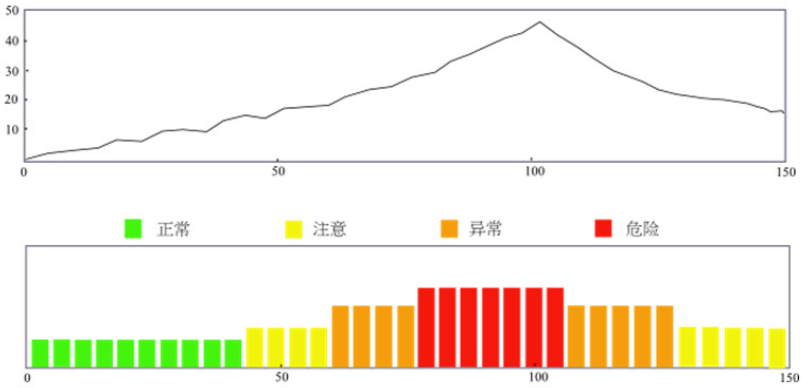


Fig.9 Test 1: Transaction response time and system output of forecast

图 9 测试 1:事务响应时间和模型预测输出

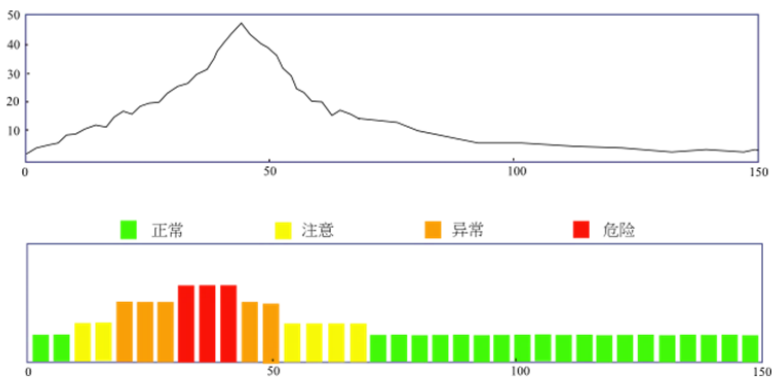


Fig.10 Test 2: Transaction response time and system output of forecast

图 10 测试 2:事务响应时间和模型预测输出

- 第 3 个测试场景为混合业务场景压力测试.

同时测试文件的修改与删除操作.每 10s 加载 10 个虚拟用户直到 50s 后,释放 30 个虚拟用户,然后在 10 秒后重新每秒加载 20 个用户.用户思考时间设置为实际值的 2%~5%.

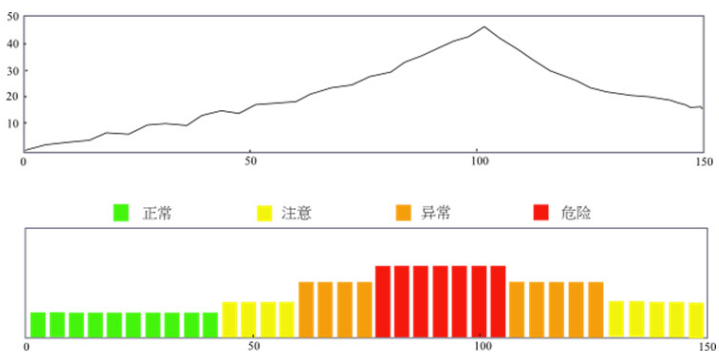


Fig.11 Test 3: Transaction response time and system output of forecast

图 11 测试 3:事务响应时间和模型预测输出

从实验数据分析我们看到:操作的平均事务响应时间在 10s 以内时,模型预测系统状态属于正常范围;10s~20s范围模型预测系统状态为注意;20s~30s范围模型预测系统状态为异常状态;30s以上为模型预测系统危险状态.另外,我们选取了系统状态预测为异常状态时的 CPU 使用率和 JVM 空间占用率数据分析,在该时间段,CPU 占用率超过了 75%,JVM 空间占用率超过 85%,系统处于严重状态.因此,模型评估和预测结果与我们的常识一致.上述实验结果表明:模型有效地评估了对系统的状态,特别是对危险状态进行了预测,因而可以较好地实时监控软件系统.

5 结束语

软件系统的状态评估机制是保障软件系统稳定运行的重要手段,然而随着软件系统规模不断增大,系统复杂性不断的提升,许多传统的评估方法的预测效果越来越不尽人意.这些方法主要的研究思路过于注重对软件系统的各个子系统的外在特征参数进行分析,建立简单模型对系统内部状态进行评估预测,这样往往不能充分挖掘大量数据中隐含的信息.本文则提出了一种基于隐马尔可夫模型的软件系统状态评估预测方法,该方法利用隐马尔可夫模型建立起系统外在状态(观测状态)和内部状态(隐藏状态)之间联系,从整体的角度对系统状态进行预测,而不拘泥于局部的特征,从而提高软件状态预测方法的准确性,便于软件系统管理和维护工作的开展.在今后工作中,我们将考虑进一步开发动态负载均衡算法以及相关程序包,并且结合软件系统的特征和本质对本算法进行优化,使其具有一定的普遍性和适用性.

References:

[1] Zhao D, Zhang WD, Liu J. The safety production information system based on B/S structure. China Public Security (Academy Edition), 2007,4:97−99 (in Chinese with English abstract).

[2] Chen DJ, Cao WG. The design and achievement of enterprise stock-sell-storage management system based on browser/server pattem. Modular Machine Tool & Automatic Manufacturing Technique, 2006,5:110 (in Chinese with English abstract).

[3] Oracle. Introduction to Oracle WebLogic Server, 11g Release. 2011.

[4] Zong Y, Jin P, Chen EH, Li H, Liu RJ. Fuzzy co-clustering algorithm for WebLogic. Journal of Electronics & Information Technology, 2012,34(3):543−548 (in Chinese with English abstract).

[5] Liu WX, Zhou JN, Zhang J. Application security of oracle database stored-procedure. Computer Systems & Applications, 2013,2: 80−83 (in Chinese with English abstract).

[6] Stephens SM, Chen JY, Davidson MG, Thomas S, Trute BM. Oracle database 10g: A platform for BLAST search and regular expression pattern matching in life sciences. Nucleic Acids Research, 2005,33(Suppl.1):D675−D679. [doi: 10.1093/nar/gki114]

[7] Rabiner LR, Juang BH. An introduction to Hidden Markov models. IEEE ASSP Magazine, 1986,3(1):4−16. [doi: 10.1109/MASSP.1986.1165342]

[8] Rabiner LR. A tutorial on Hidden Markov models and selected applications in speech recognition. Proc. of the IEEE, 1989,77(2): 257−286. [doi: 10.1109/5.18626]

[9] Li J, Najmi A, Gray RM. Image classification by a two-dimensional Hidden Markov model. IEEE Trans. on Signal Processing, 2000,48(2):517−533. [doi: 10.1109/78.823977]

[10] Sohrab PS, Xiang X, DeLeeuw RJ, Khojasteh M, Lam WL, Ng R, Murphy KP. Integrating copy number polymorphisms into array CGH analysis using a robust HMM. Bioinformatics, 2006,22(14):431−439. [doi: 10.1093/bioinformatics/btl238]

[11] Zhang JH, Zhang WB, Xu JW, Wei J, Zhong H, Huang T. Approach of virtual machine failure recovery based on Hidden Markov model. Ruan Jian Xue Bao/Journal of Software, 2014,25(11):2702−2714 (in Chinese with English abstract). http://www.jos.org.cn/1000-9825/4548.htm [doi: 10.13328/j.cnki.jos.004548]

[12] Xie JY, Gao HC. Statistical correlation and *K*-means based distinguishable gene subset selection algorithms. Ruan Jian Xue Bao/Journal of Software, 2014,25(9):2050−2075 (in Chinese with English abstract). http://www.jos.org.cn/1000-9825/4644.htm [doi: 10.13328/j.cnki.jos.004644]

[13] Wagstaff K, Cardie C, Rogers S, Schroedl S. Constrained *k*-means clustering with background knowledge. In: Proc. of the ICML. 2001. 577−584.

[14] Kanungo T, Mount DM, Netanyahu NS, Piatko CD, Silverman R, Wu AY. An efficient k-means clustering algorithm: Analysis and implementation. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2002,24(7):881−892. [doi: 10.1109/TPAMI.2002.1017616]

[15] Huang JZ, Ng MK, Rong H, Li Z. Automated variable weighting in *k*-means type clustering. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2005,27(5):657−668. [doi: 10.1109/TPAMI.2005.95]

[16] Jia B, Zhu XY, Luo YP, Hu DC. Accurate Baum-Welch algorithm free from overflow. Ruan Jian Xue Bao/Journal of Software, 2000,11(5):707−710 (in Chinese with English abstract). http://www.jos.org.cn/ch/reader/view_abstract.aspx?flag=1&file_no=20000519&journal_id=jos

[17] Miklós I, Meyer IM. A linear memory algorithm for Baum-Welch training. BMC Bioinformatics, 2005,6(1):231−238. [doi: 10.1186/1471-2105-6-231]

[18] Stratonovich RL. Conditional Markov processes. Theory of Probability and its Applications, 1960,5(2):156−178. [doi: 10.1137/1105015]

[19] Brown RG. Smoothing forecasting and prediction of discrete time series. In: Proc. of the Englewood Cliffs. Prentice-Hall, 1963.

[20] Billah B, King ML, Snyder RD, Koehler AB. Exponential smoothing model selection for forecasting. Int'l Journal of Forecasting, 2006, 22(2):239−247. [doi: 10.1016/j.ijforecast.2005.08.002]

[21] Wang GQ, Wang S, Liu HY, Xue YD, Zhou P. Self-Adaptive and dynamic cubic ES method for wind speed forecasting. Power System Protection and Control, 2014,15:117−122 (in Chinese with English abstract).

[22] Kong DG, Tan XB, Xi HS, Shuai JM, Gong T. Hidden Markov model for multi-thread programs time sequence analysis. Ruan Jian Xue Bao/Journal of Software, 2010,21(3):461−472 (in Chinese with English abstract). http://www.jos.org.cn/1000-9825/3521.htm [doi: 10.3724/SP.J.1001.2010.03521]

[23] Van BL, Garcia-Salicetti S, Dorizzi B. On using the Viterbi path along with HMM likelihood information for online signature verification. IEEE Trans. on Systems, Man, and Cybernetics, Part B: Cybernetics, 2007,37(5):237−1247. [doi: 10.1109/TSMCB.2007.895323]

[24] Dai XJ, Zhang N. Performance testing and optimization of data analysis platform based on LoadRunner. Computer Technology and Development, 2013,23(7):202−206, 210 (in Chinese with English abstract). [doi: 10.3969/j.issn.1673-629X.2013.07.052]

[25] Yang P, Li J. Using LoadRunner to test Web's load automatically. Computer Technology and Development, 2007,1(80):242−244 (in Chinese with English abstract).

[26] Levinson SE, Rabiner LR, Sondhi MM. An introduction to the application of the theory of probabilistic functions of a Markov process to automatic speech recognition. The Bell System Technical Journal, 1983,62(4):1035−1074. [doi: 10.1002/j.1538-7305.1983.tb03114.x]

[27] Baum LE, Sell GR. Growth transformations for functions on manifolds. Pacific Journal of Math, 1968,27(2):211−227.

[28] Kalekar PS. Time series forecasting using holt-winters exponential smoothing. Kanwal Rekhi School of Information Technology, 2004.

[29] Feng XP, Zhang TF. Comparison of four clustering methods. Microcomputer & Its Applications, 2010,29(16):1−3 (in Chinese with English abstract).

附中文参考文献:

[1] 赵迪,赵望达,刘静.基于 B/S 架构的安全生产监督管理信息系统.中国公共安全(学术版),2007,4:97−99.

[2] 陈帝江,曹文钢.基于 B/S 模式的进销存管理系统的设计与实现.作何机床与自动化加工技术,2006,5:110−112.

[4] 宗瑜,金萍,陈恩红,李红,刘仁金.面向 WebLogic 的模糊协同聚类算法.电子与信息学报,2012,34(3):543−548.

[5] 刘伟祥,周建宁,张捷.ORACLE 数据库存储过程应用安全.计算机系统应用,2013,2:80−83.

[11] 张建华,张文博,徐继伟,魏峻,钟华,黄涛.一种基于隐马尔可夫模型的虚拟机失效恢复方法.软件学报,2014,25(11):2702−2714. http://www.jos.org.cn/1000-9825/4548.htm [doi: 10.13328/j.cnki.jos.004548]

[12] 谢娟英,高红超.基于统计相关性与 *K*-means 的区分基因子集选择算法.软件学报,2014,25(9):2050−2075. http://www.jos.org.cn/1000-9825/4644.htm [doi: 10.13328/j.cnki.jos.004644]

[16] 贾宾,朱小燕,罗予频,胡东成.消除溢出问题的精确 Baum-Welch 算法.软件学报,2000,11(5):707−710. http://www.jos.org.cn/ch/reader/view_abstract.aspx?flag=1&file_no=20000519&journal_id=jos

[21] 王国权,王森,刘华勇,薛永端,周平.基于自适应的动态三次指数平滑法的风电场风速预测.电力系统保护与控制,2014,15:117−122.

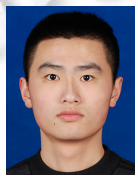
[22] 孔德光,谭小彬,奚宏生,帅建梅,宫涛.多线程程序时序分析的隐 Markov 模型.软件学报,2010,21(3):461−472. http://www.jos.org.cn/1000-9825/03521.htm [doi: 10.3724/SP.J.1001.2010.03521]

[24] 戴晓婧,张宁.基于 LoadRunner 的数据分析平台的性能测试及优化.计算机技术与发展,2013,23(7):202−206. [doi: 10.3969/j.issn.1673-629X.2013.07.052]

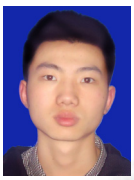
[25] 杨萍,李杰.利用 LoadRunner 实现 Web 负载测试的自动化.计算机技术与发展,2007,17(1):242−244.

[29] 冯晓蒲,张铁峰.4 种聚类方法之比较.微型机与应用,2010,29(16):1−3.

吴佳(1980－),女,四川荣昌人,博士,副教授,CCF 专业会员,主要研究领域为机器智能,数据挖掘.



陈瀚霖(1995－),男,CCF 学生会员,主要研究领域为数据挖掘,机器学习.



曾惟如(1995－),男,主要研究领域为神经网络,机器智能,数据分析.

唐雪飞(1964－),男,博士,副教授,主要研究领域为云计算,数据处理与数据可视化,代码自动化.