UNIVERSITY OF LIVERPOOL

Vertex Unique Labelled Subgraph Mining for Vertex Label Classification

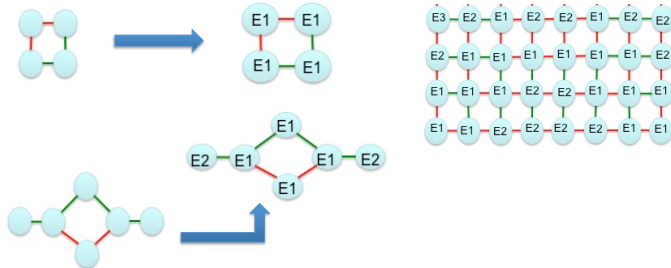*Wen Yu , Frans Coenen , Michele Zito and Subhieh El Salhi*

Department of Computer Science,
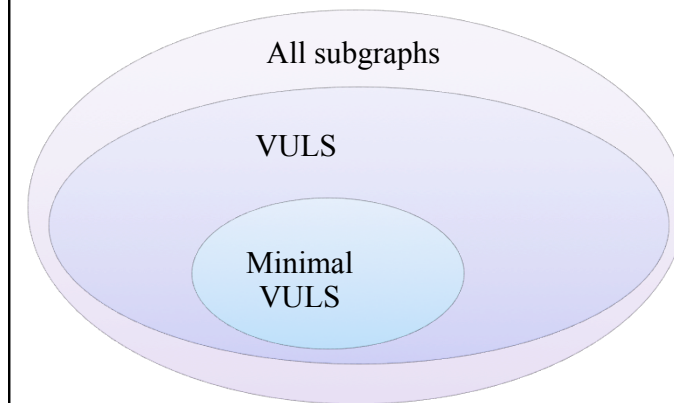
The University of Liverpool

Overview

■A novel mechanism for vertex label classification based on Vertex Unique Labelled Subgraph Mining (VULSM) is proposed.

■Training process: Minimal Right-most Extension VULS Mining (minREVULSM) algorithm.

■Classification process: the Match-Voting algorithm.

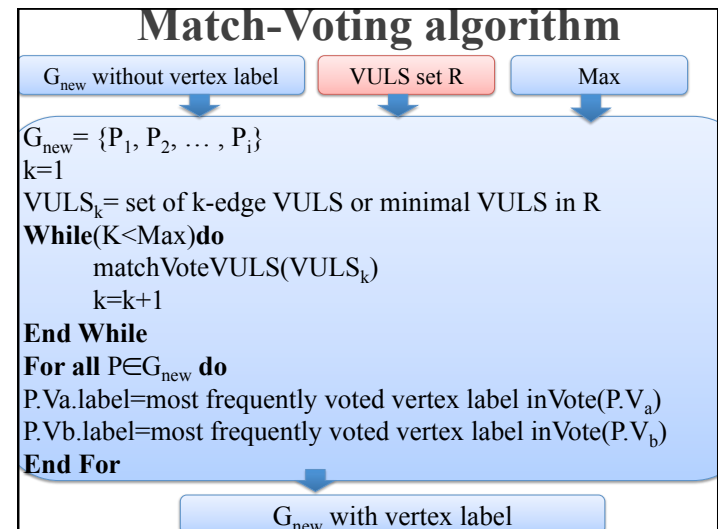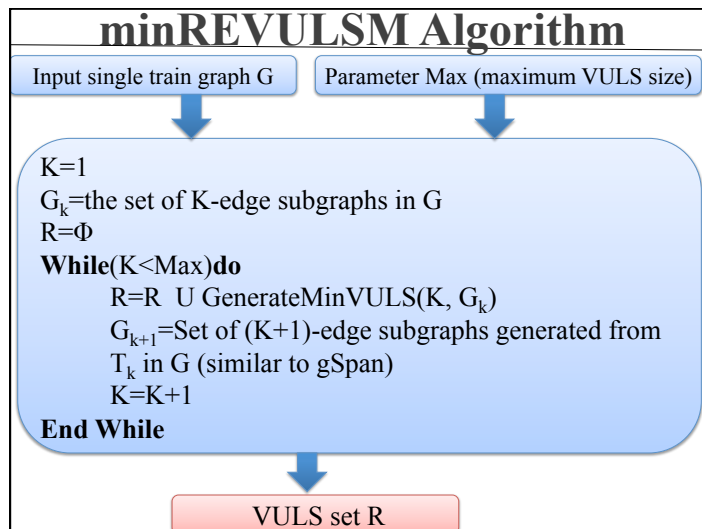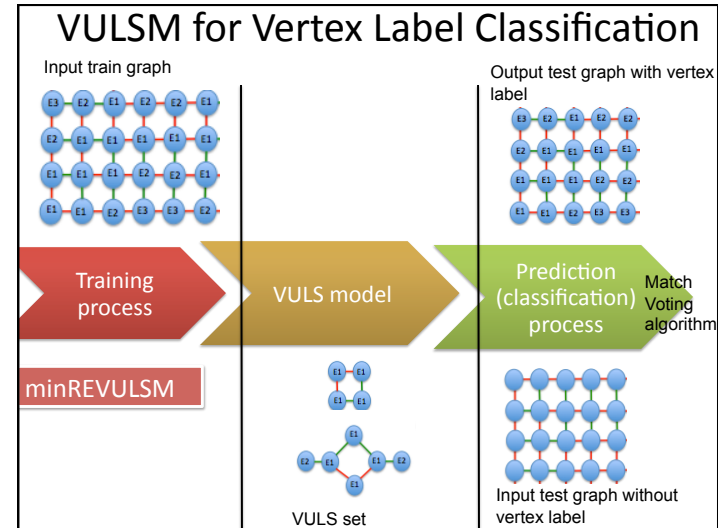■Evaluation using real data in the context of a sheet metal forming application.
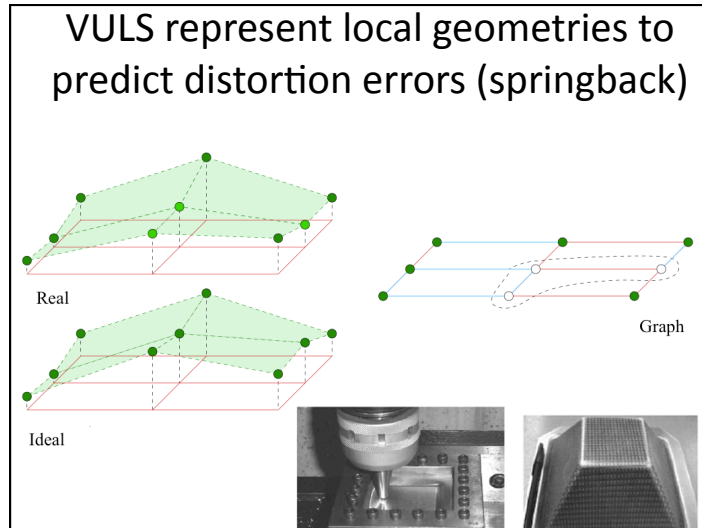

Definition of VULS

• 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

All subgraphs

VULS

Minimal VULS

VULS represent local geometries to predict distortion errors (springback)

Real

Ideal

Graph

VULSM for Vertex Label Classification

Input train graph

Output test graph with vertex label

Training process

VULS model

Prediction (classification) process

Match Voting algorithm

minREVULSM

VULS set

Input test graph without vertex label

minREVULSM 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

VULS set R

Match-Voting algorithm

G_{new} without vertex label | VULS set R | Max

G_{new}= {P_1, P_2, … , P_i}
k=1
$VULS_k$= set of k-edge VULS or minimal VULS in R
While(K<Max)**do**
    matchVoteVULS($VULS_k$)
    k=k+1
End While
For all P∈G_{new} **do**
P.Va.label=most frequently voted vertex label inVote(P.V_a)
P.Vb.label=most frequently voted vertex label inVote(P.V_b)
End For

G_{new} with vertex label

Evaluation

All VULS versus Minimal VULS

❖Number of VULS
❖Accuracy
❖AUC
❖Runtime (seconds)

Experimental results 1

max	$|L_E| \times |L_V|$	All VULS			min VULS		
		$d=10$	$d=12$	$d=14$	$d=10$	$d=12$	$d=14$
4	2×2	25	37	24	16	17	18
	3×2	137	135	112	41	35	25
	4×2	286	299	303	70	54	76
5	2×2	139	144	110	50	40	76
	3×2	611	689	573	53	113	37
	4×2	1466	1564	1628	112	113	84
6	2×2	665	696	526	118	46	204
	3×2	2732	3366	2886	87	277	42
	4×2	7137	7423	8392	205	173	89

Number of VULS Comparison

Experimental results 2

max	$|L_E| \times |L_V|$	All VULS			min VULS		
		$d=10$	$d=12$	$d=14$	$d=10$	$d=12$	$d=14$
4	2×2	29.00	9.03	43.37	38.00	26.39	83.67
	3×2	39.00	13.89	37.24	70.00	38.19	78.06
	4×2	37.00	22.92	42.35	68.00	72.22	90.82
5	2×2	30.00	11.81	51.02	70.00	65.28	88.78
	3×2	49.00	17.36	33.67	70.00	58.33	82.65
	4×2	57.00	23.61	15.31	67.00	93.75	91.33
6	2×2	24.00	15.97	47.45	70.00	72.92	92.86
	3×2	31.00	22.22	35.20	70.00	89.58	91.33
	4×2	52.00	20.83	12.76	67.00	77.78	91.33

Accuracy Comparison

Experimental results 3

max	$|L_E| \times |L_V|$	All VULS			min VULS		
		$d=10$	$d=12$	$d=14$	$d=10$	$d=12$	$d=14$
4	2×2	44.90	52.88	58.35	34.92	61.87	75.09
	3×2	47.86	7.19	33.70	50.32	19.78	50.72
	4×2	44.41	21.51	31.17	49.93	37.41	49.72
5	2×2	44.58	54.32	59.88	50.32	33.81	72.56
	3×2	52.87	8.99	31.74	50.32	30.22	53.24
	4×2	53.40	12.23	19.03	48.20	48.56	50.00
6	2×2	31.18	56.47	60.58	50.32	37.77	72.13
	3×2	37.13	11.51	32.58	50.32	46.40	57.99
	4×2	44.78	10.79	20.29	48.20	40.29	50.00

AUC Comparison

Experimental results 4

max	$\|L_E\| \times \|L_V\|$	All VULS			min VULS		
		$d=10$	$d=12$	$d=14$	$d=10$	$d=12$	$d=14$
4	2×2	0.34	0.52	0.50	0.27	0.32	0.37
	3×2	0.48	0.66	0.74	0.28	0.36	0.41
	4×2	0.59	0.82	0.87	0.40	0.38	0.41
5	2×2	0.7	1.07	0.93	0.35	0.45	0.55
	3×2	1.35	1.53	1.41	0.39	0.75	0.47
	4×2	1.50	1.69	1.64	0.42	0.60	0.46
	$4 v 3$	1.66	1.69	2.38	0.76	0.99	1.04
6	2×2	1.62	1.81	1.54	0.53	0.63	0.95
	3×2	1.91	2.78	3.24	0.46	1.15	0.56
	4×2	4.03	4.88	5.84	0.42	0.80	0.68

Runtime Comparison (seconds)

Conclusions and further study

❖ A novel mechanism for vertex label classification based on Vertex Unique Labelled Subgraph Mining (VULSM) has been described.

• Training process: Minimal Right-most Extension VULS Mining (minREVULSM) algorithm.

• Classification process: the Match-Voting algorithm.

❖ Minimal VULS mining is both efficient and effective (at least in the context of the sheet metal forming application used for the evaluation).

❖ Further work to investigate more sophisticated ways of conducting VULS based classification to improve vertex label prediction performance.

Questions and suggestions

End

Thank you for your attention!