

文章编号:1000-582X(2010)08-052-06

元胞自动机时空数据模型与预测方法

任海军^{a,b}, 张晓星^a, 周 涪^a

(重庆大学 a. 输配电装备及系统安全与新技术国家重点实验室; b. 软件学院, 重庆 400044)

摘 要:对标准元胞自动机(cellular automata, CA)模型进行扩展以满足具体行业地理信息系统(GIS)平台上时空动态模拟和预测应用要求。然后以电力行业配电网空间负荷预测为具体应用, 建立起了土地利用演化动态模型与电力负荷空间预测之间关系, 解决了时空数据管理中空间数据离散如何与属性数据进行交互问题。模拟出城市土地利用动态发展变化过程, 预测出规划区域中的各划分小区土地的未来使用类型, 建立起配电网空间负荷预测时空数据模型, 突破了传统时空分离电力预测的预测方法的局限。最后通过一个配电网空间负荷预测实例验证了该方法有效性。

关键词:元胞自动机; 时空数据模型; 土地使用决策; 电力负荷预测

中图分类号:TM 715

文献标志码:A

Cellular automata spatial-temporal data model and forecasting approach

REN Hai-jun^{a,b}, ZHANG Xiao-xing^a, ZHOU Quan^a

(a. State Key Laboratory of Power Transmission Equipment & System Security and New Technology ;
b. College of Software Engineering , Chongqing University, Chongqing 400044, P. R. China)

Abstract: The standard cellular automata(CA) model is expanded to meet requests of space-time dynamic simulation and forecast under the platform of geographic information system(GIS). Taking power load forecasting of the electric power industry as the specific application, the relations between dynamic model of the land use and power load space are established. The data and attribute data interactive discrete in spatial-temporal data management have been solved. The CA theory is practically used to simulate the process of urban land-use dynamic development, to forecast future land-use types of each small-area, to establish spatial load forecasting model. It breaks through the localization of all kinds of forecasting methods of traditional space-time separation power prediction. The effectiveness of the prediction method is verified by example.

Key words: cellular automata; spatial-temporal data model; land-use decision making; electric load forecasting

随着地理信息系统(geographic information system, GIS)在能源、交通、规划等行业大规模的应用, 使得空间数据收集相对以前容易得多, 但是由于这些系统只能存储静态的空间信息, 在反应动态变

化和数据预测推理等方面受到很大制约。时空信息系统相关技术发展进一步推动了时空数据挖掘技术相关的研究, 使得数据挖掘的研究对象也从事务型数据库扩展到了时空数据模型以及时空数据库应用

收稿日期:2010-03-20

基金项目:国家自然科学基金资助项目(50607023)

作者简介:任海军(1978-), 男, 重庆大学博士研究生, 主要从事地理信息系统与电力系统负荷预测研究,
(E-mail)jhren@cqu.edu.cn。

研究,结合时间和空间两个因素的时空数据模型和以及基于该数据模型的相关的行业应用成为了目前研究热点^[1]。

由于这些行业应用 GIS 平台都缺乏时空挖掘和时空动态模拟能力,解决的主要方法是将 GIS 和时空过程模拟模型进行集成。目前,国内外许多学者研究模拟城市的发展主要有:国内武大研究方法是先将元胞自动机(cellular automata, CA)模型进行扩展,提出基于地理实体环境的 CA 模型演变规则,再与 GIS 集成,模拟和预测长江三角洲某地区城市建设用地规模和布局^[2];信息工程大学和郑州大学采用以郑州市为具体事例,设计了城市 CA 模型,借助 VB 6.0 和 MapX 5.0 具体建立了一个与 GIS 无缝集成的 2 维 CA 模拟系统,并结合郑州市 2005—2007 年遥感影像图和土地利用图,对郑东新区进行了模拟,取得较好的效果^[3]。国外目前主要集中在扩展元胞自动机模型如何用于土地开发利用进行数据挖掘方面的研究^[4];以及时空数据模型和时空分析处理方面研究^[5]。传统的时空过程模拟模型主要是系统动力学模型,这些模型都缺乏对时间和空间的统一表达,主要是从动态性出发将时间仅作为一个单独的变量纳入到系统方程中,结果导致是时空数据分离;此外,这些应用行业 GIS 对时空的表达都是离散的,而传统的过程模拟模型基本上都是连续模型微分方程,故也很难集成。元胞自动机(CA)模型是一种离散动态模拟方法,可以很好的解决立时空一体化和过程模拟离散化的问题,将 CA 模型与行业具体应用 GIS 集成,主要是将 CA 模型作为动态空间模拟的一种框架纳入具体行业 GIS 中,一方面增强了行业 GIS 的时空动态建模功能;另一方面,GIS 提供的强大空间数据处理能力可以为 CA 模型准备数据和定义有效的元胞转换规则以及对模拟结果进行可视化挖掘,方便 CA 模型进行预测等决策支持研究。

采用扩展 CA 模型与多维 GIS 系统集成,构成了基于 CA 的多维时空数据模型,提出了基于该模型的模拟和预测原理,最后结合电力行业应用具体实例,给出了基于时空数据模型的配电网空间负荷预测应用,最后利用实例验证了该时空数据模型进行配电网空间负荷预测的方案可行性。

1 元胞自动机的定义

元胞自动机是一种构成条件相对宽松甚至近乎模糊的一种数理模型。元胞自动机是一个多学科交叉结合的“混合体”,是物理学家、数学家、计算机科

学家和以及生物学家共同研究的科学成果。因此。对元胞自动机的含义也存在不同学科上的解释,物理学家将其视为离散的、无穷维的动力学系统;数学家将其视为描述连续现象的偏微分方程的对立体,是一个时空离散的数学模型;计算机科学家将其视为新兴的智能计算方法;而生物学家则将其视为生命现象的一种抽象模型。下面给出几种常见的定义^[6]。

1.1 元胞自动机的物理学定义

元胞自动机是定义在一个由具有离散、有限状态的元胞组成的元胞空间上,并按照一定局部规则在离散的时间维上不断演化的动力学系统规则模型。具体讲,构成元胞自动机的部件被称为“元胞”,每个元胞具有一个确定状态。这个状态只取某个有限状态集中的某一个,例如或“YES”或“NO”,或者如 1024 灰度颜色数据中的一个离散数据,诸如此类;这些元胞按照确定的规则排列在“元胞空间”的空间网格上;这些元胞的各自的状态随着时间而发生变化。而依据一个动态局部规则下进行同步动态更新,结果是,一个元胞处于某时刻状态只取决于该元胞的上一时刻状态,以及以及该元胞所有邻居元胞的状态;元胞空间内的各个元胞依照确定的局部规则进行同步的状态更新和转换,于是整个空间上离散的元胞空间就表现为在离散时间维上的动态变化,这种数据模型就为时空状态变化模拟和动态演化奠定了理论基础^[7]。

1.2 元胞自动机的数学定义

美国数学家 L P Hurd 和 K Culik 等人在上世纪 90 年代初,对元胞自动机分别从集合论和拓扑学等角度进行了严格地描述和定义。

1.2.1 基于集合论的定义

设 d 代表元胞空间的维数, k 代表元胞的状态,并在一个有限集合 S 中取值, r 表元胞的邻居半径。 Z 是整数集,表示一维空间, t 代表时间。从描述和理解方便出发,选择一维元胞空间考虑元胞自动机集合定义,即假设 $d = 1$ 。那么所有元胞空间就是在—维空间,将整数集 Z 上的状态集 S 的分布则可记为 S^Z 。元胞自动机的动态演化就是离散时间上状态动态发展的变化,可以记为

$$F: S_t^r \rightarrow S_{t+1}^r. \quad (1)$$

所有动态演化过程又由各个元胞的局部演化规则 F 所决定的。这个局部函数 F 通常又常常被称为局部规则变化函数。针对一维元胞空间,元胞及其邻居可以记为 S^{2r-1} ,那么局部规则变化函数则记为

$$F: S_t^{2r-1} \rightarrow S_{t+1}. \quad (2)$$

对于局部规则 F 来讲,函数输入以及输出集均为有限集合,实际上,这个有限集合是一个有限的参照表。如: $r = 1$, 则 F 的形式表示为如下

$$\begin{aligned} & [0,0,0] \rightarrow 0; [0,0,1] \rightarrow 0; [0,1,0] \rightarrow 1; [1,0,0] \rightarrow 0; \\ & [0,1,1] \rightarrow 1; [1,0,1] \rightarrow 0; [1,1,0] \rightarrow 0; [1,1,1] \rightarrow 0. \end{aligned}$$

对元胞空间内的元胞,独立施加上述局部函数,则可得到全局的演化函数: $F(c_{i+1}^t) = f(c_i^{t-r}, \dots, c_i^t, \dots, c_i^{t+r})$, 其中 c_i^t 表示在位置 i 处的元胞, 由此就从数学意义上构成了一个元胞自动机模型。

1.2.2 元胞自动机的拓扑学定义

为描述和理解方便。同样假定维数 d 为 1。设 S 为 k 个元胞有限状态集。 Z 为整数全体的集合,称 Z 到 S 的映射的全体 S^Z 为元胞状态空间。显然 S^Z 就是用 S 中的符号组成的双侧无限的符号序列的全体,如一维元胞自动机的所有构形的集合,称 $a = (\dots a_{-1} a_0 a_1 \dots)$ 为构形空间中的点。

在 S^Z 空间中引进任意两点 X 和 Y 之间的距离为

$$d(x, y) = \sum \delta(x_i, y_i) 2^{-|i|}. \quad (3)$$

其中当 $x_i = y_i$ 时, $\delta(x_i, y_i) = 0$, 当 $x_i \neq y_i$ 时, $\delta(x_i, y_i) = 1$ 。随后,在 S^Z 中可以建立起开、闭、紧等拓扑概念。

在 S^Z 中定义移位算子 δ 为

$$\delta(x_i) = x_{i-1} \in Z. \quad (4)$$

若连续映射 $F: S^Z \rightarrow S^Z$ 产生与 δ 可交换,即 $F_\delta = \delta_F$, 或对任意的 $x \in S^Z$ 均有 $F(\delta(x)) = \delta(F(x))$, 则称 F 为元胞自动机。

对于以上数学拓扑定义,很容易将它扩展到一个多维空间,所需要要做的工作只是将 S^Z 记为 $S^{Z \wedge d}$, $S^{2^{r+1}}$ 记为 $S^{(2^{r+1}) \wedge d}$ 等,同时对其他描述作相应改变即可达到要求。

元胞自动机基本的组成为元胞、元胞空间、邻居及规则四部分。简单讲,元胞自动机可以定义为由元胞空间和定义于该空间的变换函数所组成,如图 1 所示^[8]。

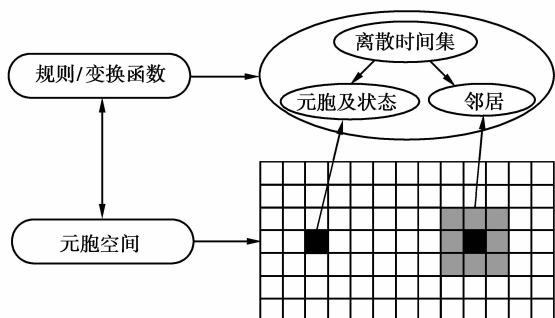


图 1 元胞自动机构成图

2 扩展元胞自动机模型

在 GIS 应用系统中,目前实际所使用的时空数据基本上是基于地理信息系统数据模型来存储数据的,就必须将标准的 CA 模型的进行扩展以满足目前具体应用的 GIS 系统,达到能在 GIS 平台上进行基于该时空数据模型的模拟与预测应用,这需要进一步建立基于 GIS 平台时空数据挖掘模型和算法,要求应用于 GIS 平台上的元胞数据模型以及定义的状态转换规则都要有明确空间信息含义,这样元胞自动机模型才有具体的时间和空间意义,就必然要求对标准 CA 模型进行物理定义上进行扩展。从关联空间和时间属性数据和结合行业应用平台出发,从以下几个方面对标准 CA 模型进行扩展^[9]。

1) 扩展元胞空间^[10]。元胞空间是具有实际时空含义的和空间坐标位置的相对应的元胞空间。根据模拟和预测研究的对象需要,可以将元胞抽象成一维或者多维元胞空间。

2) 扩展元胞状态^[11]。标准的元胞状态集是一个有限的且相对离散的数据集合,将 CA 模型用来进行时空数据挖掘时,元胞的状态集合将扩展定义为表征空间实体或实体所具有的特定空间现象的数据指标参数或者空间对象实体编码或时空特性对象特定描述,如国土利用中的适宜性评价参数指标、土地利用转换规则、土地等级评价参数等都可以非常方便的用元胞状态集合来表示。

3) 元胞状态转换规则的扩展^[12]。标准的 CA 模型的转换规则一般是局部的转换规则,是由空间具体分布位置决定各自的转换规则。但实际的应用项目中,空间对象实体的动态演化和转换原因很多规则也相对复杂,除受局部元胞间相互作用的影响,还受外部整个大环境影响,所以具有空间特性的 CA 模型中状态转换规则必须兼顾微观和宏观多个层面,需要对规则进行约简,进行梳理找出最合理的规则才能更有效模拟出空间实体的真实演化过程。

4) 转换时间上的扩展^[13]。CA 模型中的模拟时间必须和空间实体动态演化中的真实世界时间建立相应的关系,否则,时空建模就失去了应用价值。一般来讲采用时空数据模型和其他模型预测数据来建立二者之间对应联系。

3 基于 CA 模型与 GIS 系统建立时空数据模拟与预测基本原理

采用扩展 CA 模型是对空间实体动态变化发展过程的高度概括和抽象,其规则的灵活性决定了模

型的普适性^[14]。作为一种动态时空建模框架和模型结构,扩展 CA 模型主要应用在与具体行业应用 GIS 系统中的一个软件模块而集成到具体的 GIS 系统整体中,主要用来增强行业的 GIS 的时空动态模拟与预测能力,这主要是由于 CA 模型具有可迭代运算的能力,这样结合 CA 模型的 GIS 系统可以模拟空间实体和空间对象各种时空变化过程,系统中的空间元胞将能实现各个实体的时空动态变化过程。CA 模型中的元胞主要是通过与邻域内的其他元胞的相互作用关系来模拟整个动态系统内部之间的互相关系和互相作用效果,这种相互作用也是随着时间变化而动态变化发展,可以通过定义不同 CA 在空间上的规则以及定义在这些规则转换时间,来达到模拟复杂的时空演化现象和时空变化过程^[15]。具体应用案例可以将空间 CA 建模框架集成到具体的 GIS 系统中形成时空动态模拟与预测通用软件中间件,用这个软件中间件上用户可以直接定义 CA 模型相应参数,定义出该 CA 模型的局部具体转换规则,主要功能集成在 GIS 系统中作为一个动态模拟和预测软件中间件,使得应用层用户可以根据具体的应用需求建立面向行业应用的专题模型与预测模型,从时空 2 个方面,挖掘出所需要的专题数据图,做到时空数据挖掘的目的,在应用的 GIS 系统上建立起时空数据的模拟和预测就是基于这个原理^[16],其模拟和预测流程如图 2 所示。

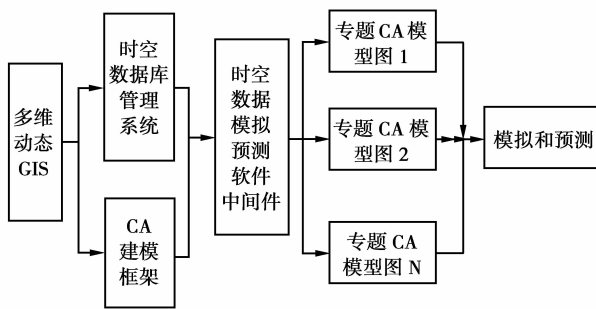


图 2 时空数据模拟和预测

4 基于扩展 CA 模型的配电网空间负荷预测

由于配电网空间负荷预测是对负荷时空特性的分布预测,是描述未来负荷的大小与时间、地点之间的关系,是城市配电网发展规划重要依据,但也是一个涉及时空方面多因素影响复杂问题^[17]。随着电力行业地理信息系统和相关多维时态地理信息系统技术成熟,以及在配网规划中应用,对配网规划的区

域的空间信息和历史负荷信息的收集和越来越方便,为空间负荷预测提供了基础条件,加上引入扩展元胞自动机来对小区未来土地使用类型进行模拟和预测。并根据城市发展的实际情况,确定空间 CA 的迭代时间和转换规则调整时间,进一步确定了动态的元胞自动机转换规则和时间,最后用预测年的用地面积乘以相应的预测出类型用地负荷密度旧可以得出该区域的空间负荷,基于时空数据模型的配电网空间负荷预测主要步骤如图 3 所示^[18]。

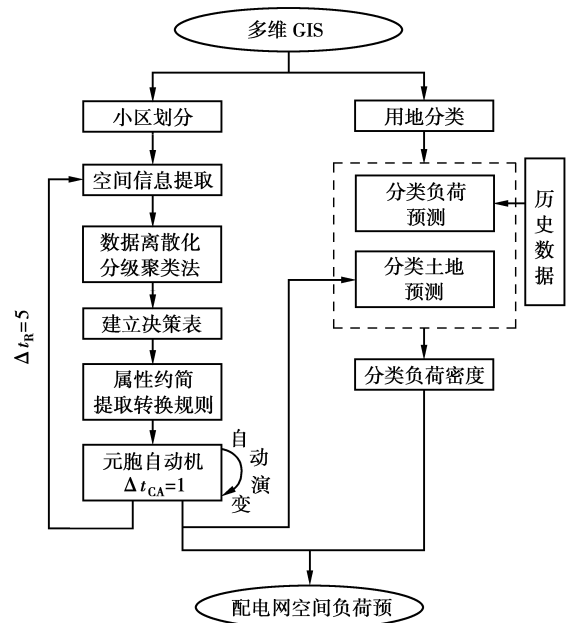


图 3 配电网空间负荷预测主要步骤

5 具体预测实际应用案例

以重庆江北某供电区域为例运用上述方法进行基于元胞自动机的时空数据模型建立时空数据模型和空间负荷预测模型。为简化元胞模型计算,根据预测区域的土地实际情况,将预测区域的划分采用规则的大小相等面积的正方形网格,共划分为 295 个 300 m×300 m 的小区,简化为其中每个小区代表一个负荷元胞,根据不同的负荷用电特性和土地利用类型,结合该预测区域的实际负荷统计数据,将该区域用户分为工业用电、商业用电、居民用电和市政用电 4 个大类,预测区域的元胞小区划分如图 4。

按照上述预测原理方法,利用获取的用地类型转换规则对新增负荷小区进行用地类型预测^[19],同时利用传统负荷预测方法进行负荷总量预测,得到各类用电负荷和用地面积的情况如表 1-2 所示。

- LI AI-MIN, LV AN-MIN, SUI CHUN-LING. Application of integrating GIS with cellular automata to dynamic models of urban growth [J]. Journal of Zhengzhou Institute of Surveying and Mapping, 2009, 26(3):165-169.
- [4] HU G Z. An extended cellular automata model for data mining of land development data [C]// The 5th IEEE/ACIS International Conference on Computer and Information Science, July 10-12, 2006, Honolulu, Hawaii, USA. [s. l.]: IEEE, 2006: 201-207.
- [5] ZANG D Y, ZHOU G Q. Area spatial object co-registration between imagery and GIS data for spatial-temporal change analysis [C]// IEEE International Geoscience and Remote Sensing Symposium, Barcelona, Spain, July 7-11, 2008, Boston, MA, USA. [s. l.]: IEEE, 2008: 2597-2600.
- [6] 周成虎,孙战利,谢一春. 地理元胞自动机研究[M]. 北京:科学出版社,1999.
- [7] Breukelaar R. Using a genetic algorithm to evolve behavior in multi dimensional cellular automata [C]// Proceedings of the 2005 conference on Genetic and evolutionary computation, June 25-29, 2005, Washington D C, USA. [s. l.]: IEEE, 2005: 107-114.
- [8] LO SC, HSU C H. Cellular automata simulation for mixed manual and automated control traffic [J]. Mathematical and Computer Modelling, 2010, 51(7/8):1000-1007.
- [9] HANSEN C, HTEBRANFUKER N, WIKENING W, et al. A method to expedite data acquisition for multiple spatial-temporal analyses of tissue perfusion by contrast-enhanced ultrasound [J]. IEEE Transactions on Ultrasonics Ferroelectrics Frequency Control, 2009, 56(3): 507-519.
- [10] SUN T D, WANG J F. A traffic cellular automata model based on road network grids and its spatial and temporal resolution's influences on simulation [J]. Simulation Modelling Practice and Theory, 2007, 15(7):864-878.
- [11] SU F H, YAO L K. Fractal cellular automata model and simulation [C]// International Conference on Computational Intelligence and Software Engineering, December 11-13, 2009, Wuhan, China. [s. l.]:IEEE, 2009:1-4.
- [12] 喻永平,陈晓勇,刘经南,等. 状态扩展元胞自动机模型在时空数据挖掘中的应用[J]. 武汉大学学报·信息科学版, 2008, 33(6): 592-595.
- YU YONG-PING, CHEN XIAO-YONG, LIU JING-NAN, et al. Application of Extended State Cellular Automata to Spatiotemporal Data Mining [J]. Geomatics and Information Science of Wuhan University, 2008, 33(6):592-595.
- [13] 杨丽徙,王金凤,陈根永,等. 基于元胞自动机理论的电力负荷空间分布预测[J]. 中国电机工程学报, 2007, 27(4):15-20.
- YANG LI-XI, WANG JIN-FENG, CHEN YONG-GEN, et al. Load spatial distribution forecasting model on cellular automata theory [J]. Proceedings of the CSEE, 2007, 27(4):15-20.
- [14] YANG X, YUAN J, ZHANG T, et al. Application of uncertainty reasoning based on cloud theory in spatial load forecasting [J]. Intelligent Control and Automation, 2006(2): 21-23.
- [15] WASILEWSKI J. Application of fuzzy cellular automata to spatial electric load forecasting[J]. Archives of Electrical Engineering, 2009, 58(229):83-96.
- [16] LIU Y, HE J. Developing a web-based cellular automata model for urban growth simulation [C]// International Symposium on Spatial Analysis, Spatial-Temporal Data Modeling, and Data Mining, October 28-30, 2009, Wuhan, China. [s. l.]:IEEE, 2009: 7492.
- [17] WASILEWSKI J, PAROL M. Prediction of spatial distribution of electric energy consumers with the use of fuzzy cellular automata [J]. Przegląd Elektrotechnology, 2009, 85(3):208-211.
- [18] LIM J U, JIANG J N. Bibliography review on applications of correlation analysis in power system operation and planning[J]. European Transactions on Electronical Power, 2010, 20(2):114-122
- [19] 周濂,李健,孙才新,等. 基于粗糙集和元胞自动机的配电网空间负荷预测[J]. 中国电机工程学报, 2008, 28(25):68-73.
- ZHOU QUAN, LI JIAN, SUN CAI-XIN, et al. Spatial load forecasting for distribution networks based on rough sets and cellular automata[J]. Proceedings of the CSEE, 2008, 28(25):68-73.
- [20] 乐欢,王主丁,肖栋柱,等. 基于空区推论的空间负荷预测分类分区实用法[J]. 电力系统自动化, 2009, 33(7): 81-85.
- YUE HUAN, WANG ZHU-DING, XIAO DONG-ZHU, et al. Vacant area inference based classification and sub area practical method for spatial load forecasting[J]. Automation of Electric Power Systems, 2009, 33(7):81-85.

(编辑 张小强)