Simulation setup:

1. Generate (static) network with specified structure (density and homophily)
2. Simulate epidemic spread over network – yields times of transitions and genetic distances
3. Sample egos from network to get “behavioral survey”
4. Define priors based on genetics and survey
5. Fit Bayesian model to epidemic data to get estimates of network properties (density and homophily)

Things that we could vary:

* ~~Network size~~
* Density
* ~~Distribution of attribute in population (fixed to data)~~
* Strength of homophily
* ~~Epidemic parameters – transmissibility, times in compartments (maybe in 2~~~~nd~~ ~~paper – informing design or when to use this approach)~~
* Sampling coverage for behavioral survey: 2 or 3 levels?
* Biased sampling for behavioral survey
* ~~Add noise to transition times to reduce accuracy of epidemic data~~
* ~~Missing transition times~~
* ~~Sampling of genetic data?~~

Genetic data to prior: inverse distance or inverse squared distance? Question for Vlad. Also could depend on how we are simulating “genetic sequences”

To make an interesting first paper, I think implementing some form of biased sampling for the behavioral survey will be important – we need to be doing something that couldn’t be achieved directly from the behavioral survey itself. Evaluating the contribution of the different data sources is another interesting question. I’m thinking we evaluate that by varying the survey coverage, adding noise to dates, and changing the variance of the prior induced by the genetics. I don’t really want to get into sampling the genetic network and inferring distances, etc.

For unbiased sampling, vary num\_samples and set strong\_prior to FALSE.

For biased sampling, set num\_samples = N, strong\_prior to FALSE. Then replace the generated Prob\_Distr\_Params with an appropriate biased sample (avoids having to reprogram CCMnet).

Add output of degree distribution, or measure of distance of distribution from truth. Don’t save off all graphs.