

УДК 004.083.72

Насыров Р.И., старший преподаватель, Набережночелнинский институт ФГАОУ ВО «Казанский (Приволжский) федеральный университет», rinasuyarov@gmail.com

Насыров И.Н., доктор экономических наук, профессор, Набережночелнинский институт ФГАОУ ВО «Казанский (Приволжский) федеральный университет», ecoseti@yandex.ru, ResearcherID A-8016-2016

ЭФФЕКТИВНОСТЬ АЛГОРИТМА ПРОГНОЗИРОВАНИЯ НАДЕЖНОСТИ НАКОПИТЕЛЕЙ ИНФОРМАЦИИ ДЛЯ КРУПНЫХ DATA-ЦЕНТРОВ¹

Аннотация: В крупных data-центрах на основании статистических данных об отказах можно сделать вывод о том, сколько накопителей выйдут из строя в ближайшее время, но какие конкретно – узнать нельзя. А для обеспечения безопасности информации нужно знать заранее какой именно накопитель откажет, т.к. безопасность технически гарантируется полнотой и своевременностью копирования информации с ненадежного на новый и надежный накопитель.

Целью исследования является повышение эффективности прогнозирования надежности конкретных накопителей информации.

Разработан алгоритм прогнозирования надежности отдельных накопителей информации, позволяющий маркировать те из них, у которых есть опасность выхода из строя.

Дополнительно с помощью моделирования получено, что если замену вышедшего из строя оборудования производить на более надежные накопители, то количество отказов с течением времени будет уменьшаться, в относительных единицах экономия за счет снижения количества требующих замен накопителей составит 5% за два года, 18% за пять лет и до 35% за десять лет.

Ключевые слова: накопитель информации; надежность; эффективность.

При оценке качества и надежности оборудования компьютерных систем управления предприятиями выявлено, что наиболее критичными элементами являются накопители информации [1]. Основные показатели качества накопителей связаны с механикой жестких дисков и среди них главным необходимо выделить переназначенные сектора [2, 3]. Выставив к этому

¹ Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 16-37-00002 мол_а.

показателю систему критериев можно ранжировать накопители по степени надежности [4, 5].

Однако при изучении зависимости количества переназначенных секторов от времени эксплуатации накопителей обнаружена ступенчатая нелинейность показателя. В этих условиях для моделирования и прогнозирования работоспособности накопителей информации наиболее подходящим оказался численный метод конечных разностей с линейной аппроксимацией [6-8]. Использование нейронных сетей ввиду неустойчивости их работы в условиях ступенчатости изменения показателя представляется нецелесообразным [9-13].

Методику оценки надежности накопителей желательно закрепить в виде некоторого свода правил аналитической (экспертной) системы [14, 15]. В результате в целом образуется алгоритм, позволяющий ранжировать накопители информации по степени опасности, в том числе и при ступенчатой нелинейности показателя надежности [16, 17].

Целевая функция алгоритма оценки и прогнозирования надежности накопителей информации в системах управления предприятиями включает затраты на приобретение накопителей, оплату труда, а также упущенную выгоду от потери заказов. Оптимизация работы предприятия достигается при минимуме целевой функции [18, 19].

Оценка потребности в программе прогнозирования надежности накопителей информации, полученная путем опроса среди частных пользователей, показывает наличие хотя и устойчивого, однако не очень значительного спроса на нее [20-23].

Таким образом, встает вопрос о поиске области более эффективного применения разработанного метода прогнозирования надежности накопителей информации [24, 25]. В частности, представляют интерес облачные технологии хранения информации с большими объемами памяти. Какая для этого необходима модификация алгоритма?

В крупных data-центрах на основании статистических данных об отказах можно сделать вывод о том, сколько накопителей выйдут из строя в ближайшее

время, но какие конкретно – узнать нельзя. А для обеспечения безопасности информации нужно знать заранее какой именно накопитель откажет, т.к. безопасность технически гарантируется полнотой и своевременностью копирования информации с ненадежного на новый и надежный накопитель. При этом актуальным является сохранение экономической эффективности функционирования data-центра.

Исходя из этого сформулирована научная задача – разработка эффективного алгоритма прогнозирования надежности отдельных накопителей информации для подобных центров.

Целью исследования является повышение эффективности прогнозирования надежности конкретных накопителей информации.

Объект исследования – накопители информации, собранные в RAID-массивы и используемые в data-центрах.

Предмет исследования – надежность накопителей информации.

Метод исследования – численное моделирование функционирования data-центра с точки зрения количества отказов накопителей информации и проведения необходимых замен оборудования.

Рассмотрены данные, собранные за период с 2013 по 2016 годы одним из крупнейших в мире коммерческих data-центров Backblaze [26]. Эти данные получены за каждый день наблюдения начиная с 10.04.2013 г., сгруппированы по годам и выложены в открытом доступе на фирменном сайте. В частности для моделирования было принято во внимание, что общее число накопителей составляло к 2016 году свыше 60000 штук, число используемых моделей накопителей доходило до 90 разновидностей, а число производителей – до 6. Надежность различных моделей накопителей варьировалась от значений близких к 0% до свыше 10% отказов в год за весь период наблюдения.

Главным фактором, определяющим эффективность всего бизнеса, является ценность самой хранимой информации. Ведь именно для обеспечения ее сохранности клиенты идут на дублирование и платят за это деньги. Да и в самих data-центрах эту информацию еще раз дублируют с помощью RAID-

массивов. Тем не менее, опасность одновременного или близкого по времени выхода из строя двух или более накопителей в одном массиве остается.

Для предотвращения возможности возникновения подобной ситуации нами разработан модифицированный алгоритм прогнозирования надежности, который позволяет по одному или нескольким параметрам опознать накопители информации с признаками опасности выхода из строя (рис. 1).

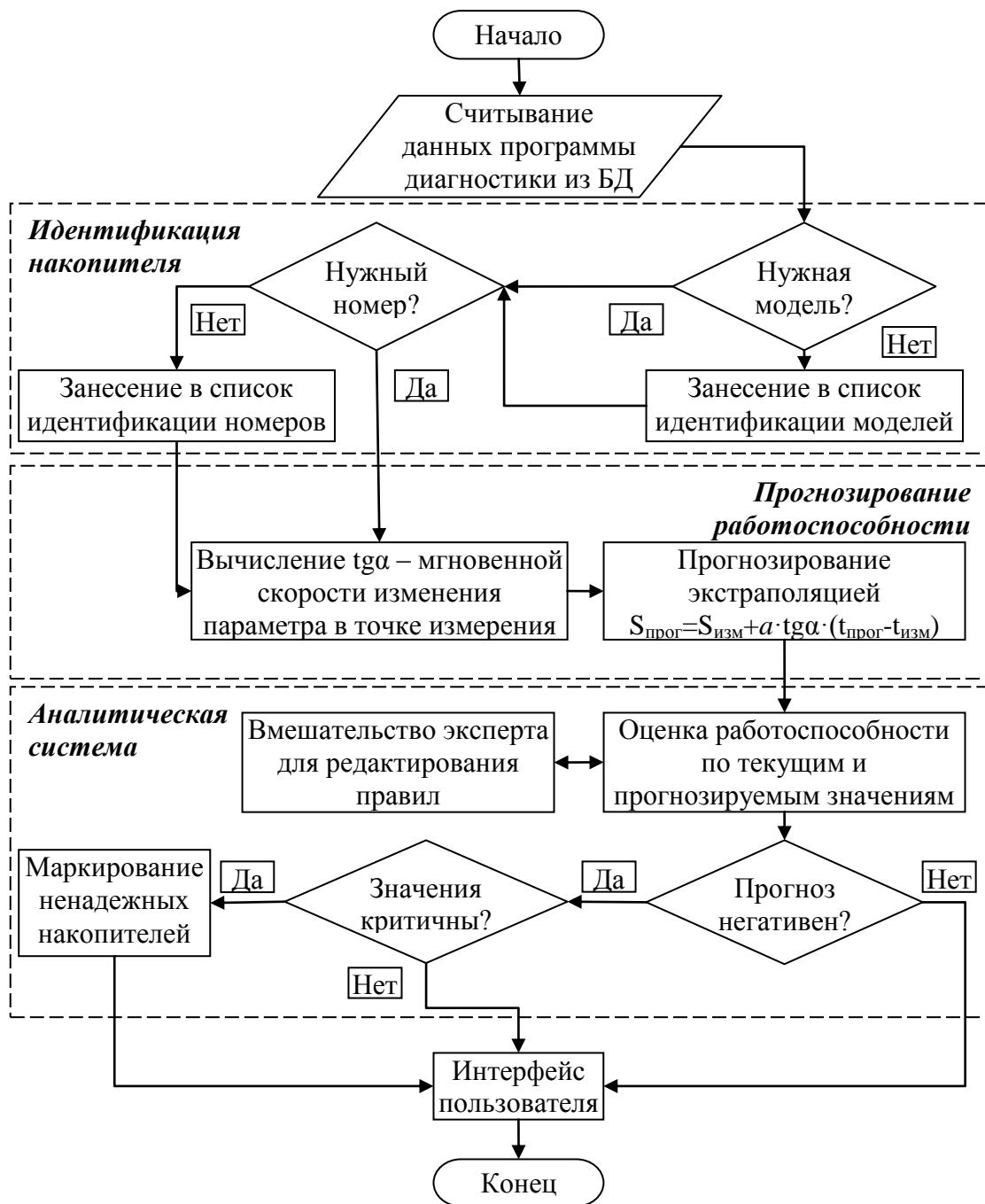


Рис. 1. Алгоритм прогнозирования надежности накопителей информации

Отличительными особенностями этого алгоритма являются:

- 1) возможность анализа по любому параметру из SMART-таблиц;
- 2) то, что анализ ведется одновременно как по текущим так и по прогнозным значениям параметра.

Результатом применения алгоритма будет маркировка накопителей как надежных или как ненадежных с разным уровнем опасности отказа. Это позволит оператору при отказе одного из накопителей в RAID-массиве опираясь на сведения о состоянии остальных дисков в массиве принять решение о замене только отказавшего накопителя или о полном копировании информации из оставшихся работоспособных дисков на новые. А накопители из этого массива утилизировать: отказавший вернуть изготовителю для обмена по гарантии, ненадежные продать по низкой цене, надежные тоже продать по цене чуть выше или продолжить их использование в data-центре.

Дополнительно к алгоритму смоделирован процесс замены вышедшего из строя оборудования применительно к реальной ситуации в Backblaze. Получено, что если отказавшие накопители заменять на аналогичные с такой же вероятностью отказов (для примера рассмотрена величина в 10% в год), то на графике зависимости количества замененных единиц оборудования от времени эксплуатации получится прямая линия (рис. 2). Но если замену производить на более надежные накопители, то количество отказов с течением времени будет уменьшаться, причем, как видно из графика, нелинейно. В пределе, при замене только на безотказные накопители с вероятностью выхода из строя 0%, количество вновь установленных накопителей через 10 лет снизится с 60000 до 39080 единиц.

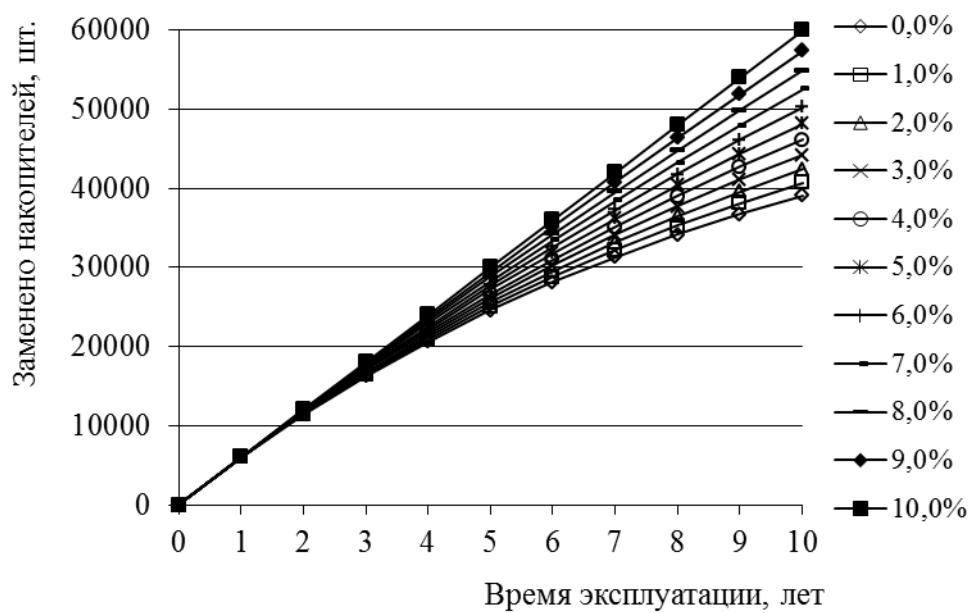


Рис. 2. Количество замен от времени эксплуатации при различной надежности накопителей информации

В относительных единицах экономия за счет снижения количества требующих замен накопителей составит 5% за два года, 18% за пять лет и до 35% за десять лет (рис. 3). Здесь также заметна нелинейность зависимости, но теперь уже не от времени эксплуатации, а от надежности оборудования.

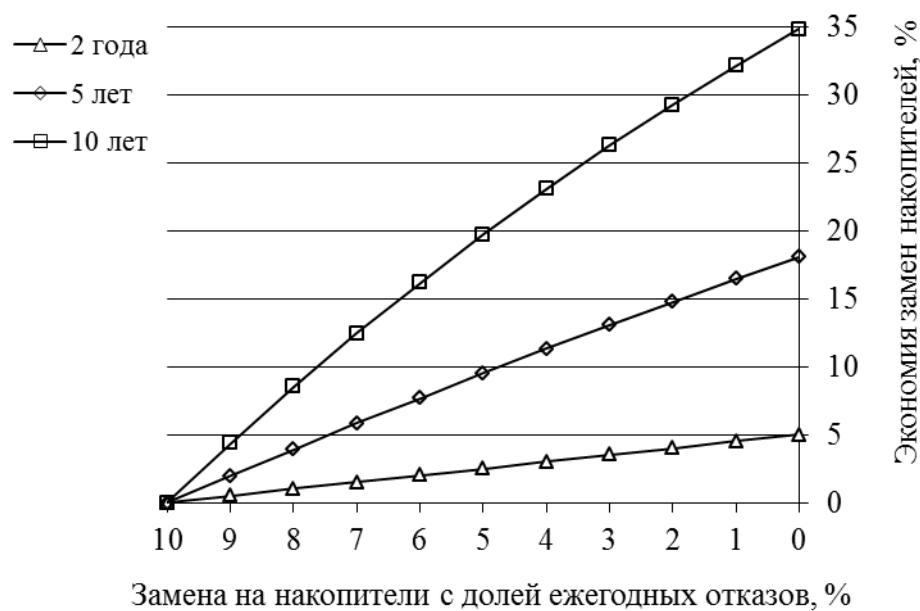


Рис. 3. Экономия при замене на накопители с различной надежностью спустя определенные периоды времени

Вместе с этим имеется влияние на эффективность применения изложенного порядка также и других факторов. Среди положительных следует отметить, в первую очередь, установленные гарантии на изделия не только продавцами, но и изготовителями, если закупки ведутся напрямую у производителя. Вторым положительным фактором является то, что с течением времени общая стоимость единицы памяти постепенно снижается и последующие закупки накопителей обходятся дешевле. Среди негативных факторов, влияющих на эффективность, надо указать то, что более надежные накопители могут стоить дороже и тогда экономия может существенно уменьшиться. Также не все производители и продавцы могут сразу поставить крупную партию накопителей, поэтому приходится закупать то, что имеется в нужном количестве в данный момент на рынке. Кроме этого технический прогресс позволяет выводить на рынок гораздо большие по объему памяти накопители. Вследствие этого происходит достаточно быстрое моральное устаревание установленного оборудования, которое для поддержания эффективности приходится также заменять, хотя оно и вполне исправно.

Таким образом, наибольший эффект от применения модифицированного алгоритма прогнозирования надежности накопителей информации для крупных data-центров достигается за счет обеспечения сохранности информации клиентов. Снижение риска ее потери без удорожания относительно низкой стоимости хранения приведет к дальнейшему расширению бизнеса. А наличие дополнительного эффекта от экономии при замене вышедших из строя накопителей на более надежные еще более упрочит конкурентные преимущества фирмы.

Литература

1. Насыров И.И., Насыров Р.И. Проблема оценки качества и надежности оборудования сложных автоматизированных систем управления производством // III Камские чтения: тезисы докл. межрег. науч.-практ. конф. (Набережные

Челны, 30 апреля 2011 г.). В 3-х ч. Часть 3. – Набережные Челны: ИНЭКА, 2011. – С.112-113.

2. Насыров Р.И. Показатели ранжирования накопителей информации по степени надежности // VIII Камские чтения: тезисы докл. всерос. науч.-прак. конф. (Набережные Челны, 22 апреля 2016 г.). В 2-х ч. Часть 1. – Набережные Челны: ИПЦ НЧИ КФУ, 2016. – С.122-124.
3. Насыров Р.И. Критерии и показатели ранжирования накопителей информации по степени надежности // Вестник компьютерных и информационных технологий. – 2016. - № 10 (148). – С.30-35.
4. Насыров И.И., Насыров Р.И. Критерии оценки надежности накопителей информации на жестких дисках // Социально-экономические и технические системы. – 2011. - № 3 (60). – С.1-6.
5. Насыров Р.И. Обоснование выбора критериев ранжирования накопителей информации по степени надежности // Информационные технологии. Автоматизация. Актуализация и решение проблем подготовки высококвалифицированных кадров (ИТАП-2016): тезисы докл. межд. науч.-прак. конф. (Набережные Челны, 16 мая 2016 г.). – Набережные Челны: Изд-во НЧИ КФУ, 2016. – С.57-60.
6. Насыров Р.И. Функциональное моделирование работоспособности оборудования систем управления машиностроительным предприятием // Автомобиле- и тракторостроение в России: приоритеты развития и подготовка кадров: тезисы докл. 77-й межд. науч.-техн. конф. ААИ. (Москва, 27-28 марта 2012 г.). – Кн. 13. – М.: МГТУ «МАМИ», 2012. – С.44-46.
7. Насыров И.И., Насыров Р.И. Прогнозирование надежности накопителей информации на жестких дисках // Образование и наука Закамья Татарстана. – 2011. - № 22. [Электронный ресурс]. – Режим доступа: <https://cloud.mail.ru/public/FKme/WGD7GtaxX> (дата обращения 09.02.2017).
8. Насыров Р.И. Численный метод прогнозирования ступенчатой нелинейности показателя надежности накопителей информации в системах управления организациями // Современные проблемы информатизации российской

экономики: тезисы докл. межд. науч.-прак. конф., посвященной 20-летию Академии информатизации РТ. (Казань, 13-14 декабря 2013 г.). – Казань: ИЭУП, 2013. – С.107-113.

9. Насыров И.И., Насыров Р.И. Анализ зависимости работы накопителей информации на жестких дисках для нечеткой нейросетевой модели оценки надежности систем управления промышленными предприятиями // Проектирование и исследование технических систем: межвуз. науч. сб. Вып. 4 (18). – Набережные Челны: ИНЭКА, 2011. – С.28-32.

10. Насыров Р.И. Критерий адекватности математических моделей прогнозирования надежности накопителей информации в системах управления машиностроительными предприятиями // Инновационные машиностроительные технологии, оборудование и материалы – 2013 (ИМТОМ-2013): тезисы докл. межд. науч.-техн. конф. (Казань, 11-13 сентября 2013 г.). – Казань: КФУ, 2013. – С.385-388.

11. Насыров Р.И., Тимергалиев С.Н. Нейросетевое моделирование ступенчатой нелинейности показателя надежности накопителей информации // Образование и наука Закамья Татарстана. – 2014. - № 28. [Электронный ресурс]. – Режим доступа: <https://cloud.mail.ru/public/4tsr/inK5gvP5t> (дата обращения 09.02.2017).

12. Насыров Р.И. Адекватность отображения нейронными сетями ступенчатой нелинейности показателя надежности накопителей информации // VI Камские чтения: тезисы докл. всерос. науч.-прак. конф. студентов, аспирантов и молодых ученых. (Набережные Челны, 25 апреля 2014 г.). В 3-х ч. Часть 1. – Набережные Челны: НЧИ КФУ, 2014. – С.115-118.

13. Насыров Р.И., Тимергалиев С.Н. Перспективы метода нейросетевого прогнозирования надежности накопителей информации в случае ступенчатой нелинейности показателя // Информационные технологии. Автоматизация. Актуализация и решение проблем подготовки высококвалифицированных кадров (ИТАП-2015): тезисы докл. межд. науч.-прак. конф. (Набережные Челны, 17 апреля 2015 г.). – Набережные Челны: Изд-во НЧИ КФУ, 2015. – С.174-179.

14. Насыров И.И., Насыров Р.И. Методика формирования базы знаний интеллектуальной системы для диагностирования систем управления промышленными предприятиями // IV Камские чтения: тезисы докл. межрег. науч.-практ. конф. (Набережные Челны, 27 апреля 2012 г.). В 3-х ч. Часть 3. – Набережные Челны: ИНЭКА, 2012. – С.105-106.
15. Насыров И.И., Насыров Р.И. Методика формирования базы знаний экспертной системы диагностики накопителей информации в системах управления промышленными предприятиями // Измерения, контроль и диагностика – 2012: тезисы докл. II всерос. науч.-техн. конф. студ., асп. и мол. ученых с межд. участием, посв. 60-летию ИжГТУ. (Ижевск, 14-16 мая 2012 г.). – Ижевск: ИжГТУ, 2012. – С.253-256.
16. Насыров И.И., Насыров Р.И. Обоснование необходимости разработки алгоритма прогнозирования надежности накопителей информации в системах управления // V Камские чтения: тезисы докл. всерос. науч.-практ. конф. (Набережные Челны, 26 апреля 2013 г.). В 3-х ч. Часть 1. – Набережные Челны: НЧИ КФУ, 2013. – С.48-50.
17. Насыров Р.И. Алгоритм программы ранжирования накопителей информации по степени опасности при ступенчатой нелинейности показателя надежности // Информационные технологии. Автоматизация. Актуализация и решение проблем подготовки высококвалифицированных кадров (ИТАП-2014): тезисы докл. IV межд. науч.-практ. конф. (Набережные Челны, 28 марта 2014 г.). – Набережные Челны: НЧИ КФУ, 2014. – С.177-181.
18. Насыров Р.И. Целевая функция при разработке программы оценки и прогнозирования надежности накопителей информации в системах управления предприятиями // Научно-технический вестник Поволжья. – 2013. - № 1. – С.226-229.
19. Насыров И.И., Насыров Р.И., Ломаев Г.В., Тимергалиев С.Н. Оптимизация оценки и прогнозирования надежности накопителей информации в системах управления предприятиями // Прикладная информатика. – 2013. - № 6 (48). – С.122-131.

20. Насыров И.И., Насыров Р.И., Ломаев Г.В., Тимергалиев С.Н. Программа ранжирования накопителей информации по степени надежности в компьютерных системах управления предприятиями // Образование и наука Закамья Татарстана. – 2012. - № 25. [Электронный ресурс]. – Режим доступа: <https://cloud.mail.ru/public/MBno/HLVFSqE4t> (дата обращения 09.02.2017).
21. Насыров Р.И. Программа ранжирования накопителей информации по прогнозируемому и фактическому количеству переназначенных секторов // Свид. о гос. рег. программы для ЭВМ № 2015617660 от 16.07.2015.
22. Насыров Р.И. Потребность в программе прогнозирования надежности накопителей информации в компьютерах частных пользователей // Научно-технический вестник Поволжья. – 2014. - № 2. – С.169-172.
23. Насыров И.И., Насыров Р.И., Зиязетдинова Г.У. Оценка потребности в программе нейросетевого прогнозирования надежности накопителей информации // Инновационные машиностроительные технологии, оборудование и материалы - 2016 (МНТК «ИМТОМ-2016»): тезисы докл. межд. науч.-техн. конф. (Казань, 7-9 декабря 2016 г.). – Казань: АО «КНИАТ», Фолиант, 2016. – В 2-х ч. Часть 1. – С.333-336.
24. Насыров И.И. Гибридный метод прогнозирования надежности накопителей информации / И.И. Насыров, Р.И. Насыров, Г.В. Ломаев. – Saarbrücken, Deutschland: Palmarium academic publishing, 2012. – 143 р.
25. Насыров Р.И. Эффективность метода прогнозирования надежности накопителей информации / Р.И. Насыров, И.Н. Насыров. – Saarbrücken, Deutschland: Lambert academic publishing, 2016. – 178 р.
26. Hard drive data and stats [Электронный ресурс]. – Режим доступа: <https://www.backblaze.com/b2/hard-drive-test-data.html> (дата обращения 09.02.2017).

Nasyrov R.I., senior tutor, Naberezhnye Chelny Institute of Kazan (Volga region) Federal University, rinasnyrov@gmail.com;
Nasyrov I.N., doctor of economic sciences, professor, Naberezhnye Chelny Institute of Kazan (Volga region) Federal University, ecoseti@yandex.ru, ResearcherID A-

THE EFFICIENCY OF ALGORITHM FOR PREDICTING THE INFORMATION STORAGES RELIABILITY FOR LARGE DATA-CENTERS

Abstract: In large data centers on the basis of statistical data about refusals it can be concluded on how many drives fail in the near future, but what exactly – it is impossible to know. But for security of information you need to know in advance what drive should fail because security is technically guaranteed by completeness and timeliness of copying information from unreliable on new and reliable drive.

The aim of the study is to increase the efficiency of predicting the reliability of a particular data storage.

An algorithm for predicting the reliability of individual data storage, which identifies those who have the risk of failure, was developed.

In addition with the help of simulation it was obtained that if replacement of faulty equipment we produce by more reliable drives, the number of failures over time will decrease, in relative terms the savings from reducing the number of required storages for replacement will be 5% in two years, 18% over five years and to 35% in ten years.

Keywords: data storage; reliability; efficiency.