

Pattern Mining in Dynamic Graphs

Karel Vaculík

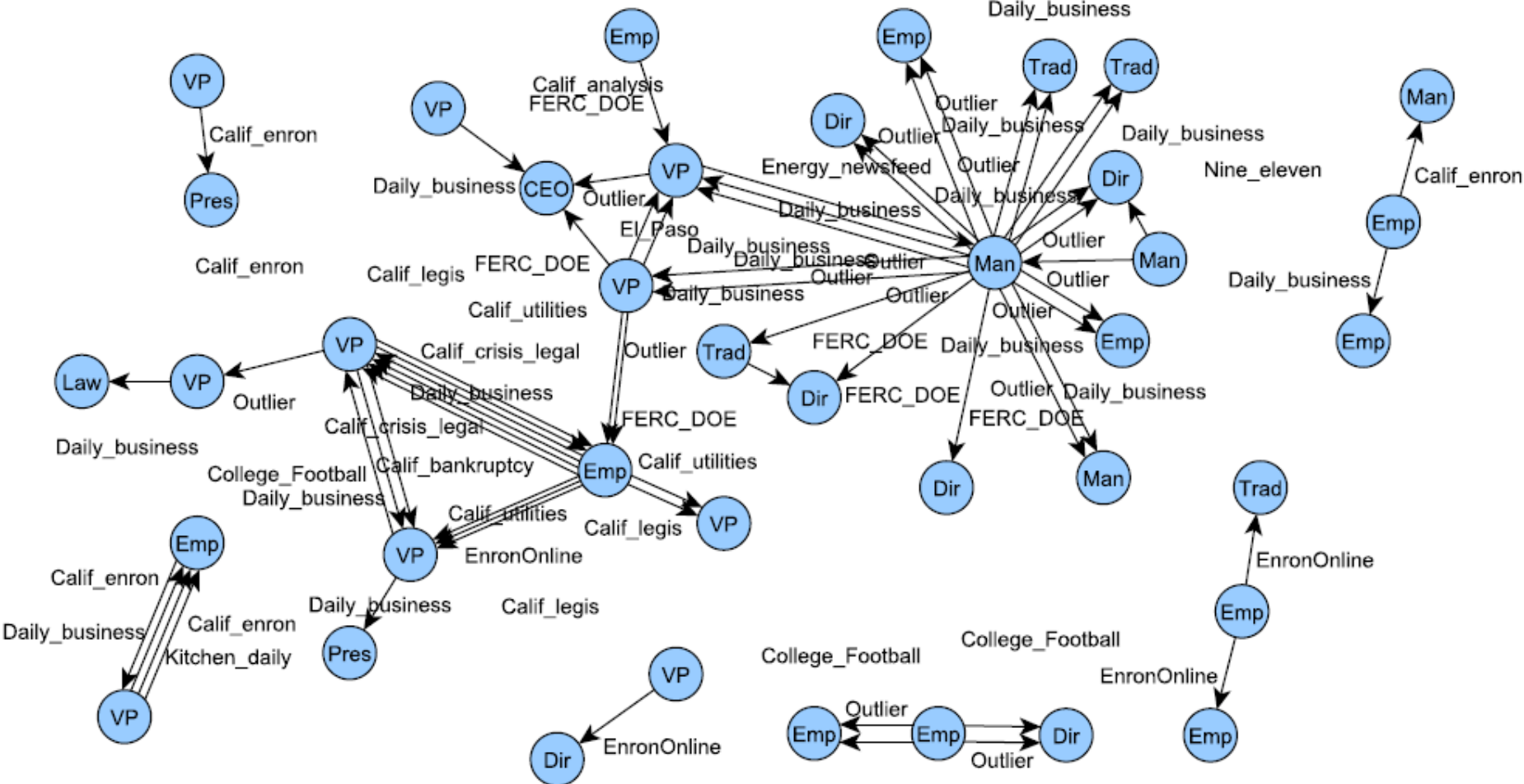
Outline

- Introduction
- Frequent patterns
- Anomalous patterns
- Future directions

Outline

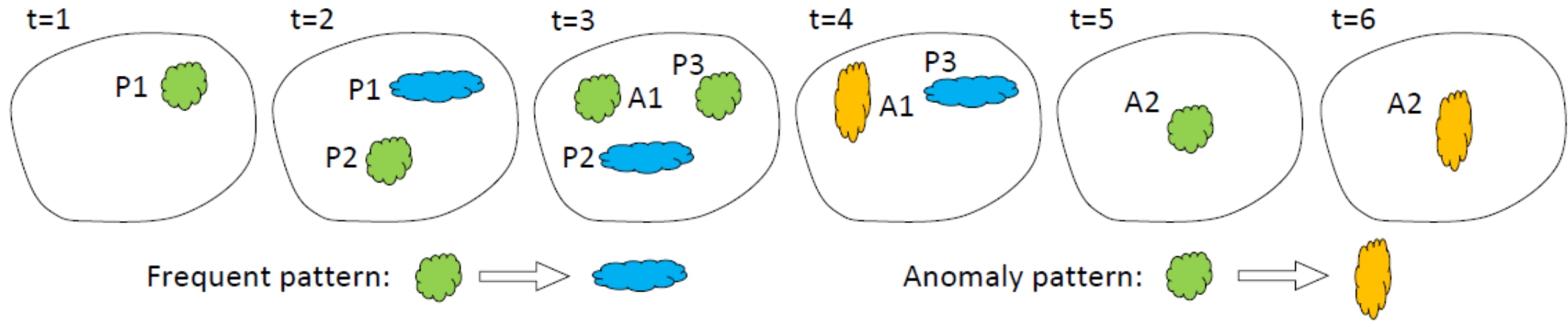
- Introduction
- Frequent patterns
- Anomalous patterns
- Future directions

Introduction



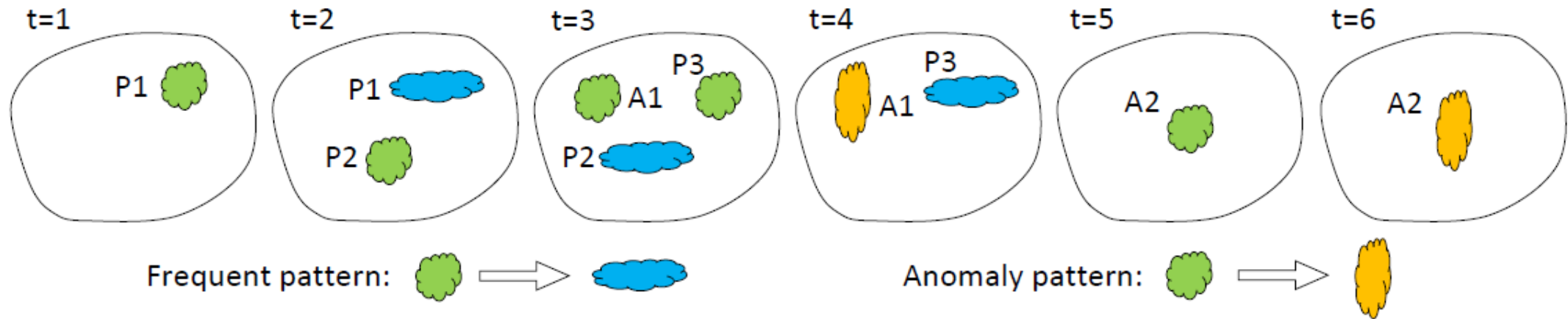
Introduction

Dynamic graph:

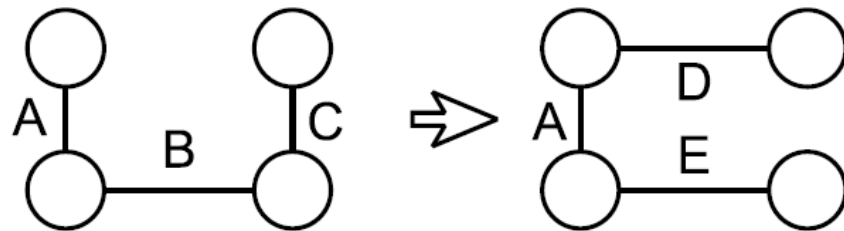


Introduction

Dynamic graph:



An example of a rule:



Applications

- Citation / bibliographic networks
 - Collaboration patterns
 - Evolution of publication behaviour
- Social networks
 - Leaving community / change of attributes after specific interactions (behaviour patterns)
- Communication networks
 - Utilization of common communication patterns for productivity increase

Outline

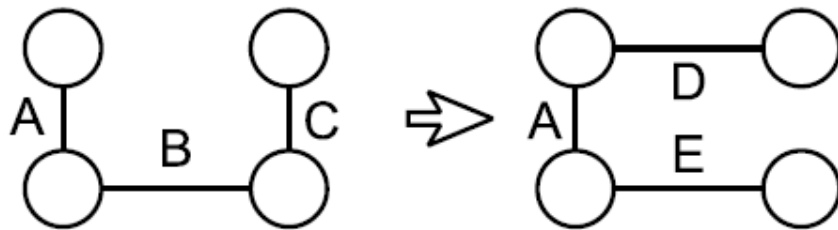
- Introduction
- Frequent patterns
- Anomalous patterns
- Future directions

DGRMiner

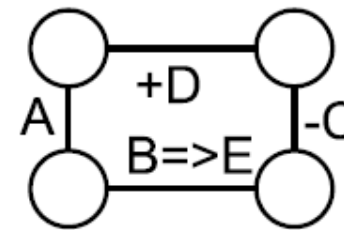
- Frequent pattern mining
- Anomaly detection and explanation
- Considers various types of changes:
 - Vertex addition / deletion
 - Edge addition / deletion
 - Change of vertex / edge labels
- Undirected as well as directed edges; multiedges
- Single dynamic graph or multiple dynamic graphs on input

DGRMiner: Union Graph Representation

a predictive graph rule:



union graph of the rule:



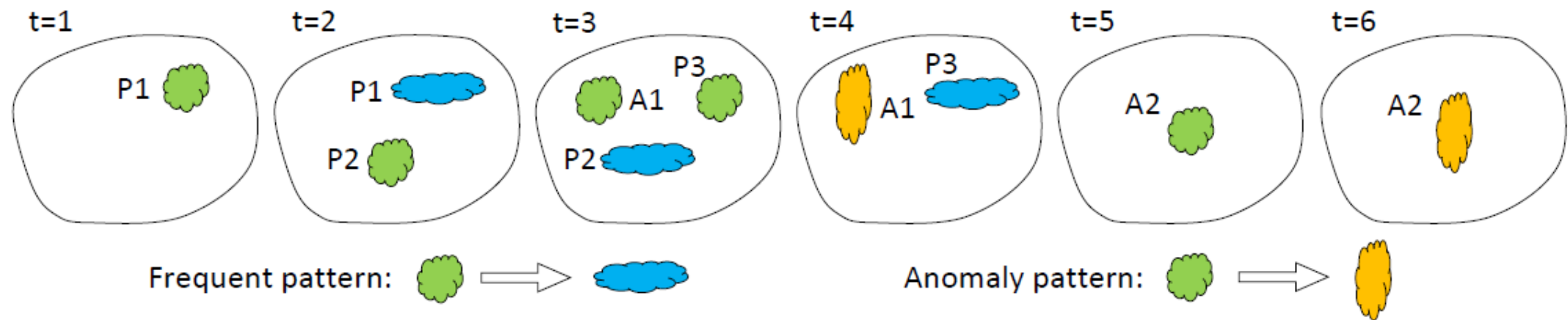
+X ... addition

-X ... deletion

Y=>X ... Change from Y to X

DGRMiner: Support and Confidence

Dynamic graph:

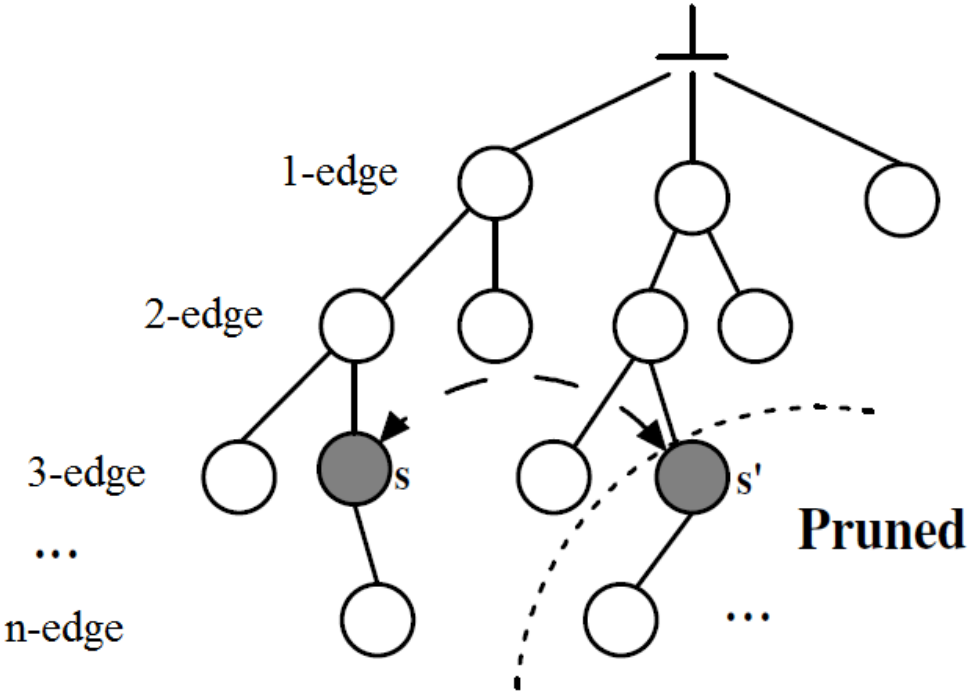
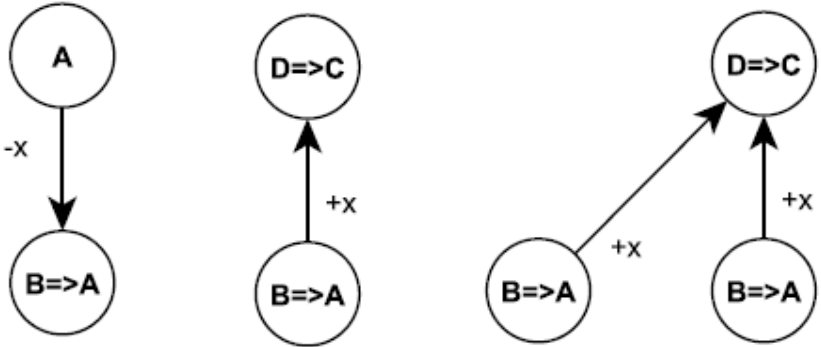
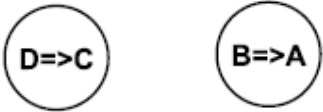


- Support of the freq. pattern: $3/5$ (or 3 as absolute)
- Confidence of the freq. pattern: $3/4$

DGRMiner: Frequent Pattern Mining

Based on gSpan algorithm:

- 1. Mine frequent change vertices
- 2. Mine frequent patterns built from change edges
 - in a depth-first-search manner
 - avoids duplicate patterns



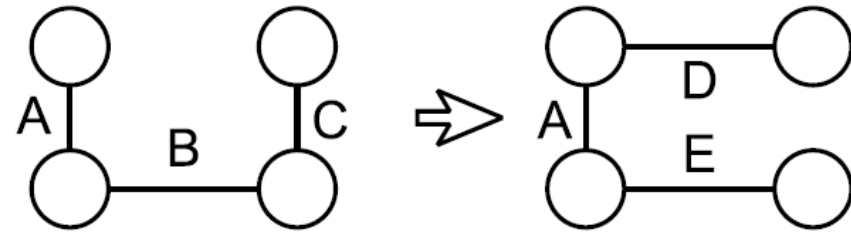
*gSpan (Yan & Han @ ICDM'02);
modified image of the tree from the same paper

Outline

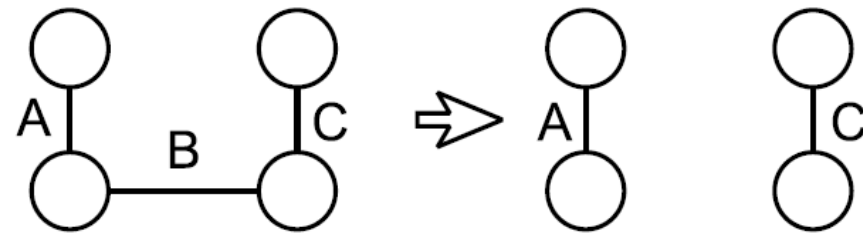
- Introduction
- Frequent patterns
- Anomalous patterns
- Future directions

DGRMiner: Anomalies

- Frequent pattern



- Anomalous pattern = deviation from the frequent pattern



Outlierness = 1 - confidence

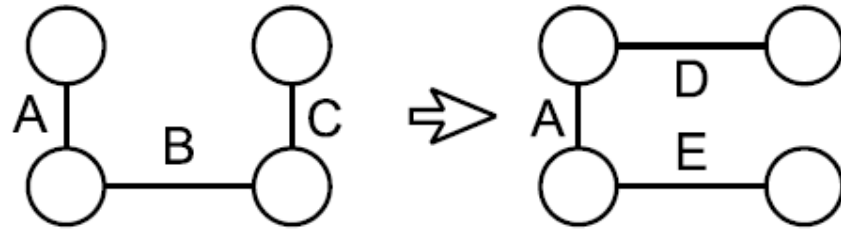
Single-vertex Anomalies

Frequent pattern	Possible anomalies
$\neg A$	$A, A \Rightarrow C$ (where $C \neq A$)
$A \Rightarrow B$	$A, \neg A, A \Rightarrow C$ (where $C \neq B$)
$+B$	$\neg B$

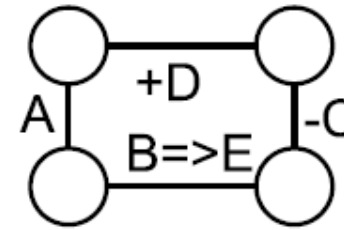
- How to compute support of these anomalies?

Non-trivial Anomalies

a predictive graph rule:



union graph of the rule:

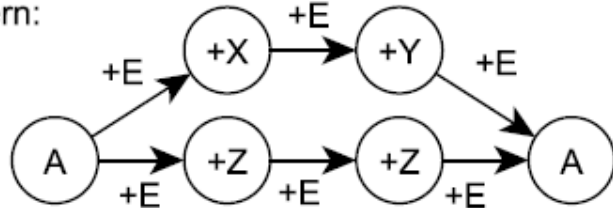


- Frequent patterns without “additions”: simple enumeration of antecedents

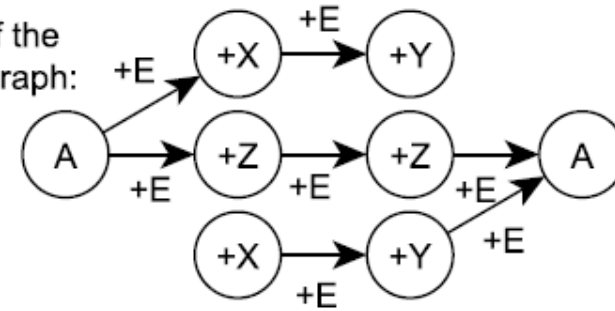
Non-trivial Anomalies

- Frequent patterns with “additions”:

Frequent pattern:



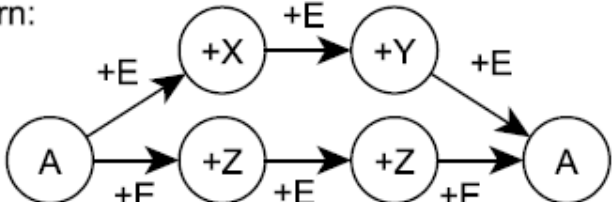
Union graph of the
input dynamic graph:



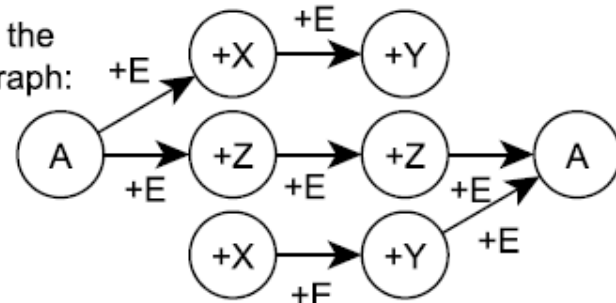
Non-trivial Anomalies

- Frequent patterns with “additions”:

Frequent pattern:

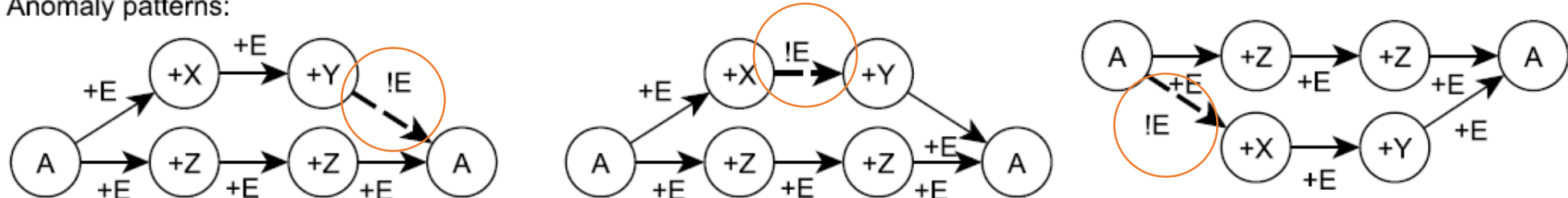


Union graph of the input dynamic graph:



- Solution: Maximal common subgraphs of the freq. pattern and the input union graph

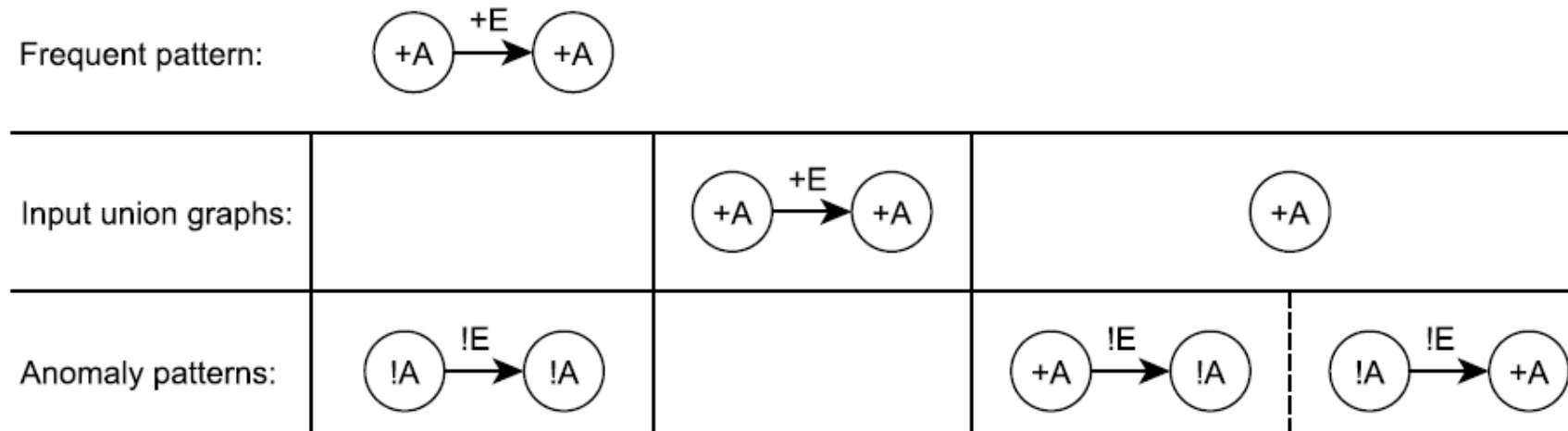
Anomaly patterns:



Non-trivial Anomalies

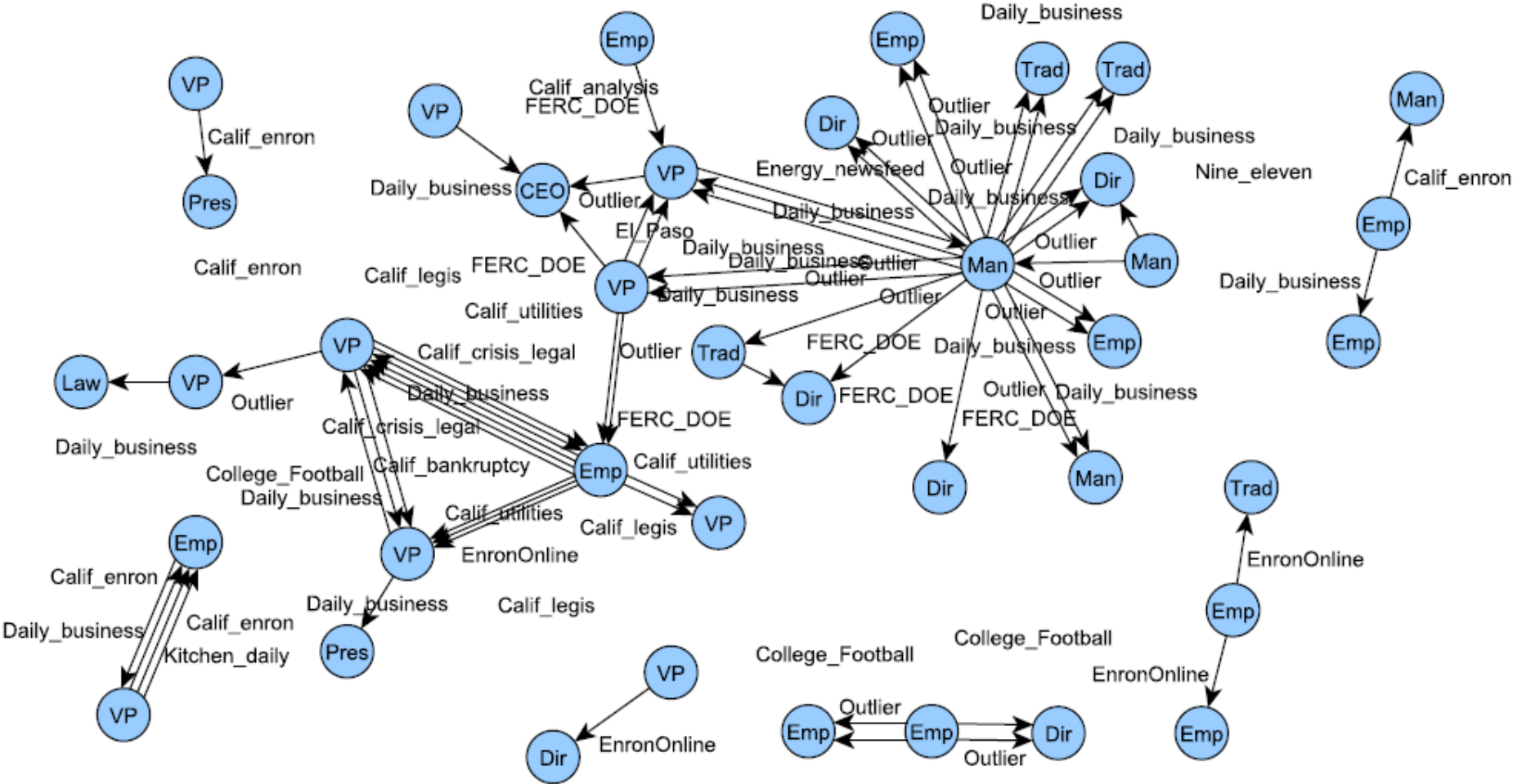
How to compute support of non-trivial anomalies?

- Simple if the corresponding frequent pattern contains “non-additions”
- Otherwise:



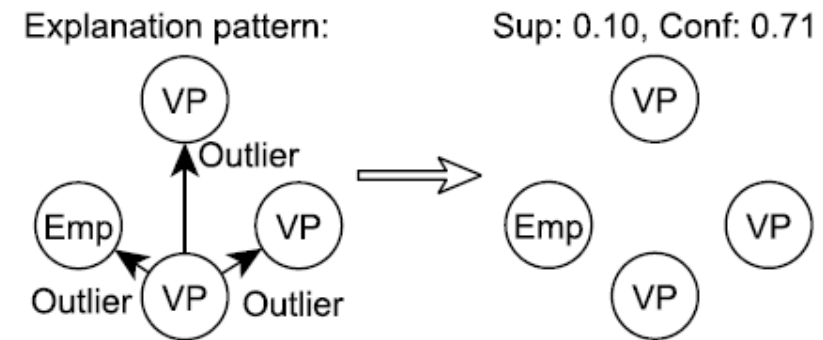
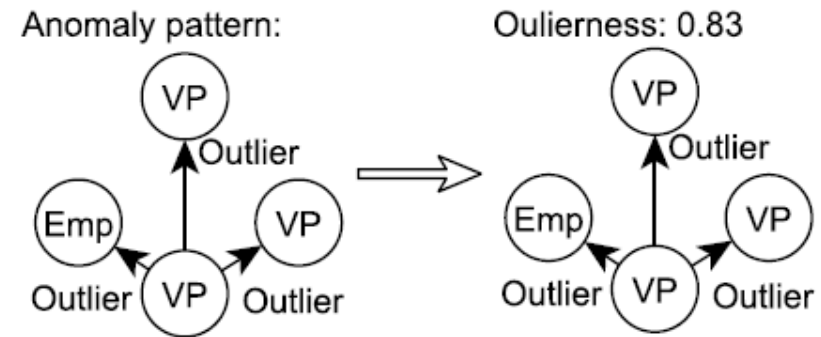
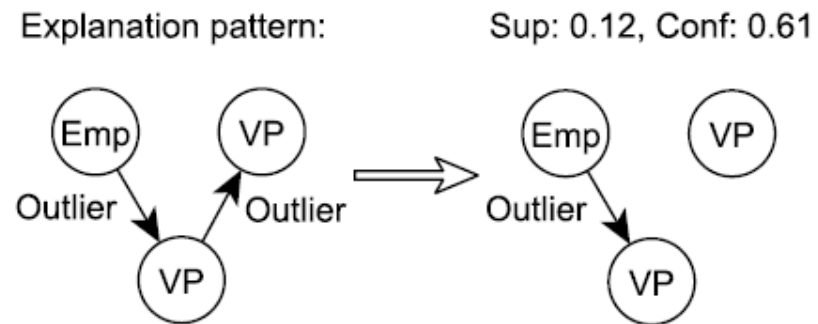
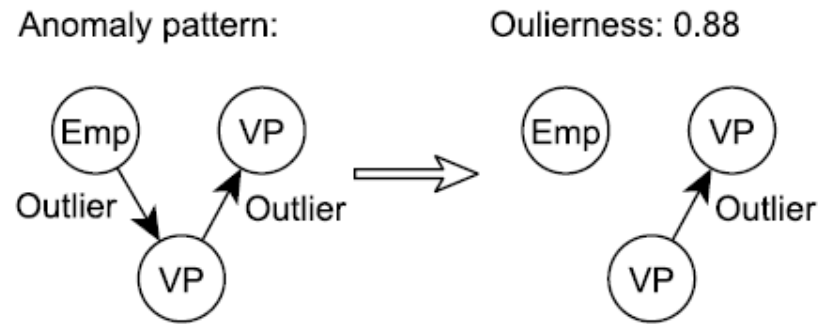
Experiments

Enron



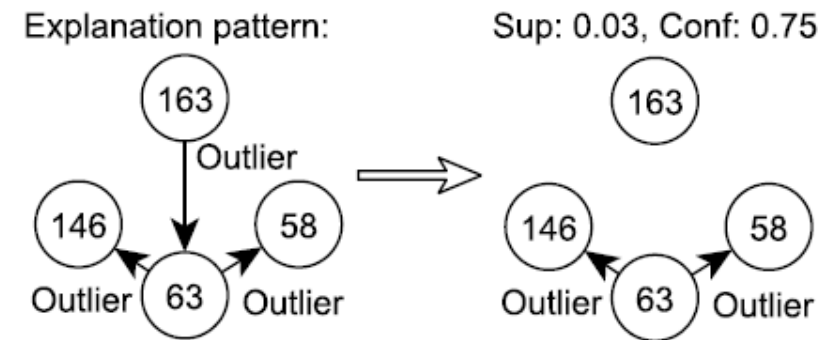
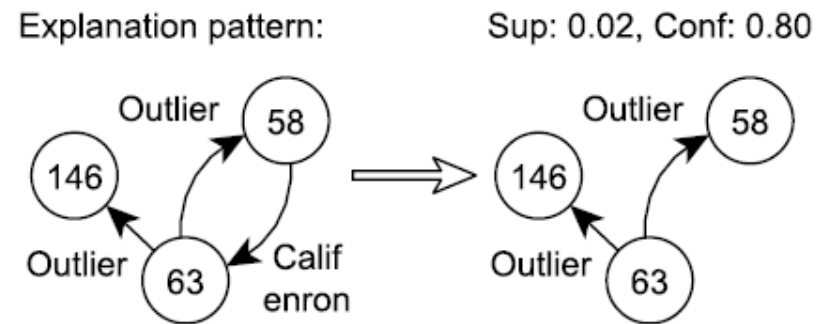
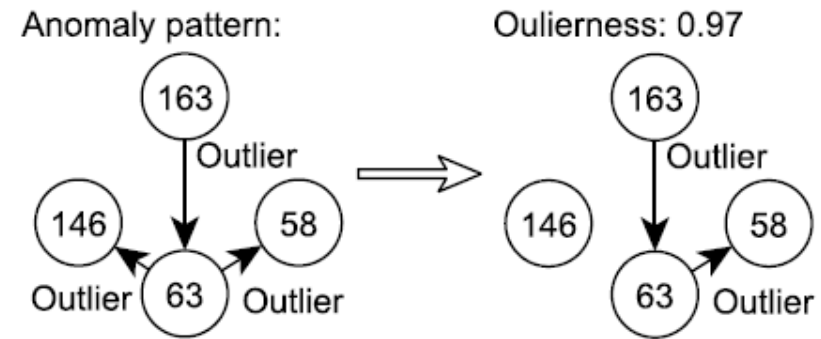
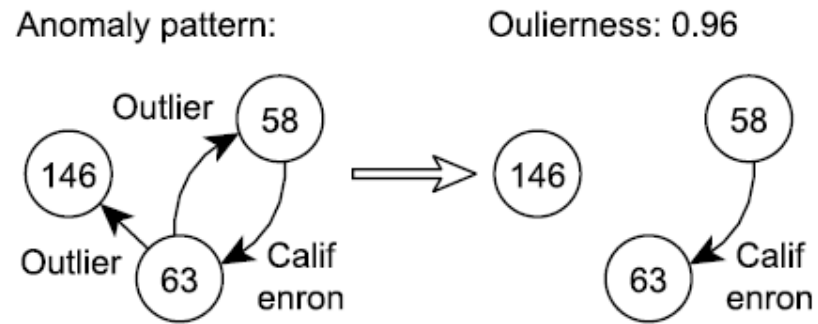
Experiments

Enron



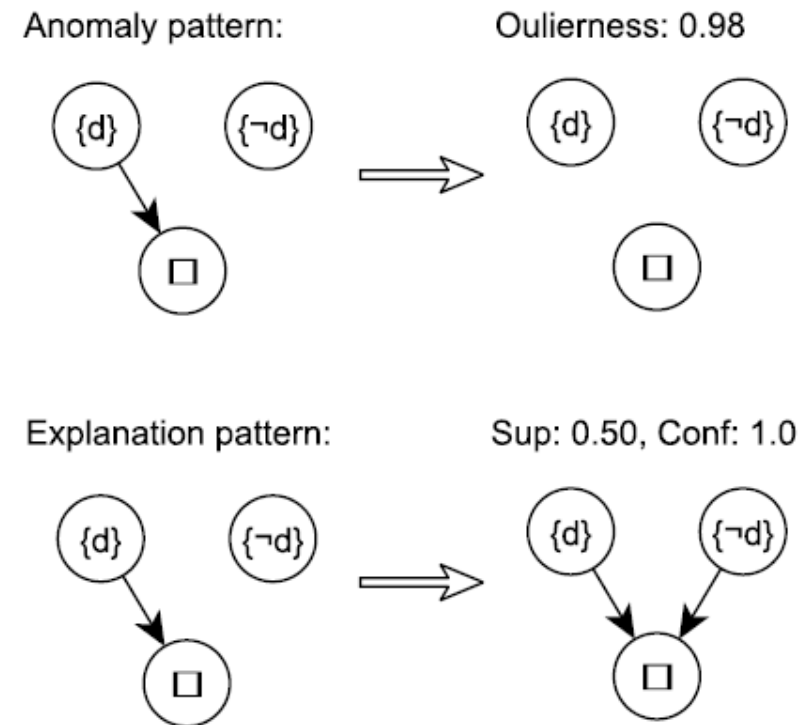
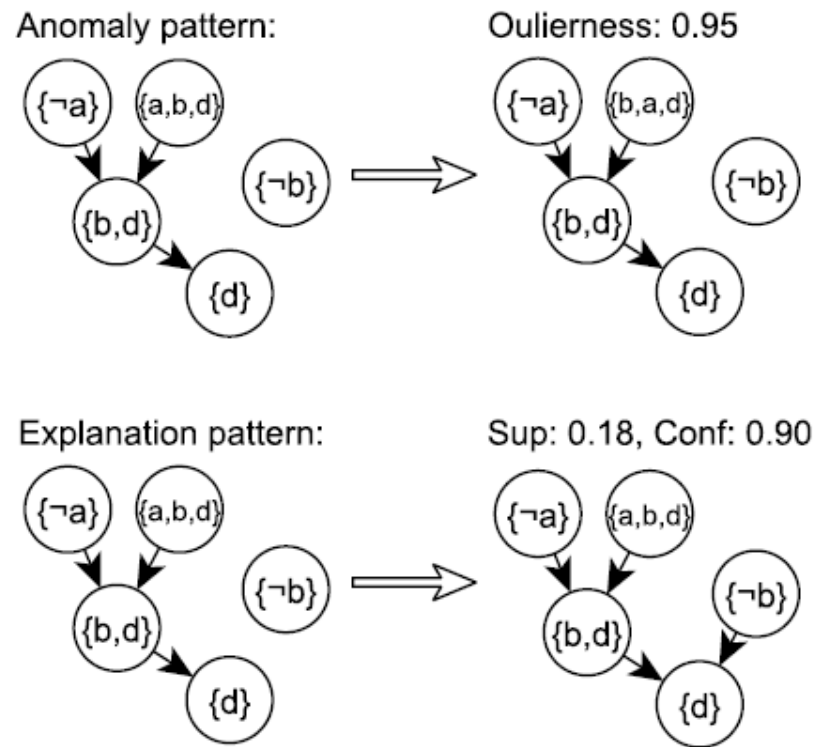
Experiments

Enron (unique vertex labels)



Experiments

Resolution proofs

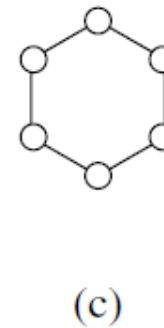
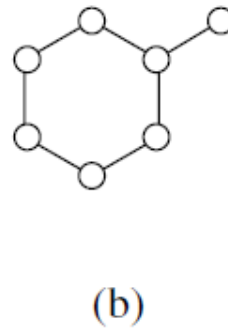
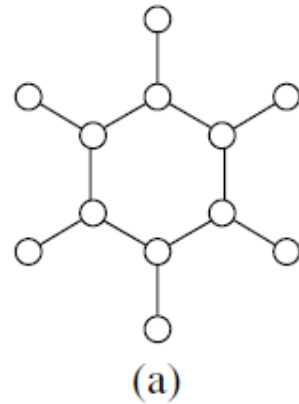


Outline

- Introduction
- Frequent patterns
- Anomalous patterns
- Future directions

Support Definition

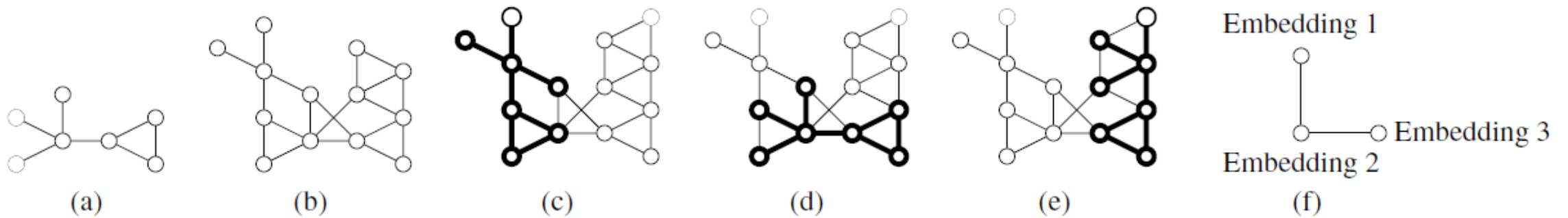
- DGRMiner at this moment: count at most one occurrence in each union graph
- Counting as many occurrences as possible:



Anti-monotonicity is broken!

Support Definition

=> solve the Maximum Independent Set problem on the graph of embeddings



Patterns with deviating “additions”