

面向论坛网络的观点传播与舆情形成模型

徐会杰^{1,2} 蔡皖东¹ 陈桂茸¹ 王剑平¹

(西北工业大学计算机学院 西安 710072)¹ (河南科技大学管理学院 洛阳 471003)²

摘 要 根据网络论坛的信息传播规律和用户交互特性,提出了一个基于节点影响力的离散时间观点传播与舆情形成模型。在该模型中,节点具有自我影响力和观点值,并且节点影响力被定义为网络中边的权值,节点在每一个周期内重新分配权值并更新它们的观点。节点观点收敛于某一定值时,形成网络舆情。仿真结果表明,所提模型可以准确地体现节点观点的动态交互特性和收敛特征,并可为舆情传播趋势的预测研究提供依据。

关键词 论坛网络,观点传播,舆情形成,影响力

中图分类号 TP391.7 **文献标识码** A

Opinion Propagation and Public Opinion Formation Model for Forum Networks

XU Hui-jie^{1,2} CAI Wan-dong¹ CHEN Gui-rong¹ WANG Jian-ping¹

(School of Computer Science, Northwestern Polytechnical University, Xi'an 710072, China)¹

(School of Management, Henan University of Science & Technology, Luoyang 471003, China)²

Abstract According to the information dissemination and the interaction characteristics of users for online forums, a discrete time model was proposed for opinion propagation and public opinion formation based on the node influence. In the model, every node has the opinion and the influence which is defined as the weight of network, and the nodes will redistribute the weights and update their opinions during the each period. When the opinion of nodes converges to a certain value, the public opinion will be formed. The simulation results show that this model can emulate accurately the characteristics of interaction and convergence. Besides, the results can provide the reference for the prediction of public opinion propagation.

Keywords Forum networks, Opinion propagation, Public opinion formation, Influence

1 引言

随着 Web2.0 技术的成熟和广泛应用,以多元化、交互式为特点的网络论坛已成为网络舆情的重要来源^[1,2]。研究论坛网络中用户观点的传播与舆情的形成和发展规律,对于网络舆情的识别与监测具有重要意义。

传统的观点传播与舆情形成模型主要采用 Sznajd 模型^[3]和 French-DeGroot 模型^[4,5],后来的很多模型都是这两种模型的扩展。Sznajd 模型是类似于 Ising 模型的一维链上的舆情演化模型,认为节点是排列在一维空间上的点,并假定节点可以选择其中小数量的离散观点。该模型由于对选举过程作出了较好的解释而受到广泛关注,并被迅速推广到小世界网络^[6,7]和无标度网络^[8]。进一步的工作还包括 Schulze^[9,10]增加相互作用的距离和约束以及考虑扩散效应的影响,Sznajd^[11]等增加观点的数量以及 Sabatelli^[12]等加入噪音的影响等;French-DeGroot 模型认为节点的观点可以在一个任意维度和结构的空間上延伸。观点在吸引力驱使下实时演化,即节点在观点空间中变换立场以朝向那些其他节点观点

已经集中的领域。假定这种动力学变化是线性的,DeGroot 又引入了观点收敛模型,其中下一时刻的观点是原来观点的线性组合。形成舆情所需的条件在稳定性定理中被阐明^[5,13]。该模型能够比较真实地反映现实网络中节点观点的倾向问题,已被应用于多种模型^[14-17]。

Sznajd 模型和 French-DeGroot 模型虽然在一定程度上表达了观点传播和舆情形成的主要特征,但采用上述模型进行分析将存在以下问题:①Sznajd 模型认为人们的观点只有两种:支持和反对,不能真实反映节点态度的模糊性和连续性;②French-DeGroot 模型假定网络中边的权值是不随时间变化的定值,不能真实反映节点间联系的亲密程度和时变特性。

为了解决以上两个问题,本文在 French-DeGroot 模型的基础上,提出了基于节点影响力的离散时间观点传播和舆情形成模型。该模型充分考虑了网络中节点观点变化的连续性,以及节点间不同的连接强度及其时变特性。实验表明,该模型能够比较真实地再现论坛网络中观点的传播和舆情形成过程,非常适合研究论坛网络舆情的形成与发展规律。

到稿日期:2012-07-12 返修日期:2012-11-30 本文受国家教育部博士点基金(200806990030),西北工业大学基础研究基金(JC201149),2012 年度洛阳市社会科学重点规划项目(2012A043)资助。

徐会杰(1980—),男,博士生,主要研究方向为网络信息安全与信息对抗,E-mail:xhj004@gmail.com;蔡皖东(1955—),男,教授,博士生导师,主要研究方向为分布式计算、网络信息安全与对抗。

2 论坛网络构建

定义 1 由网络论坛中的节点及节点间相互连通的边构成的网络称为论坛网络。

假设 1 论坛网络是一个封闭的世界, 舆情在论坛网络中产生, 并仅在论坛网络中传播。

假设 2 论坛网络中的节点数是不变的, 每一个节点在 t 时刻都有一个用实数表示的观点值 $x(t)$ 。

定义 2 论坛网络为三元组: $G=(V, E, W)$, 其中, V 是论坛中代表用户的节点集, $|V|=n$ 为网络中的节点数; E 是连接所有节点的无向边组成的集合, 代表节点观点可能的传播路径; $W=[w_{ij}]$ 是边的权值矩阵, 在本文中定义为节点 i 对其存在连接的各节点的影响力而构成的影响力矩阵。

在论坛网络中, 节点间的连接因它们观点的不同具有不同的强度和内涵, 并且是随时间动态变化的。为了描述这种特征, 假设 t 时刻网络中节点 i 和 j 的观点分别为 $x_i(t), x_j(t)$, 则节点间的观点距离 d_{ij} 可表示为^[17]:

$$d_{ij} = |x_i(t) - x_j(t)|, \forall i, j \in N, i \neq j \quad (1)$$

$$\text{令 } w_{ij} = \frac{1}{d_{ij}} = \frac{1}{\max(d, |x_i(t) - x_j(t)|)} \quad (2)$$

式中, d 是一个非常小的正数, 以代替 $x_i(t) = x_j(t)$ 时分母为零的情况。

式(2)表明两个节点间的观点距离越接近, 彼此间的影响力就越大。这是很符合实际的, 社会中两个人的关系越亲密, 当其中一个人遇到问题时更愿意从亲密的朋友那里寻找建议, 该朋友对他的影响力也就越大。

为简化问题, 令节点 i 与所有存在直接连通的节点的权值之和为 1, 即 $\sum_{j=1}^n w_{ij} = 1$ 。则对于 $t \geq 0$ ^[16], 存在直接连通的两节点在 $t+1$ 时刻的影响力值 w_{ij}^{t+1} 可表示为:

$$w_{ij}^{t+1} = \frac{w_{ij}^t}{w_i^t + \sum_{\substack{j=1, j \neq i \\ (i, j) \in V}} w_{ij}^t} \times 1 \quad (3)$$

式中, w_i^t 为节点 i 在 t 时刻自己对自己的影响力值。按照归一化准则, 在 $t+1$ 时刻, $w_i^{t+1} = 1 - \sum_{\substack{j=1, j \neq i \\ (i, j) \in V}} w_{ij}^{t+1}$ 。

3 观点传播和舆情形成模型建立

定义 3 舆情系统 S 为五元组: $S = \{n, T, x(t), W, G_t(V, E, W)\}$, 其中, n 表示系统中的节点个数; $T = \{0, 1, 2, \dots\}$ 是系统中离散的时间点; $x(t) = [x_1(t), x_2(t), \dots, x_n(t)]$ 表示 t 时刻系统的观点剖面; $W = [w_{ij}]$ 是影响力矩阵; $G_t(V, E, W)$ 是 t 时刻系统的拓扑结构图。

3.1 观点传播和舆情形成过程

网络论坛是一种不受身份限制、面对所有社会公众开放的公共话语空间。在这个空间中, 论坛用户可以通过发帖或跟帖的形式来表达个人意见、参与社会公共事务讨论。

考虑一个网络论坛中持有不同初始观点值的用户, 用户间的交互首先是以持某一观点的用户提出一个包含标题(title)的主贴(thread)展开的^[18]。如果其他用户(可以是提出主贴用户)对发帖者的观点感兴趣, 则可围绕该标题通过发一个或多个包含相应内容的帖子(post)展开讨论, 讨论的过程即是用户观点彼此交互、传播的过程, 如图 1 所示。这样, 论坛

中参与讨论的用户在外在环境的影响下, 不断更新自己的观点值, 随着时间的推移, 各种不同的用户观点值逐渐彼此相互接近直至趋于一致, 形成网络舆情。

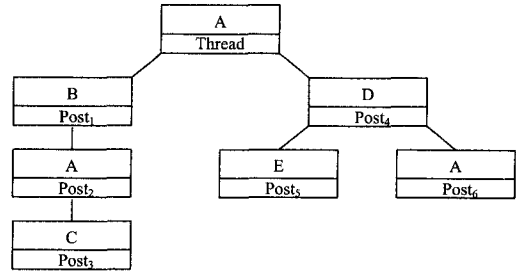


图 1 论坛 Thread 组成图

上述舆情的形成过程是一个用户观点不断递归进化的过程。在这个过程中, 用户既不简单分享, 也不随意丢弃其他人的观点, 而是在每一离散的时间里不断考虑他们的观点, 在一定程度上形成自己的观点^[19]。本文通过在用户邻居的观点值上放置不同的权值来描述外部信息源对用户观点的影响, 详细的建模见第 3.2 节。

3.2 舆情形成与传播模型

基于以上分析, 可建立观点传播和舆情形成模型, 描述论坛网络中人群观点随时间变化的趋势和网络舆情的形成过程。设节点 $i \in V$ 在 $t \in T$ 时刻的观点为 $x_i(t) \in x(t)$, 则节点 i 在 $t+1$ 时刻的观点 $x_i(t+1)$ 可表示为^[20]:

$$x_i(t+1) = \sum_{j=1}^n w_{ij}^{t+1} x_j(t) \quad (4)$$

式中, w_{ij}^{t+1} 为 $t+1$ 时刻节点 j 对节点 i 的影响力值, 在 $G_{t+1}(V, E, W)$ 中表示为节点 i 和 j 间边的权值。

节点间的联系有连接和断开两种状态, 处于连接状态的两个节点间可以进行信息交互; 而处于断开状态的两个节点则无法相互影响。用示性函数 $\delta_{ij}(t+1)$ 表示节点 i 和 j 间的状态为:

$$\delta_{ij}(t+1) = \begin{cases} 1, & i \text{ 和 } j \text{ 处于连接状态} \\ 0, & i \text{ 和 } j \text{ 处于断开状态} \end{cases} \quad (5)$$

由式(4)、式(5)得, S 在 $t+1$ 时刻的观点剖面 $x(t+1)$ 为:

$$x(t+1) = W \delta_{ij}(t+1) w_{ij}^{t+1} x(t) \quad (6)$$

式中, W 是一个由元素 w_{ij} 组成的 $n \times n$ 阶行随机矩阵, 对于所有的 $i, j \in V, w_{ij} \in [0, 1]$ 。

如此, 在论坛网络中, 随着时间的推移, 每一节点总是不断地改变与其存在直接连通的节点的影响力, 实时更新对自身的影响力, 并根据该时刻对存在直接连通的各节点的影响力及上一时刻各节点的观点来更新自己的观点。网络各节点观点的演化实质即是基于影响力矩阵的节点观点的迭代过程。

为了使式(6)既包括网络的拓扑信息, 又包括网络中各节点之间相互影响的信息, 引入 W 的 Laplacian 矩阵表示。

定义 4 R 的 Laplacian 矩阵 L 为:

$$L = \text{diag}(d) - A \quad (7)$$

式中, $\text{diag}(d)$ 为 $G_t(V, E, W)$ 的对角矩阵, $d_i = \sum_{j \in V, j \neq i} w_{ij}$; A 为 $G_t(V, E, W)$ 的邻接矩阵, $a_{ij} = \begin{cases} w_{ij}, & i \neq j \\ 0, & \text{其他} \end{cases}$ 。

则式(6)可转化为:

$$x(t+1)=[I-L\delta_{ij}(t+1)w_{ij}^{t+1}]x(t) \quad (8)$$

式中, I 为单位对角矩阵。

定义 5 对于舆情系统 $S=[S[1], \dots, S[n]]$, 如果对任意的节点 i 和 j , 满足 $|x_i(t) - x_j(t)| < \epsilon$, 则认为舆情系统 S 收敛, 即式(8)渐进收敛于某一定值。

其中, ϵ 是一个很小的实数。

如定义 5 所述, 当时间 $t \rightarrow \infty$ 时, 论坛网络中的节点观点在外在信息源的作用下发生改变。如果在某一时刻 t^* , 所有节点观点收敛于某一定值 $S(t^*)$, 就称论坛网络中节点观点达到完全一致或渐进一致, 从而形成网络舆情。

4 仿真分析

4.1 仿真环境和仿真步骤

本文的仿真平台使用一台个人计算机, CPU 为 Pentium (R) Dual-core E5300, 主频为 2.60GHz, 内存为 1GB, Windows XP 操作系统。采用 UCINET6.0 软件作为论坛网络生成与分析工具, 以 matlab 7.0 对论坛网络的观点传播与舆情形成进行仿真。仿真步骤如下:

(1) 通过 UCINET6.0 生成 n 个节点的论坛网络, 为每一个节点随机分配 $x_i(t) \in [1, 5]$ 的初始观点值和 $w_{ii} \in [0, 1]$ 的自身对自身的初始影响力值。仿真时间长度为 t_N , 初始时刻 $t=0$ 。

(2) $t=t+1$ 时, 根据式(1)一(3)计算每一个节点 i 对自身及与其存在连接的节点 j 的影响力(边的权值), 构造影响力矩阵 W_{ij}^{t+1} ($d=10^{-2}$), 并以此按照式(8)计算 $t+1$ 时刻系统的观点剖面值。

(3) $t=t^*$ 时, 如果对任意的节点 i 和 j , 其观点 $|x_i(t^*) - x_j(t^*)| < \epsilon$ ($\epsilon=10^{-3}$), 则仿真结束; 否则, 重复步骤(2)和步骤(3)。

4.2 仿真分析

为了验证所提出模型在描述观点传播和舆情形成方面的有效性, 实验中取 4 个不同规模的论坛网络进行验证分析。依据仿真步骤, 对随机生成的 4 个论坛网络进行仿真, 可以得到如图 2—图 5 所示的观点传播与舆情形成过程图, 图中 Y 轴代表各节点的观点值, X 轴代表迭代次数。

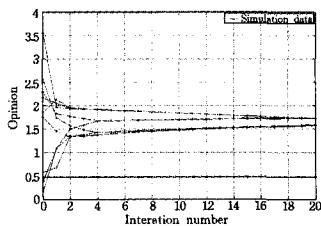


图 2 $n=10$, 观点传播与舆情形成过程

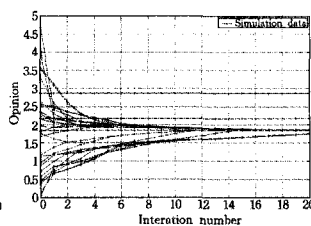


图 3 $n=30$, 观点传播与舆情形成过程

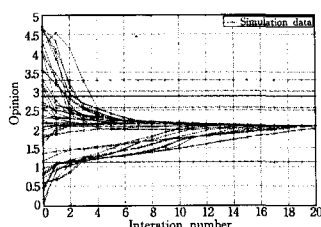


图 4 $n=50$, 观点传播与舆情形成过程

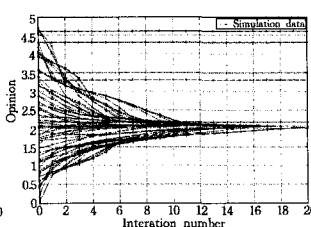


图 5 $n=70$, 观点传播与舆情形成过程

(1) 无论网络规模如何变化, 模型均能收敛。即论坛网络中存在直接或间接连通的节点通过相互作用, 各自持有的观点逐渐发生改变, 最终收敛于某一定值。在这个过程中, 观点传播在起始几个周期变化梯度较大, 随着时间推移变化趋缓, 逐渐趋向一个稳定值, 表明系统从一非平衡态趋向平衡态。这与现实网络论坛中观点传播的趋势相仿, 最初众人观点不一, 随着彼此讨论、相互影响, 最终观点趋于一致, 形成舆情。

(2) 从图 2—图 4 中可以发现, 观点最终形成两种不同的簇, 即论坛网络中的节点最终形成两种持不同观点的群。这进一步验证了论坛网络具有的社区结构特性, 即具有较高观点相似度值的节点可以划归于同一类型的社区, 它们之间存在较多的连接; 而社区与社区之间没有或有很少连接。

(3) 图 2—图 5 中表现为通过较短的观点传播周期达到最终收敛值的节点, 对照网络图可以发现这类节点存在更多的邻居节点。这与现实网络论坛中的用户相仿, 即邻居较多的用户因有更多的机会参与交互, 其持有的观点可以更易影响到其他用户, 同时也更易受到其他用户的影响, 这些用户的观点状态趋于一致的速度更快。

(4) 孤立节点因不参与交互而不会影响到其他节点和受其他节点的影响, 其观点状态不会发生改变(在图 2—图 5 中变化趋势为直线), 在现实网络论坛中这类节点对应于论坛中只浏览帖子而不发帖的用户。

(5) 对比图 2—图 5, 发现 $n=70$ 的论坛网络中节点观点的收敛速度快于其他 3 个网络, 说明网络规模并非影响网络舆情形成的主要因素, 网络规模大时, 舆情形成速度可能更快。

结束语 在对论坛网络中用户观点的传播与舆情的形成和发展规律的研究中, 本文提出了一个基于节点影响力的离散时间观点传播与舆情形成模型。仿真实验表明, 该模型与实际网络论坛中观点传播与舆情形成过程相符, 且可以得到论坛网络中各离散时间点的观点剖面和其他统计特性。

将来的工作主要包括: (1) 将本文提出的模型与论坛网络结构的时变特性相结合, 以更准确地反映现实网络论坛的动态变化特性以及与此相关联的观点传播与舆情形成规律; (2) 量化研究实际网络论坛中用户的影响力和观点, 并探讨基于量化数据的观点传播与舆情形成模型的构建, 以为网络舆情的预测研究提供必要的依据。

参考文献

- [1] 吴承荣. 骨干通道上的网络论坛通信信息监测和分析的关键技术研究[D]. 上海: 复旦大学, 2011
- [2] 刘颖, 李欲晓. 网络舆情传播特征分析[J]. 北京邮电大学学报: 社会科学版, 2011, 13(4): 1-6
- [3] Sznajd-Weron K, Sznajd J. Opinion evolution in closed community[J]. International Journal of Modern Physics C, 2000, 11(6): 1157-1165
- [4] Frech J R P. A formal theory of social power[J]. Psychological Review, 1956, 63: 181-194
- [5] DeGroot M. Reaching a consensus[J]. Journal of the American Statistical Association, 1974, 69: 118-121
- [6] Elgazzar S. Applications of small-world networks to some socio-economic systems[J]. Physica A, 2003, 324(1/2): 402-407

(下转第 176 页)

Aschoal_{SAE}共有28个AE,其中有9个AE参与了HT-TP的响应。演化图中只记录了参与演化的AE的情况,这样能集中反映与演化有关AE的相关情况,更有利于对软件本次运行的分析。图2分别利用本文提出的软件可信演化方法和基于光谱的软件演化方法描绘了Aschoal_{SAE}软件在实验中的数据。

从图2(a)和(b)可以看出,整个A和B区域都被深色填充着,这说明在此阶段中AE($i=1\sim 9$)中的Component不断进行变化,不断地增加和删除。同时可以看到,A区域中软件可信性呈上升趋势,这说明此时软件的自配置采取的演化手段有利于软件可信性的保持和增长。而B区域中,软件可信性在实施自配置操作后则发生了软件可信性降低的结果,这说明配置策略选择不当,需要着重观察此时进行演化的AE自配置策略和演化方法,以便对其进行调整。但是从图2(c)中只能看到A和B两个区域发生了演化,演化频率比较高,不能说明演化与软件可信性之间的关系,不能对演化后软件可信性的相关情况进行控制,无法对演化策略进行指导。

结束语 本文通过对软件行为的分析和对软件行为历史数据的刻画,选取软件可信演化的策略,使软件可信性保持在一个稳定且局部非降的水平上,满足用户对软件可信性的需求。使用软件自律化方法使软件中的组件具有一定的自我管理的能力,通过改变组件之间的关系进行演化操作,以软件可信演化历史反映软件可信演化对软件可信性的影响,帮助软件可信演化选择适当策略,最终在无人工干预或少人工干预情况下完成软件可信演化,以达到软件可信性的保持和增长。

参 考 文 献

- [1] 金海,李克秋.可信计算[J].中国计算机学会通讯,2010,6(2):6-7
- [2] 王怀民,尹刚.网络时代的软件可信演化[J].中国计算机学会通讯,2010,6(2):28-36
- [3] 曾晋,孙海龙,刘旭,等.基于服务组合的可信软件动态演化机制[J].软件学报,2010,21(2):261-276
- [4] 李季,刘春梅.基于本体的可信软件演化框架模型[J].计算机应用研究,2010,27(12):4551-4554
- [5] Oreizy P, Medvidovic N, Taylor R. Architecturebased runtime software evolution[C]//Proceedings of the International Conference on Software Engineering. Kyoto, Japan, April 1998:177-186
- [6] Oreizy P, Medvidovic N, Taylor R N. Runtime software adaptation: framework, approaches, and styles[C]//Proceedings of the International Conference on Software Engineering. ACM Press, 2008:899-910
- [7] Gall H, Jazayeri M, Riva C. Visualizing software release histories: The use of color and third dimension[C]//Proceedings of the International Conference on Software Maintenance. IEEE Press, 1999:99-108
- [8] Lanza M. The evolution matrix: recovering software evolution using software visualization techniques[C]//Proceedings of the International Workshop on Principles of Software Evolution. ACM Press, 2001:37-42
- [9] Wu J, Holt R, Hassan A. Exploring software evolution using spectrographs[C]//Proceedings of 11th Working Conference on Reverse Engineering. IEEE Press, 2004:80-89
- [10] Wu J, Spitzer C, Hassan A, et al. Evolution spectrographs: visualizing punctuated change in software evolution[C]//Proceedings of 7th International Workshop on Principles of Software Evolution. ACM Press, 2004:57-66
- [11] Grba T, Ducasse S. Modeling history to analyze software evolution[J]. Journal on Software Maintenance and Evolution: Research and Practice, 2006, 18(3):207-236
- [12] Robbes R, Lanza M. A change-based approach to software evolution[J]. Electronic Notes in Theoretical Computer Science, 2007, 166:93-109
- [13] Berger R L. A necessary and sufficient condition for reaching a consensus using DeGroot's method[J]. Assoc. , 1981, 76(374):415-418
- [14] DeMarzo P, Vayanos D, Zwiebel J. Persuasion Bias, Social Influence, and Unidimensional Opinions[J]. Quarterly Journal of Economics, 2003, 118(3):909-968
- [15] Donetti L, Hurtado P I, Munoz M A. Entangled networks, synchronization, and optimal network topology[J]. Physical Review Letters, 2005, 95(18):1-4
- [16] Golub B, Jackson M O. Naive learning in social networks and the wisdom of crowds[J]. Amer. Econ. J. Microecon, 2010, 2(1):112-149
- [17] Pan Z. Trust, influence, and convergence of behavior in social networks[J]. Mathematical Social Sciences, 2010, 60(1):69-78
- [18] 陈友,程学旗,杨森.面向网络论坛的高质量主题发现[J].软件学报,2011,22(8):1785-1804
- [19] Hegselmann, Rainer, Krause U. Opinion dynamics and bounded confidence: models, analysis, and simulation[J]. Journal of Artificial Sciences and Social Simulation, 2002, 5(3):1-33
- [20] Slanina F. Dynamical phase transitions in Hegselmann-Krause model of opinion dynamics and Consensus[J]. European Physical Journal B-Condensed Mater, 2011, 79(1):99

(上接第152页)

- [7] He M F, Li B, Luo L D. Sznajd model with "social temperature" and defender or small-world networks[J]. Int. J. Mod. Phys. C, 2004, 15(7):997-1003
- [8] Aleksiejuk A, Holyst J A, Stauffer D. Ferromagnetic phase transition in Barabasi-Albert networks[J]. Physica A, 2002, 310(1/2):260-266
- [9] Schulze C. Long-range interactions in Sznajd consensus [J]. Physica A, 2003, 324(3/4):717-722
- [10] Schulze C. Advertising in the Sznajd marketing model[J]. Int. J. Mod. Phys. C, 2003, 14(1):95-98
- [11] Sznajd W K, Sznajd J. Who is left, who is right? [J]. Physica A, 2003, 351(2-4):593-604
- [12] Sabatelli L, Richmond P. Non-monotonic spontaneous magnetization in a Sznajd-like consensus model[J]. Physica A, 2004, 334(25):274-280