Physica A 441 (2016) 62–68

Contents lists available at ScienceDirect

Physica A

journal homepage: www.elsevier.com/locate/physa

A diffusion perspective on temporal networks: A case study on a supermarket

Shiguo Deng, Lu Qiu, Yue Yang, Huijie Yang *

[a] *Business School, University of Shanghai for Science and Technology, Shanghai 200093, China*

HIGHLIGHTS

- Temporal network is understood from a perspective of diffusion process.
- Group IDs are used to represent trajectories of a temporal network.
- The supermarket evolves in a strong order, rather than a complete random.

ARTICLE INFO

Article history:
Received 20 February 2015
Received in revised form 21 June 2015
Available online 9 September 2015

Keywords:
Temporal networks
Diffusion process
Behaviors of Supermarket customers

ABSTRACT

From a large amount of records, one can extract behavioral characteristics of a social system at different scales. Theoretically, it can help us to know how the global behavior of a social system is formed from individual activities. Practically, it can be used to optimize and even to control the social system. Complicated relationships between the individuals form a network, which evolves with time. The behavior of the system can be accordingly understood in the framework of temporal network. In the present paper, instead of focusing on microscopic structures, we develop a framework to investigate temporal networks from the viewpoint of diffusion process, in which each snapshot network is divided into groups and the ID number of the group a node belongs to is used to measure its state. By this way trajectories of the nodes form an ensemble of realizations of a stochastic process. As an illustration, we investigate the diffusion behavior of a supermarket. One can find that with the increase of time the customers cluster and separate into different groups. Meanwhile, the system evolves in a significant order way, instead of a complete random one.

1. Introduction

Human behavior in social systems attracts persistent attentions of researchers from diverse research fields. A large amount of records collected in recent years make it possible for us to extract behavioral characteristics of a social system at different scales, based upon which one can understand how a global behavior is generated from characteristics of non-homogeneous groups or even of different individuals. Practically, understanding of human behavior is the preliminary step to optimize performance of the system. Interesting discoveries have been reported in literature, including scaling laws in human mobility [1–3], bursts in temporal activities [4–7], and (long-term) memory effect widely existing in social systems [8–16]. Generally, characteristics of individuals in a social system will change with time, which may lead to evolution of global behavior. Let us consider, for instance, customers in a supermarket. By using buying records, one can easily calculate similarity between each pair of individuals, and accordingly map the similarity relationships to a network,

* Corresponding author.
E-mail address: hjyang@ustc.edu.cn (H. Yang).

http://dx.doi.org/10.1016/j.physa.2015.08.058

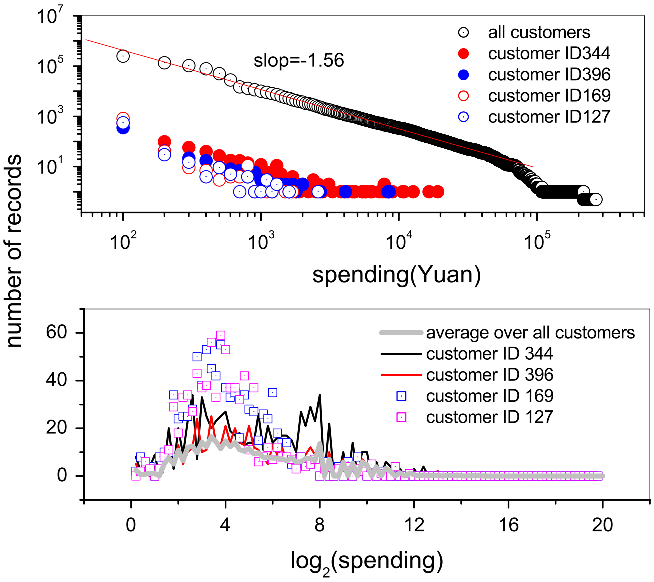

Fig. 1. (Color online) Distribution of costs. (a) The spending range is divided into many bins with width 100 Yuan each. We reckon the number of records whose costs are in every specified bin, as being the distribution corresponding to the bin. Distribution of costs for all the customers obey a perfect power-law (the records from 2001 to 2009 are used). And four typical distributions for individual customers obey also roughly power-laws (the records from 2001 to 2003 are used). (b) Distributions of the logarithm values of costs (in the duration of 2001 to 2003) for all the customers (averaged) and for the four individuals in (a).

cluster all the customers into different groups. But the network and consequent clusters depend strongly on the specific sets of records. Regarding each network constructed from records in a specific duration as a snapshot of the system's dynamical process at the corresponding time, a series of snapshots form a temporal network. Hence, in a certain degree evolutionary behavior of the customers can be investigated in the framework of temporal networks.

Novel concepts have been proposed in the present temporal network theory [17–22], but most of them are defined on every node's relationships with others and their changes with time. In the present paper, instead of focusing on changes of microscopic structures, we propose an analysis scheme to investigate evolution of groups. Technically, from each snapshot network we cluster the nodes into different groups, and take the group each node belongs to as measurement of its state. We will be interested in how the present groups split into different segments, and how the segments subsequently merge into new groups in the successive snapshot networks. Evolutionary trajectory of each node is then described by the series of IDs of groups it belongs to in the successive snapshots. A temporal network is mapped to a consequent diffusion process, and all the trajectories of the nodes form an ensemble of realizations of this stochastic process.

As an illustration, we consider a system containing a total of 563 loyal customers in a supermarket. The records cover a duration of 9 years. We find that behavior of most customers diversify dramatically, namely, customers in the same group at present time will usually join in different groups in the next snapshot network. From the viewpoint of complexity, the system is in a significant order state and evolves towards much more order.

2. Materials and methods

2.1. Data

We analyze, as a typical example, behavior of loyal customers of a supermarket [23]. The database contains 563 customers, a total of 249, 522 records in the duration from 2001 to 2009. Each record includes the customer ID, time, set of goods, and total amount of cost, etc. The costs cover a wide range from several Yuan (unit of RMB, the currency in Mainland China) to several hundreds thousand Yuan. Fig. 1(a) presents the distribution of costs for all the customers, which turns out to obey almost a perfect power-law. Several typical distributions of costs for single customers are also shown in Fig. 1(a), which distribute also in wide ranges and obey roughly power-laws, respectively.

2.2. Similarity network

In a short time duration each customer may be interested in goods whose costs are in a certain range of the wide cost spectrum. Hence, its behavior in the corresponding time can be measured by the distribution of spending. Because of the power-law behavior of the spending distributions, we use the log-binning method. Specifying a special interval of time, e.g., 2001–2003, one can extract all the records in the duration, and calculate subsequently from the records the logarithm