NetW4PPL User Interface to fill data

M. Geier, F. Zanollo, M. Ruggiero, J. P. Baraybar, I. Caridi

April 3, 2021

1 Motivation

This User's Interface is a free, open-source, standalone, and multi-platform software designed to help users fill data to formalize networks in the context of missing people. The formalization of the networks allows in the preliminary research of missing individuals [1] and, at the same time, many users are not so familiar with the networks tools. For this reason, the software becomes necessary for this problem. Therefore it was developed in a general open free tool and supported in different languages (until now English, French and Spanish). The idea is that this software could facilitate collaborative work in projects where the information is fragmented, and there are many different sources and organizations gathering data. Besides, it could be applied to other related problems.

2 User's Manual of the Interface

The Interface shows the User two windows offering one table each (one for individuals´ attributes and one for their relationships). The Interface enables the generation of the following two files from each table (in .csv format):

- file of individual's attributes (ATT)
- file of relationships between actors (LIN)

For a particular Protocol (with a specific list of attributes and options for some of the variables), the Interface allows filling a new actor in a new row (pushing the button *new row* in the table menu), or a new relationship in the same way.

ID assignation

For each new row, the Interface automatically assigns a unique code (ID) to each new row (representing the ID of the individual) starting from the initial ID chosen by the User (in the top-left button) and for the rest of the assigned IDs, following the last one in a correlative way.

To define the first individual's ID, the User should set the first desired ID in the Interface's upper left window (ID window) in this way: three letters code followed by the symbol ”-” and then a five-digit numeric code. The User can redefine the desired ID at any moment, and it will assign the ID of the following ”new row” in the attributes (ATT) table. If the User does not change the ID (in the window's ID), the ID assigned to a new row will be correlative to the one previously generated.

Attributes with special input

Users can fill in some of the attributes using the dropdown option or spinbox (both are enabled by pressing the space key or double click). Others have autocompleted option enabled and will show options as User's type. You can see the Specifications section for how to modify them.

As of this implementation, it is as follows:

- Attributes with dropdown option
  - role
  - route
  - day
  - month
- Attributes with spinbox option
  - age
  - year
- Attributes with autocompleted option
  - nationality
  - Residence

Adding new relationships

The Interface allows adding new relationships between actors automatically (undirected links until now). Cases can be selected in the attributes window (by clicking them in the first column). After choosing the desired individuals in the attributes window, and by clicking the button *add links* shown by the mouse right option, the Interface will automatically generate new connections in the LINKS window. The new links are all the possible relationships between those cases selected.

In each case, the User must choose the kind of relationship corresponding to each pair through the dropdown option in the *type of relationship* of the LIN table. The interface will automatically copy the ID of the individuals involved in the relationship and their names.

Delete options

The delete option (both in the attributes table and relationships table) is available in the software. It is necessary to select the cases to be deleted and click the button *delete* shown by the mouse right option.

Load and Save options

The interface allows saving the loaded information using the *save* buttons from the attributes and relations tables menu.

Two files will be generated (one for each table) in .csv format (which could be imported in Excel or another program). Previous tables (generated by the same user or another one) can be loaded using the *load* button from the same Menu (they must be in .csv format, which can be exported from Excel).

There's also an option of *new table.* This option will clear and delete current tables (attributes and links).

Column filters

Filters are enabled by default. There's a row right below the header for their input. It will show only rows with data starting with the text proposed; rows have to fulfill all filters to appear. The user can disable them by clicking the checkbox called *Filters.* This option will also clean and undo all filters applied.

Some observations

This implementation of the Interface does not check the data loaded by the User to determine whether they are consistent or not (both attributes and relations). In case the User load a previous attribute's table, the User should define the first new ID in the ID window. Otherwise, no ID will be generated from the loaded table but independent of this.

Future steps

The following stages of the Interface will include new options helping the User to fill in data. At this moment, the option of visualizing the network is implemented in R script, which interacts with the Cytoscape program. Moreover, this script includes the possibility of generating a list of data as an output from a given network (see section Networks application for details) regarding how to use networks to prioritize activities. In future versions, the loading of several databases simultaneously and their interaction will be allowed.

3 Developer's Specifications

The file *att_fields.csv* contains information regarding the list of variables to be used in the attributes table (called the *header*). Each variable has some associated parameters:

1. *name_TR*: internal name used by the program for the variable.

2. *name*: variable´s visible name in the interface.

3. *type*: variable´s type of data. An option among the following ones:

   - code: variable of the code type. An example of this is the case of the Individual´s ID.
   - text: when the variable is the text type, it's possible to autocomplete it using a specific dictionary's suggestions. Autocomplete suggestions dictionary is a .csv file whose name is defined in the parameter *parameter* of the same variable.
   - free: when the variable is completed with free text (such as people's names).
   - category: when it is a categorical variable that takes a fixed set of values displayed in a dropdown. The set of values defined in the .csv file found in the parameter called *parameter* of the same variable).
   - date: variable of date type, a free field is expected when variables are of this type. The User must complete the date in a specific format. Future versions of the Interface will implement a validation of the data that the User loads.
   - numeric: numerical type variable (for example, birthdate). In this case, the range of values for the variable is defined, such as ”minimum value / maximum value”. If it is a day of the month, the range is 1/31. If it is a month, the range is 1/12. This variable type will have a dropdown option for input if they are less than forty options and a spinbox otherwise.
   - boolean: a variable that can take only two categories, ”yes” or ”no” (to implement it in the future versions). For now, we define these types of variables as a category with a dictionary that has two options: yes/no.

4. essential: It refers to whether it is essential or not that this variable is complete to define a new row in the table. The current version does not implement this option. The idea is that there are only a few variables that are essential (essential = yes). It is necessary to revisit this option when the interface allows loading new data by the User. The current version of the interface does not use this option, and the User can always define new rows in both Attributes and Links tables.

5. parameter: according to the type of variable, it has different information in this column: the name of the file containing the dictionary if it is a category variable or a numeric range in the way ”minimum value/maximum value” if it is a numeric type variable.

6. default_value: A default value can be set only for numeric and category types, have in mind that the option should be valid.

In the file *lin_fields.csv* there is information about links. This file also follows the same idea as that of the attribute file to characterize the variables with type, essential, and parameter. For example, the relationship is also a category variable, and the dictionary where the possible values are is in dict_relationship.csv.

There's an *att_fields.csv* and *lin_fields.csv* for each language, divided into said languages folder. This allows for different behaviour in every language. When building the app *tr.py* script will search for the configuration files inside *EN* folder for English, *ES* for Spanish and so on.

4 Licence

Referencias

[1] A. Salgado, J.P. Baraybar, I. Caridi, NetW4PPL *Net-Works* for People: complex networks to plan activities in contexts of missings (in preparation 2021).