Improved Disk Drive Failure Warnings

G. F. Hughes, *Fellow*, J. F. Murray, K. Kreutz-Delgado *Senior Member*, and C. Elkan

Abstract— Improved methods are proposed for disk drive failure prediction. The SMART (Self Monitoring and Reporting Technology) failure prediction system is currently implemented in disk drives. Its purpose is to predict the near-term failure of an individual hard disk drive, and issue a backup warning to prevent data loss. Two experimentally tests of SMART showed only moderate accuracy at low false alarm rates. (A rate of 0.2% of total drives per year implies that 20% of drive returns would be good drives, relative to ≈1% annual failure rate of drives). This requirement for very low false alarm rates is well known in medical diagnostic tests for rare diseases, and methodology used there suggests ways to improve SMART.

Two improved SMART algorithms are proposed here. They use the SMART internal drive attribute measurements in present drives. The present warning algorithm based on maximum error thresholds is replaced by distribution-free statistical hypothesis tests. These improved algorithms are computationally simple enough to be implemented in drive microprocessor firmware code. They require only integer sort operations to put several hundred attribute values in rank order. Some tens of these ranks are added up and the SMART warning is issued if the sum exceeds a prestored limit.

These new algorithms were tested on 3744 drives of two models from one manufacturer. They gave 3-4 times higher correct prediction accuracy than error thresholds on will-fail drives, at 0.2% false alarm rate. The highest accuracies achievable are modest (40%-60%). Care was taken to test "will-fail" drive prediction accuracy on data independent of the algorithm design data.

Additional work is needed in order to verify and apply these algorithms in actual drive design. They may also be useful in drive failure analysis engineering. It may be possible to screen drives in manufacturing using SMART attributes. Marginal drives might be detected before substantial final test time is invested in them, thereby decreasing manufacturing cost and possibly decreasing overall field failure rates.

***Index terms*—disk drive, failure prediction, predictive failure analysis, SMART, magnetic recording**

Manuscript received July 10 2000, revised June 22, 2001. This work was supported by the UCSD Information Storage Industry Center, funded by the Alfred P. Sloan Foundation. The authors are with the University of California San Diego, La Jolla CA, (email: gfhughes@ucsd.edu, jfmurray@ucsd.edu, kreutz@ece.ucsd.edu, elkan@cs.ucsd.edu)

ACRONYMS

ATA	Standard drive interface, desktop computers
FA	Failure analysis of apparently failed drive
FAR	False alarm rate, 100 times probability value
MVRS	Multivariate rank sum statistical test
NPF	Drive failed, but "No problem found" in FA
RS	Rank sum statistical hypothesis test
R	Sum of ranks of warning set data
R_c	Predict fail if R> R_c critical value
SCSI	Standard drive interface, high-end computers
SMART	"Self monitoring and reporting technology"
WA	Failure warning accuracy (probability)

NOTATION:

n	Number of reference (old) measurements
m	Number of warning (new) measurements
N	Total ranked measurements (n+m)
p	Number of different attributes measured
Q(X)	Normal probability Pr(x>X)
RS	Rank sum statistical hypothesis test
R	Sum of ranks of warning set data
R_c	Predict fail if R> R_c critical value

I. INTRODUCTION

COMPUTER disk drives are reliable data storage devices with annual failure rates of 0.3% to 3% per year [1,2]. (A 1% nominal failure rate will be used for comparisons here.) Nonetheless, drive failure can cause a catastrophic loss of user data. This is often far more serious than the hardware cost of replacing the failed drive. If impending drive failure could be predicted a warning could be issued to the drive user to back up the data onto another storage device.

In 1995, the drive industry adopted a standardized specification for such failure warnings, called "SMART" (see definitions, Section IV). SMART is based on monitoring a number of internal drive technology measurements relevant to impending failure. A failure warning algorithm is run by the drive microprocessor firmware. This checks whether the measurements exceed maximum thresholds and produces a binary (won't-fail/will-fail) warning. The SMART warning time goal is 24 hours before drive failure.

Computer operating systems can issue standardized drive commands to enable and to read this failure warning. These commands are defined for the two predominant computer to drive interface standards, ATA and SCSI [3].

Additionally, the SCSI "Enclosure Services" specification allows RAID array controllers to be notified if thresholds are exceeded on drive *external* environments such as power supply voltage, current, and ambient temperature [3].

SMART technology is implemented in most 3.5-inch disk drives manufactured today, the most widest used disk drives from personal computers to supercomputers. However, it is unknown how many computer systems today enable or read the SMART warning. In some personal computers, SMART is checked on computer bootup by the CMOS/BIOS firmware. Drive manufacturers supply diagnostic programs that read the SMART warning. Information on SMART warning accuracy is anecdotal at best, and much of the drive internal monitoring technology is manufacturer proprietary.

This paper assesses the accuracy of the existing "SMART" failure warning algorithm in drives, and an improved algorithm. Experimental data is from drive design reliability testing of two different Quantum Corporation disk drive models. Tradeoff curves of the will-fail-drive correct warning accuracy "WA" are calculated, vs. the false alarm rate "FAR" (defined as the probability that a fail warning will occur in a drive that doesn't subsequently fail).

II. Failure Warning Technology

A. Background

Failure warning markedly differs from normal disk drive reliability methodology. The latter statistically predicts failure probability over an entire drive population, and assumes that all drives are equally likely to fail [2,4,5]. SMART predicts *individual* drive failure.

Failure warning technologies such as condition monitoring and predictive maintenance are also used, in process control and large motor monitoring [6,7,8].

B. "SMART" disk drive failure warning

The SMART ATA drive specification [3] allows up to 30 internal drive measurements. These are termed failure attributes and are periodically measured by a drive. Attribute values are stored in the drive reserved data area with other drive operational parameters. For a drive user to receive a SMART warning the computer system must issue specific drive interface commands to enable the algorithm and then to read the resultant "won't-fail/will-fail" warning [9]. Some drives will unilaterally shut down if internal sensors detect extreme temperature or mechanical g-shock [10].

Maximum thresholds are defined for each attribute by the drive manufacturer. The SMART warning flag is set in response to an ATA SMART "Return Status" command, if any attribute exceeds its threshold. This is a logical 'OR' operation among the several attribute threshold tests, and is used because some drive failures may be predicted by only one attribute. But this 'OR' operation can also cause a high false alarm rate, since it does not require multiple confirming attribute "signatures" to trigger the warning.

Table I lists SMART attributes, starting with basic nearly universal attributes, to proposed future attributes. The basic attributes exploit existing drive internal technology (thus allowing minimal added cost). Many were historically adopted for drive error recovery and for reliability analysis, with SMART warning thresholds added later. Most attributes are *incremental* error counts over a fixed time interval. For example, certain rates of seek and read soft errors are allowed by drive designers, and if the incremental counts of these errors remains stable, then failure is not indicated. Cumulative counting would mislead.

Power on hours (POH) is a traditional measure of drive age. Low POH may imply infant mortality failure risk and high POH may imply end of life failure risk. But for failure warning, both need corroboration by other attributes. A related attribute is contact start-stops (CSS) which is a count of drive power cycles; i.e., power on, disks spin up, heads fly, power off, heads contact disk while spinning to stop. High CSS increases the risk of head/disk sliding contact wear. (Some drives avoid head-disk contact and have no CSS attribute). These attributes are *cumulative*.

Seek errors (SKE) is an incremental count of track seeks that require a second re-seek to find the intended track. The count is reset to zero after a fixed number of thousands of seek commands. If a re-seek also fails a recalibrate retry (RRT) reinitializes the head tracking servo system, and is counted in a separate RRT attribute.

A read soft error (RSE) is a data read error detected and corrected by an error correction code. It can indicate disk defects, high head fly, or head off-track. Repeated RSE errors at the same user data disk location can invoke drive error recovery which moves the user data to a new location and records a grown defect (GDC) count. Read channel parameters such as the Viterbi detector mean-square error (MSE) can warn of an approaching GDC before an RSE occurs [11].

TABLE I: SMART ATTRIBUTES

POH	Drive aging	C
CSS	Drive power cycles	C, F
SKE	Heads seek to wrong track	I,T
RRT	Drive re-initializes	I
RSE	Errors corrected by inner ECC code	I, F T
MSE	Precursor of RSE	I
GDC	New disk defects, found after manufacture	C, F T
SUT	Power on to drive ready	I, F
TMR	Head-track misregistry	I,T
GMX	Mechanical shock	I
TAS	Thermal asperity count	I,F
FLY	By PW50, or Wallace, or read IC MSE, FIR taps	I
TMP	Drive internal limit	I

DCL	Read IC ok	I
-	Acoustics, start current, or control loop analysis	I,T
POH	Drive aging	C

Keys: C (cumulative measurement), I (incremental), F(head/disk interface indicative), T(track servo indicative)

Disk spinup time (SUT) is the elapsed time from power on to drive ready for data transfer. Increasing SUT may indicate head-disk stiction, raising the risk that the drive spin motor may be close to its maximum starting torque limit. Disk spin motor current and spin servo parameters can detect late head fly takeoff, bearing damage or runout.

TMR monitors the track servo misregistry error signal [12], which can indicate mechanical G-shock, head mechanical resonance faults, or spindle bearing runout. It is also used to inhibit writing, to eliminate the risk of corrupting data on an adjacent track.

Head/disk fly height (FLY) can be measured using magnetic recording physics, such as playback pulse width PW50, or the pulse peak amplitude normalized by pulse area, or read channel equalization parameters [11]. The Wallace spacing loss formula can also be used [13]. One head flying significantly high (referenced to the average of other heads in the drive) indicates a risk of poor writing or reading, and a low fly head increases head-disk wear risk.

Internal drive temperature (TMP) is measured by some manufacturers using a dedicated thermal sensor in the drive, or from read preamplifier outputs which indicate the magnetoresistive read sensor resistance, and hence its temperature. High temperature stresses the electromechanics of the drive, and low temperature can allow moisture condensation on the disks, leading to head stiction. G-shock can be monitored (GMX) by a G-sensor MEMS IC.

C. *The failure rarity problem*

Tests of SMART failure warning algorithms can be made using experimental data sets of periodic attribute reads over the life of drives. Times when drives appear to fail are noted and failure verified by physical failure analysis. Because drive failure rates are only about 1% per year, thousands of drives must be tested for more than a year to get statistically significant numbers (>15) of failed drives. This is larger than the number of drives in most large RAID arrays, even in many supercomputers.

Controlling the false alarm rate places the most critical demand on SMART warning algorithms. A seemingly small false alarm rate of 1% per year would double the total number of drives returned for failure, because this rate is about equal to actual annual failure rates. This requirement for very low false alarm rates is well-known in medical diagnostic epidemiology tests for rare diseases [14].

One good source of experimental data (used here) is testing new drive designs. Typically, several thousand drives of a new design are tested by a drive manufacturer to expose latent design and reliability problems. So more failures are expected than in production drives. The testing includes drives with experimental components or built under experimental conditions that are not used in full-scale mass production. Consequently, significant numbers and types of failures are likely to occur. This has the advantage of producing more failures for statistical SMART test development. However, caution is necessary to guard against failure modes caused by test conditions rather than inherent drive technology. It is felt that the test data used here are valid for SMART analysis because the failure types are representative and typical of field failures.

Failure analysis ("FA") is performed to verify failed drives and determine failure causes for corrective redesign. Typically 20%-30% of apparently failed drives are no-problem-found "NPF" drives, which operate normally when analyzed. Therefore, FA is important for a valid data set, and is also highly effective in gathering definitive failure data and statistics, which can guide attribute performance and selection.

These NPF rates also imply that disk drives have a false alarm rate of 0.2%-0.3% even in the absence of SMART (20%-30% of the 1% annual "perceived failure" rate).

Another possibility (not tried here) is to mathematically characterize the experimental data sets in order to generate *simulated* attribute data using Monte Carlo methods.

D. *Attribute data characteristics*

In addition to the SMART warning flag, the original ATA SMART specifications [3] define a 512-byte SMART data record format. This allows the drive internal SMART attributes to be read out, as 1-12 byte integers (raw, unnormalized attribute data are used here, see [3]).

Figure 1 shows histograms of GDC, SKE, and RSE attribute data from one of the two drive design tests. Data on top is from won't-fail drives, and data below from will-fail drives (which subsequently failed during the test). These histograms are *all* the attribute data from *all* the drives of one model, to illustrate the nature of the attribute data. For example, there were about 55,000 occurrences of zero grown defects among all the attribute data reads taken from all drives that did not fail, and 75 zero GDC reads from drives that did ultimately fail during the test.

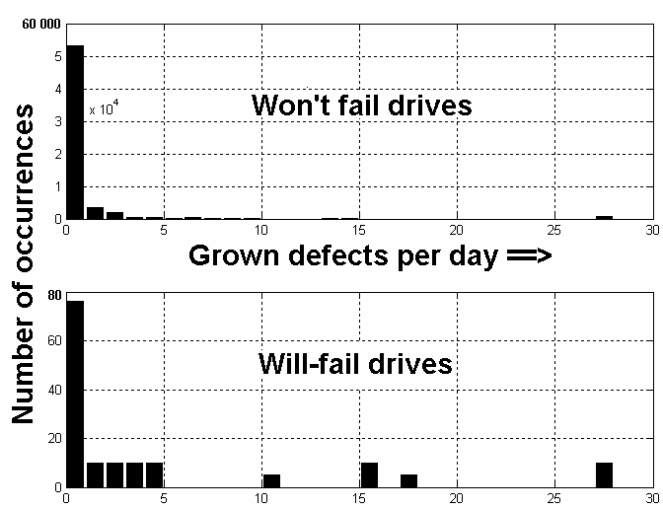

Fig. 1a. Grown-defect data histograms, all won't-fail drives (top) vs. all will-fail drives (bottom)

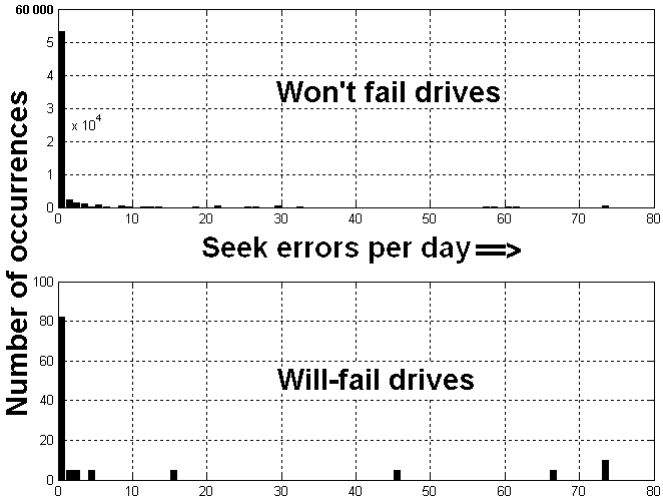

Fig. 1b. Seek Errors histograms, all won't-fail drives (top) vs. all will-fail drives (bottom)

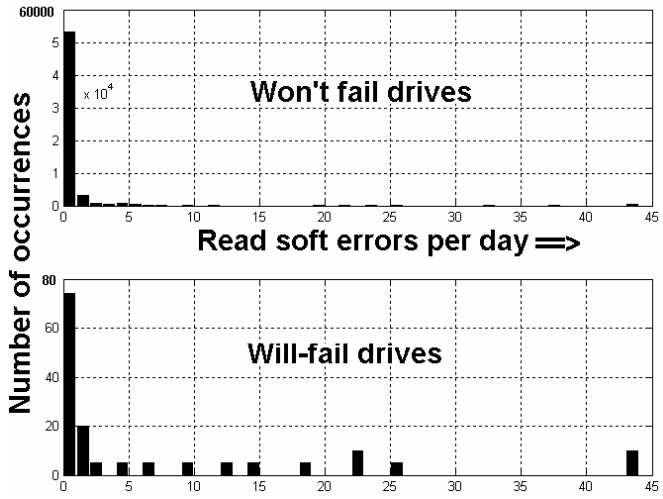

Fig. 1c. Read Soft Errors histograms, all won't-fail drives (top) vs. all will-fail drives (bottom)

The characteristic distinguishing the will-fail drives from the won't-fail drives is a *pattern* of high attribute values during a warning measurement interval. A single high value could be a statistical or transitory accident, but a "scatter dominance" pattern appears significant in Fig. 1. Precise attribute values are not as important as a pattern of scatter.

In statistical terms, this is "ordinal" data, in that increasing attribute values imply increased failure risk, but a doubled attribute value does not necessarily double the risk.

Although there appears to be randomness in these histograms, they certainly do not resemble continuous parametric distributions, such as Gaussian, Poisson or Weibull. Because drive failure is caused by significant physical changes, the only relationship between will-fail and won't-fail drive data may be simply that they significantly differ from each other, more than would be expected from statistical "noise."

E. Special data factors in drive design tests.

Low POH and CSS values are generally indicative of infant mortality risk, and high values suggest increased end-of-life failure potential. But these are not conclusive without other confirming attributes. Additionally, in drive design tests of fixed duration, these attributes can appear 100% predictive but misleading. (All won't-fail drives get the full test duration POH and CSS, will-fail drives obviously get less.)

Drive design testing is done to expose latent problems, in order to eliminate them by redesign. The drives being tested have higher failure rate than mass production drives (see Sect. II-C). This is similar to medical research conditions that accelerate the disease under study [14].

Using the same drive test data to both select a SMART algorithm and to test its accuracy can be misleading. Here, the algorithm parameters are selected to give an acceptable FAR on won't-fail drives, and then tested for WA on a different data set, namely the will-fail drives. In addition the same algorithm is tested on two independent drive models. The FAR on production drive data should be lower, since the test purpose is to remove drive design failure modes.

F. Rank sum statistical SMART tests

SMART algorithms can be regarded as statistical hypothesis tests. They use SMART data to test the hypothesis that a drive will fail against the null hypothesis that a drive is remaining stable and will not fail. The existing SMART threshold algorithm uses only the most recent attribute values, and issues a failure warning if any attribute is above its critical failure threshold. This is a logical "OR' of independent tests on each attribute.

Wilcoxon rank sum [15] statistical tests are proposed here to replace the threshold tests, to improve failure warning accuracy and lower false alarm rates.

Rank sum tests are widely recommended for rare failure situations (such as rare disease epidemiology) where false alarms are costly. They are particularly useful when the

statistical distributions are unknown and suspected of being non-gaussian [14,16].

For drive failure warning, an appropriate hypothesis test is to use a "warning" data set of recent attribute values and compare it to an original "reference" data set taken from the drive population during manufacture. If the two data sets vary only in probable statistical "noise" the null hypothesis is selected. Namely, the drive is stable and no warning is issued. Figure 1 illustrates the general idea, with the upper histogram of each attribute representing the reference set, and the lower the warning set. (But the warning histogram would be data from an *individual* drive, not all the drives as in Figs. 1)

The *warning* data set for each drive is taken to be its last five samples of each attribute (the most-recent 5 days of data for these drive models).

The *reference* data set for each attribute is taken to be 50 random samples of that attribute taken from initial SMART reads, averaged over many good drives. The optimum data set sizes will vary depending on factors such as the SMART attribute read interval. These data set sizes gave the best WA and FAR results for this test data.

The best reference sets were using the first few attribute values from the (several thousand) good drives. They were randomly divided into 50 groups, and each attribute for the group taken as the (single) average of 50 values (rounded to an integer). Will-fail drives are not included in the reference set averaging. They can be kept as independent data for predictive accuracy testing. (Even if they were included in a production drive situation, a ≈1% FAR implies that the averaging should wash out their influence.)

So if 50 typical measurements from new drives have one or two seek errors (SKE), an example SKE reference data set might consist of 48 zeros, a single "1" count, and a "2" count (in any order).

The warning and reference data sets might look like Fig. 1c, with many ties at the lowest rank (zero seek errors), and the maximum rank being one instance of 43 seek errors in this one-day time interval. (Fig. 1c is actually *all* drives.)

The rank sum test for a given drive is computed numerically by a sorting operation on the *combined* warning and reference data sets for each attribute. The reference data never changes for a given drive, but its warning data does since it is the last 5 samples before the SMART warning test is to be made.

Rank "1" is given to the smallest combined attribute value, rank "2" to the next, and so on. The rank sum statistic is just the sum of the ranks of the *warning* attribute values, among *all* the attribute values in the warning and reference data sets. If the drive is stable, the warning data ranks should intersperse randomly among the reference data ranks, since the two data sets have the same statistical distribution (which does not have to be known). If the rank sum is higher than a fixed limit (precalculated and stored in the drive), the test concludes that the two data sets have distinct statistical differences, within a specified false alarm rate (FAR). This implies that the drive attributes have statistically changed since manufacture, indicating potential failure.

G. *Numeric example of rank sum test*

Sect. V outlines the mathematics and derivation of the rank sum test. A numeric example may best demonstrate how it operates.

Consider the made-up seek error data in Table II. Each reference datum in column 2 is a seek error count over a specified SMART frame interval, randomly taken from good drives. There are 8 occurrences of 1 seek error, 3 of 2 errors, and 1 of 4 errors. This reference data never changes.

The latest warning data error counts from one drive are in column 3. They show a pattern of higher counts, qualitatively suggesting that this drive is now making more seek errors and may fail.

The total ranks of column 2 *and* 3 data are shown in column 4. When ties occur at any error count all the tied data are given the average rank of the ties. (Sect. V-E discusses why data ties are given their average rank, and why zero error counts are ignored.) The rank for the 8+1=9 error counts of one seek error is therefore the average rank of 1 and 9 for all 9 error counts. So the rank of the single seek error count of 1in the warning set is (1+9)/2=5. This is the first term in the warning set rank sum, shown in column 5. There are 3+2=5 error counts of 2, taking total ranks 10-14, 2 of which are in the warning data, so the rank sum gets two entries with the average rank of 12. There is only one warning error count of 3, so the warning rank sum gets its rank of 15. The next warning rank is at 5 errors, and its total rank of 17 goes into the rank sum. There are no 6-error counts, and the rank sum gets the next rank of 18 for its single 7-error count. The resultant rank sum is 79. Notice that no error count datum is ignored - the rank averaging of the ties keeps the total rank count equal to the total 12+6 = 18 error counts.

TABLE II: EXAMPLE RANK SUM SEEK ERROR DATA

Seek Errors	Referencedata	Warning data	Rank numbers	Warning Ranks
1	8	1	1-9	(1+9)/2=5
2	3	2	10-14	12,12
3	0	1	15	15
4	1	0	16	-
5	0	1	17	17
6	0	0	-	-
7	0	1	18	18
Sums	12	6		79

If the 18 total rank numbers independently result from the same probability distribution, then the rank sum of the 6 warning data should be the sum of six random integers from the set 1,2,…,18. Each datum has an equally random rank if all are independently drawn from any single distribution. The average rank sum should be about 6 times the average integer, or 6*(1+18)/2 = 57. The rank

sum variance should be 6 times the variance of a uniform probability distribution with range from 1 to 18, so

$$\sigma = (18-1)\sqrt{6/12} = 12.0$$

The warning rank sum of 79 is (79-57)/12 = 1.8 sigmas above its mean, significantly higher than from random probability. (These rough statistics are from Sect. V-A)

This rank sum procedure is repeated for each drive attribute. The results are combined into a single drive failure warning, if any rank sum exceeds a maximum threshold (as in present disk drive SMART), or a single overall rank sum can be computed (see Sect. II-I).

H. Rank sum test advantages

Several advantages ensue from rank hypothesis testing. First, the rank warning is based on a statistically significant test that a warning data *set* differs from the reference set instead of the single data *point* used in threshold SMART. This can lower the FAR by statistical "averaging." Second, the rank sum test makes no mathematical assumptions nor needs any information about the statistical distribution function of the data. It only assumes that the data has *some* fixed distribution if the drive is remaining stable and that the attribute samples are independent. Third, rank sum is a "stochastic dominance" test based on "ordinal" statistics. This means that failure risk is increasing if the attribute values are statistically increasing, but no numerical proportionality is assumed. Fourth, the ranks are relatively immune to errors in the attribute data. Extreme value outliers merely get the maximum rank no matter how large they are. Fifth, summing the ranks exploits the known monotonicity of the attributes (attributes are defined so larger values mean increased failure risk, see Sect. II-D). Finally, rank sum mathematics is simple enough to implement in disk drive firmware, requiring only sorting and adding of attribute integer values. The Appendix presents the mathematics of the rank sum test.

I. Multivariate tests vs. OR'ing single attribute tests.

Like the threshold SMART algorithm, rank sum SMART as just described tests attributes individually, and issues the SMART warning if *any* attribute is significantly increasing. Combining single-attribute hypothesis tests by this 'OR' operation (also used in threshold SMART) could increase the false alarm rate.

A warning algorithm based on the entire set of warning and reference attribute data could offer higher predictive accuracy at lower false alarm rate by exploiting statistical correlations between the attributes. We have developed a multivariate SMART decision rule for this purpose (Appendix, Sect. V-D). It is able to operate on variables defined so increasing values imply increasing failure risk. This covers all attributes in Table I, except for possibly using POH/CSS to capture infant mortality failures. (For that situation, the highest rank could be put on the smallest POH/CSS values, but this simple inversion would be unable to test for end-of-life failure risk.)

III. Experimental Results

Experimental data sets were obtained from drive design testing of 3744 drives of two different Quantum Corporation drive models (Table III). Each set contains 2-3 months of reliability design test data. There were 36 verified drive failures (1.0%, or 4%-6% annual rate, see Sect. II-C). The attributes found most predictive in this data were grown defects (GDC), seek errors (SKE), and read soft errors (RSE). These three attributes are also physically reasonable for the actual drive failure causes. Examples of verified failure causes in model "A" drives and their SMART attribute warnings are: grown defects from disk mechanical misalignment (GDC warning); mobile thermal asperities (foreign particles on the disk) causing grown defects (GDC); unstable servo due to head problem (SKE); head arm flex cable electrically intermittent (RSE); head instability (RSE); head burst noise (RSE). These are normal design failure types with perhaps an unusually large number of head problems.

Figs. 2a and 2b show tradeoff curves of WA vs. FAR, for the two drive models "A" and "B". In Figs. 2, the NPF drives are grouped with the won't-fail drives. But since the NPF drives did apparently suffer some transient failure during the test, internal damage might have occurred even though failure analysis found them operational and they were returned to the testing. The dotted line shows the change in the multivariate rank sum test, if NPF drives are grouped with the will-fail drives. Calling the NPF drives failures lowers the rank sum accuracy from about 32% to 20% at 0.2% FAR.

Fig. 2a shows an OR'ed rank sum test correct warning probability of about 40%, at 0.002 (0.2%) FAR probability. The multivariate rank sum is similar. Conventional SMART OR'ed thresholds have warning accuracy 3-4 times lower, at 0.2% FAR.

TABLE III: Drive Test Summary

Drive Model:	**"A"**	**"B"**
Drives in test	1936	1808
Fail drives	9	27
NPF in fail analysis	6	11
Attribute reads:	94071	63153
Significant attributes:	GDC, SKE, RSE	GDC, SKE, RSE

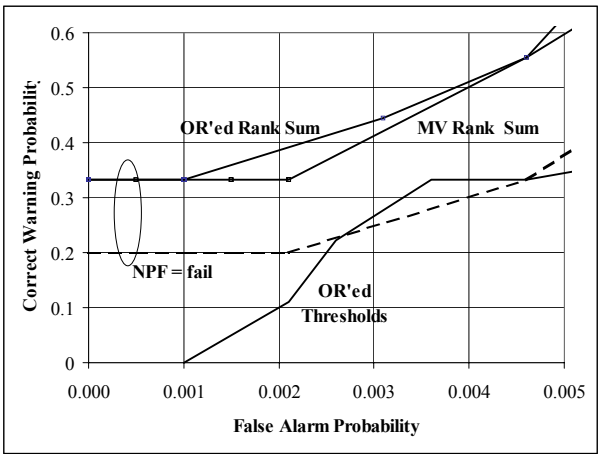

Fig. 2a: Drive model "A": Warning accuracy vs. false alarm rate. Dotted curve: MV rank sum if NPF drives are called fails.

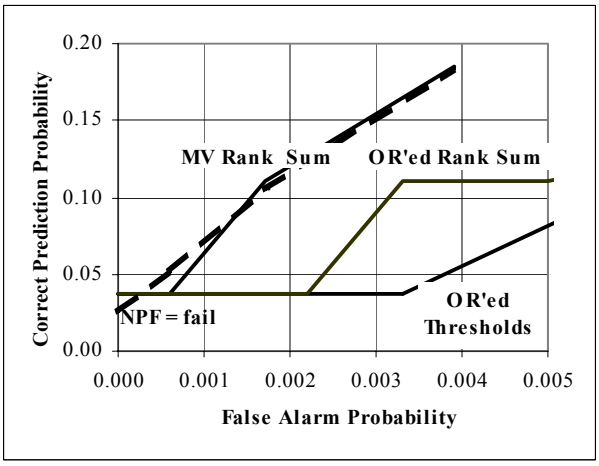

Fig. 2b: Drive model "B": Warning accuracy vs. false alarm. Dotted curve: MV rank sum if NPF drives are called fails.

Fig. 2b shows overall poorer results for drive model "B." Multivariate rank sum is best at 12% with OR'ed rank sum and OR'ed threshold tests at 4%, both at 0.2% FAR. However, this data set was difficult to analyze. Fourteen of the 1808 non-fail "B" drives had corrupted data. These were eliminated. Of the 27 will-fail "B" drives, 18 had all-zero attribute reads and their failure is unpredictable by any method. (If these 18 drives are ignored, the 12% WA increases to 36%.) Additionally, only 57 of the 1724 won't-fail drives had any nonzero attributes. This weakens the reliability of the FAR values obtained on "B."

If SMART warnings were used to signal drive replacement, the 0.2% FAR quoted above might be acceptable by drive manufacturers. It would put a 20% limit on the increase in apparently failing drives, compared to a nominal 1.0% annual drive failure rate. It is also roughly the false alarm rate with no SMART at all (20% NPF drives in 1% "perceived failure" drives per year). For a drive user, this FAR might be acceptable if the 40% WA in Fig. 2a was a useful accuracy. In a RAID array or enterprise network, higher FAR may be acceptable, if SMART is used to trigger data backup instead of drive replacement. The highest WA attainable in "A" drives was 60%, at 0.5% FAR.

These results indicate significant accuracy improvement potential over present SMART, tested on two independent drive models, using the same new SMART algorithms. However, results from drive design test data do not prove expected WA in production drives.

Additionally, disk drives have a variety of possible failure modes; not all monitored by SMART attributes. So the WA cannot be 100%. Two of the nine model "A" drive fails were unpredictable because their attribute reads were all zeros. Figs. 2 may show the highest realistic WA.

IV. Appendix

A. Rank sum hypothesis test

Consider an individual attribute x, a set of *m warning* measurements x_k, and a *reference* data set of *n* measurements. For example, the warning set in any individual drive being the last m=5 read soft error counts per SMART read interval of (say) 8 or 24 hours, stored as few-byte integers. The reference data set might be n=50 reads taken from nominal drives that passed design testing. All drives of one model made in one production configuration could have the same reference set data stored in them.

The rank sum algorithm first puts all $N=n+m$ attribute measurements in rank order, ignoring which data set they came from, with the highest rank on the numerically largest measurement. This is a simple integer sort. Then the numerical ranks of the *m warning set* measurements are added up. The resultant rank sum R is compared to a precomputed limit R_c (two-byte rank summing and a two-byte critical limit constant stored in the drive firmware code is sufficient for N=250 data samples). R_c is computed under the null hypothesis that both data sets are from the same distribution, using its mean μ and variance σ^2 [17], [18]:

$$\mu[R] = m(N+1)/2 \tag{1}$$

$$\sigma^2[R] = n(N+1)m/12 \tag{2}$$

$$\sigma/\mu = \sqrt{n/6m(N+1)} \tag{3}$$

Using μ and σ, the significance level (false alarm probability) of the rank sum test is used to calculate a critical rank sum value limit R_c, using the single-tail normal distribution $Q(X)=Pr(x>X)$, with $X= (R_c-\mu)/\sigma$.

$$FAR = Q[(Rc-\mu)/\sigma] \tag{4}$$

If a warning test rank sum R exceeds R_c, the two data sets are statistically dissimilar and failure should be predicted.

This normal distribution approximation is widely used if $n,m>20$, but was found inadequate with this SMART data, and the *FAR* was set numerically (Sect. IV-F)

B. Simplified rank-sum mathematical derivation

Under the null hypothesis (drive is remaining stable), all $N=n+m$ measurements are independent samples from the same statistical distribution. This distribution may be discontinuous or have any shape, mean, or variance.

These minimal assumptions make it equally likely that any measurement has any rank from 1 to N. Whatever the unknown underlying distribution may be, the rank of the first measurement is a random selection from the integers 1 to N. The probability of any one is $1/N$. The second measurement has equal probability of having any of the remaining N-1 ranks. (This key observation that the ranks are uniformly distributed random integers underlies many of these *distribution free* statistical methods.)

Consider a set of N balls, marked with the integers 1 through N. The rank sum is a statistic obtained by drawing m of these balls, and adding up their marked values. Ignoring for the moment that we are drawing without replacement, each integer is a random number taken with equal probability $1/N$ from the uniform discrete probability distribution with range 1 to N. Each one has expected value

$$\mu = \sum_{k=1}^{N} k / N = (N+1)/2 \tag{5}$$

The sum of m independently sampled integers strongly converges to a normal distribution, and its mean is just m times the individual mean μ, so $\mu(R)= m(N+1)/2$, which is (1). Proving (2) is more difficult, due to negative covariance between pairs of ranks caused by the non-replacement [17].

Rank sum tests as used and described here should be distinguished from other rank sum tests used for tests of location (mean) shift between two data sets, and tests of paired data [18].

C. Choosing the data sample sizes

For the normal distribution approximation (4) to be valid, the sample sizes m and n have to be sufficiently large for the central limit theorem to be valid. The "warning" data set size m = 5 used here is too small for this purpose (and made worse by the many ties in this situation of counting discrete, rare errors: see Sect. V-E. Smaller m minimizes the failure warning time, after sudden attribute changes occur. Also, the experimental data sets included some drives with only five non-zero SMART samples. It can be seen from (3), that the R test statistical variability μ/σ decreases to an asymptotic constant $\sqrt{1/6m}$ as the reference data set size n is increased. Ample "Good" drive reference experimental data was available, and a somewhat arbitrary n = 50 was chosen.

D. Multivariate rank Sum Test

Let R_i be the rank sum of attribute i considered alone, $1 \le i \le p$, p being the number of attributes. For simplicity assume that each of the attributes has the same warning and reference data set size, m and n. Then an overall rank sum for all p attributes can be defined as

$$R = \sum_{i=1}^{p} R_i \tag{6}$$

Because the attributes are defined to be monotonic (larger values mean increased failure risk, Sect. II-D), this multivariate rank sum exploits any favorable correlations among the failure attributes, because they should be positive. Replacing the error counts by their ranks automatically solves problems of scale and normalization. The individual attribute ranks can be simply added. Under the null hypothesis, the individual attribute measurements are assumed statistically independent within each attribute (as in the single-variable rank sum), and the attribute measurements are assumed independent of each other. The mean and variance of R are then:

$$\mu[R] = \sum_{i=1}^{p} \mu_i = pm(N+1)/2 \tag{7}$$

$$\sigma^2[R] = \sum_{i=1}^{p} \sigma^2[R_i] = pn(N+1)m/12 \tag{8}$$

These values can be used along with the multivariate rank sum R, as the Q-function argument in (4).

E. Data ties and zeros

The rank sum test was originally developed for continuous data, with only accidental data ties, but ties are certainly prevalent in this case of discrete-valued SMART error attributes. For example, most SMART attribute reads in Figs. 1 are tied at zero. With the experimental data used here, best results were with zero attribute values ignored, not surprising since zero error counts give little information.

The standard recipe [17 Sect. 5.1] for the rank sum test states that one imagines that tied variates be arbitrarily separated infinitesimally. Ranks are then assigned and the *average* rank of all infinitesimally close data is assigned to each of them. For example, if ten identical error counts of 1 occur in the reference plus warning data sets, then each of them gets rank (1+10)/2 = 5.5 (since we ignore the zero error counts). This rule worked well with the discrete SMART data, and best preserves the rank sum virtue that drive failure trends producing simultaneous positive shifts in the attributes will produce large changes in the rank sum, towards the failure limit.

F. Setting the rank sum failure limit parameter

Equations (1)-(4) change and lose accuracy with ties, although the rank sum remains a robust statistical test [17]. Discrete valued rare-error attributes can produce enough ties that the normality approximation leading to (4) is inaccurate. As a rough rule, the number of untied values in the smaller data should exceed 20 [18]. For the

experimental data here the significance level had to be set numerically to get a desired FAR.

A good method to do this is simply to find the rank sum limit producing the desired FAR in the "won't-fail" drives. Average over all experimental drives and all warning sets of m sequential attribute reads of each drive (all possible SMART read times). This was tested on the new rank sum tests, their OR'ed SMART flag result, and the (single) multivariate rank sum test.

Section II-E discusses why this should be a conservative estimate of production drive FAR.

ACKNOWLEDGMENT

This work was supported by the UCSD Information Storage Industry Center, funded by the Alfred P. Sloan Foundation. The authors thank Tim Nelson, Christopher Reynolds, and Claude Camp of the Quantum Corporation for assistance and for supplying the experimental data, and also thank Winston Tran, R. Bohn, and S. Schultz.

REFERENCES

[1] E. Grochowski, "Future Trends in Hard Disk Drives," *IEEE Trans. Magnetics,* vol. 32, no. 3, May 1996, pp. 1850-4.

[2] J. Yang and F-B Sun, "A comprehensive review of hard-disk drive reliability," *1999 Proceedings Annual Reliability and Maintainability Symposium,* Washington DC, pp. 403-9, Jan. 1999.

[3] ATA SMART Feature Set Commands, www.t13.org; Small Form Factors Committee SFF-8035; and SCSI "Mode Sense" code "Failure Prediction Threshold Exceeded," www.t10.org), available from American National Standards Institute, 11 W. 42nd Street, New York, New York 10036.

[4] J. Yang and X. K. Zunzanyika, "Field reliability projection during new (disk drive) product introduction," *Proceedings – Annual Technical Meeting of the Inst. of Environmental Sciences,* May 1966, pp. 58-67.

[5] Elerath, J.G. Specifying reliability in the disk drive industry: No more MTBF's,". Proceedings. International Symposium on Product Quality and Integrity , p.194-9, 2000

[6] A.S.R. Murty and V.N.A Naikan, "Condition monitoring strategy-a risk based interval selection," *International Journal of Production Research*, vol.34, (no.1), Taylor & Francis, Jan. 1996, p.285-96.

[7] B.E..Preusser and G.L Hadley, "Motor current signature analysis as a predictive maintenance tool," *Proceedings of the American Power Conference,* Chicago, IL, USA, 29 April-1 May 1991, Chicago, IL, USA: Illinois Inst. Technol., 1991, p.286-91 vol.1.

[8] Lewin, D.R., "Predictive maintenance using PCA," *Control Engineering Practice*, vol.3, (no.3), March 1995. p.415-21.

[9] "Drive industry endorses protocol for predicting disk drive failure," *Data Storage magazine,* vol. 2, no. 9, Sept. 1995, http://ds.pennwellnet.com

[10] SMART monitor software information: www.compaq.com/im/fault.html. Drive manufacturer SMART technology sources: www.seagate.com/support/kb/disc/SMART.html, www.storage.ibm.com/oem/tech/pfa.htm, www.westerndigital.com/sevice/lifeguard/dig_smart.html

[11] J.D.Coker, R.L Galbraith, and G.J.Kerwin, "Magnetic characterization using elements of a PRML channel," *IEEE Transactions on Magnetics*, vol.27, no.6, pt.1, pp.4544-8, Nov. 1991.

[12] T.J. Chainer and E.J. Yarmchuk, A technique for the measurement of track misregistration in disk file," *IEEE Transactions on Magnetics,* vol.27, (no.6, pt.2), Nov. 1991. p.5304-6.

[13] B.C. Schardt, E. Schreck, R. Sonnenfeld, Q. Haddock, and J.R. Haggis, "Flying height measurement while seeking in hard disk drives," *IEEE Transactions on Magnetics*, vol.34, (no.4,pt.1), July 1998. p.1765-7.

[14] K. Rothman and S. Greenland, *Modern Epidemiology*, 2nd ed., Philadelphia: Lippencott-Raven 2000, p509

[15] F. Wilcoxon, "Individual comparisons by ranking methods," *Biometrica*, 1, pp. 80-83, 1945.

[16] P.D Bridge, and S.S. Sawilowky , "Increasing Physicians' Awareness of the Impact of Statistics on Research Outcomes: Comparative Power of the t-Test and Wilcoxon ran-sum test in small samples applied research," *Journal Of Clinical Epidemiology,* March 1999, 52(3):229-35.

[17] W. Conover, *Practical Nonparametric Statistics,* 3rd ed., New York: Wiley, 1999, pp. 271-281

[18] E. L. Lehman and H. J. M. D'Abrera, *Nonparametrics – Statistical Methods Based on Ranks,* New Jersey, Prentice-Hall, 1999

AUTHORS

Gordon Hughes is the Associate Director of the Center for Magnetic Recording Research at the University of California, San Diego. He is the principal investigator of the UCSD SMART project on disk drive predictive failure He received his BS in Physics and Ph.D. in Electrical Engineering from Cal Tech. Before 1983 he worked for Xerox PARC, on magnetic recording research for disk drives, then joined Seagate Technology as Senior Director of Recording Technology. At Seagate he worked on recording heads, disks, and systems, was part of the team that established sputtered thin film disk media as today's standard, and formed Seagate's first Material and Process Lab and Contamination Control Group for drive reliability analysis.

Charles Elkan is an associate professor in the Department of Computer Science and Engineering at the University of California, San Diego. His main research interests are in artificial intelligence and data mining. His research has led to two best paper awards and first place in two international data mining contests. In 1998/99 Dr. Elkan was a visiting associate professor at Harvard. He earned his Ph.D. at Cornell University in computer science, and his B.A. at Cambridge University in mathematics.

Joseph Murray received the B.S. in Electrical Engineering at the University of Oklahoma in 1998, and the M.S. in Electrical and Computer Engineering from the University of California, San Diego (UCSD). Currently he is a Ph.D. student at UCSD, and his interests include pattern recognition, learning theory and computer vision.

Kenneth Kreutz-Delgado (SM '93) received the M.S. in Physics and Ph.D. in Engineering Systems Science from the University of California at San Diego (UCSD) where he is currently a professor in the Department of Electrical and Computer Engineering (ECE). Prior to joining the Faculty of UCSD in 1989, Dr. Kreutz-Delgado was a researcher at the the Jet Propulsion Laboratory, California Institute of Technology, where he worked on the development of intelligent space telerobotic systems. His

current research interests include applications of statistical data analysis and learning theory, nonlinear signal processing, and computational intelligence to communication systems, bioinformatics, predictive failure analysis, machine intelligence, and robotics.