

Single-Sample Network Entropy and Early Warning of Sudden Deterioration for Some Complex Diseases

Mingliang Li, Rui Liu

School of Mathematics, South China University of Technology, Guangzhou Guangdong
Email: 452868587@qq.com, 1622468926@qq.com

Received: Jan. 2nd, 2019; accepted: Jan. 17th, 2019; published: Jan. 24th, 2019

Abstract

For the early diagnosis of complex diseases, traditional methods require many patient samples to obtain statistical indicators. In real life, people do not go to hospitals for medical examination frequently, so it is difficult to obtain a large number of patients data, and only the result of individual single sampling has been got, which requires a single sample based algorithm to detect critical points in the development of complex diseases. Based on the theory of dynamic system and dynamic network biomarkers, a new algorithm is developed by using new individual samples combined with local biomolecular networks and the concept of entropy. The algorithm is applied to numerical simulation data and a real disease data. In these applications, the algorithm can detect the critical point in time and accurately. The new algorithm developed in this paper is stable and efficient. It can help medical workers to further understand the dynamic biomolecular mechanism of complex diseases, and can more quickly propose a reasonable treatment plan.

Keywords

Single-Sample, Dynamical System, Dynamic Biomarkers, Local Network Entropy

单样本网络熵与某些复杂疾病恶性突变的预警

李明亮, 刘 锐

华南理工大学数学学院, 广东 广州
Email: 452868587@qq.com, 1622468926@qq.com

收稿日期: 2019年1月2日; 录用日期: 2019年1月17日; 发布日期: 2019年1月24日

摘要

关于复杂疾病的早期诊断, 传统方法需要许多病人样本来获取统计指标。而现实生活中人们不会频繁出入医院进行体检, 故难以获得大量的患者数据, 只有对个体单次取样的结果, 这就需要提出基于单样本的探测复杂疾病发展过程中临界点的算法。本文基于动力系统和动态网络生物标志物的理论, 利用个体新样本结合局部生物分子网络以及熵的概念开发了一个新算法。将该算法应用于数值模拟数据和一个真实的疾病数据, 在这些应用中, 该算法能及时准确地探测出临界点。本文开发的新算法具有稳定性和高效性, 能帮助医学工作者进一步了解复杂疾病的动态生物分子机制, 能更快速地提出合理的治疗方案。

关键词

单样本, 动力系统, 动态生物标志物, 局部网络熵

Copyright © 2019 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

许多复杂系统的演变过程都存在着临界点, 比如种群生物数量系统、金融市场、气候系统、癫痫发作等系统中都存在着在某个很短的时间段内系统从一个稳定状态迅速转变到另一个稳定状态的现象。在临界点时期, 复杂系统的鲁棒性急剧降低, 状态转移风险大幅升高[1]。关于复杂疾病的临界点, 相关的研究已经取得了许多有效的成果, 比如最近有国内学者提出了利用动态网络生物标志物进行复杂疾病临界点探测的方法[2] [3] [4]。一般而言, 可以将复杂疾病(比如肝癌、糖尿病、癫痫等)的发展过程划分为三个阶段, 分别是健康阶段、临界阶段、疾病阶段, 其中临界阶段持续的时间相较于其他阶段而言非常短暂。处于健康阶段的系统比较稳定, 不易受到外界环境的扰动而改变, 具有较强的恢复力和鲁棒性。而临界阶段的系统恢复力和鲁棒性均很弱, 状态很不稳定, 经过及时治疗, 系统可以从临界阶段恢复到健康阶段, 一旦系统进入了疾病阶段, 则系统便处于了另一个较稳定状态, 即疾病已经发展到重症阶段, 再给予多大强度的治疗, 系统也难以从疾病阶段恢复到健康阶段。

预测复杂疾病进展过程中的临界点具有重要的实际意义, 但这项工作却是非常困难的。前人提出的单纯利用基因表达浓度的变化得到相关的不一致性指标的方法是静态的方法, 不适用于探测复杂疾病的临界点。而利用传统的临界慢化法则需要病人提供大量的时序列样本数据, 现实情况下难以在短时间内获取符合要求的疾病数据, 所以利用生物网络对复杂疾病进行早期诊断是大势所趋。最近有学者提出了利用单个样本检测复杂疾病临界点的方法[5], 本文则基于利用单个样本与已知样本结合局部网络[6] [7] [8] [9] [10]以及熵的概念[11] [12] [13] [14]开发出了一个新的探测复杂疾病临界点的算法。本文首先利用希尔方程建立具有 11 个结点的复杂疾病生物仿真网络, 并将我们的算法应用于仿真网络生成的数据, 算法成功探测出了仿真数据的临界点。然后对 NCBI 生物网站上的小鼠肺光气实验(GSE2565)这个真实的高通量时序列数据应用本文的算法进行研究发现, 该疾病即将恶化的预警信号出现的时间点与实验观察的结果吻合。

2. 方法

2.1. 单样本网络熵的推导

由于生理现象的复杂性, 影响生理状态的因素非常多, 目前尚无法用一个微分方程来描述。对于复杂疾病的临界点这个特殊时期, 疾病的发展依赖于某几个主导因素, 于是通过研究低维的动力系统微分方程来代换高维而复杂的疾病生理系统[2] [3] [4]。

对于一个复杂疾病临界点时的系统, 可通过建模成为离散的动力系统:

$$Z(t+1) = f(Z(t); P). \tag{1}$$

在系统(1)中的 $Z(t) = (z_1(t), z_2(t), \dots, z_n(t))$ 代表固定 t 时刻下 n 维基因表达浓度的状态变量, 参数向量 $P = (p_1, p_2, \dots, p_s)$ 代表生理系统的驱动因子, 比如基因拷贝数变异、DNA 甲基化等。

系统(1)需要满足一些条件: 1) 存在不动点 $\bar{Z}(t) = (\bar{z}_1(t), \bar{z}_2(t), \dots, \bar{z}_n(t))$ 满足 $\bar{Z} = f(\bar{Z}; P)$; 2) 存在 P_c 使得雅可比矩阵

$$\left. \frac{\partial f(Z; P_c)}{\partial Z} \right|_{Z=\bar{Z}}, \tag{2}$$

有一个或一对特征值的模等于 1; 3) 当 $P \neq P_c$ 时, 系统(1)的特征根的模不总等于 1。这三个条件意味着当 P 达到阈值 P_c 时该系统在 \bar{Z} 处经历了一个相变或一个余维数一的分岔。

在系统(1)的 \bar{Z} 附近, 当 P 到达 P_c 前, 按假设系统应该会留在一个稳定的不动点 \bar{Z} 上, 因此所有的特征值的模是在(0,1)之间的。在系统处于临界点时的参数值 P_c 称为分岔点值或临界跃迁值。

现在考虑系统(1)的线性化近似方程, 具体而言, 通过引入新的变量 $Y(t) = (y_1(t), y_2(t), \dots, y_n(t))$ 和一个可逆的转换矩阵 S , 即

$$Y(t) = S^{-1}(Z(t) - \bar{Z}). \tag{3}$$

结合系统(1)有

$$Y(t) = \Lambda(P)Y(t-1) + \xi(t-1), \tag{4}$$

(4) 式中 $\xi = (\xi_1, \xi_2, \dots, \xi_n)$ 是均值为零的微小高斯噪声, $\xi_i (i=1, 2, \dots, n)$ 的标准偏差为 $\sigma_i (i=1, 2, \dots, n)$, Λ 是关于参数 P 的对角矩阵。

根据假设, 对角化矩阵 $\Lambda = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$ 的每个 $\lambda_i (i=1, 2, \dots, n)$ 的模均在 0 和 1 之间。不失一般性, 在 Λ 的特征值中, 假设模最大的一个为 λ_1 , 当参数转换 $P \rightarrow P_c$ 发生时, λ_1 是第一个到达 1 的。特征值 λ_1 表征了系统围绕不动点改变的速率, λ_1 被称为占主导地位的特征值。系统远离临界点时的状态对应于 $|\lambda_1| < 1$, 而系统趋近临界点时的状态对应于 $|\lambda_1| \rightarrow 1$ 。设 Y 中与 λ_1 相关的变量为 y_1 , 又根据方差与均值的性质得到 $y_i (i=1, 2, \dots, n)$ 的方差为

$$\text{Var}(y_i) = \frac{\sigma_i^2}{|1 - \lambda_i^2|}, \tag{5}$$

当主导特征值 $|\lambda_1| \rightarrow 1$ 时, 从(5)式中, 可以得到变量 $z_i = s_{i1}y_1 + s_{i2}y_2 + \dots + s_{in}y_n + \bar{z}_i (s_{i1} \neq 0)$ 的标准偏差急剧增加的结论。因此, 当系统处于远离临界点的状态时, $|\lambda_1|$ 远小于 1 并且没有大的变化, $Z(t)$ 和 $Z(t-1)$ 之间的分布也没有明显的变化。然而, 当该系统处于趋近临界点时的状态, 即 $|\lambda_1|$ 接近 1 时, $Z(t)$ 和 $Z(t-1)$ 这两者的分布便有差异。

对于以第 i 个结点为中心的局部网络, 假设总共有 m 个一阶邻居 i_1, i_2, \dots, i_m , 现令随机变量 \mathcal{X} 代表局部网络的状态, 局部网络中结点的值为 0 或 1, 则 \mathcal{X} 总共取值个数为 2^{m+1} , 设 $P_{u,v}(t)$ 代表 t 时刻第 u 个

局部网络并且状态为 v 的概率, 则第 i 个局部网络的熵为

$$E_i^{(t)}(\mathcal{X}) = -\sum_v P_{i,v}(t) \log P_{i,v}(t), \quad (6)$$

整体网络熵为

$$E^{(t)}(\mathcal{X}) = -\sum_{u,v} P_{u,v}(t) \log P_{u,v}(t). \quad (7)$$

系统远离临界点时, 假设局部网络中某个结点 i_i 为显著改变状态, 即值取 1 的概率为

$$P(x_{i_i}(t) = 1) = P(|\Delta Z_{i_i}(t)| > d_i) = P_1, \quad (8)$$

其中 $\Delta Z_{i_i}(t) = Z_{i_i}^c(t) - Z_{i_i}^n(t)$, $Z_{i_i}^c(t)$ 为 t 时刻单样本第 i_i 个基因的表达浓度值, $Z_{i_i}^n(t)$ 为 t 时刻已知样本第 i_i 个基因的表达浓度值, 所以结点 i_i 为不显著改变状态即值取 0 的概率为

$$P(x_{i_i}(t) = 0) = P(|\Delta Z_{i_i}(t)| \leq d_i) = P_0 = 1 - P_1, \quad (9)$$

因为正常状态下基因表达稳定, 所以 $P_1 \rightarrow 0$, 即

$$P(x_{i_i}(t) = 1) \rightarrow 0. \quad (10)$$

令第 i 个局部网络的状态 $A_i = \{x_i = 0, x_{i_1} = 0, \dots, x_{i_m} = 0\}$, 则有

$$P(\mathcal{X} = A_i) \rightarrow 1, \quad (11)$$

而 $P(\{\dots, x_{i_i} = 1, \dots\}) \leq P(x_{i_i}(t) = 1) \rightarrow 0$, 所以局部网络 i 处于 A_i 以外的其他状态的概率是趋于 0 的,

故在正常状态有

$$E^{(t)}(\mathcal{X}) = -\sum_{u,v} P_{u,v}(t) \log P_{u,v}(t) \rightarrow 0. \quad (12)$$

而系统趋近于临界点时, 对于第 i_i 个结点, 有 $sd(Z_{i_i}) \rightarrow \infty$, 其中 Z_{i_i} 为第 i_i 个基因的表达浓度值, 而 $\Delta Z_{i_i}(t) = Z_{i_i}^c(t) - Z_{i_i}^n(t)$, 故有

$$sd(\Delta Z_{i_i}) = sd(Z_{i_i}), \quad (13)$$

也即说明 ΔZ_{i_i} 的离散程度非常高, 对于第 i 个局部网络, 有

$$E_i^{(t)}(\mathcal{X}) = -\sum_{v=1}^{2^{m+1}} P_{i,v}(t) \log P_{i,v}(t), \quad (14)$$

和

$$\sum_{v=1}^{2^{m+1}} P_{i,v}(t) = 1. \quad (15)$$

针对(14)和(15)二式引入拉格朗日乘子 λ , 得到

$$G_i^{(t)}(\mathcal{X}) = -\sum_{v=1}^{2^{m+1}} P_{i,v}(t) \log P_{i,v}(t) + \lambda \left(\sum_{v=1}^{2^{m+1}} P_{i,v}(t) - 1 \right), \quad (16)$$

令 $G_i^{(t)}(\mathcal{X})$ 对每个 $P_{i,v}$ 及 λ 求偏导并令其等于 0, 可得

$$-\log P_{i,v}(t) - 1 + \lambda = 0, v = 1, 2, \dots, 2^{m+1}, \quad (17)$$

和

$$\sum_{v=1}^{2^{m+1}} P_{i,v}(t) = 1. \quad (18)$$

所以 $\lambda - 1 = \log P_{i,v}(t) \Rightarrow P_{i,v}(t) = e^{\lambda-1} = \frac{1}{2^{m+1}}$, 则有

$$\lambda = \log\left(\frac{1}{2^{m+1}}\right) + 1, \quad (19)$$

故当 ΔZ_{ij} 是均匀分布的时候, 单样本网络熵能取到最大值, 此时有

$$E_i^{(t)}(\mathcal{X}) = \log(2^{m+1}) = m + 1. \tag{20}$$

2.2. 探测临界点

1) 针对某种复杂疾病, 收集该疾病已有的多个样本, 并将该疾病对应的全基因组映射到 STRING (STRING 数据库是一个搜寻已知蛋白质之间和预测蛋白质之间相互作用的系统)的蛋白质 - 蛋白质相互作用网络。

2) 针对某个未知状态的样本, 将该样本分别与每个已知样本建立状态网络。

3) 局部化全基因组状态网络。

4) 计算每个局部化状态网络的单样本网络熵。

5) 将所有局部状态网络的单样本网络熵排序并取前 10% 的平均值作为差异评价指标。

3. 主要结果

3.1. 仿真网络

本文利用 11 个结点网络生成疾病的仿真数据, 然后利用新开发的算法探测疾病发展过程的临界点。这种类型的基因调控网络常常应用于研究基因转录、翻译等影响基因调控的活动。利用 MM (Michaelis-Menten) 方程并加入基因自身的降解率, 可以得到 11 个结点的仿真网络, 对应的微分方程组为:

$$\left\{ \begin{aligned} \frac{dz_1(t)}{dt} &= \frac{(8-3|P|)z_4(t)}{20(1+z_4(t))} - \frac{8+3|P|}{20}z_1(t) + \zeta_1(t) \\ \frac{dz_2(t)}{dt} &= \frac{(5-3|P|)z_1(t)}{20(1+z_1(t))} + \frac{(5-3|P|)z_4(t)}{20(1+z_4(t))} - \frac{1}{2}z_2(t) + \zeta_2(t) \\ \frac{dz_3(t)}{dt} &= \frac{3|P|-6}{10} + \frac{6-3|P|}{20(1+z_1(t))} + \frac{6-3|P|}{20(1+z_4(t))} - \frac{3}{5}z_3(t) + \zeta_3(t) \\ \frac{dz_4(t)}{dt} &= \frac{(8-3|P|)z_1(t)}{20(1+z_1(t))} - \frac{(8+3|P|)}{20(1+z_4(t))}z_4(t) + \zeta_4(t) \\ \frac{dz_5(t)}{dt} &= -\frac{1}{2} + \frac{1}{20(1+z_1(t))} + \frac{7z_2(t)}{10(1+z_2(t))} + \frac{3z_3(t)}{5(1+z_3(t))} + \frac{1}{20(1+z_4(t))} \\ &\quad + \frac{1}{5(1+z_6(t))} + \frac{1}{5(1+z_7(t))} - \frac{6}{5}z_5(t) + \zeta_5(t) \\ \frac{dz_6(t)}{dt} &= -\frac{1}{5} + \frac{z_2(t)}{5(1+z_2(t))} + \frac{z_3(t)}{5(1+z_3(t))} + \frac{1}{5(1+z_5(t))} + \frac{z_7(t)}{5(1+z_7(t))} - \frac{6}{5}z_6(t) + \zeta_6(t) \\ \frac{dz_7(t)}{dt} &= -\frac{2}{5} + \frac{1}{5(1+z_2(t))} + \frac{1}{5(1+z_3(t))} + \frac{z_5(t)}{5(1+z_5(t))} + \frac{z_6(t)}{5(1+z_6(t))} - \frac{6}{5}z_7(t) + \zeta_7(t) \\ \frac{dz_8(t)}{dt} &= \frac{z_6(t)}{5(1+z_6(t))} + \frac{z_7(t)}{5(1+z_7(t))} + \frac{2z_{10}(t)}{5(1+z_{10}(t))} - \frac{8}{5}z_8(t) + \zeta_8(t) \\ \frac{dz_9(t)}{dt} &= \frac{4z_6(t)}{5(1+z_6(t))} + \frac{4z_7(t)}{5(1+z_7(t))} - \frac{9}{5}z_9(t) + \zeta_9(t) \\ \frac{dz_{10}(t)}{dt} &= \frac{z_6(t)}{(1+z_6(t))} + \frac{z_7(t)}{(1+z_7(t))} - 2z_{10}(t) + \zeta_{10}(t) \\ \frac{dz_{11}(t)}{dt} &= \frac{6z_6(t)}{5(1+z_6(t))} + \frac{6z_7(t)}{5(1+z_7(t))} - \frac{11}{5}z_{11}(t) + \zeta_{11}(t) \end{aligned} \right. \tag{21}$$

其中 P 为数值控制参数, $\zeta_i(t)(i=1,2,\dots,11)$ 是均值为 0 的高斯噪声, $z_i(i=1,2,\dots,11)$ 为第 i 个 mRNA 的浓度, 11 个 mRNA 的自降解率分别为 $\left(\frac{8+3*|P|}{20}, \frac{1}{2}, \frac{3}{5}, \frac{8+3*|P|}{20}, \frac{6}{5}, \frac{6}{5}, \frac{6}{5}, \frac{8}{5}, \frac{9}{5}, 2, \frac{11}{5}\right)$ 。微分方程组(21)完全体现了仿真网络之中结点间的调控关系, 并且有一个稳定点:

$$\bar{Z} = (\bar{z}_1, \bar{z}_2, \bar{z}_3, \bar{z}_4, \bar{z}_5, \bar{z}_6, \bar{z}_7, \bar{z}_8, \bar{z}_9, \bar{z}_{10}, \bar{z}_{11}) = (0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0). \tag{22}$$

将微分方程组(21)利用一个小 Δt 进行差分, 然后用 $J = \left. \frac{\partial f(Z(k); P)}{\partial Z} \right|_{Z=\bar{Z}}$ 代表差分方程组的雅可比

矩阵, 有

$$J = e^{A*\Delta t}, \tag{23}$$

其中 A 为微分方程组(21)线性主部的系数矩阵, A 的 11 个特征值分别为:

$$\left(-\frac{3}{10}*|P|, -\frac{1}{2}, -\frac{3}{5}, -\frac{4}{5}, -1, -\frac{6}{5}, -\frac{7}{5}, -\frac{8}{5}, -\frac{9}{5}, -\frac{10}{5}, -\frac{11}{5}\right), \tag{24}$$

所以当 $P \in (0, 1]$ 时平衡点 \bar{Z} 是稳定的, 当 $P \rightarrow P_c = 0$ 时, 系统便失去了稳定性, 并且系统状态会经历一个临界转变。

现在将雅可比矩阵 J 进行对角化, 先引入可逆矩阵 S :

$$S = \begin{bmatrix} 1 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \tag{25}$$

使得 $S^{-1}JS = \Lambda$, 对于 Λ 有:

$$\Lambda = \text{diag}\left(0.7408^{|P|}, 0.6065, 0.5488, 0.4493, 0.3679, 0.3012, 0.2466, 0.2019, 0.1653, 0.14, 0.11\right), \tag{26}$$

即 Λ 是个对角矩阵。现假设 Λ 的特征值 $\lambda_i(i=1,2,\dots,11)$ 对应的特征向量分别为:

$$(y_1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0), (0, y_2, 0, 0, 0, 0, 0, 0, 0, 0, 0), \dots, (0, 0, 0, 0, 0, 0, 0, 0, 0, 0, y_{11}). \tag{27}$$

注意到变量 y_1 与最大的特征值 $0.7408^{|P|}$ 相关并且最接近 1, 因此, 当 P 趋于 0 时, 主特征值 $0.7408^{|P|}$ 趋于 1, 系统发生临界转变。又根据:

$$Y(k) = S^{-1}(Z(k) - \bar{Z}), \tag{28}$$

可知 $(z_1, z_2, \dots, z_{11})$ 中的四个变量 z_1, z_2, z_3 和 z_4 是直接主特征值对应的 y_1 相关的, 根据 2.1 节的理论推导可知, $\{z_1, z_2, z_3, z_4\}$ 构成了 11 个结点仿真网络系统的主要部分, 即相当于真实数据中的动态网络生物标志物。

基于本文新开发的算法, 收集 11 个仿真结点的时序列数据, 并将算法应用于数值模拟数据得到的结果如下图 1:

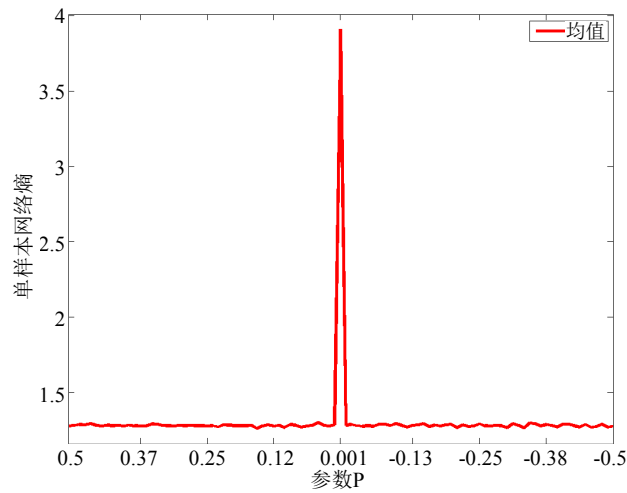


Figure 1. Curve: result of algorithm applied to simulation network data
图 1. 算法应用于仿真网络数据结果曲线

3.2. 真实数据

实验数据是从 NCBI 的 GEO 数据库(<https://www.ncbi.nlm.nih.gov/geo/>)下载得到, 数据编号为 GSE2565。氯化羰(光气)是一种有毒的工业化合物, 人或小鼠暴露于光气之中会产生肺水肿和不可逆急性肺损伤, 令 CD-1 雄性小鼠全身暴露于空气(控制组)和光气(实验组)之中, 并分别在暴露 0.5、1、4、8、12、24、48、72 小时后提取小鼠肺部 RNA 以获得基因表达浓度, 实验中观测到在 GSH 还原转化进入氧化型 GSSG 的过程起非常重要作用的 GPX-2 在暴露 4 到 12 小时的时候基因表达显著增加, 表现出显著性[15]。根据小鼠肺光气实验数据, 利用本文基于单样本开发的新算法, 得到单样本网络熵在第 4 小时显著增加并达到峰值, 即探测到的临界点为第 4 个时间点, 得到的结果与实验观测吻合, 真实数据的结果如下图 2:

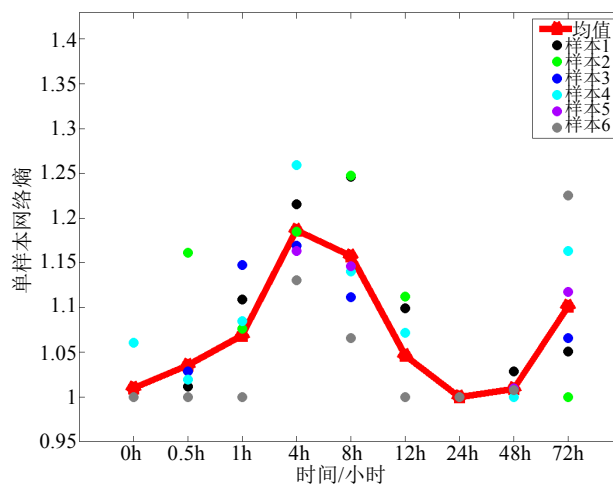


Figure 2. Curve: result of algorithm applied to data GSE2565
图 2. 算法应用于数据 GSE2565 结果曲线

4. 讨论

本文思考传统方法的局限性, 创新性地将单样本方法与熵进行结合, 提出了一种新的有效稳定的预警复杂疾病临界点的方法。基于小鼠肺光气实验数据, 本文开发的算法成功探测出了第四个时间点为临界点。复杂疾病比如前列腺癌、肝癌、肿瘤等等对人类健康造成巨大的威胁, 而能够成功探测出复杂疾病即将恶化的信号, 便能帮助人类及时预防甚至治疗复杂疾病, 本文开发出的新算法是传统方法的一个突破。

本文仍然有不足之处, 比如为了使模型简单, 只讨论了线性化系统系数矩阵的特征值全不相同的这种情况, 并没有考虑有重根和共轭复根时算法的效果, 这是以后需要改进的一个主要方向。

基金项目

广州市科技计划珠江科技新星专项项目资助(No. 201610010029)。

参考文献

- [1] Scheffer, M. (2010) Complex Systems: Foreseeing Tipping Points. *Nature*, **467**, 7314. <https://doi.org/10.1038/467411a>
- [2] Chen, L., Liu, R., Liu, Z., *et al.* (2012) Detecting Early-Warning Signals for Sudden Deterioration of Complex Diseases by Dynamical Network Biomarkers. *Scientific Reports*, **2**, 342. <https://doi.org/10.1038/srep00342>
- [3] Liu, R., *et al.* (2012) Identifying Critical Transitions and Their Leading Biomolecular Networks in complex Diseases. *Scientific Reports*, **2**, 813. <https://doi.org/10.1038/srep00813>
- [4] Chen, P., Liu, R., Chen, L., *et al.* (2016) Detecting Critical State before Phase Transition of Complex Biological Systems by Hidden Markov Model. *Bioinformatics*, **32**, 2143-2150. <https://doi.org/10.1093/bioinformatics/btw154>
- [5] Liu, X., Wang, Y., Ji, H., *et al.* (2016) Personalized Characterization of Diseases Using Sample-Specific Networks. *Nucleic Acids Research*, **44**, e164. <https://doi.org/10.1093/nar/gkw772>
- [6] He, D., Liu, Z., Honda, M., *et al.* (2012) Coexpression Network Analysis in Chronic Hepatitis B and C Hepatic Lesion Reveals Distinct Patterns of Disease Progression to Hepatocellular Carcinoma. *Journal of Molecular Cell Biology*, **4**, 140-152. <https://doi.org/10.1093/jmcb/mjs011>
- [7] Chen, L., Wang, R., Li, C., *et al.* (2010) Modeling Biomolecular Networks in Cells: Structures and Dynamics. Springer, New York. <https://doi.org/10.1007/978-1-84996-214-8>
- [8] Wang, J., Huang, Q., Liu, Z., *et al.* (2011) NOA: A Novel Network Ontology Analysis Method. *Nucleic Acids Research*, **39**, e87. <https://doi.org/10.1093/nar/gkr251>
- [9] Ideker, T. and Krogan, N.J. (2012) Differential Network Biology. *Molecular Systems Biology*, **8**, 565. <https://doi.org/10.1038/msb.2011.99>
- [10] Liu, R., Wang, X., Aihara, K., *et al.* (2014) Early Diagnosis of Complex Diseases by Molecular Biomarkers, Network Biomarkers, and Dynamical Network Biomarkers. *Medicinal Research Reviews*, **34**, 455-478. <https://doi.org/10.1002/med.21293>
- [11] Cover, T.M. and Thomas, J.A. (2006) Elements of Information Theory. Wiley-Blackwell, New Jersey.
- [12] Demetrius, L. and Manke, T. (2005) Robustness and Network Evolution an Entropic Principle. *Physica A*, **346**, 682-696. <https://doi.org/10.1016/j.physa.2004.07.011>
- [13] Manke, T., Demetrius, L. and Vingron, M. (2001) An Entropic Characterization of Protein Interaction Networks and Cellular Robustness. *Journal of the Royal Society Interface*, **30**, 51-64.
- [14] Gomez-Gardenes, J. and Latora, V. (2008) Entropy Rate of Diffusion Processes on Complex Networks. *Physical Review E*, **78**, 065102(4).
- [15] Sciuto, A.M., Phillips, C.S., Orzolek, L.D., *et al.* (2005) Genomic Analysis of Murine Pulmonary Tissue Following Carbonyl Chloride Inhalation. *Chemical Research in Toxicology*, **18**, 1654-1660. <https://doi.org/10.1021/tx050126f>

知网检索的两种方式：

1. 打开知网页面 <http://kns.cnki.net/kns/brief/result.aspx?dbPrefix=WWJD>
下拉列表框选择：[ISSN]，输入期刊 ISSN：2324-7991，即可查询
2. 打开知网首页 <http://cnki.net/>
左侧“国际文献总库”进入，输入文章标题，即可查询

投稿请点击：<http://www.hanspub.org/Submission.aspx>

期刊邮箱：aam@hanspub.org