



Борис Березін, Дмитро Ланде

**ДОСЛІДЖЕННЯ СТАНУ ОПТИЧНИХ НОСІЇВ ПРИ ДОВГОСТРОКОВОМУ
ЗБЕРІГАННІ ЦИФРОВОЇ ІНФОРМАЦІЇ**

У статті з'ясовано стан наявних проблем, моделі та стандарти у сфері прискореного та природного старіння оптичних дисків, подано результати дослідження природного старіння колекції DVD дисків, підтверджено закономірності та розподіл помилок відповідно до принципів Парето для оптичних носіїв.

Ключові слова: прискорене старіння, природне старіння, CD, DVD, розподіл помилок.

Згідно з дослідженнями компанії IDC, обсяг інформації, записаної на комп'ютерні носії різних видів у світі, зростає із швидкістю 60% на рік, а витрати на зберігання її – на 50%. За період 2009–2011 рр. продаж Blu-ray дисків виріс на понад 150%, а жорстких дисків – майже на 50%. За даними провідних компаній Verbatim, Philips & Lite-On Digital Solutions (PLDS)¹ у найближчій перспективі кількість поставок оптичних носіїв буде зростати й надалі (хоча основна динаміка розвитку очікується в секторі жорстких дисків). Із зростанням долі архівної інформації збільшується її вплив на інфраструктуру зберігання та вибір відповідних проектних рішень. Усе більші вимоги ставляться щодо характеристик оптичних носіїв та можливостей їх довгострокового зберігання.

Усе це підвищує актуальність вивчення характеристик надійності оптичних носіїв. У зв'язку з цим в Інституті проблем реєстрації інформації НАН України досліджуються та розробляються системи оптичної реєстрації інформації, створюються технології довгострокового зберігання цифрової інформації на основі високостабільних матеріалів.

Серед основних напрямків дослідження стану оптичних носіїв можна виділити дві групи моделей: прискореного та природного старіння.

Моделі прискореного старіння. Для тестування оптичних носіїв з метою оцінки строку служби при довгостроковому зберіганні, як правило, використовуються моделі прискореного старіння Ейрінга або Арреніуса².

Модель прискорення Ейрінга:

$$t = A T^{\alpha} e^{\Delta H/kT} e^{(b+C/T) \times RH}.$$

У цій моделі: t – час до відмови; A – часова константа; T^{α} – температурний фактор; ΔH – активаційна енергія молекули; k – константа Больцмана; T – температура по Кельвіну;

B, C – експоненціальні константи; RH – відносна вологість.

Для температурного режиму, що використовується в цьому методі, α та C прирівнюються нулю. Тоді модель Ейрінга спрощується до вигляду:

$$t = A e^{\Delta H/kT} e^{B \times RH}.$$

Модель прискорення Арреніуса використовує тільки температуру для прискорення старіння. Передбачається, що час до відмови може бути отримано за допомогою такого виразу:

$$t = A e^{\Delta H/kT}.$$

Стандарти. Нині найвідомішими міжнародними стандартами, що специфікують методи прискореного тестування терміну служби оптичних носіїв, є стандарт, прийнятий ISO в 2008 р. і оновлений у 2011 р., та стандарт, прийнятий ECMA International в 2010 р.³ Вважається, що стандарт ECMA-396 є еквівалентом стандарту ISO за 2008 р. ISO/IEC 10995:2011 специфікує тестовий метод прискореного старіння для оцінювання очікуваного терміну служби інформації, що зберігається на читаних або перезаписуваних оптичних дисках форматів DVD-R/-RW/-RAM, +R/+RW, може бути використаний для інших форматів оптичних дисків із відповідними змінами в специфікації та може бути оновлений для майбутніх вимог. Стандарт включає умови тестових навантажень (температура та відносна вологість) для моделей Ейрінга та Арреніуса, опис системи оцінки, підготовку режимів, процедуру підготовки даних та інтерпретацію результатів.

Розглянуто стандарт ISO/IEC 10995:2008 та процедура оцінки терміну служби на основі моделі⁴. Отримана максимальна ймовірнісна оцінка розподілення часу служби, основана на прискореній моделі життєвого циклу, моделі Ейрінга, логнормальному розподіленні, з точки зору інженерної надійності. Запропонована відмінна і проста процедура для аналізу набору даних від



тесту прискороеного життєвого циклу для перезаписуваних оптичних дисків.

Аналіз показує, що, незважаючи на актуальність теми надійності оптичних носіїв, серед присвячених їй публікацій недостатньо праць, узагальнюючих способи та результати оцінки надійності для використання їх в процесі створення систем зберігання та управління життєвим циклом зберігання даних.

Використання показників та моделей. Фундаментальні результати щодо показників надійності оптичних дисків наведено в таких роботах⁵. Найбільший обсяг результатів тестування отримано у працях Національного інституту стандартів та технологій та Бібліотеки конгресу США⁶. З використанням моделі Ейрінга було протестовано 15 типів DVD дисків (5 – DVD-R, 5 – DVD+R, 2 – DVD-RW та 3 – DVD+RW). По кожному з 15 типів протестовано по 90 зразків. 7 типів із них мали оціночний термін служби, очікуваний в зовнішніх умовах 25 градусів С та 50% відносної вологості, понад 45 років; 4 типи – 30–45 років; 2 типи 15–30 років; 2 типи – менше 15 років. Крім того, було протестовано 7 типів CD-R дисків. По кожному з 7 типів було протестовано по 90 зразків. 5 із 7 типів протестованих дисків мали оціночний термін служби, очікуваний в зовнішніх умовах, понад 45 років; 2 типи – 30–45 років.

У праці подано результати оціночного тестування очікуваного терміну служби для CD-R дисків, який при використанні моделі Ейрінга склав 15–30 років роботи⁷. Також подано результати у вигляді графіків залежності показників стану оптичних дисків від часу при підвищених значеннях температури та відносної вологості⁸. Для CD-R дисків, зразки яких відрізнялись типом покриття та фарби, вимірювались показники BLER та Jitter. Для DVD-дисків, які відрізнялись виробником, вимірювались показники PI Sum 8, PO error та Jitter.

Значно менше робіт присвячено темі природного старіння оптичних дисків. Основне дослідження цього напрямку було пов'язано з вибіркоким аналізом стану 125 CD дисків (1976–1996 рр.) із складу аудіоколекції з 60 тисяч CD дисків Бібліотеки Конгресу США⁹. Дослідження присвячено виявленню параметрів для моніторингу CD, які сигналізують про необхідність створення резервних копій або міграції даних, а також ідентифікації механізмів пошкоджень, які є наслідком недоліків виробництва.

Для дослідження природного старіння оптичних дисків в 1996 р. проведено базове тестування

параметрів CD (Baseline), а у 1999 та 2003 рр. – повторні тестування (Run 2 та Run 3) параметрів усіх 125 CD. Вивчалось значення логарифму параметру BLER (Block Error Rate – кількість блоків даних із помилками, які можна виправити декодером першого рівня), подане на рис. 1. Максимальне допустиме значення BLER – 220 помилок на секунду.

На рис. 2 подано середнє значення показника помилок BLER для усіх 125 дисків, розраховане на основі кожного з трьох тестувань Baseline, Run 2 та Run 3. Середнє значення параметру Bler з 1996 до 2003 рр. зросло з 70,5 до 74,5. На деяких із дисків невіправні помилки почали виявлятися протягом менше 10-ти років із моменту випуску.

Результати іншого тестування вибіркового масиву документів, проведеного в 2011 р., подано на сайті Федерального архівного агентства (Росархіву). Для оцінки CD дисків використано п'ять рівнів оцінки: від відмінної до дуже поганої якості залежно від значень низки показників, що визначаються при тестуванні.

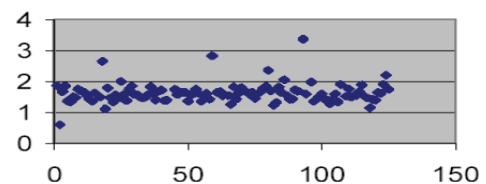


Рис. 1. Значення логарифму показника помилок BLER для кожного номера CD у колекції в 2003 р.

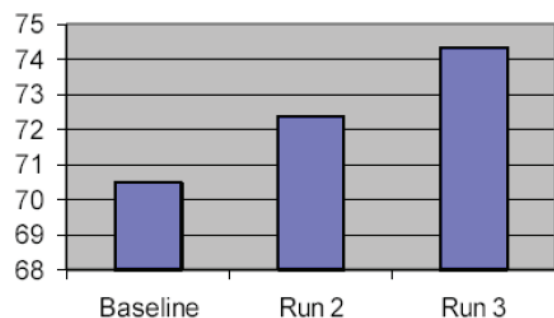


Рис. 2. Середнє значення показника помилок BLER для усіх 125 дисків, розраховане для кожного з трьох тестувань (1996 р., 1999 р., 2003 р. відповідно).

На основі результатів тестування на сайті Росархіву зроблено рекомендації щодо використання дисків від різних виробників. Інформацію з носіїв, якість яких оцінюється нижче ніж задовільна, треба перенести на нові носії. Такі носії не відповідають вимогам довгострокового зберігання і повинні бути замінені на носії від



іншого виробника. Результати тестування дають змогу побудувати діаграми для виявлення особливостей впливу старіння дисків на якість інформації, що зберігається. За результатами тестування колекції майже 400 CD дисків Російського державного архіву науково-технічної документації (РДАНТД), записаних із 2000 до 2009 рр. (тестування проводилося у 2010–2011 рр.), побудовано

діаграму (рис. 3). На основі результатів тестування колекції майже 70 CD дисків Російського державного архіву літератури і мистецтва (РДАЛМ), записаних у 2005–2006 рр. (тестування проводилося у 2011 р.), побудовано діаграму (рис. 4). Подані на них дані підтверджують наведені рекомендації.

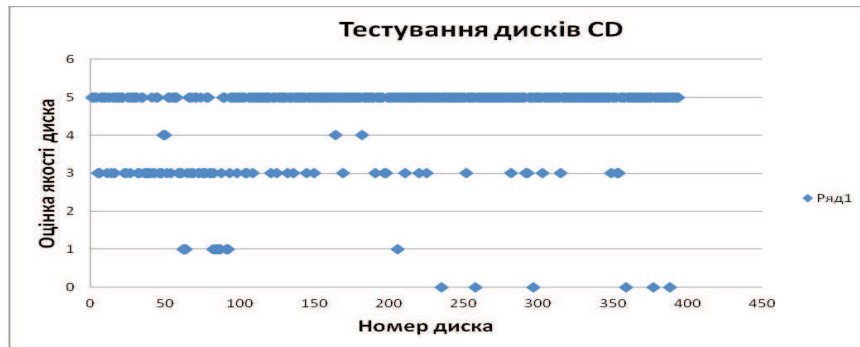


Рис. 3. Діаграма оцінок якості стану носіїв залежно від номера для 400 CD колекції РДАНТД (покращені диски).

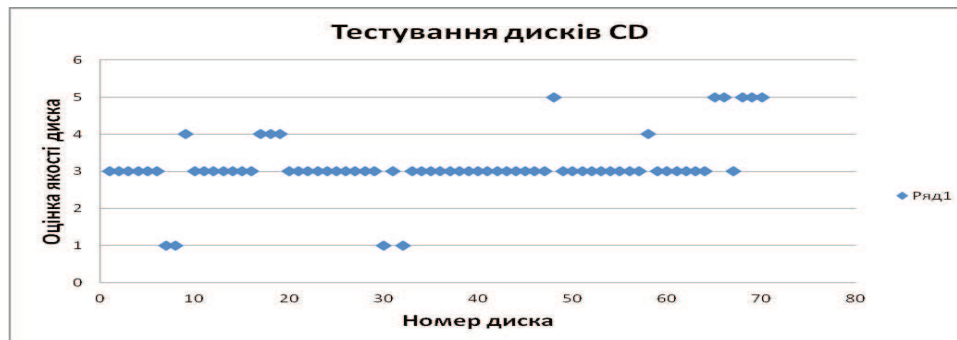


Рис. 4. Діаграма оцінок якості стану носіїв залежно від номера для 70 CD колекції РДАЛМ (звичайні диски).

Проведений аналіз досліджень із прискореного та природного старіння оптичних носіїв показав обмеженість даних про характеристики природного старіння CD та особливо DVD, BD, UDO дисків.

З метою дослідження характеристик природного старіння DVD дисків, порівняння їх із даними попередніх досліджень та визначення можливостей подальшого використання для управління інфраструктурою довгострокового зберігання інформації було створено базу даних колекції DVD дисків. По кожному диску база вміщує його номер, час запису інформації на нього, об'єм інформації, тип носія, ідентифікатор виробника диску, час тестування, значення по-

казника помилок, оцінку зовнішнього стану диску тощо. В 2012 р. проведено тестування вибіркового масиву приблизно з 150 носіїв із колекції DVD дисків, записаних у 2006–2012 рр. Для вимірювання щільності помилок використовується значення PIE (Parity Inner Error) – кількість рядків парності блоку ECC із помилками (Error Correction Code – код корегування помилок), а точніше PI Sum 8 – значення для 8 послідовних ECC із блоку. Максимальне допустиме значення PI Sum 8 складає 280 помилок. На рис. 5. подано розподіл значення показника помилок PI Sum 8 залежно від номера кожного з цих дисків, що аналізувалися авторами.



Рис. 5. Значення показника помилок PI Sum 8 (вісь ординат) для кожного (вісь абсцис) з 150 DVD дисків колекції.

Для порівняння розподілу значення показника помилок PI Sum 8 залежно від часу запису інформації на диск, на рис. 6 подано відповідні гістограми для двох вибірок по 28 дисків, записаних у 2010 р. та 2007–2008 рр. відповідно. По осі абсцис гістограми відкладено інтервали значень показника PI Sum 8, а по осі ординат – відповідні кількості дисків із вибірок 2010 р. та 2007–2008 рр. Для порівняння стану носіїв колекції, записаних у 2010 р. та 2007–2008 рр. для представлених вище рядів розподілу розраховані середні величини, які складають 87

для вибірки 2010 р. та 90,2 для вибірки з колекції 2007–2008 рр. відповідно. (Середні значення показника помилок BLER, розраховані для вибірки із 125 дисків у 1996 р., 1999 р. та у 2003 р., в дослідженні колекції Бібліотеки Конгресу США складають відповідно 70,5; 72,5 та 74,5¹⁰). Отримані значення та гістограма показують, що середнє значення показника PI Sum 8 для носіїв нової колекції знаходиться далі від граничного значення цього показника для DVD-дисків, визначеного на рівні 280¹¹, ніж середнє значення для старої колекції.



Рис. 6. Значення показника помилок PI Sum 8, отримані для двох вибірок по 28 дисків за 2010 та 2007–2008 рр.

Для виявлення особливостей розподілу характеристик DVD-дисків при природному старінні дані про вибірку з 150 носіїв було проранжировано за кількістю помилок. Отриманий розподіл подано на рис. 7. Він може бути апроксимований

за допомогою степеневої функції (Power Law) із степеневим показником – 0,827 та достовірністю апроксимації 0,87.

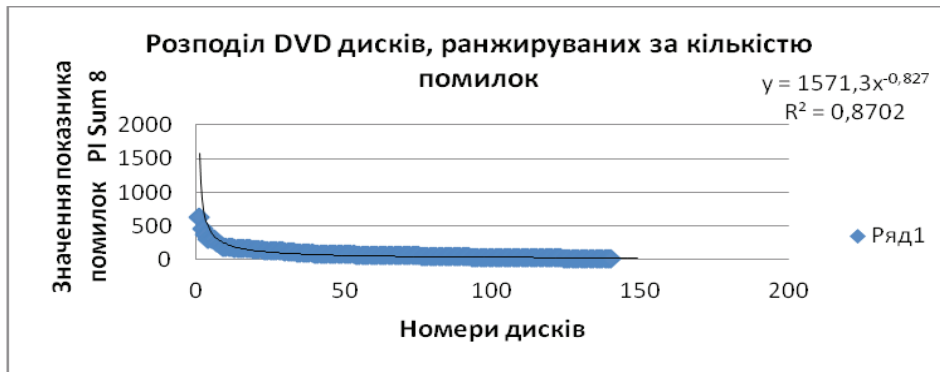


Рис. 7. Дані про вибірку з 150 DVD дисків, ранжировані за кількістю помилок з апроксимацією степеневою функцією.

Для порівняння характеристик CD та DVD дисків при природному старінні аналогічне ранжирування було виконано для частини даних, розрахованих на основі результатів дослідження колекції CD-дисків Бібліотеки Конгресу США в 1999 р.¹² Отриманий розподіл подано на рис. 8. Він теж може бути апроксимований за допомогою степеневої функції із степеневим коефіцієнтом – 0,724 та достовірністю апроксимації 0,91.

Із загальної вибірки 150 DVD дисків було виділено 33 носія, виготовлених одним виробником. Розподіл цих DVD дисків, ранжированих за кількістю помилок, подано на рис. 9. Він може бути апроксимований за допомогою степеневої функції із степеневим коефіцієнтом – 0,964 та достовірністю апроксимації 0,91.

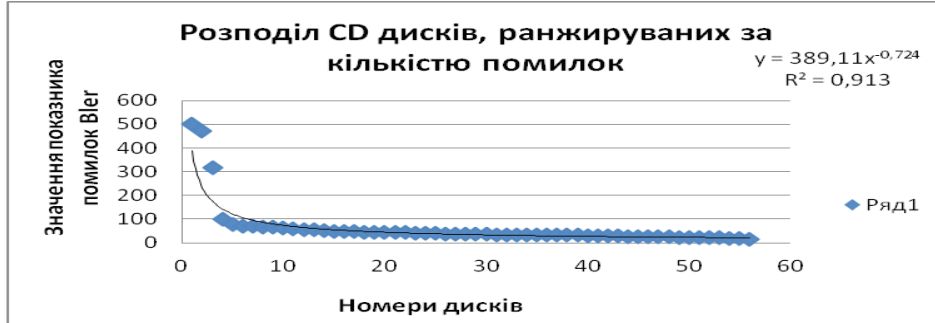


Рис. 8. Ранжировані дані про частину вибірки з 125 CD дисків (дані розраховано на основі результатів дослідження¹³) з апроксимацією степеневою функцією.



Рис. 9. Дані про вибірку DVD дисків від одного виробника, ранжировані за кількістю помилок з апроксимацією степеневою функцією.



Отримані значення показників стану носіїв для виборки колекції цифрової інформації, записаної в 2006–2012 рр., дають змогу порівнювати стан DVD-дисків колекції та вибирати необхідні моменти міграції файлів колекції на нові носії.

Висока ступінь достовірності при апроксимації отриманих розподілів степеневою функцією підтверджує відповідність старіння CD та DVD дисків закономірностям та розподілу по-

милок, що визначаються універсальною закономірністю Парето, завдяки якій можна зробити оцінку живучості інформаційних об'єктів, що розміщуються на цих носіях. Відомо, що живучість інформаційного об'єкта оцінюється як ймовірність того, що об'єкт буде непошкодженим протягом визначеного періоду часу t при визначених умовах¹⁴, тобто:

$$S(t) = \Pr \{T > t\} = 1 - \Pr \{0 < T \leq t\} = 1 - \int_0^t f(x)dx,$$

де $S(t)$ – значення живучості, $f(x)$ – функція щільності ймовірності руйнування інформаційного об'єкта.

Якщо інформаційний об'єкт зберігається частинами на n носіях інформації, то ймовірність руйнування його оцінюється як:

$$F_{lost}(t) = \prod_{i=1}^n F_i(t).$$

У цьому добутку $F_i(t)$ – ймовірності руйнування i -го носія за час t .

Відповідно живучість оцінюється як:

$$S_n(t) = 1 - F_{lost}(t) = 1 - \prod_{i=1}^n F_i(t).$$

Маючи на увазі, що ймовірність виникнення помилок на носіях пропорційна часу їхнього існування, що доведено даними вимірів, і те, що розподіл помилок має степеневий розподіл, можна вважати доцільним і обґрунтованим дослідження моделі із степеневим розподілом помилок, що принципово відрізняється від підходів, в яких використовується пуасонівський потік помилок (теорія систем масового обслуговування) та розподіл помилок за Вейбулом¹⁵. У цьому випадку живучість можна оцінювати як:

$$S_n(t) = 1 - \prod_{i=1}^n F_i(t) = 1 - \prod_{i=1}^n Ct^{-\beta} = 1 - C^n t^{-n\beta},$$

де C, β – деякі константи.

Виявлені закономірності статистичного розподілу помилок дають підстави зробити висновки, пов'язані з живучістю інформаційних об'єктів, що розміщуються на оптичних носіях даних, а саме: враховувати явища самоподібності, нерегулярності виникнення помилок, наявність «товстого хвосту» в розподілі, що характеризує надзвичайно велику кількість носіїв із незначною кількістю помилок тощо.

Подані вище залежності для оптичних дисків разом із відповідними характеристиками інших видів носіїв можна використати при побудові інструментальних засобів управління інфраструктурою довгострокового зберігання інформації для підвищення її ефективності та живучості.

¹ Будущее оптических накопителей глазами PLDS / интервью Бойса Лина // КомпьютерПресс. – 2010. – № 1. – С. 42–43.

² Standard ECMA-396. Test Method for the Estimation of Lifetime of Optical Media for Long-term Data Storage, 2010. – 44 p.; ISO/IEC 10995:2011. Information technology – Digitally recorded media for information interchange and storage – Test method for the estimation of the archival lifetime of optical media; ISO/IEC 10995:2008. Information Technology – Digitally recorded media for information interchange and storage – Test method for the estimation of the archival lifetime of optical media; Yamamoto W., Kumazaki C., Suzuki K. Estimation of Archival Lifetime Distribution for Writable Optical Disks from Accelerated Testings // Mathematical and Statistical Models and Methods in Reliability: Applications to Medicine, Finance, and Quality Control, 2010. – P. 307–313. – (Statistics for Industry and Technology); Irie M., Okino Y., Kubo T. Test method for evaluating optical disk reliability // Proceedings of the SPIE. – 2005. – Vol. 5643. – P. 205–210.

³ Standard ECMA-396. Test Method for the Estimation of Lifetime of Optical Media for Long-term Data Storage, 2010. – 44 p.; ISO/IEC 10995:2011. Information technology – Digitally recorded media for information interchange and storage – Test method for the estimation of the archival



lifetime of optical media; ISO/IEC 10995:2008. Information Technology – Digitally recorded media for information interchange and storage – Test method for the estimation of the archival lifetime of optical media.

⁴ Yamamoto W., Kumazaki C., Suzuki K. Estimation of Archival Lifetime Distribution for Writable Optical Disks... P. 307–313.

⁵ Righi M. CD and DVD preservation issues // AXMEDIS 2008 (Firenze, 17–19 novembre 2008): Proceedings / Paolo Nesi, Kia Ng, Jaime Delgado (eds.). – Firenze: Firenze University Press, 2008. – Vol 1. – P. 46–50; Final Report: NIST / Library of Congress (LC) Optical Disc Longevity Study. September 2007. – 32 p.; Irie M., Okino Y., Kubo T. Test method for evaluating optical disk reliability. – P. 205–210; Shahani C. J., Manns B., Youket M. Longevity of CD Media Research at the Library of Congress / Preservation Research and Testing Division Library of Congress. – Washington DC, USA, 2005. – 14 p.; Slattery O., Lu R. Stability Comparison of Recordable Optical Discs – A Study of Error Rates // Harsh Conditions Journal of Research of the National Institute of Standards and Technology. – 2004. – Vol. 109, № 5. – P. 517–524.

⁶ Final Report. – 32 p.

⁷ Irie M., Okino Y., Kubo T. Test method for evaluating optical disk reliability. – P. 205–210.

⁸ Slattery O., Lu R. Stability Comparison of Recordable Optical Discs – A Study of Error Rates. – P. 517–524.

⁹ Righi M. CD and DVD preservation issues. – P. 46–50; Shahani C. J., Manns B., Youket M. Longevity of CD Media Research at the Library of Congress. – 14 p.

¹⁰ Shahani C. J., Manns B., Youket M. Longevity of CD Media Research at the Library of Congress. – 14 p.

¹¹ Standard ECMA-396. Test Method for the Estimation of Lifetime of Optical Media for Long-term Data Storage, 2010. – 44 p.

¹² Shahani C. J., Manns B., Youket M. Longevity of CD Media Research at the Library of Congress. – 14 p.

¹³ Ibid.

¹⁴ Li Y., Miller E. L., Long D. D. E. Understanding Data Survivability in Archival Storage Systems // Proceedings of the 5th Annual International Systems and Storage Conference (SYSTOR 2012), June 4–6, 2012, Haifa, Israel.

¹⁵ Ibid.

В статті вивчено стан існуючих проблем, моделі і стандарти в області прискореного і природного старіння оптичних дисків, наведені результати дослідження природного старіння колекції DVD дисків, підтверджені закономірності і розподілення помилок в відповідності з принципами Парето для оптичних носіїв.

Ключові слова: прискорене старіння, природне старіння, CD, DVD, розподілення помилок.

The condition of the existing problems, models and standards in the field of accelerated and natural aging of the optical drive are presented. The results of the study of natural aging of DVD discs collection, confirmed that patterns and distribution of errors are in accordance with the principles of Pareto for optical media.

Key words: accelerated aging, natural aging, CD, DVD, error distribution.

УДК 339.17:[005.92:004.63]

Олена Дроздова

ЕЛЕКТРОННА ДОКУМЕНТАЦІЯ ОБЛІКУ ДІЯЛЬНОСТІ ТОРГОВЕЛЬНИХ ПІДПРИЄМСТВ: ОРГАНІЗАЦІЙНИЙ АСПЕКТ

У статті розглянуто питання документального забезпечення обліку діяльності торговельних підприємств та торговельного процесу в паперовому й в електронному вигляді на сучасному етапі розвитку України.

Ключові слова: документ, електронна документація, електронний підпис, електронний договір, торговельне підприємство, первинні документи.

Різноманітність форм власності в Україні в умовах ринкової економіки та розширення прав торговельних підприємств вимагають альтернативних підходів при вирішенні конкретних питань методики й техніки обліку торговельної діяльності за допомогою електронних документів.

Торговельне підприємство – первинна, основна ланка сфери торгівлі, її самостійний господарюючий суб'єкт (із правом юридичної особи), створений для закупівлі, реалізації, а також

зберігання товарів, надання різного роду супутніх послуг з метою задоволення потреб ринку та отримання прибутку. Основні ознаки підприємства: організаційна єдність, відокремлене майно, майнова відповідальність, власне ім'я¹.

З організаційно-економічної точки зору підприємство – це сукупність матеріальних, трудових та фінансових засобів. При господарюванні кожне торговельне підприємство здійснює:

© Олена Дроздова, 2012