

Social Networking Platform Competition Simulation

Hyunwoo Park
University of California, Berkeley
School of Information
102 South Hall
Berkeley, CA 94720-4600
+1-510-926-1185
hwpark@ischool.berkeley.edu

ABSTRACT

In this paper, a new simulator for social network platform competition is introduced. In addition to explaining the basic model for the simulator, three findings will be presented. Key findings from the simulation are as follow. First, a social networking platform starting with a locally concentrated user base outperforms one with a globally distributed user base. Second, topologies similar to the real world show less network effects than a random network topology. Last, existence of first mover's advantage in the social networking competition world is reaffirmed.

Categories and Subject Descriptors

I.6.3 [Simulation and Modeling]: Applications

General Terms

Management, Economics, Experimentation, Verification

Keywords

Simulation, Social, Platform, Networking

1. INTRODUCTION

Competitions among social networking platforms are simulated based on a given social relationship network topology. Therefore, network topology should be provided to run simulation using this simulator. Many network topologies have been developed. Along with these theoretic topologies, real world topologies collected from Facebook and Twitter are also used as inputs for this study. After importing a relationship topology, the simulator creates platforms based on the configuration file. Then, on every iteration, every agent decides which platform it belongs to. The basic mechanism is that every agent pays its attention as money to the platform it belongs to and receives benefits from the platform based on how many friends, agents which have relationship with the agent, are already on the platform. The number of members in each platform is recorded after a predefined number of iterations. The program is written in Java.

2. NETWORK TOPOLOGIES

Five topologies consisting of 10,000 nodes are used as inputs for the simulations in the study. All topology data was converted into

BRITE¹[3] file format before being entered to the simulation. The summary of characteristics of topologies used for the simulation is shown on Table 1 on the next page.

2.1 Waxman Topology

Waxman topology[5] is a random network generated based on the geographic layout of nodes. It decides whether two nodes are connected or not based on the probability based on the distance between the two nodes. BRITE was used for generating Waxman topology.

2.2 Barabasi-Albert Topology

Barabasi-Albert (BA) topology[1] is a network topology reflecting incremental growth and preferential edge creation properties of a real world network topology. The algorithm for creating BA topology starts with one node and then adds nodes to the existing nodes. BA topology was also generated by BRITE.

2.3 Small World Topology

Small World topology is a network topology that exhibits the clustering property like BA topology does. Since BRITE does not support Small World topology, the topology was generated by code. The algorithm[2] takes two parameters, k and p for generating a topology. k defines the edge degree of each node and p defines the probability used for deciding whether to break a certain edge.

2.4 Facebook Topology

Facebook API² (developers.facebook.com) provides two methods for creating a network topology from the Facebook database. First, it provides a function that confirms whether a certain integer is a valid Facebook user ID or not. The function is `users.getInfo`. Second, there is a function that determines whether two user IDs are friends or not. The function is `friends.areFriends`. Combining these two functions, the set of valid user IDs and the relationship matrix which shows whether two nodes are connected can be obtained. The first valid 10,000 user IDs were collected. Since 10,000 nodes were created, 50,000,000 queries had to be made to create the Facebook topology.

¹ <http://www.cs.bu.edu/brite/>

² <http://developers.facebook.com>

2.5 Twitter Topology

Twitter also provides its API³ to public. It is easier to create a topology by using the Twitter API, because it provides a method called `friends` which returns a list of friends of a certain user up to 100. By using this method, Twitter topology was created based on the breadth first search from the given seed ID. The seed was arbitrarily chosen.

Table 1. Network topologies used for the simulation

Topology Type	Wax-man	BA	Small World	Face-book	Twitter
Number of Edges	20,000	19,997	250,000	569,692	141,619
Node Placement	Heavy-tailed	N/A	N/A	N/A	N/A
k, p	N/A	N/A	50, 0.2	N/A	N/A

3. MODELS

3.1 Platforms and Scenarios

Every platform object has two properties: fixed cost and variable cost. Fixed cost is defined as the amount of attention an agent has to pay to join the platform, while variable cost is defined as the amount of attention an agent has to pay to stay on the platform. All platforms are created and join the simulated competition according to a predefined scenario configuration file. The scenario configuration file should be formatted in XML. The required elements per each scenario are Time, PlatformId, Action, FixedCost, VariableCost, InitialStake, Locality. The meanings of each element are summarized in Table 2 below.

Table 2. Scenario configuration elements

Element Name	Meaning
Time (T)	The time when the action will be performed (Integer)
PlatformId (PI)	ID number of the target platform (Integer)
Action (A)	Type of action to be performed (String) [add, modify, delete]
FixedCost (FC)	Fixed cost of the target platform (Double)
VariableCost (VC)	Variable cost of the target platform (Double)
InitialStake (IS)	Number of initial members of the target platform (Integer)
Locality (L)	Whether to start with locally concentrated user base or globally distributed user base (Integer) [0, 1]

³ <http://apiwiki.twitter.com>

The example of the scenario configuration file is presented in Figure 1 on the right page.

```
<?xml version="1.0" encoding="UTF-8"?>
<Scenarios>
  <Scenario>
    <Time>5</Time>
    <PlatformId>0</PlatformId>
    <Action>add</Action>
    <FixedCost>50.0</FixedCost>
    <VariableCost>15.0</VariableCost>
    <InitialStake>950</InitialStake>
    <Locality>1</Locality>
  </Scenario>
  ...
</Scenarios>
```

Figure 1. Example of the scenario configuration XML file

3.2 Agents

Each agent calculates its next move based on the following utility model. Fixed cost and variable cost of each platform explained earlier are included in the formula. Every agent calculates its utility against each platform. The equation and relevant variables are shown in Figure 2. The neighbor depth affecting benefit of the equation below can be configured through general configuration file.

$$u_i(t) = \sum_j \beta_{jk}(t-1) \cdot M - \alpha_{ik} \cdot c_k - (\neg \alpha_{ik}) \cdot C_k \text{ for } \forall k$$

i : Agent i

j : Neighbors of Agent i

k : Platform ID

α_{ik} : 1 if Agent i was on Platform k , 0 otherwise

β_{jk} : 1 if Agent j was on Platform k , 0 otherwise

M : Multiplier to balance cost and benefit

c_k : Variable cost of Platform k

C_k : Fixed cost of Platform k

Figure 2. Utility model equation

After finishing calculating the formula above, every agent holds its utility value for each platform for the next iteration. If multiple subscriptions are allowed, each agent joins every platform which returns positive utility. Otherwise, only the platform offering the top utility is selected. In both cases, agents break all memberships if every utility is negative.

3.3 Other Configurations

Configuration settings required to run simulations are saved in another XML file. The elements included in the general configuration file are Iteration, LogGranularity, LogIntermediate, MultipleMembership, UtilityMultiplier, InfluenceDepth. The meaning of each term is explained in Table 3 on the next page.

Table 3. General configuration elements

Element Name	Meaning
Iteration (I)	Number of iterations (Integer)
LogGranularity (LG)	Length of the interval between times when the result is saved (Integer)
LogIntermediate (LI)	Whether to record the entire membership status (Integer) [0, 1]
MultipleMembership (MM)	Whether to allow agents to be able to join multiple platforms (Integer) [0, 1]
UtilityMultiplier (UM)	Multiplier for balancing cost and benefit (Integer)
InfluenceDepth (ID)	The number of depth that determines the boundary of neighbors (Integer)

The example of the general configuration file is presented in Figure 3 below.

```
<?xml version="1.0" encoding="UTF-8"?>
<Configuration>
  <Iteration>50</Iteration>
  <LogGranularity>1</LogGranularity>
  <LogIntermediate>1</LogIntermediate>
  <MultipleMembership>0</MultipleMembership>
  <UtilityMultiplier>7</UtilityMultiplier>
  <InfluenceDepth>1</InfluenceDepth>
</Configuration>
```

Figure 3. Example of the scenario configuration XML file

4. FINDINGS

Three key findings were observed through the simulation. First, a strategy starting with locally concentrated user base is superior to one starting with globally distributed user base. Second, topologies modeling real world network showing clustering property prevent a winner from taking all by providing sheltering clusters to losers. Last, the existence of first mover's advantage was reconfirmed quantitatively.

4.1 Locality of Initial User Base

If you see Figure 4 below, you notice Platform 0, 2, 4 outperformed Platform 1, 3. The first group of platforms started with a locally concentrated user base, while the second group had a globally distributed user base at the beginning. The locally concentrated user base was created by the algorithm that first searched neighbors of the existing members, and the globally concentrated user base was generated by picking random nodes.

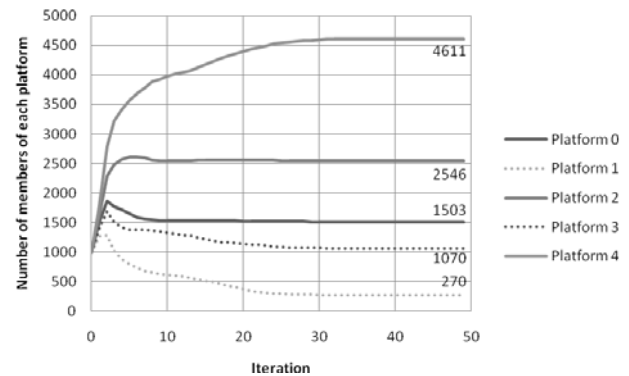


Figure 4. Simulation result for locality effect

The general configuration file was same as the example configuration file shown on Figure 3 on the left column and the scenario configuration file are presented on Table 4 below.

Table 4. Scenario configuration file for locality test

Scenario #	1	2	3	4	5
Time	0	0	0	0	0
PlatformId	0	1	2	3	4
Action	add	add	add	add	add
FixedCost	50	50	50	50	50
VariableCost	15	15	15	15	15
InitialStake	1000	1000	1000	1000	1000
Locality	1	0	1	0	1

4.2 Topology Effect

First, simulation results for Waxman, BA, Small World topologies are compared. Losers in the Waxman case lost everything, while losers in the BA, Small World cases still had their own stakes after the competition was stabilized. The suspected reason for this phenomenon is that topologies resembling a real world network provide sheltering clusters for losers since these topologies have clusters inside. If a platform takes a cluster first, then the platform seems to be able to enjoy the first mover's advantage, which will be explained in the following section, in the cluster. Each result for these topologies is presented on Figure 5-7 respectively.

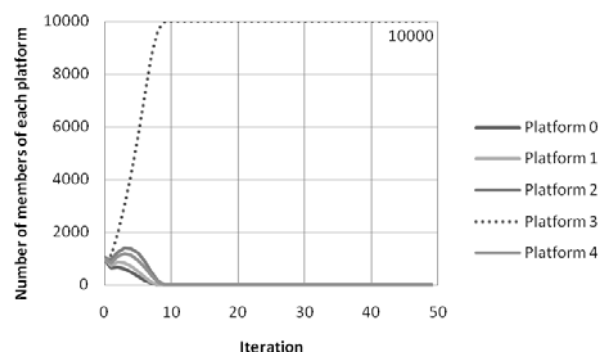


Figure 5. Result from Waxman topology

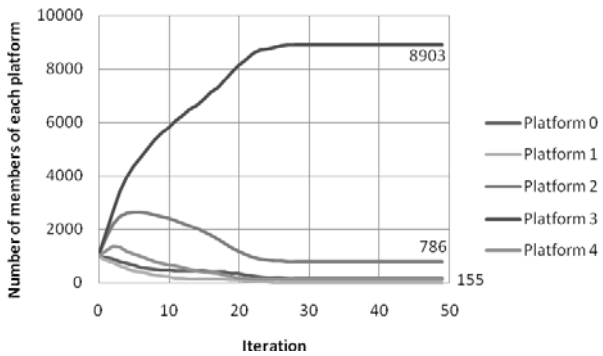


Figure 6. Result from Barabasi-Albert topology

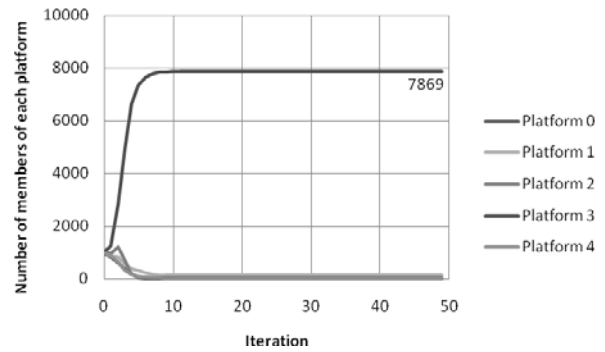


Figure 9. Result from Twitter topology

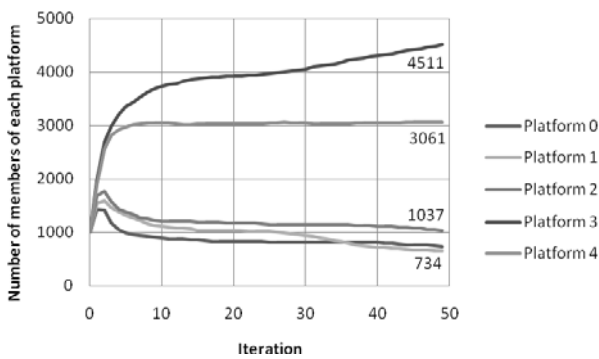


Figure 7. Result from Small World topology

After simulations with theoretically generated topologies were run, the real world topologies such as Facebook and Twitter topologies were used for next simulations. The results are shown on Figure 8 and 9. Unfortunately, these two topologies did not show what BA and Small World topologies did. This is partly because these real world topologies are too dense to be regarded as same with Waxman or BA topologies which has only 20,000 edges. Although Small World topology has more than 200,000 edges and showed stable equilibrium, modified configuration files are applied to the Small World network case. Because of time limit, the configuration files could not be applied to Facebook and Twitter cases. This could be the reason why Facebook and Twitter topologies did not show the desired results. Further simulations are needed to make more firm conclusion from the finding.

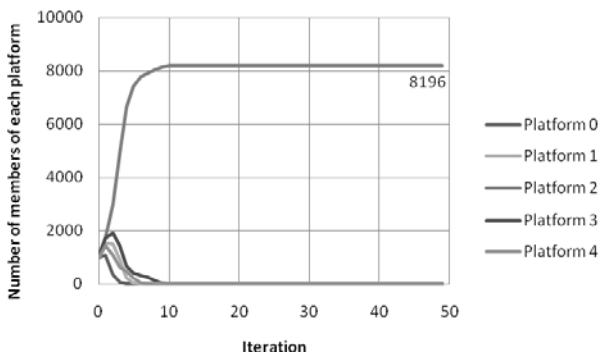


Figure 8. Result from Facebook topology

The general configuration file and the scenario configuration file used for Waxman, BA, Facebook, Twitter topology simulations are shown on Table 5 and 6 respectively. The modified ones are presented on Table 7 and 8.

Table 5. General configuration file for Waxman, BA, Facebook and Twitter topologies

Topology	I	LG	LI	MM	UM	ID
Waxman, BA	50	1	1	0	1	2
Facebook, Twitter	50	1	1	0	1	1

Table 6. Scenario configuration file for Waxman, BA, Facebook and Twitter topologies

Scenario #	1	2	3	4	5
Time	0	0	0	0	0
PlatformId	0	1	2	3	4
Action	add	add	add	add	add
FixedCost	4.89	4.47	4.37	4.32	4.30
VariableCost	1.58	1.92	1.97	1.07	1.48
InitialStake	963	991	1031	1028	974
Locality	0	0	0	0	0

Table 7. General configuration file for Small World topology

Key	I	LG	LI	MM	UM	ID
Value	50	1	1	0	1	2

Table 8. Modified scenario configuration file for the Small World topology

Scenario #	1	2	3	4	5
Time	0	0	0	0	0
PlatformId	0	1	2	3	4
Action	add	add	add	add	add
FixedCost	50	45	45	45	45
VariableCost	15	20	20	10	15
InitialStake	1000	1000	1000	1000	1000
Locality	0	0	0	0	0

4.3 First Mover’s Advantage

The existence of the well known effect called first mover’s advantage was verified in this social networking platform simulation environment. Two different platform profiles were given to the simulator. One has 19% higher fixed and variable costs compared to the other. The more expensive platform is Platform 0 in the figures below. Although Platform 1 beat when the two were introduced to the competition simultaneously, the two could coexist in a half-and-half equilibrium when Platform 0 was introduced earlier than Platform 1 was. The amount of the time delay in the second simulation was 5 iterations long. The simulation results are shown on Figure 10 and 11 below.

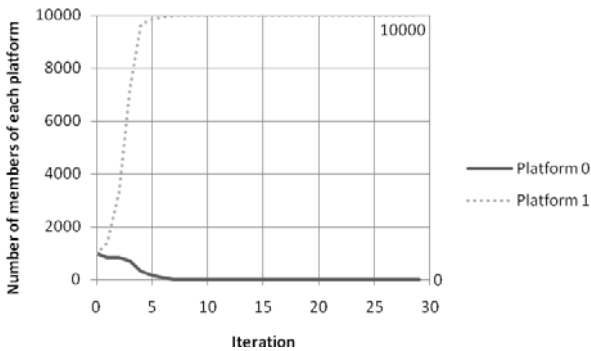


Figure 10. Case when two platforms enter the competition simultaneously

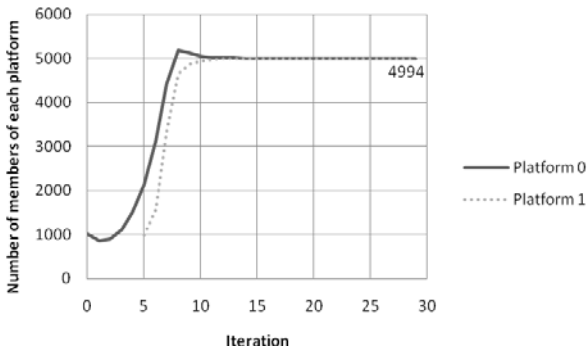


Figure 11. Case when two platforms enter the competition sequentially

The general configuration used for these simulation is presented on Table 9 and the scenario configuration file is on Table 10 and 11.

Table 9. General configuration file for the two simulation

Key	I	LG	LI	MM	UM	ID
Value	30	1	1	0	5	1

Table 10. Scenario configuration file Simultaneous Entrance

Scenario #	T	PI	A	FC	VC	IS	L
1	0	0	add	50	15	1000	0
2	0	1	add	42	12.6	1000	0

Table 11. Scenario configuration file Sequential Entrance

Scenario #	T	PI	A	FC	VC	IS	L
1	0	0	add	50	15	1000	0
2	5	1	add	42	12.6	1000	0

5. FUTURE WORKS

First of all, statistical verification has not been made for the findings. In the future, simulations should be run more times so that statistical valid conclusion can be made.

Second, internal situations such as a membership status should be looked at on each iteration in order for reasons why these findings and phenomena occurred to be explained. It would be also helpful to visualize the simulation process so that how membership status is changed and which parts of network belongs to which platform can be observed.

Last, the whole simulation package needs to be well organized so that it can be open to research community and everyone can use and improve it.

6. ACKNOWLEDGMENTS

First of all, I would like to appreciate Professor John Chuang for giving me invaluable feedback and input for this study. I also thank to Facebook and Twitter for allowing me to collect social network topology using their APIs. Java Universal Network Graph (JUNG) framework[4] was used for implementing the simulator. Lastly, thanks to ACM SIGCHI for allowing me to modify templates they developed.

7. REFERENCES

- [1] Barabasi, A., and Albert, R. 1999. Emergence of Scaling in Random Networks. Science.
- [2] Li, D., and Du, Y. 2008. Artificial Intelligence with Uncertainty. CRC Press, p 333.
- [3] Medina, A., Lakhina, A., Matta, J., and Byers, J. 2001. BRITE: An Approach to Universal Topology Generation. Proceedings of MASCOTS. DOI= <http://doi.ieeecs.org/10.1109/MASCOT.2001.948886>
- [4] O’Madadhain, J., Fisher, D., White, S., and Boey, Y. 2003. The JUNG (Java Universal Network/Graph) Framework. University of California, Irvine, California
- [5] Waxman, B. 1988. Routing of multipoint connections. Selected Areas in Communications. IEEE Journal on.