Social psychology testing platform leveraging Facebook and SNA techniques

Testing platform integrated with Facebook to introduce Social Network Analysis techniques in social studies.

Andrea Biancini INFN - Sezione Milano Bicocca,
Università degli Studi di Milano Bicocca - Edificio U2
Piazza della Scienza, 3 - 20126 Milano, Italy.
andrea.biancini@mib.infn.it

June 7, 2012

In psychology, testing is largely used in order to assess mental constructs, such as cognitive and emotional functioning, of individuals. Tests permit researchers to easily reach a large population of subjects. Tests introduce also quantitative aspects to the psychological introspection mechanisms.

To design tests and run them an open source platform is commonly available. This platform, called OpenSesame, provides researchers with an advanced and easy to use experiment builder to support their researches.

This article presents a platform designed to provide additional social information to the tests implemented on the OpenSesame platform. The platform presented is able to compute sociological indexes on the Facebook networks of the subjects using SNA techniques. These indexes are meant to integrate the data coming from test execution and enrich it with social information about the subjects.

Keywords: Social sciences psychological tests OpenSesame Facebook social network analysis Software

1 Introduction

The word psychology literally means, the study of the soul. This gives the idea of the vastity of the eld of study of psychological sciences: the mind. Studying the mind is quite a complex task because of the immateriality and unobservability of the subject of study. To circumvent this limitation psychologists have developed different research techniques that aim to analyze constructs not directly observable. Amongst these tools, psychology makes a signicant use of tests to perform its analysis (like in [4]).

A psychological test is an instrument created and designed to measure psychological constructs, modeled through so called latent variables. Psychological tests are typically, but not necessarily, a series of tasks or problems that the respondent has to solve. Many different test techniques have been developed over time, one of the most structured and quantitative is represented by personality inventories (with the different approaches compared in [3]) or structured multiple-choiche tests. In these tests, the subject is requested to go through a series of questions and provide answers choosing between different options presented. These choices could be structured on different scales (one of the most commonly used scales is the [8] scale).

The possible elds of application of psychological studies are very numerous and broad. Together with the study of singular cases and the generalization of mental constructs, psychology very often refers to the sociological world. Social psychology ([11]) is dened as the scientic study of how peoples thoughts, feelings, and behaviors are inuenced by the actual, imagined, or implied presence of others. Social psychologists typically work to explain human behaviors interpreting them as the result of the interaction of internal mental states and the immediate social situations. Social psychology is an empirical science and, as psychology in general, can leverage tests in its analysis (social psychology relies on a articulated methodology, as exposed in [10]). Tests, as seen, can be easily diffused and can potentially hit a large population. These wide numeric data is often very useful in describing social trends and interactions.

The main problems in dening, submitting and executing tests in the eld of social psychology are about these aspects: test creation, test submission and the collection of relevant test data. To support the ease of test creation, the OpenSesame project ([**?**]) offer a graphical experiment builder for the social sciences. OpenSesame is an opensource experimental software which is easily able to be used in combination with existing sofware. The OpenSesame application has been extended to be integrated with a testing platform, presented in this paper, that runs as a Facebook application and collects and compute social indexes on the subjects networks of contacts. With such integration it is possible to achieve a better diffusion of the tests (leveraging Facebook social features) and to enrich test data with relevant social information about the subjects (computed on their Facebook networks). The platform developed, called the Rorschach Test Platform, has the aim to reach these goals. This paper will present the platform and explain from a functional and technical point of view the benets provided by this new tool.

The rest of this paper is structured as following: in the next section the benets of using Facebook as a mean of distribution is described. The second section will present social network analysis (SNA) techniques and will describe all the indexes computed within the proposed platform. The third section will present the overall functionalities developed in the platform and how they can be used in conjunction with the OpenSesame platform to create and run real tests. The last section will show conclusions and possible future works on this topic. The appendix will present some detail about the technical design and implementation of the platform.

2 Leveraging Facebook social data and means of distribution

The advancements in technology and in computer science permit the growth of social medias like Facebook and other social networking paltforms. These tools, leveraging on modern advancements in computer science technologies, permit to gather and store lots of data about users and their activities.

Social network sites (SNSs) allow individuals to present themselves, articulate their social networks, and establish or maintain connections with others ([2]). Great attention has been cast upon these tools. They in fact offer large quantities of user data that could be used in sociological studies (as, for example, in [5]).

Approaching these technologies, it is important not to be dragged by the stimulant opportunities, but to take into serious consideration the scientic design and validity of the research. Its however undoubt that a conscious and careful use of these tools could be strongly benecial in the sociological eld.

Objective of this work. This article has the objective to present a platform designed with the aim to give sociological testers new tools and information leveraging Facebook as the most successful SNS to date. The underlying idea is that, with a smart integration with Facebook, two main benets can be obtained:

- On one side, Facebook has a very high permeation, expecially in in the western world. This popularity could permit a tremendous diffusion with a very limited cost. This makes very attractive the usage Facebook as a platform to diffuse psychological tests.
- On the other side, Facebook is also a very powerful tool able to mock the social life of its participants. As in [7], different datasets of information can be extracted from Facebook and used for sociological studies. This could permit researchers to enrich the information coming from the analysis of psychological test answers, and obtain higher levels of insight.

3 Social network analysis techniques

The social eld, as the psychological one, is an interdisciplinary eld of study and knowledge where different scientic methods and approaches work together. In this eld the mathematical study of networks has brought relevant results (as described in [**?**]). The techniques that permit this sort of quantitative analysis over networks and social interactions go under the name of Social Network Analysis (SNA) [**?**].

These tecniques are very promising and, if well used in scientic experiments, could bring tremendous results in analyzing the social relationships and behaviors of subjects. As described in [9], users information based on SNA indices increases the awareness of their social context, their collaboration and social interaction, and creates a sense of social presence at a group level.

All the sociological analysis are performed starting from a representation of the network of contacts for the user.

As shown in figure 1, a Facebook network is represented as a graph where the nodes are the friends of the given users and the connection amongst the friends are represented as arcs between the nodes of the graph.

The mathematical representation of the network graph permits to apply all the techniques of Graph Analysis (a branch of algorithmic mathematics), as described in [1]. Mathematically, we can dene a Graph as in 1:

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

where V is a set of nodes and E a set of edges. The graphs representing Facebook networks are undirected graphs (if A is friend of B, B is by denition friend of A). This implies that the elements of E are unordered pairs of elements of V.

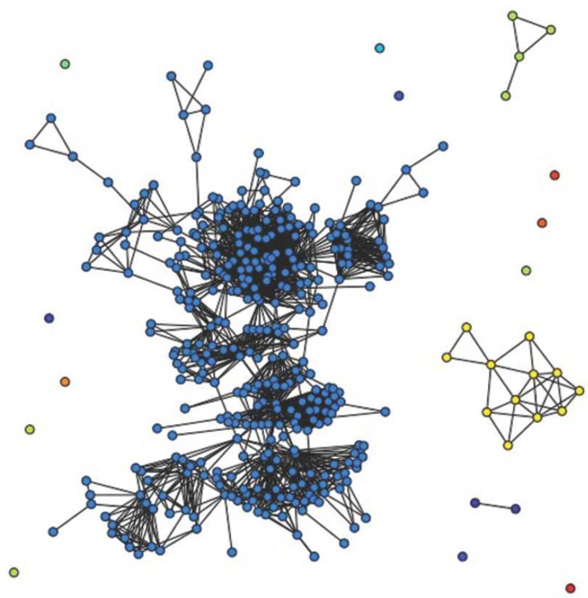

Figure 1: Visualization of a Facebook network

With this denition we can apply graph theory to study the graph traversal and generation and the complexity of these operations. Upon this mathematic denition, Social Network Analysis is at this point able to compute measures about social connections. Different measures can be computed, in the following the measures computed by Rorschach Test Platform will be described. They can roughly be divided into three sets:

1. basic indexes;
2. centrality indexes;
3. subgroups indexes.

3.1 Basic indexes

Basic indexes are measures computed on the more direct and visible structural characteristics of the graph. They can be easily measured from the graph data (nodes and edges) and produce information on the network as a whole (i.e. these indexes do not refer to any specic node or subset of the network but they refer to the complete network).

The Rorschach Test Platform computes four such indexes:

- *Density*: This index measures the density of the network intended as the number of edges actually present in respect to the total theoretical number of edges. The density of the network is the ratio of edges present in the network and the total number of possible edges of the network. A network of N nodes can have a maximum of N * (N-1) edges.

The formula for this index is, then, the one presented in 2.

$$density(G) = \frac{|E|}{|N| * (|N| + 1)} \tag{2}$$

Facebook networks are usually very sparse, so very low values of this index are expected. Typical values can be around .05 or even less.

- *Geodesic*: In graph theory, the geodesic distance is the smallest path length between two nodes. This index, then, measures the average of all the geodetic distances between all connected nodes in the network.

  This value represents the number of connections each node has to contact in order to reach every other node in the network.

  Given the denition of geodesic distance between two nodes as in 3,

$$\begin{aligned} distance(v_1, v_2) = min\left(d(v_1, v_k) + d(v_k, v_2)\right), \forall v_k \in V \\ \text{where } d(v_i, v_j) = \begin{cases} 0 \text{ if } v_i = v_j \\ 1 \text{ if } v_i, v_j \in E \\ +\infty \text{ otherwise} \end{cases} \end{aligned} \tag{3}$$

  the formula for this index can be expressed as in 4.

$$geodesic(G) = \frac{\sum_{v_1, v_2 \in V} distance(v_1, v_2)}{|E|} \tag{4}$$

  This index is lower the more connected the network is. In fact connected networks offer short paths between two randomly choosen nodes. Typical values for quite strongly connected Facebook networks are around 4.

- *Fragmentation*: The fragmentation index shows how much the network is fragmented, that is how many node pairs of the network cannot reach other node in the network. Thisindex is the ratio of the number of node pairs for which the network does not have a path to connect them. Nodes not connected may be isolated nodes or nodes belonging to two non connected subgraphs.

  This index gives an indication about how uniformely the network can be divided. Very low values indicates networks that shows very few unconnected subparts. Higher values indicate networks in which different components and subparts can be identied.

This index can be computed starting from a matrix of geodesic distances between all the nodes. This matrix can be computed using the Floyd-Warshall algorithm [6].

The formula for this index can then be expressed as in 5.

$$\begin{aligned} fragmentation(G) = count(n_1, n_2), \forall n_1 \in V, n_2 \in V \\ \text{where } distance(n_1, n_2) = +\infty \end{aligned} \tag{5}$$

Usual Facebook values could be quite high, usual social networks have more than one set of nodes (i.e. parents, friends form work, friends from university, ...). A typical real value for this index could be around .20.

- *Diameter*: The diameter of a network is the longest geodesic between two nodes randomly choosen inside the network. This index measure the longest path that can be followed inside the network between its nodes without starting a circle. The lower this value, the more interconnected the network is.

  This index is computed from the matrix obtained with Floyd-Warshall algorithm which computes the geodesic distance between each pair of nodes in the graph.

  The formula for this index can be expressed as in 6.

$$\begin{aligned} diameter(G) = max\left(distance(v_1, v_2)\right), \forall v_1 \in V, v_2 \in V \\ \text{where } distance(v_1, v_2) \neq +\infty \end{aligned} \tag{6}$$

  Typical values may be around 10. Networks with smaller values are more connected.

3.2 Centrality indexes

3.3 Subgroups indexes

References

[1] Biggs N, Lloyd E, Wilson R (1999) Graph theory, 1736-1936. Clarendon Press

[2] Boyd D, Ellison N (2007) Social network sites: Definition, history, and scholarship. Journal of Computer-Mediated Communication 13(1):article 11

[3] Burisch M (1984) Approaches to personality inventory construction: A comparison of merits. American Psychologist 39(3):214–227

[4] Cohen R, Swerdlik M, Smith D (1992) Psychological testing and assessment: an introduction to tests and measurement. Mayfield Pub. Co.

[5] Ellison N, Steinfield C, Lampe C (2007) The benefits of facebook "friends": Social capital and college students' use of online social network sites. Journal of Computer-Mediated Communication 12(4):1143–1168

[6] Lawler E (2001) Floyd-warshall method. Combinatorial optimization: networks and matroids 1:86–92

[7] Lewis K, Kaufman J, Gonzalez M, Wimmer A, Christakis N (2008) Tastes, ties, and time: A new social network dataset using facebook.com. Social Networks 30(4):330–342

[8] Likert R (1932) A technique for the measurement of attitudes. Archives of Psychology 22(140):1–55

[9] Martino F, Miotto A, Davide F, Gamberini L (2007) Exploring social network indices as cues to augment communication and to improve social practices. In: Proceedings of the 1st International Workshop on Maps Based Interaction in Social Networks (MapISNet), Rio de Janeiro, Brazil

[10] Reis H, Judd C (2000) Handbook of research methods in social and personality psychology. Cambridge University Press

[11] Smith P, Bond M, Kâğıtçıbaşı Ç (2006) Understanding social psychology across cultures: living and working in a changing world. Sage social psychology program, SAGE