

时序阵发性对信息传播的影响

邓冬梅 朱建 陈端兵 高辉

(电子科技大学计算机科学与工程学院 成都 611731)

摘要 近年来,传播动力学是网络研究的一个热门话题,传统的信息传播都是在静态网络上进行研究的,但现实生活中许多网络是有序时的。到目前为止,已有很多学者针对时序阵发性对信息传播影响做了研究,结果表明对不同的数据集、不同的节点感染方式,阵发性对信息传播呈现的作用是不同的。针对此现象,通过构建 DCW 空模型消除原数据的阵发性特征,分析信息在原数据和空模型数据上的传播情况,以找到阵发性对传播呈现不同影响的原因。

关键词 时序网络,传播,阵发性,空模型

Influence of Bursty on Information Diffusion

DENG Dong-mei ZHU Jian CHEN Duan-bing GAO Hui

(Computer Science and Engineering Institute, University of Electronic Science and Technology, Chengdu 611731, China)

Abstract In recent years, the transmission dynamics is a hot topic in network research, traditional information dissemination researches are based on static network, but many actual networks are temporal. So far, many scholars have researched the impact of bursty on information dissemination, and the results show that bursty presents different roles for information dissemination on different data sets and different nodes infection way. Regarding to this phenomenon, studying the influence of bursty on information dissemination by using DCW null model, and analyzes information spreading on the original data and the null model data, and to find out the reason why bursty shows different effects on the spreading.

Keywords Temporal network, Propagation, Bursty, Null model

1 简介

传播动力学研究是复杂网络领域的一个热门话题。通过模拟生活中信息、疾病等在网络上的传播情况,可以在疾病暴发的初期就找到关键传播点,从而进行预防;也可以在产品推广的初期找到重要用户;还可以预测一则消息最后能传播的规模,等等^[1,2]。

但是,传统的信息传播都是在静态网络上进行研究的,而多数实际的网络是动态的、有时序性的,即为时序网络^[3]。时序网络有许多独特的特征,如阵发性、周期性以及事件发生的等待时间呈现重尾现象^[4-6]等等。近年来,信息传播研究是时序网络领域的一大热点^[7-12],有相当一部分学者研究了阵发性对信息传播的影响,结果显示对不同的数据集、不同的节点感染方式,阵发性呈现的作用不同。

在一部分研究中,采用的节点感染方式为:若感染节点与易感染节点接触,易感染节点变为感染节点。文献[13]中,作者 M. Karsai 等针对节点感染方式为接触即感染的 SI 传播模式,通过构建网络的泊松分布,提出泊松分布与幂律分布的信息传播比值公式: $r = \frac{(\alpha-2)^2}{2(\alpha-1)(\alpha-3)}$,其中 α 是幂律分布指数。用 matlab 工具画出该函数,可知, α 为 0~1 时, r 递增(大于 1); α 为 1~3 时, r 小于 0; α 接近 1 和 3 时, r 为无穷大; α 为 3~3.4 时, r 递减(大于 1); α 大于 3.4 时, r 就小于 1。若 $r < 1$,时序阵发性减缓信息的传播;若 $r > 1$,时序阵发性加快

信息的传播。常见的社交网络幂律指数都介于 3~3.4^[15],因此对这种节点感染方式,时序阵发性一般称呈现的是缓减作用。Petter 认为,这是因为胖尾现象(幂律分布)的存在,导致信息从一个点传到另一个点所需要的平均时间比服从指数型的网络更长^[16]。但与此同时,对不同的网络以及不同的信息传播模式,阵发性可能起到增强作用^[17,18]。

一般情况下,感染个体与易感染个体一接触就将病毒传染给易感染个体的概率较小,通常需要经过多次接触易感染个体才会被感染,即依赖历史型的节点感染方式。在文献[16]中,Petter 对依赖历史型的 SI 传播模型进行研究,表明对这种节点感染方式的传播,阵发性起到了加速信息传播的作用。

到目前为止,阵发性对信息传播呈现不同影响的具体原因还不是很确定。本文主要选取了 Sexual escorts 和 Infectious 两个网络数据集,对依赖历史型的节点感染方式,针对阵发性对不同数据集、不同传播模型和不同的节点感染阈值的影响都做了对比分析,以找出阵发性对不同数据集,不同传播模型呈现不同影响的原因。

2 技术方法

对静态网络,研究人员构建相应的随机网络(ER 等),然后对比分析信息在原网络和随机网络中的传播情况,以研究网络的不同结构特性对信息传播的影响。要研究时序网络特

征对信息传播的影响,也应构建相应的随机模型,称之为空模型(Null Model)。相比静态网络,时序网络有独特的特征,分别有阵发性、事件间相关性、时序相关性、周期性、日常模式等等。因此,要分别研究时序网络不同特征对信息传播的影响,需要构建的空模型就更复杂。文献[13]中介绍了几种常见的空模型,如:DCWB(equal-weight link-sequence shuffled)——消除事件间相关性;DCB(link-sequence shuffled)(打乱原网络拓扑结构);DCW(time-shuffled)——消除事件间相关性和阵发性;D(configuration model)——只保留日常模式。M. Karsai^[13]等通过实验分析,表明事件间相关性对信息的传播影响不大,因此本文主要用 DCW 空模型来消除时序网络的阵发性特征,然后研究信息分别在原时序数据和 DCW 空模型数据中的传播情况,对比得出结论。

DCW 空模型:将原时序网络所有事件时间重新排列,然后重新随机分配给每个事件。DCW 空模型又称 RP(randomly permuted times)空模型^[3],此空模型消除了网络的时序阵发性和事件间相关性特征。

依赖历史的节点感染方式:为每个节点 i 设一个 v_i 标志,表示其对信息的兴趣值。随着节点 i 与感染节点接触的数次增加,它被感染的概率 v_i 增大,但是,如果节点 i 未与已感染节点接触, v_i 随着时间的推移呈指数型下降。 v_i 的变换公式如下:

$$v_i(t) = \sum \exp(-\frac{t-t_e}{\tau_d}) \quad (1)$$

式中, t_e 是节点 i 与感染节点接触的时间, τ_d 是一个延迟时间,是信息的衰减时间(式(1)来源于文献[16])。

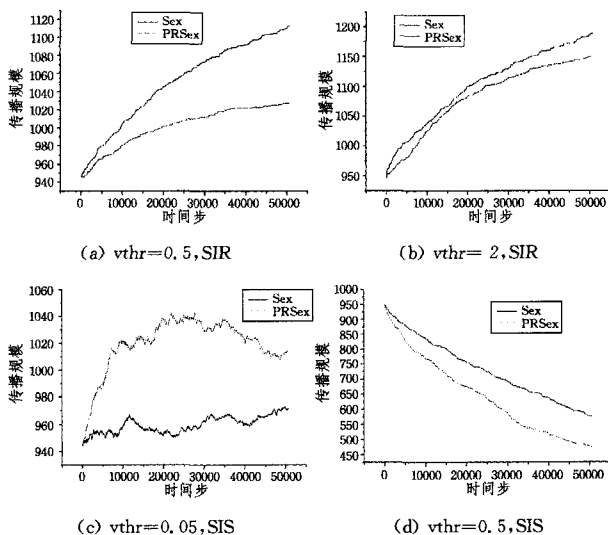
本文主要采用 SIR、SIS 传播模型,节点感染方式为依赖历史型。对 SIR 模型,已感染节点在一定时间后以随机概率变成免疫状态,对 SIS 模型,已感染节点在一定时间后以随机概率恢复成易感染状态。将所有节点对信息的兴趣值 v_i 初始化为 0。节点感染阈值用 v_{thr} 表示,为简单起见,同一次实验,将所有节点的感染阈值 v_{thr} 设为相同。初始时网络中的所有节点为易感染状态,然后选取一个感染节点集,将感染节点状态修改为已感染。将事件按发生时间的先后顺序读入,每读入一个事件,根据情况改变事件目标节点状态。如果当前节点的累积兴趣值 v_i 超过了阈值 v_{thr} ,则节点感染,将其状态修改成已感染,否则利用式(1)修改 v_i 的值。本实验 τ_d 取 10000000, v_{thr} 取 0.05、0.2、0.5、1、1.2、2 等变化值。

3 实验分析

本实验主要选取了两个数据集,都来自于 KONECT(the Koblenz Network Collection)网站。一个是都柏林科学画廊展览(INFECTIOUS;STAY AWAY)期间,人群面对面行为的数据集 Infectious,另一个是性交易网络 1997 年 11 月 8 日到 2000 年 1 月 2 日的 Sexual escorts。数据集描述如表 1 所列,其中 Sexual escorts 数据集的初始感染节点选取了 958 个,Infectious 数据集选取了 20 个。

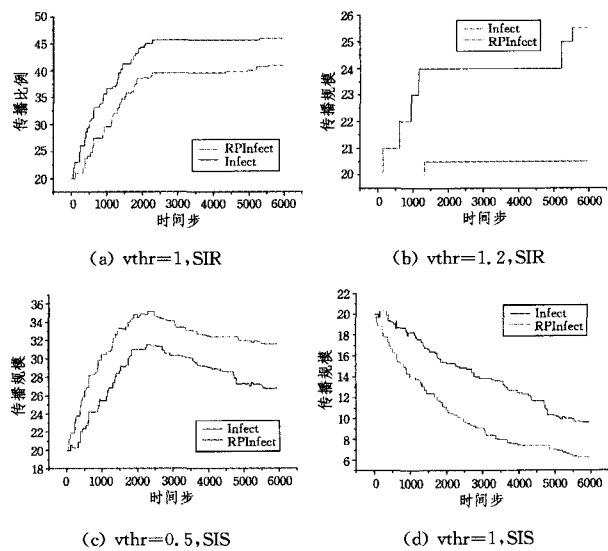
表 1 数据集描述

Name	Type	Nodes	Edges	Description
Sexual escorts	Undirected	16730	50632	(graphs) from 1997. 11. 8—2000. 1. 2
Infectious	Undirected	10972	415912	(graphs) from 2009. 4. 28—2009. 7. 16



其中:(a)为节点感染阈值为 0.5,传播模式为 SIR 的传播对比图;(b)为节点感染阈值为 1,传播模式为 SIR 的传播对比图;(c)为节点感染阈值为 0.5,传播模式为 SIS 的传播对比图;(d)为节点感染阈值为 2,传播模式为 SIS 的传播对比图。

图 1 sexual escorts 数据集,针对不同传播模式和不同节点感染阈值,原始时序数据和 DCW 空模型数据的传播结果对比图



其中:(a)为节点感染阈值为 1,传播模式为 SIR 的传播对比图;(b)为节点感染阈值为 1.2,传播模式为 SIR 的传播对比图;(c)为节点感染阈值为 0.5,传播模式为 SIS 的传播对比图;(d)为节点感染阈值为 1,传播模式为 SIS 的传播对比图。

图 2 Infectious 数据集,针对不同传播模式和不同节点感染阈值,原始时序数据和 DCW 空模型数据的传播结果对比图

实验结果如图 1 和图 2 所示。对比图 1(a)和(b):同一个数据集 Sexual escorts,对传播模型 SIR,当 $v_{thr}=0.5$ 时,阵发性减缓了信息的传播,当 $v_{thr}=2$ 时,阵发性加快了信息的传播。对比图 1(c)和(d):同一个数据集 Sexual escorts,对传播模型 SIS,当 $v_{thr}=0.05$ 时,阵发性减缓了信息的传播,当 $v_{thr}=1$ 时,阵发性加快了信息的传播。对比图 2(a)和(b):同一个数据集 Infectious,对传播模型 SIR,当 $v_{thr}=1$ 时,阵发性减缓了信息的传播,当 $v_{thr}=1.2$ 时,阵发性加快了信息的传播。

对比图 2(c)和(d);同一个数据集 Infectious,同一种传播模型 SIS,当 $v_{hr}=0.5$ 时,阵发性减缓了信息的传播,当 $v_{hr}=1$ 时,阵发性加快了信息的传播。对比图 2(a)和(d),当 $v_{hr}=1$ 时,阵发性对 SIR 传播模型起减缓信息传播作用,对 SIS 传播模型起加速信息传播作用。对比图 1(d)和图 2(c),当 $v_{hr}=0.5$ 时,对 SIS 传播模型,阵发性对 Infectious 数据集起减缓信息传播作用,对 Sexual escorts 数据集起加速信息传播作用。

综上,对同一数据集,同一传播模型,当节点感染率阈值不同时,阵发性呈现出的作用是不一样的;对同一传播模型,同一节点感染率阈值,不同数据集,阵发性呈现出的作用是不一样的;对同一数据集,同一节点感染率阈值,不同传播模型,阵发性呈现出的作用是不一样的;对同一种传播模型,同一节点感染率阈值,不同数据集,阵发性呈现出的作用是不一样的;同一数据集,对同一种传播模型,同一节点感染率阈值,当感染节点规模不同,阵发性呈现出的作用也是不一样的。

结束语 对常见的复杂网络,由于其幂律指数大多分布在 1~3 之间,因此复杂网络中事件发生的平均等待时间就会比均匀网络更长,所以,对接触即感染的传播方式,阵发性起减缓信息传播速度的作用。

对依赖历史的节点感染方式,同一种传播模型,当数据集、已感染节点的规模和节点感染阈值不同时,阵发性体现出的作用都可能会发生变化。比如在传播前期,阵发性对依赖历史的感染方式传播起加速作用,但是当感染节点达到一定规模时,未感染节点的邻居就会有更多已感染节点,此时短时间内易感染节点与已感染节点接触的概率增大,节点就更容易被感染,阵发性就体现不出它的优势,反而退化成了接触即感染的方式,那么阵发性起的作用便是减缓信息传播。

对不同的传播模型,同一个感染阈值,可能在 SIS 模式中阵发性起的是增强作用,在 SIR 模式中起的却是衰减作用。因为一般情况下,SIR 模型中节点的规模很快就会增加,但是 SIS 模型中却不一定,甚至可能在减少。

综上所述,对于依赖历史节点感染方式的传播,当节点感染阈值较低或网络中感染节点规模较大时,就接近接触即感染方式的情况,此时阵发性减缓信息的传播,但是,对节点感染阈值较高或网络中感染节点规模较小的情况,阵发性加速信息的传播。因为对节点感染率与已感染接触次数和时间(依赖历史)相关的情况,短时间内多次发生事件会导致节点与已感染节点接触的次數增加,并且时间相对集中,从而节点感染率就比均匀分布中更大,更容易超过节点感染阈值,因此阵发性对这种感染方式的传播起的是加速作用。而当网络中感染规模大到某种程度时,阵发性就体现不出它的增强作用了,反而因为网络中事件发生的平均等待时间大于均匀网络,使得传播速度变低。但是,对不同的数据集、不同的传播模型,节点感染阈值和感染节点的规模的界限是多少才能使阵发性呈现的作用不同,这还需要进一步的研究。

- [1] 周涛,傅忠谦,等. 复杂网络上传播动力学研究综述[J]. 电子科技大学,中国:自然科学进展,2005,15(5):513-518
- [2] 丁军平,蔡皖东. 面向 P2P 特定信息的传播动力学模型研究[J]. 中国:计算机科学,2011,38(11)
- [3] Petter H, Jari S. Temporal networks[J]. Phys. Rep, 2012, 519: 97-125
- [4] Alexei V, Joao G, Zoltan D, et al. Modeling bursts and heavy tails in human dynamics[J]. Phys. Rev. E 73, 036127, 2006
- [5] 陈琳,刘维奇. 重尾分布族及其关系图[J]. 高校应用数学学报, 2009, 24(2):166-174
- [6] 樊超,郭进利,韩筱璞. 人类行为动力学研究综述[J]. 复杂系统与复杂性科学, 2011, 1(17):1672-3813
- [7] Alessandro M C, Sebastian G. Epidemics scenarios in the "Romantic network"[J]. PLoS ONE, 2012, 7(11):e49009
- [8] Duygu B, Vittoria C, Bruno G, et al. Multiscale mobility networks and the large scale spreading of infectious diseases[J]. PNAS, 2009, 106(51):21484-21489
- [9] Manuel G, Bernahard S. Influence maximization in continuous time diffusion networks[J]. ICML, 2012, 1205-1682; 10750658
- [10] Jiang Z Q, Xie W J, Li M X, et al. Calling patterns in human communication dynamics[J]. PNAS; 2013, 110(5):1600-1605
- [11] Adrien G, Hakim H, Cecile F. Predicting the temporal dynamics of information diffusion in social networks [J]. arXiv: 1302. 5235v2 [cs. SI], 2013
- [12] Petter H. Eepidemiologically optimal static networks from temporal network data[J]. arXiv:1302 [physics. soc-ph], 2013
- [13] Karsai M, Kivela M, Pan R K, et al. Small but slow world: how network topology and burstiness slow down spreading [J]. Phys. Rev. E 83, 025102(R), 2011
- [14] Alexei V, Balazs R, Ras L, et al. Impact of non-Poisson activity patterns on spreading processes [J]. Phys. Rev. Lett. 98, 158702, 2007
- [15] 郭进利,汪丽娜. 幂律指数在 1 与 3 之间的一类无标度网络[J]. 中国:物理学报, 2007, 56(10):5635-05
- [16] Taro T, Naoki M, Petter H. Bursty communication patterns facilitate spreading in a threshold-based epidemic dynamics [J]. nlin. physics. soc-ph, 2012
- [17] Crocha L E, Liljeros F, HOIME P. Simulated epidemics in an empirical spatiotemporal network of 50,185 sexual contacts[J]. PLoSComput. Biol. 7, e1001109, 2011
- [18] Karimi F, Holme P, unpublished, 2012
- [19] Lorenzo I, Ciro C, Wouter V. What's in a crowd? Analysis of face-to-face behavioral networks [J]. J. Theor. Biol. , 2011 (271):166-180
- [20] Seth M, Chen Guang-zhu, Jure L. Information diffusion and external influence in networks [J]. Stanford University: ACM, 2012:33-41