A Fuzzy-Interval Based Approach For Explicit Graph Embedding

Muhammad Muzzamil Luqman[1,2], **Josep Llados**[2], **Jean-Yves Ramel**[1], **Thierry Brouard**[1]

[1] Laboratoire d'Informatique, Université François Rabelais de Tours, France
[2] Computer Vision Center, Universitat Autònoma de Barcelona, Spain
{mluqman, josep}@cvc.uab.es, {ramel, brouard}@univ-tours.fr

ICPR'2010 contest GEPR
Sunday, 22 August 2010

Graph embedding (whats and whys)

" … is a methodology aimed at representing a whole graph (possibly with attributes attached to its nodes and edges) as a point in a suitable vectorial space … "

" Graph embedding, in this sense, is a real bridge joining the two worlds: once the object at hand has been described in terms of graphs, and the latter represented in the vectorial space, all the problems of matching, learning and clustering can be performed using classical Statistical Pattern Recognition algorithms. "

ICPR'2010 contest GEPR
http://nerone.diiie.unisa.it/contest/description.shtml

Explicit graph embedding

"… one vector per graph … "

ICPR'2010 contest GEPR
http://nerone.diiie.unisa.it/contest/description.shtml

- **Encode details on the structure, topology and geometry of underlying graph**

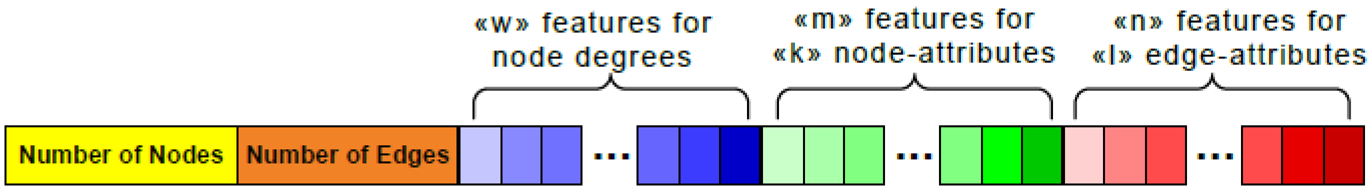
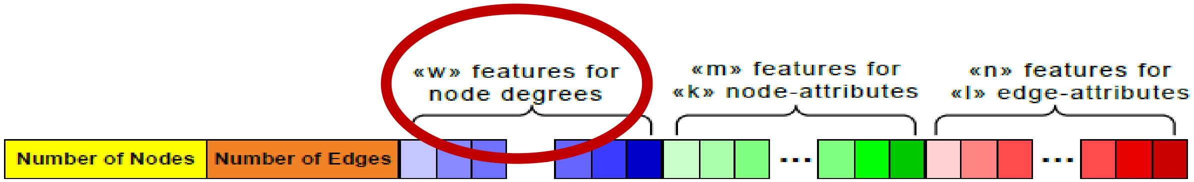


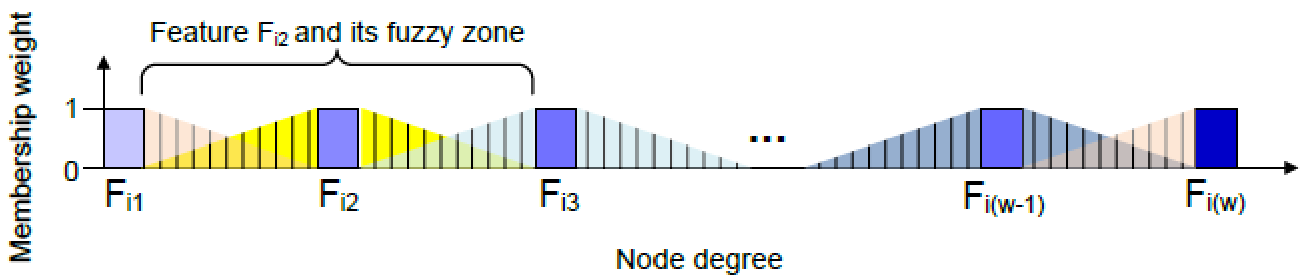
Fig. The proposed feature vector

Length of feature vector: 1+1+w+(m)+(n)

Features for node degrees



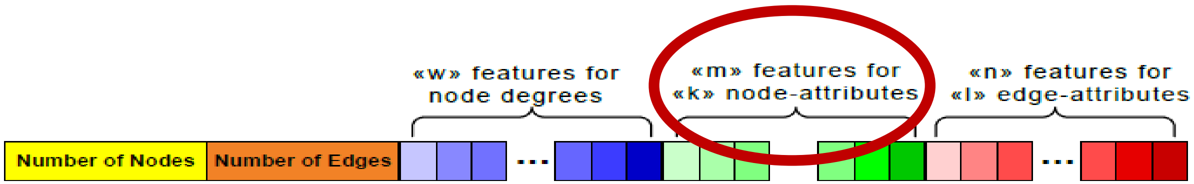
1. Start with node-degrees of all the graphs in dataset (learning - if available)
2. Histogram based binning technique for obtaining an initial set of bins [1]
3. Arrange these initial set of bins to define fuzzy intervals (i.e. features)
    1. 3-bins forms 1-fuzzy interval (as used for GEPR)
    2. Fuzzy zones of (full, medium, low) membership around each fuzzy interval
    3. Membership weights of 1.00, 0.66, 0.34



4. For each graph in dataset (testing) compute the "w" features for its node degrees (i.e. the cardinality of each fuzzy interval)

Number of features for node-degrees = w

Features for node-attributes

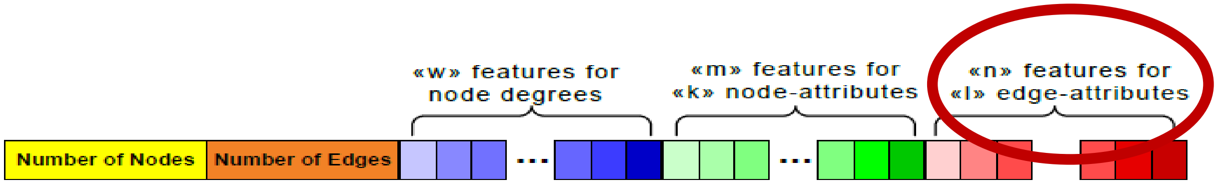


For each node-attribute (K_i)

1. Start with the values of the "K_i" node-attribute of all the graphs in dataset (learning – iff)
2. Histogram based binning technique for obtaining an initial set of bins [1]
3. Arrange these initial set of bins to define fuzzy intervals (i.e. features)
   1. 3-bins forms 1-fuzzy interval (as used for GEPR)
   2. Fuzzy zones of (full, medium, low) membership around each fuzzy interval
   3. Membership weights of 1.00, 0.66, 0.34
4. For each graph in dataset (testing) compute the "Kx_i" features for its "K_i" node-attribute (i.e. the cardinality of each fuzzy interval)

Number of features for node-attributes = m = $\sum Kx_i$

Features for edge-attributes

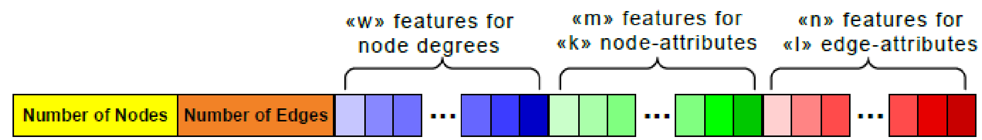


For each edge-attribute (L_i)

1. Start with the values of the "L_i" edge-attribute of all the graphs in dataset (learning – iff)
2. Histogram based binning technique for obtaining an initial set of bins [1]
3. Arrange these initial set of bins to define fuzzy intervals (i.e. features)
    1. 3-bins forms 1-fuzzy interval (as used for GEPR)
    2. Fuzzy zones of (full, medium, low) membership around each fuzzy interval
    3. Membership weights of 1.00, 0.66, 0.34
4. For each graph in dataset (testing) compute the "Lx_i" features for its "L_i" edge-attribute (i.e. the cardinality of each fuzzy interval)

Number of features for edge-attributes = n = $\sum Lx_i$

Conclusion



- Explicit graph embedding for undirected attributed graphs
- Proposed feature vector encodes details on,

  1. the structure of graph
  2. attributes of nodes
  3. attributes of edges

- **A basic fuzzy approach to incorporate robustness against graph-deformations**
- In future planning to extend this work for directed attributed graphs

References

[1] Colot, O., Olivier, C., C., P., A., E.M.: Information criteria and abrupt changes in probability laws. In: Signal Processing VII : Theories and Applications. pp. 1855-1858 (1994)