14th International Conference on Structural Safety and Reliability - ICOSSAR'25
1-6 June, 2025 - Los Angeles California, USA
R. Ghanem, E. Johnson, S. Masri and A. Olivier (Eds)

ATTRIBUTED GRAPHS PRESERVE USER IMPACTS ON NETWORK PERFORMANCE COMPUTATION

RAUL RINCON[1], JAMIE E. PADGETT[2] AND LEONARDO DUEÑAS OSORIO[3]

[1] Department of Civil and Environmental Engineering, Rice University
6100 Main St, Houston, TX 77005, USA
raul.rincon@rice.edu

[2] Department of Civil and Environmental Engineering, Rice University
6100 Main St, Houston, TX 77005, USA
jamie.padgett@rice.edu

[3] Department of Civil and Environmental Engineering, Rice University
6100 Main St, Houston, TX 77005, USA
leonardo.duenas-osorio@rice.edu

Key words: Attributed graphs, network fidelity, socio-physical modeling, connectivity-based metrics

Abstract. High-fidelity modeling of population, sectors, or goods is required for impact assessment of infrastructure systems, particularly to distribute resilience investments. Two common challenges arise in such a context. First, aggregated metrics at the system level mask the impact of service disruption on potentially vulnerable subpopulations, economic sectors, or types of goods. Second, estimating networked systems' performance over high-fidelity models becomes computationally expensive in probabilistic assessments. This paper proposes the use of *attributed networks* to improve the formulation and analysis of community resilience. Herein, 'users' are assigned to the physical network while preserving their labels and associated weights. Using the new user-physical model, we propose the extension of connectivity metrics from their network-only version to a user-weighted one, and to find 'user-physical communities', which helps uncover hidden interactions between the physical and social systems. We further use the 'communities' to reduce computational demands while keeping intact the infrastructure model. To illustrate, we use computational experiments on hypothetical infrastructure and social systems. Comparisons of the attributed-graph performance estimates with their equivalent unweighted (traditional) versions show that the former avoids the 'masking' problem. Also, the proposed metrics showed that while fidelity reduction using network-only approaches can bias decision-making for some labels (e.g., a social group), the community-based reduction conserves the information of user-physical communities, thus the expected performance per user type. By enabling the quantification of subpopulation impacts, this contribution advocates for more equitable estimates while guiding resilience investments.

1 INTRODUCTION

Analyzing infrastructure systems' performance when components suffer damage entails a large computational cost due to the combinatorial number of system states [1]. For probabilistic assess-

ments, the analysis can become impractical, and finding efficient approaches becomes of utmost importance. A common practice is to reduce the network fidelity or network representation, inducing a shrinkage of the number of system states to consider, thus lessening the computational burden. Different approaches have been proposed for such a task, including ranking-based approaches, hierarchical representations, and importance-based (user-defined) approaches [2, 3, 4].

In general, the most common techniques use ranking-based approaches. These use network topology to determine the 'importance' of the physical components. Besides, their simplicity and use of objective metrics related to structural features of the network make them desirable. These ranking-based, plus heuristics, methods have been used to inform the selection processes of network resolution [3]. After sorting the components by their rank, a set of user-defined rules may be needed to determine the final network representation. For example, defining how many of the elements must be kept for the system representation, assuring minimum connectivity, or guaranteeing reachability from special nodes (e.g., generators or origins).

Recent studies have measured the impact of modeling resolution and its relation to biased measurement of network performance [4, 5]. However, the implications of these modeling choices for final decision-making are not well understood and can even hide the needs of the final users the network serves [5]. Indeed, reducing the network fidelity has focused mostly on leveraging network information only, but the social systems' characteristics are not commonly explicitly considered. Studies that have indirectly considered users during this task have utilized the footprint of the population served or the amount of service demanded throughout the network [6]. Still, biased outcomes could go unnoticed, mostly because aggregated metrics at the system level mask the impact of service disruption on potentially vulnerable subpopulations, economic sectors, or types of goods.

In this paper, we aim to tackle the two above-mentioned challenges. That is, reducing the burden of network resilience computation while preserving the accuracy of impact estimates when measured disaggregated by user characteristics. To achieve this, we propose in Section 2 the use of *attributed networks* to improve the formulation of socio-physical systems' resilience. We present how to use labels (or classes) and weights (per label) to describe the users associated with the physical system nodes. In this way, our formulation is able to represent the heterogeneous spatially distributed users and their particular attributes into a single attributed network. Subsequently, we propose translating connectivity metrics, the connectivity loss in this study, currently focused on the physical system alone, to attribute-aware ones (noted herein as *attributed-graph performance metrics*). In Section 3 we propose finding 'user-physical communities' informed by both the characteristics of the population and the intrinsic topology of the infrastructure system. We leverage these communities to reduce the representation of the network to a subset of informative nodes, lessening the computational burden while preserving disaggregated impact estimates. We present an illustrative example of the proposed attributed-graph performance metrics and the use of user-physical communities to reduce the computational burden (see Section 4). Finally, we present conclusions and our vision on how the proposed methods advocate for more equitable estimates while guiding resilience investments.

2 CHANGING SYSTEM-LEVEL NETWORK METRICS TO PRESERVE INFORMATION OF IMPACTS ON ATTRIBUTED NODES

Infrastructure systems are represented using a graph $G = (V, E)$, with a set of nodes V and a set of edges E. The adjacency matrix A of such a network presents values $A_{ij} = 1$ if there is an edge between nodes ij, and $A_{ij} = 0$ otherwise. Due to natural hazards, progressive damage (e.g., aging), or human-induced damage, components of the network (nodes, edges, or both) may fail, inducing a loss

of functionality. While losses may differentially affect the groups of users of the network, traditional connectivity-based metrics do not capture this fact.

In Section 2.1, we present a socio-infrastructure system modeling proposal based on attributed graphs that aims to preserve the users' impact estimates. Subsequently, we select connectivity loss as a representative metric of network performance under component damage [7, 8] (Section 2.2). Finally, in Section 2.3, we demonstrate how attributed graphs enable the modification of connectivity loss to account for differential performance between user types.

2.1 Attributed graphs to represent socio-physical infrastructure systems

Let the network now be represented by $G = (V, E, K_V, K_E)$, where the node and edge weighed set of attributes are represented by K_V and K_E, respectively. In this paper, the attributes are considered to be weighed by the total number of 'users' (or amount of flow). Given a set of attributes K_V, the labels-weights for node v are described by the function $K_V(v) = [(a^{(j)}, w^{(j)})]$, $j = 1, \ldots, |K_V|$. That is, $K_V(v)$ consists on an array of $|K_V|$ tuples defined by labels $a^{(j)} \in A_v^{(j)}$ and associated weights $w^{(j)}$; $A_v^{(j)}$ is the set of labels $a^{(j)}$ can take. An equivalent description applies for $K_E(e)$.

The description of $K_V(v)$ depicts the case where (physical) nodes are associated with a unique user (see node v_5 in Figure 1a). In the socio-physical systems studied in network resilience, multiple users may rely on the functionality of their most proximal infrastructure node, resulting in node attributes represented by tuples of vectorized label-weights, i.e., $K_V(v) = [(\mathbf{a}^{(j)}, \mathbf{w}^{(j)})]$, $j = 1, \ldots, |A_V|$. Another case of aggregated node attributes occurs when the resolution of the physical network is reduced, requiring the assignment of multiple users to the new (and possibly skeletonized) representation.

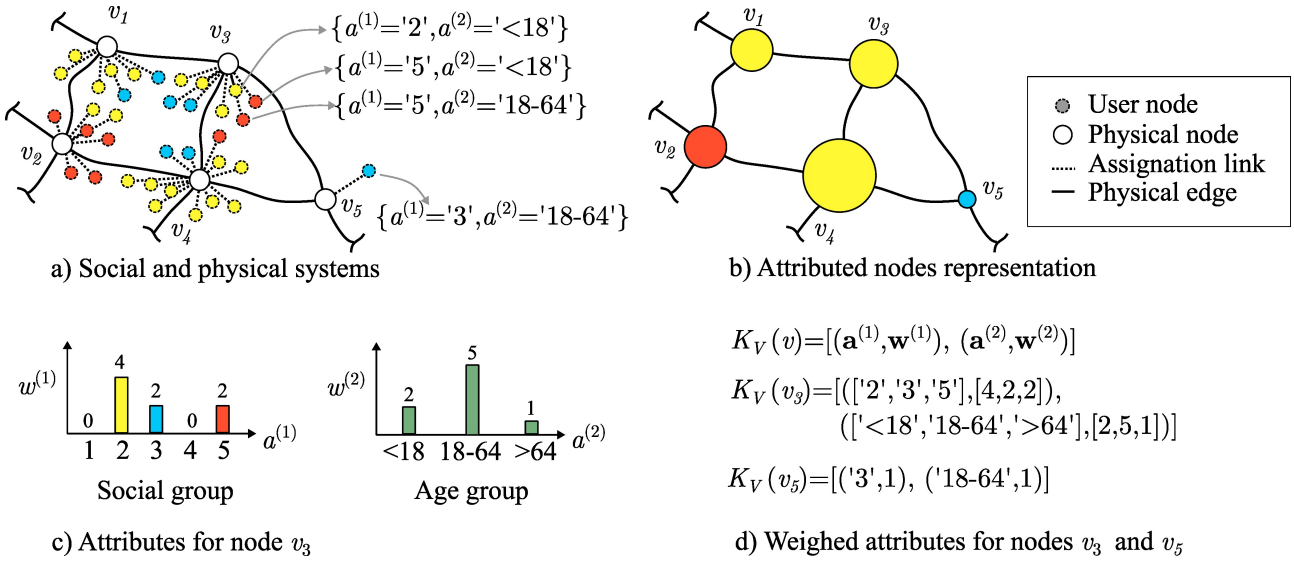

Figure 1: Attributed-graph representation of socio-physical infrastructure systems. a) Linkage between physical nodes and users based on proximity. b) Resulting node attributes for node v_3. c) Attributed graph using the colors of the most frequent user label of attribute $a^{(1)}$. d) Examples of label-weights in $K_V(v)$ for the new model

Figure 1a portrays a portion of a physical system with the surrounding population and the proposed attributed representation is presented in Figure 1b. The population is assigned to node v based on their physical proximity, preference, or any other defined type of attachment. In this example, only the nodes are weighed, but the extension to edges is obvious. The set of attributes is

$K_V = \{a^{(1)}, a^{(2)}\}$, where $a^{(1)}$:=*social group* and $a^{(2)}$:=*age group*. The attribute labels sets $A_v^{(1)}$ and $A_v^{(2)}$ are $\{1, 2, 3, 4, 5\}$ and $\{< 18, 18 - 65, > 65\}$, respectively.

As shown in Figure 1c, the proposed representation keeps all information within the node. For example, node v_3 is used by $0, 4, 2, 0$, and 2 users of attribute labels a_1 (i.e., that belong social groups 1 through 5, respectively); if attribute a_2 is observed, node v_3 users are respectively $2, 5$, and 1. Note that when aggregated into physical nodes, the dense representation of node attributes is equivalent to the set of possible labels $\mathbf{a}^{(j)} = A_v^{(j)}$. However, the vectorized tuples also enable a convenient sparse representation (as presented in Figure 1d).

The attributed network could be visualized in many different ways, leveraging the new data carried. For the sake of simplicity, in Figure 1 we use the colors of the most frequent user label of the attribute $a^{(1)}$, and the node size represents the total weight associated with this attribute (i.e. $\mathbf{w}^{(1)} \cdot \mathbf{1}$). We propose extending the concept of average connectivity loss when components are weighed using the attributed graph representation.

2.2 Network performance measured using the connectivity loss C_L

Connectivity loss C_L is a common metric that depicts the system-level impacts due to component failure. This metric depicts the average reduction in the connectivity between distribution nodes or 'target nodes' (i.e., destination of the flow) and generation or 'source' nodes (i.e., source of the flow) [8, 7]. The current definition of connectivity loss is

$$C_L = 1 - \frac{1}{|T|} \sum_{i \in T} \frac{S_i}{|S|} \tag{1}$$

where S_i represents the number of generation (source) nodes able to supply node i, where node i belongs to the set of distribution or target nodes T; S represents the set of generation nodes; and $|S|$ and $|T|$ represent the sizes of these node sets.

2.3 Attributed-graph connectivity loss metric $C_L^{(a)}$

The connectivity loss C_L metric was proposed by Albert et. al (2004) in the context of power grid systems, separating 'generator' nodes from 'distribution' nodes [8]. Herein, we work on the particular case of systems where all the nodes can be generators $S = V$ and distributors $T = V$, e.g., for transportation networks. In such a case, the number of source nodes able to supply node i (i.e., S_i) is obtained by inspecting the size of each connected component H_i, $i = 1, \ldots, h$. The nodes reachable from node i are $S_i = |H_i| - 1$, where subtracting 1 avoids double-counting the source node i (see Eq. 2) :

$$C_L = 1 - \frac{1}{n} \sum_{i=1}^{h} \frac{(|H_i| - 1)|H_i|}{(n-1)} = 1 - \sum_{i=1}^{h} \frac{(|H_i|^2 - |H_i|)}{(n^2 - n)} \tag{2}$$

where a weighted average across all the subgraphs H_i, $i = 1, \ldots, h$ is done. Note that the weighted average avoids going through each node to count their connected neighbors, but it is equivalent. In Equation 2, the normalization of $S_i = |H_i| - 1$ must be $n - 1$, representing the initial number of nodes connected to node i. Importance (or node weight) is 1 in this generalized version.

To account for the *label-weights* representation proposed in Section 2.1, we propose the attributed-graph connectivity loss metric $C_L^{(a)}$, computed for a given attribute $a \in K_V$. First, we present its version where only the total weight per node is used. That is, the estimate is not disaggregated by

user type but considers the heterogeneous users' location (through the node weights). Consider a node v is weighed by the summation of all the users (or flow type) $w_v = \mathbf{w}_v \cdot \mathbf{1}$, then the connectivity loss is computed as:

$$C_L^{(a)} = 1 - \frac{1}{W_G} \sum_{i=1}^{h} \sum_{v \in V(H_i)} \frac{W_{H_i} - w_v}{W_G - w_v} \cdot w_v \tag{3}$$

where $W_G = \sum_{v \in V} w_v$ is the total weight of the network, $W_{H_i} = \sum_{v \in V(H_i)} w_v$ is the total weight of the connected component H_i, and $V(H_i)$ is the set of nodes of H_i or the so-called neighborhood.

The attributed-graph connectivity loss $C_L^{(a)}$ can be disaggregated per label ℓ of the attribute. However, the computation in this case requires using the vectorized weights as follows:

$$C_L^{(a)} = \mathbf{1} - (\mathbf{1} \oslash \mathbf{w}_G) \odot \sum_{i=1}^{h} \sum_{v \in H_i(V)} [(\mathbf{w}_{H_i} - \mathbf{w}_v) \oslash (\mathbf{w}_G - \mathbf{w}_v)] \odot \mathbf{w}_v \tag{4}$$

where the element-wise division and product operators are $\oslash$ and $\odot$, respectively, and $\mathbf{w}_G$ and $\mathbf{w}_{H_i}$ are the vectorized versions of the total weight per label for the entire network and the connected component H_i, respectively; $\mathbf{1}$ is a row vector of ones of size L.

Equations 3 and 4 represent the relative loss of ability to communicate users, entities, or units, in general (not disaggregated), or particular to their 'type', respectively. These equations allow understanding the global performance of 'users' with shared characteristics when the network is exposed to damage. This means we have converted a physical-only metric into an attribute-aware one thanks to the proposed attributed network representation.

3 NETWORK PERFORMANCE COMPUTATION BY PRESERVING IMPACT ON USER TYPES BY MEANS OF ATTRIBUTED GRAPHS

We propose to leverage the attributed graph model to represent socio-physical systems exposed to stressors. Such representation enables investigating the complex interplay between structural features of the infrastructure, attributes of its users, and the damage and recovery of the system. Using this representation, we propose the concept of 'user-physical communities', that is, communities of attributed nodes that share commonalities in their user type as well as in the physical features of the network portion they belong to (Section 3.1). As an application of such discovered clusters, we use them to reduce the number of network analyses when computing system-level risk and resilience estimates (Section 3.2).

3.1 Finding user-physical communities in socio-physical infrastructure systems models

Let the nodes V be divided into p communities C_i, $i = 1, \ldots, p$. A given set of communities $C = \{C_1, ...C_p\}$ of the attributed network G is noted as a partition. In general, existing methods for finding communities in attributed networks commonly differ in how they balance the information given by the network connectivity and the information contained in the labels of the nodes. Depending on the assumptions and algorithm used, a partition C can also be comprised of either disjoint or overlapping sets of nodes. Once any method has been used, we expect to obtain communities $C_i \in C$ where the users within the community share attributes that relate them. For example, some user similarities can give insights into their common ability to react, respond, recover, or adapt to the loss of functionality of the physical network. More interestingly, we expect these communities to uncover

complex interactions between user attributes and the advantages (or vulnerabilities) of the physical infrastructure nodes they belong to. That is, we expect to discover hidden structural interactions between the physical and social systems.

3.2 Computation of network resilience using subsets of community members

Assume a networked infrastructure system has edges with two possible states (failure-functional) with reliability ϕ (i.e., failure probability $1-\phi$). Such a system with m edges has 2^m possible edge-network states. Without loss of generality, we assume the networked system to be an almost planar graph, where m approximately ranges between n and $3n$ (thus approximately 2^n to 2^{3n} possible states). Metrics beyond connectivity loss can involve the computation of distances among all nodes, e.g., network efficiency involves computing the total $n(n-1)$ shortest weighted paths between the n users for each network state [9]. Used for estimating the impact on users due to loss of network functionality, such a metric naively requires $n(n-1)2^m$ operations. Given that it becomes impractical, the reduction of network fidelity (say, to $n^{(low)}$ nodes and $m^{(low)}$ edges) becomes a practical way to manage the number of operations. We propose using the discovered communities to guide the selection of $n^{(low)}$ representative nodes for the computation of the *attributed-graph performance metrics*, but no reduction of the network fidelity is performed.

To guide the definition of $n^{(low)}$, we propose using $\lceil \gamma |C_i| \rceil$ as the number of nodes selected per community C_i. The parameter $\gamma \in (0,1]$ is a user-defined fraction of nodes per community. The ceiling function in the expression guarantees that, at minimum, one node is selected for the community. Hence, $n^{(low)} \in [p, n]$, where $p = |C|$ is the number of communities and n is the number of nodes in the original network. We propose selecting the $\lceil \gamma |C_i| \rceil$ nodes with the largest 'total weight' within each cluster. Then, the label-weights from the rest of the nodes are 'sent' through the network to their closest –within community– selected node. The resulting network in our approach keeps the physical network as is, but assumes the communities are agglomerated into $\lceil \gamma |C_i| \rceil$ nodes. Hence, only operations with these nodes are required when computing network performance. In this way, the computational cost is reduced (i.e., replacing n with $n^{(low)}$) while keeping intact the physical system representation.

4 ILLUSTRATIVE EXAMPLE

In this section, we illustrate the use of the proposed network representation, the computation of attributed-graph performance metrics, and how to leverage user-physical communities to reduce the computational burden. We use a hypothetical community located within the boundaries of latitude $[29.5°, 30.0°]$ and longitude $[-95.75°, -95°]$. The information of the population consists of a single attribute with six possible classes or categories $A_v^{(1)} = \{1, 2, 3, 4, 5, 6, \}$. There is a total of 1250 users per label, except for group '6' which comprises 3750 users (see Figure 2a, where each user is represented by a dot and the color represents its label). Figure 2b portrays the infrastructure system of $n = 944$ nodes and $m = 1306$ edges that serve the population. Additionally, the relevance of the edges is represented by their thickness in Figure 2c. While other methods could be used, we select the edge betweenness [10] as a method to rank the edges' relevance. In both subfigures, the social system footprint is presented behind the network.

We compare the implications of reducing the resolution of networked systems using network-only information and the case of selecting representative nodes within the user-physical communities (proposed in Section 3). With this purpose, we first determine two network representations with

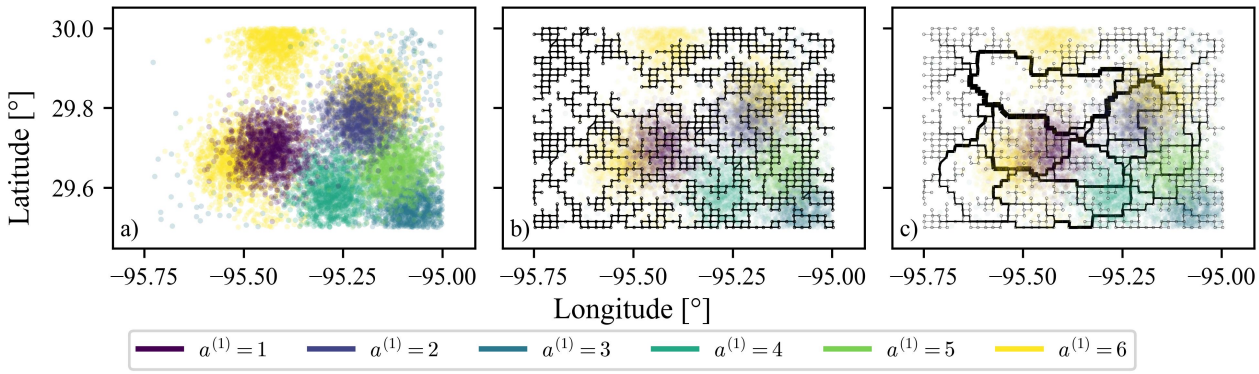

Figure 2: Hypothetical models of a) population, b) infrastructure system, and c) edge betweenness metric used to determine the physical importance of edges. Colors represent the different user categories, as depicted by the accompanying labels in a).

reduced fidelity (Figures 3b and c). The reduction is defined by selecting two different thresholds of 'edge betweenness relevance'. Figure 3a presents again the full-resolution network for comparative purposes. As shown in Figure 3, skeletonized networks result from this process, as commonly done in network resilience analysis. This example could represent a generic "origin-destination" network where representative edges are highways, for the lowest resolution, and highways plus secondary roads, for the intermediate resolution. The total number of nodes in these new representations is $n = 327$ and $n = 192$, and the number of edges is $m = 374$ and $m = 203$, respectively.

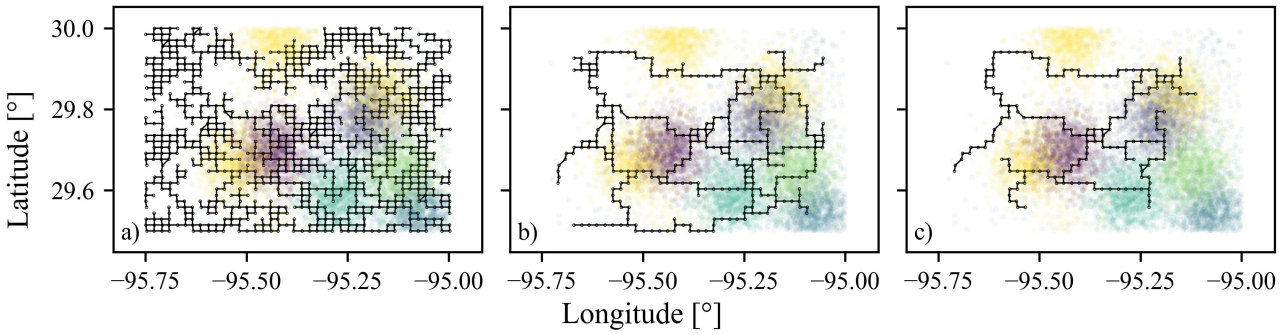

Figure 3: Physical network represented with three levels of fidelity: a) full, b) intermediate, and c) lowest resolution. The population footprint is shown for illustrative purposes.

By using networks with different levels of resolution, the initial relationship between users and the physical system may be compromised. Figures 3b and 3c show some regions with a good match between the population footprint and the networked systems, but there are others much less covered by the physical system. This means that the initial 'origin' or 'destination' nodes will not exist for some users, thus outcomes from these reduced fidelity models will assume that the population is mainly served through its most proximal nodes. The next subsection presents the attributed graphs for these reduced fidelity models and the case of the full-resolution network.

4.1 Assigning populations to networks of different levels of fidelity

In a full resolution model, one can assume users will use their most proximal physical node (see Figure 1). The resulting assignment is presented in Figure 4a. The model captures the heterogeneous and spatially distributed users and their particular attributes into a single attributed network. Note that each node contains the information disaggregated into arrays of label-weights, but a simpler strategy was used for its visualization. The colors are representative of the most frequent user label at node $v \in V$, and the node size is representative of the total node weight $\mathbf{w}_v \cdot \mathbf{1}$ (as illustrated in Figure 1).

The same must occur in a reduced resolution model. The users are attached to the most proximal physical node in the two reduced resolution networks, obtaining the networks presented in Figures 4b and c. Figure 4 shows that some nodes in the reduced fidelity models become more important in terms of the number of users served (i.e., their total weight). Such importance is artificial and a by-product of reducing the physical system resolution to a smaller number of nodes. Far from reality, this assumption may induce incorrect decision-making over populations with shared labels. For example, most users with label '4' are distant from any physical node in the lowest resolution model (Figure 3c), thus they will be mostly assigned to a few nodes (see Figure 4c). When these few edges fail, the entire community of user type '4' will become –unrealistically– disconnected. A better way is to keep infrastructure resolution and reduce the number of operations by selecting a subset of relevant nodes, as we present in the upcoming subsections.

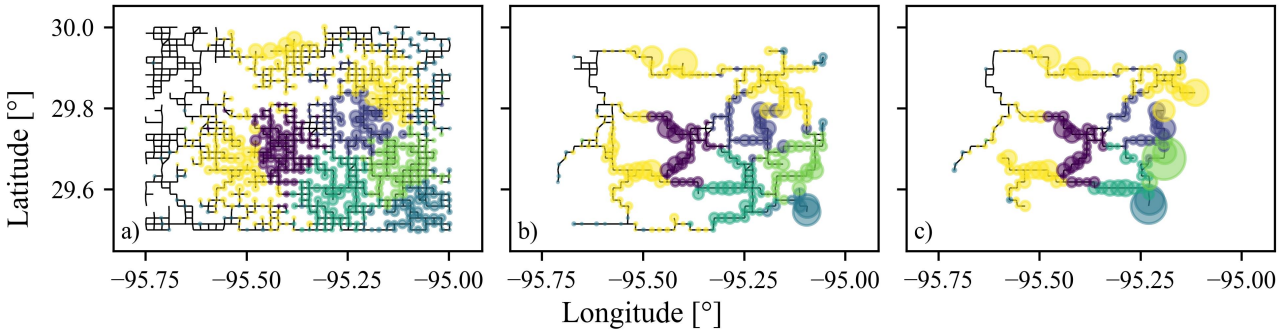

Figure 4: Population assignment to physical networks of a) full, b) intermediate, and c) low resolution.

4.2 Assigning populations based on community detection algorithms

We argue that selecting the physical nodes (for model fidelity reduction) should also consider the population served and their complex interaction with the full-resolution network model. To achieve this, community detection algorithms can be leveraged to uncover informative clusters within the attributed network representation. Different methods exist, and we propose using the Louvain Extended to Vertex Attributes (EVA) algorithm [11]. The EVA algorithm optimizes the *modularity* Q with the *purity* P of a partition C. Modularity is a well-known quality function that measures how good a partition of a network is (into 'structural modules') [12]. This metric considers modules as communities of nodes with many edges within them and fewer connections across other communities. In turn, the purity of the partitions depicts the intra-community label dispersion, i.e., the (dis)similarities in the labels of the nodes within communities [11]. In the EVA method, well-defined communities are those that present a larger frequency of the node labels (i.e., more shared characteristics) while presenting structural modularity.

The EVA method balances between the importance of modularity and purity through their linear combination using the partition 'score' Z computed as follows [11]:

$$Z = \alpha P + (1 - \alpha)Q \tag{5}$$

where α is a tuning parameter for such a balancing purpose. The authors in [11] found that $\alpha \in [0.8, 0.9]$ achieves a good balance for most of their experiments. We select $\alpha = 0.80$; increasing α in our case study produced more communities, but also a reduction in the score Z was observed.

The EVA algorithm uses node labels, but not their weight. In our problem, each node may contain users of different labels, but one label class is typically the most dominant (given the use of the high-resolution model). Hence, in order to use the EVA algorithm, we assume each node label corresponds to the most representative label within the node. A total of $p = 50$ communities are obtained from this approach (shown using different colors in Figure 5a). The obtained modularity $Q = 0.88$ and purity $P = 0.86$ are considerably high when compared with similar experiments (see for reference [11]). This means the user-physical communities (or socio-physical communities if working with network resilience) proposed herein are promising objects.

Using these communities and the procedure presented in Section 3.2, we propose reducing the number of representative nodes. Three different cases for γ are selected: $\gamma_1 = 0.20$, $\gamma_2 = 0.10$, and $\gamma_3 = 0.001$. Respectively, the different values for γ result in selecting 214, 123, and 50 nodes. Figures 5b and 5c present the selected nodes for the γ_1 and γ_3 cases, respectively. That is, each community C_i in 5a results represented by $\lceil \gamma |C_i| \rceil$ nodes in Figures 5b and 5c. Compared with Figures 4b and 4c, the distributions of most relevant nodes (in terms of their weight and label) greatly differ. Moreover, by zooming in on Figure 5, one can corroborate that there are a few nodes in the periphery that have been preserved by our method. These observations imply that connectivity loss (and other metrics such as network efficiency) are expected to diverge, or even contradict, due to the different amount of information preserved by each model reduction approach.

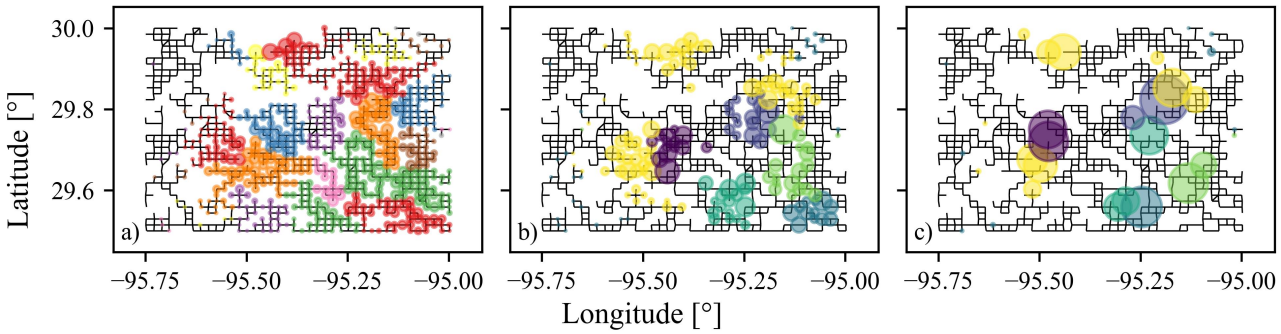

Figure 5: Community detection algorithms to discover a) socio-physical clusters in full-resolution networks and to find representative nodes with sizes dictated by b) $\gamma_1 = 0.20$ and c) $\gamma_3 = 0.001$.

4.3 Differences on attributed-graph connectivity loss due to network model fidelity

The attributed-graph performance metric proposed in this study is used to measure the interplay of the varying levels of resolution on impact estimates. The networks are subject to loss of functionality of their edges with probability $1 - \phi$. We simulate the loss of functionality of edges several times for

each $1-\phi$ value and compute, for each simulation, the attributed-graph version of connectivity loss $C_L^{(a)}$ (Figure 6). Following Rincon and Padgett (2025) [5], the different network representations are subjected to synchronous component failures to obtain objective comparisons. The synchronization is made with respect to the highest fidelity model. We first simulate the failure of edges in this model, and then we check if the same failed edges exist in the other models. If so, we eliminate those in the latter. This technique not only guarantees a systematized and objective comparison but also achieves a faster convergence in the compared estimates [13, 5].

Figure 6 presents the attributed-graph connectivity loss. The outcomes from the full-resolution model (Figure 6a) are considered as the benchmark model; hence, these are repeated across all the plots using transparent lines. The second and third sub-figures (Figures 6b and 6c) correspond to reduced resolutions presented in Figure 3b and 3c, respectively. Finally, the last three sub-figures (Figures 6d, 6e, and 6f) correspond to the outcomes using the proposed community-based representative nodes for the three values of γ. Colored lines represent the vectorized version of $C_L^{(a)}$ (see Equation 4); the black dashed line represents the weighted alone version in Equation 3. On each plot, the number of nodes n used to model the system is presented, accompanied by the number of used nodes $n^{(low)}$ (in parentheses) for the computation of the metric.

Notably, differences in the performance of separate users, in terms of the connectivity loss, are easily observable. The benchmark model portrays that the user type with the largest connectivity loss, compared to the rest, is the one shown in (user type '6'). This is somewhat expected due to their location towards the periphery of the rest of the users; however, traditional (unweighted) metrics are not capable of capturing such discrepancies. Also, note that the size of this user type is dominating the social system. Hence, the connectivity loss when only weights are observed (Equation 3) results in a very similar pattern (black dashed line). Such results demonstrated again that the attributed graph performance metrics avoid masking the impact of service disruption in the different subpopulations of users. Therefore, disaggregating the outcomes by user type is expected to contribute to debiasing the final results for informing decision-makers.

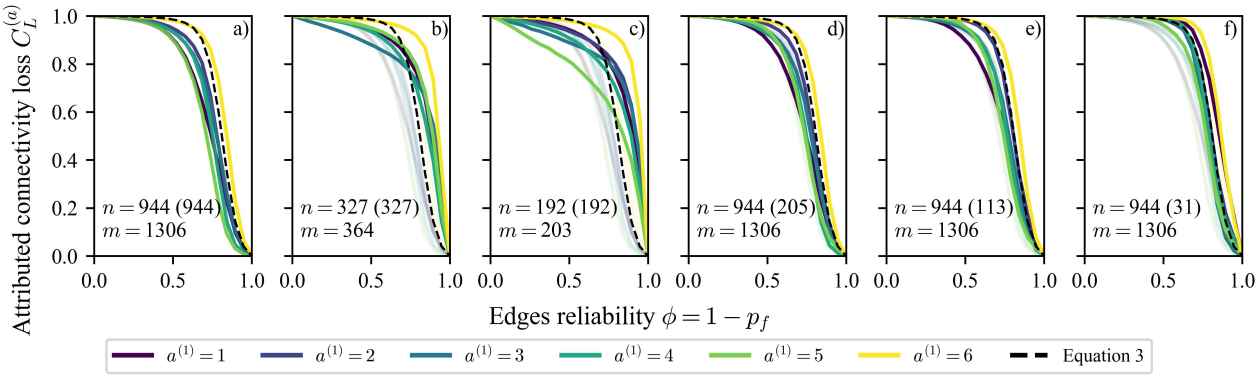

Figure 6: Attributed-graph connectivity loss $C_L^{(a)}$ as a function of the edges reliability ϕ

The above-mentioned differences can drastically vary as we change the model fidelity to represent the same socio-physical system. Figures 6b and 6c show that for the reduced fidelity in physical networks, the outcomes largely differ from the benchmark model. Not only do they diverge, but they also differentially impact the outcomes of different user types. For example, for the intermediate resolution model ($n = 327$), the connectivity loss of category 6 users is higher than expected in the full

resolution model (see Figure 6b). More critically, for the lowest resolution model ($n = 192$), values of $C_L^{(a)}$ from some groups are lower than the benchmark model, as shown in Figure 6c. Such outcomes are considered non-conservative errors and may affect future decision-making for this population. For instance, resilience planning for user type '5' could result in omitted interventions due to the alluded good performance when, in fact, they may need them.

5 CONCLUSIONS

Reductions in network fidelity may be needed for computationally expensive analyses, such as network resilience analysis for informing decision-making. Existing practices focus on using network information alone to define the low-resolution model, possibly impacting the final estimates in a disproportionate manner. Such biased model outcomes may affect the populations for which the decisions are made. This is not easily observable, given that common performance metrics do not account for disaggregated impacts by user types. To solve this, we propose the formulation of network performance computation by using attributed networks. The proposed representation method weights the nodes of the physical system with the information of the users, i.e., their attributes and the weights associated with them. Moreover, this new representation allows us to propose 1) a connectivity-based metric based on attributed graphs and 2) the use of community detection algorithms to find insightful 'user-physical communities'. The former, noted as attributed-graph connectivity loss, allows modelers and decision-makers to measure the impacts of loss of functionality, now disaggregated by user type. The latter were found to be representative of complex interplays within socio-physical systems; herein, we presented their ability to guide model fidelity reduction.

We tested the proposed network representation, the attributed-graph performance metric, and the use of user-physical communities using an illustrative example. The connectivity loss metric on attributed graphs was found capable of capturing the discrepancies between the expected impact on users who share characteristics. Moreover, this metric demonstrated that model fidelity reduction based on network-only information can hinder the outcomes observed in some subpopulations. That is, thanks to the ability to disaggregate the outcomes, discrepancies were found to be artificially accentuated by the set of modeling choices (i.e., when the network fidelity is reduced directly by eliminating components from the real system). Also, the proposed model fidelity reduction based on detected user-physical communities was shown to be a reliable method to reduce computational burden. The use of representative nodes of user-physical communities enables reducing the number of operations needed for quantifying system-level impacts without altering the expected results, especially when measured impacts can be presented disaggregated by types of users.

While the proposed user-physical communities were used to reduce computational burden, this is only one possible application. The proposed attributed-graph representation for socio-physical systems can open a path to the investigation of the relationship between heterogeneous users' features, their spatially inhomogeneous density, and the structural features of the physical network. For example, we envision investigating the complex interplay of different user attributes related to resilience (e.g., the users' ability to respond, absorb, recover, or adapt to system failures) with the topological features of the physical infrastructure. With this study, we expect to provide new tools to investigate the quantification of subpopulation impacts on networks, advocating for more equitable modeling approaches and, in turn, guiding more objective resilience investments.

6 ACKNOWLEDGMENTS

We gratefully acknowledge the support provided by the National Science Foundation under Award Numbers CMMI-2227467 and CMMI-2037545. Raul Rincon would also like to acknowledge the grant support from the Fulbright-Minciencias scholarship program.

REFERENCES

[1] Won-Hee Kang, Junho Song, and Paolo Gardoni. Matrix-based system reliability method and applications to bridge networks. *Reliab. Eng. Syst. Saf.*, 93(11):1584–1593, November 2008.

[2] Camilo Gómez, Mauricio Sanchez-Silva, Leonardo Dueñas-Osorio, and David Rosowsky. Hierarchical infrastructure network representation methods for risk-based decision-making. *Struct. Infrastruct. Eng.*, 9(3):260–274, March 2013.

[3] Fabrizio Nocera and Paolo Gardoni. Selection of the modeling resolution of infrastructure. *Comput Aided Civ Infrastruct Eng*, 37(11):1352–1367, September 2022.

[4] Navya Vishnu, Sabarethinam Kameshwar, and Jamie E. Padgett. Road transportation network hazard sustainability and resilience: correlations and comparisons. *Struct. Infrastruct. Eng.*, pages 1–21, July 2021.

[5] R. Rincon and J.E. Padgett. Bias-based comparison of sub-model fidelity and its compounded effect on infrastructure resilience estimates. *J. Eng. Mech*, 2024.

[6] Roberto Guidotti, Paolo Gardoni, and Nathanael Rosenheim. Integration of physical infrastructure and social systems in communities' reliability and resilience analysis. *Reliab. Eng. Syst. Saf.*, 185:476–492, May 2019.

[7] Leonardo Dueñas-Osorio, James I. Craig, and Barry J. Goodno. Seismic response of critical interdependent networks. *Earthq Eng Struct Dyn*, 36(2):285–306, 2007.

[8] Réka Albert, István Albert, and Gary L. Nakarado. Structural vulnerability of the North American power grid. *Physical Review E*, 69(2):025103, February 2004. American Physical Society.

[9] Vito Latora and Massimo Marchiori. Efficient Behavior of Small-World Networks. *Physical Review Letters*, 87(19):198701, October 2001. Publisher: American Physical Society.

[10] LiYing Cui, Soundar Kumara, and Réka Albert. Complex Networks: An Engineering View. *IEEE circuits and systems magazine (New York, N.Y. 2001)*, 10(3):10–25, 2010. IEEE.

[11] Salvatore Citraro and Giulio Rossetti. Eva: Attribute-Aware Network Segmentation. In Hocine Cherifi, Sabrina Gaito, José Fernendo Mendes, Esteban Moro, and Luis Mateus Rocha, editors, *Complex Networks and Their Applications VIII*, pages 141–151. Springer International Publishing, 2020.

[12] M. E. J. Newman and M. Girvan. Finding and evaluating community structure in networks. *Physical Review E*, 69(2):026113, February 2004. American Physical Society.

[13] P. L'Ecuyer. Efficiency improvement and variance reduction. In *Proceedings of Winter Simulation Conference*, pages 122–132, December 1994.