

# 基于决策论元胞自动机的网络舆情传播研究

黎永壹

(钦州学院 电子与信息工程学院, 广西 钦州 535099)

**摘 要:** 在舆情传播研究中, 传统元胞自动机采用多数演变规则, 以多数情感代表倾向度, 忽略了舆情个体间的相互影响, 会导致错误的预测结果, 为此, 将决策论结合元胞自动机应用于舆情传播问题中. 首先介绍元胞自动机和决策论的基本概念及特点; 然后建立基于决策论的元胞自动机模型, 以相对收益模型作为演变规则, 设计元胞状态演变的遍历算法; 最后通过实例和模拟仿真表明算法可实现舆情传播倾向度趋向明显, 结果验证了算法的有效性和可行性.

**关键词:** 元胞自动机; 舆情传播; 决策论; 竞争收益; 演变规则

**中图分类号:** TP181

**文献标识码:** A

**文章编号:** 1000—7180(2016)06—0126—04

## Research on Network Public Opinion Spread Based on Decision Theory Cellular Automata

LI Yong-yi

(School of Electronics and Information Engineering, Qinzhou University, Qinzhou 535099, China)

**Abstract:** In the study of public opinion spread, the evolution rules of the traditional cellular automaton are the majority principle which provides most emotional represents the tendency, but ignoring public opinion between individuals influence each other, it can lead to wrong prediction results, therefore, this paper combines decision theory and cellular automata and applied to the public opinion spread. First of all, introduce the basic concept and characteristics of cellular automata and decision theory. Cellular automata model based on Decision theory is set up, the algorithm of cellular state evolution is designed with relative income model as the evolution rules. Finally examples and simulation indicates that algorithm can realize public opinion spread tendency obviously, the results verify the feasibility and effectiveness of the algorithm.

**Key words:** cellular automata; public opinion spread; decision theory; competitive benefits; evolution rules

### 1 引言

在网络发达的环境下, 对网络舆情的研究有以下三个方面, 即舆情案例数据库技术、舆情事件分析和舆情倾向性分析. 舆情案例数据库技术包括数据的采集、统计、存储和预处理等<sup>[1]</sup>; 舆情事件分析包括如何确定主题及其聚类分析等, 属于统计和数据挖掘的范畴<sup>[2]</sup>, 舆情倾向性分析根据网上言论的有用信息, 围绕某些有明显倾向性的信息, 研究网络舆

情演化过程, 建立舆情传播的模型, 并作为预测的依据, 本文主要针对舆情传播的问题进行研究.

网络舆情传播的形式有很多种, 如论坛、博客和手机微信等, 多种网络交流工具的发展决定了网络舆情传播的多元化发展<sup>[3]</sup>, 网络舆情已经成为社会舆情非常关键的一个方面, 网络舆情通常包括主体和本体, 主体为网民, 本体为网民的意愿<sup>[4]</sup>, 与常见的社会舆情相比, 网络舆情存在开放性、虚拟性等特点, 舆情的发展随着网络规模的发展对社会的影响

**收稿日期:** 2015—08—07; **修回日期:** 2015—09—30

**基金项目:** 广东省科技计划项目(2012b010100019); 中央大学的基础研究基金(2013 zz0047); 广西教育科学“十二五”规划课题(2015C435); 广西高校科研项目(ZD2014137)

日益明显,目前的研究主要集中于元胞自动机模型<sup>[5]</sup>、社会网络<sup>[6]</sup>或因特网概率统计模型<sup>[7]</sup>和隐马尔可夫模型<sup>[8]</sup>三个方面.舆情的结构与元胞结构的及其关系有一定的相似性,因此将元胞自动机的变化过程应用于舆情传播是可行的.

一般的元胞自动机都采用多数规则的演变方法,如文献<sup>[9]</sup>所示,但它们往往忽略了邻居状态对舆情传播的影响.为了解决上述问题,本文提出基于决策论元胞自动机的网络舆情传播模型,建立了竞争收益的模型.在 MATLAB 环境下对模型进行仿真,实验结果表明模型具有很高的合理性和可行性.

2 基于决策论的元胞自动机模型

2.1 基于决策论的舆情传播元胞自动机模型

根据舆情传播的特点,以元胞自动机系统为基础,结合决策论的基本概念,构建一个基于决策论的舆情传播元胞自动机模型.

2.1.1 元胞形态

表示网络舆情主题 C,为每个主题定义两种倾向性态度  $A=\{\text{支持}, \text{反对}\}$ ,利用元胞状态表示舆情主体 C 的态度,即  $A=\{1, -1\}$ .

2.1.2 元胞空间

将舆情传播结构描述为虚拟空间,利用一个二维正方网格结构,每个网格表示一个主体,应用八邻域,  $(i, j)$  表示空间中心,定义网络空间的边界条是周期型,  $S_{(ij)}$  表示空间点  $(i, j)$  的情感倾向度.

2.1.3 邻居

决策论在元胞自动机里的邻居即为两个参与者,中心元胞  $(i, j)$  的邻居可表示为  $(i, j-1), (i, j+1), (i+1, j), (i+1, j+1), (i+1, j-1), (i-1, j), (i-1, j+1), (i-1, j-1)$ ,如图 1 所示.

|              |            |              |
|--------------|------------|--------------|
| $(i-1, j-1)$ | $(i-1, j)$ | $(i-1, j+1)$ |
| $(i, j-1)$   | $(i, j)$   | $(i, j+1)$   |
| $(i+1, j-1)$ | $(i+1, j)$ | $(i+1, j+1)$ |

图 1 中心元胞邻域集合

2.1.4 演变规则

用一个元胞代表一个个体,随机选择一个元胞  $(i, j)$ ,将其看作是中心元胞并作为研究对象,元胞状态的演变规则主要是以决策论的收益为根据,即某个元胞对某个舆情保持一个态度,该元胞和相邻元胞竞争后,对自身以及邻居元胞产生一定影响,根据决策论,通过判断竞争收益,决定个体的情感倾

向,采用一个二维表对中心元胞与任一邻居元胞竞争的结果进行描述.表 1 描述的是中心元胞与其邻域竞争产生各自的收益.

表 1 中心元胞与某一邻居对某一舆情竞争后的收益情况

| 中心态度 \ 相邻态度<br>收益 | 支持       |          | 反对       |          |
|-------------------|----------|----------|----------|----------|
|                   | 支持       | 反对       | 支持       | 反对       |
| 支持                | $E_1(C)$ | $E_1(N)$ | $E_2(C)$ | $E_2(N)$ |
| 反对                | $E_3(C)$ | $E_3(N)$ | $E_4(C)$ | $E_4(N)$ |

当中心元胞和邻居对某个舆情都持支持态度时,  $E_1(C)$  表示中心元胞的收益,  $E_1(N)$  表示邻域元胞的收益;当中心元胞对某个舆情持支持态度,而邻居对该舆情持反对态度时,  $E_2(C)$  表示中心元胞的收益,  $E_2(N)$  表示邻居的收益;当中心元胞对某个舆情持反对态度,而邻居对该舆情持支持态度时,  $E_3(C)$  表示中心元胞的收益,  $E_3(N)$  表示邻居的收益;当中心元胞和邻居对某个舆情都持反对态度时,  $E_4(C)$  表示中心元胞的收益,  $E_4(N)$  表示邻域元胞的收益;

中心元胞  $(i, j)$  与每一个邻居元胞进行一次竞争后收益的表达式如式(1) 所示.

$$E_N(C) = \begin{cases} E_1(C) & (S_{(i,j)} = 1, S_N = 1) \\ E_2(C) & (S_{(i,j)} = 1, S_N = -1) \\ E_3(C) & (S_{(i,j)} = -1, S_N = 1) \\ E_4(C) & (S_{(i,j)} = -1, S_N = -1) \end{cases} \quad (1)$$

中心元胞  $(i, j)$  与所有邻居元胞进行一次竞争后的平均收益如式(2) 所示.

$$W(i, j) = \frac{1}{8} \sum E_N(C) \quad (2)$$

式中  $N$  表示中心元胞的某个邻居,即  $N \in \{(i, j-1), (i, j+1), (i+1, j), (i+1, j+1), (i+1, j-1), (i-1, j), (i-1, j+1), (i-1, j-1)\}$

当中心元胞、相邻元胞均竞争一次后,即可获取该竞争后所有元胞的平均收益,同事求出中心元胞加 8 邻居的总平均收益,如式(3) 所示.

$$AE(i, j) = \frac{1}{9} \left( \sum_{i=-1}^{+1} \sum_{j=-1}^{+1} W(i, j) \right) \quad (3)$$

整个平面内所有元胞的总平均收益即为  $AE(n, n)$ ,并以此与  $W(i, j)$  作比较,作为元胞下一个时段状态值改变的主要依据,根据最大可能性决策准则,即最大收益原则,如果平均收益大于总平均收益,元胞  $(i, j)$  在下一时刻改变其状态,如果平均收益等于总平均收益,在下一时刻清空元胞  $(i, j)$ ,否则状态不改变.根据元胞自动机的基本构成,演变

规则可表示为式(4).

$$S_{(i,j)}^{t+1} = \begin{cases} -S_{(i,j)}^t, & (W(i,j) > AE(n,n)) \\ \emptyset, & (W(i,j) = AE(n,n)) \\ S_{(i,j)}^t, & (W(i,j) < AE(n,n)) \end{cases} \quad (4)$$

为了衡量收益的大小,需要对每一次竞争后的收益进行量化,假设竞争者中赞成者的比例为  $\rho$ ,反对者的比例为  $1 - \rho$ ,则赞成者的收益可描述成:

$$PC = \rho R + (1 - \rho)S \quad (5)$$

反对者的收益可描述成:

$$PD = \rho T + (1 - \rho)P^{[10-11]} \quad (6)$$

以囚徒困境为例,依据 Nowak 和 May 所使用的竞争参数  $R = 1, T = b(1 < b < 2), P = S = 0^{[12]}$  代入上式.

确定收益值的方法有两种,即绝对收益值和相对收益值. 绝对收益值是由元胞状态决定,与邻居状态无关;相对收益值是由元胞及其邻居的状态比较决定. 表 2、表 3 分别表示绝对收益和相对收益的赋值情况,对应的赋值公式分别为式(7) 和式(8).

表 2 中心元胞与邻居元胞对某一舆情竞争后的绝对收益值

| 收益<br>中心态度 \ 相邻态度 |  | 支持 |    | 反对 |    |
|-------------------|--|----|----|----|----|
|                   |  | 支持 | 反对 | 支持 | 反对 |
| 支持                |  | PC | PC | PC | PD |
| 反对                |  | PD | PC | PD | PD |

表 3 中心元胞与邻居元胞对某一舆情竞争后的相对收益值

| 收益<br>中心态度 \ 相邻态度 |  | 支持 |    | 反对 |    |
|-------------------|--|----|----|----|----|
|                   |  | 支持 | 反对 | 支持 | 反对 |
| 支持                |  | PC | PC | PD | PD |
| 反对                |  | PD | PD | PC | PC |

$$E_N(C) = \begin{cases} E_1(C) = PC \ (S_{(i,j)} = 1) \\ E_2(C) = PC \ (S_{(i,j)} = 1) \\ E_3(C) = PD \ (S_{(i,j)} = -1) \\ E_4(C) = PD \ (S_{(i,j)} = -1) \end{cases} \quad (7)$$

$$E_N(C) = \begin{cases} E_1(C) = PC \ (S_{(i,j)} = 1, S_N = 1) \\ E_2(C) = PD \ (S_{(i,j)} = 1, S_N = -1) \\ E_3(C) = PD \ (S_{(i,j)} = -1, S_N = 1) \\ E_4(C) = PC \ (S_{(i,j)} = -1, S_N = -1) \end{cases} \quad (8)$$

2.2 算法模型

2.2.1 相关定义

(1) $n$ :元胞空间的行列数;(2) $n \times n$  二维元胞空间;(3) $A$ :倾向性态度集合  $\{1, -1\}$ ;(4) $(i, j)$ :元胞空间某个舆情主体坐标;(5) $S_{(i,j)}^t$ : $t$  时刻,主体  $(i, j)$

的情感倾向度;(6) $AE(n, n)$ :整个元胞空间所有舆情个体的总平均收益;(7) $N$ :表示中心元胞的某个邻居;(8) $W(i, j)$ :中心元胞  $(i, j)$  与所有邻居元胞进行一次竞争后的平均收益;(9) $CAS(n, n)$ :元胞自动机系统;(10) $E_N(i, j)$ :中心元胞  $(i, j)$  与邻居元胞  $N$  进行一次竞争后的收益;(11) $r$ :竞争遍历次数.

2.2.2 算法实现

输入: $CAS^t(n, n)$ ;

输出: $CAS^{t+1}(n, n), S_{(i,j)}^{t+1}$ ;

(1) 初始化:二维空间  $n$ ;定义元胞自动机系统  $CAS^t(n, n); t = 0$ ;中心元胞初始收益  $E_N(i, j) = 0$ ;通过统计初始支持数计算初始支持者收益及反对者收益;总平均收益  $AE(n, n) = 0$ ;

(2) 开始遍历元胞  $(i, j)$ ;

(3) 将中心元胞  $(i, j)$  与邻居竞争;

(4) 累加中心元胞  $(i, j)$  与每个邻居元胞进行一次竞争后的收益  $E_N(i, j)$ ;

(5) 如果未遍历所有邻居,转(3);

(6) 计算每个中心的平均收益  $W(i, j) = \frac{1}{8} \sum E_N(i, j)$ ;

(7) 累加元胞空间的总平均收益  $AE(i, j)$ ;

(8) 遍历下一个元胞,转(2);

(9) 计算整个元胞空间所有舆情个体的总平均

收益  $AE(n, n) = \frac{1}{9} \sum AE(i, j)$ ;

(10) 将每一个元胞  $(i, j)$  与总平均收益相比;

(11) 如果  $(W(i, j) > AE(n, n))$  则  $S_{(i,j)}^{t+1} = -S_{(i,j)}^t$ ;

(12) 如果  $(W(i, j) = AE(n, n))$  则  $S_{(i,j)}^{t+1} = \emptyset$ ;

(13) 否则  $S_{(i,j)}^{t+1} = S_{(i,j)}^t$ ;

(14) 相比下一个元胞,转(10);

(15) 输出  $CAS^{t+1}(n, n)$ ;

其中,步骤(3)~(7)是每个中心元胞与所有邻居竞争并计算收益的过程;步骤(10)~(14)是比较绝对元胞状态的变化过程.

3 实验及结果分析

利用基于决策论元胞自动机舆情传播算法,其中以相对收益方法作为演变规则,借助 MATLAB 软件进行模拟仿真,验证算法的有效性和合理性.

(1)状态变化

假设  $n=30$ ,按一定的概率随机地取 1 值(或-1值),格子取 1 值的概率实际上代表了初始时刻系统对某种意见的倾向率,如图 2(a)中,支持数为 524;每个格子的状态根据与邻域竞争后的收益演变,图 2(b)表示 50 次演变后的终态图形,可以发现,具有相同值的格子都聚集在一起,这似乎体现了“物以类聚、人以群分”的舆情传播现象.对不同初始支持数比例进行了演变,模拟结果体现了现实舆情传播的特征.

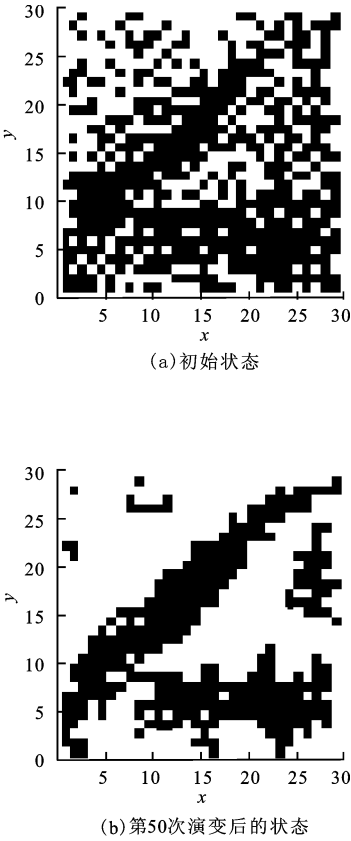


图 2 利用决策论元胞自动机演变舆情状态图

(2)判断舆情的倾向度

在二维元胞空间中,取 1 值和取-1 值的格子数目就分别代表了对两种意见的倾向度,为此,定义对某个观点的支持倾向度(简称支持度),表示为式(9).

$$R_1 = N_1 / N \tag{9}$$

$N$  是系统总的格子数目, $N_1$  是系统取 1 值的格子数目;显然反对倾向度(简称反对度)为  $R_0 = 1 - R_1$ .

元胞空间的演变状态如图 3 所示,由于初始二维空间中支持数大于反对数,利用决策论在一定次数演变后,反对元胞数趋向零,支持者在空间中分布趋向集中,演变次数 50 次以后,空间中元胞状态的

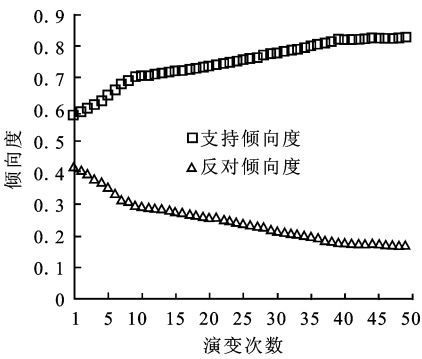


图 3 舆情演变次数与倾向度的变化趋势

变化趋向稳定;从演变次数与倾向度的关系曲线(图 3)可以看出,一旦传播开始,由于决策论竞争规则的作用,元胞总数逐渐减少,支持倾向度和反对倾向度区别逐渐明显,这给予我们的启示是:舆情传播的趋势不一定与多数倾向有必然关系,而与元胞及其邻居的状态变化有关系,网络舆情中多数倾向不会一直决定整个网络舆情的发展.

4 结束语

本文将决策论结合元胞自动机应用于舆情传播问题中,解决了传统元胞自动机多数规则演变方法所引起的问题,建立基于决策论的状态演变规则模型.考虑邻居对中心元胞的影响,使用相对收益的计算模型,将舆情传播个体放在虚拟空间中,设计元胞状态演变的遍历算法,在实例中演示元胞竞争演变过程.通过模拟仿真,对  $n=30$  的二维空间进行试验,经过多次演变后,在决策论竞争规则的作用下,总元胞数会减少并趋向稳定,弱势元胞数逐渐减少甚至为零,强势元胞数也会减少,但在空间总的比例增多,即舆情传播的倾向度会逐渐明显,舆情倾向度观点靠拢,呈现高度认同性.因此,在元胞自动机舆情传播中引入决策论是一种有效的研究方法,这对舆情的分析与预测研究提供了帮助.

参考文献:

[1] 殷卫东,朱晓华,赵俊凯.网络舆情技术浅析[J].信息安全,2012,30(3):5-9.  
[2] 栾霞,马晨辰.基于 Single-Pass 的部队医院网络舆情监控系统设计[J].电子设计工程,2015,23(4):60-63.  
[3] 刘毅.网络舆情研究概论[M].天津:天津人民出版社.2007(7).  
[4] 曾润喜.我国网络舆情研究与发展现状分析[J].图书馆研究,2009,17(8):2-6.

显的下降,且随着任务数量增长,下降趋势越来越明显。

本文算法的负载均衡性较高,这是因为本文算法不仅将完成时间作为约束条件,而且在任务调度中还进行了任务资源服务能力约束,在达到资源优化分配的情况下,也使云平台任务更均衡地调度到不同的资源上,增强了云平台的负载均衡性。

采用本文算法的服务总费明显低于最小最大算法,这是由于在对云平台海量任务进行调度时,本文算法将费用作为任务调度的约束条件,使其最终总费用较低。

## 4 结束语

本文提出一种新的云平台海量任务的多约束调度算法,将任务优先级、资源服务能力、完成时间和费用作为约束条件建立多约束调度模型,通过效用函数将云平台海量任务多约束调度模型转换成效用,将最大化全部用户的总效用作为目标函数。依据蚁群原理,将云平台中被频繁调度的计算资源与数据资源之间的传输链路看作是蚂蚁经过的路径,将信息素引入云平台任务调度中,获取符合约束条件的计算资源,达到调度目标最优。实验结果表明,所提算法不仅效率高,而且负载均衡性强,所需总费用较低。

## 参考文献:

[1] 朱映映,陈阳,明仲.云系统中面向海量多媒体数据的动态任务调度算法[J].小型微型计算机系统,2013,34(4):783-788.

[2] 姚卫红,黄小远,方仁孝.基于车联网应用的云平台任务调度算法[J].计算机仿真,2015,31(10):165-169.  
[3] 吴洲.云计算环境下基于用户和资源约束的免疫效用均衡任务调度算法[J].计算机系统应用,2015,24(10):176-180.  
[4] 刘丹琦.移动云计算环境下多 DAG 工作流的节能调度算法及能耗测量研究[D].乌鲁木齐:新疆大学,2014.  
[5] 熊磊,陈宏伟,王淑平.CloudSim 中基于偏序关系的调度算法研究[J].湖北工业大学学报,2013,28(5):79-81.  
[6] 徐惠萍,刘光明,孙涛,等.基于云计算技术的电信经营分析系统中对海量数据的研究[J].计算机光盘软件与应用,2013,18(12):267-269.  
[7] 王祺元,闫宏印.基于改进 RM 算法的云环境任务调度研究[J].计算机测量与控制,2013,21(6):1612-1614.  
[8] 马艳,龚斌,邹立达.网格环境下基于复制的能耗有效依赖任务调度研究[J].计算机研究与发展,2013,50(2):420-429.  
[9] 何鹏举,吴来斌,宋凯华,等.基于粒子群算法的嵌入式云计算资源调度[J].电子设计工程,2014,22(10):88-90.  
[10] 王鹏,黄焱,李坤,等.云计算集群相空间负载均衡度优先调度算法研究[J].计算机研究与发展,2014,51(5):1095-1107.

## 作者简介:

曾薇 (1981-),女,硕士研究生,讲师.研究方向为算法设计及优化、数据挖掘. E-mail:389713653@qq.com.

(上接第 129 页)

[5] 方薇,何留进,孙凯,等.采用元胞自动机的网络舆情传播模型研究[J].计算机应用,2010,30(3):751-755.  
[6] 聂哲,李粤平,温晓军,等.个体间相互影响的网络舆情演变模型[J].计算机工程应用,2009,45(14):220-222.  
[7] 陆题佳.因特网中危机信息传播规律及应对模式研究[D].合肥:中国科学技术大学,2010.  
[8] Zeng Jianping, Zhang Shiyong, Wu Chengrong, et al. Modelling topic propagation over the Internet [J]. Mathematical and Computer Modelling of Dynamic Systems, 2009, 15(1): 83-93.  
[9] 王杨,尤科本,王梦瑶,等.基于博弈论的网络社区舆情

传播模型[J].计算机应用研究,2013,30(8):2480-2482.

[10] Hofbauer J, Sigmund K, Evolutionary games and population dynamics[D]. cambridge: Cambridge University press, 1998.  
[11] Hauert C, Szab G, Game theory and Physics[J]. Am J Phys, 2005, 73(5): 405-414.  
[12] Nowak M A, May R, Evolutionary games and spatial-chaos[J]. Letters to Natwre, 1992, 359(11): 826-829.

## 作者简介:

黎永壹 男, (1979-), 硕士, 高级工程师. 研究方向为信息系统与知识发现. E-mail: qzxylyy@163.com.