**Ego-Centric Networks:**

**From: Networks and the epidemiology of infectious diseases (Danon et al 2011)**

* [10.1155/2011/284909](https://dx.doi.org/10.1155/2011/284909)
* NB: May be useful for references when thinking about simulating transmission on the Tanzania market network

Given the huge logistical difficulties of capturing the full network of interactions between individuals within a population, a variety of methods have been developed to generate synthetic networks from known attributes. Generally, such methods fall into two classes: those that utilise egocentric information and those that attempt to simulate the behaviour of individuals.

Egocentric data generally consists of information on a number of individuals (the egos) and their contacts (the alters). As such the information gathered is very similar to that collected in the sexual contact network studies in Manitoba and Colorado Springs, but with only the initial step of the snowball sampling was performed; the difference is that for the majority of egocentric data the identity of partners (alters) is unknown and therefore connections between egos cannot be inferred (Figure [1(c)](https://www.hindawi.com/journals/ipid/2011/284909/fig1/#c)). The data, therefore exists as multiple independent “stars” linking the egos to the alters, which in itself provides valuable information on heterogeneities within the network. Two major studies have attempted to gather such egocentric information: the NATSAL studies of sexual contacts in the UK [13–16], and the POLYMOD study of social interactions within 8 European countries [17]. The key to generating a network from such data is to probabilistically assign each alter a set of contacts drawn from the information available from egos; in essence, using the ego data to perform the next step in the snowball sampling algorithm. The simplest way to do this is to generate multiple copies of all the egos and to consider the contacts from each ego to be “half-links”; the half-links within the network can then be connected at random generating a configuration network [18–20]; if more information is available on the status (age, gender, etc.) of the egos and alters then this can also be included and will reduce the set of half-links that can be joined together. However, in the vast majority of modelling studies, the egocentric data have simply been used to construct WAIFW (who-acquires-infection-from-whom) matrices [15, 17, 21] that inform about the relative levels of transmission between different groups (e.g., based on sexual activity or age) but neglect the implicit network properties. This matrix-based approach is often reliable: for STIs it is the extreme heterogeneity in the number of contacts (which are close to being power-law or scale-free distributed; see Section [3.2](https://www.hindawi.com/journals/ipid/2011/284909/#sec3.2)) that drives the infection dynamics [22] although larger-scale structure does play a role [23]; for social interactions, it is the assortativity between (age-) groups that controls the behaviour, with the number of contacts being distributed as a negative binomial [17]. The POLYMOD matrices have therefore been extensively used in the study of the H1N1 pandemic in 2009, providing important information about the cost-effective vaccination of different age-classes [21, 24].

*References:*

* *POLYMOD - 17,*
* *See refs 18-20 for the general configuration model approach (half-links, above)*
  + *18, 19: examples from spread of STI on sexual networks*
  + *See also 27 for a more general approach, and 23 for application of the model to defining high-risk groups for intervention.*
* *WAIFW H1N1 - 21, 24*

Erdos-Renyi