Social Networks and Mathematical Models: A Research Commentary on "Critical Mass and Willingness to Pay for Social Networks" by J. Christopher Westland

Andrew Odlyzko

School of Mathematics
University of Minnesota
odlyzko@umn.edu
http://www.dtc.umn.edu/~odlyzko

Abstract. This brief research commentary on Westland's (2010) article in this issue of *Electronic Commerce Research and Applications* is intended to add two cautionary notes. He attempts to answer two questions: What is critical mass? And how one can manage a network's growth to reach that mass? He proposes a model for social networks, and shows how the model might possibly be managed. What needs to be remembered is that models are always imperfect. And even when they do describe well the natural behavior of some system, they may not fit as well, once some external influence begins affecting that system. Hence Westland's contribution should be regarded as a definite step forward, but not necessarily the definitive answer to either of the questions it attempts to solve.

Keywords: percolation theory, research commentary, social networks

1 Introduction

Social networks have always been important for human life and commerce. In recent years, they have become much more prominent because of the rapid growth and huge size of systems such as MySpace, Facebook, and Twitter. These have excited speculation about reshaping of society in general, and potentially producing huge profits in particular.

A major reason for the increased attention paid by researchers to the recent social networks such as Facebook is that they are far more measurable than earlier ones. All interactions take place online, and therefore, at least in principle, can be monitored and quantified. However, at the moment we have far too little both of theory and of empirical evidence about social networks.

Empirical evidence is sparse. This is largely because the entities that do possess it, namely commercial social networks, either do not collect detailed enough data, or, more frequently, are extremely reluctant to share it for combinations of proprietary and privacy reasons.

Theoretical models are also lacking. Curiously enough, the two deficiencies, in data and theory, feed on each other. Lack of data makes it harder to devise or validate theoretical models. And lack of models makes it harder to decide what data is relevant.

Most observers agree that the concept of critical mass is indeed critical for understanding growth of social networks. Defining critical mass is harder, though, and harder still is finding ways to influence its arrival. Westland's contribution is to propose a model for social networks that offers potentially very valuable insights into this issue.

The purpose of this commentary is add two cautionary notes to Westland's work. His paper attempts to answer two questions, namely what is critical mass, and how one can manage a network's growth to reach that mass. What it really does is propose a model for social networks, and shows how the model might possibly be managed. What needs to be remembered is that models are always imperfect. And even when they do describe well the natural behavior of some system, they may not fit as well once some external influence begins affecting that system. Hence Westland's contribution should be regarded as a definite step forward, but not necessarily the definitive answer to either of the questions it claims to solve.

2 All models are wrong, but some are useful

Anyone aiming to engage in modeling should keep in mind the Box–Draper dictum (Box and Draper 1987, p. 424):

Essentially, all models are wrong, but some are useful.

Those who suffered from the crash of 2008 can place much of the blame on this dictum being ignored. There was far too much slavish adherence to financial models. Those models were useful for a long time, as they did provide guidance under normal conditions. But they did not envisage practically all financial valuations getting out of whack, and the result was a catastrophic collapse.

As Box and Draper observe, (Box and Draper 1987, p. 14), in common with other experts in modeling,

A mechanistic model has the following advantages:

- 1. It contributes to our scientific understanding of the phenomenon under study.
- 2. It usually provides a better basis for extrapolation (at least to conditions worthy of further experimental investigation if not through the entire range of all input variables).
- 3. It tends to be parsimonious (i.e, frugal) in the use of parameters and to provide better estimates of the response.

Models, whether explicit ones that can be rigorously specified, or vague mental ones, is what people always use. Getting a detailed record of all the interactions on Facebook, some petabytes of it, would be of no use. One needs to distill that information, and fit it into some model that is appropriate for the purpose at hand.

Briscoe, Tilly, and I proposed (Briscoe, Odlyzko, and Tilly 2006, Odlyzko and Tilly) that networks that allow interactions among users, such as most social networks, as well as communication networks, are likely to have values that grow like $n \log(n)$. (And so it is naturally gratifying to see that Westland's work leads to the same formula!) The goal in that research was to achieve maximum parsimony, by coming up with as simple a formula as possible that appeared to reflect observed behavior as well as some theoretical considerations. However, it was understood and emphasized that this was not to be regarded as a rigid rule, but a convenient rule of thumb, a first approximation. One of the deficiencies of that rule is that it was static, and did not provide any information on how growth rates might be affected. (And it was recognized that the rule concealed a variety of more complex phenomena taking place inside a network, such as those discussed in (Kilkki and Kalervo 2004).)

Westland proposes a more sophisticated model, one that draws on the vast literature of percolation theory. Since that theory has been extensively developed, one can draw on the results from it without having to develop the full machinery from scratch, thus again (as in applying models to complex systems) suppressing much of the complexity. That theory allows Westland to provide a more dynamic model of a social network. However, it should be remembered that this is still just a model. In particular, the basic assumptions of Westland's percolation-based model do not fit social networks very well, as in the assumption there are no cycles in the graph of relations. This does not mean that the model should be disregarded, but its limitations should be kept in mind.

A good recent example of the power and limitations of models is provided by the paper (Willinger, Alderson, and Doyle 2009). There has been extensive work over the last decade on the topology of the Internet, resulting in some interesting models based on statistical physics approaches similar to those used by Westland. These models do have the advantages that Box and Draper list, and can be applied usefully in some cases. However (see Willinger, Alderson, and Doyle 2009 for discussion and references), the underlying mechanisms that have produced Internet's real topology are quite different from those of the models that have been developed. As a result, in some cases the models produce visibly incorrect results.

3 The observer effect

Even when we have a model of some complex system that has been extensively validated, caution is necessary in applying that model for some purposes, in particular for influencing or manipulating that system. A system with the addition of a new agent, especially of a powerful one, is no longer the same system, and the model may not apply.

In quantum mechanics we have the solidly established uncertainty principle, which says that the act of observing some physical quantity unavoidably affects it. In social sciences we have the less rigorous, but still solid notion of an observer effect, in which an anthropologist visiting some newly-discovered tribe in New Guinea, or a sociologist sitting inside a high-tech startup in Palo Alto, will affect the interactions that take place. Hence what is observed is not what would have taken place in the absence of such investigators. In finance, we have somewhat similar concepts, such as Soros' concept of "reflexivity" (Soros 1987). We saw this unfold in the bubble that burst in 2008, when, for example, rating agencies' evaluations of securities were distorted by investment banks packing those securities to meet known criteria. We can see this even in academia, where the use of journal impact factors to determine promotion, tenure, grants, and the like, as well as which journals to subscribe to, is leading to manipulation of those impact factors (Reedij and Moed 2008).

In modern social networks, observers have the advantage that they can operate undetected. Since interactions are taking place through photons and electrons that are transmitted by operators, those can be duplicated and studied without the knowledge of the members of the social network. However, once the observers start using the results of their observations to influence the behavior of the social network, problems can quickly arise. Members of the social network will often detect manipulation, since they also have modern information technology tools at their disposal. And they may take action, either individually, by masking their activities, or collectively, by leaving for another network, or by protesting publicly, or by calling for government intervention. They may not do any of this, and may regard some types of manipulation as just good clean fun. But one cannot assume that a manipulated system will behave according to the model of a natural one.

4 Conclusions

Social networks call for far more study, whether the valuations currently attached to Facebook and a few other systems turn out to be justified or not. Empirical research is currently growing rapidly, but there is too little theoretical work, and far too little investigation that spans theory and actual evidence. Westland's paper opens up new avenues of research by proposing to use results from percolation theory to model social networks. This is an interesting approach, as the percolation literature is vast, and might provide useful tools and results beyond those in Westland's paper. The implications that Westland already draws about critical mass, for example, are intriguing.

However, one should treat the results of Westland's paper with caution. They need to be validated by comparison to empirical data. And even if they are confirmed, one will have to be careful about applying his suggestions for spurring growth of social networks, since a complicated system, especially one that involves people, can change its behavior when an outside force is brought in.

References

Box, G. E. P., and Draper, N. R. *Empirical Model–Building and Response Surfaces,* John Wiley and Sons, New York, NY, 1987.

Briscoe, B., Odlyzko, A., and Tilly, B. Metcalfe's Law is wrong. *IEEE Spectrum,* July 2006, 26–31. Available at www.spectrum.ieee.org/jul06/4109. Last accessed on November 11, 2009.

Kilkki, K., and Kalervo, M. KK-Law for group forming services. Presented at XVth International Symposium on Services and Local Access, Edinburgh Conference Centre, Heriot-Watt University, Edinburgh, Scotland, March 21-26, 2004. For information about the conference, see conferences.iee.org/ISSLS2004/. Available at kotisivukone.fi/files/50ajatelmaa.ajatukset.fi/tiedostot/Others/kilkki_kk-law.pdf. Last accessed on November 11, 2009.

Odlyzko, A., and Tilly, B. A refutation of Metcalfes Law and a better estimate for the value of networks and network interconnections. Available at www.dtc.umn.edu/~odlyzko/doc/metcalfe.pdf. Last accessed on November 11, 2009.

Reedij, J., and Moed, H. F. Is the impact of journal impact factors decreasing? *Journal of Documentation,* 64, 2, 2008, 183–192.

Soros, G. *The Alchemy of Finance: Reading the Mind of the Market,* Simon and Schuster, New York, NY, 1987.

Willinger, W., Alderson, D., and Doyle, J. C. Mathematics and the Internet: A source of enormous confusion and great potential. *Notices of the American Mathematical Society,* 56, 5, 2009, 586–599. Available at www.ams.org/notices/200905/rtx090500586p.pdf. Last accessed on November 11, 2009.