Point of View Based Clustering of Socio-Semantic Networks

Juan David Cruz*, Cécile Bothorel*, François Poulet**

*Département LUSSI – Télécom – Bretagne
{juan.cruzgomez, cecile.bothorel}@telecom-bretagne.eu,
http://www.telecom-bretagne.eu/
**Université de Rennes 1 – IRISA
francois.poulet@irisa.fr
http://www.irisa.fr/texmex/

1 Introduction

Socio–semantic networks contain two different types of information, the social relationships and information related to the actors such as thematic profiles.

In order to analyze such augmented network from different perspectives, we propose to influence the communities detection process with semantic information, the clustering process is divided into two phases. In the first one, the point of view is clustered using Self – Organizing Maps (SOM) (Kohonen (1997)) to obtain groups based on the similarity of the node features and changing the weights of the graph according to the semantic groups found. Then, in the second phase, a classic community detection algorithm is used, based on the topology of the network.

2 Using the Point of View to Influence the Clustering

A social network can be represented as a non–directed graph $G(V, E)$ where V is the non-empty set of vertices representing actors and E is the set of edges representing the relationships among them.

Given a graph $G(V, E)$, $\mathcal{C} = \{C_1, C_2, \ldots, C_k\}$ is a partition of the set V into k non–empty disjoint subsets C_i.

Let F_V be the set of semantic features of the actors of the social network, which can be represented by a matrix of size $|V| \times |F_V|$.

Let $F_V^* \in \mathcal{P}(F_V) \setminus F_V$, where $\mathcal{P}(A)$ is the powerset of the set A, be a non–empty set of features to be used to define the point of view PoV.

For each vertex $v_i \in V$ we assign a binary vector $\xi_i = v_i \times F_V^*$ of size $|F_V^*| = f$. If the vertex i has the feature $p, 1 \leq p \leq f$ from F_V^*, then $\xi_{i,p} = 1$ or 0, otherwise.

A point of view is defined as the set of all instances derived from the set F_V :

$$PoV_{F_V^*} = \bigcup_{i=1}^{|V|} \xi_i \qquad (1)$$