

DOI:10.20079/j.issn.1001-893x.220725001

基于 SMART 数据模式的 HDD 硬盘状态预测方法*

万成威,王霞,王猛

(北京航天飞行控制中心,北京 100094)

摘要:硬盘广泛应用于各类信息系统中,其工作状态预测对信息系统的正常运行管理有着重要意义。现有基于 SMART(Self Monitoring Analysis and Reporting Technology)属性的机器学习预测算法为保证其通用性,普遍选取部分典型属性作为特征,带来一定的信息丢失。在分析 SMART 数据特点的基础上,提出数据模式分类后再进行机器学习预测的 SMART 数据处理方法。实际测试结果表明,经分类处理后,采用简单的机器学习算法即可获得与强分类器接近的性能,同时,该方法可有效简化 SMART 数据机器学习时的特征选择过程,有效降低算法的资源消耗。

关键词:HDD 硬盘;状态预测;SMART 数据模式;机器学习

开放科学(资源服务)标识码(OSID):



微信扫描二维码
听独家语音释文
与作者在线交流
享本刊专属服务

中图分类号:TP302.7 文献标志码:A 文章编号:1001-893X(2024)02-0310-06

HDD State Prediction Based on SMART Data Modes

WAN Chengwei, WANG Xia, WANG Meng

(Beijing Aerospace Control Center, Beijing 100094, China)

Abstract: The disk is widely adopted in different information systems, and its state prediction is significant for the daily operation and management of those systems. Present self monitoring analysis and reporting technology (SMART)-based machine learning methods usually chose some typical SMART attributes as features for the prediction to ensure the generality. However, some information is lost by this way. A hard drive disk (HDD) state prediction method based on SMART data modes is proposed according to its characteristics, which classifies the SMART records into different subsets according to the data modes. Then, The general machine learning methods are evaluated for these subsets. The test result shows that the proposed method provides the similar performance compared with some strong predictors. At the same time, the feature selection is effectively simplified, and the resource consumption is effectively decreased.

Key words: hard drive disk; state prediction; data mode of SMART records; machine learning

0 引言

数据存储子系统是各类信息系统或设备必不可少的组成部分之一,硬盘是数据存储子系统的核心部件。然而,硬盘也是影响信息系统或设备性能和工作状态的重要因素^[1]。固态硬盘(Solid State Disk, SSD)可极大提高硬盘的数据读写性能,然而硬盘的经常性失效问题仍不可避免,业界普遍采用

独立冗余磁盘阵列(Redundant Array of Independent Disk, RAID)技术实现硬盘失效时的数据完整性保护。同时,硬盘健康状态的预测研究也日渐受到业界的重视^[2-4],期望通过硬盘自身的状态日志、设备运行状态、机房位置等信息,准确预测硬盘的健康状态^[5],最终实现硬盘剩余寿命的准确预测,为实现数据管理、设备运维等提供相应的支撑手段。

* 收稿日期:2022-07-25;修回日期:2022-10-02
通信作者:万成威 Email:wancheng_2002@163.com

机械硬盘 (Hard Drive Disk, HDD) 在大规模数据中心、大容量数据存储系统中广泛应用,其故障一般分为两种:可预测的和不可预测的。不可预测的硬盘故障偶尔发生,如芯片突然失效、机械振动或撞击等,但如磁盘磁头磨损、盘片磁介质性能下降等,均属于可预测情况,通过磁盘监控技术,检测磁盘某些重要的状态参数实现磁盘健康状态的预警。1995 年,康柏公司联合多家硬盘生产商提出该技术方案,并于 1996 年正式成为行业技术标准。然而,准确地实现硬盘健康状态预测仍然具有很强的挑战性,主要原因来自两个方面^[6]:一是软件系统优化设计后可较好地适配硬件系统部分功能失效情况,导致系统操控管理人员对包含硬盘故障在内的硬件系统部分功能失效情况的重视程度不够;二是各种不同品牌和型号规则的硬盘在数据中心内部广泛使用,但考虑到商业利益因素,大量的研究仅公开了少量的采样获取的自我检测分析与报告 (Self Monitoring Analysis and Reporting Technology, SMART) 数据,难以支撑大规模的开放性研究,而且研究重点也主要着眼于不同类型硬盘的某些特性,如结构、器件等等对异常产生原因的分析。

近年来,随着 BackBlaze、Alibaba 等云服务商逐步开放其超大数据中心硬盘的 SMART 数据,基于机器学习、深度学习等方式的硬盘健康状态预测研究^[7-11]也随之逐年增加,针对硬盘失效问题的深度分析也逐步展开。清华大学交叉信息研究院的徐葳团队^[6]联合百度公司对百度数据中心 2012—2016 年以来的 290 000 个硬件失效问题进行了研究,从失效问题发生的时间、产品线、空间位置、部件以及用户的响应等 5 个维度进行了分析。从部件维度看,HDD 硬盘故障占硬件系统故障的 82%,其中约 30% 的故障直接由 SMART 属性值反应出来,其他部分故障也可由相关的 SMART 属性与之对应,同时,研究结果表明,HDD 失效与高负载、人员操作等均具有一定的潜在关联性。Lu 等人^[5]针对 HDD 硬盘的 SMART 数据、硬盘的性能数据、硬盘的空间位置信息等 3 类数据,采用朴素贝叶斯分类、随机森林^[12]、梯度增强决策树 (Gradient Boosted Decision Tree, GBDT)^[13-14]、长短时记忆网络 (Long Short Term Memory Network, LSTM) 等深度学习方法,实现对 HDD 硬盘工作状态的预测,并详细比较了上述 3 类信息不同组合条件下的预测准确性,结果表明,3 类信息联合情况下可获得最优的预测性能,原因在

于硬盘的性能数据体现了其工作负载,硬盘的空间位置体现了其运行环境,这些因素均会影响硬盘的工作状态。

随着 SSD 硬盘在数据中心应用的逐步深入,针对其工作状态的预测研究也逐步展开^[15-16],这些研究方法对 HDD 硬盘工作状态预测具有较好的借鉴意义。

本文将 HDD 作为研究对象,基于 BackBlaze 公布的 HDD 硬盘 SMART 日志数据,采用机器学习方法,完成硬盘健康状态的分类预测。

1 SMART 数据

SMART 技术为硬盘行业标准,依托该技术,硬盘会自动检测其磁头、盘片、马达、电路等运行状态,并通过不同的属性值记录下来,实现其工作状态的记录,当属性值超出安全值范围时,就会自动向用户发出告警。一般而言,硬盘会周期性产生 SMART 记录数据。考虑到硬盘大部分时间均处于正常工作状态,已有研究表明,对硬盘工作状态预测而言,每天一条 SMART 数据记录即可满足应用要求^[5]。目前,Alibaba、BackBlaze 等大型云服务提供商均开源了其部分数据中心硬盘的 SMART 数据。典型的 SMART 属性如表 1 所示。

表 1 典型 SMART 属性值
Tab. 1 Typical SMART attributes

属性 ID	属性名称
1	Read Error Rate
4	Start/Stop Count
5	Reallocated Sectors Count
7	Seek Error Rate
9	Power-On Hours
193	Load/Unload Cycle Count
194	Temperature
199	UltraDMA CRC Error Count

SMART 属性数据以二进制代码的形式记录于硬盘的系统保留区内,除标准属性外,厂商可根据自己产品的特性提供不同的 SMART 属性数据。目前公开的 SMART 数据集中,除 SMART 属性的原始数据外,同时包含每个属性的归一化值,联合起来作为硬盘的 SMART 属性。单条 SMART 记录包含日期、序列号 (或 ID)、硬盘型号、工作状态及属性值等。一般情况下,硬盘的大部分属性值为空,不同厂商、

不同型号规格的硬盘,为空的属性值也有所不同。为此,根据硬盘 SMART 属性是否非空,定义 SMART 数据模式如下:给定的 SMART 数据记录中,若某一属性非空,则该属性的数据模式为“1”;否则,该属性的数据模式为“0”,由所有属性数据模式组成的二进制字符串即为该 SMART 记录的数据模式。

从已开放的 SAMRT 数据看,尽管不同厂商、不同型号规格硬盘的 SMART 数据各不相同,但仍体现出明显的特点,具体如下:

1) SMART 数据具稀疏特性。尽管最新标准中定义了多达 256 项 SMART 属性,然而,一般硬盘在日志中仅记录了其中几十项属性值,且其中大部分属性为空,形成典型的稀疏特性。

2) SMART 数据模式类型有限。SMART 数据模式指硬盘所有非空 SAMRT 属性的组成规律。大型数据中心中存在不同品牌、不同型号规格的硬盘,然而,这些硬盘的 SMART 数据模式类型较为有限,如本文后文所述,并非如直观印象中有大量不同的数据模式类型。

3) SMART 数据具有明显的不均衡性。SMART 数据的不均衡性主要表现在硬盘状态正常与失效的 SMART 数据数量相差极大。硬盘技术和制造工艺的进步,使得硬盘失效概率明显降低,大部分时间内硬盘处于正常工作状态,因此 SMART 数据中失效数据比例也较低。

考虑 SMART 数据模式类型有限的特点,且同一数据模式下硬盘 SMART 属性完全一致,更趋于表现出一致的运行状态,基于此,本文提出基于 SMART 数据模式分类的 HDD 硬盘工作状态预测方法。

2 硬盘工作状态预测

现有研究主要以 SMART 属性值及其归一化属性值作为特征,采用各种机器学习、深度学习算法,实现硬盘工作状态的预测。尽管针对大型数据中心硬件系统故障问题的长期统计分析表明,硬盘负载、机房环境等对其工作状态具有一定的影响^[5],从表 1 也可以看到,上述信息在 SMART 属性数据中也有一定的体现。另外,现有基于 SMART 数据的硬盘工作状态预测研究中,一般考虑算法的通用性,期望算法可广泛适用于所有类型的硬盘,通常仅选择十几项典型的属性值作为硬盘工作状态预测算法

的特征,然而,除常用的典型属性外,不同型号规格的硬盘可定义不同 SMART 属性,这将导致部分信息的丢失^[16]。为此,本文在深入分析 SMART 数据特点的基础上,提出相应的 HDD 硬盘工作状态预测算法。

2.1 算法思想

为了避免现有算法选择少量典型 SMART 属性作为特征引起的信息丢失问题,结合 SMART 数据的特点,本文提出基于 SMART 数据模式的机器学习 (SMART Data Mode Based Machine Learning, SDM-ML) 算法,以实现 HDD 硬盘工作状态的准确预测,其基本思想如下:

SMART 技术作为硬盘行业标准,反映了硬盘的各种工作状态,其属性的设计均具有一定的针对性。直观理解,选取的属性越多,包含的硬盘信息也相应增多,利于机器学习算法更充分识别硬盘的工作状态,对其做出更加准确的预测。经统计分析,硬盘 SMART 数据模式较为有限,为此,SDM-ML 算法先根据数据模式类型,将 SMART 数据分为不同的子集,不同子集的数据分别采用有监督机器学习算法进行训练和验证测试,综合多种机器学习算法测试结果,选择性能最优算法的输出结果作为最终预测结果。

2.2 算法流程

如图 1 所示,算法分为数据预处理、数据模式分类、按数据模式分类机器学习预测。实际过程中,根据数据模式类型,将预处理后的 SMART 数据输入训练好的机器学习组件中,即可完成 HDD 硬盘状态的预测。

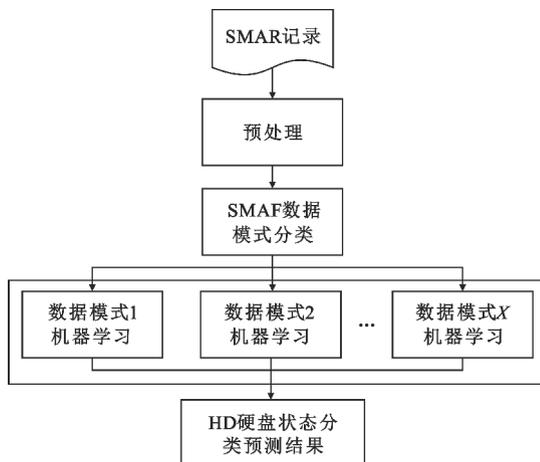


图 1 基于数据模式分类的 HDD 硬盘工作状态预测
Fig. 1 HDD state prediction based on the SMART data modes

数据预处理主要完成原始 SMART 数据的规范化处理,包含:以序列号为关键字,将单个硬盘的 SMART 数据按照日期先后顺序归档整理;仅以 SMART 属性的实际值作为研究对象,不考虑数据提供方给出的归一化属性值;将属性值为空的 SMART 属性统一填充 0,若硬盘当日 SMART 属性均为空,即表示当日未成功获取 SMART 日志记录,则删除当天 SMART 记录。

数据模式分类主要根据硬盘 SMART 记录中的原始属性数据,确定该硬盘的 SMART 数据模式类型。所有数据模式类型中均包含该 SMART 数据的所有属性,不对其进行筛选、过滤、聚合等处理。

并行训练主要将分类后的 SMART 数据分别按照数据模式输入监督机器学习算法进行算法的训练。对于每块硬盘,均选择 N 天的 SMART 记录依次拼接组合作为机器学习的输入数据。其中,对于发生失效问题的硬盘,选择从失效当天起的前 N 天 SMART 数据作为故障类数据(类型为“1”);对于未发生失效问题的硬盘,随机选择连续 N 天的 SMART 数据作为正常数据(类型为“0”)。所有数据根据对应属性值,按式(1)进行归一化处理:

$$\bar{x}_i = \frac{x_i - x_{i\min}}{x_{i\max} - x_{i\min}} \quad (1)$$

式中: $x_{i\max}$ 和 $x_{i\min}$ 分别为第 i 个属性的最大值和最小值; x_i 和 \bar{x}_i 分别为第 i 个属性的原始值和归一化值。将归一化处理后的 N 条 SMART 属性记录依次前后连接,作为机器学习算法的一条输入数据。考虑 SMART 数据的不均衡性,采用 EasyEnsemble 算法^[17],按照故障数据规模,将正常数据分为 M 份,分别与故障数据合并,共同用于机器学习算法的训练与验证测试,取 M 个分类器测试结果的平均值作为机器学习算法的评估结果。

2.3 算法性能评估

2.3.1 数据集

本文以 BackBlaze 开放的 2020 年 3 季度云存储系统中硬盘 SMART 日志作为数据集,用于评估 SDM-ML 算法性能。该数据集包含 2020 年 7 月 1 日—9 月 30 日期间共 151 166 块 HDD 硬盘的 SMART 日志,其中发生失效问题的硬盘 324 块。按照 SMART 数据模式,上述记录中共包含 18 种不同的数据模式,各类 SMART 数据模式统计信息如表 2 所示。

表 2 SMART 数据模式统计
Tab.2 Statistics of SMART data modes

数据模式	正常硬盘数量	失效硬盘数量
1	15 244	28
2	886	29
3	13 342	13
4	2 600	2
5	15 793	7
6	461	6
7	19 062	60
8	274	1
9	887	2
10	96	0
11	98	1
12	10	0
13	60	1
14	24 254	66
15	1198	3
16	33 524	61
17	17 321	37
18	6 056	7
合计	151 166	324

表 2 中主要统计工作时间不少于 10($N=10$) 天的 HDD 硬盘,考虑到发生失效的硬盘数量远小于正常硬盘数量,为进一步增加失效硬盘 SMART 样本数量,对于 SMART 记录多于 10 天的失效硬盘,依次从失效前第 10, 15, 20... 天开始,并取连续 10 天的 SMART 数据,均作为失效硬盘 SMART 数据。考虑到 HDD 硬盘状态预测的应用场景,扩充的失效硬盘 SMART 数据最长不超过失效前 30 天。尽管如此,仍有部分 SMART 数据模式的失效样本数量过少,如第 4, 8, 9, 10, 11, 12, 13, 15 类数据模式,难以形成有效的训练、测试数据集,本文暂不将其作为 HDD 硬盘工作状态预测研究的数据集。

2.3.2 测试结果

本文采用 K 近邻 (K-Nearest + Neighbor, KNN)^[12]、线性回归 (logreg)^[10]、支持向量机 (Support Vector Machine, SVM)^[8] 等 3 种机器学习算法,分别按照不同的 SMART 数据模式类型,从预测准确率、F1 指数、马修斯相关系数 (Matthews Correlation Coefficient, MCC) 等 3 个方面对 HDD 硬盘工作状态预测性能进行了评估。 M 根据预处理后正常样本与失效样本的比例关系确定。测试结果如图 2 所示。

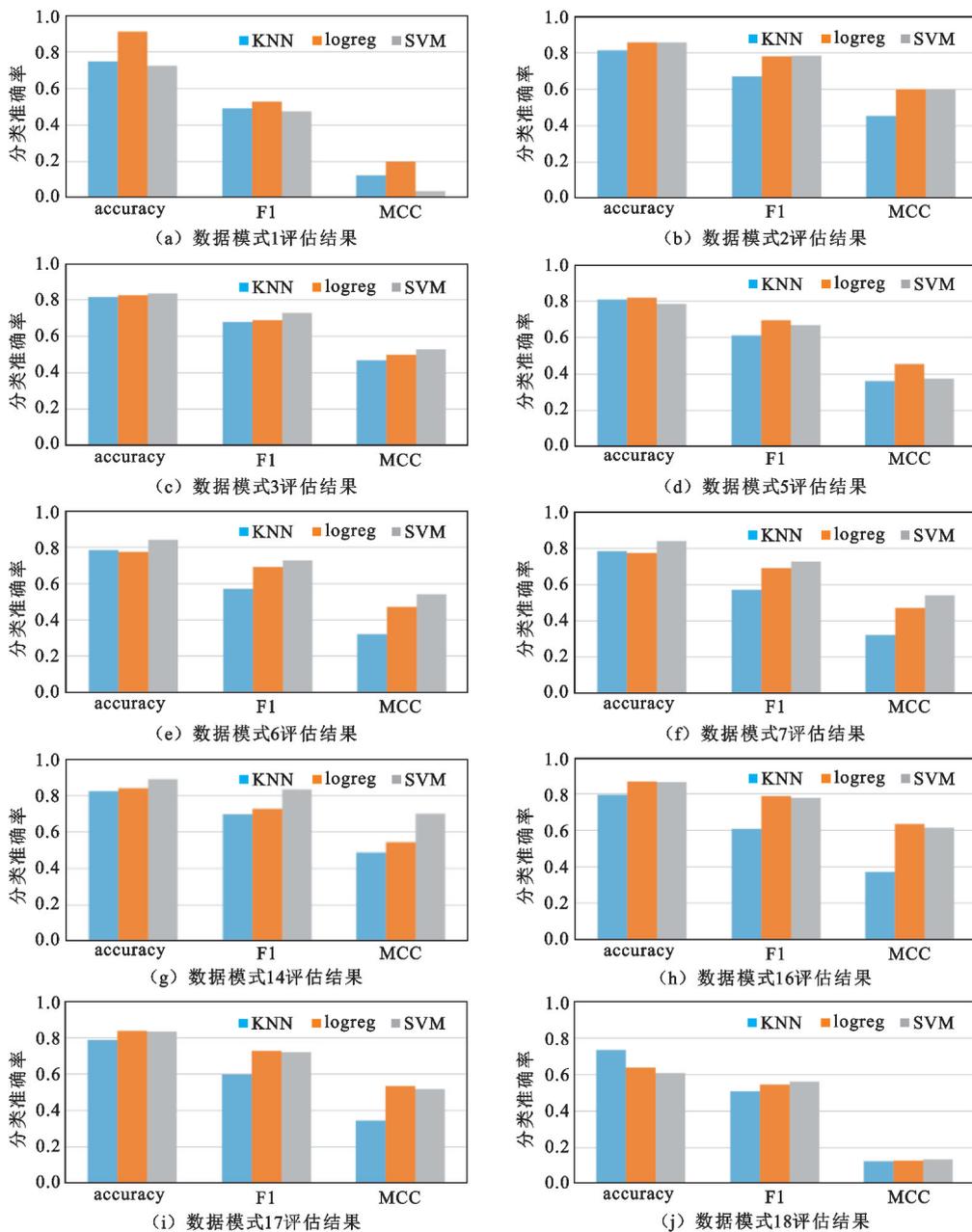


图 2 SDM-ML 算法评估结果

Fig. 2 Evaluation results of SDM-ML

可以看到,SDM-ML 算法普遍可实现 80% 以上的分类准确率,达到与现有强分类器相同的预测精度。从 F1 指数和 MCC 系数看,SDM-ML 算法表现出相同或类似的测试性能,即按照 SMART 数据模式分类后,再采用常见的机器学习算法,即可获得强分类器相近的预测结果。从机器学习算法看,线性回归和 SVM 算法针对不同类型的 SMART 数据模式均表现出较为一致的预测性能,且优于 KNN。从 SMART 数据模式看,第 1 和 17 类数据模式的 MCC 系数测试结果相对较差,这也反映出不同型号规格

硬盘产品在 SMART 属性上的差异及其对机器学习算法特征选择的影响,与文献[16]中针对 SSD 硬盘 SMART 属性的研究结果一致。同时,这也说明特征选择对机器学习算法性能的影响。本文提出在 SMART 数据模式分类后再进行机器学习预测 HDD 硬盘工作状态,主要目的也是为了尽可能简化众多 SMART 属性条件下的特征选择过程,测试结果也表明了本文方法的有效性。同时,本文方法在通用机器学习算法下即可获得较好的预测效果,可有效降低因使用强机器学习算法带来的计算、存储等资源

消耗,具有较好的实际应用效果。

3 结 论

本文研究了基于 SMART 数据的 HDD 硬盘工作状态预测算法,针对现有算法普遍以少量通用 SMART 属性作为机器学习特征,存在信息丢失现象的问题,提出先按照数据模式对 SMART 数据进行分类,然后采用机器学习算法进行硬盘工作状态预测,在保证 SMART 信息完整性的同时,简化了机器学习算法的特征选择过程,有效降低了算法的资源消耗。实际测试结果表明,在这种数据处理方式下,采用常用的机器学习算法即可获得与强分类器相近的预测结果,充分证明了算法的有效性。

参考文献:

- [1] JIANG W, HU C, ZHOU Y, et al. Are disks the dominant contributor for storage failures; a comprehensive study of storage subsystem failure characteristics [J]. ACM Transactions on Storage, 2008, 4(3): 7-31.
- [2] PREETHI A, QIAO M, DIYYESH J. Large scale predictive analytics for hard disk remaining useful life estimation [C]// Proceedings of 2018 IEEE International Congress on Big Data. San Francisco: IEEE, 2018: 251-254.
- [3] MIRELA M, IOANA G, JASMINA B, et al. Predicting disk replacement towards reliable data centers [C]// Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. San Francisco: ACM, 2016: 39-48.
- [4] IAGO C, MANOEL R, LUCAS G, et al. Hard disk drive failure prediction method based on a Bayesian network [C]// Proceedings of 2018 International Joint Conference on Neural Networks. Rio de Janeiro: IEEE, 2018: 1-7.
- [5] LU S, LUO B, TIRTHAK P, et al. Making disk failure predictions SMARTer! [C]// Proceedings of the 18th USENIX Conference on File and Storage Technologies. Santa Clara: USENIX, 2020: 151-167.
- [6] WANG G, XU W, ZHANG L. What can we learn from four years of data center hardware failures? [C]// Proceedings of the 47th Annual IEEE/IFIP International Conference on Dependable Systems and Networks. Denver: IEEE, 2017: 25-36.
- [7] HAMERLY G, ELKAN C. Bayesian approaches to failure prediction for disk drives [C]// Proceedings of the 8th International Conference on Machine Learning. Williamstown: ACM, 2001: 202-209.
- [8] ZHU B, WANG G, LIU X, et al. Proactive drive failure prediction for large scale storage systems [C]// Proceedings of IEEE the 29th Symposium on Mass Storage Systems and Technologies. Long Beach: IEEE, 2013: 1-5.
- [9] QIAN J, SKELTON S, MOORE J, et al. Priority based proactive prediction for soon-to-fail disks [C]// Proceedings of 2015 IEEE International Conference on Networking, Architecture and Storage. Boston: IEEE, 2015: 1-8.
- [11] LI J, JI X, JIA Y, et al. Hard drive failure prediction using classification and regression trees [C]// Proceedings of the 44th Annual IEEE/IFIP International Conference on Dependable Systems and Networks. Atlanta: IEEE, 2014: 383-394.
- [12] XU C, WANG G, LIU X, et al. Health status assessment and failure prediction for hard drives with recurrent neural networks [J]. IEEE Transactions on Computers, 2015, 13(9): 1-8.
- [13] SHEN J, WAN J, LIM S, et al. Random-forest-based failure prediction for hard disk drives [J]. International Journal of Distributed Sensor Networks, 2018, 14(11): 1-15.
- [14] JEROME H F. Stochastic gradient boosting [J]. Computational Statistics and Data Analysis, 2002, 38(4): 367-378.
- [15] JERRY Y, JYH-HERNG C, JIANG C, et al. Stochastic gradient boosted distributed decision trees [C]// Proceedings of the 18th ACM Conference on Information and Knowledge Management. Hong Kong: ACM, 2009: 2061-2064.
- [16] XU E, ZHENG M, QIN F, et al. Lessons and actions: what we learned from 10K SSD-related storage system failures [C]// Proceedings of 2019 USENIX Annual Technical Conference. Renton: USENIX, 2019: 961-976.
- [17] XU F, HAN S, LEE P, et al. General feature selection for failure prediction in large-scale SSD deployment [C]// Proceedings of the 51st Annual IEEE/IFIP international conference on Dependable Systems and Networks. Taipei, China: IEEE, 2021: 1023-1030.

作者简介:

万成威 男, 1982 年生于湖北安陆, 2012 年获博士学位, 现为工程师, 主要从事航天测控计算机系统设计工作。

王霞 女, 1970 年生于河北安国, 1996 年获硕士学位, 现为正高级工程师, 主要从事航天测控计算机系统论证、计算机应用系统设计等工作。

王猛 男, 1982 年生于山东东明, 2011 年获硕士学位, 现为高级工程师, 主要从事航天测控计算机系统设计工作。