

RESEARCH

Open Access

# Modeling population growth in online social networks

Konglin Zhu<sup>1\*</sup>, Wenzhong Li<sup>1,2</sup> and Xiaoming Fu<sup>1</sup>

\*Correspondence:

zhu@cs.uni-goettingen.de

<sup>1</sup>Institute of Computer Science,  
University of Goettingen,  
Goettingen, Germany

Full list of author information is  
available at the end of the article

## Abstract

**Purpose:** Online social networks (OSNs) are now among the most popular applications on the web offering platforms for people to interact, communicate and collaborate with others. The rapid development of OSNs provides opportunities for people's daily communication, but also brings problems such as burst network traffic and overload of servers. Studying the population growth pattern in online social networks helps service providers to understand the people communication manners in OSNs and facilitate the management of network resources. In this paper, we propose a population growth model for OSNs based on the study of population distribution and growth in spatiotemporal scale-space.

**Methods:** We investigate the population growth in three data sets which are randomly sampled from the popular OSN web sites including Renren, Twitter and Gowalla. We find out that the number of population follows the power-law distribution over different geographic locations, and the population growth of a location fits a power function of time. An aggregated population growth model is conducted by integrating the population growth over geographic locations and time.

**Results:** We use the data sets to validate our population growth model. Extensive experiments also show that the proposed model fits the population growth of Facebook and Sina Weibo well. As an application, we use the model to predict the monthly population in three data sets. By comparing the predicted population with ground-truth values, the results show that our model can achieve a prediction accuracy between 86.14% and 99.89%.

**Conclusions:** With our proposed population growth model, people can estimate the population size of an online social network in a certain time period and it can also be used for population prediction for a future time.

**Keywords:** Spatiotemporal scale-space; Population distribution; Population growth; Online social networks

## Background

Nowadays online social networks (OSNs) are considered as the most popular applications on the web, which offer platforms for people to interact, communicate and collaborate with others. The user population of online social networks is growing expeditiously. It is reported that Facebook (2013) has reached 900 million users in April 2012. Meanwhile, Twitter (2013) has also surpassed 500 million users in July 2012. The rapid development of OSNs facilitates people's daily communications. However, the growth of user population

also causes problems to service providers, such as overload of servers. One example is the “fail whale” phenomenon in Twitter, where the requested page returns a “fail whale” image when too many burst requests occur.

The issues and patterns of population growth in OSNs have drawn much attention from the academia and many works have been done in the past years. A study of micro evolution on OSNs (Leskovec et al. 2008) captured the best fits of population growth in four different OSNs and showed that the growth tendency varies with time. Torkjazi et al. (2009) and Rejaie et al. (2010) observed S-shaped population growth: it experiences a slow growth in the beginning, following a period of exponential growth and finally a significant and sudden slow down in the growth of the population. However, most of these studies fail to provide a theoretical model to describe the population growth in OSNs. Besides, existing works study population growth only in the temporal dimension, and they lack concern of the dynamics in the geographic scale.

In this paper, we investigate the population growth of OSNs from spatiotemporal scale-space. Our investigation is based on three data sets randomly sampled from the popular OSN website including Renren (2013), Twitter (2013) and Gowalla (2011), from which we explore their population distributions over various geographic locations and time-varying properties on population growth. We find out that in the spatial scale, the population size follows a power-law distribution over geographic locations. In the temporal scale, the population growth in the largest populated location is revealed to fit a power function increasing with time. The number of populated locations also increases as a power function as time. Based on these observations, we propose an aggregated population growth model by integrating the population growth over geographic locations and time. Theoretical analysis is presented to derive this model and comprehensive experiments are conducted to verify its effectiveness. It is shown that the proposed model fits well for population growth in large scale rapid growing OSNs such as Facebook (2013) and Sina Weibo (2013). As an application, we further utilize the model to predict population growth in three data sets, which illustrates that our model can achieve a prediction accuracy between 86.14% and 99.89%.

There are several applications of our work. It has significant meanings for Internet Service Providers (ISPs) to understand the population growth of OSN users, which will further reveal the user interaction patterns and network traffic patterns. It is also benign for the OSN web sites to deploy servers and cast advertisements on the base of population growth model. The third-party service providers can analyze the service market by the model and further optimize their resource deployment and investment.

## Data

To conduct our analysis, we collect data from three online social network sites: *Renren*, a social-based application service, *Twitter*, a social-based media service and *Gowalla*, a location-based online social service. Renren, established in December 2005 and now with 160 million users, is a Chinese online social network which organizes users into membership-based networks representing schools, companies and geographic locations. It allows users to post short messages known as status, blogs and pictures. It also allows people to share contents such as videos, articles and pictures. Twitter, with over 500 million users, launched in July 2006, is known as its microblogging services by which users

can write any topic within the 140 characters limit. Such kind of short message is known as tweet. A follower can follow any other users and receive any kind of tweets from his/her followings. Varied from above-mentioned two online social networks, Gowalla is a location-oriented online social service. People are allowed to check-in their visiting places via mobile devices. It is launched in 2007 and closed in 2012 with approximately 600,000 users.

We collect the Renren and Twitter data sets by crawling from their sites. We start our crawling with randomly selected users from the largest weakly connected component (WCC). Following friends' links in the forward direction in a breadth first search (BFS) fashion, we collect a sample of each social network. To eliminate the degree bias caused by BFS, we launch the BFS-bias correction procedure described in (Kurant et al. 2011). Furthermore, according to the estimation method of the size of social networks by Katzir et al. (2011), we believe the quality and quantity of our data sets are enough to reveal the laws of population growth in OSNs.

In order to capture the growth of population in different geographic locations, we need to know the account creation time and geographic information of each user. We trace user account creation time in Twitter from user profile. However, we cannot explicitly retrieve account creation time from user profile in Renren. To estimate the account creation time precisely, we use the time of a user's first activity such as updating status, posting a blog or interacting with friends as the time point when the account creates. Meanwhile, we seek users' geographic locations from user profiles and choose users with valid geographic information to compose our data sets.

The Gowalla data set, obtained from public source (Cho et al. 2011), contains more than 100,000 users, as well as their social relations and check-in histories. We find the user registration time by their first check-ins. To reveal users' geographic information, similar as (Cho et al. 2011), we infer a user's location by compartmentalizing the globe into 25 by 25 km cells and defining the location as the cell with the most check-ins.

The statistics of three data sets are shown in Table 1. Renren data set (Renren) contains around 1 million users and covers 10,039 locations. It records user growth over the period of January 2006 to December 2010 (60 months). Twitter data set (Twitter) consists of more than 250 thousand users covering 8,929 locations. We collect user population growth over the period of August 2006 and October 2010 (51 months). Gowalla data set (Gowalla) has around 100 thousand nodes, with 5,088 populated locations. It contains the population growth ranging from February 2009 to October 2010 (21 months).

**Table 1 Statistics of data sets**

	Renren	Twitter	Gowalla
Nodes	997,849	257,929	107,089
Locations	10,039	8,929	5,088
Period	01/06 – 12/10	08/06 – 10/10	02/09 – 10/10

The table presents the detail information of Renren, Twitter and Gowalla data sets. It contains the number of nodes, the number of locations and the time period of the data set.

## Methods

We present the methods for modeling of population growth in online social networks in this section.

### Basic approach

To study the population growth in OSNs, we first illustrate the basic approach of modeling the population growth in spatiotemporal dimension.

The population in OSNs grows over locations and time. In spatial aspect, people from different geographic locations may register as users in an OSN, thus people from more and more locations join in the network. The OSN expanding from locations to locations, leads to the growth of population spatially. At the same time, the population in each geographic location grows in temporal scale. People in a geographic location may be attracted to join in the network from time to time and thus the location will have more and more people. Therefore, combining spatial and temporal effects, we model the population growth as the accumulation of populations in different geographic locations, while population in each location changes as a function of time. We describe the population growth in spatial and temporal dimensions as follows.

As the first step, we consider the population growth model in the spatial dimension. The aggregated population  $GP$  is a sum of populations in all populated locations on a certain time point. It can be formulated as an accumulation function:

$$GP = \sum_{x=1}^M S_x, \quad (1)$$

where  $S_x$  denotes the population size in location  $x$ .  $M$  is the total number of populated locations. To calculate this formulation, one needs to know the population size of every location, which is not scalable in large online social networks. We take one step further. Instead of numerating all population sizes, we use the proportion of the number of locations with population size  $s$  over the total number of populated locations, known as population distribution, which is denoted as  $P_s$ , the total number of populated locations  $L$  and the largest population size  $N$  to construct the formulation of aggregated population:

$$GP = \sum_{s=1}^N P_s * L * s, \quad (2)$$

This formulation describes the aggregated population on a certain time point in spatial aspect.

Now we consider the population growth process with temporal factor. The aggregated population growth is a dynamic process, with population in each location growing as a function of  $t$ . Therefore, we consider  $N$ ,  $L$  and  $P_s$  are time-dependent functions, where  $N = n(t)$ ,  $L = l(t)$  and  $P_s = P(s, t)$ . The discrete accumulation of Eq. 2 can be approximated as integral as following:

$$GP(t) = \int_1^{n(t)} (P(s, t) * l(t) * s) ds. \quad (3)$$

The formulation reflects not only the spatial characteristics that it is aggregation of populations in different geographic locations, but also temporal factor that the population growth is a dynamic process as a function of time.

So far, we propose a population growth model in spatiotemporal perspective. To specify this model, we need to study three time-dependent functions: the dynamics of population distribution  $P(s, t)$ , the growth function of populated locations  $l(t)$  and the growth function of the largest population  $n(t)$  in the following subsections.

#### **Population distribution: $P(s, t)$**

The population distribution reveals the proportions of different population sizes in an OSN. We investigate the population distribution by drawing the log-log plots of probability density function (PDF) of populations in different geographic locations of three data sets as shown in Figure 1 with respect to different time points. The observation to the figures shows that the population distributions are close to each other in different time periods. In particular, the population distributions of Renren are close to each other from 10th month to 50th month, as shown in Figure 1a. The similar phenomenon can also be observed from Twitter from 8th month to 48th month in Figure 1b and Gowalla from 5th month to 20th month in Figure 1c. It allows us to use one distribution curve approximately fit all population distributions in one data set, which means the population distributions of various periods are roughly can be fitted with one identical distribution function. Therefore,  $P(s, t)$  is a time-independent function. Then we find that the population distribution of each data set is seemingly a straight line in the log-log scale, which indicates that the distribution can be fitted with a power-law function. To confirm this observation, we conduct the hypothesis testing as described in (Clauset et al. 2009), which uses a *goodness-of-fit test* to determine plausibility of the power-law fit. It generates a  $p$ -value to quantify the plausibility. If  $p$ -value is close to 1, the fit is considered as plausible to the empirical data. Otherwise, it is considered as implausible. The results of the hypothesis tests show that  $p = 1.0$  for Renren,  $p = 0.90$  for Twitter and  $p = 0.95$  for Gowalla. All three values are close to 1. Therefore, we say that the power-law function is plausible to fit three data sets.

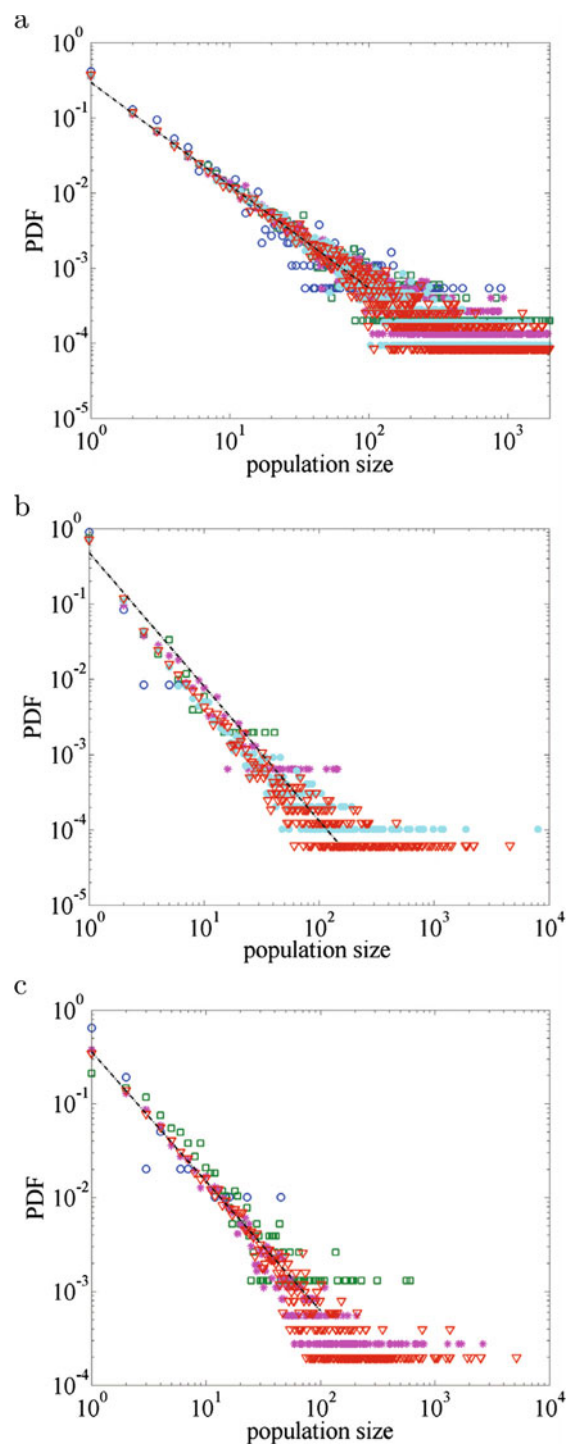
Alternatively, we conduct alternative hypothesis testing regarding the population distribution by the *likelihood ratio test* (Clauset et al., 2009), which suggests that the distribution is a power-law if the likelihood ratio between the alternative and power-law distribution is positive. We calculated the likelihood ratio of exponential distribution compared with power-law distribution, which is 2.23, and the likelihood ratio of log-normal distribution compared with power-law distribution, which is 0.12. The results suggest that power-law is the best distribution to represent population distribution.

We fit each distribution in figures with maximal likelihood estimation (MLE) (Newman 2005, Clauset et al. 2009). The fitting results are shown in dashed lines. It shows that Renren data set has a power-law exponent of 1.4, Twitter has a power-law exponent of 1.78 and Gowalla has a power-law exponent of 1.4.

We give the mathematical model of the population distribution in different time periods as:

$$P(s, t) = \varphi * s^{-\lambda}, \quad (4)$$

where  $\varphi$  is the scaling factor and  $\lambda$  is the power-law exponent. The equation reveals that the population distribution in different time periods is a power-law function, and it is independent from time.



**Figure 1 Population distribution of various periods in three data sets. (a)** Renren, the distribution of population size in 10th month (circle), 20th month (square), 30th month (star), 40th month (solid circle) and 50th month (triangle). Five curves are close to each other and seemingly follow the same distribution. The dashed line is the power law distribution:  $y \sim x^{-1.4}$ . **(b)** Twitter, the distribution of population size in 8th month (circle), 18th month (square), 28th month (star), 38th month (solid circle) and 48th month (triangle). Five curves are close to each other and seemingly follow the same distribution. The dashed line is:  $y \sim x^{-1.78}$ . **(c)** Gowalla, the distribution of population size in 5th month (circle), 10th month (square), 15th month (star) and 20th month (triangle). Four curves are close to each other and seemingly follow the same distribution. The dashed line is:  $y \sim x^{-1.4}$ .

### Populated locations: $l(t)$

To model the population growth in OSNs, one important aspect is to understand the growth of populated locations. In this subsection, we investigate the growth of populated locations.

The growth of populated locations is a function of time  $t$  and defined as  $l(t)$ . To formulate  $l(t)$ , we draw the number of populated locations of three data sets as a function of time in the log-log scale, as shown in Figure 2a, 2b and 2c. The numbers of populated locations are approximately straight lines. Again, we use MLE to fit them shown as dashed lines in figures. The fitting function can be formulated as:

$$l(t) = \eta * t^\varepsilon, \tag{5}$$

with scaling parameter  $\eta$  and the power exponent  $\varepsilon$ . The power exponent of Renren is 1.26, of Twitter is 1.96 and of Gowalla is 1.62.

In a summary, we find that the growth of the number of populated locations in OSNs is a power function of time.

### The largest population: $n(t)$

As we model the population growth as an accumulation of populations in various locations, the largest population as the upper bound of the formulation also needs to be investigated.

To measure the largest population size, we select the location with the largest population sizes in 10th month from three data sets and study the growth of their populations as a function of time, respectively. We define the growth function of the largest population as  $n(t)$ . To obtain  $n(t)$ , we grab the largest population size from month to month in each data set and plot the growth of the largest population size as a function of time. The growth trends of them are shown in Figure 3a, 3b and 3c. Similar to the analysis of populated location growth, the growth of the largest population size can also be fitted using a power function as the following:

$$n(t) = a * t^b + c. \tag{6}$$

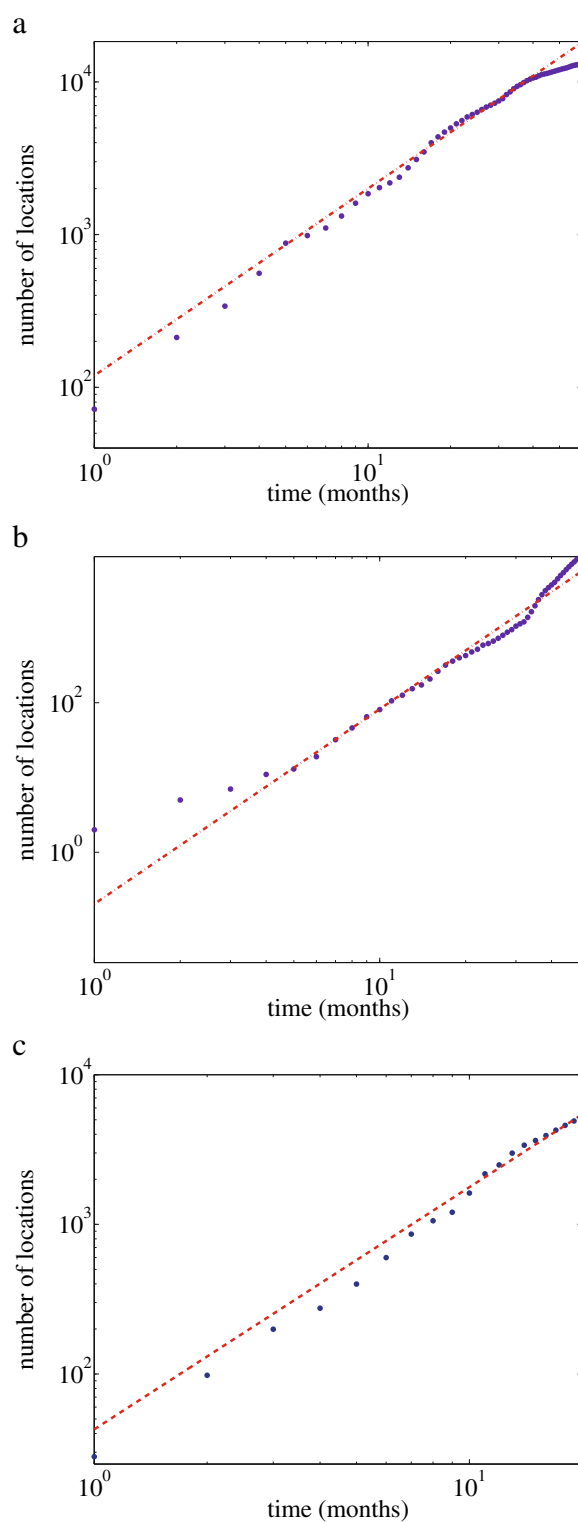
Besides the power component  $a * t^b$ , there is a constant number  $c$  added to the power function. We use this function to fit the largest population size in each data set shown as the solid lines in the figures. Specifically, the power parameter of Renren is 1.31, of Twitter is 2.97 and it is 1.61 for Gowalla.

Therefore, the growth function of the largest population size is a power function. The population growth of a location will affect the aggregated population growth. We will give the detailed model of the population growth in the following subsection.

### The population growth model

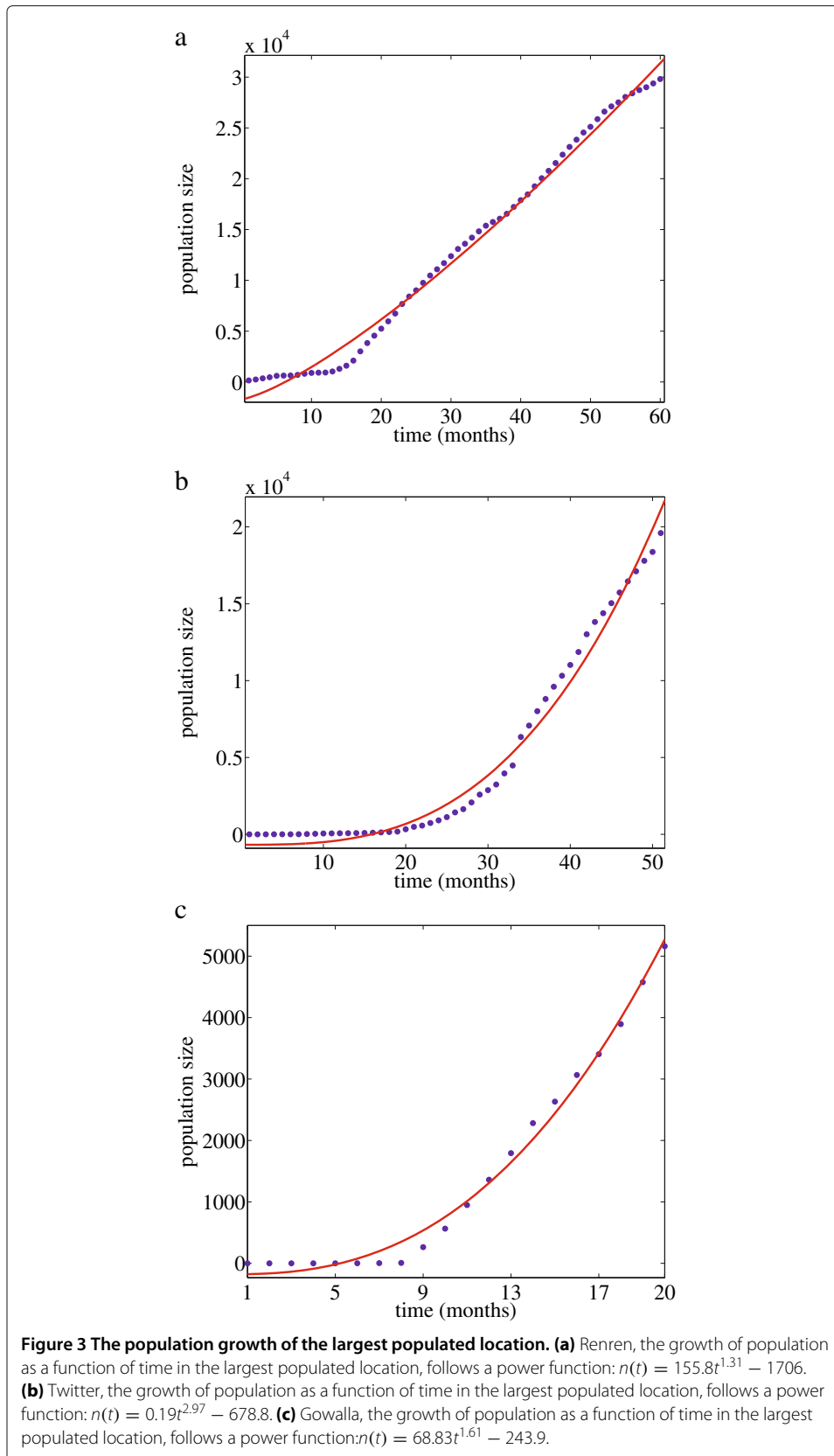
Given the distribution of population, the growth function of populated locations and the largest population size, we insert the above expressions of  $p(s, t)$ ,  $l(t)$ , and  $n(t)$  as shown in Eq. 4, Eq. 5 and Eq. 6 into Eq. 3. Then we have

$$\begin{aligned} GP(t) &= \int_1^{a*t^b+c} ((\varphi * s^{-\lambda}) * (\eta * t^\varepsilon) * s) ds \\ &= \frac{\eta\varphi t^\varepsilon}{2-\lambda} [(a * t^b + c)^{2-\lambda} - 1]. \end{aligned} \tag{7}$$



**Figure 2 The growth of populated locations. (a)** Renren, the growth of populated location as a function of time follows a power function  $y \sim x^{1.26}$ . **(b)** Twitter, the growth of populated location as a function of time follows a power function  $y \sim x^{1.96}$ . **(c)** Gowalla, the growth of populated location as a function of time follows a power function  $y \sim x^{1.62}$ .





The above equation reveals that the population growth is a function of time, and it is similar to power function. The model describes the aggregated population growth of online social networks in both temporal dimension and spatial dimension.

## Results

To present the effectiveness of our population growth model, we evaluate our proposed model from three aspects. First, we verify the model in the early stage population growth of three data sets. Then we evaluate the full population growth in Facebook (2013) and Sina Weibo (2013) by our model. Finally, as an application of our model, we use it to predict the population growth on the latter part populations of three data sets.

### Model verification in three data sets

We verify our population growth model by estimating the early stages of population growth in three data sets (i.e. first 35 months' population growth of Renren, the first 40 months' population growth of Twitter and the first 14 months' population growth of Gowalla).

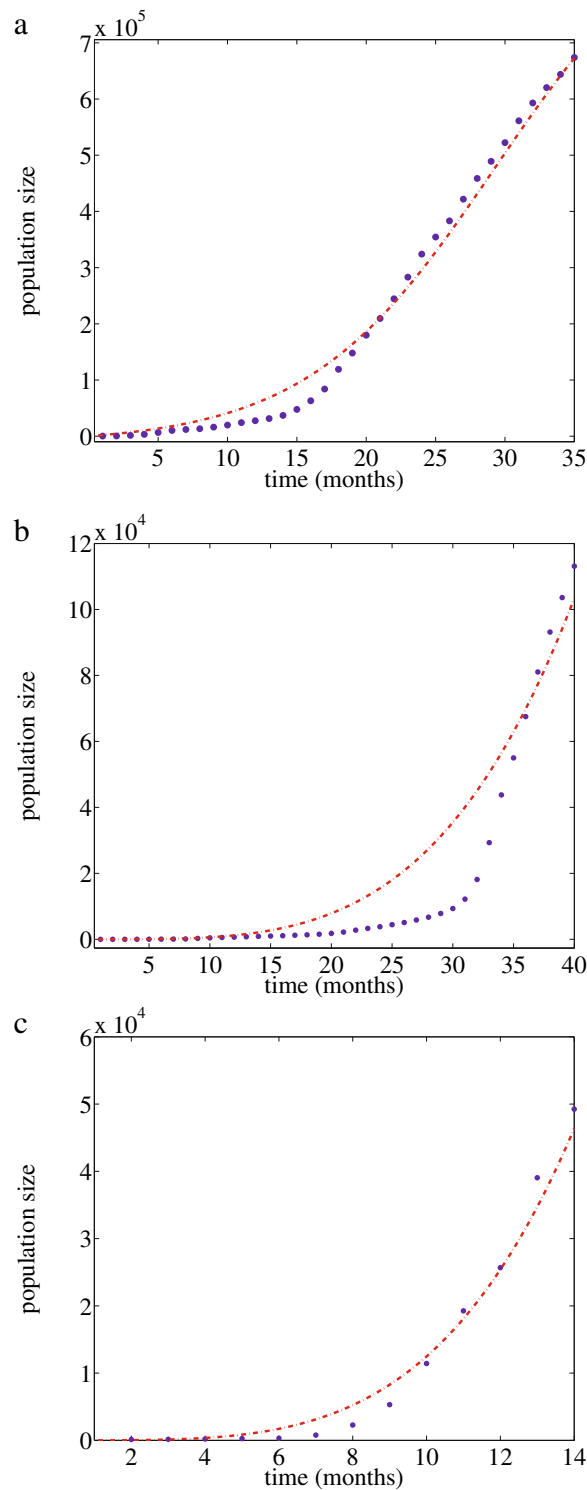
We use the population growth function as shown in Eq. 7 to fit the early stages of population growth in three data sets. We plot the values estimated by the model and the monthly ground-truth population values in three data sets as shown in Figure 4. The points in the figure denote the monthly ground-truth values of population size in each data set, and the dashed lines are the estimated value by the growth model. We compare the estimated value with the value from the data sets. Figure 4a shows the comparison of Renren. The minor difference indicates the model can represent population growth in Renren well. Although the modeled curve of Twitter has the difference with the data set value in the middle of the time period, as shown in Figure 4b, we observe that the model tends to fit the data set population very well at the end of the period. The model shows very good performance in Gowalla as shown in Figure 4c.

In a summary, the verification of the population growth model in three data sets validate the correctness of the model.

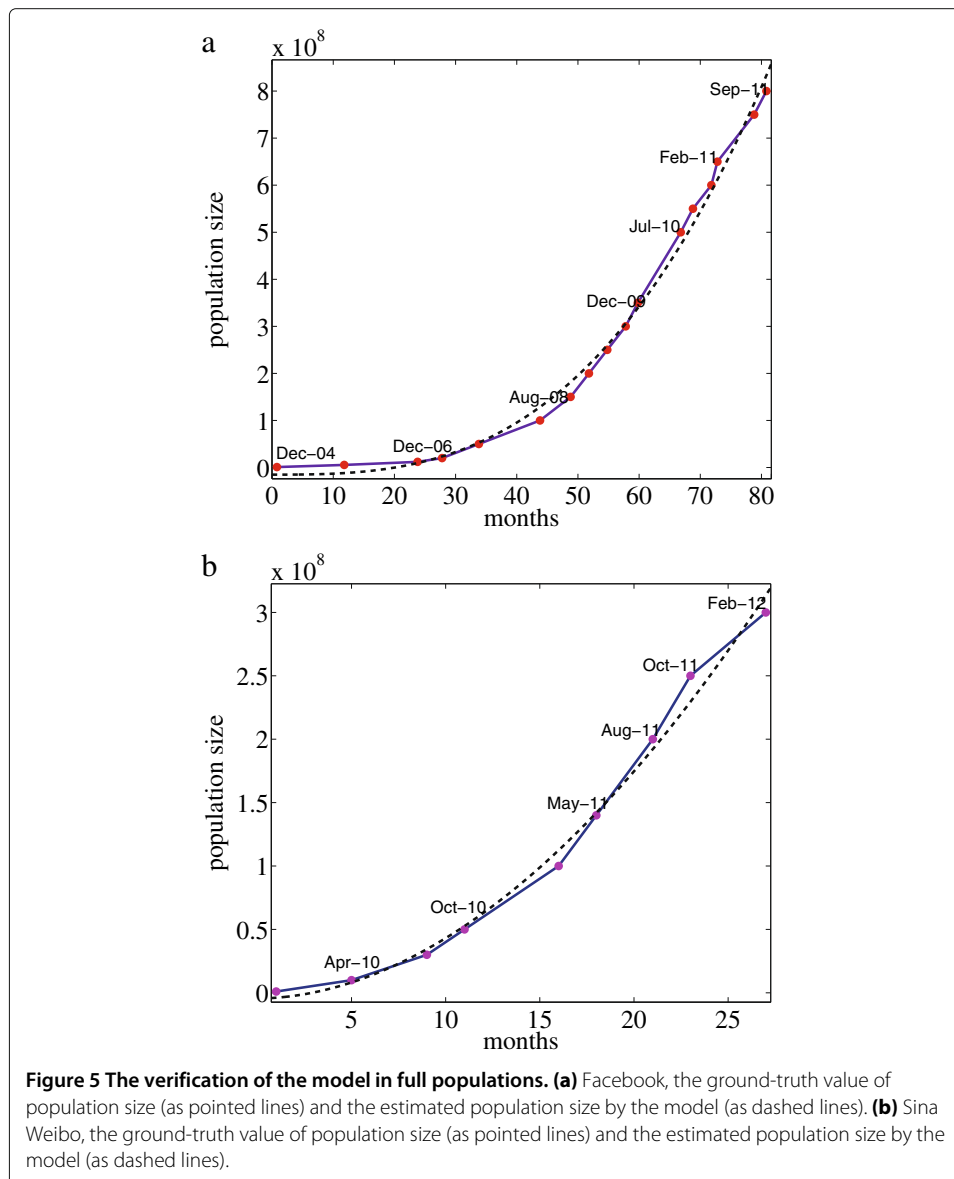
### Model verification in full OSN populations

In this subsection, we verify our model in full population of Facebook (2013) and Sina Weibo (2013) to show its effectiveness in two OSNs with integral populations.

We conduct experiments on the full user populations in Facebook and Sina Weibo. Facebook was launched in 2004 and reached 800 million in September 2011. The population growth trend is shown in Figure 5a. Sina Weibo, a Chinese microblogging web site, was launched in 2009. Its population has reached 300 million in February 2012. The growth trend of the population is shown as Figure 5b. For each population growth, we use Eq. 7 to fit the population. We plot the value estimated by the model shown as dashed lines in Figure 5a and Figure 5b, respectively. By comparing the estimated values with the ground-truth population values, we find that the estimated values are close to the real values of populations, which suggests that our model is valid for population growth of Facebook and Sina Weibo.



**Figure 4** The verification of the model in three data sets. **(a)** Renren, the ground-truth value of the first 35 months' population size (as points) and the estimated population size by the model (as dashed lines). **(b)** Twitter, the ground-truth value of the first 40 months' population size (as points) and the estimated population size by the model (as dashed lines). **(c)** Gowalla, the ground-truth value of first 14 months' population size (as points) and the estimated population size by the model (as dashed lines).

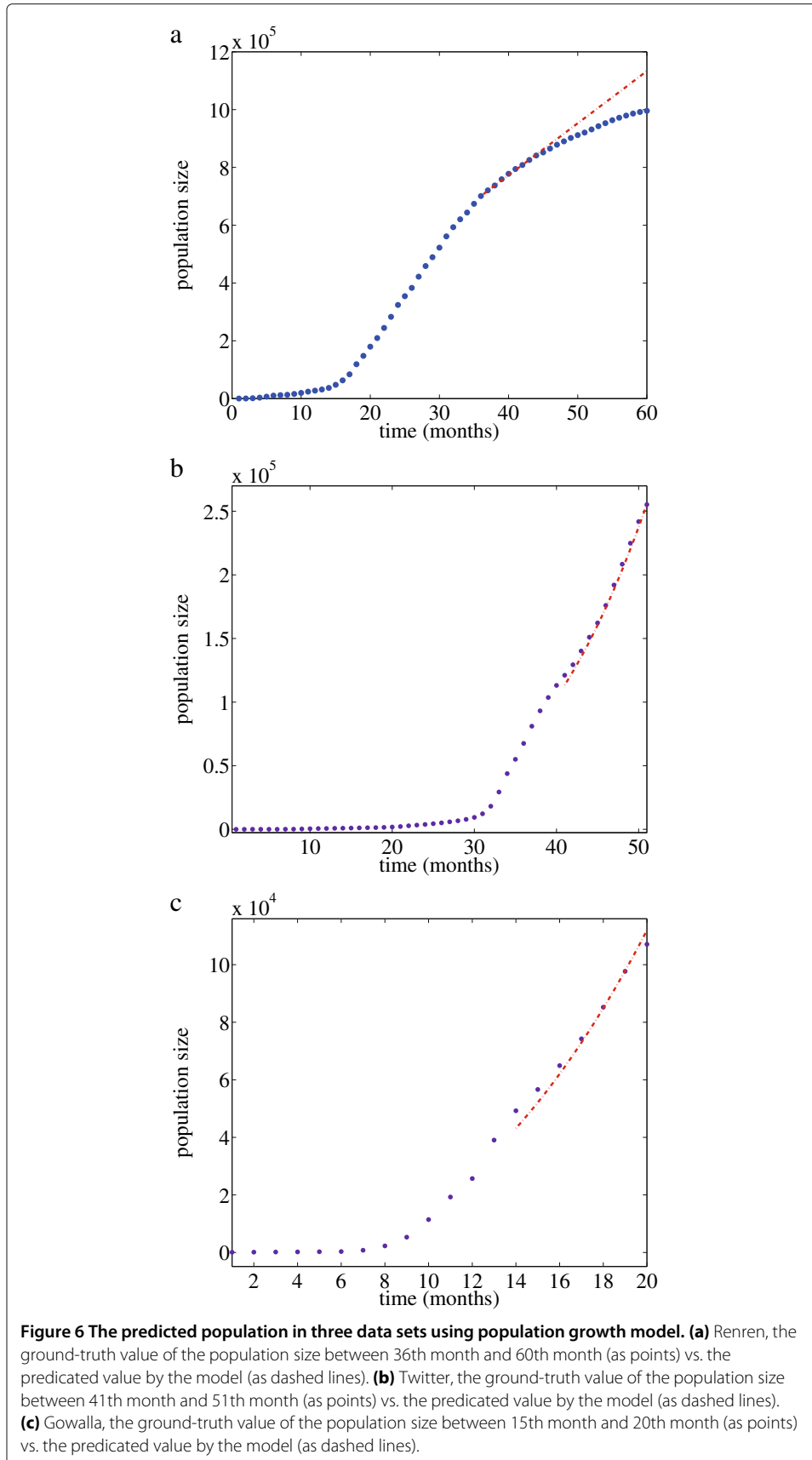


**Figure 5** The verification of the model in full populations. **(a)** Facebook, the ground-truth value of population size (as pointed lines) and the estimated population size by the model (as dashed lines). **(b)** Sina Weibo, the ground-truth value of population size (as pointed lines) and the estimated population size by the model (as dashed lines).

### Using the model for population prediction

As an application, we use our model to predict the latter part population growth of three OSN data sets.

Through estimating the early stage population size in three data sets, we have built the population growth model with parameters. Now, we use the established multi-step ahead model to predict the population growth from the 36<sup>th</sup> to 60<sup>th</sup> month in Renren, from the 41<sup>st</sup> to 51<sup>st</sup> month in Twitter, and from 14<sup>th</sup> to 20<sup>th</sup> month in Gowalla. We plot Figure 6 to show the ground-truth values of three data sets and their prediction results. The points are the population size of each data set at different times. The dashed lines are the predicted population size. Figure 6a presents the results of Renren. The predicted population fits the data set population well at the beginning. The data set population growth becomes slower than the predicted value as time changes. The prediction results of Twitter and Gowalla are also close to real populations as shown in Figure 6b and Figure 6c.



To quantify the accuracy of the prediction, We use prediction accuracy (PA) and define it as follows:

$$PA = 1 - \left| \frac{\text{predicted value} - \text{actual value}}{\text{actual value}} \right| \quad (8)$$

According to Eq. 8, we present the prediction accuracy of our model in several months in Table 2. The results in the table show that our model achieves 94.30% average prediction accuracy for Renren, 98.94% average prediction accuracy for Twitter, and 97.73% average prediction accuracy for Gowalla. These results suggest that our model can perform very well in OSN population prediction.

**Table 2 The prediction accuracy of the population growth model**

Renren		Twitter		Gowalla	
40th	99.36%	42th	96.90%	16th	95.55%
45th	98.71%	44th	99.04%	17th	98.25%
50th	95.50%	46th	99.35%	18th	99.62%
55th	91.77%	48th	99.89%	19th	99.80%
60th	86.14%	50th	99.53%	20th	95.43%

The table is the prediction accuracy of the population growth model of selected months in three data sets.

## Discussion

In this section, we discuss the impact of methods used for data collection and processing, the effects of populations acquirement in OSNs and the scope of our growth model.

### Data collecting and processing

In this paper, we use BFS started from random selected nodes to collect the data samples from online social networks. The population growth model based on random sampling data may cause inaccurate population estimation and prediction. To avoid this issue, we conduct several actions to make the data sets fair enough. First of all, we launch a BFS-bias correlation procedure (Kurant et al. 2011) to eliminate the biases caused by random walking. Secondly, according to the estimation method of social network sizes by random sampling presented in (Katzir et al. 2011), we argue that the quality of data sets cannot be affected by random selected nodes. Finally, we use full population size in Facebook and Sina Weibo to validate the effectiveness of the proposed population growth model. All these efforts are made to let the data samples collected from online social networks be accurate enough for modeling population growth.

### Population acquirement in OSNs

We use registered users as the population in OSNs. Thus, we count every registered user in the network as a member of the total population. It contains both active users and inactive users. In our model, we consider these inactive users as one part of the aggregated population for the following reasons: (1) Inactive users are also one part of the population. We cannot say a user who is not active in the network does not belong to this network. (2) Detecting active users is a complex process which cannot be done simply. For example, people may find the active users from the activity that the user conducts on the web site. However, many people may only browse the web site without any explicit activity. They perform inactive in interacting with others in OSNs, but they are also active users. By

these two reasons, we consider registered users instead of active users as the population of an OSN.

### Scope of the growth model

Our population growth model focuses on the growing stage of OSNs. We do not intend to track the life circle of an OSN. When an OSN's population stops growing, our model will not take effect on it. To specify the growing stage of an OSN, we define the population monthly growth rate as

$$r(t) = \frac{\text{population}(t) - \text{population}(t - 1)}{\text{population}(t - 1)},$$

if  $r(t) > 0$ , we consider the population is in growing stage between time  $t - 1$  and  $t$ . We say an OSN is in growing stage if its monthly growth rates are all greater than 0 in the observed time period. Actually, our crawled data sets and two full population OSNs (Facebook and Sina Weibo) are all in population growing stage, which adapt to our study. Besides, most popular OSNs (such as Facebook and Twitter) are currently still in the growing stage. Therefore, our population model focuses in the stage of population growth in OSNs.

### Conclusions

In this paper, we propose a population growth model for online social networks. We investigate the population growth in spatiotemporal perspective. By studying the population growth over locations and time in three data sets of Renren, Twitter and Gowalla, we find out the population distribution is a power-law function over various locations. The growth of populated locations and the largest population are both power functions of time. By integrating the temporal and spatial characteristics of population growth, we conduct the general population growth model. Extensive experiments show that our model can fit the population growth in Facebook and Sina Weibo. As an application, we use the model for population predication in three data sets, and it can achieve a prediction accuracy between 86.14% and 99.89%.

#### Competing interests

The authors declare that they have no competing interests.

#### Authors' contributions

All authors contributed equally to the paper. All authors read and approved the final manuscript.

#### Acknowledgements

The authors acknowledge the funding from Alexander von Humboldt Foundation and DAAD Foundation. We would like to thank Mr. Cong Ding for the help to crawl Renren data sets. We also appreciate the comments from anonymous reviewers for improving the quality of the paper.

#### Author details

<sup>1</sup>Institute of Computer Science, University of Goettingen, Goettingen, Germany. <sup>2</sup>State Key Laboratory for Novel Software and Technology, Nanjing University, Nanjing, China.

Received: 14 March 2013 Accepted: 27 May 2013

Published: 24 June 2013

#### References

- Clauset, A, Shalizi CR, Newman MEJ: **Power-law distributions in empirical data**. *SIAM Rev* 2009, **51**(4):661—703.
- Cho, E, Myers SA, Leskovec J: **Friendship and mobility: user movement in location-based social networks**. In *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '11)*. New York: ACM; 2011:1082–1090.
- Facebook; 2013. [<http://www.facebook.com>]

- Katzir, L, Liberty E, Somekh O: **Estimating sizes of social networks via biased sampling**. In *Proceedings of the 20th International Conference on World Wide Web (WWW '11)*. New York: ACM; 2011:597—606.
- Kurant, M, Markopoulou A, Thiran P: **Towards unbiased BFS sampling**. *IEEE J Selected Areas Commun* 2011, **29**(9):1799–1809.
- Leskovec, J, Backstrom L, Kumar R, Tomkins A: **Microscopic evolution of social networks**. In *Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '08)*. New York: ACM; 2008:462–470.
- Newman, MEJ: **Power laws, Pareto distributions and Zipf's law**. *Contemp Phys* 2005, **46**(5):323–351.
- Rejaie, R, Torkjazi M, Valafar M, Willinger W: **Sizing up online social networks**. *IEEE Netw: Mag Global Internetworking* 2010, **24**(5):32–37.
- Renren net; 2013. [http://www.renren.com]
- Sina weibo; 2013. [http://www.weibo.com]
- Torkjazi M: **Hot today, gone tomorrow: On the migration of MySpace users**. In *Proceedings of the 2nd ACM Workshop on Online Social Networks (WOSN '09)*. New York: ACM; 2009.
- Twitter; 2013. [http://www.twitter.com]

doi:10.1186/2194-3206-1-14

**Cite this article as:** Zhu et al.: Modeling population growth in online social networks. *Complex Adaptive Systems Modeling* 2013 **1**:14.

Submit your manuscript to a SpringerOpen<sup>®</sup> journal and benefit from:

- ▶ Convenient online submission
- ▶ Rigorous peer review
- ▶ Immediate publication on acceptance
- ▶ Open access: articles freely available online
- ▶ High visibility within the field
- ▶ Retaining the copyright to your article

---

Submit your next manuscript at ▶ [springeropen.com](http://springeropen.com)

---