A Model for the Generation of Social Network Graphs

Marco Conti, Andrea Passarella, Fabio Pezzoni
CNR-IIT, Via G. Moruzzi, 1 - 56124 Pisa, Italy
{*m.conti, a.passarella, f.pezzoni*}@iit.cnr.it

Abstract—In this paper we present and evaluate a social network model which exploits fundamental results coming from the social anthropology literature. Specifically, our model focuses on ego networks, i.e., the set of active social relationships for a given individual. The model is based on a function that correlates the level of emotional closeness of a social relationship to the time invested in it. The size of the social network is limited by the time budget a person invests in socializing. We exploit the model to define a constructive algorithm to generate synthetic social networks. Experimental results show that our model satisfies, on average, known properties of ego networks such as the size, the composition and the hierarchical structure.

Keywords-social networks; ego networks; model;

I. Introduction

The emerging pervasive and social networks are drastically changing the (information) society. First of all, we are experiencing a convergence between the cyber/virtual and physical worlds. The convergent cyber/physical world will be *content-centric* where content generated in the physical space is immediately transferred to the cyber space (e.g., multimodal sensing), and cyber outcomes have immediate impact on physical space. Humans are at the core of this convergence; each person has several (mobile) devices through which he/she can interact with the virtual world thus linking the physical world and the electronic world of users devices [1]. In this scenario, *human and online social networks* have a very important role for accessing and circulating the massive scale of content that is circulating in the network/society. By translating human relationships in the electronic world, we embed in electronic devices the key characteristics that enable humans to effectively handling and sharing large amount of information.

Human relationships can be exploited in the virtual world for fast and effective circulation of data with spatial temporal value and for content provision and personalized context, such as by sharing information of mutual relevance.

There is significant evidence suggesting that *human* social networks (i.e. the set of social relationships people maintain with each other) are not particularly affected by specific communication technologies [2]. Therefore, it is reasonable to see the properties and structures of human social networks as an invariant with respect to the evolution of the underlying means supporting social interactions.

Human social networks exhibit remarkable dynamism and structural properties that may significantly affect the quality of the information (i.e., trust and reputation, relevance, reliability, etc.) and the way information may circulate; it is conjectured necessary to transverse only a small number of human social relationships in order to connect *any* pair of people resulting in the "small world concept". Therefore, understanding and modeling human social networks is a fundamental step in designing efficient protocols for data dissemination in the cyber-physical world. In this paper we present a first important step in this direction. Specifically by exploiting social anthropology results we have developed a model of the ego network, i.e., the model describing the set of active social relationships of an individual. Results from Dunbar et al. [3], [4] indicate that human relationships have a hierarchical structure and, on average, an individual has up to 150 active social relationships, i.e., the Dunbar number. These results constitute the bases for the model developed in this paper.

The properties of the ego networks are summarized in the next section and our model is presented in Section III. In Section IV we define the functions and the parameters that characterize the model while in Section V we validate the model and formulate the conclusions.

II. Ego Networks

Ego networks are a particular category of social networks made up of an individual (an "*ego*") and the people ("*alters*") with which the ego has some kind of social relationship.

There are limits to the amount of social relationships that an individual can maintain, this is due to cognitive and time constraints [5]. In fact, keeping social relationships demands cognitive resources and time available to invest on them and both resources are limited. Different studies about ego network size have been conducted (e.g. in [4], [6], and [7]). It has been demonstrated that ego networks have a hierarchical structure that consists of a series of concentric layers of acquaintanceship with increasing sizes. Dunbar et al. suggests that the layers in an ego network are: "*support clique*", "*sympathy group*", "*band*" and "*active network*" (the whole network) with sizes ~ 5, ~ 12, ~ 35 and ~ 150 respectively, [3], [4]. The layers are hierarchically inclusive, so that each layer includes all inner levels. This structure is depicted in Figure 1. Sometimes in this paper, we use the term *external part* of a layer in order to refer to the part of the layer not overlapped with its inner levels.

Support clique and sympathy group are made up by a relatively small number of alters the ego is emotionally closest to. On the other hand the alters connected to the ego by weak ties, which represent the greatest part of the network, are included in the external layers. Each layer of the network has specific characteristics: support clique and sympathy group are well-defined in size and composition (see [8] and [9]) as well as the active network is ([4]), while no accurate information is currently available in literature about the band level. Therefore in this paper we do not explicitly model the band level and we consider it merged within the active network layer.

Regarding the correlations among the layers' sizes, the study in [8] suggests that there is a linear correlation between support clique and sympathy group. On the contrary there is no information in literature about possible correlations of their sizes and the size of the active network layer.

Relationships in social networks may be classified into different categories such as: kin, friends, neighbors, work colleagues, etc.. Moreover alters may be characterized by their gender, age, education level, marital status and so on. In social networks each relationship is also characterized by a level of *emotional closeness*. Strong relationships have a higher level of emotional closeness compared with weak ties. As suggested by Hill & Dunbar, the emotional closeness level may be the key parameter to consider in order to select in which layer a relationship has to be included [4].

The level of emotional closeness is positively correlated with the "frequency of contact", which is estimated with the inverse of the "time since last contact" [4]. The latter also reflects the time invested in a particular relationship [5], therefore it is generally assumed that there is a relation between the time invested in a relationship and the level of emotional closeness. Maintaining a relationship at high level of emotional closeness requires a lot of time invested in it, for both friends and kin. On the contrary, for low levels of emotional closeness, kin relationships require less invested time than the relationships with friends [5].

III. THE MODEL

Our model allows to define ego network graphs that, on average, have the properties described in literature, such as the size, the hierarchical structure and the composition of each layer. The model is based on parameters and functions, defined in Section IV, that are obtained exploiting results in the reference literature about the average ego network.

As previously said, the size of the network is limited by cognitive and time constraints. Since cognitive constraints are not easily quantifiable, our model focuses on time constraints, associating to each ego a certain time budget for handling his/her social relationships. In the model each relationship requires a specific amount of time, therefore the size of the ego network is constrained by the time budget.
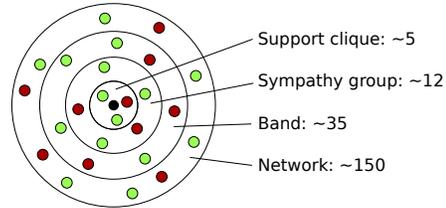

Figure 1. Hierarchical ego networks' structure. The black circle represents the ego; dark red circles refer to the kin; light green ones refer to non-kin.

In order to know the time requested by each relationship, the model exploits a function that, given the level of emotional closeness of a relationship, returns the related amount of time to handle it. The level of emotional closeness is distributed according to known probability distributions, and identifies the layer a relationship belongs to. Each layer of the ego network is related to specific interval of values of emotional closeness. The function that correlates emotional closeness and time is defined in order to obtain, on average, networks with a specific expected size.

The size of the sympathy group follows a specific distribution and it is independent of the network size. On the contrary it is linearly correlated to the support group according to a ratio defined by a probability distribution.

As previously said, the literature proposes different categorizations of relationships and alters. Our model only considers the kinship with the ego and the gender of the alters because there are many data available about these categories [8], [4]. Therefore, each relationship in the model is characterized by the type (kin or not-kin) and by the gender of the alter according to the composition of an average ego network. Our model simply considers static ego networks. Including the evolution over time, studied in [10], represents an interesting future work.

In the following subsection we present an algorithm for the generation of ego network graphs that are based on the presented model.

A. The Algorithm

The algorithm generates an ego network graph iteratively, following the proposed model. It adds relationships to the network from the inner to the outers layers, until the time budget is completely spent. To construct the ego graph, the algorithm exploits a set of functions (h_d, f_S, f_W, f_B, $f_{A,D}$ and f_E) and parameters (μ_l and m) whose values are obtained in Section IV from the analysis of an average net.

The first step is the creation of an ego and the assignment of its gender. The gender of the ego is saved in the variable g that can take values M (male) and F (female). The algorithm extracts g form a Bernoulli distribution $Ber(m)$ where m is the probability that $gen = $ M (line 2-3).

The next step is the extraction of the sympathy group size s_{sym} from the known probability density function f_S (line 4). The mean value of f_S is μ_{sym} that is the size we expect to obtain, on average, by the algorithm execution.

```
 1: procedure CREATEEGONETMODEL
 2:     g ← EXTRACTFROM(Ber (m))
 3:     ego ← CREATEEGO(g)
 4:     s_sym ← EXTRACTFROM(f_S)
 5:     w ← EXTRACTFROM(f_W)
 6:     s_sup ← s_sym · w
 7:     bdg ← EXTRACTFROM(f_B)
 8:     done =← False, tot ← 0, i ← 0
 9:     repeat
10:         l ← SELECTLAYER(i, s_sup, s_sym)
11:         a, d ← EXTRACTFROM(f_{A,D|L=l,G=g})
12:         e ← EXTRACTFROM(f_{E|D=d} in (low_{l,d}, up_{l,d}))
13:         t ← h_d (e)
14:         if t/2 < bdg − tot then
15:             r ← CREATERELATIONSHIP(l, a, d, e, t)
16:             ADDRELATIONSHIP(ego, r)
17:             tot ← tot + t
18:             i ← i + 1
19:         else
20:             done ← True
21:         end if
22:     until done
23:     return ego            ▷ s_net is the final value of i
24: end procedure
```

Figure 2.   The algorithm pseudocode

Once the algorithm knows the value of s_{sym}, it can obtain the size of the support clique s_{sup}. In order to do this the algorithm randomly extracts the ratio w between the two layers' sizes from the density function f_W. Once w is extracted, the algorithm sets $s_{sup} = s_{sym} \cdot w$ (lines 5-6). The expected value of s_{sup} is μ_{sup}.

Since the probability density functions used in the model return continuous values, but layer sizes have to be natural numbers, values are rounded using the dithering method [11]. Moreover each negative value is converted into a zero.

In the next step the algorithm assigns the time budget bdg. This amount is extracted from the known probability density function f_B (line 7).

At this point the main loop starts (lines 9-22). For each iteration the algorithm sets the parameters for a new relationship that is created only if there is enough time available. The total amount of time spent by the created relationships, is kept in the variable tot, that is updated after each relationship addition. The variable tot is initialized before the loop begins together with the control variable $done$ and the counter i, which represents the current size of the network (line 8).

By the knowledge of the current size i and the layer sizes s_{sup} and s_{sym}, the algorithm infers the current layer l. The variable l takes the values in the set $\mathcal{L}$: sup (support clique), sym (sympathy group) or net (active network) (line 10).

For each relationship, the algorithm has to set the type of the relationship d and the gender of the alter a. The variable

d takes the values K and NK, in case of kin and non-kin relationship respectively. The variable a, such as g, takes values M (male) or F (female). The algorithm randomly extracts the values of a and d from the joint probability mass functions $f_{A,D}$. Since each layer has a different composition, which also depends on the gender of the ego, there is a specific function $f_{A,D|L=\bar{l},G=\bar{g}}$ for each layer $\bar{l}$ and for each gender $\bar{g}$. The functions refer only to the composition of the external part of the layers. Considering the current layer l and the gender of the ego g, the algorithm extracts a and d from the function $f_{A,D|L=l,G=g}$ (line 11).

For each relationship, the algorithm has to assign a level of emotional closeness to the variable e using the probability density functions f_E. There are two different f_E functions, one to use in case of kin relationship $f_{E|D=K}$, and the other for non-kin relationship $f_{E|D=NK}$, therefore the algorithm selects the proper function according to d. The extraction from $f_{E|D=d}$ is limited in an interval of emotional closeness $(low_{l,d}, up_{l,d})$ related to the current layer l and the type of the current relationship d (line 12). The method to infer these intervals is described in the Subsection IV.G.

To relate the emotional closeness e to the time required to handle it, the algorithm is based on functions h_d that return an amount of time given a level of emotional closeness. There are two different functions h_K and h_{NK}, for kin and non-kin relationships respectively. Using the proper function h_d the algorithm sets the amount of time t given the level of emotional closeness e (line 13). Functions h_d must satisfy some properties listed in the Subsection IV.H.

The current relationship has to be added to the ego network only if there is enough time available. However if the algorithm discards a relationship when $t > bdg - tot$, the final value of tot is always less than the time budget bdg. Since we want that $\mathbf{E}[tot] = \mathbf{E}[bdg]$ the condition to add a relationship to the network is $t/2 < bdg - tot$ (lines 14-18). When the previous condition gets false, the boolean control variable $done$ becomes equal to True, therefore the loop ends and the algorithm returns the object ego with the related ego network (lines 19-23).

The final value of the counter i represents the network size s_{net}. If the functions and the parameters of the model are defined satisfying the properties given in the following subsections, the algorithm generates, on average, ego networks with the expected size μ_{net}.

IV. PARAMETERS AND FUNCTIONS

In this section we define all the parameters and functions the model uses exploiting results in the reference literature.

A. Layer sizes

In the literature there are different values for the layer sizes, often with significant differences. In [7], the authors collected all the required data about layer sizes and extracted the mean value for each layer. Therefore, basing on this

Table I
COMPOSITION OF SYMPATHY GROUP

a, d	$g = $ M		$g = $ F	
$a = $ M, $d = $ K	2.28	15.98%	2.38	16.64%
$a = $ F, $d = $ K	2.47	17.26%	3.53	24.72%
$a = $ M, $d = $ NK	7.38	51.61%	2.02	14.14%
$a = $ F, $d = $ NK	2.17	15.15%	6.36	44.51%
sum	14.3	100%	14.3	100%

Table II
COMPOSITION OF ACTIVE NETWORK LAYER (EXTERNAL PART)

a, d	$g = $ M		$g = $ F	
$a = $ M, $d = $ K	11.46	9.70%	17.35	14.68%
$a = $ F, $d = $ K	18.00	15.23%	17.18	14.53%
$a = $ M, $d = $ NK	52.50	44.41%	38.90	32.91%
$a = $ F, $d = $ NK	36.24	30.66%	44.78	37.88%
sum	118.2	100%	118.2	100%

work, we set the mean support clique size $\mu_{\mathrm{sup}} = 4.6$, the mean sympathy group size $\mu_{\mathrm{sym}} = 14.3$ and the mean active network size $\mu_{\mathrm{net}} = 132.5$.

B. Parameter m

Parameter m is the probability to have a male ego, that is $gen = $ M. We can reasonably assume that $m = 0.50$.

C. Function f_S

The sympathy group size distribution is presented in a histogram format ([8]) which can be fitted by a *Gamma* distribution. As f_S must be consistent with the mean size of the sympathy group μ_{sym}, we obtained $f_S = Gamma(4.1, 3.49)$ with mean 14.3.

D. Function f_W

The ratio between the support clique and the sympathy group sizes is given by the function f_W. Since we have set the mean sizes μ_{sup} and μ_{sym}, we define f_W thought a normal distribution with mean equal to $\mu_{\mathrm{sup}}/\mu_{\mathrm{sym}} = 0.3217$. We have no explicit information about the standard variation of the distribution, however it can be experimentally approximated, using the scatter plot proposed in [8]. A good approximation is obtained by setting the standard variation to half of the mean, therefore the function is defined as $f_W = Normal(0.3217, 0.1608)$.

E. Function f_B

We have no exact information about the distribution of time spent in socializing but we know that on average a person spends for it about the 20% of the time [12]. Therefore we define f_B with a mean value equal to $8760 \cdot 0.2 = 1752$ where 8760 is the number of hours in a year. In this way expected value of time budget is $\mathbf{E}[bdg] = 1752$.

The probability function f_B directly influences the distribution of the network sizes, therefore we chose its distribution and parameters experimentally, after we have done some tests, in order to obtain a network size distribution close to the one presented in [4]. The function we selected is $f_B = Gamma(205.48, 8.5264)$.

F. Functions $f_{A,D}$

Dunbar & Spoors in [8] studied the composition of the sympathy group for male and female egos. Considering the given mean size μ_{sym}, that is independent of the gender of the ego, the resulting compositions are reported in Table I.

In the same work, the authors studied the support clique and they observed that there are not significant differences between the compositions of the two layers. For this reason we can set the function $f_{A,D|L=\mathrm{sym}}$, that refers to the external part of the layer, with the values in Table I, related to the whole sympathy group. Moreover we can set $f_{A,D|L=\mathrm{sup}} = f_{A,D|L=\mathrm{sym}}$.

Regarding the external part of the active network layer we can indirectly estimate its composition starting from results in [6]. Specifically, we set $f_{A,D|L=\mathrm{net}}$ with the results presented in Table II.

G. Emotional closeness intervals and functions f_E

As shown in Figure 3 the value of emotional closeness will be extracted by different range of the f_E distribution depending on the layer. The intervals of emotional closeness can not be chosen arbitrarily but they must be consistent with the expected layer sizes (μ_{sup}, μ_{sym} and μ_{net}) and with the probability density functions f_E. The probability to extract a value of emotional closeness in an interval must be equal to the proportion of the network the related layer represents.

Our model uses two different density functions for kin $f_{E|D=\mathrm{K}}$ and non-kin $f_{E|D=\mathrm{NK}}$, therefore, in order to define the intervals, we need to know the mean proportion of kin for each layer. Using the Equation (1) we obtain the probability k'_l to have a kin in the external part of a layer l.

$$k'_l = \sum_{a \in \{\mathrm{M,F}\}} \left(m \cdot f_{A,D|L=l,G=\mathrm{M}}(a, \mathrm{K}) \right.$$
$$\left. + (1-m) \cdot f_{A,D|L=l,G=\mathrm{F}}(a, \mathrm{K}) \right) \qquad , \forall \, l \in \mathcal{L} \tag{1}$$

Using the values k'_l it is possible to obtain the probability to have a kin, k_l, in the whole layer l by the Equation (2), where c is a sublayer of l.

$$k_l = \sum_{c \subseteq l} \frac{\mu'_c}{\mu_l} \cdot k'_c \qquad , \forall \, l \in \mathcal{L} \tag{2}$$

For example, the probability to have a kin in the whole network, k_{net}, is:

$$k_{\mathrm{net}} = \frac{\mu'_{\mathrm{net}} \cdot k'_{\mathrm{net}} + \mu'_{\mathrm{sym}} \cdot k'_{\mathrm{sym}} + \mu'_{\mathrm{sup}} \cdot k'_{\mathrm{sup}}}{\mu_{\mathrm{net}}} \tag{3}$$

Considering a type of relationship d, the probability to extract a value from $f_{E|D=d}$ in the interval $\left(\mathtt{low}_{l,d}, \mathtt{up}_{l,d} \right)$ related to a layer l, must be equal to the expected proportion
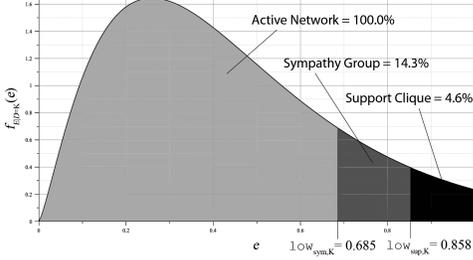
Figure 3. Distribution of emotional closeness for kin with the proportion of the network for each layer and the related limits.

of the network the layer l represents, considering only relationships with type d.

Knowing the cumulative distribution functions F_E of the densities f_E, it is possible to calculate the limits of the intervals of emotional closeness, considering them as quantiles that satisfy the following equations:

$$F_{E|D=K}\left(\texttt{low}_{\text{sup,K}}\right) = 1 - \frac{\mu_{\text{sup}} \cdot k_{\text{sup}}}{\mu_{\text{net}} \cdot k_{\text{net}}} \quad (4)$$

$$F_{E|D=NK}\left(\texttt{low}_{\text{sup,NK}}\right) = 1 - \frac{\mu_{\text{sup}} \cdot (1 - k_{\text{sup}})}{\mu_{\text{net}} \cdot (1 - k_{\text{net}})} \quad (5)$$

$$F_{E|D=K}\left(\texttt{low}_{\text{sym,K}}\right) = 1 - \frac{\mu_{\text{sym}} \cdot k_{\text{sym}}}{\mu_{\text{net}} \cdot k_{\text{net}}} \quad (6)$$

$$F_{E|D=NK}\left(\texttt{low}_{\text{sym,NK}}\right) = 1 - \frac{\mu_{\text{sym}} \cdot (1 - k_{\text{sym}})}{\mu_{\text{net}} \cdot (1 - k_{\text{net}})} \quad (7)$$

For example, considering kin relationships and the support clique layer, the limit $\texttt{low}_{\text{sup,K}}$ defines an area in f_E whose size is equal to $\frac{\mu_{\text{sup}} \cdot k_{\text{sup}}}{\mu_{\text{net}} \cdot k_{\text{net}}}$ (the dark area in Figure 3) where $\mu_{\text{sup}} \cdot k_{\text{sup}}$ is the number of kin relationships in the support clique and $\mu_{\text{net}} \cdot k_{\text{net}}$ is the number of kin relationships in the whole network.

The lower limits for the active network layer are $\texttt{low}_{\text{net},d} = 0$ while the upper limits are $\texttt{up}_{\text{sup},d} = e_{max}$, where e_{max} is the max value of emotional closeness, $\texttt{up}_{\text{sym},d} = \texttt{low}_{\text{sup},d}$ and $\texttt{up}_{\text{net},d} = \texttt{low}_{\text{sym},d}$, for each type of relationship d.

Distributions of emotional closeness for kin and non-kin are presented in [6]. As we do not have the exact distributions' values, we can only approximate them. Setting the maximum level of emotional closeness $e_{max} = 1$, obtained distributions are $f_{E|D=K} = Gamma(0.2, 2.296)$ and $f_{E|D=NK} = Normal(0.5, 0.172)$, both defined only in the interval $(0, e_{max})$. Considering the cumulative distributions F_E it is possible to solve the Equation (4), (5), (6) and (7), obtaining the limits of the intervals of emotional closeness: $\texttt{low}_{\text{sup,K}} = 0.8582$, $\texttt{low}_{\text{sup,NK}} = 0.8185$, $\texttt{low}_{\text{sym,K}} = 0.6852$ and $\texttt{low}_{\text{sym,NK}} = 0.7247$.

H. Functions h_d

h_d functions correlates the level of emotional closeness to the time spent in a relationship. Considering the studies [4] and [5] we know that h_d functions are increasing with the level of emotional closeness and that h_K returns lower

or equal values than h_{NK}. The latter observation is due to the fact that kin relationships demand less time invested on them than non-kin relationships. However, for high level of emotional closeness, the invested time in social relationships is equal for both kin and non-kin, therefore we set the following constraint:

$$h_{\text{K}}\left(e_{max}\right) = h_{\text{NK}}\left(e_{max}\right) \quad (8)$$

where e_{max} is the maximum level of emotional closeness.

Since the network size s_{net} is limited by time constraints, it is fundamental to properly define the functions h_d in order that $\mathbf{E}[s_{\text{net}}] = \mu_{\text{net}}$. In order to do this we impose that, in an average network with size μ_{net}, the total amount of time spent in relationships is equal to the main value of the time budget $\mathbf{E}[s_{bdg}]$, obtained from the density function f_B. Considering the given density functions of emotional closeness f_E and the proportion of kin in the network k_{net}, the constraint can be expressed by the Equation (9). In this equation, the value of the integral is the weighted sum of the expected values of the functions h_{K} and h_{NK}, multiplied for the probability to have a kin or a non-kin respectively.

$$\mu_{\text{net}} \cdot \int \left[h_{\text{K}}(e) \cdot f_{E|D=K}(e) \cdot k_{\text{net}} \right. \\ \left. + h_{\text{NK}}(e) \cdot f_{E|D=NK}(e) \cdot (1 - k_{\text{net}})\right] \mathrm{d}e = \mathbf{E}[bdg] \quad (9)$$

Through the graphics in [5] and in [4], we presume that h_d functions have an exponential trend therefore we define a generic h function: $h(e) = c^e + t_0 - 1$. The parameter t_0 is the value returned by $h(0)$. It can be considered as the minimum amount of time spent in a relationship in order to keep it active.

h_{K} and h_{NK} have the same form as h but have different values for the parameters c and t_0: respectively c_{K} and $t_{0\text{K}}$ in h_{K}, and c_{NK} and $t_{0\text{NK}}$ in h_{NK}.

As previously said h_K has to return lower or equal values than h_{NK} therefore $t_{0\text{K}}$ must be less or equal than $t_{0\text{NK}}$. We have no any indication on how estimate t_0 parameters, therefore we assume to be reasonable to set $t_{0\text{K}} = 0.5$ and $t_{0\text{K}} = 2$. In order to extract parameters c we can put in a system the Equation (8) and (9) where $\mu_{\text{net}} = 132.5$, $k_{\text{net}} = 0.2817$ and $\mathbf{E}[bdg] = 1752$.

With numeric methods we can solve the system of equations with a very good approximation obtaining $c_{\text{K}} = 95.3275$ and $c_{\text{NK}} = 93.8275$. Finally we can define the functions:

$$h_{\text{K}}(e) = 95.3275^e - 0.5 \quad (10)$$

$$h_{\text{NK}}(e) = 93.8275^e + 1 \quad (11)$$

V. RESULTS

We have implemented the algorithm presented in Section III.A in Java programming language and we performed 100.000 run tests creating as many ego network graphs. Results are presented in the Tables III and IV.

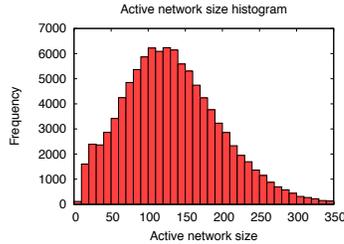|	min	max	avg	st. dev.
s_{net}	3	510	132.84	65.80
s_{sym}	0	74	14.06	7.25
s_{sup}	0	43	4.62	3.55
bdg	195.62	5197.87	1748.40	598.42

Figure 4.   Network Sizes Distribution in Simulations

As we can see, the average network size converges to a value close to the expected value 132.5. The little gap is due to approximation errors.

Also the mean average of the sympathy group is very close to the reference value 14.3. In this case the gap is due to the correlation between the time budget and the size of the layer. The algorithm extracts s_{sym} values from the distribution f_S but in a few cases the algorithm exhausts the time budget before completing to populate the sympathy group layer, making lower its mean size. This happens especially when the algorithm extracts a low value for bdg. In our tests, the sympathy group size is constrained by time budget in the 3.17% of the runs.

The average size of the support clique meets perfectively its expected value. Such as in case of the sympathy group, the time budget extracted can constrain the size of the layer however, in case of the support clique, this happened only in the 0.38% of the runs.

The shapes of the layer size distributions are similar to the distributions in the reference literature. See for example the shape of the network size distribution presented in Figure 4.

In Table IV we can see that the composition of the network is coherent with the $f_{A,D}$ functions we set. Male egos have smaller network than females. This is due to female egos have a little more kin relationships which request less time that non-kin relationships.

We have validated the model, demonstrating that it allows generating ego network graphs that are coherent with the results in the reference literature. In future work, the model may be extended in order to consider evolution over time and the connections among different ego networks.

ACKNOWLEDGMENT

a_i, d_i	$g = \mathrm{M}$ (49.85%)		$g = \mathrm{F}$ (50.15%)	
$a_i = \mathrm{M}, d_i = \mathrm{K}$	13.63	10.35%	19.97	14.89%
$a_i = \mathrm{F}, d_i = \mathrm{K}$	20.37	15.48%	20.93	15.61%
$a_i = \mathrm{M}, d_i = \mathrm{NK}$	59.44	45.17%	41.50	30.96%
$a_i = \mathrm{F}, d_i = \mathrm{NK}$	38.16	29.00%	51.68	38.54%
sum	131.59	100%	134.08	100%

of Oxford, which have been fundamental to deeply understand the structures of human social networks, and how to correctly model them.

REFERENCES

[1] M. Conti *et al.*, "Looking ahead in pervasive computing: challenges and opportunities in the era of cyber-physical convergence," *Pervasive and Mobile Computing Journal*, 2011, in press.

[2] T. V. Pollet, S. G. B. Roberts, and R. I. M. Dunbar, "Use of social network sites and instant messaging does not lead to increased social network size, or to emotionally closer relationships with offline network members," *Cyberpsychology, Behavior, and Social Networking*, 2010, in press.

[3] R. I. M. Dunbar, "The social brain hypothesis," *Evolutionary Anthropology: Issues, News, and Reviews*, vol. 6, no. 5, pp. 178–190, 1998.

[4] R. A. Hill and R. I. M. Dunbar, "Social network size in humans," *Human Nature*, vol. 14, no. 1, pp. 53–72, 2003.

[5] S. G. B. Roberts and R. I. M. Dunbar, "Communication in social networks: Effects of kinship, network size, and emotional closeness," *Personal Relationships*, 2010, in press.

[6] S. G. B. Roberts *et al.*, "Exploring variation in active network size: Constraints and ego characteristics," *Social Networks*, vol. 31, no. 2, pp. 138–146, May 2009.

[7] W. X. Zhou *et al.*, "Discrete hierarchical organization of social group sizes," *Proceedings of the Royal Society B: Biological Sciences*, vol. 272, no. 1561, pp. 439–444, 2005.

[8] R. I. M. Dunbar and M. Spoors, "Social networks, support cliques, and kinship," *Human Nature*, vol. 6, no. 3, pp. 273–290, Sep. 1995.

[9] J. Stiller and R. I. M. Dunbar, "Perspective-taking and memory capacity predict social network size," *Social Networks*, vol. 29, no. 1, pp. 93–104, Jan. 2007.

[10] R. S. Burt, "Decay functions," *Social Networks*, vol. 22, no. 1, pp. 1–28, May 2000.

[11] L. Schuchman, "Dither signals and their effect on quantization noise," *Communication Technology, IEEE Transactions on*, vol. 12, no. 4, pp. 162–165, Dec. 1964.

[12] R. I. M. Dunbar, "Theory of mind and the evolution of language," in *Approaches to the Evolution of Language: Social and Cognitive Bases*, J. R. Hurford, M. Studdert-Kennedy, and C. Knight, Eds. Cambridge: Cambridge University Press, 1998.