Maximal Cliques in Association Graphs

Éireann Leverett

February 9, 2005

1 A definition of graphs and K-Cliques

A graph can be represented by two sets, V and E.

$$G = (V, E) \tag{1}$$

V is the set of vertices or nodes, or numbers representing those nodes.

$$V = \{1, 2, ..., n\} \tag{2}$$

E is the set of edges connecting those nodes, represented by two element subsets of V. Essentially, this is a pair of nodes and edge exists between them.

$$E \subseteq \{2v \subseteq V : |2v| = 2\} \tag{3}$$

So if we have a set of vertices $V = \{1, 2, 3\}$ and edges $E = \{(1, 2), (1, 3), (2, 3)\}$, we have a triangle.

A K-clique is a graph or sub-graph that is all-connected, for example our triangle has every possible edge, and represents a 3-clique. If we added more nodes and edges, without deleting any, we would still have a 3-clique subgraph, perhaps even a larger K-clique.

2 Association Graphs

An association graph is constructed when trying to match model features to target features. This essentially represents the search space of any matching algorithm, though not as a tree. If for every model feature that is potentially matchable with a target feature using unary constraints (same length within some threshold for example); we construct a node. Ideally we have eliminated many model-target feature matches already, or our model to target mapping thresholds are very liberal.

From these nodes, we extend an edge to every other node, that is consistent with this node via a binary contraint (such as the angle between two lines, or the ratio of lengths of two lines) within a certain threshold. We now have a graph representing all possible model-target matches within our thresholds, and we want to find the best fit within it.

This best fit is the maximal K-clique subgraph, of our search-space graph. The maximal K-clique smaller than the size of the search space graph (subgraph) is almost always present.

It should be clear that the largest all connected section, or K-Clique subgraph, is an optimal model-feature matching, although there are perhaps cases where two subgraphs match equally well, and some other heuristic is needed to decide between them. Again this is an odd case, and should not be expected very often, when good constraints, and thresholds exist.

3 Efficiency

Finding a K-clique subgraph is an NP-Hard problem (computationally very inefficient, intractable very quickly as the input grows). This means that the logic of this method is sound, but for large graphs, and large numbers of features it rapidly becomes computationally intractable.

This is not the end however, as approximation algorithms for K-cliques or maximal cliques can be quite efficient. Recent research has had success in implementing these techniques using an approximation algorithm instead.

4 Other Resources

- Attributed tree homomorphism using association graphs:
  http://ieeexplore.ieee.org/xpl/abs_free.jsp?arNumber=906033
- Definition of NP-Hard:
  http://www.nist.gov/dads/HTML/nphard.html
- Definition of Cliques:
  http://www2.toki.or.id/book/AlgDesignManual/BOOK/BOOK4/NODE172.HTM