Data Collection

Web Mining 2013

18/03/2013

1 Individual data collection

Each student has to collect data describing his own multidimensional social network. In a multidimensional social network, multiple links can exist between two persons representing different kinds of relationships [1, 2].

- We are interested in four types of interactions: i) **online interactions** (Facebook, Twitter, Google+); ii) **real life interactions**.

  Data must be collected on a weekly basis: every week you must specify with whom of your colleagues you interacted on Facebook, Twitter, Google+ and in real life.

- For online interactions, specify the colleagues with whom you interacted. An online interaction is a post or a comment (Facebook/Google+), and a tag or reply (Twitter). Do not consider "Likes" (Facebook) and "+1" (Google+) as interactions.

  Don't specify the number of posts/comments or tag/replies, just indicate whether there was at least one post/comment or tag/reply during the week.

- For real life interactions, specify the persons in the class with whom you interacted outside the University context (person with whom you participated together in leisure/social activities like nightlife, sport, music and so on).

- For each online and real life interaction, indicate the **strength** of the interaction: 1 (low), 2 (medium-low), 3 (medium), 4 (medium-high), 5 (high).

  The strength is defined as the intensity of the social relationship (in terms of quantity and quality of interactions).

  For online interactions, compute the strength based on the quantity and the quality of post/comment or tag/reply.

  For real life interactions, compute tie strength based on the quantity and the quality of your leisure/social experience.

- **Research only (optional)**: Indicate also the trust of each interactions (use the same ranking from 1 to 5). The trust indicates at what extent you consider a relationship reliable (in terms of honesty, fairness, and benevolence).

- **Warning:** Only interactions with the other persons attending the Web Mining course should be taken into account. We are **not** interested in the interactions between you and a person outside the Web Mining course.

2 Individual network reconstruction

Use the collected data to construct your individual multidimensional network.

- Together with the other students, create a shared map in which each person is associated to an alphanumeric alias of 4 characters. For example:

Paulo Sinscia	14D3
Salvador Rinzivigio	TY76
Pablo Escobar	PDL2
...	...

- Use the shared map to create a file in which each row represents information about the interaction between you and your colleagues. The file must have the following format:

id_user1	**id_user2**	**network_id**	**week_id**	**strength**	**[trust]**
yourID	PDL2	2	4	1	2
yourID	GH65	1	4	5	4
...	...	...	...	...	...

Table 1: Example of file representing an individual multidimensional network

  Words in the file must be delimited through the "tab" character.
  The network identifiers to be used are the following:

  1 Real life

  2 Facebook

  3 Twitter

  4 Google+

- Now, name the file `network_yourID`, and send it by mail with subject [WM2013 - Data Collection] to the following addresses:

  lpappalardo@di.unipi.it

  giulio.rossetti@isti.cnr.it

  pedre@di.unipi.it

References

[1] Wikipedia: `en.wikipedia.org/wiki/Multigraph`

[2] M. Berlingerio, M. Coscia, F. Giannotti, A. Monreale, D. Pedreschi, Foundations of Multidimensional Network Analysis, ASONAM 2011:485-489.