

Posibles Aproximaciones para una Herramienta de Visualización de Redes Sociales

Fabio González, Juan David Cruz
Departamento de Ingeniería de Sistemas y Computación
Universidad Nacional de Colombia

Febrero 6 de 2006

Se debe definir cual va a ser la utilidad diferenciadora de la herramienta desarrollada. Para esto se encontraron cuatro preguntas que se pueden responder:

1. ¿Con quién me comunico?
2. ¿Sobre qué?
 - a) ¿Quién sabe qué?
3. ¿Cuanto y con qué frecuencia?

La herramienta va a ser diseñada de modo que muestre de manera gráfica las interrelaciones entre personas, como un grafo, cuyo centro será la persona dueña de los mensajes. Se pueden utilizar diferentes convecciones, por ejemplo, la longitud entre nodos, el grosor del enlace y el diámetro de cada nodo para simbolizar diferentes niveles de información. Adicionalmente, se deben representar los diferentes temas involucrados, pero guardadndo una relación la red de correos.

Para esto se plantean dos estrategias:

1. Ubicar los nodos y con base en ellos, ubicar los temas/regiones:

- a) Definir una métrica: Medir que tan cerca está un punto de otro:

$$\left\{ \begin{array}{ll} yo - punto & \text{VolumenyFrecuencia} \\ punto - punto & \text{Temas Comunes} \end{array} \right.$$

- b) Aplicar Multidimensional Scaling (MDS)
- c) Asignar temas a pixeles
- d) Segmentar
- e) Asignar etiquetas a regiones

2. Definir temas/regiones y con base en esto ubicar los puntos:

- a) Entrenar SOM utilizando correos. Incializarlo utilizando los temas más frecuentes: definir características (Ver adelante).

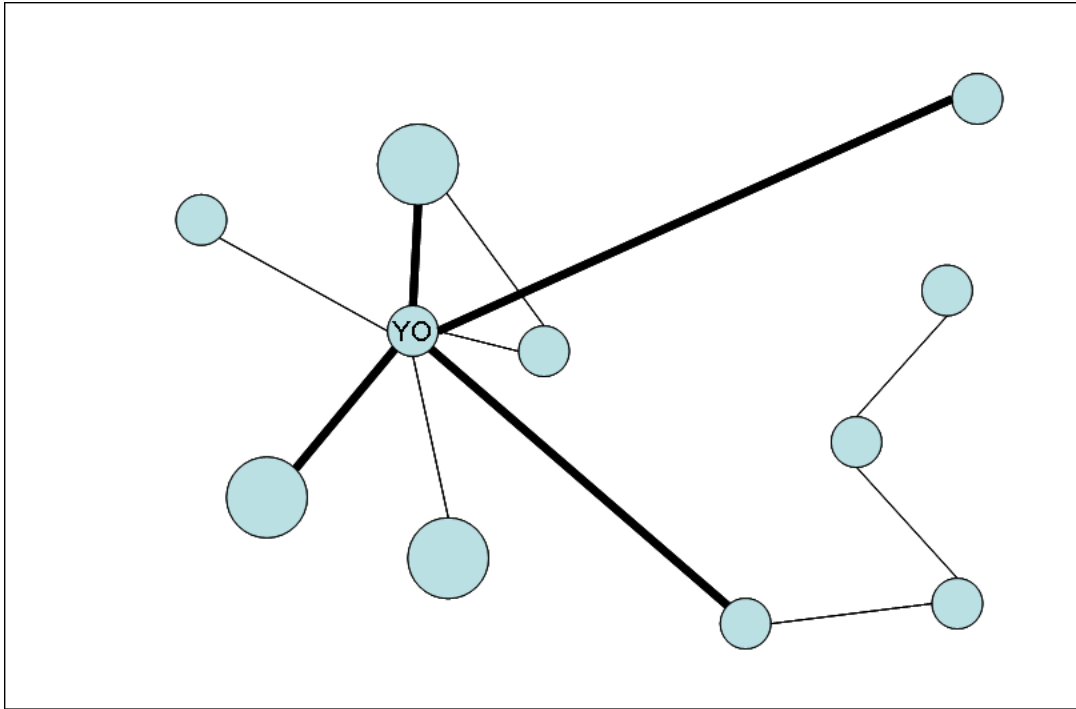


Figura 1: Representación de la red (Sin regiones de temas)

- b) Definir regiones usando las técnicas estándar para SOM
- c) Asignar puntos a regiones usando el SOM como clasificador.

Cada correo tiene unas características no explícitas, es decir, diferentes a las de destinatario y remitente. Estas son corresponden a las palabras claves y temas relevantes. Para esto se plantean los siguientes pasos:

1. Cada contacto (correo?) se representa por frecuencias de palabras clave.
2. Encontrar temas más relevantes: agrupamiento de palabras claves.
3. Reducción de representación de vectores de palabras a vectores de temas.
4. Cada contacto (correo?) se representa por el centroide de los temas de los correos asociados.