, Hospitals provide a healthcare service, Schools provide an education service, and Fuel Stations provide a power service. Figure 3(a) shows how the elements are connected into a graph. The authority and hub values adopted in this illustrative example have been computed using the R igraph package. The full library of functions adopted are available in Nepusz and Csard (2018).

In Figure 3(b), the size of the elements is proportional to their authority values. Blocks 6, 18, 19 and 20 have higher authority values than the other elements of this typology because they receive a service from the Hospital (node 16), which is an important hub. Fire Station 5 and School 9 have high values of authority because they are serviced by Bridge 3, which is also an important hub. The importance of a node in Graph Theory is closely connected with the concept of topological centrality. Referring to the illustrative example, Block 6 has the highest authority value; if a flood hit it, it would therefore affect the most central node of the network, or in other words, the node which is implicitly more privileged by the system.

In Figure 3(c), the major hubs are the elements with largest diameters: Hospital 16, Bridge 3, School 7 and Fuel Station 15. Bridge 3 is an important hub since it provides its service to Block 6, which has the highest authority value, and to Fire Station 5 and School 9. Fuel Station 15 and School 7 are also important hubs because they provide services to Block 6. The elements in the south-east part of the network inherited a relative importance (i.e. authority) from the most important hub in that area (i.e. Hospital 16). Bridge 3 is an exception to this aspect; in fact, this Bridge connects the south part (i.e. Block 6) with the north part of the city (i.e. Fire Station 5 and School 9). A flood event in the south-east part of the network would likely generate a major indirect impact on the whole system compared to other parts of the network.

This illustrative example shows how the single elements can be considered as part of the whole network and not as single separate entities. This holistic approach adds information to the traditional approach: it considers the exposed asset as whole system and it exploits the properties of single elements in order to make decisions for risk mitigation strategies.

Note that similar analyses could be carried out for other properties of the graph (e.g. betweenness) in order to obtain additional insight into the properties of the system, which could be useful for the purpose of a risk assessment. For the sake of brevity, such analyses have not been included here. A complete study of all relevant graph properties discussed above for a selected case study are presented in Part II of this manuscript.

4. Discussion

The proposed approach may bring some advantages to collective risk assessment: it provides systemic and holistic perspective, it is suitable for multi-hazard assessment, it introduce a common base for an integrated risk assessment and it promotes the study of second-order impacts and cascade effects.

The analogy between graph properties and traditional risk variables (exposure, vulnerability, and resilience) provides useful information that can enrich traditional reductionist approaches to RA. These consider collective risk as the combination of each individual element at risk, without accounting for the interaction or interdependencies between them. The new approach based on Graph Theory through which a system can be modelled as a graph, and the analysis of its properties within the RA framework, provides a systemic and holistic perspective that is missing in the traditional RA. This perspective, innovative in the context of collective risk assessment, uses both the information contained in the vertices and in the whole network.

The proposed approach is suitable for multi-hazard assessment, as the graph properties of the system are independent of the type of hazard to be analysed. Moreover, these properties can be easily integrated with the properties of the single node, already well estimated by the reductionist approaches.

The graph-based approach also introduces a common base for integrated risk assessment in terms of different features, not only in relation to the physical component of the elements but also with regard to social aspects that express the capacity of the system to respond to perturbations. Therefore, the use of Graph Theory in this field can be applied to physical, as well as social or integrated risk. In the first case, the analysis can focus on physical aspects if the graph has physical elements (buildings). The second case focuses on social aspects if the nodes represent the population with characteristics that reflect different types of vulnerability (e.g. age, education).

Lastly, the new holistic perspective introduces an important paradigm shift in the risk conceptualization: the most widely accepted risk concepts of hazard, vulnerability, exposure and resilience do not lose their validity but are integrated into a systemic perspective rather than considered separately. The graph shows properties relevant for all factors together: in fact, in the graph it is possible to study some exposure characteristics, some nodes bring information on potential cascading effects, and others convey vulnerability information of the system as a whole. The properties of the whole

graph show the studied area as a unique entity, and how the whole system together is vulnerable to an external perturbation, such as a hazardous event that can affect part of it. Beside this whole system perspective, it is also possible to assess the properties of the single parts of the graph (e.g. nodes) and detect which element, or set of elements, are more critical for the entire system.

415 One of the most important features and consequences promoted by this approach is the possibility to investigate cascade effects. The intrinsic network perspective can be applied to model cascade effects, showing the dynamic consequences of disruptive events. The links between nodes allow to pass from the direct physical damage to a broader economic and social indirect impacts. Depending on the type of disruptive event and how the network is setup in terms of nodes and links, the impact can spread throughout the network. Furthermore, during the cascade effect, it is interesting to analyse
420 the structural evolution of the network, the main properties of which are emphasised above, and study how those measurements change during the propagation of loss. If a cascade process needs few or high number of propagation steps to reach most of the nodes, this shows a lower or higher capacity to cope and adapt to the perturbation, and therefore to be less or more resilient.

5.   Final Considerations

425 This paper proposes a new approach to model the risk of complex systems based on Graph Theory. By leveraging certain analogies that can be established between graph properties and risk concepts, this approach allows obtaining a more thorough knowledge of a system compared to traditional approaches, in terms, for example, of its structure and vulnerabilities. It also allows understanding certain risk mechanisms, such as how the impacts of a hazard are propagated or where they are exacerbated, and therefore assessing the disaster risk of the system as a whole, including second-order
430 impacts and cascade effects.

The natural continuation of this study, which focuses mainly on theoretical aspects, is to implement and test the approach in case studies, verifying its feasibility. Therefore, part II of this paper presents an application to the case of urban flooding in Mexico City. Further research will aim to fully implement and integrate the graph-based approach in quantitative risk assessments, both at scenario and probabilistic level.

435 A possible extension of this framework is to model the physical hazard as one or more nodes linked to the elements at risk, rather than through a traditional approach where elements are overlapped with hazard footprints. This approach may be advantageous, as it would allow including all the factors of risk directly into the topology structure of the graph, and will be explored in further research.

Moving forward, one of the challenges that will need to be addressed is related with data requirements and availability.
440 Currently, most exposure and vulnerability databases focus on the properties of single elements, and tend to contain little to no information on the connections between them. As we have discussed, this information is key for more thoroughly understanding and assessing the risk of a system. For this reason, developing and collecting data with information related to the connections between the elements is paramount. To promote this perspective, it is necessary consider shifting the RA from using traditional relational databases to so-called graph databases. In such databases, each node contains, further
445 to the traditional characteristics, also a list of relationship records which represent its connections with other nodes. The information on these links is organized by type and direction, and may hold additional attributes.

The introduction of the network perspective of Graph Theory into the RA for collective disaster risk aims, in the long term, to be a first step for future developments of Agent Based Models and Complex Adaptive Systems in collective risk assessment. In this perspective, the nodes of the network are agents, with defined state (e.g. level of damage), and the
450 interaction between the other agents is controlled by specific rules (e.g. vulnerability and functional functions) inside the environment within they live (e.g. natural hazard phenomena).

References

Abele, W. I. and Dunn, M.: International CIIP handbook 2006 - (Vol. I) - An inventory of 20 national and 6 international critical information infrastructure protection polices, Zurich, Switzerland., 2006.

455 Albano, R., Sole, A., Adamowski, J. and Mancusi, L.: A GIS-based model to estimate flood consequences and the degree of accessibility and operability of strategic emergency response structures in urban areas, Nat. Hazards Earth Syst. Sci., 14(11), 2847–2865, doi:10.5194/nhess-14-2847-2014, 2014.

Alexoudi, M. N., Kakderi, K. G. and Pitilakis, K. D.: Seismic risk and hierarchy importance of interdependent lifelines. Methodology and important issues., in 11th ICASP International Conference on Application of Statistics and Probability
460 in Civil Engineering, August 1- 4, Zurich, Switzerland., 2011.

Arrighi, C., Brugioni, M., Castelli, F., Franceschini, S. and Mazzanti, B.: Urban micro-scale flood risk estimation with parsimonious hydraulic modelling and census data, Nat. Hazards Earth Syst. Sci., 13(5), 1375–1391, doi:10.5194/nhess-13-1375-2013, 2013.

Balbi, S., Giupponi, C., Gain, A., Mojtahed, V., Gallina, V., Torresan, S. and Marcomini, A.: The KULTURisk
465 Framework (KR-FWK): A conceptual framework for comprehensive assessment of risk prevention measures - Project deliverable 1.6., 2010.

Bazzurro, P. and Luco, N.: Accounting for uncertainty and correlation in earthquake loss estimation, 9th Int. Conf. Struct.

Saf. Reliab., 2687–2694, 2005.

Bergström, J., Uhr, C. and Frykmer, T.: A Complexity Framework for Studying Disaster Response Management, J. Contingencies Cris. Manag., doi:10.1111/1468-5973.12113, 2016.

Boccaletti, S., Latora, V., Moreno, Y., Chavez, M. and Hwang, D. U.: Complex networks: Structure and dynamics, Phys. Rep., 424(4–5), 175–308, doi:10.1016/j.physrep.2005.10.009, 2006.

Börner, K., Soma, S. and Vespignani, A.: Network Science, in Annual Review of Information & Technology, vol. 41, edited by Medford, pp. 537–607, New Jersey., 2007.

Bosetti, L., Ivanovic, A. and Menaal, M.: Fragility, Risk, and Resilience : A Review of Existing Frameworks, [online] Available from: http://i.unu.edu/media/cpr.unu.edu/attachment/2232/Assessing-Fragility-Risk-and-Resilience-Frameworks.pdf, 2016.

Bruneau, M., Chang, S. E., Eguchi, R. T., Lee, G. C., O'Rourke, D., Reinhorn, A. M., Shinozuka, M., Tierney, K., Wallace, W. A. and Winterfeldt, D. Von: A framework to quantitatively assess and enhance the seismic resilience of communites, in 13th World Conference on Earthquake Engineering, Vancouver Canada., 2004.

Bunde, A. and Havlin, S.: Fractals and Disordered Systems, Springer-Verlag Berlin Heidelberg., 1991.

Burton, C. G. and Silva, V.: Assessing Integrated Earthquake Risk in OpenQuake with an Application to Mainland Portugal, Earthq. Spectra, Vol.32(3), 1383–1403, doi:10.1193/120814EQS209M, 2015.

Cardona, O. D.: The Need for Rethinking the Concepts of Vulnerability and Risk from a Holistic Perspective : A Necessary Review and Criticism for Effective, in Mapping vulnerability: Disasters, development and people, vol. 3, pp. 37–51., 2003.

Carreño, M. L., Cardona, O. and Barbat, A.: A disaster risk management performance index, Nat. Hazards, 41(1), 1–20, doi:10.1007/s11069-006-9008-y, 2007a.

Carreño, M. L., Cardona, O. D. and Barbat, A. H.: Urban seismic risk evaluation: A holistic approach, Nat. Hazards, 40(1), 137–172, doi:10.1007/s11069-006-0008-8, 2007b.

Carreño, M. L., Cardona, O. D. and Barbat, A. H.: New methodology for urban seismic risk assessment from a holistic perspective, Bull. Earthq. Eng., 10(2), 547–565, doi:10.1007/s10518-011-9302-2, 2012.

Crowley, H. and Bommer, J. J.: Modelling seismic hazard in earthquake loss models with spatially distributed exposure, Bull. Earthq. Eng., 4(3), 249–273, doi:10.1007/s10518-006-9009-y, 2006.

Cutter, S. L., Barnes, L., Berry, M., Burton, C., Evans, E., Tate, E. and Webb, J.: A place-based model for understanding community resilience to natural disasters, Glob. Environ. Chang., 18, 598–606, doi:10.1016/j.gloenvcha.2008.07.013, 2008.

Cutter, S. L., Burton, C. G. and Emrich, C. T.: Journal of Homeland Security and Disaster Resilience Indicators for

Benchmarking Baseline Conditions Disaster Resilience Indicators for Benchmarking Baseline Conditions, J. Homel.
Secur. Emerg. Manag., 7(1), doi:10.2202/1547-7355.1732, 2010.

Dueñas-Osorio, L. and Vemuru, S. M.: Cascading failures in complex infrastructure systems, Struct. Saf., 31(2), 157–
167, doi:10.1016/j.strusafe.2008.06.007, 2009.

Dueñas-Osorio, L., Craig, J. I. and Goodno, B. J.: Probabilistic response of interdependent infrastructure networks, in 2nd
annual meeting of the Asian-pacific network of centers for earthquake engineering research (ANCER), Honolulu, Hawaii.
[online] Available from:
http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.123.2036&rep=rep1&type=pdf, 2004.

Falter, D., Schröter, K., Dung, N. V., Vorogushyn, S., Kreibich, H., Hundecha, Y., Apel, H. and Merz, B.: Spatially
coherent flood risk assessment based on long-term continuous simulation with a coupled model chain, J. Hydrol., 524,
182–193, doi:10.1016/j.jhydrol.2015.02.021, 2015.

Gallina, V., Torresan, S., Critto, A., Sperotto, A., Glade, T. and Marcomini, A.: A review of multi-risk methodologies for
natural hazards: Consequences and challenges for a climate change impact assessment, J. Environ. Manage., 168, 123–
132, doi:10.1016/j.jenvman.2015.11.011, 2016.

Gao, J., Liu, X., Li, D. and Havlin, S.: Recent progress on the resilience of complex networks, Energies, 8(10), 12187–
12210, doi:10.3390/en81012187, 2015.

Grossi, P. and Kunreuther, H.: Catastrophe Modeling: A New Aproach to Managing Risk Catastrophe Modeling, Springer
Science + Bussiness Media, Inc., Boston., 2005.

Van Der Hofstad, R.: Percolation and Random Graphs, in New Perspectives in Stochastic Geometry, edited by Wilfrid S.
Kendall and Ilya Molchanov., 2009.

IPCC: Managing the risks of extreme events and disasters to advance climate change adaptation., 2012.

Kakderi, K., Argyroudis, S. and Pitilakis, K.: State of the art literature review of methodologies to assess the vulnerability
of a "system of systems" - Project deliverable D2.9., 2011.

Karagiorgos, K., Thaler, T., Hübl, J., Maris, F. and Fuchs, S.: Multi-vulnerability analysis for flash flood risk
management, Nat. Hazards, 82, 63–87, doi:10.1007/s11069-016-2296-y, 2016.

Lane, J. A. and Valerdi, R.: Accelerating system of systems engineering understanding and optimization through lean
enterprise principles, 2010 IEEE Int. Syst. Conf. Proceedings, SysCon 2010, 196–201,
doi:10.1109/SYSTEMS.2010.5482339, 2010.

Liu, B., Siu, Y. L. and Mitchell, G.: Hazard interaction analysis for multi-hazard risk assessment: A systematic
classification based on hazard-forming environment, Nat. Hazards Earth Syst. Sci., 16(2), 629–642, doi:10.5194/nhess-
16-629-2016, 2016.

530   Menoni, S.: Chains of damages and failures in a metropolitan environment : some observations on the Kobe earthquake in 1995, J. Hazard. Mater., 86, 101–119, 2001.

Menoni, S., Pergalani, F., Boni, M. . and Petrini, V.: Lifelines earthquake vulnerability assessment: a systemic approach, Soil Dyn. Earthq. Eng., 22(9–12), 1199–1208, doi:10.1016/S0267-7261(02)00148-3, 2002.

Mingers, J. and White, L.: A Review of the Recent Contribution of Systems Thinking to Operational Research and
535   Management Science - Working paper series n.197., 2009.

Nepusz, T. and Csard, G.: Network Analysis and Visualization Author, [online] Available from: https://cran.r-project.org/web/packages/igraph/igraph.pdf, 2018.

Newman, M. E. J.: The structure and function of complex networks, Soc. Ind. Appl. Math., 45(2), 167–256, doi:10.1137/S003614450342480, 2003.

540   Ouyang, M.: Review on modeling and simulation of interdependent critical infrastructure systems, Reliab. Eng. Syst. Saf., 121, 43–60, doi:10.1016/j.ress.2013.06.040, 2014.

Pescaroli, G. and Alexander, D.: Critical infrastructure, panarchies and the vulnerability paths of cascading disasters, Nat. Hazards, 82(1), 175–192, doi:10.1007/s11069-016-2186-3, 2016.

Rinaldi, S. M., Peerenboom, J. P. and Kelly, T. K.: Identifying, understanding, and analyzing critical infrastructure
545   interdependencies, IEEE Control Syst. Mag., 21(6), 11–25, doi:10.1109/37.969131, 2001.

Schwarte, N., Cohen, R., Ben-Avraham, D., Barabási, A. L. and Havlin, S.: Percolation in directed scale-free networks, Phys. Rev. E - Stat. Nonlinear, Soft Matter Phys., 66(1), 1–4, doi:10.1103/PhysRevE.66.015104, 2002.

SFDRR: Sendai Framework for Disaster Risk Reduction 2015 - 2030. [online] Available from: https://www.unisdr.org/we/inform/publications/43291, 2015.

550   Trucco, P., Cagno, E. and De Ambroggi, M.: Dynamic functional modelling of vulnerability and interoperability of Critical Infrastructures, Reliab. Eng. Syst. Saf., 105, 51–63, doi:10.1016/j.ress.2011.12.003, 2012.

Tsuruta M. Shoji Y., Kataoka S., G. Y.: Damage propagation caused by interdependency among critical infrastructures, 14th World Conf. Earthq. Eng., (1988), 8, 2008.

555
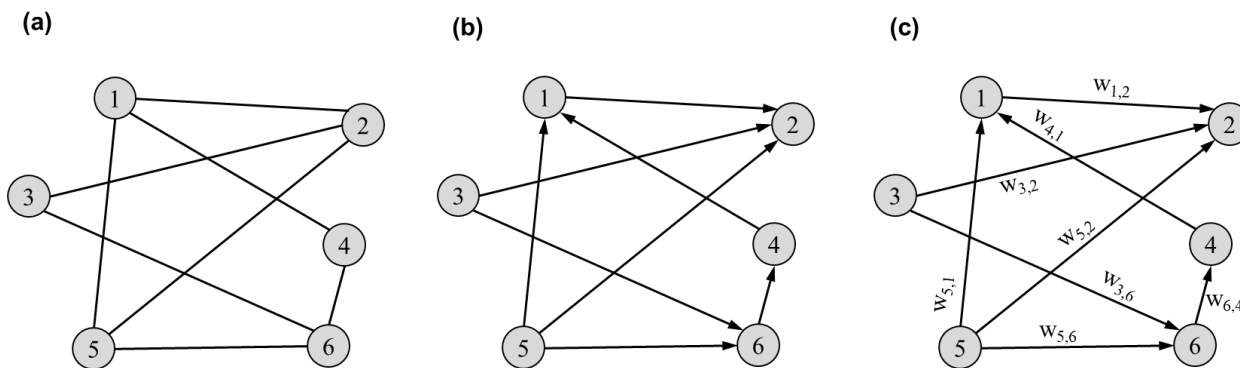
Figure 1: Graph representation of a network. (a) Undirected. (b) Directed. (c) Weighted directed.

Figure 2: Workflow.

560
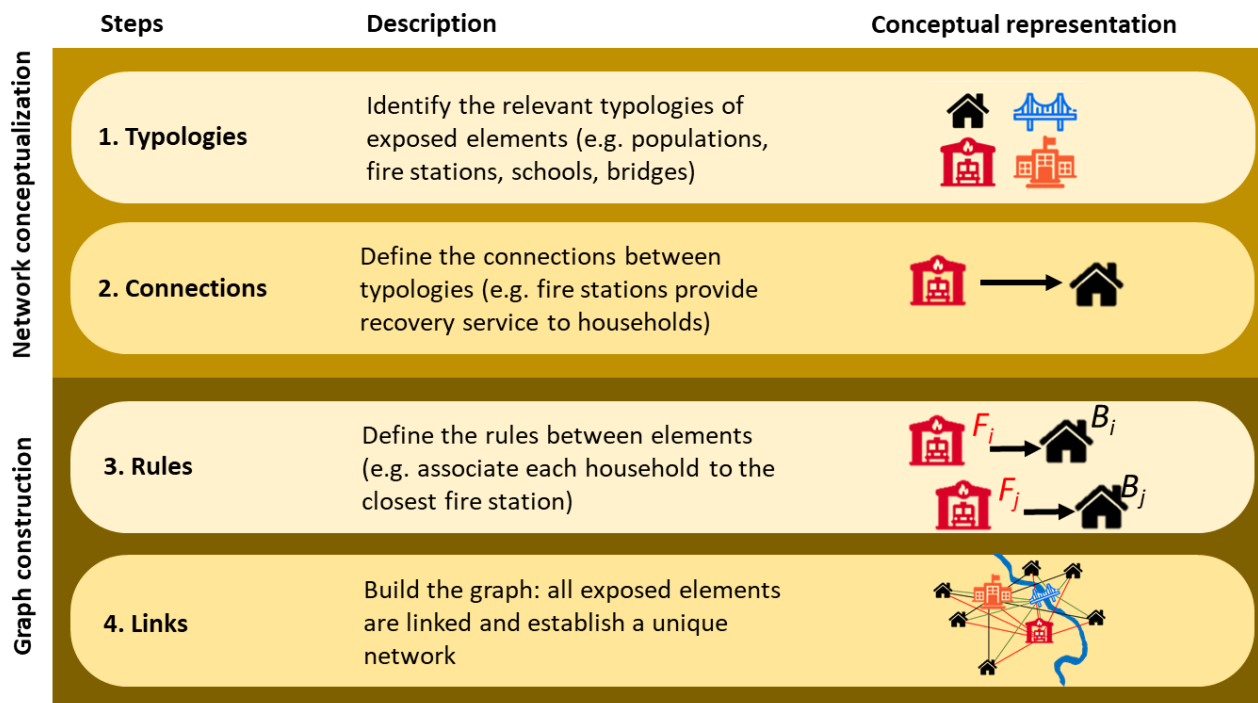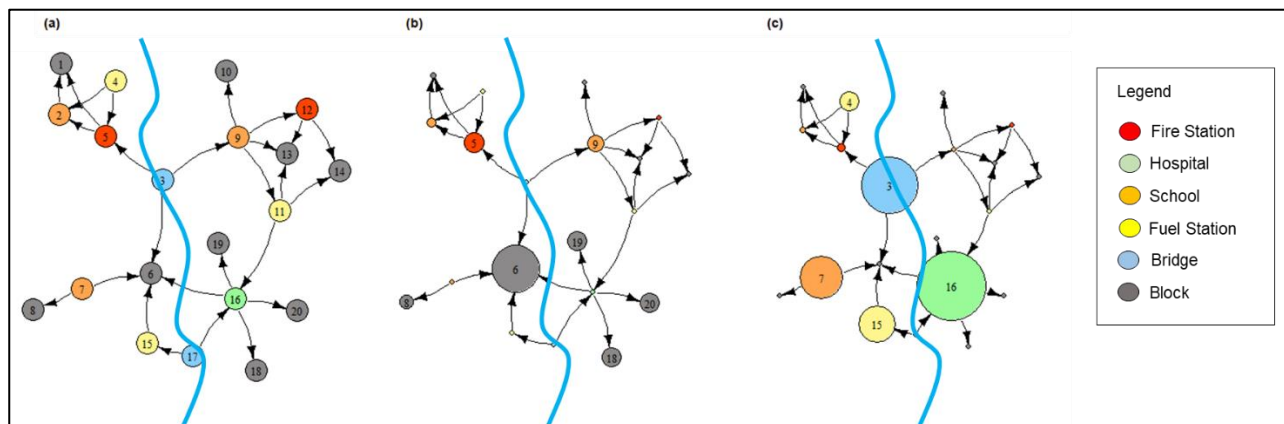
Figure 3: (a) Map of the various elements of a hypothetical municipality in a flood-prone area; (b) Same, with node sizes proportional to authority values; (c) Same, with node sizes proportional to hub values.

Table 1: Analogy of risk variables with graph properties.

Risk variables	Analogy with graph properties
Exposure	The authority represents how the system privileges the nodes, conferring them more or less importance compared with others, according to the connections established in the system.
Vulnerability	The propensity of parts of the network to be isolated because of hazard events. The closeness of a node is a measure of the single node vulnerability within the system, while degree distribution, hub, and betweenness are measures of vulnerability of the system as a whole.
Resilience	The percolation threshold, together with the network fragmentation analysis, explain the resilience of the network after a perturbation.

565