Technological Forecasting & Social Change 105 (2016) 167–178

Contents lists available at ScienceDirect

Technological Forecasting & Social Change

Forecasting new product diffusion with agent-based models

Yu Xiao [a,*], Jingti Han [b,c]

[a] School of Business Information, Shanghai University of International Business and Economics, Shanghai 201620, China
[b] School of Information Management and Engineering, Shanghai University of Finance and Economics, Shanghai 200433, China
[c] Laboratory Center, Shanghai University of Finance and Economics, Shanghai 200433, China

ARTICLE INFO

Article history:
Received 7 August 2015
Received in revised form 16 January 2016
Accepted 27 January 2016
Available online 23 February 2016

Keywords:
Agent-based model
Bass model
Social network
Combined prediction
Forecasting model

ABSTRACT

Agent-based model (ABM) has been widely used to explore the influence of complex interactions and individual heterogeneity on the diffusion of innovation, while it is seldom used as a forecasting tool in the innovation diffusion literature. This paper introduces a novel approach of forecasting new product diffusion with ABMs. The ABM is built on the hidden influence network (HIN) over which the innovation diffuses. An efficient method is presented to estimate non-structural parameters (i.e., p, q and m) and a multinomial logistic model is formulated to identify the type of the HIN for diffusion data. The simulation study shows that the trained logistic model performs well in inferring the HINs for most simulated diffusion data sets but poorly for those generated by ABMs with similar HINs. Therefore, to reduce the possible prediction loss arising from the misspecification of the HIN, three methods, namely, the predicted HIN, the weighted averaging and simple averaging, are developed to forecast new products diffusion. Their performances are evaluated by using a data set composed of 317 time series on consumer durables penetration. The results show that most identified HINs have moderate topology, and that our methods outperform four classical differential equation based diffusion models in both short-term and long-term prediction.

1. Introduction

Modeling and forecasting new product diffusion has been the focus of marketing research since the publication of a few seminal works, i.e., Rogers and Shoemaker (1971), Bass (1969), Fourt and Woodlock (1960). Bass (1969) supposes that the diffusion process of a new product is driven by the innovative and imitative inclination of the target population. The innovative inclination is influenced by external factors such as advertisement or pricing strategy, and the imitative inclination by those factors determining the strength of neighborhood influence. The Bass model is extended in a number of ways, such as adding market variables (Kalish, 1985), pricing and adverting variables (Bass, 2004), and seasonal trends (Peers et al., 2012; Fernández-Durán, 2014), allowing cross-country markets (Talukdar et al., 2002) and multi-generation of products (Jiang and Jain, 2012), or dividing the potential market into two segments—influentials and imitators (Van den Bulte and Joshi, 2007). For more detail, refer to literature reviews such as Chandrasekaran and Tellis (2007); Meade and Islam (2006) and Peers et al. (2012).

Most of the above diffusion models are differential equation based. An inherent flaw is their incompetence of flexibly incorporating network structure or individual heterogeneity, because they gain ground on the premise that a potential market is perfectly mixing or can be divided into several perfectly mixing segments. This premise has been increasingly challenged since the discovery of some common social network properties, such as “small world” effect (Watts and Strogatz, 1998), power-law degree distribution (Barabási and Albert, 1999), and social components (Maslov et al., 2004). These factors' influences on the diffusion process are widely exploited theoretically (Choi et al., 2010; López-Pintado, 2008) and validated empirically (Centola, 2010). A diffusion model not taking into account network structure would get biased penetration forecasts for a new product.

In this paper, we model the diffusion of new product over a network using a parsimonious agent-based model (ABM). The ABM is built on the hidden influence network (HIN), whose edge is defined as the influence that plays a role in the adoption of a new product between neighboring customers. The HIN does not refer to the overt network via which individuals communicate with each other in daily life, but rather refers to the active subset of the full social network for the diffusion process. This concept has been also taken by Watts and Dodds (2007) to examine the influence of influential nodes on the cascade of public opinion formation, and by Trusov et al. (2013) to improve the pre-launch diffusion forecasts.

We assume that an agent's adopting decision is influenced by its innovativeness and susceptibility to the neighboring agents' adoption, a rule consistent with the Bass (1969) model. We differ from, however, most existing models in applying an asynchronous state update rule (Harvey and Bossomaier, 1997; Cornforth et al., 2005). The asynchronous rule defines that, in each time step, all agents decide whether to

* Corresponding author.
E-mail address: jxndtbxiaoyu@163.com (Y. Xiao).

http://dx.doi.org/10.1016/j.techfore.2016.01.019
0040-1625/

adopt the innovation in a random and exclusive sequence. Intuitively, if the time interval were very long, i.e., a season or year, the early adopters in this interval, as the adopters in the previous intervals, would also impact the decision of potential customers. From this view of point, the ABM with this rule applies to aggregate-level penetration time series composed of yearly or monthly data points.

To this point, it's essential to note that our focus is not on developing a completely new ABM that is meant to replace others, but rather on providing a novel way of calibrating the ABM and consequently making it an efficient and economical forecasting tool. Most existing ABMs in innovation diffusion studies are developed to analyze the influence of individual heterogeneity or network structure on diffusion processes. Only a limited number of researchers have incorporated network properties in their new product forecasting models (Dover et al., 2012; Trusov et al., 2013). Their methods, however, pose rigid restrictions on diffusion data and are thus not suited for most common empirical diffusion time series.

We introduce a novel procedure to calibrate the ABM using aggregate-level diffusion data. The procedure consists of a two-stage method for estimating the non-structural parameters (i.e., p, q and m) and a discriminate model for identifying the best HIN among a group of networks. For the estimation of non-structural parameters, the first stage is to store a wide range of diffusion rate curve generated by the ABM, along with the corresponding parameter values, in a data set; the second stage is to search out the optimal potential market for each curve to best fit with the target diffusion data and, finally, to take the one with the least sum of squares of errors (SSE) as the estimates of the ABM. The data set can be repeatedly used, thus greatly reducing the computing consumption when there are overwhelming fitting tasks.

Unlike the non-structural parameters, the HIN is a complex component of the ABM. It is impossible to find the optimal HIN from the space composed of all possible networks. Therefore, similar to Trusov et al. (2013), we identify the best HIN among a group of HINs, which have a few common structural properties of empirical social networks. In detail, we define three discriminative indices dependent on the relative explanatory and predictive performances of a group of ABMs, whose relationship with the class of potential HIN is formulated by a multinomial logistic model. This model enables us to identify the best HIN (or its analog) for diffusion data, thus improving the penetration forecasts.

We evaluate the performance of the above methods through conducting extensive simulation experiments. As shown in the simulation study, the recovery of the "true" HIN is not successful in some situations. Therefore, we further propose three forecasting methods, namely, the predicted HIN, the weighted averaging and the simple averaging. The predicted HIN method takes the forecast solely from the ABM with the predicted HIN, while the weighted or simple averaging method combines forecasts from all the ABMs.

Our work meets the recent call for a rigorous use of ABM in the study of diffusion of innovation (Rand and Rust, 2011). Both the simulation and empirical studies show that our approach has better predictive performance than classical differential equation based diffusion models, i.e., the Bass model, Gamma/Gompertz model, Gamma/Shifted Gompertz model and Weibull model. This result is consistent with Dover et al. (2012), but it comes from fitting market-level penetration data used mostly by classical aggregate diffusion models. Moreover, we find that most of the predicted HINs for empirical penetration data are similar to a random regular network, and only a small proportion is similar to a lattice network or a random network with high variance of degree distribution.

The rest of this paper is organized as follows. In Section 2, we demonstrate the relevance of our work for the innovation diffusion literature. In Section 3, we build an asynchronous ABM on the HIN. In Section 4, we present a fast method to estimate non-structural parameters, and develop a discriminate model to identify the HIN of the ABM. In Section 5, we perform a number of simulation experiments to show the efficiency and consistency of our methods, and also propose three forecasting methods based on the ABM. In Section 6, we conduct the empirical study to test the performance of the forecasting methods. We discuss our findings and make conclusions in Section 7.

2. Prior work

Many new product or technology forecasting models can be traced back to the Bass (1969) model. This model can be interpreted in a number of ways, one of which is that all potential buyers are located in a fully connected HIN (Fibich and Gibori, 2010), and the behavior of a potential buyer is determined by its innovativeness and susceptibility. Similar to the Bass (1969) model, most market-level diffusion models can also be transferred into a fixed HIN version, but they do not provide a mechanism to vary the HINs in their models.

Agent-based model provides that kind of mechanism, and has been used widely in the innovation diffusion literature (Garcia and Jager, 2011; Kiesling et al., 2012). For example, Watts and Dodds (2007) use a parsimonious agent-based model to examine the "influential hypothesis", i.e., whether influential are important to the formation of public opinion; Hanaki et al. (2007) study the cooperative behavior emerging in an environment where individual behaviors and interaction structures coevolve; Choi et al. (2010) base a simple computational model on small-world graphs to examine the role of network structure and network effects in the success of global diffusion; Goldenberg et al. (2010) challenge the conventional wisdom that network externalities lead to faster diffusion due to the network effect; Kuandykov and Sokolov (2010) examine the effect of network structure on S-shaped diffusion curves; Stummer et al. (2015) build an ABM that deals with repeat purchase decisions, the competitive diffusion of multiple products, and both the temporal and the spatial dimension. The above studies use ABM to examining the possible influences of network structure or individual heterogeneity on diffusion phenomena, rather than to quantitatively predicting the diffusion of new products or technologies. Here, we focus on developing a valid and efficient calibrating method for the ABM using only aggregate-level diffusion data, and consequently take the ABM as a forecasting tool.

Recent work highlights the importance of forecasting new product diffusion by taking network structure into consideration (Iygengar et al., 2011; Stephen and Toubia, 2010). A few researchers have built forecasting methods based on ABMs, and showed their superiority over classical market-level models in prediction performance. For example, Dover et al. (2012) propose a two stepwise estimation procedure for their network diffusion forecast model. The first step is to estimate the constraints among network parameters and the p, q; the second step is to simulate adoption patterns over random networks with the identified constraints, which are used to fit to the empirical diffusion data. They show that the incorporation of network information in diffusion models can significantly improve the predictive performance. Their success in detecting the network, however, depends on the quality of diffusion data. For example, in their empirical study, the selected diffusion data comprises at least 50 points of data and has an identifiable peak with relatively smaller fluctuations, which is not typical for common market-level penetration data.

Trusov et al. (2013) fit the Bass model to diffusion data generated in a social network, and compare the joint probability distribution of estimates for a number of empirical diffusion time series with the joint probability distributions for several simulated data sets. The optimal network of the empirical diffusion data set can be identified by this approach, thus improving the prediction of early diffusion. However, this method depends on a hypothesis that all the empirical diffusion data are generated in the same social network, limiting its practical application.

Our approach is different from the above studies, because we fit the ABM to the very common empirical penetration data that own a relatively large disturbance and only a few data points and that are also