I. Introduction

Ensembles of graphs arise in several natural applications, like disease spreading, mobility tracking, and social networks. Mining such ensembles of graphs is an increasingly important problem, especially in the current 'big-data' era. A typical issue faced by many existing techniques is to identify subgraphs and nodes of interest in these graphs. A well-studied problem is to compute frequent dense subgraphs (see Section II). For example, a group of people who are well-connected in a large fraction of the graphs may indicate a close-knit community.

Here, we study the opposite problem. Consider a set of nodes such that the subgraphs induced by this group show considerable variation across an ensemble of networks. We can think of such a group as an "unstable community." The following example involves the social network of three characters in the popular TV show *Friends*: Chandler, Joey and Kathy. Chandler and Joey experience breakup and renewal of their friendship because of their dating and/or breakup with Kathy, which is reflected in their social networks (see Figure 1). Over time, these three characters induce as many as five of the eight possible subgraphs among them. Such unstable communities can arise in many different applications including contact networks, communication networks and citation networks (see Experiments in Section VI). Moreover, several such groups may be present in an ensemble of networks. Formalizing the concept of unstable communities and computing them efficiently are challenging tasks that have not been addressed by the data mining community.

The frequency distribution of the five subgraphs in Figure 1 is almost uniform (three graphs occur once and two graphs occur twice). Intuitively, an unstable community should induce highly varying subgraphs in an ensemble of networks and these subgraphs should have near-uniform frequency distribution. Hence, the "instability" of such a community may be captured by computing the difference between the frequency distribution of its induced subgraphs and the uniform distribution. We use this idea to propose the first formal definitions of unstable communities in the context of network ensembles. We make the following contributions in this paper:

(a) **Problem Formulation:** We introduce a new class of graph mining problems based on two novel definitions of unstable communities (or, UCs, in short) (Section III). We also compare these definitions and discuss their trade-offs (Section IV).
(b) **Algorithm:** We show that our definitions have the desirable property of anti-monotonicity, which we leverage to develop effective and efficient algorithms to mine UCs (Section V).
(c) **Results:** We apply our algorithms to diverse datasets such as phone call, citation, and communication networks. We show the usefulness of our UCs in summarizing structural variations in these networks (Section VI-C3).

II. Related Work

We survey related work in this section. Briefly, none of the research discussed below considers structural variation in network ensembles (including time-varying social networks) to find informative and interesting sets of nodes.

Mining Static Graphs: This work can be grouped into three levels. First, on the graph level, representative works include

978-1-4673-8493-3/15 $31.00 © 2015 IEEE
DOI 10.1109/ICDMW.2015.87