城市道路复杂度的浮动车自适应采样算法

曹 闻 $^{ ext{0}}$,彭 煊 $^{ ext{2}}$,孟伟灿 $^{ ext{0}}$

(①信息工程大学测绘学院,郑州 450052; ②61922 部队,北京 100120)

【摘 要】基于浮动车的城市交通信息采集技术是智能交通系统获取实时交通信息的重要手段之一。针对浮动车的城市交通信息等间距采样的不足,本文设计了一种基于城市道路复杂度的自适应采样算法:①根据道路属性定义道路结点对城市道路网络复杂度的影响因子;②利用四叉树对城市道路网络复杂度进行描述;③根据浮动车的瞬时速度和道路复杂度自适应计算浮动车的采样周期。通过仿真和试验表明,新算法能够在不同复杂程度的道路情况下自适应提供有效、可靠的采样周期。

【关键词】智能交通系统;浮动车;自适应采样;道路复杂度

【中图分类号】P208; P285.3

【文献标识码】A

【文章编号】1009-2307(2012)04-087-04

1 引言

城市交通信息是智能交通系统的基础核心数据,而交 通信息的采集则是智能交通系统的首要任务。目前交通信 息采集技术大致可以分为基于人工监视的采集技术、基于 交通监控设施的采集技术、基于浮动车的采集技术。基于 人工监视的采集技术主要依靠特定人员通过视频监控或实 地交通协管进行人工判定交通流量、平均车速、车辆数目 等信息,这种方法在一定程度上充分利用了人类经验,但 是采集信息的准确度取决于作业员的业务水准,同时又存 在工作量大以及不能全天候工作的缺点;基于交通监控设 施的采集技术主要依托环形线圈检测[1]、雷达测速仪[2]、 无线射频识别[3]、视频自动检测[4]等技术手段进行采集相 关交通信息。虽然这些技术各有特点,但是都存在一定缺 陷,同时硬件投入比较大、成本高;基于浮动车的采集技 术[5-7] 主要依靠安装在车辆上的 GPS 和通讯模块进行采集车 辆行驶信息,配合车辆管理系统和地理信息系统通过数据 挖掘技术获取动态实时的交通信息,同时这种技术还具备 成本低、范围广、全天候、便捷等优点。如表1所示,基 于浮动车的交通信息采集技术更优于前两种采集技术,因 此其也成为了智能交通系统发展的热点和趋势。

表 1 交通信息采集技术的比较

影响条件	人工采集	交通监控设施				
		环形线 圈检测	无线射 频识别	雷达 测速仪	视频自 动检测	浮动车
交通流量采集	难	易	易	易	中	易
车辆信息描述	难	难	中	难	易	易
车速采集难度	难	易	难	易	中	易
全天候、全天时	×	$\sqrt{}$	$\sqrt{}$	$\sqrt{}$	×	\checkmark
成本状况	中	高	高	高	高	低

浮动车的行驶状态是连续的,基于浮动车的应用需要



作者简介: 曹闻(1979), 男,山东泰安人,讲师,博士,主要研究方向: 遥感图像处理、车载导航地理信息系统。 E-mail: speechofsilva@126.com

收稿日期: 2011-2-18

对浮动车反映的连续信息进行离散采样,而采样周期又决定着交通信息的质量和系统服务器的负载: 采样周期过小会造成较大的数据冗余和无线网络的负载过重; 采样周期过大则不能准确表达交通信息。因此,如何确定浮动车的采样周期是交通信息获取技术中的重点分支。

目前交通信息获取中采用的浮动车采样方法一般采用主观经验^[6] 或概率统计^[5] 6] 确定采样周期,然后以等距方式进行离散化采样,而浮动车的运动状态是复杂的(譬如有的速度快,有的速度慢),对不同目标或同一目标不同时段采取相同的采样率会造成较大的资源浪费或降低数据精度。因此,本文根据浮动车瞬时行驶速度以及交通道路网络状况(如道路稀疏程度)设计了一种自适应算法,在不影响采样分辨率的基础上适时动态调节采样率,从而达到资源利用最大化的目的。

2 基于城市道路复杂度的自适应采样算法

浮动车信息采样率的确定理想原则应该是: 既能准确地描述交通信息,又能避免数据的冗余。因此对于不同类型的浮动车、不同复杂程度的城市交通需要制定自适应的采样周期,即保证一个合适的空间采样距离 S: 当目标运动过快时,采样周期应该比较小; 反之应该比较大。假设 t 时刻浮动车的速度为 v ,则浮动车的采样周期 I(t) 定义为 I

$$I(t) = \frac{S}{T} \tag{1}$$

为了防止因某个位置点上速度存在较大误差导致的采样周期过分抖动,同时考虑到浮动车的行驶状态具备马尔科夫和短时平稳性,因此本文选择一定时间间隔内历史轨迹的平均速度修正浮动车的采样周期。

$$I(t) = \frac{S}{\bar{v}} \tag{2}$$

其中,空间采样距离 S 是根据城市道路的复杂度进行确定,即不同复杂度的城市选取不同的空间采样距离 (道路交通发达的城市空间采样距离 S 相对较小;道路交通滞后的城市空间采样距离 S 相对较大)。

自适应采样的目的是针对不同类型的浮动车、不同复杂程度的城市交通制定自适应的采样周期,因此公式(2)中空间采样距离 S 的确定以城市道路交通网络的复杂度或稀疏稠密程度来确定。考虑到城市道路的复杂度与地形分析中从等高线提取 DEM 过程中地形复杂度 (S^{10}) 相似,因此参考地形复杂度的描述确定空间采样距离 S,其基本思想加下

1) 道路的离散化

假设某城市有 n 条道路 $\{R_1, R_2, \dots, R_n\}$, m 个交叉口结点。通常城市内道路交叉口越多,那么该城市的道路

网络就越复杂,那么这些交叉口结点 $\{N_1',N_2',\cdots,N_m'\}$, j = 0 的属性就可以被认为城市道路的复杂度的评价依据。另一方面,城市内不同等级、数量、属性的道路也反映着该城市的道路复杂度。例如,如果立交桥、高架桥较多,那么某种意义上该城市的道路网络也就越复杂。因此,我们可以像道路交叉口那样对这些道路进行离散化,传来。对应的属性作为评价城市道路复杂度的依据。考虑们先按照等距原则将没条道路分为三段:前段、中段、后段;然后再按照等距原则将没条道路分为三段:前段、中段、后段;然后再按照等距原则将没条道路分为三段:前段、中段、后段;然后再按照等距原则将没条道路分为三段:前段、中段、后段;然后再按照等距原则将没条道路分为三段:前段、中段、后段;然后再按照等距原则将没条道路分为三段:前段、中段、后段;然后再按照等距原则将没条道路分为三段:前段、中段、后段;然后段的离散点; N_{m+1} , N_{m+2} , N_{m+n} ,

2) 道路离散点的权值确定

经过数字化的道路不仅有空间坐标,而且有详细的属性信息,如: 道路等级、车道数、单双向等。众所周知,城市内道路的数量及属性的不同决定着该城市道路网络的复杂度,因此由道路本身离散化后的离散结点对道路网络复杂度的影响因子 c 就可以参考这些道路属性。

定义离散点所关联道路的实际道路拥塞程度(可根据实时路况信息系统的历史信息对每条道路定期更新)、个数因子、等级因子、车道因子、宽度因子等为 f_{ij}^l , f_{ij}^e , $f_{ij}^$

$$f_{ij}^2 = \frac{C(R_i)}{C_{max}} \tag{3}$$

离散点关联道路属性的权值分别为 λ_j^1 , λ_j^2 , λ_j^3 , λ_j^4 , λ_j^5 , \cdots , 其中权值满足如下条件:

- ① $\lambda_0^k \ge \lambda_1^k$, λ_2^k , λ_5^k , $\lambda_6^k \ge \lambda_3^k$, λ_4^k ; $k = 1, 2, 3, 4, \cdots$ ② $\lambda_{ii}^1 \ge \lambda_{ii}^2 \ge \lambda_{ii}^3 \ge \lambda_{ii}^4 \ge \lambda_{ii}^5 \ge \cdots$
- 则每个道路离散结点对道路网络复杂度的影响因子 c_{α} 为:

 $c_{ij} = \lambda_{j}^{1} f_{ij}^{1} + \lambda_{j}^{2} f_{ij}^{2} + \lambda_{j}^{3} f_{ij}^{3} + \lambda_{j}^{4} f_{ij}^{4} + \lambda_{j}^{5} f_{ij}^{5} + \cdots \qquad (4)$ 其中 i=1 , 2 , \cdots , n , \cdots , m+n 表示离散结点的个数 , j=1 , 1 , \cdots , 6 表示结点在道路的位置属性。

3) 基于四叉树的自适应复杂度

在完成城市内部道路的离散化之后,将城市的最小外接矩形扩展为正方形,然后如图 2 所示对城市区域按照四等分原则^[9,10]进行剖分,基本过程是:

①首先分析剖分的每级格网内的离散点的分布情况,计算落在格网内部所有离散点对城市道路网络复杂度的影响因子 c_{ij} ,从而得到该格网的道路网络复杂度影响因子 c_{ki} :

$$c_{kt} = \begin{cases} \frac{1}{N_{kt}} \sum c_{ij}; \ N_{kt} \neq 0 \\ T - 1; \ N_{kt} = 0 \end{cases}$$
 (5)

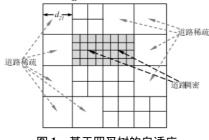


图 1 基于四叉树的自适应 复杂度原理示意图

的 4 个子格网,再将格网内所有的道路离散点配置到新的子格网内,然后将该格网的属性修改为根,而新的 4 个子格网的属性配置为叶;③重复①和②步骤,直到所有的格网的道路网络复杂度影响因子 c_{μ} < T 为止。

经过如上四叉树的剖分,可以构造城市多分辨率的道路复杂度,即: 道路稀疏的地方,格网间距大一些; 道路稠密的地方,格网间距大一些; 道路

经过基于四叉树的自适应复杂度分析之后,定义自适应采样原理的空间采样距离 S 为第 k 级第 t 个格网的间隔距离 d_k 的 n 分之一(本文试验分析中 n=3)。

$$S = \frac{d_{kt}}{n} \tag{6}$$

其中 t = 1 , 2 , … , 2^{2k} 。

这样当所有的浮动车在启动时,应用服务器向浮动车发送每个浮动车所在城市道路网络复杂度四叉树;然后根据浮动车的运动位置和四叉树结构实时更变公式(2)中自适应采样的空间采样距离S,在实时更新空间采样距离S时自动将当前浮动车位置作为采样点,以达到根据城市道路网络复杂度最优自适应采样的目的。

3 试验分析与结论

浮动车的属性信息主要应用于车辆监控管理和实时路况信息,其属性信息主要包括位置、时间、速度、行驶方向等,其中位置和行驶方向主要应用于判断浮动车处于什么位置或者在什么道路上沿什么方向行驶,而速度是获取交通信息的重要依据之一。浮动车采样评价标准应该根据应用综合考虑数据量、数据精度、网络负载等因素,本文根据实际应用和浮动车的属性信息制定了道路覆盖百分比、地图匹配精度和采样周期速度方差作为评价标准。

3.1 评价标准

1) 道路覆盖百分比 P。

将城市内每条道路等分为 N_p (本文 $N_p=5$) 段,假设浮动车原始数据点正确匹配到其相应的道路上。统计每条道路 R_i 等分段上存在原始数据点及采样点的数量 N_0 (R_i) 和 N_s (R_i),则道路覆盖百分比 P_s 定义为:

$$P_s = rac{\displaystyle\sum_{i=1}^n N_s(R_i)}{\displaystyle\sum_{i=1}^n N_0(R_i)}$$
 (7)
或大表明采样数据更接近于原始数据的

通常意义上 P_s 越大表明采样数据更接近于原始数据的运动轨迹。

2) 地图匹配精度 ρ_m

由于地理环境的影响,定位设备的定位精度一般存在着系统误差和随机误差,同时描述道路交通空间关系的电子地图存在着几何位置误差,因此行驶在道路上的浮动车实际位置往往不在电子地图的道路上^[11]。地图匹配技术的基本思想是通过特定的模型和算法将浮动车位置与电子地图上的道路相关联,是解决如上问题的技术手段。目前的地图匹配算法更多的依赖于数据采样频率(理想状态下的地图匹配算法更多的依赖于数据采样频率(理想状态率起大,算法匹配精度越低。由于浮动车采样频率通常大于 1s,因此需要制定合理的地图匹配算法进行匹配。如果仅采用单点进行地图匹配,匹配精度多低,并且无法对各种浮动车采样算法进行评价,因此本文采用基于 Curve-Curve 距离的匹配算法^[11]作为评价算法。

地图匹配精度 ρ_m 定义为采样数据在某种地图匹配算法下的正确匹配数据点个数 N_r 与所有采样数据个数 N_a 之比,即:

$$\rho_m = \frac{N_r}{N_c} \tag{8}$$

3) 采样周期速度方差 σ_{ij}

假设浮动车原始数据点速度集合为 $\{v_1, \cdots, v_i, \cdots, v_i\}$ v_n , 经过时间间隔 I(t) 采样后的数据点速度集合为{ v'_1 , \cdots , v'_{i} , \cdots , v'_{m} , 采样样本点速度 v'_{i} 在原始数据点集合 中的数据点速度为 v_i ,该点所属采样周期范围内的数据点 速度集合为{ $v_{i-I(i)/2}$,..., v_i ,..., $v_{i+I(i)/2}$ },则采样样本点速 度 v'_i 在该采样周期内的速度均值 $\mu(v'_i)$ 为:

$$\mu(v_{j}^{\prime}) = \frac{1}{I(t)} \sum_{k=-I(t)/2}^{I(t)/2} v_{i+k}$$
 (9) 则采样周期速度方差 σ_{v} 为:

$$\sigma_{v} = \frac{1}{m-1} \sqrt{\sum_{j=1}^{m} (v'_{j} - \mu(v'_{j}))^{2}}$$
 (10)

3.2 试验与结论

本文试验集是采集于 AM 06: 30 - PM 21: 00 时间段、 覆盖河南省郑州市市区的 GPS 信息点集(采样周期为 1s)。 通过3.1 叙述的评价标准对本文所提算法与传统采样算 法[5 6](采样周期为3s、5s、8s、10s、15s、20s和30s)进行 了如下试验比较(如表2所示)。

- 1) 道路覆盖百分比方面: 新算法的道路覆盖百分比在 等间距 8s 和 10s 之间,其采样数据量相对比较小,能够达 到减少数据冗余的目的,但是由于新算法认为道路对城市 道路网络复杂度影响主要反应在道路交叉口连接处,而连 接之间的中间段可以看作为整体,其对应的空间采样距离 相对比较大,因此道路覆盖率会有所降低,同时在这一部 分的数据量也会得到降低。
- 2) 地图匹配精度方面: 随着采样周期的增加,采样点 的地图匹配精度会大大降低,其原因是仅使用单点信息进 行地图匹配的精度相对比较低。新算法虽然采样点的整体 个数得到了减少,但是在道路交叉口附近的采样点相对增 多,那么在地图匹配算法中就可以引入历史轨迹信息提高 匹配精度; 而在每条道路交叉口两个连接段之间的路段采 样点个数较少,但在地图匹配中可以依托道路交叉口匹配 结果加入行驶在同一条道路上的经验规则来提高匹配 精度。

表 2 浮动车采样算法评价参数

采柏	算法	道路覆盖 百分比	地图匹配 精度	采样周期 速度方差
	3s	99.3%	91.6%	0. 07
等 间	5s	96.7%	88.3%	0. 10
距	8s	93. 2%	82.0%	0.11
采	10s	88.6%	80.9%	0. 13
样	15 s	82.4%	75.3%	0.16
算 法	20s	74.0%	69.6%	0. 20
74	30s	65.2%	61.2%	0. 24
新自适应	Z采样算法	89.8%	88.7%	0.08

3) 采样周期速度方差方面: 随着采样周期的增加,采 样周期速度方差会成递增趋势。速度方差越大表明采样样 本运动趋势与原始运动趋势越之间存在较大的差异,反之 就越接近。那么在交通信息提取时,就会引起较大的奇异, 产生错误的结果。通常在道路交叉口附近的速度变化相对 比较大,而中间路段速度相对稳定的因素,而新算法正是 考虑到了这一特点,将交通道路信息引入采样算法,可以 大大提高采样结果与原始结果运动趋势的相似性。

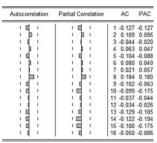
结束语

综上所述,新算法不仅考虑到浮动车速度变化因素对采 样周期的影响,而且将城市内道路分布及道路属性等因素引 入采样算法,使得确定的自适应采样周期可以获得较高的数 据精度,为浮动车采样周期提供了一种新的解决方式。

参考文献

- [1] Lelitha Vanajakshi ,L R Rilett. Loo Pdetector data diagnostics based on conservation-of-vehicle principle [J]. Transportation Research Record: Journal of the Transportation Research Board 2004 ,1870: 162-169.
- Hao Zhang ,Wen Yu ,Xiaowei Sun. A Novel Method for Background Suppression in Millimeter-Wave Traffic Radar Sensor [C]//11th International IEEE Conference on Intelligent Transportation Systems 2008, 2008: 699-704.
- Mijeom Kim ,Jinsoo Park ,Jaeyoung Oh ,Hakjin Chong , Yoonkee Kim. Study on Network Architecture for Traffic Information Collection Systems Based on RFID Technology [C]//Asia-Pacific Services Computing Conference 2008, APSCC '08 JEEE 2008: 63-68.
- [4] Guohui Zhang ,P. R. Avery ,Yinhai Wang. Video-Based Vehicle Detection and Classification System for Real-Time Traffic Data Collection Using Uncalibrated Video Cameras [J]. Transportation Research Record: Journal of the Transportation Research Board 2008 ,1993: 138-147.
- Zhang Wei , Chang Ande , Jiang Guiyan. Sampling and Transmitting Intervals Optimization Based on GPS Equipped Floating Car [C]//Second International Conference on Intelligent Computation Technology and Automation ,Volume: 4 2009: 97-100.
- [6] 张存保 杨晓光 严新平. 浮动车采样周期优化方法研 究[J]. 交通运输系统工程与信息 2007 7(3).
- Jun Hong Xuedan Zhang Zhongya Wei Li Li Yong Ren. Spatial and Temporal Analysis of Probe Vehicle-based Sampling for Real-time Traffic Information System [C]// Proceeding of the 2007 IEEE intelligent vehicles symposium Jstanbul Turkey June 2009: 1234-1239.
- 刘春 孙伟伟 吴杭彬. DEM 地形复杂因子的确定及与 [8] 地形描绘精度的关系 [J]. 武汉大学学报·信息科学 版 2009 34(9):1014-1019.
- Jiang Wen ,Baoshan Zhu ,Fan Wang. Real-Time Rendering Of Large Terrain On Mobile Device [C]//ISPRS2008, BeiJing Volume: VII July 2008: 693-697.
- [10] Wen Jiang ,Wu Yuguo ,Wang Fan. An Approach for Navigation in 3D Models on Mobile Devices [C]//Proceedings of 3D City Models ,Road Databases and Traffic Monitoring-Concepts , Algorithms and Evaluation , ISPRS , Paris , France Volume: VIII September 2009 2009: 109-114.
- [11] QUDDUS M A , NOLAND R B , OCHIENG W Y. The effects of navigation sensors and spatial road network data quality on the performance of ma Pmatching algorithms [J]. Geoinformatica 2009, 13:85-108.

(下转第98页)



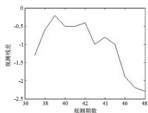


图 3 模型残差的自相关图 和偏相关图

图 4 预测残差图

预测残差见表 7 , 预测残差变化见图 4。从表 7 可以看出,预报值和实际观测值很接近,预测残差较小,说明模型是有效性的。但从图 4 可以看出,预测残差随着时间的推移越来越大,即模型预测的有效性逐渐降低,因此模型适合做短期预测,不适合长期的预测。

表7 建模数据

期数	观测值 (mm)	预测值 (mm)	预测残差 (mm)	期数	观测值 (mm)	预测值 (mm)	预测残差 (mm)
37	52. 5	53.8	-1.3	43	60. 0	61.0	-1.0
38	54. 3	54. 9	-0.6	44	61.5	62. 3	-0.8
39	55.9	56. 1	-0.2	45	62. 6	63.6	-1.0
40	56. 7	57. 2	-0.5	46	63.0	64. 9	-1.9
41	58.0	58. 5	-0.5	47	64. 1	66. 3	-2.2
42	59. 3	59. 7	-0.4	48	65. 3	67. 6	-2.3

4 结束语

时间序列分析是一种有效的变形预报方法,只需利用 已有观测序列即可对未来的变形进行有效的预测。时间序列 建模是基于平稳的序列建模,因此平稳性检验至关重要。 LM 检验、ADF 检验、残差序列图、自相关图函数和偏自相 关函数均是有效手段,应综合应用,确保对序列平稳性做 出准确判断,尤其应充分利用自相关图和偏自相关图的直 观性。

变形趋势的稳定或外界环境的稳定是确定时间序列预测的有效性是的前途,若外界环境发生重大改变,将导致整个变形轨迹出现根本性的改变,此时其预测将与实际情形产生较大差异,其差异对于工程灾害预防有着重要意义。

参考文献

- [1] 张正禄 横全义 文鸿雁 等. 工程的变形分析与预测 [M]. 北京: 测绘出版社 2007: 155-459.
- [2] 王惠穗. 变形监测数据分析与处理的几个关键: 问题 博士论文[D]. 上海: 同济大学 2006: 101-104.
- [3] 黄声享,尹辉,蔣征.变形监测数据处理[M].武汉:武 汉大学出版社 2003:101-104.
- [4] 张震海.灰色-时间序列分析组合模型在建筑物变形 预测中的应用研究[D].天津:天津大学 2007:17-22.
- [5] 叶斌.变形模型的分析研究以及变形的预测[D].上海:同济大学 2007:11-15.
- [6] 黄贤源 隋立芬 ,范澎湃 ,等 . 基于随机模型改正的有 色状态噪声处理新方法 [J] . 测绘学报 2009 34(1).
- [7] 王红 张正禄 张军.一种新的时间序列变点算法在变形数据分析中的应用[J].测绘科学 2008 33(3).
- [8] 刘娜 栾元重 横晓阳 等. 基于时间序列分析的桥梁 变形监测预报研究[J]. 测绘科学 2011,36(6).
- [9] 熊俊楠 汪泽根 徐忠印 等. 高层建筑物变形监测数据 处理方法研究及工程应用[J].测绘科学 2010 35(1).
- [10] 张健雄 蔣金豹 张建霞. 高层建筑沉降监测与灰色预测[J]. 测绘科学 2007 32(4).

Model of time series analysis for deformation forecast

Abstract: Time-series model is an effective method for deformation forecast. But in much application of time-series model for deformation forecast, the modeling process is not strict, lacking of statistics test and model test. So for the first time the paper cited LM test to determine the correlation of sequence, ADF test to determine the stability of sequence, the autocorrelation figure and partial autocorrelation figure to determine the type and the order of model preliminarily, and AIC criterion to determine the final model in deformation forecast, to assure every modeling process have strict statistical significance. With a grou Pof deformation monitoring data, it established the forecast model and forecasts. Comparing forecast results with monitoring data, it verified the effectiveness of forecast model, could be used in actual deformation forecast.

Key words: time series analysis; LM test; ADF test; autocorrelation function; stationarity

 $XIE~Rui^{\odot 2}$, LIU~Xiao- $qiang^{\odot}$, $LI~Yong^{\odot}$ (① Department of Civil Engineering , Henan Institute of Engineering , Zhengzhou 451191 , China; ② Department of Surveying and Geo-informatics Engineering , Tongji University , Shanghai 200092 , China; ③ Qingdao Research Institute of Survey and MappingShandong Qingdao 266071 , China)

(上接第89页)

Adaptive sampling algorithm of floating car based on the complexity of urban routes

Abstract: The technologies of traffic information collection based on GPS equipped floating car have become one of the main important means for real-time collecting traffic information in intelligent transportation system. According to the lack of the intervals of traditional algorithms based on equidistant sampling, an adaptive sampling algorithm based on urban routes' complexity was suggested in the paper. Firstly, according to the attributes of the route, the contributing gene of route's node was defined that is described as the complexity of urban routes' network; secondly, the complexity of urban routes' network was suggested using quarter tree; finally, the sampling intervals were adaptively obtained based on the instantaneous speed of floating car and the complexity of urban routes. The results of simulation and experiment showed that the efficient and reliable sampling interval could be adaptively determined under the circumstantialities of different routes' complexity.

Key words: intelligent transportation system; floating car; adaptive sampling; complexity of routes

CAO Wen[©], PENG Xuan[©], MENG Wei-can[©] (①Institute of Surveying and Mapping, Information Engineering University, Zheng-zhou 450052, China; ②Troops 61922, Beijing 100120, China)