

Модель для многопараметрической оценки жестких дисков по риску отказа

И.Н. Насыров (ФГАОУ ВО КФУ), И.И. Насыров (АО «Автоградбанк»),
Р.И. Насыров (ООО «Мединдекс»)

Представлена матрица распределения жестких магнитных дисков в зависимости от значений SMART-параметров с границами диапазонов ячеек согласно установленным критериям надежности. Рассмотрены параметры, при увеличении значений которых возрастает вероятность выхода из строя дисковых накопителей информации. Задачей исследования является разработка модели матричной многопараметрической оценки риска отказа накопителей информации различных производителей. В ходе исследования были проанализированы накопители торговых марок, эксплуатировавшиеся в крупнейшем коммерческом data-центре Backblaze. В результате анализа предложены критерии надежности для следующих наиболее важных в этом отношении параметров: 5 Reallocated sectors count, 196 Reallocation event count, 197 Current pending sector count, 1 Read error rate, 9 Power-on hours, 7 Seek error rate, 3 Spin-up time, 10 Spin-up retry count. Принято во внимание, что само наличие значений рассмотренных параметров зависит от производителя жестких дисков. Новизна полученных результатов состоит в том, что в матричной модели обнаружены периферийные области значений параметров, в которые никакие из отказавших накопителей не попадают.

Ключевые слова: жесткий диск, накопитель, информация, отказ, параметр, моделирование.

Введение

Практическую проверку теоретических моделей часто выполняют на наборах фактических данных, полученных при эксплуатации реальных устройств. В качестве одного из таких примеров данных рассмотрена матрица распределения накопителей информации в зависимости от значений параметров SMART (self-monitoring, analysis and reporting technology — технология самоконтроля, анализа и отчетности), применяемых для внутренней оценки состояния жесткого диска компьютера, а также для предсказания возможного выхода его из строя. Границы диапазонов ячеек выбирались согласно установленным критериям надежности накопителей. Использовались параметры, являющиеся критическими в том смысле, что если их значения увеличиваются, то возрастает вероятность выхода из строя накопителей информации на жестких дисках. Научной задачей исследования является разработка модели матричной многопараметрической оценки риска отказа накопителей информации различных производителей.

В качестве объекта исследования были выбраны накопители, эксплуатировавшиеся в крупнейшем коммерческом data-центре Backblaze. Это жесткие диски торговых марок HGST (Hitachi Global Storage Technologies), Hitachi (позднее HGST), Samsung, ST (Seagate), Toshiba, WDC (Western Digital). В результате исследования предложены критерии надежности для следующих наиболее важных в этом отношении параметров (в порядке убывания степени важности): 5 Reallocated sectors count (число переназначен-

ных секторов), 196 Reallocation event count (число успешных и неуспешных попыток переназначения), 197 Current pending sector count (число секторов, являющихся кандидатами на замену), 1 Read error rate (частота ошибок при чтении данных с диска), 9 Power-on hours (число часов, проведенных во включенном состоянии), 7 Seek error rate (частота ошибок при позиционировании блока магнитных головок), 3 Spin-up time (время раскрутки пакета дисков из состояния покоя до рабочей скорости), 10 Spin-up retry count (число повторных попыток раскрутки дисков до рабочей скорости в случае, если первая попытка была неудачной). Принято во внимание, что само наличие значений рассмотренных параметров зависит от производителя жестких дисков.

Методы

Были проанализированы зависимости значений параметров от времени эксплуатации у вышедших из строя накопителей информации на жестких магнитных дисках, приведенные на сайте компании Backblaze [1]. Изучено изменение 45 параметров SMART 92530 накопителей 93 моделей торговых марок HGST, Hitachi, Samsung, ST, Toshiba, WDC за период с 10 апреля 2013 г. по 31 декабря 2016 г. Выявлено, что на конец исследуемого периода продолжали нормально работать 79,58% накопителей, были досрочно сняты с эксплуатации 14,74%, вышли из строя 5,68%.

Известны сведения о смысловом значении более 80 параметров SMART из 256, однако большинство из них производителями не используются [2]. В свя-

зи с этим специалисты Backblaze записывали в 2013–2014 гг. только 40 из них, а начиная с 2015 г — 45 с номерами 1–5, 7–13, 15, 22, 183, 184, 187–201, 220, 222–226, 240–242, 250–252, 254, 255 (в 2015 г. были добавлены 22, 220, 222, 224, 226).

Результаты и обсуждение

Для отображения состояния поверхности жестких дисков наиболее важной является группа параметров 5, 196, 197, 1, связанных с состоянием пространства памяти [3]. Эти параметры имеются в наличии всегда для всех накопителей всех типов почти всех производителей (196 отсутствует у жестких дисков Samsung и Seagate (торговая марка ST)).

Вторым по важности является тоже всегда имеющийся в наличии параметр 9 Power-on hours [4]. Он соответствует стандартному определению надежности как времени наработки до отказа.

Третьим по важности является параметр 7 Seek error rate, характеризующий частоту ошибок при позиционировании блока магнитных головок [5].

Четвертой по важности является группа параметров, связанных с механикой приводов вращения жестких дисков: 10 Spin-up retry count, 3 Spin-up time [6].

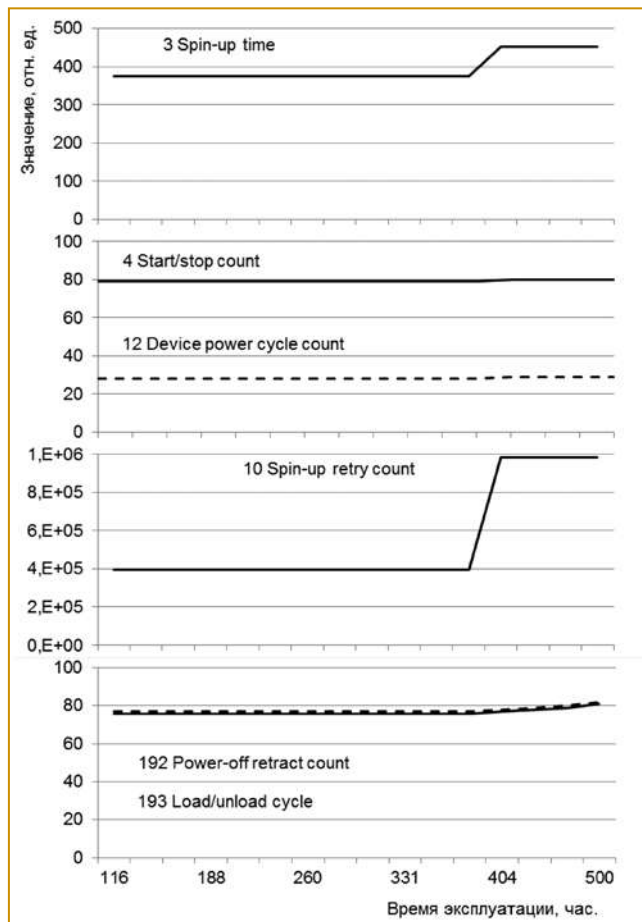


Рис. 1. Значения параметров 3, 4, 10, 192 (сплошные линии), 12, 193 (пунктирные) от времени эксплуатации для отказавшего жесткого диска марки HGST HMS5C4040BLE640 с номером PL2331LAGH007J емкостью 4 ТБ

Обычно у вышедших из строя дисков набор значений параметров, зафиксированный на момент отказа, различается между собой. Более того, различаются даже наборы самих параметров, имеющие в наличии хоть какие-то значения, и зависят от производителя дисков. Тем не менее, даже в таких условиях можно получить некоторую полезную информацию, если использовать для анализа те диски, у которых на момент отказа известны значения сразу нескольких параметров. Анализируя конкретные графические зависимости изменения значений таких параметров можно установить совпадения по времени и пригодность этих параметров для прогнозирования сбоев.

Например, при анализе параметра 10 у вышедшего из строя жесткого диска торговой марки HGST было обнаружено совпадение по времени изменений с параметрами 3, 4, 12, 192, 193 (рис. 1). Причем увеличение значения параметра 10 четко совпадает с увеличением значения параметра 3, хотя такое поведение наблюдается не всегда. Синхронное увеличение значений параметров 4, 12, 192, 193 проявлено гораздо слабее.

Такое же совпадение изменения параметров по времени наблюдалось и для некоторых вышедших из строя дисков марки Hitachi (рис. 2). Здесь относительные из-

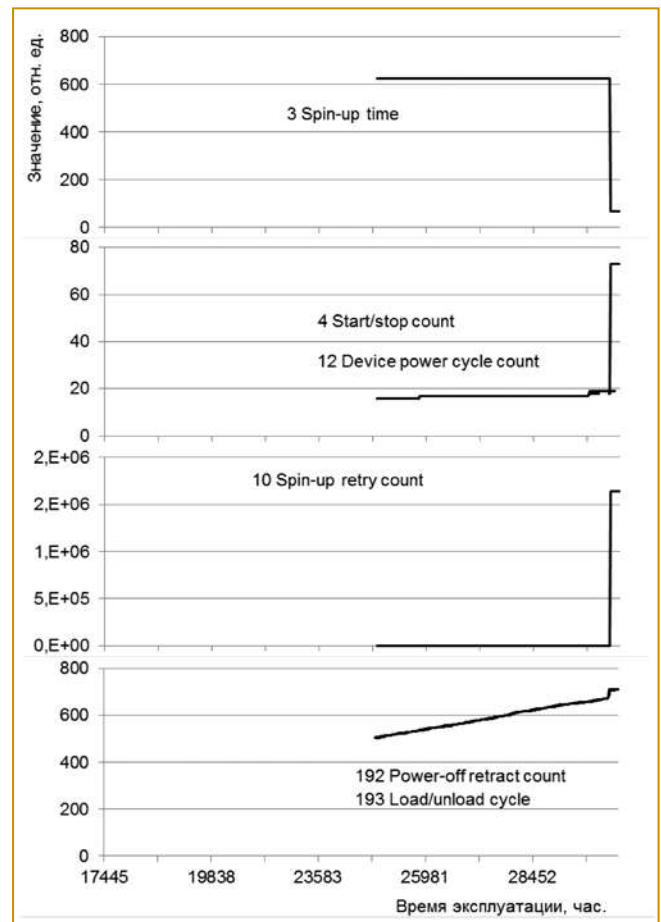


Рис. 2. Значения параметров 3, 4, 10, 192 (сплошные линии), 12, 193 (пунктирные) от времени эксплуатации для отказавшего жесткого диска марки Hitachi HDS722020ALA330 с номером JK1104B8JRWZNNW емкостью 2 ТБ

менения значений параметров 12, 192, 193 так же, как и на рис. 1, гораздо менее выражены, чем у остальных. В связи с этим прямое использование указанных параметров для ранжирования накопителей по степени опасности отказов затруднительно.

Как видно из рис. 1 и 2 параметры 4, 12, 192, 193, которые характеризуют общее число запусков/остановок диска с незначительными смысловыми оттенками, являются накапливаемыми, значения которых от времени эксплуатации только увеличиваются. Но для оценки механической надежности накопителей они не очень удобны, так как зависят не только от состояния приводов дисков, и к тому же относительные изменения значений не всегда четко выражены. Зато параметры 3 и 10, которые оказались взаимосвязанными, вполне могут быть использованы для оценки надежности. Причем у параметра 3 важны именно изменения значений, а у параметра 10 — в основном его абсолютные значения. К тому же ненулевые значения параметра 3 присутствуют у 74 дисков марки HGST из всего 167 вышедших из строя, у 409 из 510 для Hitachi, у 1 из 1 для Samsung, у 16 из 4156 для ST, у 10 из 12 для Toshiba, у 290 из 404 для WDC. При этом параметр 10 хотя и более удобен для ранжирования накопителей по степени опасности отказов, так как он является (или предполагается быть) накапливаемым, однако применять его следует в сочетании с параметром 3 ввиду отсутствия его значений у части накопителей.

Анализ отказов накопителей по относительным [7] или абсолютным [8] значениям всех указанных выше групп параметров подтверждает их важность для разработки модели оценки вероятности выхода из строя жестких дисков. Следовательно, исходя из смысла и наличия значений параметров при отображении накопителей по степени риска отказа, нужно использовать сортировку сначала по параметрам 5, 196, 197 и 1, затем по 9, далее по 7, 3 и 10.

Наибольшее число вышедших из строя накопителей в зависимости от времени эксплуатации имеет низкие или нулевые значения параметров. Также обнаружен факт переназначения секторов целыми дорожками (кратно 8). Исходя из этого, вначале для оценки

вероятности выхода из строя накопителей была проведена кластеризация по четырем диапазонам последних известных значений пока только параметров $p = 1, 5, 9, 197$ и построена матрица размером $4 \times 4 \times 4 \times 4$. Вероятность W^p определялась как среднеарифметическая величина путем деления числа отказавших накопителей со значениями параметров p , попадающих в определенную ячейку матрицы, на число всех накопителей с указанными значениями параметров:

$$W^p = \frac{\sum_{k=1}^n v_k^p}{\sum_{i=1}^l s_i^p + \sum_{j=1}^m u_j^p + \sum_{k=1}^n v_k^p}, \quad (1)$$

где s, u, v — нормальные, снятые досрочно, отказавшие накопители со значениями параметра из диапазона, l, m, n — их число, i, j, k — индексы суммирования.

Таблица. Матрица вероятности отказов накопителей информации, используемых в data-центре компании Backblaze (параметры $p1, p5, p197$ в ед., параметр $p9$ в часах, вероятность в относительных единицах)

	p197=0				p197=1				2≤p197≤8				p5=0				p197>8			
	p5=0	p5=1	2≤p5≤8	p5>8	p5=0	p5=1	2≤p5≤8	p5>8	p5=0	p5=1	2≤p5≤8	p5>8	p5=0	p5=1	2≤p5≤8	p5>8	p5=0	p5=1	2≤p5≤8	p5>8
p9<8760	норма	6658	1	2	8	0	0	0	0	3	0	0	0	2	0	0	0	0	0	0
	досрочно	877	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
	отказ	138	2	4	15	3	0	0	0	3	1	0	3	5	0	2	14			
	p1=0	0,018	0,6667	0,6667	0,625	1	0,3333	0,6667	1	0,5	1	0,6667	1	0,625	1	1	1	1		
	норма	37	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	досрочно	4	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
	отказ	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	p1=1	0,0682	0,3333	0,6667	1	0,3333	0,3333	0,6667	1	0	0,6667	0,6667	1	1	1	1	1	1	1	1
	норма	80	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	досрочно	5	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	отказ	11	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	2≤p1≤8	0,1146	0,6667	0,6667	1	0,6667	0,6667	0,6667	1	0,6667	0,6667	0,6667	1	1	1	1	1	1	1	1
норма	15245	0	21	8	0	0	0	0	35	1	1	3	10	0	0	0	0	0	0	
досрочно	221	0	2	12	0	0	0	1	1	0	1	1	4	1	1	5				
отказ	579	3	26	70	15	0	0	3	89	5	9	15	175	0	11	129				
p1>8	0,0361	1	0,5306	0,7778	1	1	1	1	0,75	0,712	0,8333	0,8182	0,7895	0,9259	0	0,9167	0,9627			
8760≤p9<17520	норма	856	0	1	2	0	0	0	0	1	0	0	0	0	0	0	0	0	0	
	досрочно	326	2	0	0	17	0	0	0	23	1	1	0	22	0	0	0	0	0	
	отказ	85	1	2	10	1	0	0	2	3	0	0	2	5	0	2	6			
	p1=0	0,0671	0,3333	0,6667	0,8333	0,0556	0,3333	0,6667	1	0,1111	0	0	1	0,1852	1	1	1	1	1	
	норма	27	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
	досрочно	8	0	0	0	3	0	0	0	1	0	0	1	4	0	0	0	0	0	
	отказ	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	
	p1=1	0,1026	0,3333	0,6667	1	0	0,3333	0,6667	1	0	0,6667	0,6667	0	0	0	1	1	1	1	
	норма	112	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
	досрочно	9	2	0	0	2	0	0	0	3	0	0	0	2	1	0	0	0	0	
	отказ	9	0	0	1	2	0	0	0	0	0	0	0	0	0	0	0	0	1	
	2≤p1≤8	0,0692	0	0,6667	1	0,5	0,6667	0,6667	1	0	0,6667	0,6667	1	0	0	1	1	1	1	
норма	18617	0	13	10	0	0	0	0	71	0	2	1	19	0	1	1				
досрочно	559	0	4	77	10	0	0	0	28	0	0	7	38	0	0	33				
отказ	456	2	14	142	13	1	1	1	75	4	7	34	279	0	16	250				
p1>8	0,0232	1	0,4516	0,6201	0,5652	1	1	1	0,431	1	0,7778	0,8095	0,8304	1	0,9412	0,8803				
17520≤p9<43800	норма	17000	27	23	88	2	0	0	0	9	0	1	1	0	0	0	0	0		
	досрочно	495	11	9	8	11	0	1	0	23	1	1	0	10	0	0	2			
	отказ	101	12	23	44	7	0	8	11	12	2	2	19	8	2	1	20			
	p1=0	0,0057	0,24	0,4182	0,3143	0,35	0,6667	0,8889	1	0,2727	0,6667	0,5	0,95	0,4444	1	1	0,9091			
	норма	78	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0		
	досрочно	27	0	1	0	1	0	0	0	3	0	0	0	0	0	0	0	0		
	отказ	5	0	0	0	1	0	0	0	1	0	0	0	1	0	0	1			
	p1=1	0,0455	0,6667	0	0	0,5	0,6667	0,6667	1	0,25	0,6667	0,6667	1	1	1	1	1	1		
	норма	97	1	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0		
	досрочно	33	1	1	1	1	0	0	0	3	0	0	0	2	0	0	0	0		
	отказ	6	0	0	3	1	0	0	0	1	0	0	0	3	0	0	0	1		
	2≤p1≤8	0,0441	0	0	0,6	0,5	0,6667	0,6667	1	0,25	0,6667	0,6667	1	0	1	1	1	1		
норма	11852	3	40	33	5	0	0	0	61	0	3	7	23	0	1	2				
досрочно	3155	23	88	712	12	1	1	12	88	1	3	65	93	0	1	102				
отказ	627	7	19	403	12	0	2	41	88	0	9	110	333	4	14	281				
p1>8	0,0401	0,2121	0,1293	0,351	0,4138	0	0,6667	0,7736	0,3713	0	0,6	0,6044	0,7416	1	0,875	0,7299				
p9>43800	норма	1952	9	12	33	3	0	0	0	2	0	0	2	0	0	0	0	0		
	досрочно	3874	204	184	89	7	3	7	7	7	1	1	3	0	0	0	0	0		
	отказ	35	4	3	19	3	1	1	1	3	0	0	0	1	0	1	0	0		
	p1=0	0,006	0,0184	0,0151	0,1348	0,2308	0,25	0,125	0,125	0,25	0	0	0	1	1	1	1	1		
	норма	46	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
	досрочно	71	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0		
	отказ	5	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0		
	p1=1	0,041	0,5	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1		
	норма	26	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0		
	досрочно	158	1	1	1	1	0	0	0	2	0	0	0	0	1	0	0	0		
	отказ	7	2	2	0	1	0	0	0	0	0	0	0	0	0	0	0	0		
	2≤p1≤8	0,0366	0,5	0,6667	0	0,5	1	1	1	0	1	1	1	1	1	1	1	1		
норма	270	1	1	14	3	0	0	1	0	0	0	0	0	0	0	0	0			
досрочно	1300	26	93	296	1	0	1	5	8	1	0	5	1	0	0	1	0			
отказ	36	1	3	54	1	0	1	9	5	0	2	20	6	0	0	10				
p1>8	0,0224	0,0357	0,0309	0,1484	0,2	1	0,5	0,6	0,3846	0	1	0,8	0,8571	1	1	0,9091				

Даже с учетом широты диапазона кластеризации (< 8760 (один год), 8760...17520, 17520...43800, >43800 ч для времени эксплуатации; 0, 1, 2...8, >8 для остальных параметров) не все ячейки матрицы были заполнены примерами, в некоторых пришлось использовать предположительные значения 1/3, 2/3, 1. В таблице эти ячейки выделены серым, светло-коричневым и коричневым цветом соответственно.

В процессе составления таблицы дополнительно были выявлены 52 работоспособных диска (2 ед. марки WDC WD10EADS емкостью 1 ТБ и 50 ед. марки Seagate ST4000DM000 емкостью 4 ТБ), а также два вышедших из строя диска марки Seagate ST4000DM000 емкостью 4 ТБ, которые были помечены как снятые с эксплуатации досрочно. Однако было установлено, что это произошло по причине сбоя на 24.01.2015 для 2 ед. WDC, на 19.06.2015 для 29 ед. ST, на 31.01.2016 для 22 ед. ST, на 24.03.2016 для 1 ед. ST. Сбой заключался в том, что никакие данные по ним за эти дни не были указаны. В связи с этим были взяты данные на последний день работы накопителей и добавлены в таблицу. Однако таблица ни на одну из выделенных цветом ячеек не уменьшилась. То есть значения добавились в уже известные позиции.

Выводы

Показано, что на базе наиболее важных параметров надежности, значения которых имеются в наличии для жестких магнитных дисков всех производителей, можно разработать модель матричной многопараметрической оценки надежности накопителей информации с учетом вероятности их отказа.

В перспективе, при переходе крупных data-центров к хранению информации в твердотельных накопителях информации, для характеристики состояния пространства памяти можно применить точно такие же параметры практически с тем же самым смыслом, за исключением естественно параметров позиционирования головок записи/считывания или механизмов вращения дисков, хотя даже у них нумерация остается прежней. Следовательно, никаких дополнительных изменений кроме изъятия параметров 3, 7 и 10 в разработанную модель вносить не придется.

Новизна полученных результатов состоит в том, что в матрице модели обнаружены периферийные области значений параметров, в которые никакие из отказавших накопителей не попадают. Это означает, что для теоретических моделей оценки надежности оборудования не всегда можно получить экспериментальные подтверждения в реальных условиях технической эксплуатации. Остается использовать предполагаемые значения.

Заключение

Для определения наиболее подходящего способа прогнозирования отказов в [9] протестировали

пять алгоритмов машинного обучения на тех же данных, при этом рассматривались жесткие диски только одной, но самой многочисленной марки Seagate ST4000DM000 за период с 2015 по 2018 гг. Для исследования из 44 параметров были выбраны пять с номерами 5, 187, 188, 197, 198. Еще одно допущение — были линейно интерполированы пропущенные в отдельных промежутках времени данные. Отмечено, что три алгоритма имеют довольно близкую и правильную производительность, однако искусственная нейронная сеть, по мнению авторов, больше подходит для предлагаемой в работе облачной архитектуры.

В [10] также подчеркнута проблема, заключающаяся в том, что технология SMART не полностью стандартизирована. Измеряемый набор атрибутов и смысловые детали реализации различны для каждого производителя жестких дисков. В указанной работе рассматривались девять параметров с номерами 5, 12, 187, 188, 189, 190, 198, 199, 200. С точки зрения машинного обучения нет никакой гарантии, что модель, обученная предсказывать отказы конкретной марки жесткого диска, сможет точно предсказать отказы другой марки жесткого диска. По этой причине для того, чтобы сделать выводы о прогнозировании отказов жестких дисков в целом, важно обеспечить, чтобы предлагаемые прогностические модели оценивались на основе различных марок жестких дисков различных производителей, а также тестировались на различных марках жестких дисков. До сих пор это ограничение не принималось во внимание должным образом, вероятно потому, что сбор репрезентативного набора данных был проблемой для многих предыдущих исследований, что ухудшало общность их выводов.

Обнаруженный в настоящем исследовании факт единичных или полностью отсутствующих значений параметров в значительной части ячеек с бесконечной границей нигде в других работах не упоминался. Возникает закономерный вопрос, на каких же данных в этих диапазонах тренируются алгоритмы машинного обучения и нейронные сети?

Таким образом, рассматриваемая тема является актуальной, а изложенная в статье модель обладает научной новизной и позволяет задачу многопараметрической оценки надежности накопителей информации решать с применением матрицы распределения жестких магнитных дисков в зависимости от значений, выявленных в результате исследования наиболее важных с этой точки зрения параметров.

Список литературы

1. Backblaze. Hard Drive Data and Stats. <https://www.backblaze.com/b2/hard-drive-test-data.html>
2. Nasyrov I.N., Nasyrov I.I., Nasyrov R.I., Khairullin B.A. Study of Failure Hazard Degree in Large Data Centers // Helix. 2019. V. 9. Is. 5. PP. 5345-5349.
3. Nasyrov I.N., Nasyrov I.I., Nasyrov R.I., Khairullin B.A. Reallocated sectors count parameter for analysing hard disk drive reliability // Journal of Computational and Theoretical Nanoscience. 2019. V. 16. Is. 12. PP. 5298-5302.

4. *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. V. 7. Is. 4.7 Special issue 7. PP. 200-203.
5. *Nasyrov I.N., Nasyrov I.I., Nasyrov R.I., Khairullin B.A.* Positioning errors indication by Seek error rate and other HDD parameters // Journal of Advanced Research in Dynamical and Control Systems. 2019. V. 11. Is. 8 Special issue. PP. 1797-1805.
6. *Nasyrov I.N., Nasyrov I.I., Nasyrov R.I., Khairullin B.A.* Spin retry count relation with other HDD parameters // Journal of Computational and Theoretical Nanoscience. 2019. V. 16. Is. 12. PP. 5303-5306.
7. *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. V. 7. Is. 4.7 Special issue 7. PP. 204-208.
8. *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. V. 10. Is. 2 Special issue. PP. 2248-2254.
9. *Ragmani A., Elomri A., Abghour N., Moussaid K., Rida M., Badidi E.* Adaptive fault-tolerant model for improving cloud computing performance using artificial neural network//Proceedings of the 11th International Conference on Ambient Systems, Networks and Technologies ANT/3rd International Conference on Emerging Data and Industry 4.0 EDI (Warsaw, Poland, April 06–09, 2020). Elsevier Science BV: 2020. PP. 929-934.
10. *Aussel N., Jaulin S., Gandon G., Petetin Y., Fazli E., Chabridon S.* Predictive models of hard drive failures based on operational data // Proceedings of the 16th IEEE International Conference on Machine Learning and Applications ICMLA (Cancun, Mexico, December 18-21, 2017). IEEE: 2017. PP. 619-625.

Насыров Искандар Наилевич – д-р экон. наук, проф. ФГАОУ ВО «Казанский (Приволжский) федеральный университет», Насыров Ильдар Искандарович – канд. техн. наук, руководитель отдела технических систем и телекоммуникаций, АО «Автоградбанк», Насыров Рустам Искандарович – руководитель проектов, ООО «Мединдекс». E-mail: ecoseti@yandex.ru

Робот-мерчандайзер на основе RFID технологии

Компания R2Robotics (Россия) осуществила разработку робота-мерчандайзера, способного проводить инвентаризацию при считывании RFID-меток на товарах. R2 – универсальный помощник для инвентаризации товаров в магазинах розничной торговли.

Преимущества робота перед работой человека

Во-первых, точность результатов автономной инвентаризации составляет $\geq 98\%$, что не ниже, чем у ручного считывания. С такой же точностью на складах и в торговых залах работают люди. Но сотрудники тратят на инвентаризацию аналогичного объема товара ≥ 6 ч.

R2Robotics создала и протестировала первый прототип робота в магазине международного ритейлера спортивной одежды в Новосибирске. Тесты показали, что вместо целого дня на инвентаризацию можно потратить всего 1 час работы робота и освободить персонал для продаж в магазине. Полученные данные тестов позволили оптимизировать алгоритм локализации RFID меток в пространстве. Компания оформила ПО на систему определения местоположения пассивных RFID меток в пространстве с точностью до $0,5 \text{ м}^3$ в качестве результата интеллектуальной деятельности. Высокая точность определения координат товара (местоположения) по пассивным RFID меткам является отличительным преимуществом R2 перед аналогами.

Во-вторых, R2 выдает подсказки по месторасположению товарных позиций на полках магазина (например, предупреждает, если товар находится не на своем месте или же закончился на прилавке), что зачастую не замечает менеджер магазина или продавец-консультант. Проводя сканирование тэгов на товарах в магазине, робот создает полную планограмму, которая в дальнейшем позволит отслеживать статистику продаж и предупреждать о необходимости дополнительных поставок.

Порядок работы R2

Для сбора данных R2 использует комбинацию автономной навигации, машинного зрения и детекции RFID-меток. Ультразвуковые и лазерные датчики измерения расстояния позволяют R2 ориентироваться в пространстве, чтобы определить препятствия (полки, стеллажи, примерочные, а также зеркала и стеклянные витрины). В это же время, машинное зрение позволяет построить схему помещения, в котором робот будет передвигаться. Далее робот собирает информацию об окружении и находящихся рядом объектах и планирует свой маршрут, по которому собирает информацию о расположении товара в магазинах. Затем информация в обработанном виде передается на компьютер клиента, при помощи чего уже можно наладить выкладку товаров и пополнить запасы.

Робот R2 был специально разработан с учетом особенностей рынка розничной торговли и в любой момент может подстроиться под изменения окружающей среды, например, присутствие людей в торговом зале, стеклянные и зеркальные объекты, витрины и стеллажи. Робот самостоятельно создает карту помещения, на которой менеджер отмечает маршрут инспектирования. Модель определяет местоположение чипированного товара с точностью до единицы торгового оборудования. В отличие от человека, R2 делает фотоснимки в одинаковых точках с фиксированными ракурсами, что позволяет проводить быстрый машинный анализ выкладки. R2 точно определяет местонахождение каждой метки, благодаря чему может быть создана качественная реальная планограмма торгового зала. Робот был создан таким, чтобы клиенту было легко использовать полученные данные для более эффективной торговли, мерчандайзинга и инвентаризации.

[Http://www.r2robotics.ru](http://www.r2robotics.ru)