

基于模拟退火的移动业务分布预测算法

孙娟娟 黄宗潘 杨大成
(北京邮电大学 93 信箱 北京 100876)

摘要: 该文提出了一种基于模拟退火的业务分布预测算法, 该算法采用模拟退火算法求解超定方程组得到业务密度值。并且提出了业务密度权值的计算方法, 从而在用模拟退火算法进行业务分布预测的基础上, 给出了利用业务密度权值进行业务分布预测的算法。采用实际统计数据计算结果表明, 该方法能够有效地预测业务的分布。

关键词: 无线网络仿真; 业务分布; 模拟退火; 权值

中图分类号: TN915.06

文献标识码: A

文章编号: 1009-5896(2008)09-2225-04

Algorithms of Mobile Traffic Distribution Forecasting Based on Simulated Annealing

Sun Juan-juan Huang Zong-pan Yang Da-cheng

(P.O.BOX 93 Beijing University of Posts and Telecommunications, Beijing 100876, China)

Abstract: An algorithm of traffic distribution forecasting based on simulated annealing is proposed, which is used to calculate the traffic density by simulated annealing. Furthermore, the weight value of traffic density is taken into account in this paper, which is introduced in traffic forecasting based on simulated annealing. The results of real statistical traffic show that the algorithms can forecast the mobile traffic distribution effectively.

Key words: Wireless network simulating; Traffic distribute; Simulated annealing; Weight value

1 引言

在无线网络仿真时, 必须根据业务分布情况, 才能够得到准确的网络容量和覆盖范围, 这需要更为准确的业务分布预测算法, 以预测出与实际网络运营状况相符合的业务分布。在无线网络的初期建设, 后期扩容和优化中, 无线网络仿真均起着重要作用, 而无线网络仿真是在业务分布预测的基础上进行的^[1-4]。

在无线网络仿真分析过程中有时假定业务分布是简单的均匀分布^[5], 有时所采用的业务分布尽管是采用非均匀分布, 但是却只是简单地假设业务密度随着与基站距离的增大而降低的函数分布, 比如负指数函数^[6]。通常的业务分布预测往往采用小区内业务分布平均, 或者分析区域内根据经验设置加权值进行业务分布, 或者采用综合计算法进行计算^[7, 8]。小区内业务平均分布只是一种简单处理, 不符合实际情况; 根据经验设置业务密度加权值要受到主观因素的影响; 对于综合计算法, 在本文的第2节进行了详细介绍, 分析了综合计算法存在的问题。为了解决综合计算法存在的问题, 本文提出了对根据基站平均业务密度分类后的基站簇构成的超定方程组, 采用模拟退火算法求解的办法, 从而给出了一种基于模拟退火算法的业务分布预测方法。

为了得到更方便易行的预测业务分布的算法, 本文提出

了一种业务密度权值的计算方法, 从而在用模拟退火算法进行业务分布预测的基础上, 给出了利用业务密度权值进行业务分布预测的算法。

本文在第2节给出了一个无线网络仿真中常用的业务分布预测算法—综合计算法, 分析其存在的问题。第3节引入模拟退火算法来求解超定方程组。第4节采用实际业务统计数据, 对根据平均业务密度进行分类后的基站簇采用模拟退火算法来进行业务分布预测。第5节提出利用业务密度权值进行业务分布预测, 并且将利用业务密度权值得到业务分布结果和模拟退火算法的结果进行比较。

2 业务分布预测方法

2.1 问题提出

业务分布预测基于数字电子地图进行。电子地图的精度由分辨率决定, 例如分辨率为 $20\text{m} \times 20\text{m}$ (地理单元 bin)。业务分布预测是在已知业务量或者用户数的情况下, 按照某种算法将业务量或者用户数分配到电子地图中的各个地理单元(bin)中。

移动用户的分布和地区分类密切相关, 如商业区、娱乐区的用户密度较高, 郊区的用户密度较低, 因此可将规划区分为若干个类别, 使每种类别具有不同的业务密度值。但是, 一般的数字电子地图并不包括上述地区类型信息, 通常只包括地物类型, 高度等数据。

尽管这些地物类型数据和地区类型是有区分的,但同时又是紧密联系的,所以业务分布预测往往是基于电子地图中地物类型数据进行。常用的方法是对同一类地物赋予相同的业务密度值。

2.2 业务分布预测方法

不同精度的电子地图可识别的地物类型种类有所不同,对地物类型的划分方法也存在差别。对于应用最广泛的20m精度数字地图,典型地物可分为:1(水域),2(海),3(湿地),4(郊区开阔地),5(城区开阔地),6(绿地),7(树林),8(40m以上高层建筑群),9(20-40m规则建筑群),10(20m以下高密度建筑群),11(20m以下中密度建筑群),12(20m以下低密度建筑群),13(郊区乡镇),14(城市公园)。

在进行业务分布预测过程中,常用的方法是综合计算法^[7, 8],其基本思路是将相同的地物类型对应相同的业务密度值。设电子地图中的地物种类为 n , t_1, t_2, \dots, t_n 代表不同地物的业务密度值,系统中的基站数量为 $m(m \gg n)$, b_1, b_2, \dots, b_m 是不同基站的忙时业务量统计值, $s_{ij}(i=1, 2, \dots, m; j=1, 2, \dots, n)$ 是第 i 个基站中第 j 类地物的面积,则有下式:

$$\mathbf{S}\mathbf{T}_R = \mathbf{B} \quad (1)$$

其中 $\mathbf{T}_R = [t_i]_{n \times 1}$, $\mathbf{S} = [s_{ij}]_{m \times n}$, $\mathbf{B} = [b_i]_{m \times 1}$ 。

由于实际网络中基站数量是庞大的,地物种类比较少,以20m精度的地图为例,一般地物种类为14类。在求解上述方程组时,由于该方程是线性超定方程,一般无解,通常是采用最小二乘法来计算,即

$$\min \left\{ F(t_1, t_2, \dots, t_n) = \sum_{j=1}^m \left(b_j - \sum_{i=1}^n t_i s_{ij} \right)^2 \right\} \quad (2)$$

根据多元函数微分学,式(2)的求解是多元函数 $F(t_1, t_2, \dots, t_n)$ 的驻点。

通过求解 n 元一次方程组,可得不同地物类型对应的业务密度值 t_1, t_2, \dots, t_n 。

以北京市900个GSM基站的话音业务统计结果为例,计算得不同地物类型的业务密度值如表1所示,分析区域不包括海(2)这种地物,故表中没有列出。

表1 最小二乘法计算不同地物的业务密度值

地物类型	1	3	4	5	6
业务密度(Erl/km ²)	-12.7	10.6	-23.6	84.5	3.56
地物类型	7	8	9	10	11
业务密度(Erl/km ²)	-1.37	191	74	35.7	40.6
地物类型	12	13	14	-	-
业务密度(Erl/km ²)	45.3	27.1	9.5	-	-

根据表1计算结果,计算出来的业务密度值有3项为负值,与实际情况不相符,实际业务密度值最小为0,不可能为负值。原因有两个,第一就是我们事先假设同一地物的业

务密度值是相同的,但是实际当中,同样为高层建筑,密集城区和郊区的业务密度值是不同的。第二个原因就是由最小二乘法算法本身所导致的,该算法是为了确保式(2)最小,业务密度的取值是没有限制的,可以为负值。正是因为这两个原因,导致计算结果与实际情况不相符,所以这种方法在实际应用中受到限制。

3 模拟退火算法

高温状态下的物质降温时其内部能量随之下降,如果降温过程充分缓慢,则在降温过程中物质体系始终处于平衡状态,从而降到某一低温时,其内能可以达到最小,称这种降温为退火过程。模拟退火过程的寻优算法称为模拟退火算法^[9, 10]。模拟退火算法的一般步骤如下:

- (1)随机给定初始状态,设定合理的退火策略;
- (2)对系统状态进行蒙特卡洛随机扰动,计算系统能量;
- (3)以某种方法确定新的系统状态;
- (4)重复步骤(2),步骤(3)直至系统达到平衡状态。

在具体应用时,对不同的问题需要设计不同的更详细的计算规则。按照模拟退火算法的一般原则,结合本文的实际问题,下面给出应用模拟退火算法求解式(2)的具体步骤^[9]:

(1)选取初始温度 T 和最低温度 T_0 ,给定式(2)的初始近似解 t ,设定退火策略为 $T(k) = \alpha T(k-1)$,其中 $\alpha \in [0.85, 0.98]$,计算 $E(\mathbf{T}_R) = \|\mathbf{S}\mathbf{T}_R - \mathbf{B}\|_2^2$,其中 $\|\cdot\|_2$ 为向量的欧氏范数;

(2)令 $\mathbf{T}'_R = \mathbf{T}_R + \Delta \mathbf{T}_R$,其中 $\Delta \mathbf{T}_R$ 是在 $[a, b]$ 上均匀分布的随机变量,计算 $E(\mathbf{T}'_R) = \|\mathbf{S}\mathbf{T}'_R - \mathbf{B}\|_2^2$ 和 $\Delta E = E(\mathbf{T}'_R) - E(\mathbf{T}_R)$;

(3)如果 $\Delta E < 0$,接受 \mathbf{T}_R 为新的近似解,否则随机选取 $\beta \in [0, 1]$,计算 $P = \exp(-\Delta E/T)$,如果 $P > \beta$,则接受 \mathbf{T}'_R 为新的近似值,否则仍选取近似解 \mathbf{T}_R ;

重复步骤(2),步骤(3),直至到达最低温度 T_0 ,或者满足 $E(\mathbf{T}_R) < \xi$,其中 ξ 是预先给定的充分小的正数。

为了保证业务密度值为正值,在步骤(2)中需要增加条件 \mathbf{T}'_R 必须要大于0,如果 $\mathbf{T}'_R < 0$,则该解不满足要求,需要重新寻找新解。

4 应用模拟退火的业务分布预测

为了保证不同类型地物的业务密度值与实际情况更相符,首先对所有的基站根据其平均业务密度值(业务量/基站覆盖面积)进行分类,平均业务密度值高的基站属于一类,平均业务密度值低的基站属于一类。分成 p 簇基站,每簇有 $m_k(k=1, 2, \dots, p)$ 个基站,即

$$\left. \begin{aligned} b_1 &= t_{k1}s_{11} + t_{k2}s_{12} + \dots + t_{kn}s_{1n} \\ b_2 &= t_{k1}s_{21} + t_{k2}s_{22} + \dots + t_{kn}s_{2n} \\ &\dots \\ b_{m_k} &= t_{k1}s_{m_k 1} + t_{k2}s_{m_k 2} + \dots + t_{kn}s_{m_k n} \end{aligned} \right\} \quad (3)$$

由于地物种类比较少,基站数量很大,上述方程组往往

是超定方程组。对该方程组的求解采用模拟退火算法。

本文采用第 3 节的模拟退火算法来求解式(3)。该算法还是采用局部搜索算法来寻找到一个高质量的解, 而且最终解往往与初始解存在某种依赖关系^[10], 为了提高最终解的质量, 并降低算法的运算时间, 本文采用模拟退火算法与最小二乘法相结合的办法求解超定方程组(3), 即模拟退火算法中的初始解 T_R 选择采用最小二乘法的解。相关参数设定如下: $T=1000$, $T_0=0$, $\alpha=0.92$, $[a,b]=[-1,1]$, $\xi=1$ 。最后计算出第 $j(j=1,2,\dots,p)$ 簇基站中不同地物的业务密度值 $[t_{j1}, t_{j2}, \dots, t_{jn}]$ 。

需要说明的是, 模拟退火算法是一种随机算法^[10], 每次迭代计算结果不同, 单个模拟的结果不能直接应用, 较好的方法之一是从若干次模拟的平均情况来观察, 本文选择了 10 次仿真结果的平均值作为业务密度值结果。

5 基于业务密度权值的业务分布预测

为了进一步分析不同地物之间业务密度之间的差异, 本文定义了业务密度权值。例如, t_{j1} 和 t_{j2} 是 1 类地物和 2 类地物的业务密度值, 则定义 k_{j1} 和 k_{j2} 分别是上述两类地物的业务密度权值, 具体关系如下:

$$t_{j1}/t_{j2} = k_{j1}/k_{j2} \quad (4)$$

业务分布预测的常用方法是给定某区域的业务总量, 从而预测不同地物类型的业务密度值。这里需要的一系列关键指标是不同地物的业务密度权值。这些权值描述了不同地物之间的业务密度值的差异。比如, 为了预测某一基站范围内的业务密度, 已知其承载(或计划承载)业务量为 b_i , 不同的地物面积 s_i , 权值为 k_i , t_i 代表业务密度常数。

$$b_i = k_1 t_i s_{i1} + k_2 t_i s_{i2} + \dots + k_n t_i s_{in} \quad (5)$$

只要上式中 $k_j(j=1,2,\dots,n)$ 的值确定后, 就可以确定 t_i , 则该小区内不同地物的业务密度值即可以确定, 下面需要研究 k_j 的计算方法。

共 p 簇基站, 计算出 p 簇业务密度值

$$[t_{ij}]_{p \times m} = [k_{ij}]_{p \times n} \mathbf{T} \quad (6)$$

其中 $\mathbf{T} = [t_i]_{p \times 1}$, 另外, 还需要对业务密度权值做下述约定: $k_{j1} + k_{j2} + \dots + k_{jn} = 1$, $0 \leq k_{j1}, k_{j2}, \dots, k_{jn} \leq 1$ 。

求解上述方程组, 得到 $t_j, k_{j1}, k_{j2}, \dots, k_{jn}$ 的取值。

对于第 j 簇基站中的第 i 类地物, 其业务密度权值为 k_{ji} 的概率为

$$P(k_i = k_{ji}) = \sum_{m_j} s_{ij} / \sum_{i=1}^m s_{ij} \quad (7)$$

而且

$$\sum_{j=1}^p P(k_i = k_{ji}) = \sum_{j=1}^p \left(\sum_{m_j} s_{ij} / \sum_{i=1}^m s_{ij} \right) = 1 \quad (8)$$

则第 i 类地物的业务密度权值为

$$k_i = E(k_{ji}) = \sum_{j=1}^p k_{ji} P(k_i = k_{ji}) \quad (9)$$

$k_i(i=1,2,\dots,n)$ 描述了不同地物类型之间业务密度的相对差异。

以第 2 节中综合计算法所采用的北京市 GSM 网络基站数据为例, 计算得不同地物类型的权值如表 2 所示:

表 2 不同地物的业务密度权值

地物类型	1	2	3	4	5	6	7
业务密度权值	0	-	0	0.04	0.07	0.04	0
地物类型	8	9	10	11	12	13	14
业务密度权值	0.24	0.18	0.12	0.12	0.11	0.08	0

为了对采用业务密度权值进行的业务分布预测效果进行衡量, 本文把利用业务密度权值计算的业务分布与第 4 节的业务分布预测结果进行比较。

设 t'_{ij} 为采用业务密度权值得到某 bin 点的业务密度值, t_{ij} 是根据第 4 节方法计算的业务密度值, 则残差 e_{ij} 定义为

$$e_{ij} = t_{ij} - t'_{ij} \quad (10)$$

图 1 所示是对第 2 节中综合计算法所采用北京市的 GSM 网络的基站统计数据, 采用第 4 节的业务分布预测方法与采用业务密度权值预测的的业务分布残差分布图。

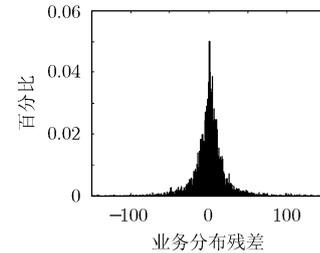


图 1 业务密度残差分布图

业务密度残差值在 $[-10\text{Erl}/\text{km}^2, 10\text{Erl}/\text{km}^2]$ 之间的 bin 数占总 bin 数的 58.7%。业务密度残差值在 $[-20\text{Erl}/\text{km}^2, 20\text{Erl}/\text{km}^2]$ 之间的 bin 数占总 bin 数的 80.6%。这说明采用基于模拟退火算法所得到的业务密度权值进行业务分布预测, 可以保证其业务密度值与直接采用模拟退火算法进行业务分布预测相近, 即基于模拟退火算法的业务密度权值, 可以用于业务分布预测。

基于模拟退火算法得到的业务密度权值, 用于业务分布预测, 比采用模拟退火算法直接进行业务分布预测, 更方便, 更具有一般性, 可以应用于不同的无线网络仿真区域。

6 结束语

业务分布预测是无线网络仿真和优化中的实际问题。本文提出的基于模拟退火算法和由此引伸的采用业务密度权值方法进行业务分布预测均可以有效地解决这一课题。尤其是根据业务密度权值的方法进行业务分布预测的方法更具

有一般意义,可以采用本文计算的业务密度权值对其他的区域进行业务分布预测,为无线网络仿真和优化提供业务分布基础数据。

参考文献

- [1] Shun-Ping Chung and Jin-Chang Lee. Performance analysis and overflowed traffic characterization in multiservice hierarchical wireless networks. *IEEE Trans. on Wireless Communications*, 2005, 4(3): 904-918.
 - [2] Santiago R C, Lyandres V, Raymond A, and Kontorovitch V Y. Effective base stations location and frequency assignment in mobile radio networks, EMC '03. 2003 IEEE International Symposium on Electromagnetic Compatibility, 11-16 May 2003, Vol.1: 653-656.
 - [3] 张长钢, 孙保红, 李猛等. WCDMA 无线网络规划原理与实践. 北京: 人民邮电出版社, 2005, 第9章.
 - [4] Laiho J, Raivio K, Lehtimaki P, Hatonen K, and Simula O. Advanced analysis methods for 3G cellular networks. *IEEE Trans. on Wireless Communications*, 2005, 4(3): 930-942.
 - [5] Spilling A R and Nix A R. Performance enhancement in cellular networks with dynamic cell sizing, PIMRC 2000, The 11th IEEE International Symposium on Personal. Indoor and Mobile Radio Communications, 18-21 Sept. 2000, Vol.2: 1589-1593.
 - [6] Ganesh R and Joseph K. Effect of non-uniform traffic distributions on performance of a cellular CDMA system. 1997 IEEE 6th International Conference on Universal Personal Communications Record, 12-16 Oct. 1997, Vol.2: 598-602.
 - [7] 张传福, 彭灿, 胡敖等. CDMA 移动通信网络规划设计与优化. 北京: 人民邮电出版社, 2006, 第7章.
 - [8] 啜钢等. CDMA 无线网络规划与优化. 北京: 机械工业出版社, 2005, 第5章.
 - [9] 李宝家, 刘昊阳. 超定方程的一种解法. 沈阳工业大学学报, 2002, (1): 76-77.
 - [10] 唐立山, 谢云, 尤矢勇等. 非数值并行算法(第一册)模拟退火算法. 北京: 科学出版社, 1994, 第2章.
- 孙娟娟: 女, 1972年生, 博士生、讲师, 研究方向为移动通信系统仿真、网络规划与优化.
- 黄宗潘: 男, 1982年生, 硕士生, 研究方向为移动通信网络规划与优化.
- 杨大成: 男, 1951年生, 教授, 主要研究方向为移动通信理论研究与应用.