UNIVERSITY OF LIVERPOOL

Minimal Vertex Unique Labelled Subgraph Mining

Wen Yu , Frans Coenen , Michele Zito and Subhieh El Salhi

Department of Computer Science,
The University of Liverpool

Overview

- Vertex Unique Labelled Subgraph Mining (VULSM), a novel form of subgraph mining.

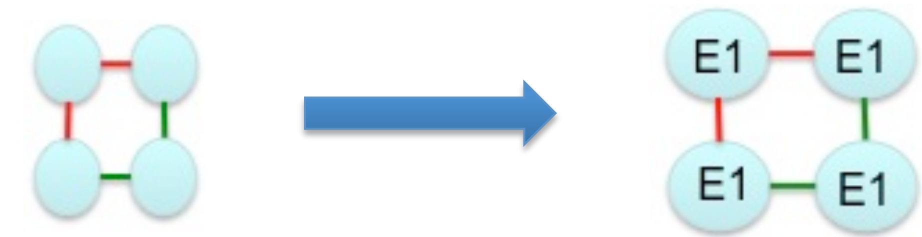
- Motivation for VULS (Sheet Metal Forming).

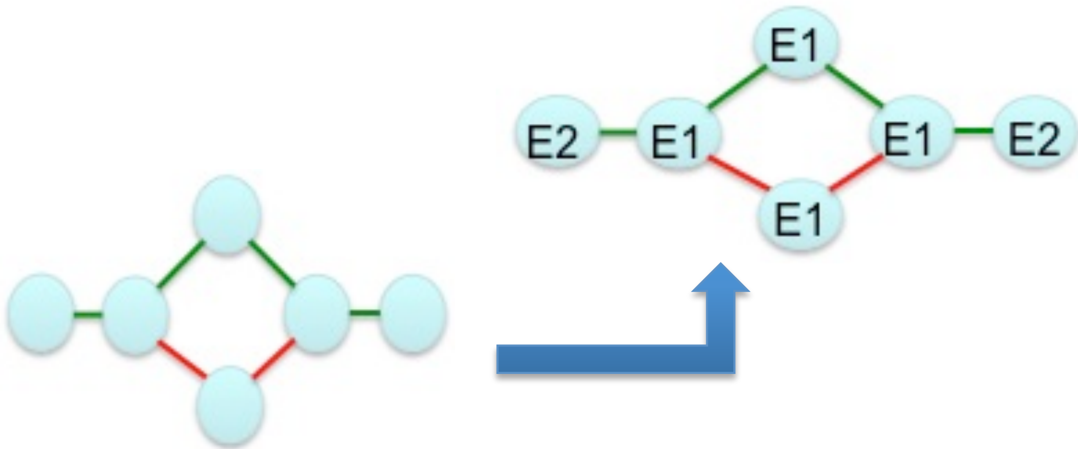
- Min-BFS-REUSMA(Minimal VULS, Breath First-Search, Right-most Extension Unique Subgraph Mining).

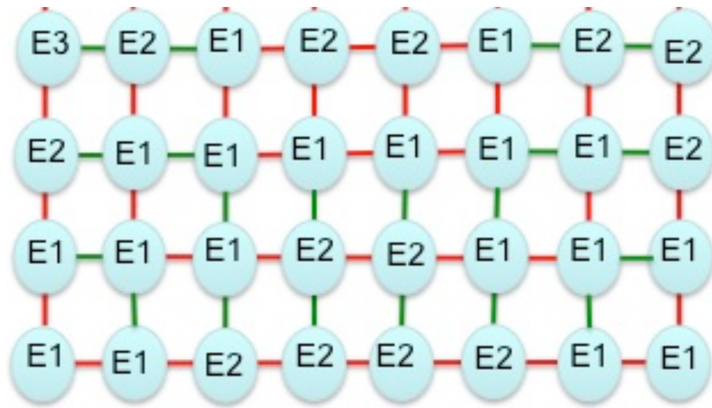
- Evaluation using real data.

Definition of VULS

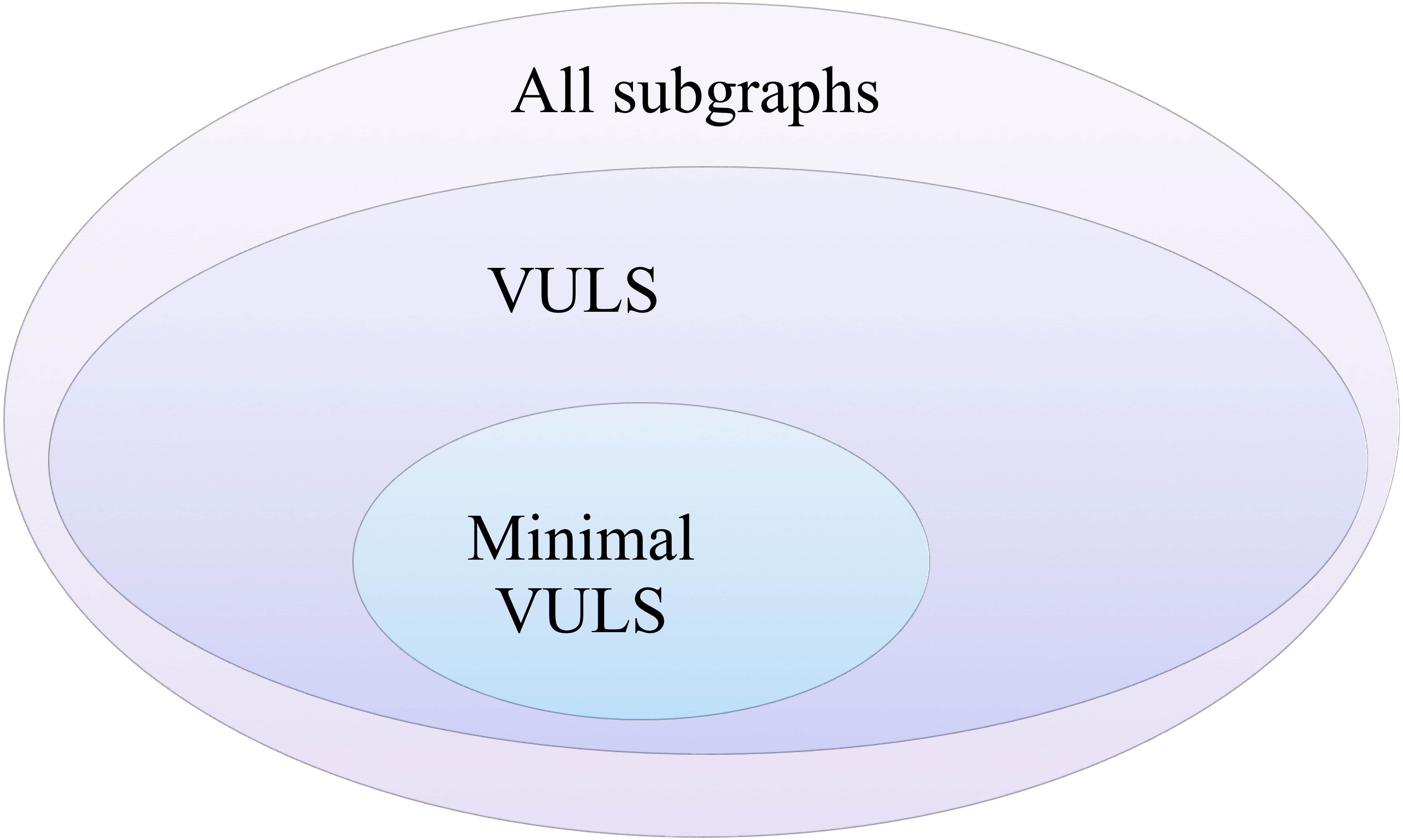
- A VULS is a subgraph with a particular structure and a specific edge labelling that has a unique vertex labelling associated with it.
- A minimal VULS is a VULS which is not a super-graph of any other VULS.







Two Types of VULS



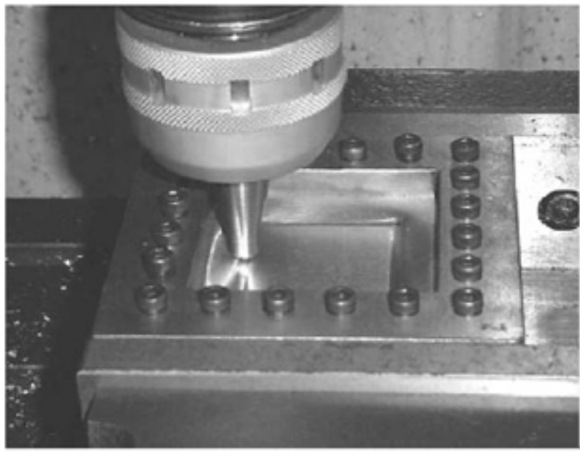
Motivation

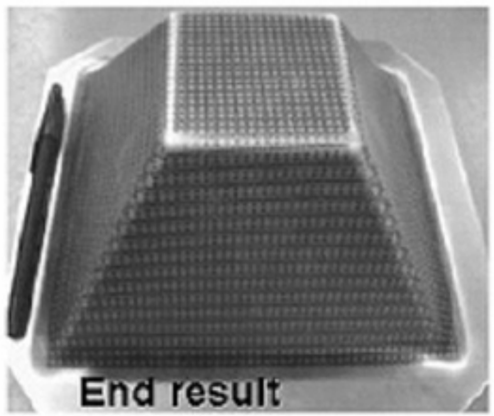
One application of VULSM is in the field of sheet metal forming where the concept can be used for error prediction.

distortions — relate to — Local geometries

VULS — Predict — Errors associated with new parts to be manufactured

Mitigating error correction





Min-BFS-REUSMA Algorithm

Input single train graph G

Parameter Max (maximum VULS size)

K=1

G_k=the set of K-edge subgraphs in G

R=Φ

While(K<Max)**do**

R=R U GenerateMinVULS(K, G_k)

G_{k+1}=Set of (K+1)-edge subgraphs generated from T_k in G (similar to gSpan)

K=K+1

End While

Minimal VULS set R

Experimental results

graph set	# vertices	# edge labels	# vertex labels
AISF1	36	3	2
AISF2	36	2	2
AISF3	36	2	3
AISF4	100	3	2
AISF5	100	2	2
AISF6	100	2	3
AISF7	441	3	2
AISF8	441	2	2
AISF9	441	2	3
AISF10	441	4	2

Summary of AISF evaluation graph sets

Coverage

$$coverage = \frac{num.\ vertices\ covered\ by\ VULS}{num.\ vertices\ in\ G_{input}} \times 100$$

We need a set of VULS that covers the test data so that the matches could be found, thus the vertex label in testing graph could be predicted more accurately. We need a set of VULS that covers the test data if we don’t , will get a poor classification result.

Results 1(Coverage and # Min VULS)

graph set	max value					
	3	4	5	6	7	8
AISF1	52.8	100	100	100	100	100
AISF2	86.1	91.7	97.2	100	100	100
AISF3	13.9	47.2	100	100	100	100
AISF4	59	87	98	99	99	99
AISF5	80	81	81	100	100	100
AISF6	19	40	67	85	99	99
AISF7	64.2	92.1	93	99.6	99.8	99.8
AISF8	24.9	70.5	71.7	71.7	71.7	OME
AISF9	7.0	27.7	58.7	69.8	76.0	OME
AISF10	76.4	86.9	87.3	87.3	87.3	88.2

Coverage (%) for a range of max values

graph set	max value					
	3	4	5	6	7	8
AISF1	9	24	24	24	24	24
AISF2	7	13	29	37	37	37
AISF3	1	11	69	69	69	69
AISF4	5	16	48	75	93	104
AISF5	13	32	51	79	79	79
AISF6	6	22	94	334	356	395
AISF7	11	27	100	226	318	369
AISF8	4	22	103	267	717	OME
AISF9	4	26	125	272	401	OME
AISF10	26	70	96	136	193	279

Total number of minimal VULS discovered for a range of max values

Results 2 (Run Time)

graph set	max value					
	3	4	5	6	7	8
AISF1	0.07	0.12	0.13	0.13	0.13	0.13
AISF2	0.08	0.1	0.16	0.17	0.16	0.21
AISF3	0.09	0.26	0.34	0.37	0.39	0.41
AISF4	0.16	0.19	0.27	0.29	0.32	0.47
AISF5	0.14	0.23	0.23	0.28	0.35	0.35
AISF6	0.17	0.29	0.51	0.68	0.77	0.97
AISF7	0.33	0.42	0.61	0.65	0.73	0.87
AISF8	0.31	0.41	0.7	0.81	1.69	OME
AISF9	0.33	0.56	0.92	1.34	2.16	OME
AISF10	0.32	0.45	0.6	0.74	0.88	1.43

Run time (seconds) comparison for a range of max values

Conclusions and further study

- The concept of VULSM has been introduced together with the Min-BFS-REUSMA.

- Has application with respect to error prediction in sheet metal forming.

- Good results were produced (high coverage).

- Further work to be conducted with respect to Frequent VULSM.

End

Thank you for your attention!

Thank you for questions and suggestions