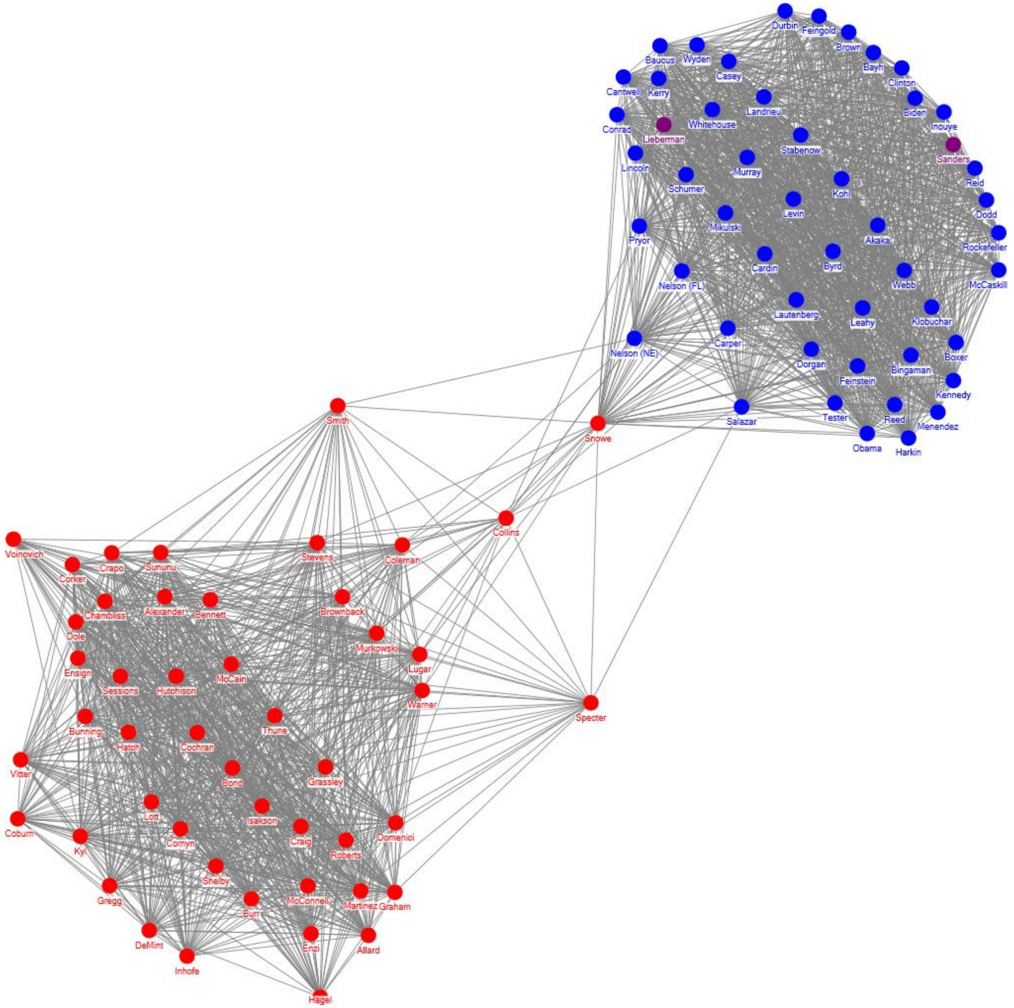
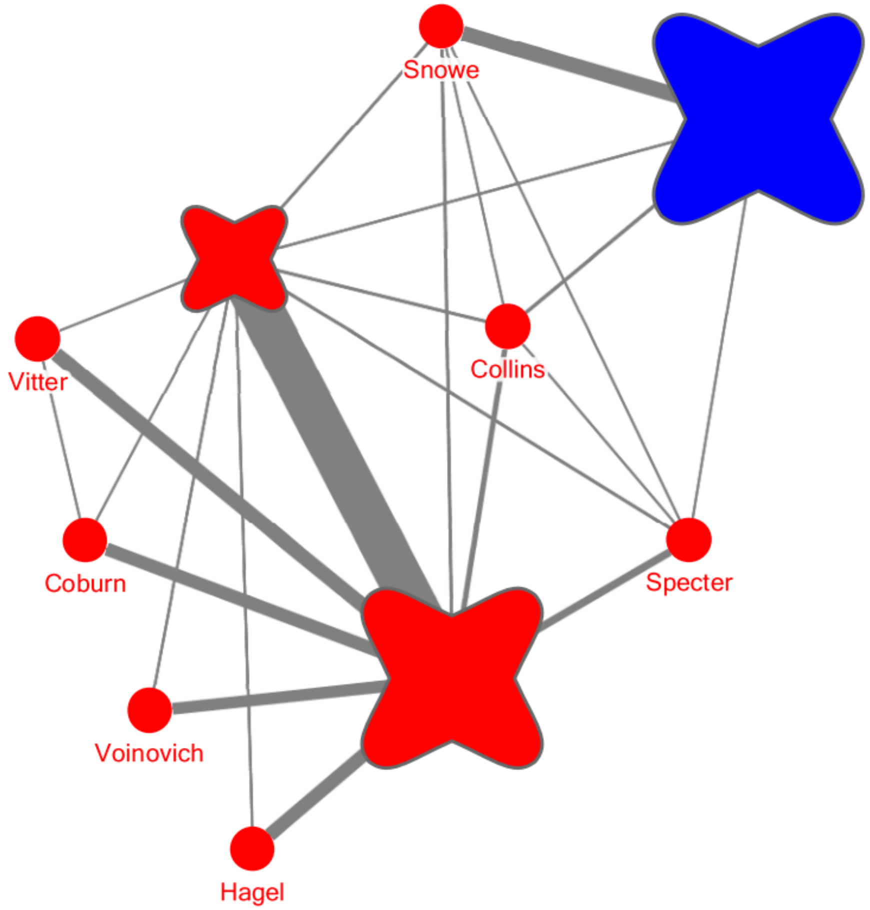


CS 7280-03 Special Topics on Visualization in Network Science

Lecture 9



Professor Cody Dunne

https://codydunne.github.io/cs7280-f16/
c.dunne@northeastern.edu

Homework 3

https://codydunne.github.io/cs7280-f16/hw/Homework-3-D3-spring-layout

D3 Force Layout Tutorial

Project Discussion

https://codydunne.github.io/cs7280-f16/project

Discussion: MatLink

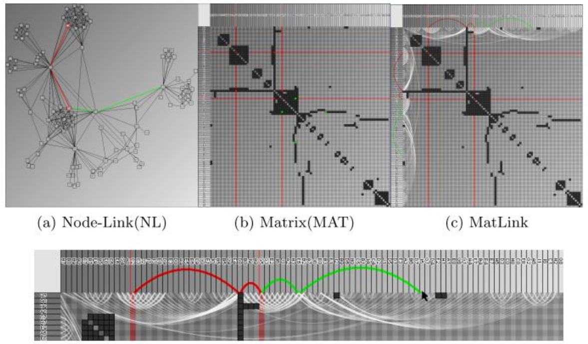
MatLink: Enhanced Matrix Visualization for Analyzing Social Networks

Nathalie Henry[1,2] and Jean-Daniel Fekete[1]

[1] INRIA/LRI, Bat. 490, Univ. Paris-Sud, F91405, Orsay, France
[2] Univ. of Sydney, Australia

Abstract. Visualizing social networks presents challeges for both node-link and adjacency matrix representations. Social networks are locally dense, which makes node-link displays unreadable. Yet, main analysis tasks require following paths, which is difficult on matrices. This article presents MatLink, a hybrid representation with links overlaid on the borders of a matrix and dynamic topological feedback as the pointer moves. We evaluated MatLink by an experiment comparing its readability, in term of errors and time, for social network-related tasks to the other conventional representations on graphs varying in size (small and medium) and density. It showed significant advantages for most tasks, especially path-related ones where standard matrices are weak.

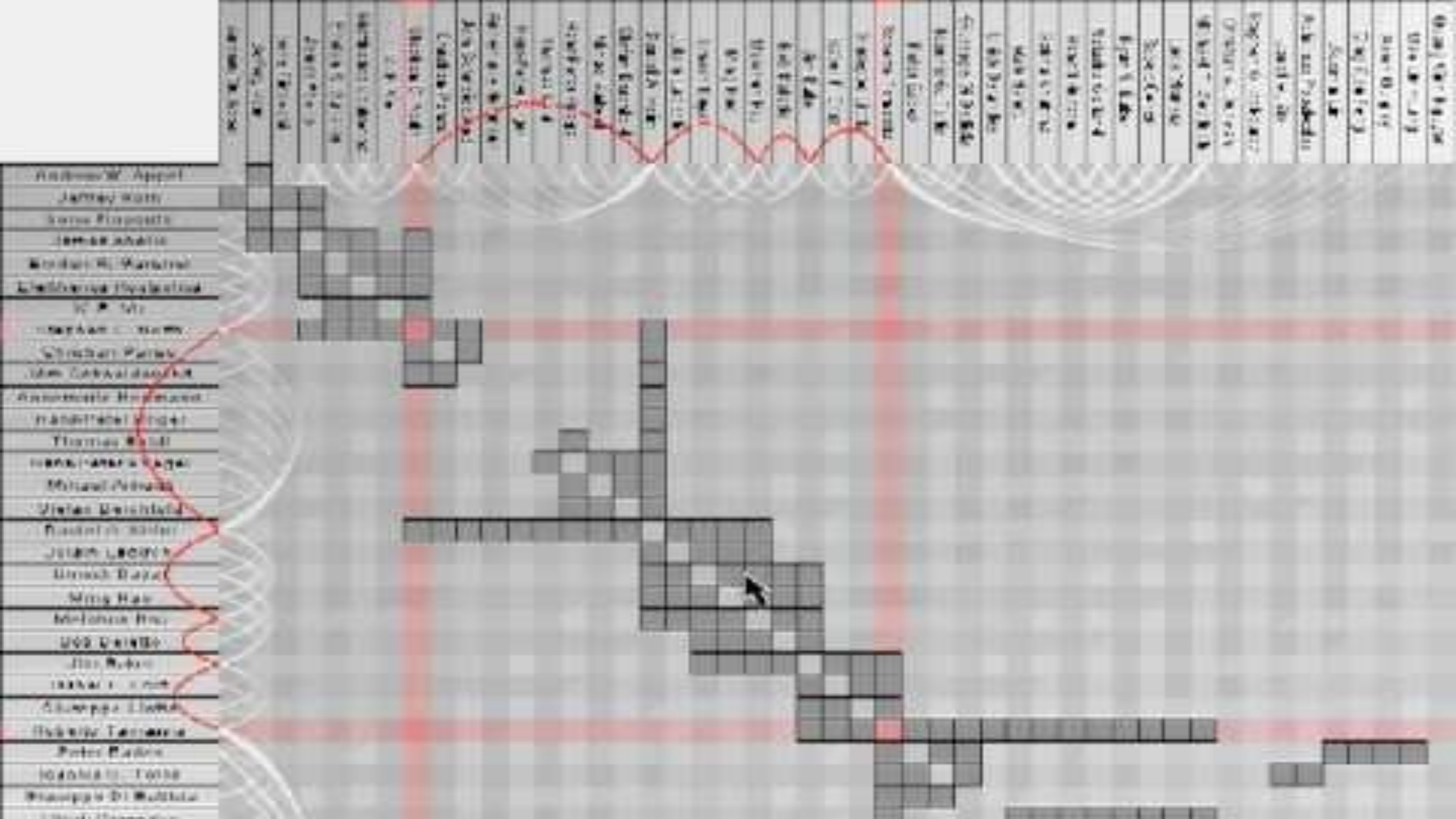
Key words: Node-Link Diagram, Matrix Visualization, Social Network.

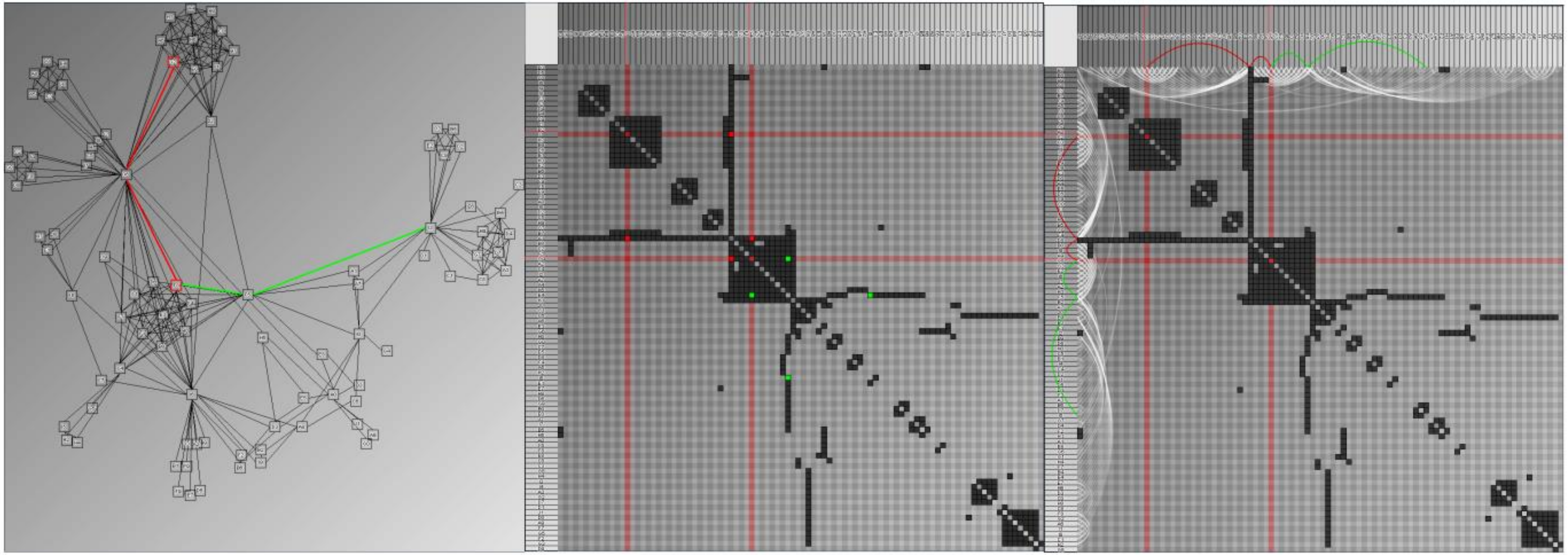


(a) Node-Link(NL) (b) Matrix(MAT) (c) MatLink

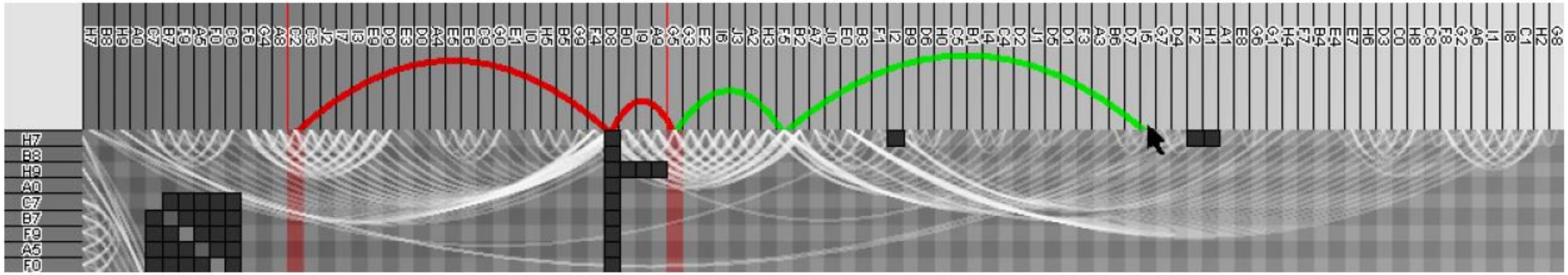
(d) Zoom on MatLink

Fig. 1: Three representations of a social network.





(a) Node-Link(NL) (b) Matrix(MAT) (c) MatLink



(d) Zoom on MatLink

MatLink slides.pdf

Henry (Riche) et al., 2007

Discussion: NodeTrix

NodeTrix: A Hybrid Visualization of Social Networks

Nathalie Henry, Jean-Daniel Fekete, and Michael J. McGuffin

Abstract— The need to visualize large social networks is growing as hardware capabilities make analyzing large networks feasible and many new data sets become available. Unfortunately, the visualizations in existing systems do not satisfactorily resolve the basic dilemma of being readable both for the global structure of the network and also for detailed analysis of local communities. To address this problem, we present NodeTrix, a hybrid representation for networks that combines the advantages of two traditional representations: node-link diagrams are used to show the global structure of a network, while arbitrary portions of the network can be shown as adjacency matrices to better support the analysis of communities. A key contribution is a set of interaction techniques. These allow analysts to create a NodeTrix visualization by dragging selections to and from node-link and matrix forms, and to flexibly manipulate the NodeTrix representation to explore the dataset and create meaningful summary visualizations of their findings. Finally, we present a case study applying NodeTrix to the analysis of the InfoVis 2004 coauthorship dataset to illustrate the capabilities of NodeTrix as both an exploration tool and an effective means of communicating results.

Index Terms—Network visualization, Matrix visualization, Hybrid visualization, Aggregation, Interaction.

1 Introduction

Social network analysis is a growing area of the social sciences. Vast new datasets are becoming available as people conduct ever more of their social lives electronically. Online projects such as Wikipedia or open-source software development are creating new social networks on a global scale. At the same time, the challenges of a more integrated world generate new demands for analysis such as monitoring terrorist networks or the spread of potentially pandemic diseases. Social network visualization is becoming a popular topic in information visualization, generating more and more tools for analysts. In 2006, 10 network-related articles (or 30%) were presented at the InfoVis Symposium and 6 at the VAST symposium. The large majority of network visualization systems use the node-link representation: 54 (out of 55) node-link based systems are referenced in the Social Network Analysis Repository (`http://www.insna.org/`), and 49 (out of 52) on the Visual Complexity website (`http://www.visualcomplexity.com/`). This representation is well suited to show sparse networks, but social networks are known to be globally sparse and locally dense. Therefore, social network visualization faces a major challenge: obtaining a readable representation for both the overall sparse structure of a social network and its dense communities.

In this article, we propose a novel visualization called *NodeTrix* to address this challenge. NodeTrix integrates the best of the two traditional network representations by using node-link diagrams to visualize the overall structure of the network, within which adjacency matrices show communities.

The article is organized as follows: after the related work section, we describe the NodeTrix representation and the data structure we rely on. We then detail the interaction techniques we designed for creating a NodeTrix hybrid, either by starting from a standard node-link diagram or from a standard adjacency matrix. Finally, we describe a case study using NodeTrix to explore and present the results of a co-authorship social network.

- *Nathalie Henry is with INRIA Futurs/University Paris-Sud, France and University of Sydney, Australia, E-mail: nathalie.henry@lri.fr*
- *Jean-Daniel Fekete is with INRIA Futurs, France, E-mail: jean-daniel.fekete@inria.fr.*
- *Michael J. McGuffin is with Ontario Cancer Institute and University of Toronto, Canada, E-mail: mjmcguff@cs.toronto.edu.*

Manuscript received 31 March 2007; accepted 1 August 2007; posted online 2 November 2007.
For information on obtaining reprints of this article, please send e-mail to: tvcg@computer.org.

Fig. 1: NodeTrix Representation of the largest component of the InfoVis Co-authorship Network

2 Related Work

In the rest of this article, we use "graph" to refer to the topological structure with no associated attributes and "network" for a graph with an arbitrary number of attributes associated with its vertices and edges. "Vertices" and "edges" refer to topological features while "nodes" and "links" refer to their visual counterparts in node-link diagrams. For matrices, "rows" and "columns" refer to the visual representation of vertices and "cells" are the visual representation of edges.

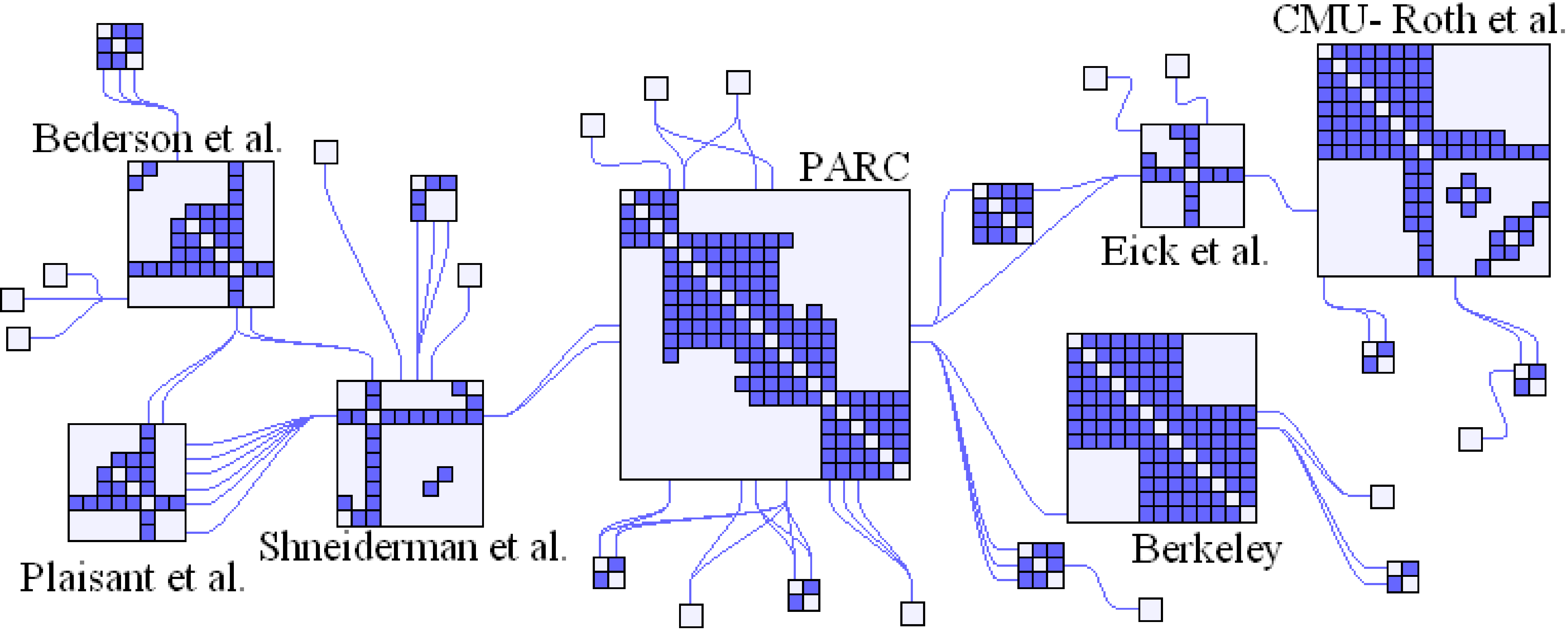
2.1 Social Network Analysis

Social networks are graphs where the vertices are actors (people) and the edges are relationships. They vary from very sparse (genealogical trees) to very dense (exports and imports between countries). *Small-world networks* belong to an intermediate category that occurs very frequently in social networks, including many acquaintanceship networks as well as the global Internet. They are the focus of many studies [23, 20] because of their interesting properties [33]. For social network visualization, the most relevant of these properties are a high clustering coefficient, corresponding to the presence of many *locally dense* clusters, and a small cross-section, caused by a small number of hub vertices connecting communities in a graph that is *globally sparse*.

Social network analysis relies on three important tasks [31, 27]:

- (T1) identify *communities*, *i.e.* cohesive groups of actors that are strongly connected to each other;
- (T2) identify *central actors*, *i.e.* actors linked to many others or that bridge communities together;
- (T3) analyze roles and positions — these are higher level tasks relying on the interpretation of groups of actors (positions) and connection patterns (roles).

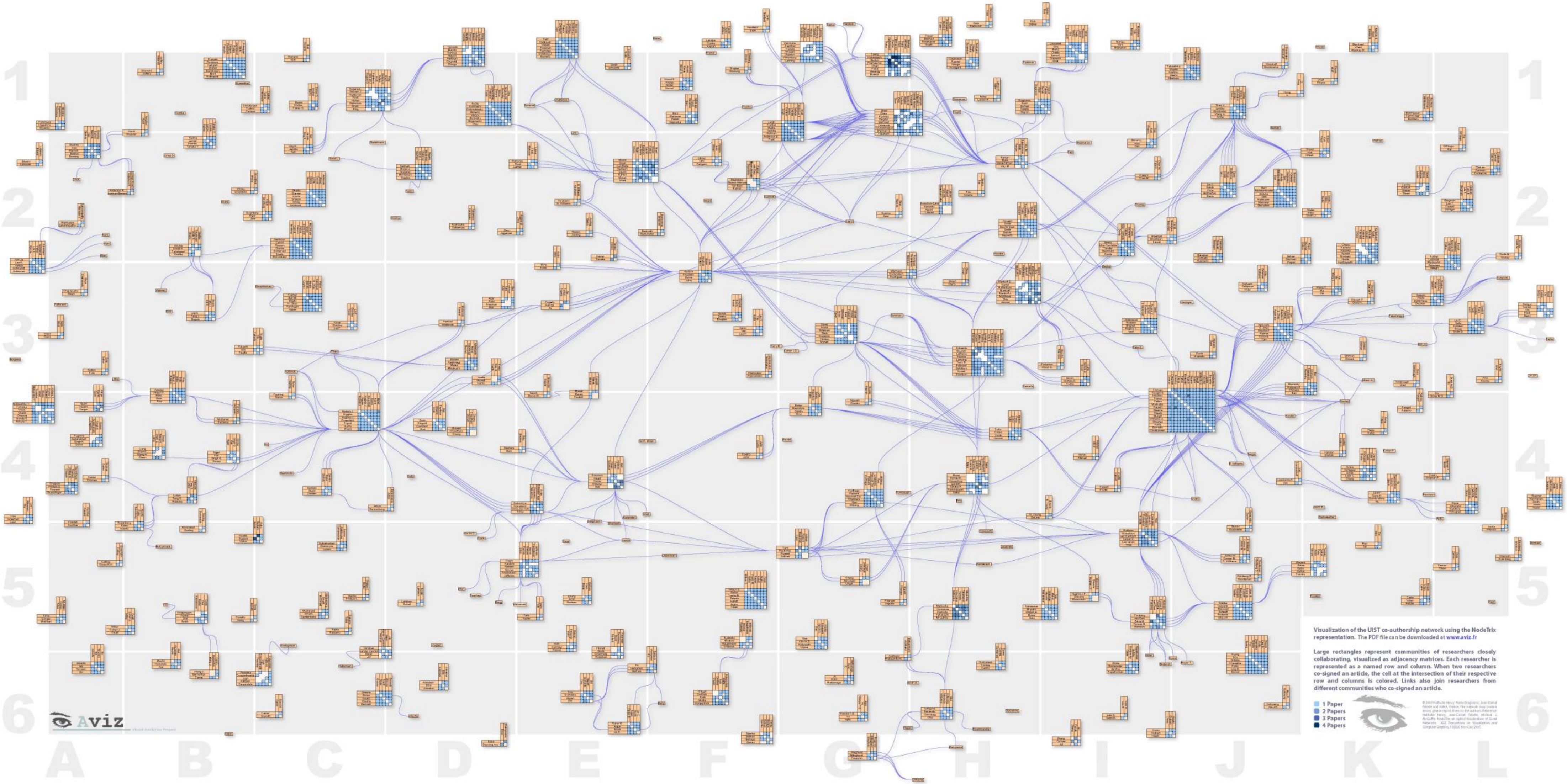

Video Expert Yann Riche
NodeTrix based on the Infovis Toolkit



Henry (Riche) et al., 2007

UIST Coauthorship Network

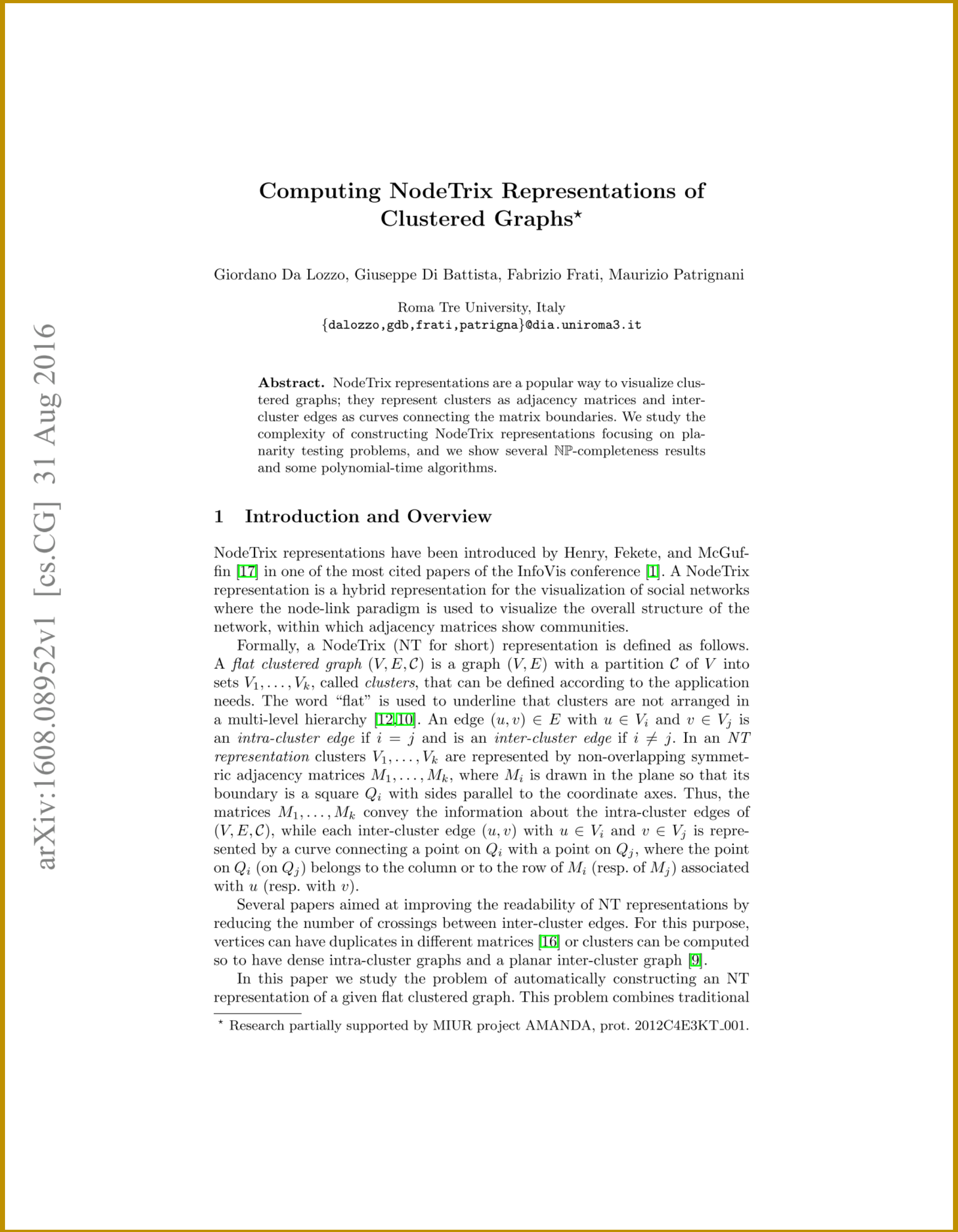
20 Years of Collaboration



Visualization of the UIST co-authorship network using the NodeTrix representation. The PDF file can be downloaded at www.aviz.fr

Large rectangles represent communities of researchers closely collaborating, visualized as adjacency matrices. Each researcher is represented as a named row and column. When two researchers co-signed an article, the cell at the intersection of their respective row and columns is colored. Links also join researchers from different communities who co-signed an article.

Discussion: NodeTrix Computation



Computing NodeTrix Representations of Clustered Graphs*

Giordano Da Lozzo, Giuseppe Di Battista, Fabrizio Frati, Maurizio Patrignani

Roma Tre University, Italy
{dalozzo,gdb,frati,patrigna}@dia.uniroma3.it

Abstract. NodeTrix representations are a popular way to visualize clustered graphs; they represent clusters as adjacency matrices and inter-cluster edges as curves connecting the matrix boundaries. We study the complexity of constructing NodeTrix representations focusing on planarity testing problems, and we show several $\mathbb{NP}$-completeness results and some polynomial-time algorithms.

1 Introduction and Overview

NodeTrix representations have been introduced by Henry, Fekete, and McGuffin [17] in one of the most cited papers of the InfoVis conference [1]. A NodeTrix representation is a hybrid representation for the visualization of social networks where the node-link paradigm is used to visualize the overall structure of the network, within which adjacency matrices show communities.

Formally, a NodeTrix (NT for short) representation is defined as follows. A *flat clustered graph* $(V, E, \mathcal{C})$ is a graph (V, E) with a partition $\mathcal{C}$ of V into sets $V_1, \dots, V_k$, called *clusters*, that can be defined according to the application needs. The word "flat" is used to underline that clusters are not arranged in a multi-level hierarchy [12,10]. An edge $(u, v) \in E$ with $u \in V_i$ and $v \in V_j$ is an *intra-cluster edge* if $i = j$ and is an *inter-cluster edge* if $i \neq j$. In an *NT representation* clusters $V_1, \dots, V_k$ are represented by non-overlapping symmetric adjacency matrices $M_1, \dots, M_k$, where M_i is drawn in the plane so that its boundary is a square Q_i with sides parallel to the coordinate axes. Thus, the matrices $M_1, \dots, M_k$ convey the information about the intra-cluster edges of $(V, E, \mathcal{C})$, while each inter-cluster edge (u, v) with $u \in V_i$ and $v \in V_j$ is represented by a curve connecting a point on Q_i with a point on Q_j, where the point on Q_i (on Q_j) belongs to the column or to the row of M_i (resp. of M_j) associated with u (resp. with v).

Several papers aimed at improving the readability of NT representations by reducing the number of crossings between inter-cluster edges. For this purpose, vertices can have duplicates in different matrices [16] or clusters can be computed so to have dense intra-cluster graphs and a planar inter-cluster graph [9].

In this paper we study the problem of automatically constructing an NT representation of a given flat clustered graph. This problem combines traditional

* Research partially supported by MIUR project AMANDA, prot. 2012C4E3KT_001.