

Community Embeddings with Bayesian Gaussian Mixture Model and Variational Inference

Anton I. N. Begehr

Graduate School of Business

National Research University Higher School of Economics

Moscow, Russia

a.begehr@fu-berlin.de

Prof. Dr. Petr Panfilov

Graduate School of Business

National Research University Higher School of Economics

Moscow, Russia

ppanfilov@hse.ru

Abstract—Graphs, such as social networks, emerge naturally from various real-world situations. Recently, graph embedding methods have gained traction in data science research. The graph and community embedding algorithm ComE aims to preserve first-, second- and higher-order proximity. ComE requires prior knowledge of the number of communities K . In this paper, ComE is extended to utilize a Bayesian Gaussian mixture model with variational inference for learning community embeddings (ComE BGMM+VI), similar to ComE+. ComE BGMM+VI takes K as the maximum number of communities and drops components through the trade-off hyperparameter weight concentration prior. The advantage of ComE BGMM+VI over the non-Bayesian ComE for an unknown number of communities K is shown for the small Karate club dataset and explored for the larger DBLP dataset.

Index Terms—graph, embedding, community embedding, ComE, Bayesian, variational inference, Gaussian mixture, expectation maximization

I. INTRODUCTION

REFERENCES