XSwap analysis for the Hetmech project

Michael Zietz

October 2018

1 Zietz Markov Chain Model

Let $q$ be the probability that one XSwap iteration creates an edge where there was previously none. Let $p$ be the probability that one XSwap iteration destroys an edge where previously an edge existed. These probabilities are both conditional on the existence or non-existence of an edge previously.

This gives the following state diagram, in which 0 represents a state with no edge, while 1 represents the state in which there is an edge between the two nodes:

![State Diagram](image)

This diagram corresponds to a transition matrix $P$. With "no edge" represented as $[1, 0]^T$ and "edge" represented as $[0, 1]^T$, $P$ is given by,

$$P^T = \begin{bmatrix} 1 - q & r \\ q & 1 - r \end{bmatrix} \tag{1}$$

The stationary distribution of this system should correspond to the distribution when the number of swaps goes to infinity. It can be found by computing the eigenvectors of the system, as we know that the stationary distribution vector, $v$ satisfies

$$P^T v = v$$

The normalized eigenvector $v$ is given by

$$v = \frac{1}{\sqrt{1 + r^2/q^2}} \begin{bmatrix} r/q \\ 1 \end{bmatrix}$$

Now we use $q$ and $r$ as derived by Kyle Kloster, that is,
\[ q = \frac{d_u d_v}{S} \]
\[ r = \frac{m - d_u - d_v + 1}{S} \]

Now let \( k = r/q \). Substituting this value, we get the stationary distribution of XSwap for a single node pair as
\[ k = \frac{m - d_u - d_v + 1}{d_u d_v} \]
\[ v = \frac{1}{\sqrt{1 + k^2}} \begin{bmatrix} k \\ 1 \end{bmatrix} \]

Thus the probability of an edge existing after XSwap has reached its stationary distribution is
\[ P_{u,v} = \frac{1}{\sqrt{1+k^2}} = \frac{d_u d_v}{\sqrt{(d_u d_v)^2 + (m - d_u - d_v + 1)^2}} \]

2 Kloster Model

For a given metaedge, let \( s = \{s_i\} \) be the set of source nodes and \( t = \{t_i\} \) be the set of target nodes. Denote the degree of node \( u \) as \( d_u \). For any \((s_i, t_j) \in s \times t\), the Cartesian product of \( s \) and \( t \), let \( P_{ij} \) be the number of edges observed over \( N \) permutations. Let \( m \) by the number of edges along the metaedge. Since XSwap is degree-preserving, \( m \) is constant.

"[T]he observed probability of an edge appearing is," \( P_{ij}/(N \times m) \). The analytic probability of an edge appearing between nodes \( u \in s \) and \( v \in t \) is
\[ P(u, v) = \left( \frac{1}{2} \right) \binom{m}{2} \left( \binom{d_u}{2} \right) - \sum_{i \in s \cup t} \binom{d_i}{2} - m + d_u + d_v - d_u d_v - 1 \]

For easier comparison with the other methods, we can scale this probability by a factor of \( m \), giving
\[ P(u, v) = \left( \frac{m}{2} \right) \binom{m}{2} \left( \binom{d_u}{2} \right) - \sum_{i \in s \cup t} \binom{d_i}{2} - m + d_u + d_v + d_u d_v - 1 \]

This can be directly compared to the Zietz/Kloster Markov model described earlier, and can be more easily plotted as the range should be roughly from zero to one. As y-values now, we will use \( P_{ij}/N \).
Note that in the above figure, we are not concerned by the fact that some "probabilities" rise above 1. This is because we have re-scaled probabilities to more precisely represent the expected number of edges allocated to a single node pair.