CX 4230 Project 3B: Literature Review

Karan Shah and Kartikeya Kumar
For Dr.Vuduc, Spring 2016

For the Literature Review, the papers we read can be classified into 3 broad categories: Traditional sociological methods, models based on graph theory and agent based modeling and models derived from physical theories (mostly statistical mechanics).

Ravenstein's theory of Migration forms the basis of modern migration theory. First published in 1885 in the Journal of The Statistical Society, it established some basic laws of migration based on statistics of migration in England and Scotland. These laws have been revised and you can find the original paper[1] and the revised laws[2] in the references. Ravenstein based his "laws" on insightful, careful scrutiny of census tables.[2]

Almost a 100 years after Ravenstein's laws of migration were published, two scientists set out to create a mathematical framework of the laws[3]. Treating the whole world as a grid, the authors used "Pushing" and "Pulling" factors to model migration. Pushing factors push the population from their original location while pulling factor entices it to a new location. A distance decay factor makes migrating to distant places more costly. We found this study to be very useful for our model. This study assumes constant population and there are no births nor deaths.

We also read a survey of migration models that walked through some famous traditional and historical models of migration[4]. 2 models stuck out to us as easier to understand and intuitive. The 1st one was called the Human Capital Model where a person looks at the expected benefits accruing to him/her over a specified time interval in a certain area discounted at present value net the cost of moving to that area and living there over the same time interval. Whichever area has the highest net utility, the human migrates there. This model incorporates both a spatial dimension in the form of a cost function (cost of living in an area) as well as a temporal dimension in the form of a time interval, however, there is a fundamental assumption that all information about every area is available to every human being, which is not true of real life. To improve upon this, a Random Utility Model is proposed which introduces a stochastic term based on a Weibull Probability Distribution to simulate individual idiosyncrasies. So, here, net utility of an area is calculated as the sum of the non-random utility term (like we found through the Human Capital Model) added to this stochastic/random term. This random term tries to account for individual randomness in migration.

Moving on, a very interesting paper that we read tried to simulate migration by adopting a network model based approach where each country was represented as a node[5]. Accordingly, a directed graph representing all 226 countries was made, based on a comprehensive dataset of immigration all across the world. We found the network based approach to be very good for simulating migration on a macroscopic global scale. The paper used an approach called ”maximization of communities”, which means that neighboring countries are grouped into communities and treated as a single unit. These Communities are decided by finding the most optimized state (highest "community" score) using Monte Carlo methods. This is helpful in reducing computational workload and finding the relationship between different countries which wouldn't be apparent by looking at countries individually. Over time steps, as migration numbers change, these clusters break and new clusters are formed, allowing for very interesting patterns to be noticed. While we can use these techniques to see migration patterns on a large scale, the paper uses real data to study patterns but does not predict future events. So it is more of an analysis of the current state of the world than a proper simulation.

After the network model based approach, we tried to look for a more dynamic modeling technique and read a paper on Agent Based Modeling (ABM)[6]. Since migration is essentially a complex and heterogeneous human behavior at an individual level and cannot be deterministically described by differential equations, treating each human being as an agent with a set of rules governing their interactions with other agents would make sense. ABM, since it is a bottom up model, would also be great at capturing any aggregate emergent phenomena. However, as the paper itself says, treating each human or even subsets of them as individual agents and then simulating their interactions ground up can be computationally intensive. Apart from that, to decide the rules that govern an agent based modeling system itself is more of an 'art than a science'.

The wave of advance of Advantageous genes[7]: This 1937 paper by Dr. R. A. Fisher, one of the founders of the field of population genetics, was the first to establish the concept of the "wave of advance" which has been used by many other researchers below. It uses the wave equation to show the spread of advantageous genes across a population. Fisher was also a great statistician and he used statistics to model gene spread in this paper too. This gene spread can be used to model migration over large time scales.

The Measuring the Rate of Spread of Early Farming in Europe paper[8] This looks at the spread of agriculture in ancient Europe and uses diffusion to model it and fit it with the archeological evidence over a span of 10000 years. The authors test different variations in the diffusion constant (variable with time only, space only and other combinations). The authors make their model around historical data and come up with an average rate of spreading. The authors of this paper have been cited in multiple papers below.

The Synergetics book[9] attempts to quantify social science using systems of differential equations (like the Romeo Juliet lab we did in class). Chapter 1 is not directly relevant but it is a fun read about analogies in the sciences. For simulating migration, the authors use coupled differential equations (more complicated than the simple LV model) and make the population move. They also add stochasticity and chaos to the mix. They show use cases for both macroscopic (countries) and microscopic (urban dynamics) systems. This approach seems versatile, but we also don't understand large parts of this approach at this point.

The two papers[10,11] by Kehner on using Statistical Mechanics to model populations lay a comprehensive mathematical foundation for simulating population movement. He starts with Lotka Volterra mechanics and adds a lot of statistical mechanics techniques. Then he defines analogs of state variables like Gibbs energy, temperature and pressure for the populations. He then used thermodynamics laws like the gas equation and boltzmann distribution law on the system. In the conclusion of the first paper, he also states that we could use conduction to model migration and heat engines(like Carnot's) to model periodic migration. We found this very intriguing. The advantage of his approach is that he uses properties of each member of the population to predict the macroscopic behavior of the population. The disadvantage is that his approach is very math heavy and some of his statements require the knowledge of Hamiltonian mechanics (which none of the project members have).

The econophysics model paper[12] proposes a relatively straightforward model. The authors apply an ”economic” potential field to the entire region under study and treating population as charge, which moves under the influence of an inverse square force from economic attractors and repellers. This is a simpler model but it breaks down because people in real world are not binary and don't migrate for a single reason (say economics).

The random walk model[13] does a comparative study of different random walk models. The author treats each human as an actor and simulates different types of random walks on a lattice. He starts with an unweighted random walk and then moves on to Lotka Volterra models, then adds the concept of utility and radially dependent utility functions. Using all these models, he studies urban models and confirms the linear relationship between the spatial size and the population of different cities. The author does not take into account the development in transportation technologies that allows people to move more than one step in space in a time step.

Hagerstrand's Monte Carlo diffusion model[14]: The author studies a small system of farms in Sweden and shows the diffusion pattern that emerges. The farms are given a choice to accept a new government plan. At first, onlu a small number of farms (that form a central core) accept that plan. Then the farm owners communicate using pairwise telephone calls and other farmers adopt the plan as they are convinced. When the simulation is run, the adoption of the new plan spreads in a pattern that looks like diffusion. It is a well written paper and has many quotable quotes too. My favorite is "*The dice is the motive power of life in the model*", which is a poetic way to describe Monte Carlo methods (and life in general if you overthink it).

Apart from these papers, we also read some other papers which we felt were not as useful for modeling our system. These papers can be found in the "`NotInTheLiteratureReview`" folder.

References:

[1] Ravenstein, E. G.. 1885. "The Laws of Migration". Journal of the Statistical Society of London 48 (2). [Royal Statistical Society, Wiley]: 167–235. doi:10.2307/2979181.

[2] Anon. Wikipedia. Retrieved April 5, 2016 from https://en.wikipedia.org/wiki/ernst˙georg˙ravenstein?oldformat=true

[3] Guido Dorigo and Waldo Tobler. 1983. Push-Pull Migration Laws. Annals of the Association of American Geographers 73, 1 (1983), 1–17.

[4] Molho, I. (1986), THEORIES OF MIGRATION: A REVIEW. Scottish Journal of Political Economy, 33: 396–419. doi: 10.1111/j.1467-9485.1986.tb00901.x

[5] Kyle F. Davis, Paolo D'odorico, Francesco Laio, and Luca Ridolfi. 2013. Global Spatio-Temporal Patterns in Human Migration: A Complex Network Perspective. PLoS ONE 8, 1 (2013).

[6] E. Bonabeau. 2002. Agent-based modeling: Methods and techniques for simulating human systems. Proceedings of the National Academy of Sciences 99, Supplement 3 (2002), 7280–7287.

[7] R.A. Fisher. 1937. The Wave Of Advance Of Advantageous Genes. Annals of Eugenics 7, 4 (1937), 355–369.

[8] A.J. Ammerman and L.L. Cavalli-Sforza. 1971. Measuring the Rate of Spread of Early Farming in Europe. Man 6, 4 (1971), 674.

[9] Wolfgang Weidlich. 1991. Physics and social science — The approach of synergetics. Physics Reports 204, 1 (1991), 1–163.

[10] Edward H. Kerner. 1957. A statistical mechanics of interacting biological species. The Bulletin of Mathematical Biophysics Bulletin of Mathematical Biophysics 19, 2 (1957), 121–146.

[11] Edward H. Kerner. 1959. Further considerations on the statistical mechanics of biological associations. The Bulletin of Mathematical Biophysics Bulletin of Mathematical Biophysics 21, 2 (1959), 217–255.

[12] Anca Gheorghiu and I. Spânulescu, An Econophysics Model for the Migration
Phenomena, arXiv: 12020996 (2012).

[13] David A. Young. 2002. A New Space-Time Computer Simulation Method for Human Migration. American Anthropologist 104, 1 (2002), 138–158.

[14] Torsten Hägerstrand. 1965. A Monte Carlo Approach to Diffusion. Eur J Soc European Journal of Sociology 6, 01 (1965), 43.