

## SUMMARY

The internet, navigation, power grids, social interactions, cells etc. are complex real-world systems that can be modeled with networks. A network consists of nodes (e.g., computers on the internet, people in a society, or genes in a cell) and interactions (or relationships) between the nodes. Network science, which deals with analyses of complex real-world networks, is an incredibly exciting field. One type of network analysis is network alignment, which aims to compare two or more networks in order to find regions of their topological or functional similarities. In our study, network alignment is defined as an injective (one-to-one) mapping between the compared networks [1]. Network alignment can only be approximated as it falls into the category of NP-hard problems. Existing (approximate) network alignment methods allow for studying only static networks [2]. However, most of real-world complex systems are dynamic. With the recent increase of temporal network data available about such dynamic systems, it only seems natural to use the time information in the process of network alignment [3]. Hence, we aim to generalize an existing state-of-the-art method for alignment of static networks to its dynamic counterpart, in order to allow for aligning dynamic networks.

## NETWORK SCIENCE

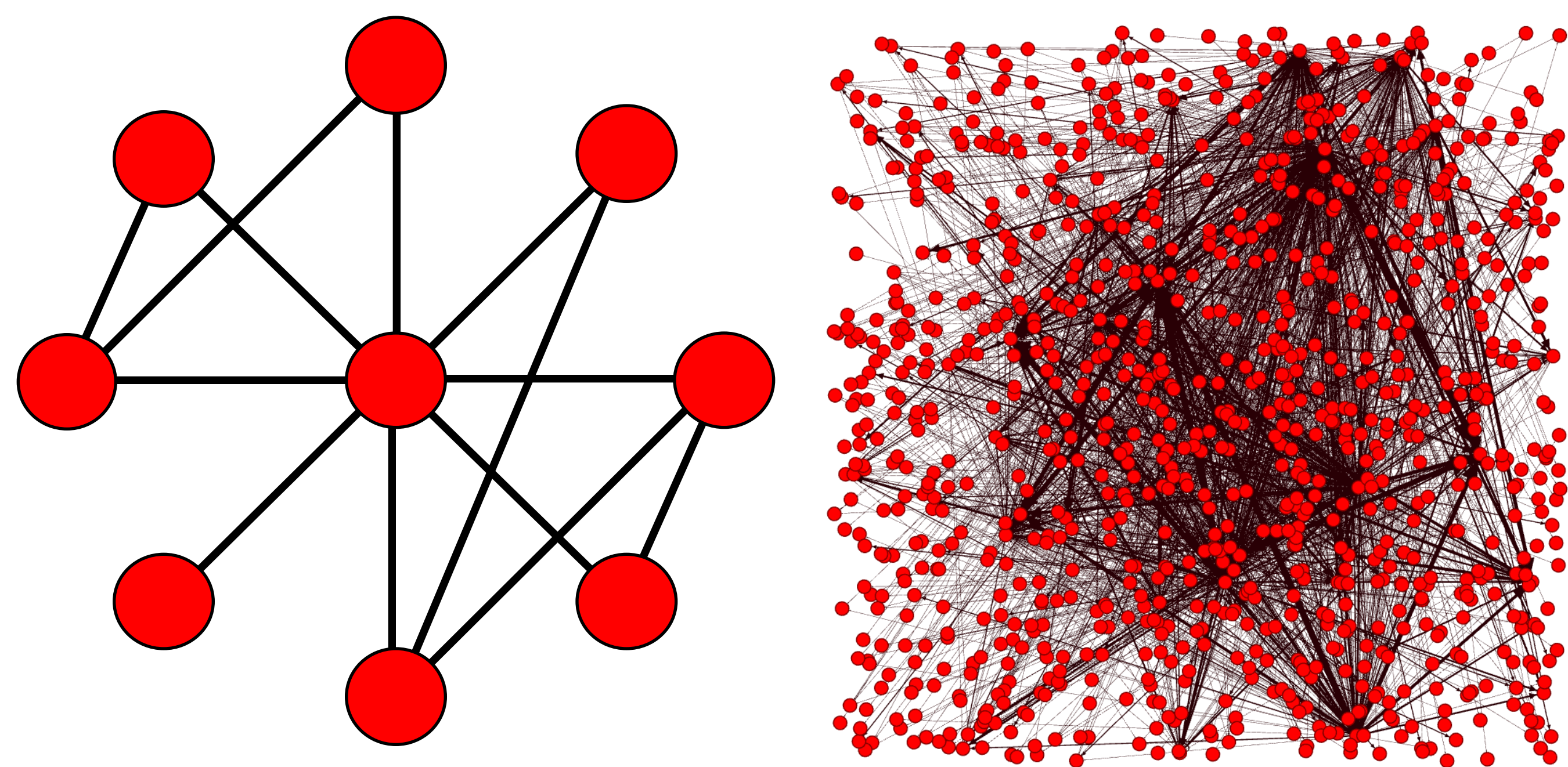


Figure 1: (a) Simple network with 9 nodes and 14 edges, used for demonstrations and concept building which can then be applied to complex networks. (b) Complex network with 1000 nodes and 2000 edges, representative of a real world social network with interactions between people.

Most real world systems can be modeled as networks making Network Science incredibly important. This field has endless problems, where one is more interesting than the next. Although there are many ways of analyzing networks, we focus on the problem of network alignment.

## TRADITIONAL (STATIC) NETWORK ALIGNMENT

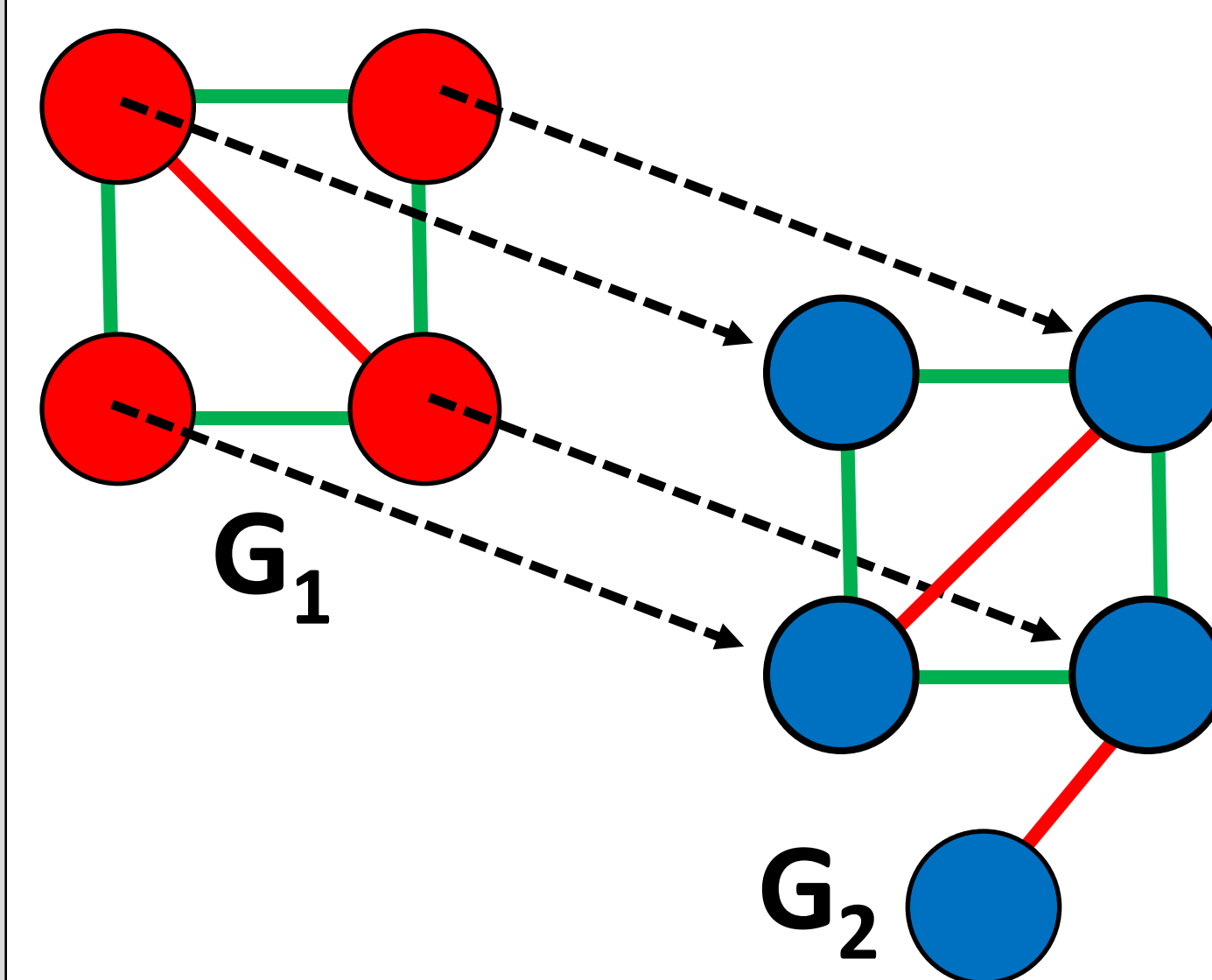


Figure 2: Network alignment between static network  $G_1$  and  $G_2$  is an injective mapping between their nodes, as illustrated in the diagram. Edges marked in green represent conserved edges and edges in red are non-conserved edges. The  $S^3$  quality alignment measure [1] quantifies the amount of conserved edges, as the ratio of the number of such edges out of both conserved and non-conserved edges.

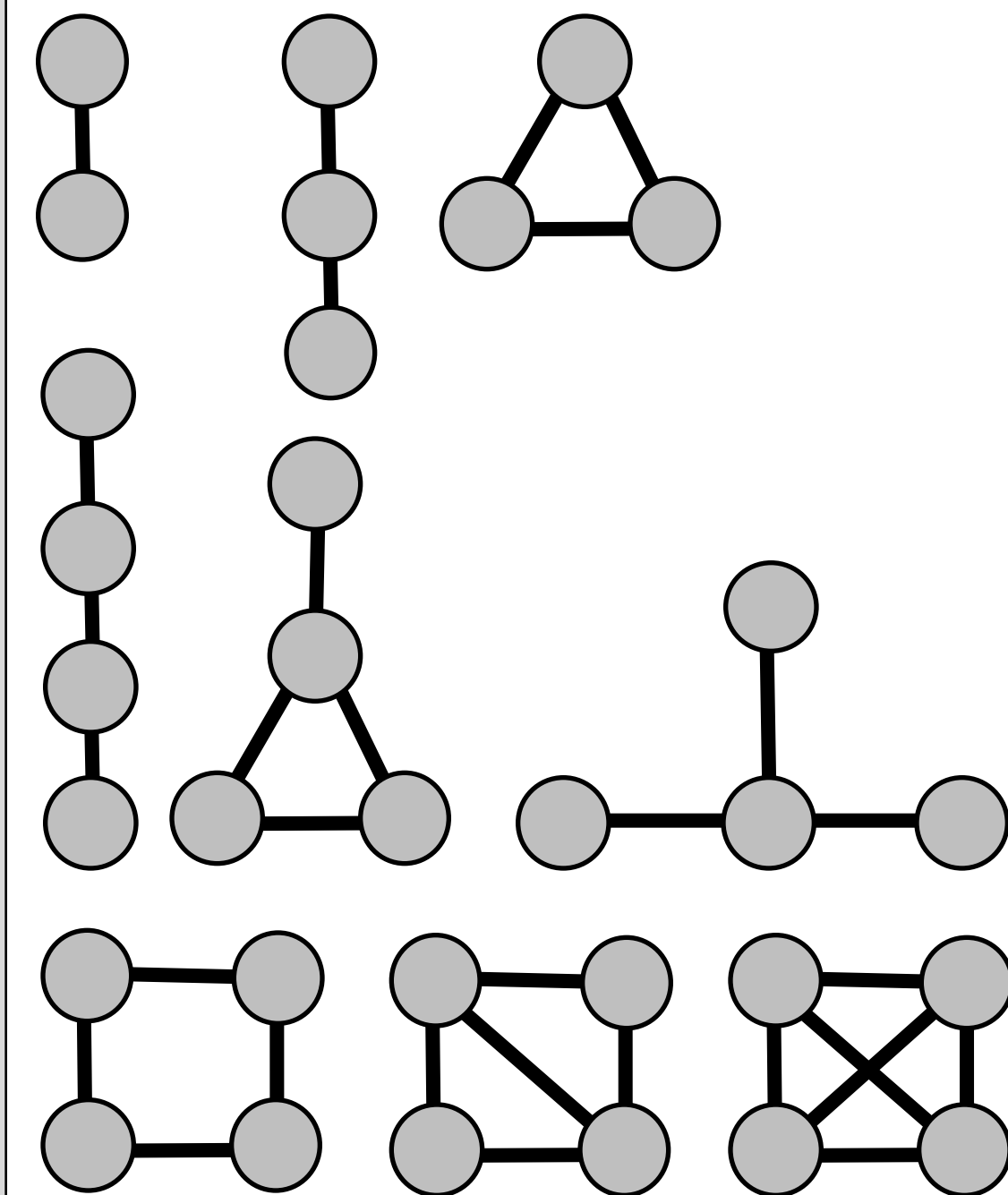


Figure 3: Induced subgraphs called graphlets [4]. All possible graphlets with 2, 3, and 4 nodes are shown. In addition to edged conservation (Figure 2), network alignment aims to optimize node conservation, which can be captured by comparing "graphlet signatures" of two nodes from the networks being aligned. This idea can be extended to a dynamic approach. In this case we would compare "dynamic graphlet signatures" of two nodes.

## DYNAMIC (TEMPORAL) NETWORKS

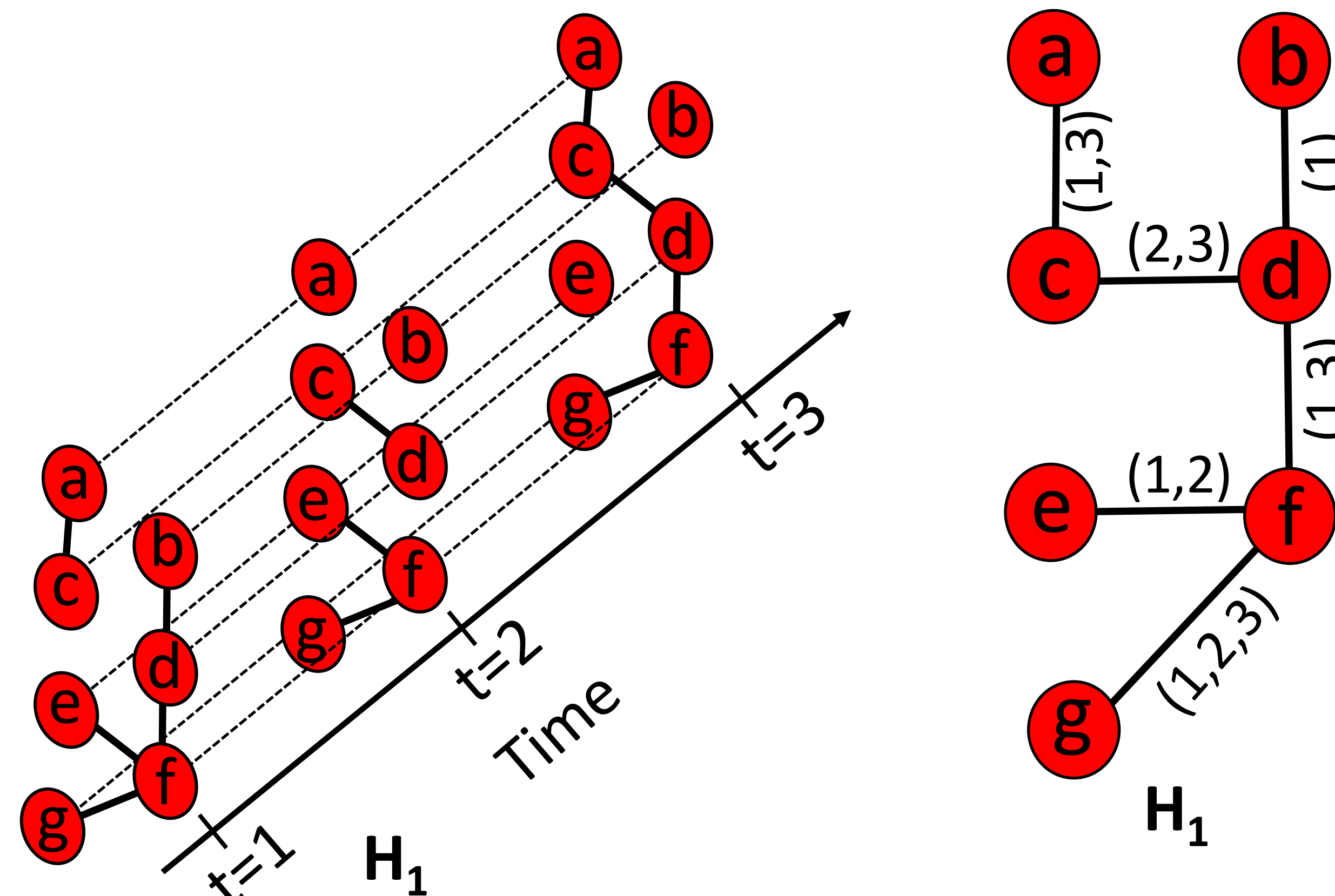


Figure 4: (a) A snapshot-based representation of a temporal network ( $H_1$ ). At each time-step, nodes that are present and edges that are active at that particular point in time are shown [4]. (b) Time-stamp decorated, aggregate graph representation of ( $H_1$ ) [3]. An edge is existing between two nodes if at least one event occurs between those nodes and the edges are labeled with the timestamps of all events that occur between the corresponding nodes. For example, edge (a,c) is active at times 1 and 3 but not at 2.

## OUR DYNAMIC NETWORK ALIGNMENT METHOD

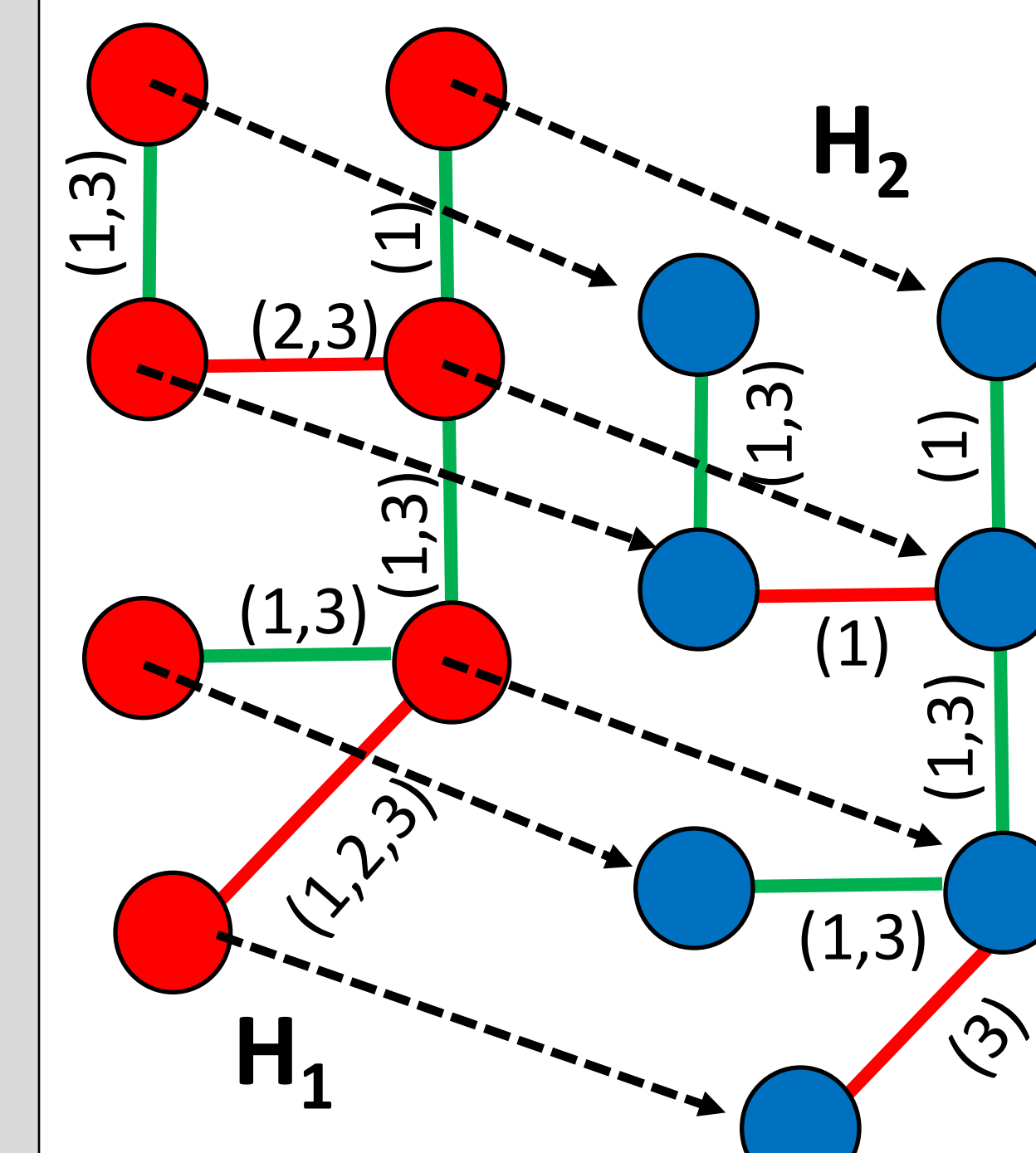


Figure 5: Illustration of dynamic network alignment between temporal networks  $H_1$  and  $H_2$ . Each network is shown as a time-stamp decorated, aggregate graph. Clearly, the static "backbones" of  $H_1$  and  $H_2$  are the same, but the networks are different when accounting for the temporal information about the events. Thus, while aligning the static network backbones would mistakenly identify the networks as identical, we propose to align the dynamic network representations by optimizing dynamic edge (i.e., event) conservation as well as dynamic graphlet-based node similarities. Edges marked in green are conserved durations and edges marked in red are non-conserved durations.

conservation as well as dynamic graphlet-based node similarities. Edges marked in green are conserved durations and edges marked in red are non-conserved durations.

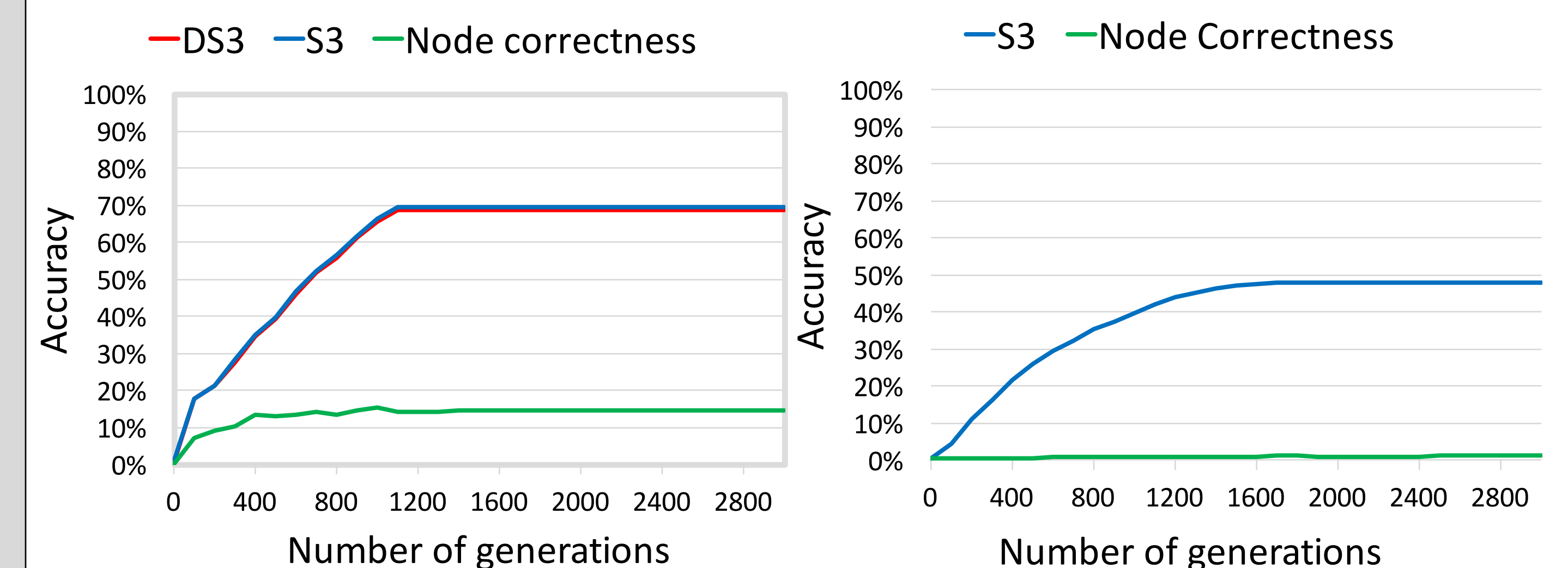


Figure 6: Preliminary results of (a) our proposed approach for dynamic network alignment that optimizes dynamic graphlet-based node similarities and dynamic edge conservation ( $DS^3$ ) and (b) its static counterpart. These results indicate that: 1) Our dynamic approach converges faster than its static counterpart, as it reaches optimal performance in fewer generations. 2) Our approach results in higher accuracy than its static counterpart with respect to both the measures that it optimizes, such as  $DS^3$ , and those that the static approach optimizes, such as  $S^3$ , as well as with respect to independent measures that none of the two approaches optimizes, such as node correctness with respect to the actual (true) node mapping.

## REFERENCES

1. Saraph, V. and Milenković, T. (2014). Bioinformatics, 30(20).
2. Vijayan, V., Saraph, V., and Milenković, T. (2015). Bioinformatics, 31(14).
3. Holme, Petter (2015). Eur. Phys. J. B.
4. Hulovatyy, Y., Chen, H., and Milenković, T. (2015) Bioinformatics, 31(12)

## ACKNOWLEDGEMENTS

- Dr. Tijana Milenković for her guidance and opportunity to work on this project as a part of the Complex Networks Lab
- Data Intensive Scientific Computing REU
- National Science Foundation Grant No. 1560363