Learning Graph Matching Substitution Weights based on a Linear Regression

Shaima Algabli and Francesc Serratosa
Universitat Rovira i Virgili,Tarragona, Catalonia, Spain
{shaima.ahmed@estudiants.urv.cat, francesc.serratosa}@urv.cat

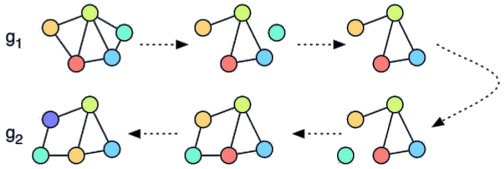
UNIVERSITAT ROVIRA i VIRGILI

Introduction

Attributed graphs are structures that are useful to represent objects through the information of their local parts and their relations.
Nodes represent local parts of the object. Edges represent relations between local parts.
We present a method to learn the weights on nodes and edges.
These weights gauge the importance of each attribute while computing the distance between graphs.

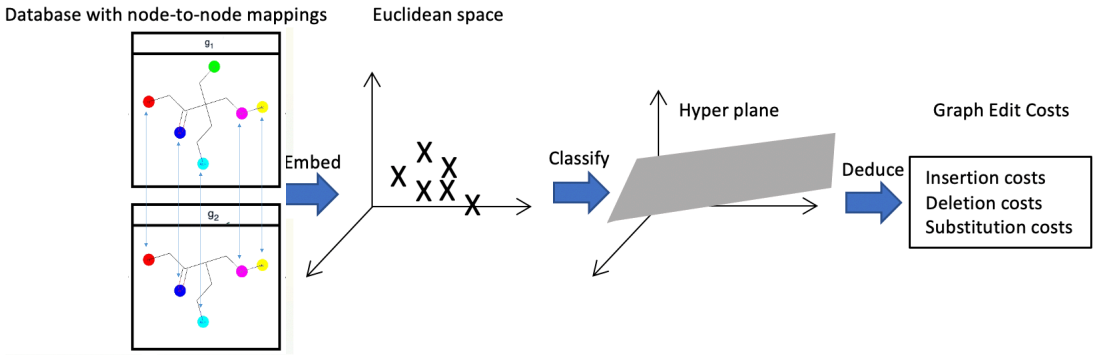
Graph Edit Distance

The Graph Edit Distance between two attributed graphs is defined as the transformation from one graph into another through edit operations.
These edit operations are: Substitution, deletion and insertion on nodes and edges. Every edit operation has a cost depending on their attributes.

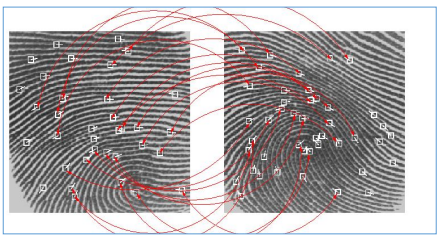


Learning model

Our learning method learns the weights on nodes and edges in two steps:
-Embedding the node-to-node mappings: *It* embeds the ground truth node-to-node mappings into a Euclidean space.
-Deducing a hyperplane: *It* computes a linear regression of the embedded points. The hyperplane constants are the weights we want to learn.



Experimental evaluation

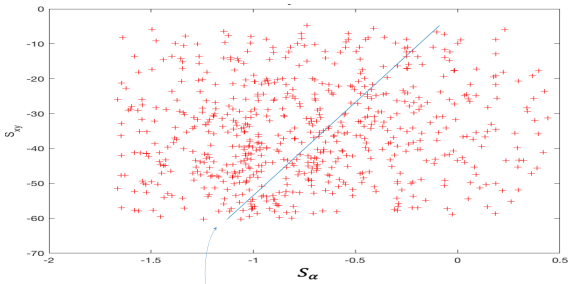


Weights to be learn

$$D(m'_\alpha, m_i) = \alpha \cdot sd(m'_\alpha, m_i) + \beta \cdot dd(m'_\alpha, m_i)$$

$$sd(m'_j, m_i) = \sqrt{(x'_j - x_i)^2 + (y'_j - y_i)^2}$$

$$dd(m'_j, m_i) = \min\{|\theta'_j - \theta_i|, 360^\circ - |\theta'_j - \theta_i|\}$$



Method	DB	1	2	3	4	5	6	7	8	9	10	11	12
Our method 0.05sec	SL	30	39	43	53	36	40	40	42	35	38	47	56
	CR	**1**	**1**	0.98	0.91	**1**	**1**	**0.91**	0.80	**1**	0.91	**0.78**	**0.57**
	H	**0.01**	**0.05**	0.06	0.19	**0.07**	**0.10**	0.15	0.20	0.07	**0.12**	0.19	0.25
Leordeano 10 min	SL	28	39	46	52	38	40	40	44	29	37	45	63
	CR	.9	1	0.95	0.90	1	0.97	0.89	0.81	1	0.92	0.81	0.57
	H	0.03	0.15	0.16	0.22	0.09	0.11	0.12	0.18	0.07	0.2	0.18	0.23
Caetanu 5 min	SL	38	24	10	4	250	142	37	42	406	72	55	59
	CR	1	1	1	1	1	0.92	0.91	0.87	1	0.93	0.77	0.57
	H	0.01	0.04	0.03	0.08	0.08	0.14	0.15	0.18	0.06	0.12	0.19	0.25
Cortés 8 min	SL	30	39	Inf	Inf	36	41	Inf	42	-	-	48	64
	CR	1	1	0.12	0.03	1	1	0.01	0.8	-	-	0.78	0.55
	H	0.01	0.05	0.66	0.67	0.07	0.10	0.63	0.20	-	-	0.19	0.26

$$SL = -\frac{\alpha}{\beta}$$

CR: classification ratio
HD: Hamming distance

This research is supported by NanoInformaTIX: Development and Implementation of a Sustainable Modelling Platform for NanoInformatics, H2020-NMBP-14-2018-814426.