Sequential Modeling and Structural Anomaly Analytics in Industrial Production Environments

Martin Atzmüller [1] and Andreas Schmidt [1] and David Arnu [2]

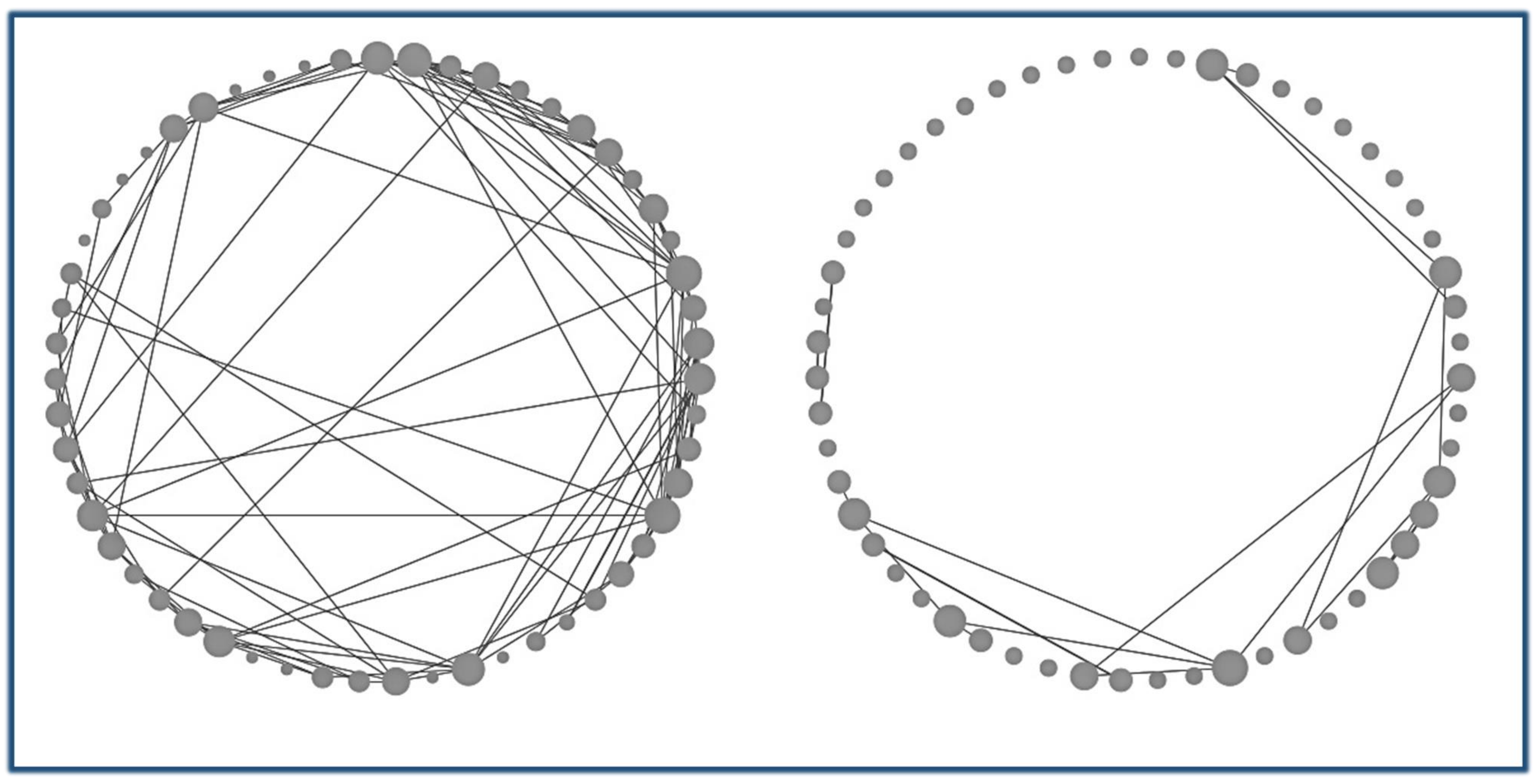
[1] University of Kassel, Research Center for Information System Design {atzmueller, schmidt}@cs.uni-Kassel.de
[2] RapidMiner GmbH, darnu@rapidminer.com

FEE
Early Detection and Decision Support for Critical Situations in Production Environments

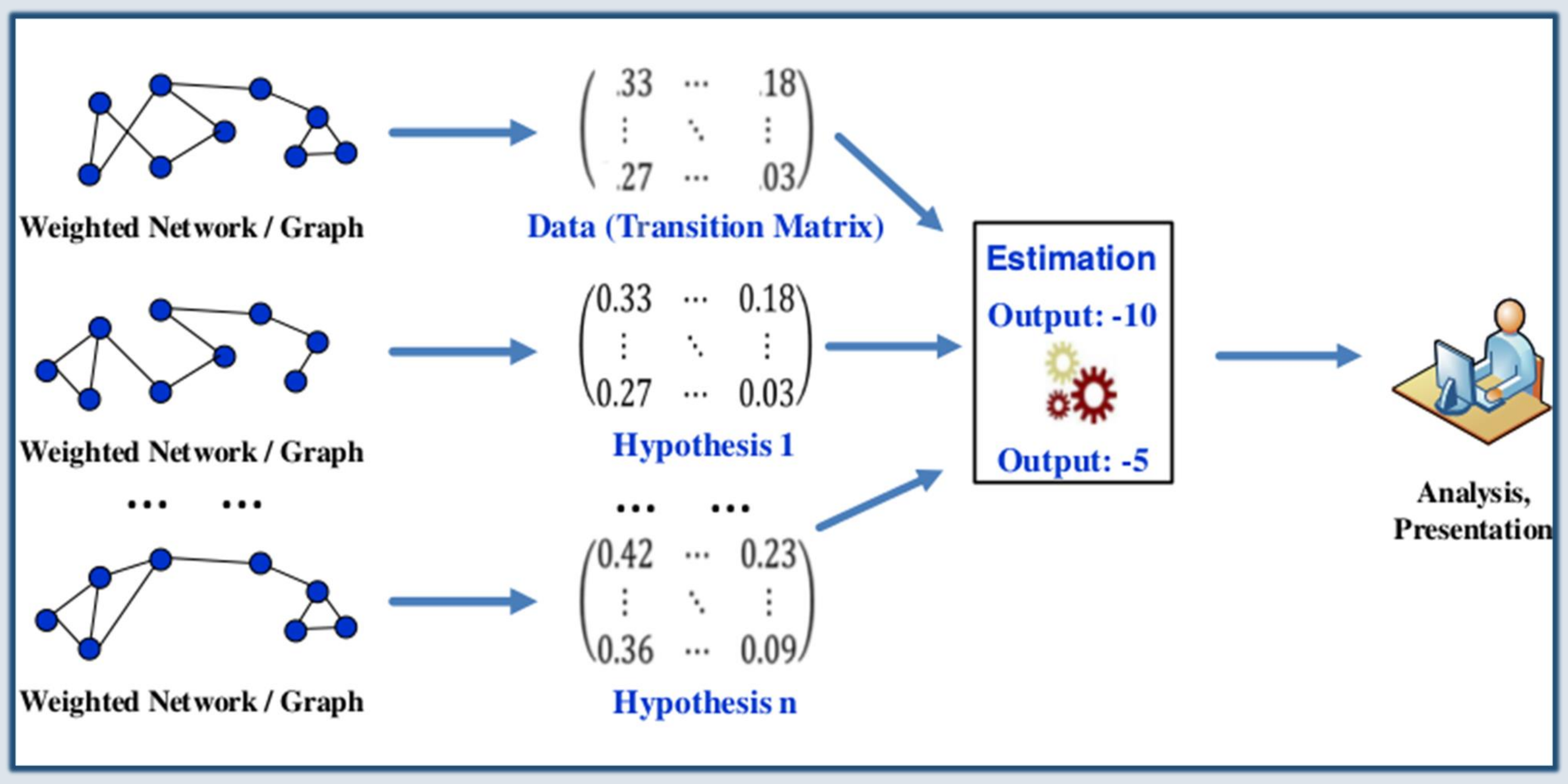
www.fee-projekt.de

Introduction

In many industrial areas, production facilities have reached a high level of automation. Here, knowledge about the respective processes is crucial, e. g., targeting the *topological structure* of a plant, sequences of operator notifications (alarms), and unexpected (critical) situations. The analysis of (exceptional) *sequential patterns* is an important task for obtaining insights into the process and for modeling predictive applications.
Our application context is given by (abstracted) *alarm sequences* in industrial production plants in an Industry 4.0 context. Specifically, we consider the analysis of the plant topology and anomaly detection in alarm logs. We formulate the "reference behavior" collecting normal episodes as sequences of normal situations.



Network visualization of exemplary transition matrices: Normal state (left) vs. Anomaly (right). The nodes denote different plant units (weighted by their share of aggregated outgoing alarm notifications) - and the edges in the example indicate sets of transitions that are normal or abnormal, respectively.



Overview of the HYPGRAPHS modeling and analysis process. Based on the weighted network graph a probabilistic matrix is build for each system state and apply HypTrails to calculate evidence scores. In summary, we utilize Bayesian interference on a first-order Markov-Chain model.

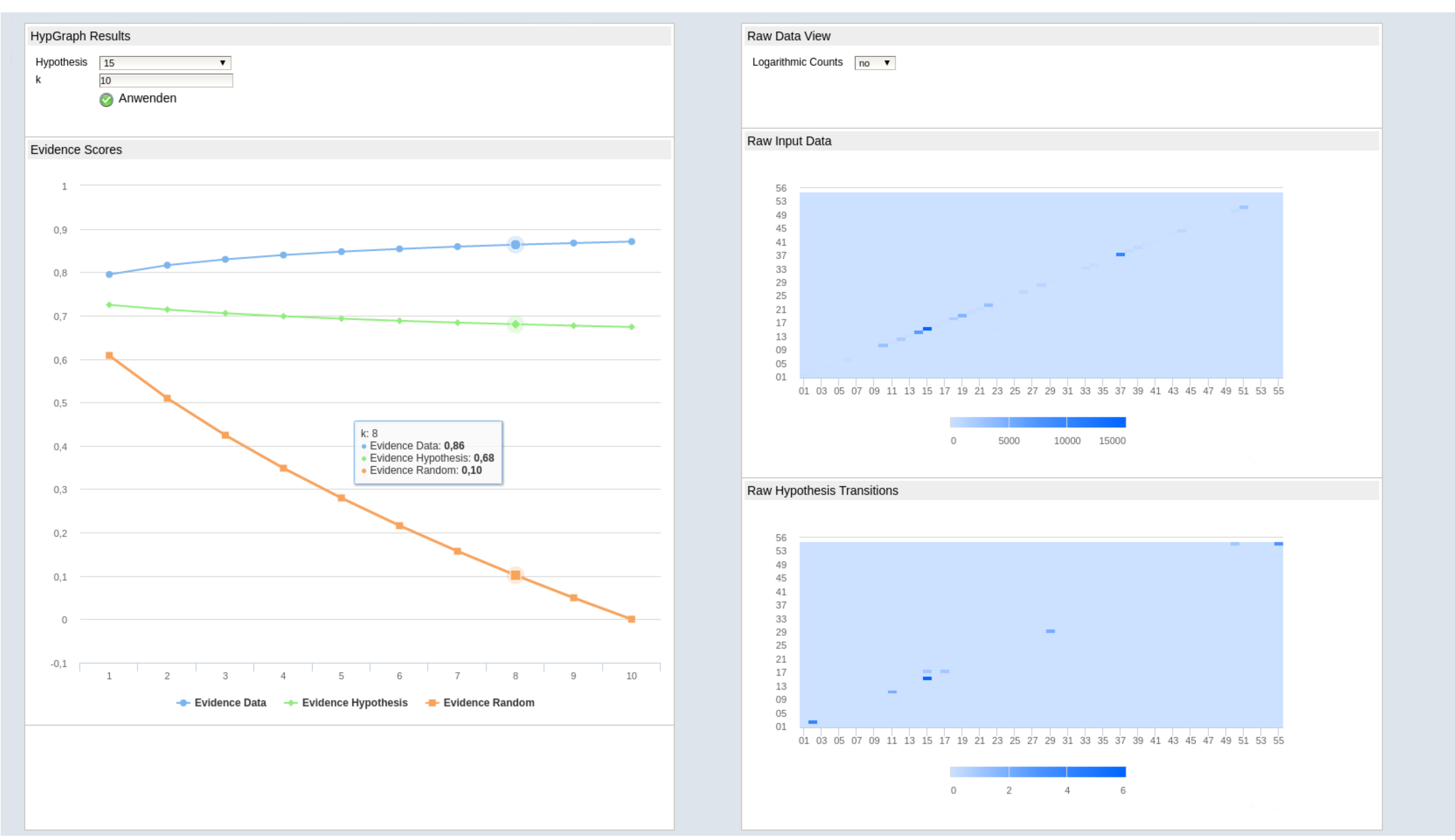
Method

Following the DASHTrails [1] and the HypGraphs [2] approaches, we model transition matrices given a probability distribution of certain states. We assume a discrete set of such states Ω corresponding to the nodes of a network (w.l.o.g. $\Omega = \{1, \dots, n\}, n \in \mathbb{N}, |\Omega| = n$).
The three steps of the process are:

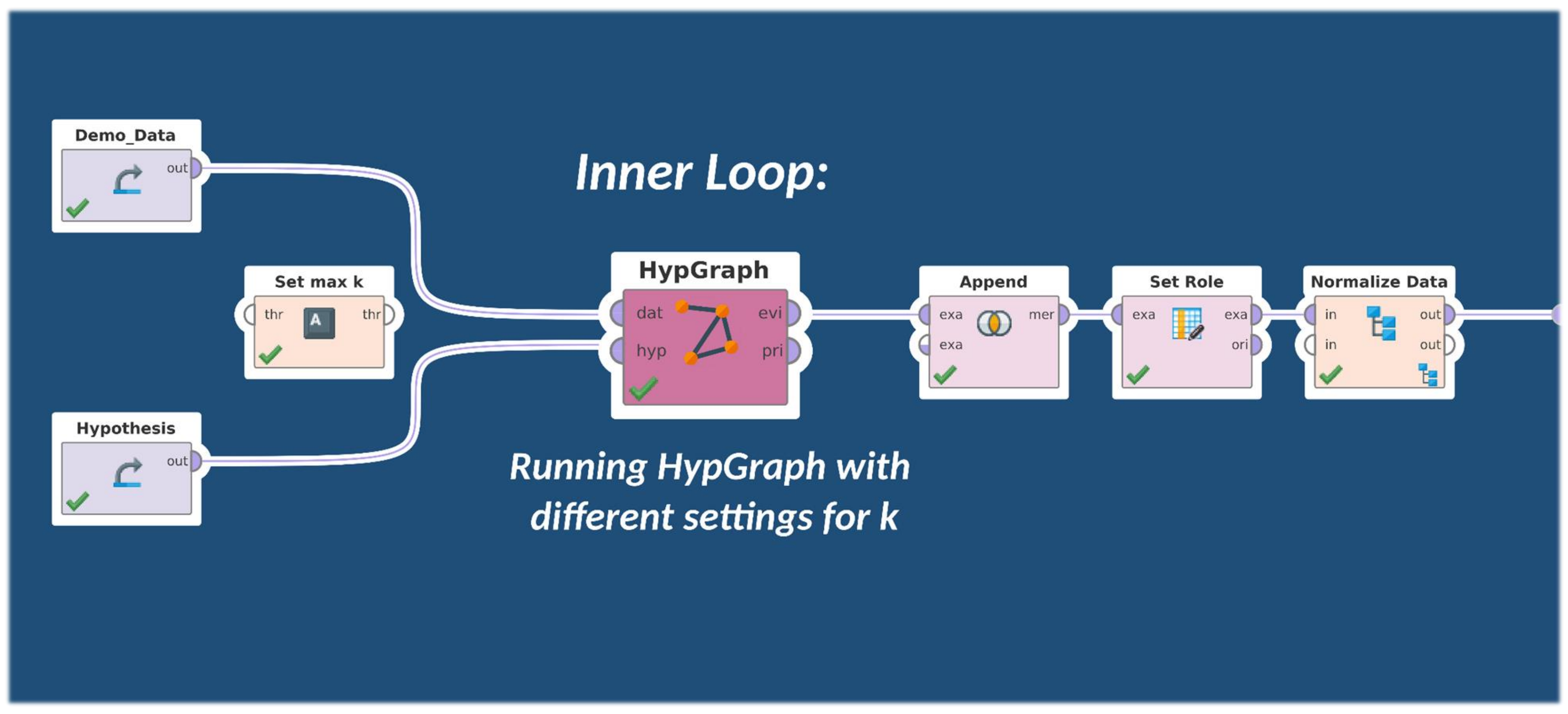
1. **Modeling:** Determine a transition model given the weighted network using a *transition modeling function* $\tau: \Omega \times \Omega \rightarrow \mathbb{R}$. Transitions between sequential states $i, j \in \Omega$ are captured by a transition matrix M, i.e., $m_{i,j} = \tau(i,j)$
2. **Estimation:** Apply HypTrails on the given transition matrix and the respective hypotheses, and return the resulting evidence
3. **Analysis:** Present the results for semi-automatic introspection and analysis, e.g., by visualizing the network as a Heatmap

Process Model & Implementation

Starting with process data from a streaming data source (or historic data, e.g., as flat files). For different believe weights *k* the evidence scores are calculated for the presented hypothesis. As upper and lower bounds the evidence scores for the data are compared against itself. These are calculated for purely random transitions as well. The evidence scores presented in an interactive dashboard or can be embedded in an Operator Support System [3].



RapidMiner Server Dashboard showing evidence values for the hypothesis compared to a randomized hypothesis



RapidMiner process performing the HypGraph calculation for different believe weights *k*

Literature

1. Atzmüller, M., Schmidt, A., Kibanov, M.: DASHTrails: An Approach for Modeling and Analysis of Distribution-Adapted Sequential Hypotheses and Trails. In: Proc. WWW 2016 (Companion). IW3C2 / ACM, New York, NY, USA (2016)
2. Atzmüller, M., Schmidt, A., Klöpper, B., Arnu, D.: HypGraphs: An Approach for Modeling and Comparing Graph-Based and Sequential Hypotheses. In: Proc. ECML-PKDD Workshop on New Frontiers in Mining Complex Patterns (NFMCP). Riva del Garda, Italy (2016)
3. Klöpper, B., Dix, M., Schorer, L., Ampofo, A., Atzmüller, M., Arnu, D., Klinkenberg, R.: Defining Software Architectures for Big Data Enabled Operator Support Systems. In: Proc. INDIN. IEEE Press, Boston, MA, USA (2016)

Duration: 09/2014 – 08/2017

Project partners:

ABB AG
BASF SE
INEOS Köln GmbH
PCK Raffinerie GmbH
RapidMiner GmbH
Technische Universität Dresden
Universität Kassel

Contact:

ABB AG (Consortial Leader)
Dr. Benjamin Klöpper
University Kassel (Scientific leader)
Univ.-Prof. Dr.-Ing. Andreas Kroll
Univ.-Prof. Dr. Gerd Stumme
http://www.fee-projekt.de/

SPONSORED BY THE
Federal Ministry of Education and Research

FKZ: 01|S14006A-E