

An Approach to Improve Diffusion Coefficient of Geospatial Information Model

Chongfu Huang

Academy of Disaster Reduction and Emergency Management, Faculty of Geographical Science
Beijing Normal University
Beijing, China
hcongfu@bnu.edu.cn

一种改进地理空间信息扩散系数的方法

黄崇福

北京师范大学地理科学学部减灾与应急管理研究院
北京 100875, 中国
hcongfu@bnu.edu.cn

Abstract—This paper proposes an approach to improve the diffusion coefficient of the geospatial information diffusion model. The diffusion coefficient calculated by the average distance formula is appropriately amplified to become the initial diffusion coefficient. Employing a search method, we take two test points in the search interval consisting of 0 and the initial diffusion coefficient. Comparing the errors of the two test points used in the geospatial information diffusion model, we adjust the search interval: if the error of the left test point is small, the left point of the new search interval is unchanged, and the original right point of search interval is replaced with the right test point; if the error of the right test point is small, The right point of the new interval is unchanged, and the original left point of search interval is replaced with the left test point. Repeatedly, the search interval is continuously narrowed until the distance between the two test points is less than a given value, then the search is stopped. Meanwhile, the test point with a small error will be an optimized diffusion coefficient. A case constructing a relationship between the background data and disaster, with a sample size of 30, shows that the diffusion coefficient can reduce error approximately 17%.

Keywords—geospatial information diffusion, diffusion coefficient, search interval, test point, background data, disaster

摘要—本文提出了一种改进地理空间信息扩散模型中扩散系数的方法。将平均距离公式计算的扩散系数进行适当放大,成为初始化扩散系数。使用某种搜索法,在由0和初始化扩散系数构成的搜索区间中取两个测试点。将两个测试点分别用于地理空间信息扩散模型,比较它们的误差,调整搜索区间:如果左测试点的误差小,则新搜索区间左端点不变,将原搜索区间的右端点换成右测试点;如果右测试点的误差小,则新搜索区间的右端点不变,将原搜索区间左端点换成左测试点。如此反复,不断缩小搜索区间,直到两个测试点的距离小于一个给定的值,则停止搜索时,并以误差较小的小测试点,为优化的扩散系数。用容量为30的样本,构建背景数据和灾情间关系的算例表明,扩散系数优化后,大约能减小17%的估计误差。

关键词—地理空间信息扩散, 扩散系数, 搜索区间, 测试点, 背景数据

I. 引言

大灾中的信息孤岛,比比皆是。由于具有非线性识别能力,且能学习矛盾样本,地理空间信息扩散模型^[1],较之加权地理回归^[2]和人工神经网络^[3],能更好地推测出空白地理单元上的灾情,有效解决信息孤岛的问题。优化模型中的扩散系数,是进一步提高推测结果精度的一个重要途径。

地理空间信息扩散模型,是将灾区已观测的地理单元上的背景数据和灾情形成的样本,视为小样本,用正态信息扩散公式^[4],对其进行集值化处理,构造出“背景数据”和“灾情”之间的因果关系。据此,我们用空白地理单元上的背景数据,可推导出该地理单元上灾情。

地理空间信息扩散模型,是一个集值统计回归模型。扩散公式中的扩散系数,决定着样本点的集值化程度,对预测结果有明显的影 响。扩散系数较大时,较多的监测点从一个样本点获得有效信息;反之,点较少。理论上,样本足够大时,扩散系数为零,样本点的信息,没有扩散。传统统计回归,就是在没有扩散的情况下进行。换言之,传统统计回归,依赖于大样本。

目前,信息扩散理论的基础比较稳固,应用涉及面较广。人们常用式(1)的平均距离公式计算正态信息扩散中的扩散系数 h ^[5]。

$$h = \begin{cases} 0.8146(b-a), & n=5; \\ 0.5690(b-a), & n=6; \\ 0.4560(b-a), & n=7; \\ 0.3860(b-a), & n=8; \\ 0.3362(b-a), & n=9; \\ 0.2986(b-a), & n=10; \\ 2.6851(b-a)/(n-1), & n \geq 11. \end{cases} \quad (1)$$

式中, $b = \max_{1 \leq i \leq n} \{x_i\}$, $a = \min_{1 \leq i \leq n} \{x_i\}$, x_i 是样本点, n 是样本容量。在多维情况下, 分别计算各分量的扩散系数。

显然, 样本点之间的距离, 常常差异很大。由平均距离假设, 对两点择近原则方程进行简化, 进而推导出的式 (1), 存在很大改进空间。

2000 年, 刘新立等用各样本点与其他样本点的距离中的最远距离, 乘以式 (1) 计算出的扩散系数, 使各样本点使用的扩散系数不一样, 改进了其所研究问题的估计精度^[6]。但于由于过程比较复杂, 也难以证明在一般情况下是否有效, 此项研究成果, 并没有得到广泛应用。2001 年, 王新洲和游扬声使用非参数估计理论中的最优核函数, 提出了一个迭代公式, 计算最优窗宽^[7]。由于只有当样本容量较大时, 最优窗宽的优越性才比较显著, 加之最优核函数其实只有在估计正态分布时才有效等天生不足, 他们建议的迭代公式, 对一般的小样本问题, 优越性并不突出。2007 年, 沈赟琦和陆余楚将解可压缩流体流动问题的 MacCormack 法引入扩散函数的研究, 提出了二步二阶差分格式法^[8]。这种从初值问题中解出扩散函数的方法, 理论上可同时解决扩散系数问题, 但应用上有困难, 主要是其相关方程的假设, 条件比较苛刻。2007 年, Karimia 和 Hüllermeier^[9]提出用经验贝叶斯法改进信息扩散理论中的内集-外集模型^[10], 但其经验分布, 不易获得。

值得一提的是, 随着互联网的广泛应用, 从 2016 年起, 在研究社会网络的国外文献中, Information diffusion 这一术语, 呈井喷式地出现。国内文献, 将其译为“信息扩散”, 其实指的是信息传播, 早年称之为“通讯”(Communication), 已有几千年的研究史。现代通讯技术, 就是信息传播的研究成果。无论从 20 世纪 60 年代人们借鉴传染病模型提出的谣言传播数学模型^[11], 还是如今人们将传播场理论的信息波动扩散与衰减模型引入社会网络信息传播, 提高在线广告效率^[12], 媒介都决定着信息传播的方式。

本文涉及的地理空间信息扩散, 并不考虑地下、地表或空气中的媒介。而是为弥补信息不足, 将没有几何大小的样本点, 进行集值化的“自由”信息扩散, 并非“受控”的信息传播。换言之, 没有“媒介”性质, 可供调节式 (1) 中的扩散系数。

类似地, 数模糊数学中的“云计算”和当今基于互联网的云计算, 也极易混淆。早年, 数模糊数学中没有“云计算”这一概念, 只有“云模型”, 是一个表达定性概念的模型^[13]。后来, 人们将用其进行的计算, 称为云计算。而基于互联网的云计算, 是指共享软硬件资源和信息, 按需提供给计算机和其他设备。这种云计算, 其实是并行计算发展到格计算, 加上大数据后的一种计算。尽管云模型的创始人在一次与作者的交流中认为两种“云计算”有关, 但过于牵强。从某种意义上讲, “云模型”与本文涉及的信息扩散, 关系反而密切一些。因为正态信息扩散中的扩散系数, 与正态云模型中的标准差是同一个功能。云模型效果如何, 大受其影响。

II. 地理空间信息扩散模型的监督学习

以数学模型计算的输出与训练样本的实际输出之误差, 来控制模型进行样本学习的进程, 称有监督学习。例: 线性回归、回传学习人工神经网络, 都是用最小均方差来监督学习。

用观测样本构建数学模型时, 如果没有任何准则来控制建模进程, 称无监督学习。例: 对一些观察对象进行聚类、用随机样本点进行概率密度函数的估计, 没有任何准则可用监督学习, 是无监督学习。凡是目标未知, 构建的模型无好坏之分的, 均是无监督学习。以聚类为例, 假如我们去参观一个画展, 我们对艺术完全一无所知, 但是欣赏完多幅作品之后, 我们也能把它们分成不同的派别。类, 可分多, 也可分少, 没有标准。

作为集值统计回归的地理空间信息扩散模型, 其功能是通过学习训练样本, 近似估计样本所来自母体中的输入-输出关系。例如, 使用该模型学习灾区已观测的地理单元上的背景数据和灾情形成的样本, 可以近似估计它们的输入-输出关系, 从而可由空白地理单元上的背景数据, 推测出灾情。信息扩散方法的优势, 是能从小样本中较准确地推测其母体中的输入-输出关系。如果样本容量足够大, 大到与母体在统计特征意义上没有差别, 就没有通过扩散对样本点进行集值化处理的必要。换言之, 由大样本支撑的传统统计回归模型, 是信息扩散统计回归模型的特例。既然是统计回归, 地理空间信息扩散模型当然就是一个有监督学习模型, 可用最小均方差来监督学习样本。

不失一般性, 设学习样本 W 为:

$$W = \{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\} \quad (2)$$

式中, $x_i, y_i, i=1, 2, \dots, n$ 分别是样本点中的输入向量和输出向量。进一地, 可假设输入向量 x_i 和输出向量 y_i 分别有 λ 和 τ 个分量组成, 即

$$\begin{cases} x_i = (x_{i1}, x_{i2}, \dots, x_{i\lambda}) \\ y_i = (y_{i1}, y_{i2}, \dots, y_{i\tau}) \end{cases} \quad (3)$$

$\forall j \in \{1, 2, \dots, \lambda\}$ 和 $\forall k \in \{1, 2, \dots, \tau\}$, 由学习样本 W 可得第 j 个输入分量观测值集合 X_j 和第 k 个输出分量观测值集合 Y_k , 如式 (4) 所示。注意, 式 (3) 的是小括号用于表达向量, 式 (4) 的大括号用于表达集合。各分量的第一个下标表达在 W 中的序号, 第二个下标表达分量的序号。

$$\begin{cases} X_j = \{x_{1j}, x_{2j}, \dots, x_{nj}\} \\ Y_k = \{y_{1k}, y_{2k}, \dots, y_{nk}\} \end{cases} \quad (4)$$

分别对各 X_j 和各 Y_k 用式 (1) 计算出扩散系数, 记为 h_{X_j} 和 h_{Y_k} 。

信息扩散, 是将样本点在其定义域上进行集值化处理。由于此工作是在计算机上进行, 只能进行等步长离散化处理。在定义域中选取的点, 称为信息监控点, 这些点构成信息监控空间。对 X_j 和 Y_k 的信息监控空间, 分别记为 U_{X_j} 和 V_{Y_k} , 由式 (5) 表之。

$$\begin{cases} U_{X_j} = \{u_{1j}, u_{2j}, \dots, u_{m(X_j)j}\}, j = 1, 2, \dots, \lambda \\ V_{Y_k} = \{v_{1k}, v_{2k}, \dots, v_{q(Y_k)k}\}, k = 1, 2, \dots, \tau \end{cases} \quad (5)$$

式中, j 和 k 个是分量标识, 整数 $m(X_j)$ 和 $q(Y_k)$ 是监控点个数, 通常取相同的值。此数决定了监控点的密集程度, 可根据计算量适当选取。当针对一个分量的监控点步长小于该分量观测值间非零最小距离时, 更小的步长, 徒增计算量, 并不能改进精度。

组合各分量的监控点, 构成对学习样本 W 的 $\lambda \times \tau$ 维信息监控空间 Ω :

$$\Omega = U_{X_1} \times U_{X_2} \times \dots \times U_{X_\lambda} \times V_{Y_1} \times V_{Y_2} \times \dots \times V_{Y_\tau} \quad (6)$$

$U_{X_1} \times U_{X_2} \times \dots \times U_{X_\lambda}$ 中的一个任意点和 $V_{Y_1} \times V_{Y_2} \times \dots \times V_{Y_\tau}$ 中的一个任意点, 分别记为 u 和 v , 即:

$$\begin{cases} u = (u_{l_1}, u_{l_2}, \dots, u_{l_\lambda}), l_j \in \{1, 2, \dots, m(X_j), j = 1, 2, \dots, \lambda \\ v = (v_{t_1}, v_{t_2}, \dots, v_{t_\tau}), t_k \in \{1, 2, \dots, m(Y_k), k = 1, 2, \dots, \tau \end{cases} \quad (7)$$

集值化处理, 就是将学习样本 W 中的各点 (x_i, y_i) , 在式 (6) 的 $\lambda + \tau$ 维信息监控空间 Ω 中, 分别用式 (8), 进行信息扩散。

$$\begin{aligned} \mu(x_i, y_i, u, v) \\ = \left(\prod_{j=1}^{\lambda} \exp\left[-\frac{(x_{ij} - u_{l_j})^2}{2h_{X_j}^2}\right] \right) \times \left(\prod_{k=1}^{\tau} \exp\left[-\frac{(y_{ik} - v_{t_k})^2}{2h_{Y_k}^2}\right] \right) \end{aligned} \quad (8)$$

累加 W 中的各样本点分别向 Ω 中监控点 (u, v) 扩散的信息, 得信息增量:

$$Q_{uv} = \sum_{i=1}^n \mu(x_i, y_i, u, v) \quad (9)$$

由 Ω 中所有监控点的信息增量构成的 $\lambda + \tau$ 维信息矩阵

$$Q = (Q_{uv})_{(u,v) \in \Omega} \quad (10)$$

称为给定样本 W 在信息监控空间 Ω 上的原始信息矩阵。

根据 Q 的性质, 选择适当的方式, 对其进行归一化处理, 可转变为一个表达因果律的模糊关系矩阵, 从而可近似推理, 由输入计算出输出。具体算法, 参见文献 [14]。显然, 原始信息矩阵 Q 的功能, 类似于一个函数, 决定着用模型计算的输出与训练样本的实际输出之误差。

基于 Q , 由输入向量 x_i 计算得的输出向量, 记为 \tilde{y}_i 。其与训练样本的实际输出之误差为:

$$\begin{aligned} \varepsilon_i = |\tilde{y}_i - y_i| \\ = \sqrt{(\tilde{y}_{i1} - y_{i1})^2 + (\tilde{y}_{i2} - y_{i2})^2 + \dots + (\tilde{y}_{i\tau} - y_{i\tau})^2} \end{aligned} \quad (11)$$

式中, $(\tilde{y}_{i1}, \tilde{y}_{i2}, \dots, \tilde{y}_{i\tau})$ 和 $(y_{i1}, y_{i2}, \dots, y_{i\tau})$ 分别为 \tilde{y}_i , y_i 两个点的坐标。于是, 用学习样本 W 得到的原始信息矩阵 Q 计算的输出, 其平均误差是

$$E = \frac{1}{n} \sum_{i=1}^n \varepsilon_i \quad (12)$$

式 (12) 中的 E , 就是地理空间信息扩散模型计算的输出与训练样本的实际输出之误差, 以减小此误差为目标, 用来控制模型进行样本学习的进程, 进行监督学习。

III. 扩散系数的优化算法

给定学习样本 W 和扩散系数向量 $h = (h_{X_1}, h_{X_2}, \dots, h_{X_\lambda}, h_{Y_1}, h_{Y_2}, \dots, h_{Y_\tau})$, 地理空间信息扩散模型的误差 E 即已确定。将地理空间信息扩散模型记为 Φ , 则监督学习用的误差可表为:

$$E = \Phi(W, h) \quad (13)$$

优化扩散系数的算法, 就是通过调整式 (13) 中的扩散系数 h , 尽可能地使误差 E 趋向于 0 的计算过程。根据 E 的变化趋势, 尝试性地调整 h , 称为搜索法。

本文建议, 用式 (1) 的平均距离公式, 分别计算学习样本 W 中各分量的扩散系数, 再分别优化它们。也就是说, 在地理空间信息扩散模型 Φ 中, 当我们用搜索法对某一分量的扩散系数 h 进行优化时, 其它分量在模型中的扩散系数仍用平均距离公式计算出来的结果。分别进行优化后形成的向量, 即为扩散优化系数向量。

不失一般性, 设 h 是平均距离公式计算的扩散系数, 在其它分量扩散系数不变的情况下, 地理空间信息扩散模型 Φ 的误差可记为 $e = \Phi(W, h)$, 用搜索法优化 h 的具体过程如下:

第1步: 确定停止搜索的准则

设 h 和 h' 是搜索时用到的两个扩散系数, 其差定义为

$$\rho = |h - h'|$$

取定一个很小的数 δ (例如, $\delta = 0.0001$)。当 ρ 小于或等于 δ 时, 停止搜索。

第2步: 初始化扩散系数

用平均距离公式 (1) 处理式 (2) 中的学习样本 W , 获得初始化扩散系数 h_0 。

第3步: 初始化搜索区间

令 $I_0 = 0$, $I_1 = 2h_0$, 将 (I_0, I_1) 选定为初始化搜索区间。

第4步: 选取两个测试点

使用某种搜索法, 在区间 (I_0, I_1) 中选取两个测试点。位于左边的点, 记为 h_L ; 位于右边的点, 记为 h_R 。

第5步: 计算测试点误差

分别将 h_L 和 h_R 代入地理空间信息扩散模型 Φ 中, 产生两个误差:

$$e_L = \Phi(W, h_L), \quad e_R = \Phi(W, h_R)$$

第6步: 判断是否继续搜索

如果 $\rho = |h_L - h_R| > \delta$, 跳到第8步, 继续搜索。

第7步: 优化扩散系数赋值

如果 $e_L < e_R$, 优化扩散系数取 $h = h_L$; 否则,

$h = h_R$ 。

输出优化扩散系数，停止搜索。

第8步：调整搜索区间

如果 $e_L < e_R$ ，令 $I_0 = I_0$ ， $I_1 = h_R$ ；否则，令 $I_0 = h_L$ ， $I_1 = I_1$ 。返回第4步，继续搜索。

IV. 算例演示

表 1 给出了一个输入和输出均为一维的学习样本 W 。输入值组成的集合 X 和输出值组成的集合 Y 是

$$X = \{x_1, x_2, \dots, x_{30}\} = \{2.6621, 12.6357, \dots, 17.5259\},$$

$$Y = \{y_1, y_2, \dots, y_{30}\} = \{70.0631, 164.0516, \dots, 97.8588\}$$

它们的容量，均为 $n=30$ 。 X 的最大值 $b_X=20.1385$ ，最小值 $a_X=0.5087$ ；对 Y 则有 $b_Y=240.7038$ ， $a_Y=35.5073$ 。

表 1. 供地理空间信息扩散模型使用的学习样本 W

No.	输入-x	输出-y	No.	输入-x	输出-y	No.	输入-x	输出-y
1	2.6621	70.0631	11	15.3787	158.5034	21	20.1385	240.7038
2	12.6357	164.0516	12	12.2721	218.1278	22	4.1030	100.4524
3	11.0625	113.2266	13	3.5797	63.7786	23	10.1387	218.0500
4	14.7150	149.1018	14	19.2520	181.3052	24	16.4955	141.1978
5	11.6647	141.3266	15	4.5223	134.9897	25	16.5713	231.7838
6	12.1015	171.9499	16	17.2097	124.9565	26	16.1310	129.9850
7	3.9335	82.4262	17	17.8633	99.2571	27	12.2141	198.6481
8	9.8980	226.6308	18	11.1781	182.3973	28	1.4417	35.5073
9	13.4709	171.3853	19	2.3603	79.7147	29	15.4163	164.9526
10	0.5087	77.2911	20	13.9994	174.1485	30	17.5259	197.8588

用式 (1) 中的 $h=2.6851(b-a)/(n-1)$ ，得扩散系数：

$$h_X = \frac{2.6851(b_X - a_X)}{n-1} = \frac{2.6851(20.1385 - 0.5087)}{30-1} = 1.8175$$

$$h_Y = \frac{2.6851(b_Y - a_Y)}{n-1} = \frac{2.6851(240.7038 - 35.5073)}{30-1} = 18.9991$$

(14)

取 $\delta=0.0001$ 为确定停止搜索的准则，我们以优化 h_X 为例，演示上述算法。

1. $2h_X=3.635$ ，得初始化搜索区间 $(0, 3.635)$ 。
2. 使用某搜索法，在区间 $(0, 3.635)$ 中选取两个测试点： $h_L=1.3886$ ， $h_R=2.2465$ 。
3. 维持 $h_Y=18.9919$ 不变，分别用两个测试扩散系数 h_L 和 h_R 替代地理空间信息扩散模型中的初始化扩散系数 $h_X=1.8175$ ，产生的误差分别是 $e_L=25.5366$ ， $e_R=27.8688$ 。
4. $\rho = |h_L - h_R| = |25.5366 - 27.8688| = 2.3322 > \delta = 0.0001$ 。将调整搜索区间，继续搜索。
5. 由于 $e_L < e_R$ ，搜索区间调整为 $(0, h_R) = (0, 2.2465)$ 。
6. 用同样的搜索法，在区间 $(0, 2.2465)$ 中选取两

个测试点： $h_L=0.8581$ ， $h_R=1.3883$ 。

7. 将新测试扩散系数代入地理空间信息扩散模型，分别计算它们相应的误差。
8. 如果 $\rho > \delta$ ，继续调整搜索区间，并进行搜索。
9. 进行到第 23 次选取测试点时，我们得到 $h_L=0.6086$ ， $h_R=0.6086$ ；相应的误差为 $e_L=21.7059$ ， $e_R=21.7059$ 。此时， $\rho = |h_L - h_R| = |0.6086 - 0.6086| = 0 < \delta$ 。
 $h_X=0.6086$ 即为优化后的扩散系数。

同理，得关于 Y 的优化扩散系数 $h_Y=17.82686$ 。

当我们用地理空间信息扩散模型学习表 1 中的样本 W 时，采用式 (14) 给出的扩散系数，产生的误差是 $e=26.3722$ ；而采用优化扩散系数 $h_X=0.6086$ 和 h_Y

$=17.82686$ 后，误差是 $e'=21.7529$ 。

由于

$$\varphi = \frac{|e' - e|}{e} = \frac{|21.7529 - 26.3722|}{26.3722} = 0.1752 \quad (15)$$

我们说，就此算例而言，扩散系数优化后，大约能减小 17% 的估计误差。

V. 结论与讨论

由于地理空间信息扩散模型是一个集值统计回归模型，模型计算的输出与训练样本的实际输出之间存在误差，因此，可用最小均方差来监督模型学习样本。此模型的误差，不仅与扩散方式有关，也受扩散系数的影响。用监督学习优化扩散系数，简单、有效。

将人们常用的平均距离公式用于扩散系数初值计算，由其两倍之长，与 0 点一同构成优化扩散系数的搜索区间，以某种搜索法逐步缩小搜索区间，即可找到非常满意的扩散系数。

由于具有非线性回归特性的扩散模型优于线性回归模型，而能学习矛盾样本的能力解决了人工神经网络不收敛的问题，扩散系数优化后的地理空间信息扩散模型，能更好地解决大灾中的信息孤岛问题。

必须有适当的背景数据支持，地理空间信息扩散模型才能发挥应有的作用。只有充分发挥互联网的作用，并有大量的利益相关者参与，才能在高风险地区获得与灾害有关的、可用的、实实在在的背景数据。这是防灾减灾救灾领域中一项基础性的工作。

事后的地理空间信息扩散不难，难的是第一时间用信息扩散模型填补空白处信息。只有充分发挥互联网的作用，第一时间将可进行灾情调查的地区之灾情，传到嵌入了地理空间信息扩散模型的网络平台上，才能解决大灾中的信息孤岛问题，为及时有效的应急救灾，提供科技支撑。

致谢

本项目受国家重点研发计划课题“重大自然灾害多层次精准救助关键技术研究”（编号：2017YFC1502902）和国家自然科学基金项目“独立多灾种年度财产损失综合风险信息扩散合成范式研究”（编号：41671502）联合资助。

参考文献

- [1] 黄崇福, “地理空间上的信息扩散及其在风险分析中的应用”, 一带一路背景下的风险分析和危机反应---中国灾害防御协会风险分析专业委员会第八届年会论文集(西安, 2018年10月20-21日), pp.1-7, October 2018.
- [2] D. J. Lieske, and D. J. Bender, “A robust test of spatial predictive models: geographic cross-validation,” *J. Environ. Inf.*, vol. 17, issue 2, pp. 91-101, 2011.
- [3] B. Purkait, S. S. Kadam, and S. K. Das, “Application of artificial neural network model to study arsenic contamination in groundwater of Malda District, Eastern India,” *J. Environ. Inf.*, vol.12, issue 2, pp.140-149, 2008.
- [4] C. F. Huang, “Information diffusion techniques and small sample problem,” *Int. J. Info. Tech. Dec. Mak.*, vol. 1, issue 2, pp.229-249, 2002.
- [5] 黄崇福, 自然灾害风险分析与管理. 北京: 科学出版社, 2012年, pp.220-221.
- [6] 刘新立, 黄崇福, 史培军, “对不完备样本下风险分析方法的改进及应用——以湖南省农村种植业水灾为例,” *自然灾害学报*, vol. 7, issue 2, pp.10-16, 1998.
- [7] 王新洲, 游扬声, “论信息扩散估计的窗宽,” *测绘科学*, vol. 26, issue 1, pp.16-19.
- [8] 沈赞琦, 陆余楚, “二维非均匀信息扩散及其应用,” *应用数学与计算数学学报*, vol. 21, issue 2, pp. 91-96, 2007..
- [9] I. Karimi, E. Huellermeier, “Risk assessment system of natural hazards: a new approach based on fuzzy probability,” *Fuzzy Sets and Systems*, vol. 158, issue 9, pp. 987-999, 2007.
- [10] C. F. Huang, C. Moraga, “A fuzzy risk model and its matrix algorithm,” *Internat. J. Uncertainty Fuzziness and Knowledge-based Systems*, vol.4, pp. 347-362, 2002.
- [11] D. J. Daley and D. G. Kendall, “Epidemics and rumors,” *Nature*, vol. 204, issue 4963, pp. 1118, 1964.
- [12] D. Margaritis, C. Vassilakis, and P. Georgiadis, “Recommendation information diffusion in social networks considering user influence and semantics,” *Soc. Netw. Anal. Min.*, vol. 6, issue 1, 108, 2016.
- [13] 刘禹, 李德毅, “正态云模型雾化性质统计分析,” *北京航空航天大学学报*, vol. 36, issue 11, pp.1320-1324, 2010.
- [14] 黄崇福, “借助背景数据的地理空间信息扩散技术,” *Journal of Risk Analysis and Crisis Response*, vol. 9, issue 1, pp.2-10, 2019.