We study the effect of social network structures on alcoholism and alcohol related problems. Our goal is to reveal the effect of network topology on the spread and persistence of alcohol dependence, and the policy conclusions that can be drawn from understanding this effect. We start by studying some very basic models of the evolution of alcohol dependence on random and small-world networks. We simulate the evolution of different initial conditions on a set of social network realizations and examine the resulting mean and variance of the initial fraction of alcohol dependent people. After understanding the evolution of a basic epidemiological model, we study the effect of treatment regimes.
1 Problem Description
The objective of this problem is to study the effect of social network structures on alcoholism and alcohol related problems. We want to develop epidemiological models [1] of alcohol dependence on models of social networks [2-5]. Our goal is to reveal the effect of network topology on the spread and persistence of alcohol dependence, and the policy conclusions that can be drawn from understanding this effect. We start by studying some very basic models of the evolution of alcohol dependence on random and small-world networks.

We simulate the evolution of different initial conditions on a set of social network realizations and examine the resulting mean and variance of the initial fraction of alcohol dependent people. After understanding the evolution of a basic epidemiological model, we study the effect of treatment regimes.

We also make use of data that are available from random surveys of populations regarding alcohol dependence. This information is relatively reliable compared to many types of hard drug abuse. The data are available in relatively small geographic areas such as zip codes. We discuss how these data can be used to evaluate the ecology of the population and to help make effective policy decisions.

2 Model
The modeling problem is twofold. First, we decide on the type of the network that best resembles the social interactions among individuals. Second, we formulate a mathematical model that captures the time-evolution of the alcoholism on the society with the given network structure. As for the network type, we pick a small-world network, which possesses both randomness and clustering inherently. We consider an unweighted and an undirected graph, where each vertex corresponds to an individual, and edge corresponds to a relationship/influence between connected individuals. That is to say that there is an edge between two vertices if and only if the corresponding persons have some relation between each other. Such network can be represented by an adjacency matrix $A$ whose entry $A_{ij} = 1$ if and only if person $i$ and person $j$ know each other, and zero otherwise.

As for the state variable we choose $u$ to be the probability of an individual to involve in an alcohol related problem. A logistic-like model governs the dynamics of $u$, which also capture the spread or persistence of alcoholism in the society:

$$
\dot{u}_i = ku_i (1-u_i)(r \deg(i) - A_{i,:} u_i)
$$

where $k$ is gain, $r$ is a threshold value which measures individual's strength/resistance to the effect of others, $\deg(i)$ denotes the degree of node $i$, i.e number of people that individual $i$ has relation with, and $A_{i,:}$ is the $i$th column of the adjacency matrix $A$. 

To get some idea of the behavior of our model in (1) we pick a 4-node network with the given topology and corresponding adjacency matrix:

\[
A = \begin{bmatrix}
0 & 1 & 1 & 1 \\
1 & 0 & 1 & 0 \\
1 & 1 & 0 & 1 \\
1 & 0 & 1 & 0
\end{bmatrix}
\]

For the above network topology, the model in (1) becomes:

\[
\begin{align*}
\dot{u}_1 &= ku_1(1-u_1)(3r-u_2-u_3-u_4) \\
\dot{u}_2 &= ku_2(1-u_2)(2r-u_1-u_3) \\
\dot{u}_3 &= ku_3(1-u_3)(3r-u_1-u_2-u_4) \\
\dot{u}_4 &= ku_4(1-u_4)(2r-u_1-u_3)
\end{align*}
\]

This model gives at least 2 expected steady states at \(u=0\) and \(u=1\). Another trivial steady state is \(u=r\). In \(\mathbb{R}^4\), this steady state lies on the segment connecting the origin \((0; 0; 0; 0)\) and the point \((1; 1; 1; 1)\). If the initial data is on this segment and close to the origin, \(u\) will tend to the \(u=0\) steady state. If the initial data is close to the other end of the segment, \(u\) will tend to the \(u=1\) steady state. In Section 3, we consider a different threshold for each vertex. Simulation results for this simple network with uniformly distributed \(u(0)\), and with the model parameters \(k=1\), and \(r=0.3, 0.7,\) and \(0.5\) are given in Figure 1 below.

**Figure 1.** Trajectories of the 4-node network. For \(r=0.7\) (upper left), all nodes converge to \(u \to 0\), for \(r=0.3\) (upper right), they all go to \(u \to 1\); and when \(r=0.5\), depending on the initial conditions, all nodes go to \(u \to 0\) or \(u \to 1\) with equal probability.
3 Simulations of the Model

We next simulate our model in large scale networks of different interconnection structures: Connected-Caveman, Small World and Random. For example, for 100 individuals, the corresponding network topologies are:

![Network Topologies](image)

**Figure 2.** Three network topologies each with 100 nodes. Upper left: Connected-Caveman, Upper right: Random, Bottom: Small-World

Based on the different structures, we get different mean value of $u$ after enough experiments.

First we keep the threshold $r$ as a fixed number chosen from the interval $[0 \ 1]$. Then in order to make the model more realistic, we pick $r$ to be a normally distributed random vector with mean=0.5 and variance=0.1, and then truncate the distribution at 0 and 1, to have $r$ in the interval $[0 \ 1]$.

3.1 Experiments with different values of $r$

In order to get the probability that $u$ goes to 1 in different network with different threshold value $r$, we use Quasi Monte-Carlo Method. For each network, we did 2000 experiments for different values of $r$ and uniformly distributed initial conditions of $u$ ($u(0)$ in $[0 \ 1]$), and get the result as follows:
Figure 3. r=0.5 Quasi Monte-Carlo simulations for the probability of $u \rightarrow 1$
Blue: random network; Green: Cavemen network, Red: Small-world network

Figure 4. 0.5<r<1 Quasi Monte-Carlo simulations for the probability of $u \rightarrow 1$
Blue: random network; Green: Cavemen network, Red: Small-world network
From the statistical results above, we see that when \( r=0.5 \), the probability of \( u \) converging to 1 is about 0.5; when 0<\( r<0.5 \), the probability of \( u \) becoming 1 is about 1; and when 0.5<\( r<1 \), this probability is about 0. This result makes sense since changing \( r \) changes the direction of the response of that vertex to the value of its neighbors.

Also, the more connections in the network, the closer the probability of \( u \) becoming 1 is to 0 or 1. From the graphs above we can see that after the same number of experiments, the result of random network case keep closest to 1 or 0 when \( r<0.5 \) or \( r>0.5 \), since there are stronger influences between individuals than those in the other two networks.

### 3.2 Modify ‘\( r' \) to a vector of distribution based on the survey data

To get some idea about the distribution of “\( r' \)” may look like, we employ the survey data, and analyze the distribution of level of education. As seen in Figure 6, the level of education has a skewed normal distribution, and therefore we choose the threshold ‘\( r' \)’ with a normal distribution. This certainly is not the best choice, or education level is not the only factor in individuals’ resistance to the effect of others, but it is a reasonable identifier to get some idea of the distribution of ‘\( r' \).
Figure 6. Histogram of the education level of the individuals surveyed.

To simulate a more reasonable situation in the real world, we make $r$ as a variable for each individual and make it obey normal distribution $N(0.5, 0.1)$:

Figure 7. Distribution of the threshold $r$ picked for simulations

Then, after 2000 experiments, we get the following result:
Figure 8. r: unif. dist., mean=0.5, var=0.1 Quasi Monte-Carlo simulations for the probability of $u \rightarrow 1$. Red: random network; Green: Cavemen network, Blue Small-world network

We see that the probability that $u$ goes to 1 is around 0.5, which is the mean value of the normal distribution $N(0.5,0.1)$.

4 Treatment

After the modeling and simulation, we studied ways to treat individuals so that the society reaches a clean state, that is $u=0$. When certain individual has a high probability of causing a problem, i.e. $u$ is close to 1; we institutionalize that individual and treat for some specified time. In the network graph this corresponds to removing the vertex with high $u$, and all its edges, so that that individual has no effect on the network. We then insert the individual to the society with a lower probability $u$. That is, the vertex is inserted back into network with its entire previous links, but with a smaller $u$ value.

For simulation purposes, we pick a 6 cluster network with 6 nodes each. The initial topology of the network is given in Figure 9 below.

Figure 9. 36-node 6-cluster network. Initial topology in treatment simulations.
In our model, we pick the threshold $r$ as a normally distributed random variable with mean 0.5 and variance 0.1, truncated in the interval $[0, 1]$, gain $k=1$, and the initial condition for $u$ is a uniformly distributed random variable in the interval $[0, 1]$. When the network is let to evolve according to the model with no treatment the trajectories are given in Figure 10. Most of the nodes go to $u=1$, which means the alcohol problem is spreading throughout the network due to the effect of highly problematic, i.e. $u_i \sim 1$, nodes. We note that our equations are dimensionless; however, to interpret the results we assume time $t$ is in months.

![Figure 10. No treatment](image)

We then apply our treatment strategy to the network with same topology, $r$, and initial $u$. We set our treatment threshold to be $u=0.8$, treatment time $T=3$ months, and recovery value $u_{i_{\text{new}}}=0.4$, that is all individuals with $u_i>0.8$ will be institutionalized, and treated for 3 months, and released with alcohol related problem probability $u_{i_{\text{new}}}=0.4$. The results are given in Figure 11. We see that almost all individuals in this case converge to $u=0$. We note, however, the two individuals with black trajectories have a very strong coupling. When one of them is isolated the other one starts to decrease, but once that removed one is released back it immediately pulls the other one up; but, eventually they both converge to zero.
Lastly, we make another simulation with a longer treatment time $T=6$ months, and see in Figure 12 that convergence of those two individuals to zero is achieved faster.

Figure 12. Longer treatment with parameters $T=6$ months, $u$ threshold=0.8, $u^{\text{new}}=0.4$

5 Conclusions

In this work, we formulated a mathematical model for the spreading of alcohol related problems in the population. We modeled the society and the interactions within the members by both small-world and random networks, with unweighted and undirected interconnection. We also made use of the survey data, and related level education information to our model parameter ‘$r$’. Finally, a basic treatment strategy is proposed.
The cost of the treatment is an important factor in decision/policy-making, and in our formulation it depends on threshold value for treatment, time of treatment, and returning value. To reduce the cost, we simulated our model with an improved criterion, which basically sets two treatment threshold values \( u_1 \) and \( u_2 \), \( u_1 > u_2 \), such that if \( u \) satisfies \( u_2 < u < u_1 \), but has only a few connections, so that it does not have much effect on the society, we simply keep it on the network and watch its evolution. However, if it gets \( u > u_1 \), then irrespective of the number of its connections we remove that individual from the society for treatment. This approach also results in successful treatment with a lower cost.

Another task that we want to pursue with the given data set was to develop a geographic network of zip codes. The data set contains information about the adjacency of zip code areas as well as other demographic information about the zip code population. Information regarding alcohol consumption as well as information regarding hospital discharges for alcohol related admissions and alcohol related crimes per zip code are also available. This information could form the basis for a geographic network of the zip codes.

One interesting piece of information is the correlation between the numbers of people who have moved in a zip code area with the amount of drinking. Using this information along with other information in the survey data, it might be possible to examine the flow of heavy drinkers. Do they stay within their own zip code, or do they change zip code areas?

For future work, we believe there is valuable information to be gained in examining a network based on the survey data that would allow focused treatment as well as examining the dynamics and interrelationships within the zip codes.

Once developed, a generic algorithm or set of algorithms could then be applied to other areas around the country or world to make similar predications about alcohol and other socially related behaviour.

References


