复杂系统数据挖掘的多尺度混合算法

康 卓, 黄竞伟, 李 艳, 康立山

康 卓1, 黄竞伟2, 李 艳1, 康立山3 1(武汉大学 计算中心,湖北 武汉 430072)2(武汉大学 计算机学院,湖北 武汉 430072)
3(武汉大学 软件工程国家重点实验室,湖北 武汉 430072)
第一作者: 康卓(1970－),男,湖北武汉人,博士生,副教授,主要研究领域为演化算法,智能计算.
联系人: 康立山 E-mail: kang@whu.edu.cn
Received 2002-09-07; Accepted 2003-01-20

Abstract

Any complex system must be controlled by some basic laws, including macroscopic level submicroscopic level and microscopic level laws. How to discover its necessity-laws from these contingency - phenomena (observed data) is the most important task of data mining (DM) and KDD, and it is the goal of this paper too. Based on the evolutionary computation and natural fractal, a multi-scale dynamic prediction system is proposed, which models the macro-behavior of the system by ordinary differential equations while models the micro-behavior of the system by natural fractals. The financial data such as the stock market data of Jun'an stock price and the scientific observed data such as rainfall data of Wuhan in flood season are used as the test data for simulated test of analysis and prediction. The experimental results show that this system fits the data very well, and the simulated prediction is good too, even for modeling the time series with large undulating.

摘要

任何复杂系统都要受到某些基本规律的约束,包括宏观、中观与微观的多层次规律的约束.怎样从一个系统的这些偶然现象(观测数据)中找出它的必然规律,是知识发现(KDD)与数据挖掘(DM)的首要任务,也是研究目标.建立了一个基于演化计算与自然分形相结合的多尺度的动态预测系统.它以微分方程描述系统的宏观行为,以自然分形刻画系统的微观行为.同时,以股票市场数据(君安证券股票数据)和科学观测数据(武汉汛期雨量数据)为例,进行了分析与预测模拟.数值实验表明,该系统的描述(拟合)性能优越,即使是对起伏波动很大的时间序列,也能拟合得很好,预测效果也较好.

References:

[1] Iba H, Sasaki T. Using genetic programming to predict financial data. In: Angeline PJ, ed. Proceedings of the Congress on Evolutionary Computation. Piscataway: IEEE Press, 1999. 244~251.

[2] Hafner C, Frohlich J. Generalized function analysis using hybrid evolutionary algorithms. In: Angeline PJ, ed. Proceedings of the Congress on Evolutionary Computation. Piscataway: IEEE Press, 1999. 287~294.

[3] Yoshihara I, Numata M, Sugawara K, Yamada S, Abe K. Time series prediction model building with BP-like parameter optimization. In: Angeline PJ, ed. Proceedings of the Congress on Evolutionary Computation. Piscataway: IEEE Press, 1999. 295~301.

[4] Ferreira AR, da Silva. Evolving best-basis representations. In: Angeline PJ, ed. Proceedings of the Congress on Evolutionary Computation, Vol 1. Piscataway: IEEE Press, 1999. 302~309.

[5] Kang LS, Li Y, Chen YP. A tentative research on complexity of automatic programming. Wuhan University Journal of Natural Sciences, 2001,6(1-2):59~62.

[6] Cao HQ, Kang LS, Chen YP. Evolutionary modeling of systems of ordinary differential equations with genetic programming. Genetic Programming and Evolvable Machines, 2000,1(4):309~337.