International Journal of Psychosocial Rehabilitation, Vol. 23, Issue 03, 2019
ISSN: 1475-7192

Reallocated Sectors Count Parameter for Analysing Hdd Reliability

Iskandar Nailovich Nasyrov[1], Ildar Iskandarovich Nasyrov[1], Rustam Iskandarovich Nasyrov[1] and Bulat Askarovich Khairullin[1]

Abstract--- *The dependence of the SMART parameter 5 Reallocated sectors count value change on the operating time characterising the number of reallocated sectors is considered. This parameter is critical in the sense that if the attribute value increases, this may indicate deterioration in the state of the disk surface. The scientific task of the study is to establish relationships in the failed hard drives between the specified parameter and the values of other reliability parameters for information stores of various manufacturers. In the course of the study, the drives of the HGST, Hitachi, Samsung, ST, Toshiba, WDC trademarks operated in the Backblaze largest commercial data centre were analysed. The analysis revealed a relationship between the specified parameter and the parameters 1 Read error rate (frequency of errors (when reading data from the disk), the origin of which is due to the hardware of the disk), 196 Reallocation event count (number of reallocation operations), 197 Current pending sector count (number of sectors that are candidates for reallocation). It is shown that the nature of the change in the values of the considered parameters depends on the manufacturer of information storage devices. It is proposed to perform an individual assessment of the reliability of hard drives using the parameters identified as a result of the study.*

Keywords---*Reallocated Sector, Hard Drive, Reliability, Information, Security, Drive.*

I. INTRODUCTION

To ensure the security of information, it is necessary to timely and completely copy data from an unreliable drive to a new and reliable drive. For this purpose, they usually use SMART technology (self-monitoring, analysis and reporting technology [1]) for internal assessment of the computer's hard drive, as well as a way to predict its possible failure.The paper considers the dependence of the change on the operating time of the parameter 5 Reallocated sectors count, which characterizes the number of reallocated sectors. This parameter is critical in the sense that if the value of this attribute increases, this may indicate deterioration in the state of the disk surface. The scientific task of the study is to establish for failed hard drives a connection of the specified parameter with the values of other reliability parameters of information storage devices from various manufacturers.

In the course of the study, the parameters of failed drives of the HGST, Hitachi, Samsung, ST, Toshiba, WDC brands operated in the Backblaze largest commercial data centre were analysed. The analysis revealed a connection between parameter 5 Reallocated sectors count and parameters 1 Read error rate (frequency of errors when reading data from the disk, the origin of which is due to the hardware of the disk), 196 Reallocation event count (the number of reallocation operations), 197 Current pending sector count (number of sectors that are candidates for reallocation).

[1]*Kazan Federal University, E-mail: ecoseti@yandex.ru*

Received: 10 Mar 2019 | Revised: 27 Mar 2019 | Accepted: 25 Apr 2019

It is shown that the nature of the change in the values of the considered parameters depending on the operating time of the information storage devices depends on the manufacturer. It is proposed to perform an individual assessment of the reliability of hard drives using the parameters identified as a result of the study.

II. METHODS

To analyse the dependence of the parameter values on the operating time of information storage devices on hard magnetic disks that have failed, we studied the SMART data provided on the Backblaze website [2]. 45 parameters of SMART 92530 type were examined for 93 drives and 6 brands: HGST (Hitachi Global Storage Technologies), Hitachi (later HGST), Samsung, ST (Seagate), Toshiba, WDC (Western Digital) for the period from April 10, 2013, until December 31, 2016. It was found that at the end of the study period 79.58% of drives continued to work normally, 14.74% were prematurely decommissioned, and 5.68% failed.

In total, information about the semantic value of more than 80 SMART parameters is available, but most of them are not used by manufacturers. Therefore, Backblaze specialists recorded only 40 of them in 2013-2014, and since 2015, 45 with numbers 1-5, 7-13, 15, 22, 183, 184, 187-201, 220, 222-226, 240 -242, 250-252, 254, 255 (in 2015, 22, 220, 222, 224, 226 were added).

: the group which values accumulate (of the “count” type), the values of which reflect the rate of change (of the “rate” type or similar in meaning), the group which values are associated with other parameters (also of the “count” type or similar). According to another classification, three groups of parameters can also be identified: those associated with the state of the memory space, or the surface of the disks; those associated with the positioning of the write / read heads; and those associated with the mechanical part of the hard drive design.

Data storage defects can also be divided into two large groups: physical and logical. Physical defects include surface defects, servo errors, and bad hardware sectors.

Surface defects occur due to mechanical damage to the magnetic coating inside the sector space, for example, due to scratches caused by dust, ageing of disk plates or careless handling of a hard disk.Such a sector should be marked as unusable and excluded from circulation.

All modern drives use a system called “moving coil” to move the heads, which, unlike a stepper motor, does not have any discreteness of movement.For an accurate hit of the heads on the tracks, a feedback system is used;it is guided by special magnetic servos applied to the disk.Servo tags are available on each side of each drive.They are evenly spaced along all tracks forming a servo format. It does not belong to the lower level format, but it is available for all modern hard drives and plays a crucial role. According to servo marks, the engine speed is stabilised and the head is held on a given track regardless of external influences and thermal deformation of the elements.

However, during the operation of drives, some servo tags may be destroyed.If there are too many faulty servo tags, failures will occur at these points when accessing the information track: the head, instead of taking the position it needs and reading the data, will begin to move from side to side.It will look like a bad sector or even a group of them.The elimination of such defects is possible only with special programs, by disabling defective tracks, and sometimes the entire disk surface. For these purposes, some drives have a defect list that stores information about

bad servo marks. The servo-defect list is not used by the translator, but by the entire firmware of the hard drive.Sectors with defective servo tags are blocked even by physical parameters;this helps to avoid knocks and breakdowns when accessing them. A hard drive cannot restore the servo format on its own; it is done only at the factory.

3) Hardware bad sectors appear due to a malfunction in the mechanics or electronics of the drive.Such problems include breakage of the heads, displacement of the disks or a bent shaft as a result of an impact, dusting of the hermetic zone, as well as various interruptions in the operation of electronics. Errors of this type are usually catastrophic and cannot be corrected programmatically.

Logical defects arise not because of surface damage, but because of violations of the logic of the sector.They can be divided into correctable and uncorrectable.Logical defects have the same external manifestations as physical ones, and they can be distinguished only indirectly, according to the results of various tests.

The essence of the research method is to compare changes in the values of SMART parameters for failed storage media and to identify coincidences in time.

III. RESULTS AND DISCUSSION

The parameter 5 Reallocated sectors count is the best option for displaying the surface state of hard drives. It is always used for all drives and is equally interpreted by all manufacturers. Also, its changes coincide with changes in parameters 1, 196, 197. The example of suchcoincidence for the failed hard drive brand HGST is shown in Figure 1.A similar coincidence is observed for failed drives of the Hitachi brands (Figure 2), Toshiba (Figure 3), WDC (Figure 4).For the only Samsung manufacturer that lost its working capacity, parameters 5 and 196 are equal to zero. Nevertheless, there is still a coincidence of the change in parameters 1 and 197 (Figure 5).Parameters 13 and 183 are additionally shown here in order to indicate another reason for the failure of the drive. There is a coincidence for failed drives of the ST brand, which is similar to the previous ones, but without parameter 196, which is absent in all Seagate drives, and with the chaotic nature of parameter 1 change, (Figure 6).

As can be seen from Figures 1-6, usually parameter 1 changes earlier than all the others.Here,it characterizes the appearance of the very first write / read errors. Then, parameter 197 changes;it shows the number of sectors in which these write / read difficulties are observed. Further, parameters 196 and 5 are changed with all or only successful reallocation attempts.

Taking these circumstances into account when analysing relative [3] or absolute [4] values allows one to predict the failure of a hard disk.

Received: 10 Mar 2019 | Revised: 27 Mar 2019 | Accepted: 25 Apr 2019

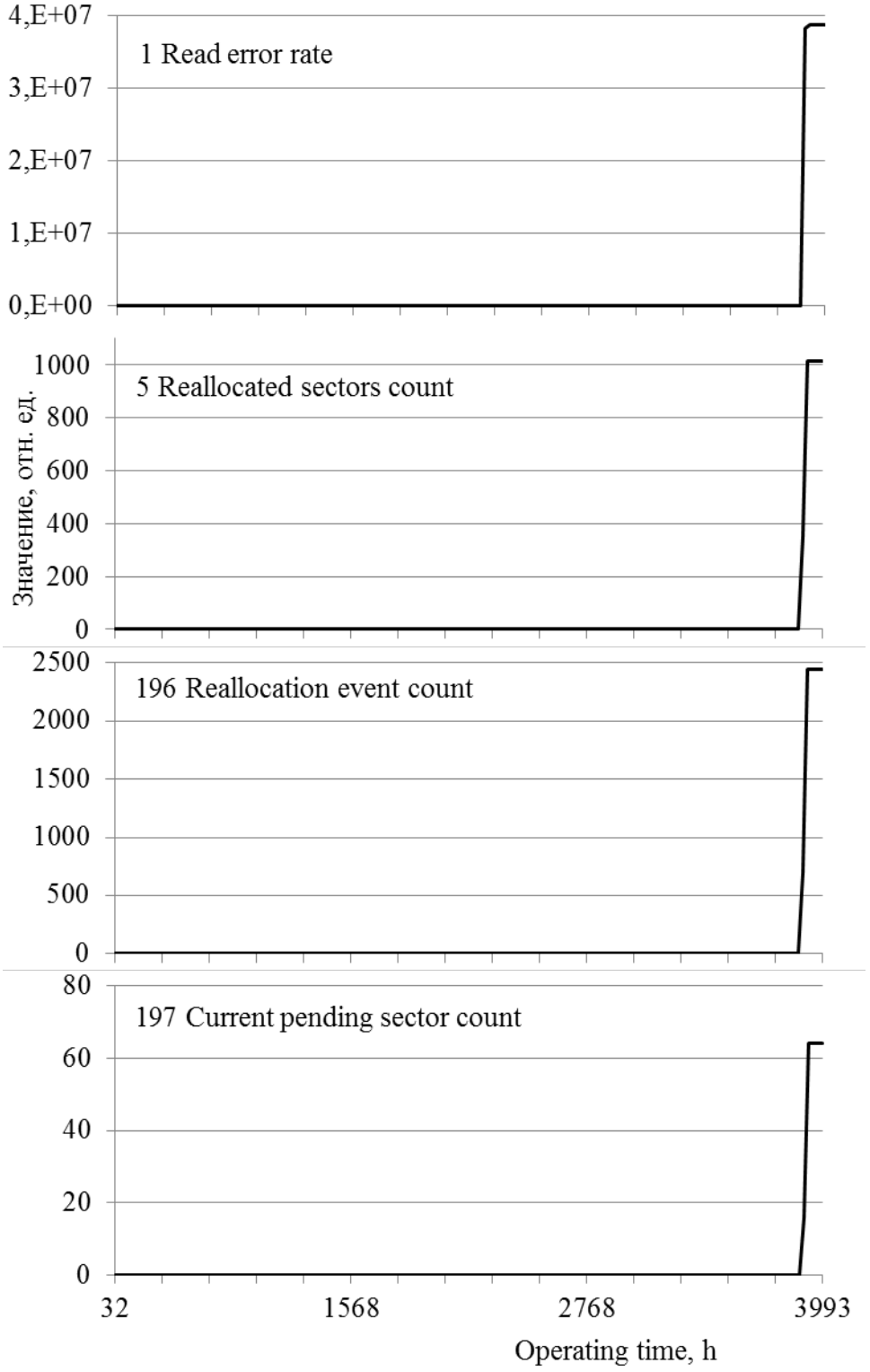

Figure 1: Parameter 1, 5, 196, 197 values depending on the operating time for a failed hard drive of model HGST HMS5C4040ALE640 with the number PL1331LAGRTU8H and with a capacity of 4 TB

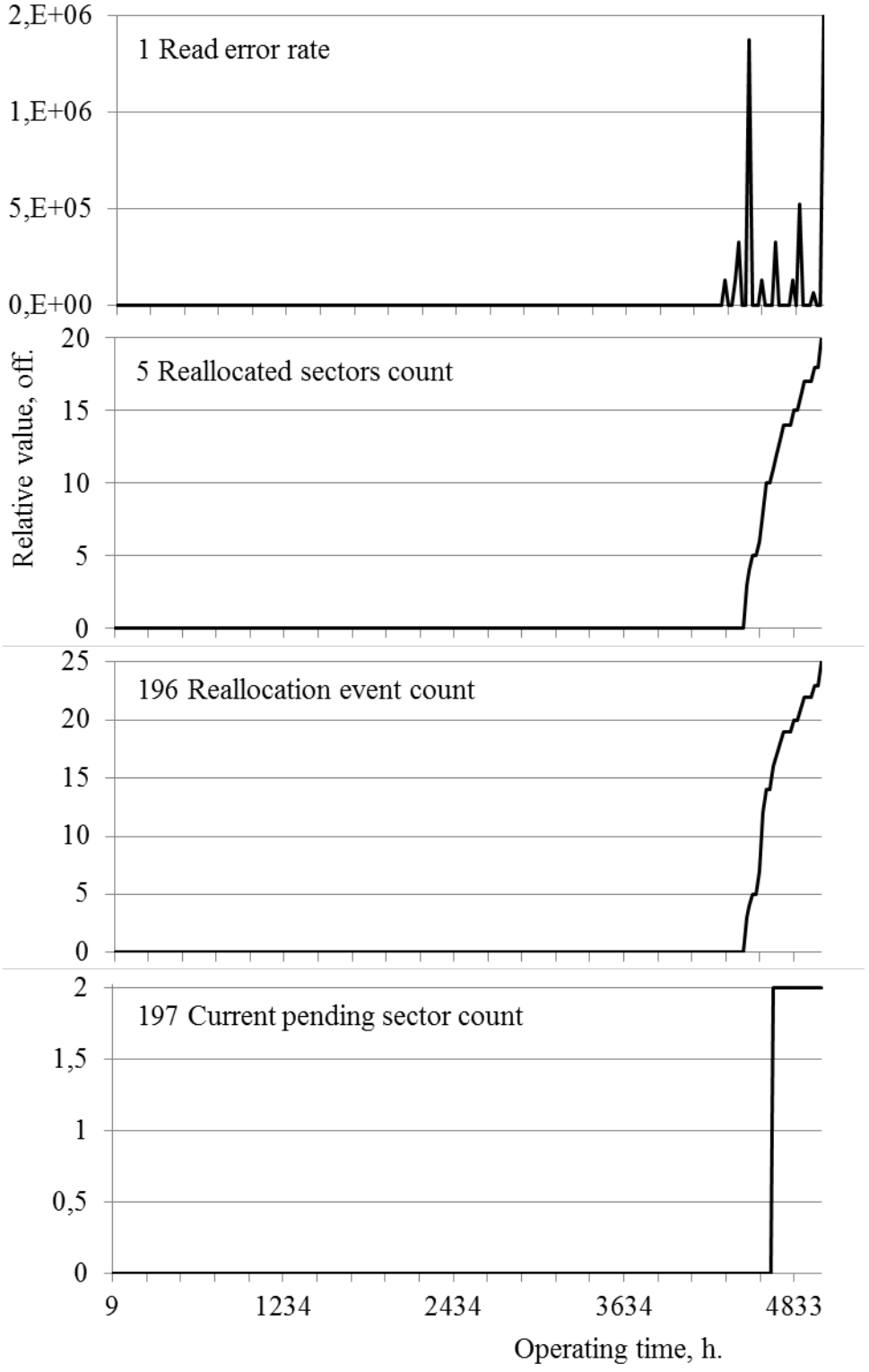

Figure 2: Parameter 1, 5, 196, 197 values depending on the operating time for the failed Hitachi HDS5C3030ALA630 model hard drive with the number MJ1311YNG6EH8A and with a capacity of 3 TB

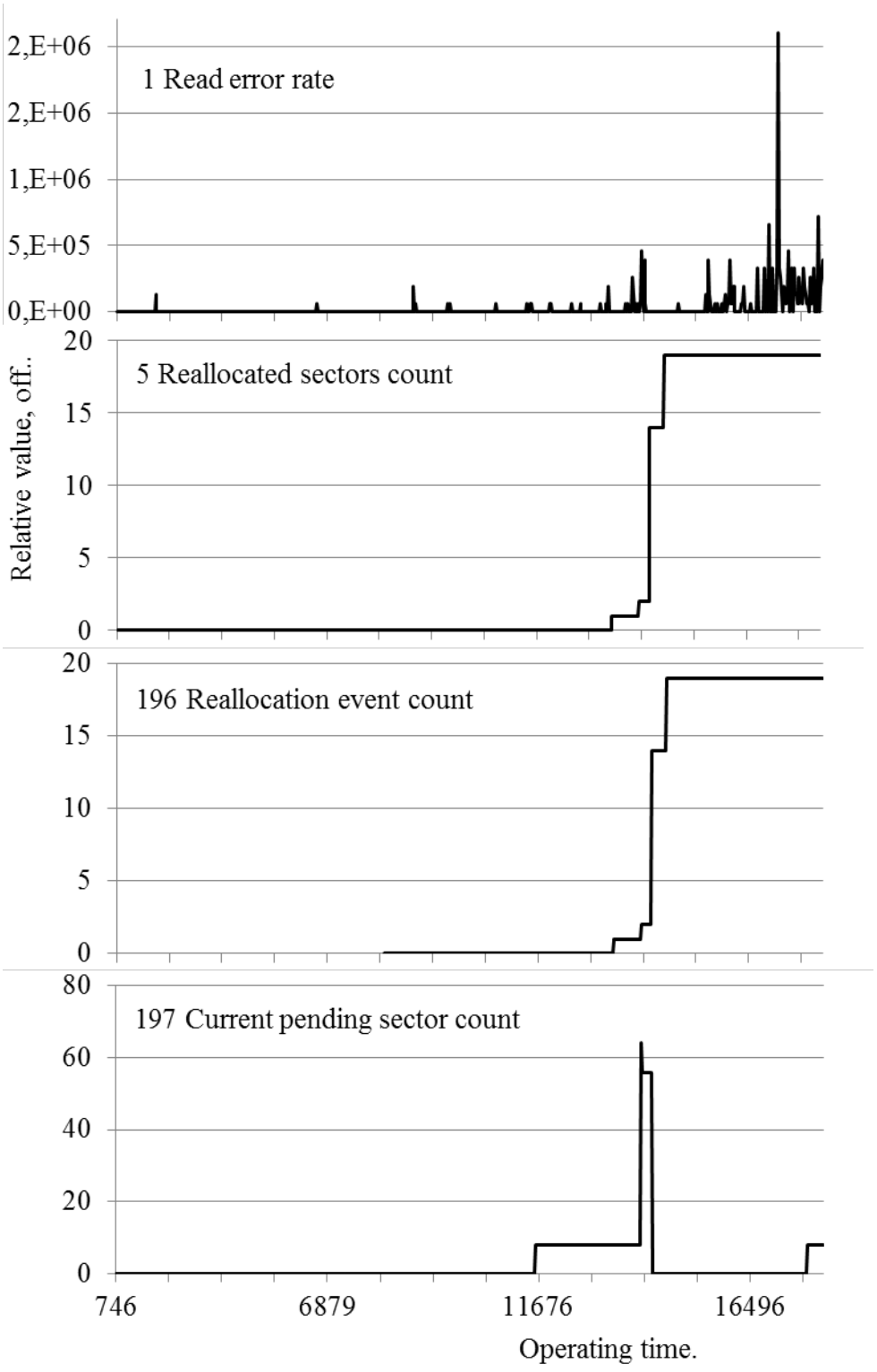

Figure 3: Parameter 1, 5, 196, 197 values depending on the operating time, in hours, for the failed Toshiba DT01ACA300 model hard drive with the number Z262EBNAS and with a capacity of 3 TB

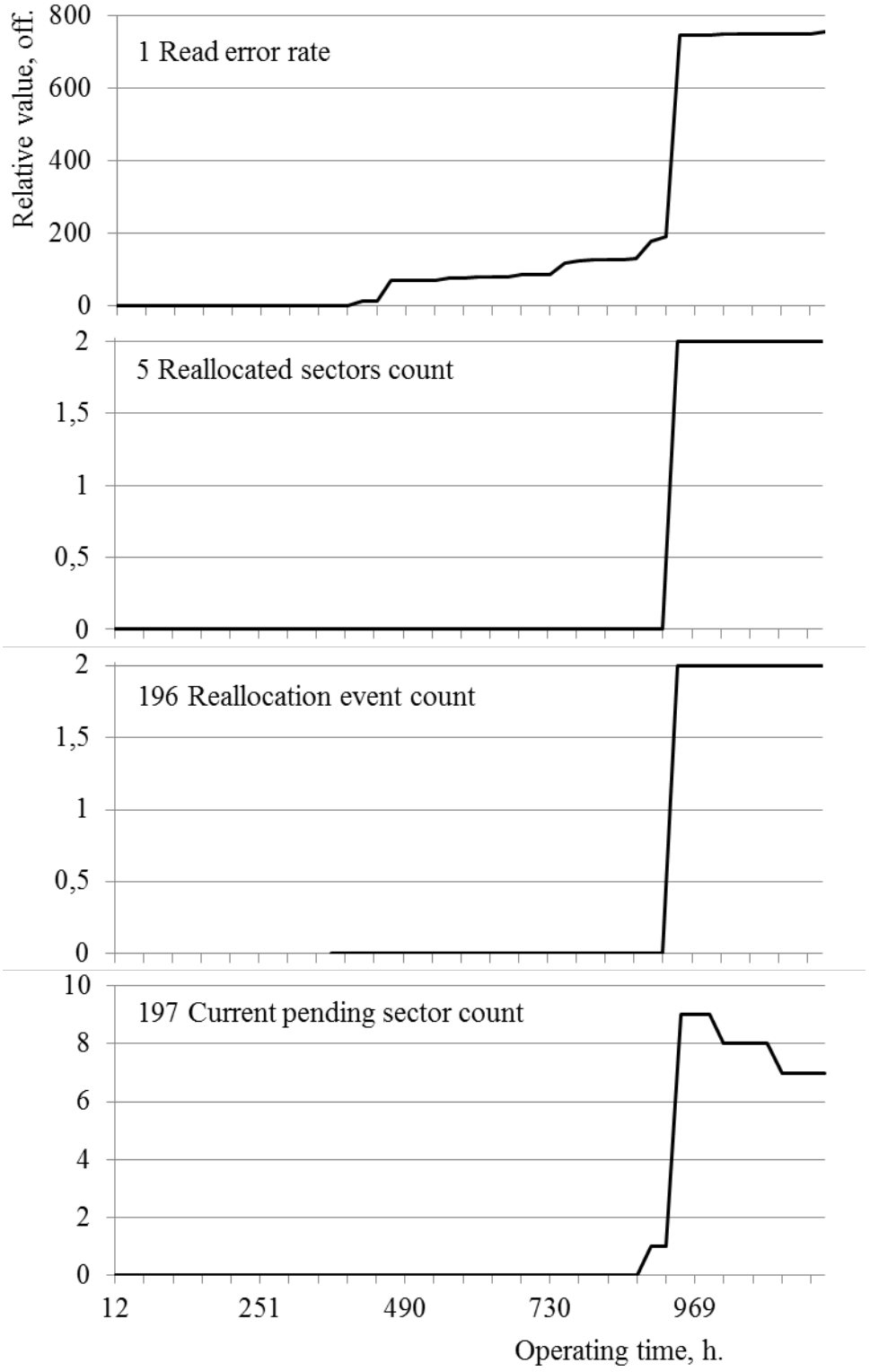

Figure 4: Parameter 1, 5, 196, 197 values depending on the operating time for a failed hard drive model WDC WD30EFRX with the number WD-WCC4N0299367 and with a capacity of 3 TB

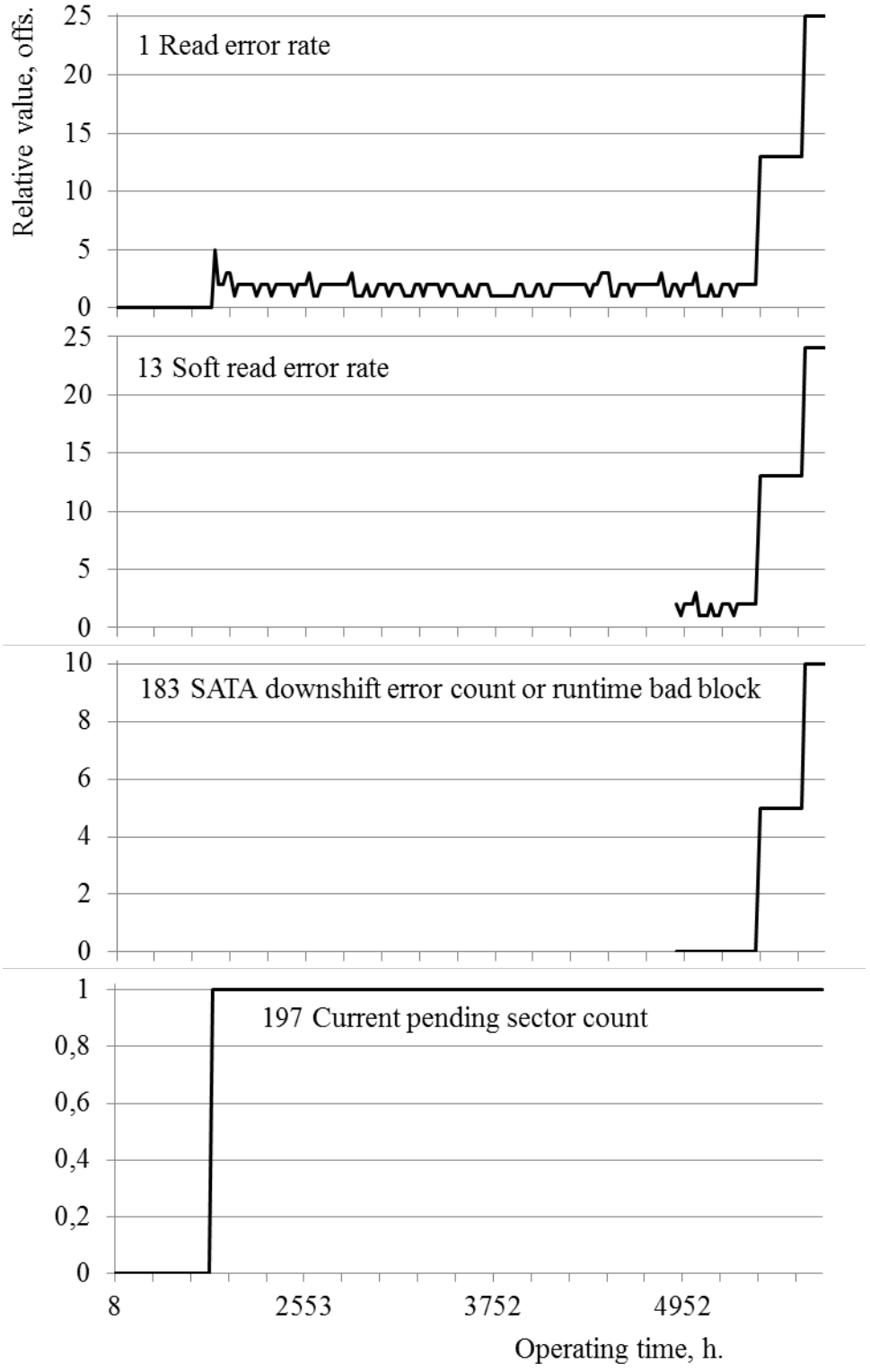

Figure 5: Parameter 1, 13, 183, 197 values depending on the operating time for the failed SAMSUNG HD154UI model hard drive with the number S2CHJR0Z900286 and with a capacity of 1.5 TB

Received: 10 Mar 2019 | Revised: 27 Mar 2019 | Accepted: 25 Apr 2019

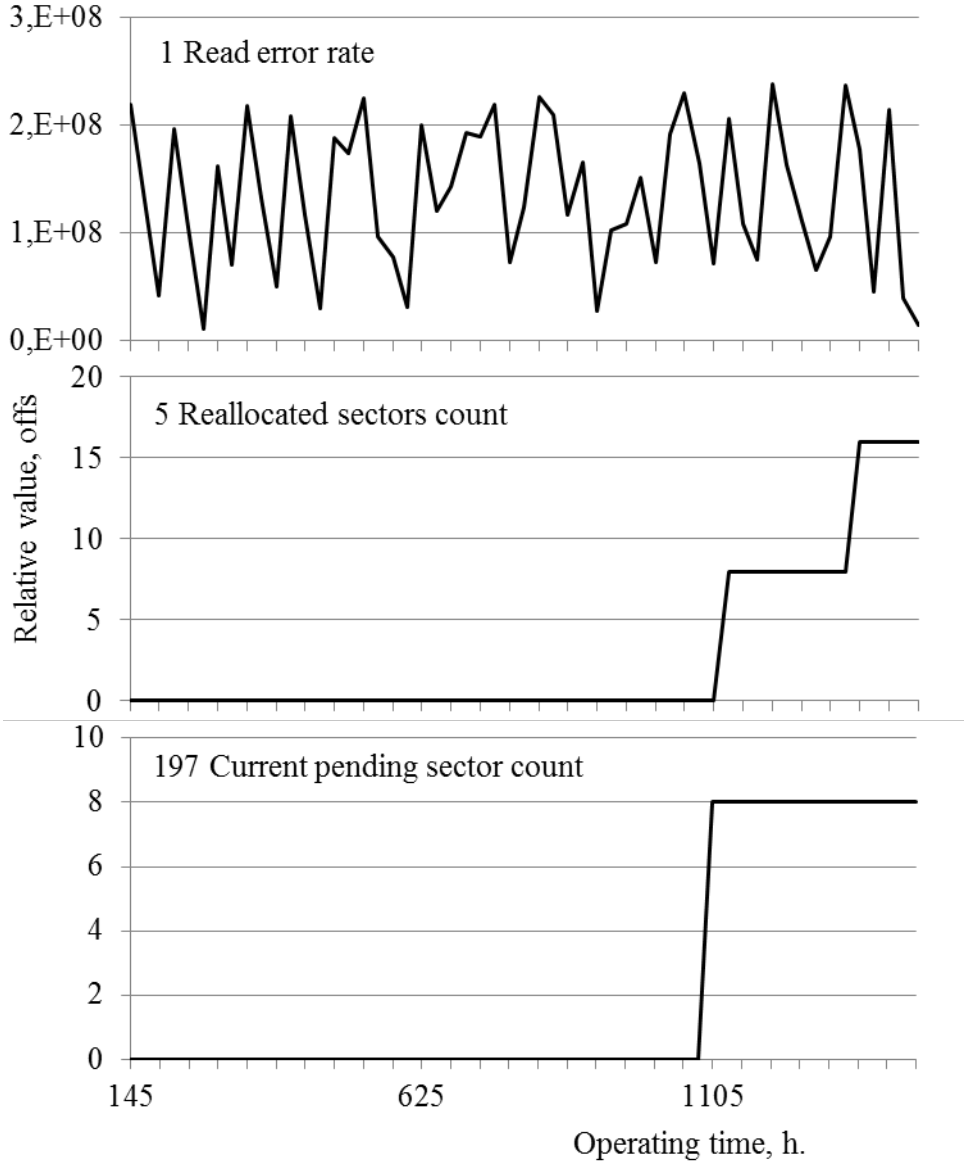

Figure 6: Parameter 1, 5, 197 values depending on the operating time for a failed hard drive model ST4000DM000 with the number Z300X6Y6 and with a capacity of 4 TB

Here, parameter 1 Read error rate characterizes the frequency of errors, the origin of which is due to the hardware of the disk when reading data from a disk. This is the number of internal data corrections made before being issued to the interface. Frighteningly huge numbers may come up. Parameter 5 Reallocated sectors count characterizes the number of reallocated sectors. When a disk detects a read/write error, it marks the sector as "reallocated" and transfers the data to a dedicated backup area. An increase in the value of this attribute may indicate deterioration in the surface of the disk platters. Parameter 196 Reallocation event count shows the number of reallocation operations. The attribute characterizes the total number of attempts to transfer information from reallocated sectors to the backup area. Both successful and unsuccessful attempts are taken into account. Parameter 197 Current pending sector count shows the number of sectors that are candidates for reallocation. They have not yet

been identified as bad, but reading from them is different from reading a stable sector, these are the so-called suspicious or unstable sectors. In case of successful subsequent reading of the sector, it is excluded from the list of candidates. In the case of repeated erroneous readings, the drive tries to restore it and performs a reallocation operation. An increase in the value of this attribute may indicate physical degradation of the hard drive.

A feature of parameters 1, 5, 197 (as well as 9 Power-on hours, which forms the horizontal time axis in the figures, and 194 Temperature, which characterizes the case temperature, but not yet used in the analysis) is that they are available for all drives of all manufacturers. Moreover, they can all be used to evaluate the reliability of hard drives. Therefore, this group of parameters is the most important and mandatory for use in the analysis.

As can be seen from Figures 1-6, parameters 5 and 196 are cumulative; their values only grow depending on the operating time. Although parameter 197 is cumulative, however, if a sector moves from the category of candidates for reallocation to reallocated sectors, then its value decreases. Parameter 1 characterizes the rate of change in the values of other parameters and has the character of a derivative.

The number of failed drives depending on the operating time has two types: falling and dome-shaped [5]. The first type means that disk failures occur immediately after the start of their use, the so-called "infant mortality". The second type is associated with wear and tear and occurs mainly after the expiration of the two-year warranty period. As can be seen from the above figures, failures associated with the surface state of the disks can be of both types.

IV. SUMMARY

As a result, according to the results of the study, it was found that four SMART parameters have time-coinciding changes in values of parameters for failed hard drives. These are parameters with numbers 1, 5, 196, and 197.Some of them, namely, parameters 5, 196, 197, characterize the total number of reallocated sectors, attempts to reallocate them, and candidates for reallocation. Parameter 1 characterizes the rate of change of these parameters.

The scientific novelty of the results is that we can develop criteria for the risk of drive failures based on the identified parameters characterizing the state of the surface of the hard drives. The justification of these criteria is based on the fact of coincidence in time of the change in the values of the specified parameters detected as a result of the analysis.

V. CONCLUSIONS

Similar studies based on the same data with heterogeneous disk groups were carried out in [6], where a search was made for universal predictors of disk failures that could be applied to disks of all makes and models. The main problem was a significant number of SMART parameters, data for which were not available for most brands and models of disks. As a result, the authors were forced to discard parameters that were absent in at least 90% of the disks, after which only 21 parameters remained.

In [7–11], the SMART parameters of the specified data set of the Backblaze data centre were also used to determine the intensity and predict failures of disk information storage devices.

Therefore, the issue of assessing the reliability of information storage devices by the values of SMART parameters is really important for ensuring data security in any organization. It is proposed to solve the problem of individual assessment of the reliability of information storage devices using the identified parameters, based on the detected coincidence in time of changing the values for the parameters of failed hard drives.

ACKNOWLEDGEMENTS

The work is performed according to the Russian Government Program of Competitive Growth of Kazan Federal University.

REFERENCES

[1] S.M.A.R.T. From Wikipedia, the free encyclopedia. URL: https://en.wikipedia.org/wiki/S.M.A.R.T. Checked 24.03.2019.

[2] Hard Drive Data and Stats / Backblaze. URL: https://www.backblaze.com/b2/hard-drive-test-data.html. Checked 24.03.2019.

[3] Nasyrov I.N., Nasyrov I.I., Nasyrov R.I., Khairullin B.A. Data mining for information storage reliability assessment by relative values // International Journal of Engineering and Technology (UAE). 2018. Vol.7, Is.4.7 Special Issue 7. P.204-208. URL: https://www.sciencepubco.com/index.php/ijet/article/view/20545.

[4] Nasyrov I.N., Nasyrov I.I., Nasyrov R.I., Khairullin B.A. Parameters selection for information storage reliability assessment and prediction by absolute values // Journal of Advanced Research in Dynamical and Control Systems. 2018. Vol.10, Is.2 Special Issue. P.2248-2254. URL: http://jardcs.org/backissues/abstract.php?archiveid=5363.

[5] Nasyrov I.N., Nasyrov I.I., Nasyrov R.I., Khairullin B.A. Dependence of reallocated sectors count on HDD power-on time // International Journal of Engineering and Technology (UAE). 2018. Vol.7, Is.4.7 Special Issue 7. P.200-203. URL: https://www.sciencepubco.com/index.php/ijet/article/view/20544.

[6] Rincón C.A.C., Paris J.-F., Vilalta R., Cheng A.M.K., Long D.D.E. Disk failure prediction in heterogeneous environments // Proceedings of the International Symposium on Performance Evaluation of Computer and Telecommunication Systems, SPECTS 2017. Seattle, WA, USA, July 9-12, 2017. URL: http://ieeexplore.ieee.org/document/8046776/.

[7] Qian J., Skelton S., Moore J., Jiang H. P3: Priority-based proactive prediction for soon-to-fail disks // Proceedings of the 10th IEEE International Conference on Networking, Architecture and Storage, NAS 2015. Boston, MA, USA, August 6-7, 2015. – 7255224. – p. 81-86. URL: http://ieeexplore.ieee.org/document/7255224/.

[8] Botezatu M.M., Giurgiu I., Bogojeska J., Wiesmann D. Predicting disk replacement towards reliable data centres// Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '16. San Francisco, California, USA, August 13-17, 2016. – p. 39-48. URL: https://dl.acm.org/citation.cfm?doid=2939672.2939699.

[9] Chaves I.C., de Paula M.R.P., Leite L.G.M., Queiroz L., Pordeus J.P., Machado J.C. BaNHFaP: A Bayesian Network-Based Failure Prediction Approach for Hard Disk Drives // Proceedings of the 5th Brazilian Conference on Intelligent Systems, BRACIS 2016. Recife, Pernambuco, BR, October 9-12, 2016. – 7839624. – p. 427-432. URL: http://ieeexplore.ieee.org/document/7839624/.

[10] Gaber S., Ben-Harush O., Savir A. Predicting HDD failures from compound SMART attributes // Proceedings of the 10th ACM International Systems and Storage Conference, SYSTOR '17. Haifa, Israel, May 22-24, 2017. – Article No. 31. URL: https://dl.acm.org/citation.cfm?doid=3078468.3081875.

[11] Gopalakrishnan P.K., Behdad S. Usage of product lifecycle data to detect hard disk drives failure factors // Proceedings of the ASME International Design Engineering Technical Conference. Cleveland, Ohio, USA, August 6–9, 2017. URL: http://proceedings.asmedigitalcollection.asme.org/proceeding.aspx?articleid=2662132.