SiaDFP: A Disk Failure Prediction Framework Based on Siamese Neural Network in Large-Scale Data Center

Xiaoyu Fang, Wenbai Guan, Jiawen Li, Chenhan Cao, and Bin Xia

***Abstract*—With the rapid development of cloud services, service providers increasingly rely on a dependable storage system equipped with large-capacity disks to ensure data availability. The primary source of unreliability in such storage systems attributes to disk failures. In recent years, some proactive methods base on machine learning models have emerged, aiming to predict impending disk failures by leveraging the SMART attributes of disks. These methods enable service providers to timely back up storage data. While the methods prove more effective and efficient in disk failure prediction, they still face challenges, such as inadequate mining of abnormal information and imbalanced classification. In this paper, we mainly analyzed the change of data distribution in hard disks. From the data analysis, we observed that the distribution change in the failed disk is obvious during the period before the disk damage, while that in the healthy disk is insignificant during running time. Motivated by the observation, we propose a novel framework named SiaDFP, based on Siamese neural network, designed to predict impending disk failures by capturing the distribution changes in failed disks. Additionally, we observed that the failed disks exhibit some change points as an abnormal feature by analyzing the disk data trend. To fully mining abnormal information inhere in failed disks, we propose CP-MAP mechanism and 2D-Attention mechanism. Furthermore, we present a subsampling approach named Region Balanced Sampling to address the challenge of imbalanced classification. Experiments on the real-world dataset Backblaze and Baidu demonstrate that the performance of SiaDFP is outstanding in the task of disk failure prediction.**

***Index Terms*—Attention mechanism, change point detection, disk failure prediction, siamese neural network.**

I. Introduction

THE proliferation of cloud service technology has led numerous organizations to favor the deployment of projects on cloud platforms. These endeavors necessitate substantial hard disks for data center of cloud platforms [1]. The dependability of the data center is of paramount importance, as it significantly influences the choice of cloud service providers by users [2]. However, a substantial proportion, ranging from 76% to 95%, of disk replacements in large-scale data centers can be attributed to hard disk failures [3]. Such failures not only result in service downtime but also huge data loss, potentially incurring financial losses for both users and service providers [4], [5].

The classification of disk failures including two primary categories: unpredictable and predictable. Unpredictable failures are chiefly attributed to physical damage and external influences [6], which tend to appear suddenly and cannot be premonitored early, posing an impossible challenge to predict. The failure could be prevented by the reactive methods (e.g., Erasure coding [7], [8] and Redundant Array of Independent Disks (RAID) [9]). In contrast, predictable failures are not impulsive behaviors but gradual ones, primarily caused by component wear. The characteristics of the failures could be monitored by the Self-Monitoring Analysis and Reporting Technology (SMART), which makes predicting the failure possible. For the failure, combined proactive methods with the reactive methods could effectively ensure the data security. In this paper, we mainly focus on addressing the predictable failure in disks with proactive method.

To augment the reliability of data centers, researchers have introduced many proactive methods for predictable disk failures. The proactive methods are designed to predict impending disk failures, affording service providers the opportunity to safeguard their data in advance. The methods mainly relies on the analysis of disk attributes, as recorded by the SMART, including metrics such as power-on time and seek error rate. SMART is capable of discerning abnormal disks by comparing attribute values to predefined thresholds [10]. Nevertheless, the approach is straightforward [11], resulting in the failure detection rate (FDR) of merely 3%–10%, coupled with a meager false alarm rate (FAR) of 0.1% [12]. Several traditional machine learning-based models have been proposed to predict disk failures by using SMART attributes as features. These models include regression trees [13], random forests [14], and Bayesian networks [15]. However, the models fall short of capturing the temporal characteristics of SMART attributes, which contain crucial information for identifying the failed disks. Some deep learning models, such as Recurrent Neural Networks (RNNs) [16] and Long Short-Term Memory (LSTM) networks [17], [18], have demonstrated the ability to address the limitations of traditional machine learning-based models. However, the deep learning models lack an explanation for the reason behind the disk failure. To overcome the limitation of explanation of disk failure,

Manuscript received 14 October 2023; revised 12 April 2024; accepted 16 April 2024. Date of publication 29 April 2024; date of current version 9 October 2024. *(Corresponding author: Bin Xia.)*

Xiaoyu Fang, Wenbai Guan, Jiawen Li, and Chenhan Cao are with the Nanjing University of Posts and Telecommunications, Nanjing 210049, China.

Bin Xia is with the Jiangsu Key Laboratory of Big Data Security and Intelligent Processing, Nanjing University of Posts and Telecommunications, Nanjing 210049, China (e-mail: bxia@njupt.edu.cn).

This article has supplementary downloadable material available at https://doi.org/10.1109/TSC.2024.3394692, provided by the authors.

Digital Object Identifier 10.1109/TSC.2024.3394692

1939-1374 © 2024 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See https://www.ieee.org/publications/rights/index.html for more information.

Authorized licensed use limited to: Nanjing Univ of Post & Telecommunications. Downloaded on March 09,2025 at 02:06:28 UTC from IEEE Xplore. Restrictions apply.

attention mechanism [19], [20] has been introduced for disk failure prediction. The mechanism excels at capturing significant abnormal features within disk data, providing administrators with a convincing understanding of the root causes of failures. However, the attention mechanism can only identify the importance of features based on single-dimensional information, typically the time dimension. The feature of disks extracted by the mechanism is redundant and inaccurate. Furthermore, the existing methods fail to harness the valuable information residing in change points within SMART attributes, which is a pivotal feature for distinguishing between failed and healthy disks [19], [21].

In this study, we discover that the distribution of SMART attributes greatly changes during the period before the disk becomes damaged, while the distribution change is not obvious in the running period of the healthy disk. Motivated by the observation, we explore a novel framework **SiaDFP**, which could effectively capture the distribution change based on Siamese neural network. Meanwhile, to address the above limitations in existing methods, we propose CP-MAP mechanism and 2D-Attention mechanism to capture abnormal features. In real-world dataset, the number of healthy disks and failed disks is severly imbalanced, which results in the model being biased in predicting the disk failure. To address the issue, we proposed the Region Balanced Sampling (RBS) to select the representative disks. Extensive experiments are conducted to validate the effectiveness of **SiaDFP** on the real-world dataset Backblaze and Baidu. The contributions of this work are as follows:

- We proposed a novel deep framework **SiaDFP** based on Siamese neural network, which proved to be highly effective in capturing distribution changes as abnormal information for disk failure.
- We proposed CP-MAP mechanism and 2D-Attention mechanism to mine abnormal features of failed disks. The CP-MAP is to capture the feature of change points that appeared in SMART attributes, while the 2D-Attention mechanism evaluates the significance of abnormal features by considering information from two distinct dimensions.
- We proposed the Region Balanced Sampling to address the imbalance classification in disk failure prediction. The Region Balanced Sampling could select the representative disks as the dataset for training deep learning models.

II. Related Work

A. Disk Failure Prediction

In the field of disk failure prediction, the related research can be divided into (1) Imbalanced classification, (2) Granularity of prediction, (3) Data-driven task, (4) Fail-Slow detection, and (5) Explanation of disk failure. Real-world scenarios often present a substantial disparity between the number of failed disks and healthy ones, which poses challenges to the effectiveness of disk failure prediction methods. To solve this problem, Liu et al. [22] proposed a semi-supervised prediction method for disk failure based on a model that combines variational autoencoder and long short-term memory network. The model can predict disk failure by capturing the pattern of healthy disks. Xu et al. [23] proposed a data augment method to minimize the limitation of the number of failed disks. This method generates failed disk data for training by shifting the data on the time axis forward or backward with a certain time step. In addition, the granularity of disk failure prediction has been a subject of investigation. Züfle et al. [24] explored the prediction of disk failure within various specific future periods. Their approach involves classification to predict the impending failure period of disks and regression to determine the exact day when a disk is likely to fail. Liu et al. [16] considered the disk failure prediction as the disk status prediction, where the disk status is split into several healthy levels based on expert knowledge. The administrators are capable of estimating the disk failure based on the specific healthy level of disks. The traditional disk failure predictions are totally conducted based on the SMART data, which constitutes only a fraction of the comprehensive information available in data centers. To this end, Lu et al. [18] constructed a data-driven method that supplied the SMART data with the system-performance information to improve the performance of disk failure prediction. Furthermore, Luo et al. [20] considered the information from adjacent disks as additional features in their model, exploring the interplay among disks for improving predictive accuracy. For the fail-slow problem in disk failure, Lu et al. [25] implemented a light regression-based model to fast pinpoint and analyze fail-slow failures at the granularity of drives. The explanation of disk failure is also a significant task that provides convincing advice for administrators to decide whether the disk should be replaced. Yu et al. [19] implemented the attention mechanism to identify abnormal features leading to disk failure based on the information of time dimension, aiding administrators in conducting comprehensive failure analyses. However, the above methods still encounter some challenges: (1) The substantial number of healthy disks far exceeds that of failed disks in real-world scenarios. However, the existing methods could not effectively select representative healthy disks to assist models in fully understanding the characteristics of healthy disks. (2) Some abnormal information (e.g., change points, distribution offset) exhibited in the SMART attributes before disk failed. However, the existing methods could not effectively utilize the abnormal information to help models adequately learning the characteristics of failed disks. (3) The importance of various SMART attributes varies for disk failure prediction. However, the existing methods could not effectively distinguish the importance of the attributes. The challenges result in the unsatisfactory performance of the existing methods in disk failure prediction.

B. Siamese Neural Networks

Siamese neural network, comprising two parallel networks sharing structures and parameters, serves the purpose of calculating the similarity between two objects [26]. This model is widely used in face verification, natural language processing (e.g., semantic similarity), and object tracking. In face verification, Yann et al. [27] proposed a siamese convolutional neural network to identify faces. The model outputs higher similarity between

Authorized licensed use limited to: Nanjing Univ of Post & Telecommunications. Downloaded on March 09,2025 at 02:06:28 UTC from IEEE Xplore. Restrictions apply.

the faces from the same person while outputs lower similarity between the faces from different people, based on the proposed discriminative loss function. For the practical scenario, Khalil-Hani et al. [28] proposed a light-weight siamese convolutional neural network for the face verification task, which converges faster and has better generalization. Semantic similarity assessment is a popular task in natural language processing, involving comparisons between documents, sentences, or words. Mueller et al. [29] proposed a siamese-structure model based on LSTM to calculate the similarity between sentences within different lengths. Bolucu et al. [30] introduces a novel framework that combines an attention network and a Siamese neural network for the task of textual similarity. The attention network is employed to capture the semantic representation of a sentence. The Siamese neural network is applied to perform semantic textual similarity tasks using these semantic representations. Different from the aforementioned tasks, object tracking aims to locate the same object across consecutive frames. In other words, Siamese neural network is used to find out the most similar region (i.e., the tracked object) in subsequent frames. Bertinetto et al. [31] proposed a fully convolutional neural network based on the siamese architecture, which locates the position of the tracked object in the partial regions of the subsequent frame. However, the performance of the proposed model is sensitive to the size of the tracked object in different frames. To address this issue, Li et al. [32] combined the original classification branch with a bounding box regression branch (i.e., a region proposal network) to improve the robustness of the model for the diverse size of objects.

III. Data Investigation & Analysis

Analyzing the data from hard disks contributes to identifying the difference between the healthy disk and the failed disk, and gaining deep insights into the characteristics of disk failure. In this section, we will explore the characteristics of the hard disk based on two real-world datasets sampled from Backblaze and Baidu. 87 SMART attributes are collected per day over consecutive 6 years in Backblaze dataset, while 12 SMART attributes are collected at much finer-granularity (i.e., per hour) over consecutive 20 days in Baidu dataset. In this study, our main focus will be on analyzing SMART trends and distribution offset based on the two datasets.

A. SMART Trend

The SMART trend represents the direction of change in SMART values over a specific period on a disk, which is a crucial factor in distinguishing failed disks from healthy ones. In this section, our primary focus is on tracking the trends of SMART attributes over time to identify differences between healthy and failed disks. After investigating SMART attributes across all disks from the Backblaze dataset and the Baidu dataset, we found that the trends of various SMART attributes behave differently but stable in healthy disks, as shown in Figs. 1 and 2: (1) the value of Reallocated Sectors Count remains constant; (2)the value of Hardware ECC Recovered remains seasonal change with fixed range; (3)the value of Current Pending Sector remains constant; (4)the value of Total LBAs Read remains linear change. However, in failed disks, the trends of SMART attributes suddenly change before the disks become damaged as shown in Figs. 1 and 2. The sudden changes in trend are primarily caused by firmware wear, physical damage, or software issues, and are clear abnormal behaviors for distinguishing failed disks from healthy ones.

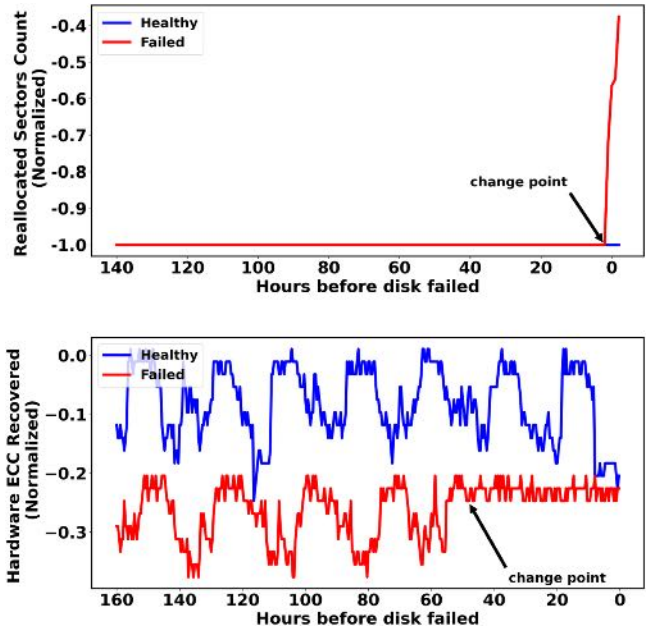


Fig. 1. Trend of SMART attributes in healthy disk and failed disk from Baidu dataset.

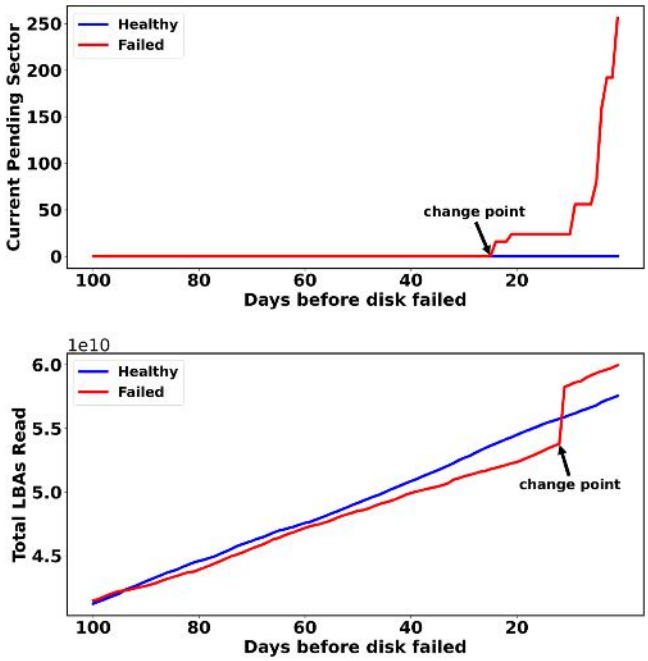


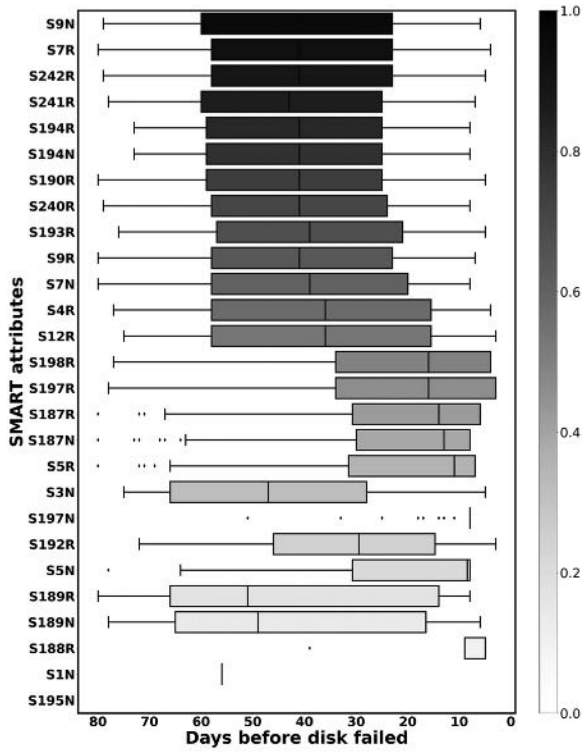
Fig. 2. Trend of SMART attributes in healthy disk and failed disk from Backblaze dataset.

Authorized licensed use limited to: Nanjing Univ of Post & Telecommunications. Downloaded on March 09,2025 at 02:06:28 UTC from IEEE Xplore. Restrictions apply.

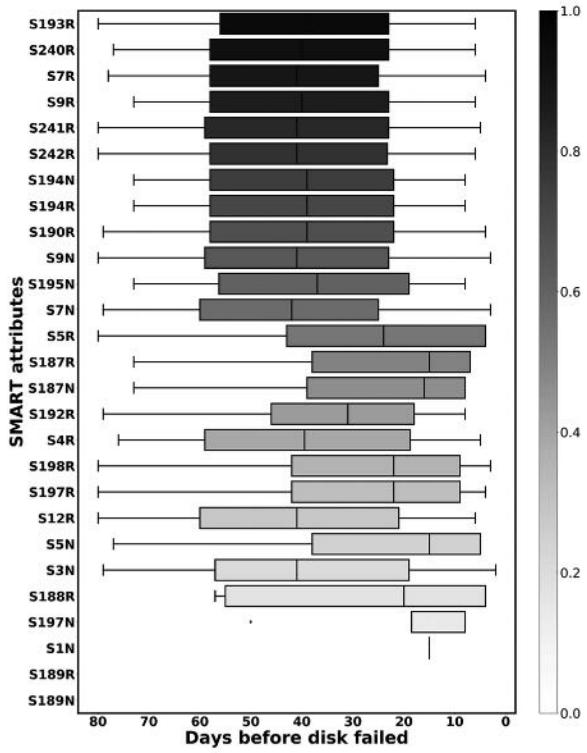
Furthermore, the change points displayed in Figs. 1 and 2 serve as indicators of change in the trends of SMART attributes. To investigate the characteristics of trend changes among various SMART attributes in pre-damage states of failed disks, we utilized the Bayesian Change Point Detection [33] method to capture change points in SMART attributes. Figs. 3(a), (b), and 4 illustrate the time intervals at which change points appeared in SMART attributes before disk failed among Seagate ST4000DM000 and ST8000DM002 disks from the Backblaze dataset and that among Seagate ST31000524NS disks from Baidu dataset. The time appearing change points in SMART attributes varies across different disk models as shown in Figs. 3(a), (b), and 4. Meanwhile, within a single model of disk (e.g., Seagate ST4000DM000), the time appearing change points in various SMART attributes exhibit dissimilarity, implying the abnormal behaviors (i.e., trend changes) in various SMART attributes occur at distinct time before disk failed. For instance in Seagate ST4000DM000 from Backblaze dataset, the attribute SMART_197_RAW typically appears change points around 10 days before disk failed. Conversely, attributes SMART_7_NORM and SMART_193_RAW appear change points around 40 days before disk failed. In addition, it is noteworthy that attributes SMART_9_NORM and SMART_7_RAW tend to display significantly more change points compared to attributes SMART_188_RAW and SMART_5_NORM among all Seagate ST4000DM000 disks. This observation suggests that abnormal behaviors (trend changes) are frequently observed in certain SMART attributes (e.g., SMART_9_NORM and SMART_7_RAW), which should be focused. Furthermore, merely 90 percent of disks within the Baidu dataset exhibit change points (predominantly observed in SMART_194_NORM) as shown in Fig. 4. Few disks from Baidu dataset do not appear any change point before the disk failed [18]. However, the data of SMART attributes from Baidu dataset are collected during 20 days before disk failed, missing the data over 20 days. Therefore, conclusions about trend change drawn from analyzing the Baidu dataset are constrained. In the next section, we primarily focus our analysis on the disks from Backblaze dataset.

B. Distribution Offset

Distribution offset refers to the change of data distribution in SMART attributes during a consecutive period. In this section, we mainly investigate the difference of distribution offset in healthy disks and failed disks. First, to denote the difference between the distribution in healthy disks and that in failed disks, we implement time windows with the length of 30 days sampling from healthy disk and failed disk. For healthy disks, distribution is sampled randomly within the lifespan of the disks, whereas for failed disks, distribution is sampled from the period leading up to failure. Fig. 5 illustrates the data distribution of various SMART attributes in healthy and failed disks from ST4000DM000. The comparison of data distribution between healthy and failed disks reveals distinct differences. Fig. 6 depicts the distribution offset of SMART_241_RAW in different statuses of failed disks before disk failed. We can learn that the distribution in



(a) ST4000DM000



(b) ST8000DM002

Fig. 3. Time appeared change points before disk failed among all failed disks of Seagate ST4000DM000 and ST8000DM002 in Backblaze dataset. S*R in the y-axis represents the attribute of SMART_*_RAW, and S*N in the y-axis represents the attribute of the SMART_*_NORM (e.g., S4R denotes the SMART_4_RAW attribute and S190 N denotes the SMART_190_NORM attribute). The colorbar denotes the ratio of the number of disks captured change points to the total number of disks.

Authorized licensed use limited to: Nanjing Univ of Post & Telecommunications. Downloaded on March 09,2025 at 02:06:28 UTC from IEEE Xplore. Restrictions apply.

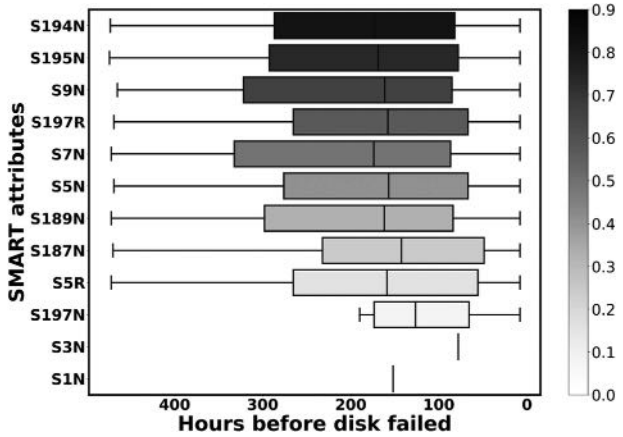
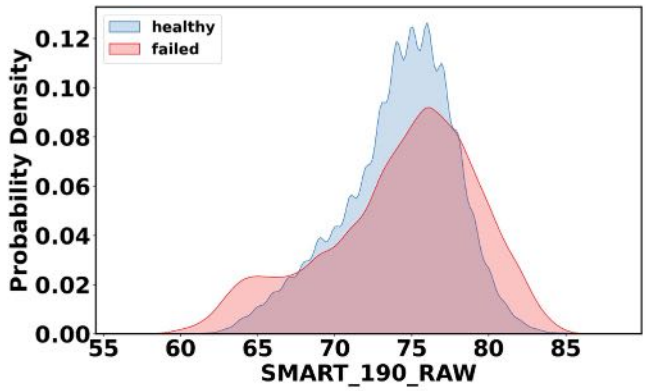
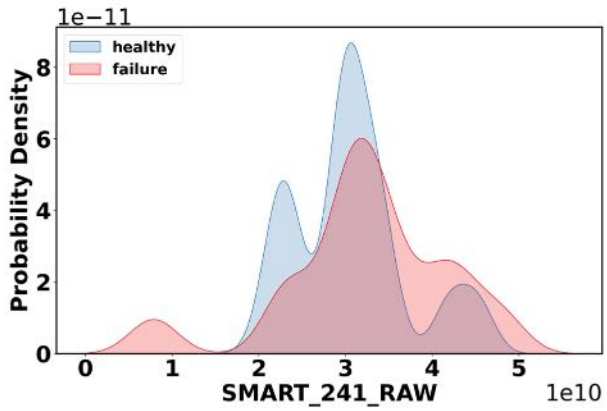


Fig. 4. Time appeared change points before disk failed among all failed disk of Seagate ST31000524NS in Baidu dataset.

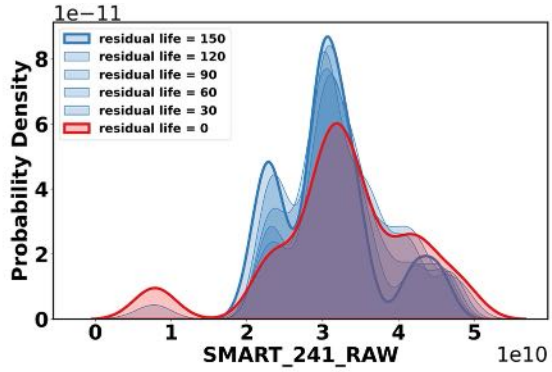


(a) Probability density of SMART_190_RAW

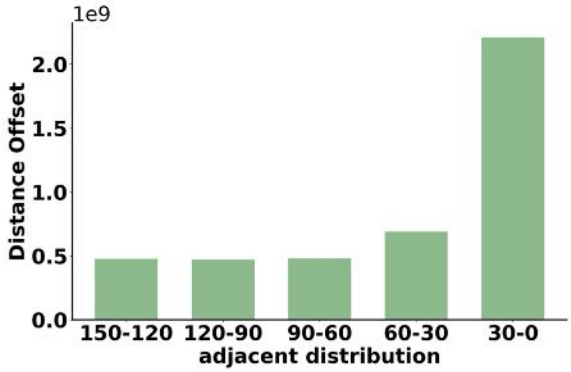


(b) Probability density of SMART_241_RAW

Fig. 5. Distribution offset in various SMART attributes.

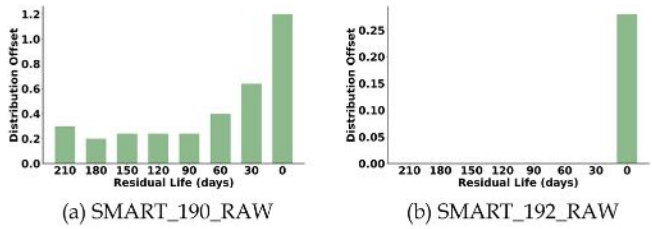


(a) Processing of distribution offset in SMART_241_RAW



(b) Euclidean Distance of various distribution in SMART_241_RAW

Fig. 6. Distribution offset in SMART_241_RAW.



(a) SMART_190_RAW (b) SMART_192_RAW

Fig. 7. Distribution offset before the disk W300KW25 failed within the length of 60 days time window.

healthy status (i.e., residual life >30) is similar as shown in Fig. 6(a). Meanwhile, the euclidean distance of the distribution in healthy status (i.e., residual life >30) is small as shown in Fig. 6(b). However, the distribution is greatly changed when the disk gets abnormal (i.e., residual life <30) before disk failed, and the difference of the distribution between abnormal status and healthy status is huge. As mentioned in III-A, the trends of SMART attributes remain stable in healthy disks, while in failed disks, they suddenly change due to abnormal behaviors. Consequently, distribution offsets tend to be small in healthy disks while larger in failed disks. To capture the distribution offset in various SMART attributes of a disk, we implement fixed-length time windows to sample SMART attributes during consecutive periods (i.e., 60 days). In addition, we utilize the euclidean distance between the distribution of SMART attributes in the first half of the time window and that in the second half of the time window (1) to quantify the distribution offset.

$$E\text{-}D(\boldsymbol{S}_i) = (\boldsymbol{S}_i^{h+} - \boldsymbol{S}_i^{h-})^2. \quad (1)$$

$\boldsymbol{S}_i$ represents a time series vector of ith attribute. $h-$ and $h+$ denote the first half time and the second half time, respectively. $\boldsymbol{S}_i^{h-}$ represents a sub-vector of $\boldsymbol{S}_i$ in the first half of the time and $\boldsymbol{S}_i^{h+}$ represents a sub-vector of $\boldsymbol{S}_i$ in the second half time.

Fig. 7 shows distribution offset of the disk W300KW25 from ST4000DM000 in various periods before the disk failed within

Authorized licensed use limited to: Nanjing Univ of Post & Telecommunications. Downloaded on March 09,2025 at 02:06:28 UTC from IEEE Xplore. Restrictions apply.

the length of 60 days time window. The distribution offset is relatively small in the healthy status of the disk (i.e., residual life >30 days), since the trend of SMART attributes are stable. However, the distribution offset is very large when the disk being abnormal status (i.e., residual life <30 days).

The phenomenon of distribution offset is widely observed in various models of modern commercial disks (See Fig. A1 in Appendix A, available online). We selected the specific attributes for each disk model that could be observed the phenomenon in the failed disks (See Table A1 in Appendix A, available online). In fact, most of the selected attributes are implemented to record abnormal events occurring during disk operations. The occurrence of such abnormal events is unusual when a disk is healthy, and the distribution of the SMART attributes remains stable during the period, with a slight distribution offset. However, when the disk becomes abnormal, the abnormal events occur more frequently. Meanwhile, the distribution of related SMART attributes changes, which results in a significant distribution offset. Table A1 shows that SMART_5_RAW (Reallocated Sectors Count), SMART_197_RAW (Current Pending Sector Count), and SMART_198_RAW (Off-Line Uncorrectable Sector Count) are selected in most disk models. These attributes are related to the status of disk sector which is always becoming abnormal suffering from the frequent read/write operations. For example, when the disk sectors are healthy, the SMART_5_RAW value remains consistently stable, and the distribution offset is small. However, when a disk sector consistently experiences abnormalities (i.e., read/write errors), the disk firmware program will redirect the address of the sector to a pre-reserved healthy spare sector [10]. Meanwhile, the distribution of SMART_5_RAW changes when the number of reallocated sectors becomes large, indicating that the disk becomes abnormal. Similarly, the significant distribution offset of SMART_197_RAW and SMART_198_RAW also indicates the potential fault of the corresponding disks.

The length of the time window could impact the quantified value of distribution offset in various SMART attributes. To explore the relationship between the two, we implement distinct lengths of time window sampling healthy disks and failed disks among various SMART attributes. Meanwhile, to denote the difference between quantified value of distribution offset in healthy disks and that in failed disks, the rate of quantified value of distribution offset in failed disks compared to that in healthy disks is used (i.e., a large rate means the difference is large). The length of time window with the largest rate varies across SMART attributes as shown in Fig. 8.

To explore an optimal and unified length of time window for a disk with various SMART attributes (i.e., distribution offset in failed disks and that in healthy disks are markedly different), we implement (2) to quantify the distribution offset in a disk with the selected SMART attributes :

$$D\text{-}O_{disk} = \sum_{i}^{n} E\text{-}D(\boldsymbol{S}_i), \tag{2}$$

where $D\text{-}O_{disk}$ represents the Distribution Offset and $E\text{-}D$ represents the euclidean Distance. Fig. 9 illustrates the rate of

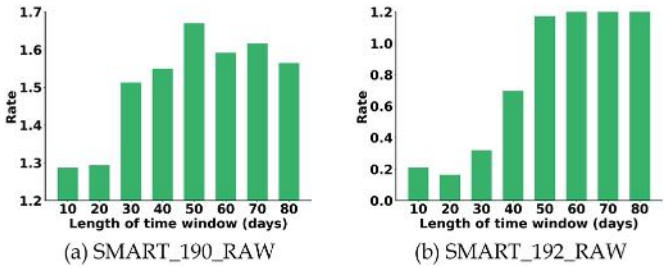


Fig. 8. Rate of distribution offset in failed disks to that in healthy disks among various SMART attributes with different length of time windows in ST4000DM000.

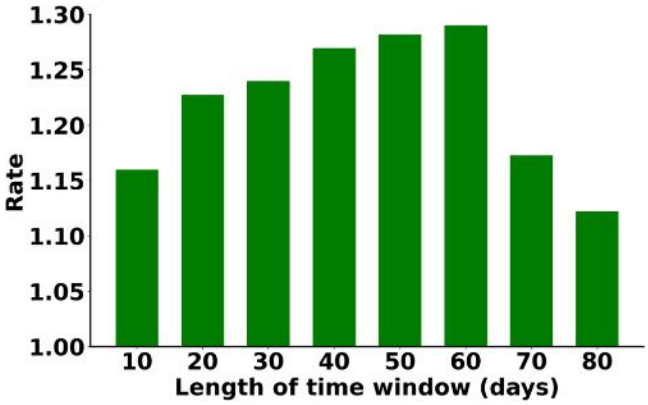


Fig. 9. Rate of distribution offset in failed disk to healthy disk with various length of time window in ST4000DM000.

distribution offset in failed disks compared to healthy disks for ST4000DM00 disks among the SMART attributes with various time window lengths. To effectively distinguish failed disks and healthy disks, we selected the length with largest rate as an optimal time window length. Notably, we observe that the time window length of 60 days appears to be optimal for distinguishing the distribution offset in failed disks from that in healthy disks in ST4000DM000 as shown in Fig. 9. Additionally, we explored the optimal window length for various disk models (See Fig. B1 in Appendix B, available online). We found that the rate within most time window lengths (i.e., 20–80 days) for various disk models is >1, signifying a distinguishable distribution offset between failed and healthy disks within the time window lengths. However, the optimal length of time windows varies among different disk models (e.g., 40 days for ST8000NM0055, 30 days for ST12000NM0008). The discrepancy arises from the differing design and application principles employed by manufacturers for various disk models, leading to variations in the distribution of the same SMART attributes among different disk models. Consequently, the optimal time window lengths differ across various disk models.

Based on the data investigation and analysis regarding SMART trends and distribution offset, several key insights emerge:

1) The SMART trend undergoes changes before disk failure, and the identification of change points becomes a crucial

Authorized licensed use limited to: Nanjing Univ of Post & Telecommunications. Downloaded on March 09,2025 at 02:06:28 UTC from IEEE Xplore. Restrictions apply.

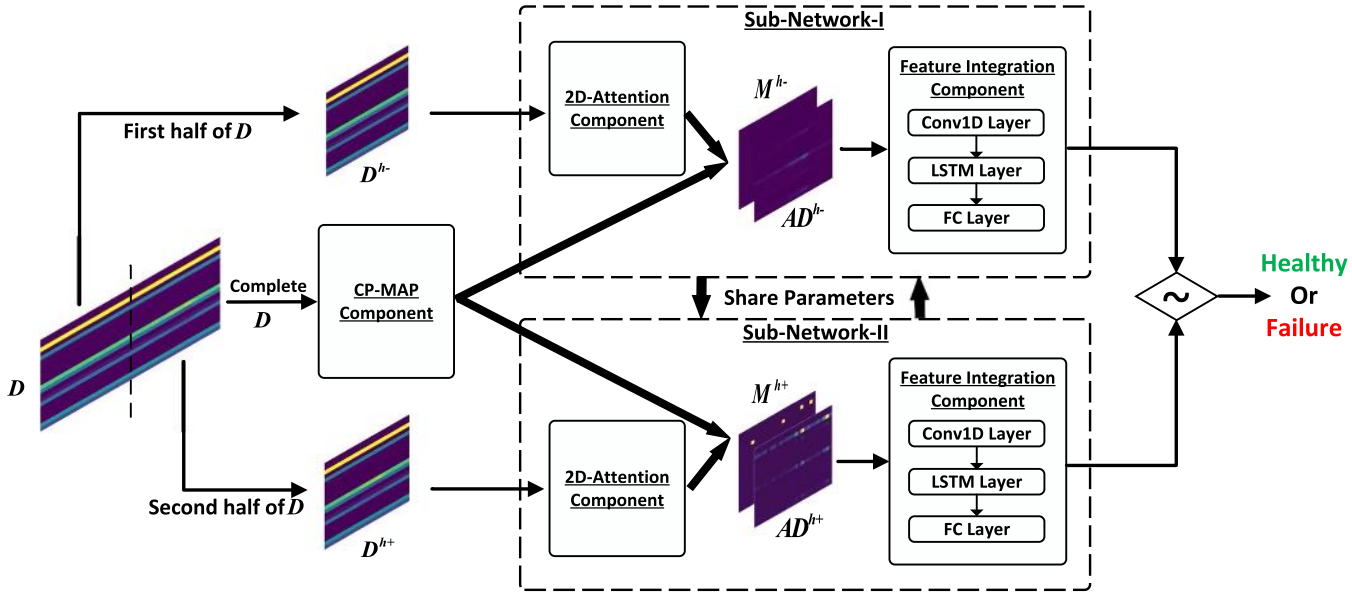


Fig. 10. Framework of the SiaDFP.

feature for predicting disk failure. It is essential to develop a mechanism to capture these change points effectively. Moreover, the time of trend change varies across different SMART attributes, indicating that the significance of abnormal features may differ at different times. Additionally, some SMART attributes exhibit frequent trend changes, while others do not, suggesting varying importance of SMART attributes in predicting disk failure. As a result, a mechanism to distinguish the importance of these features across different time and SMART attributes becomes essential.

2) The distribution offset between healthy and failed disks exhibits distinct differences. Creating a model that effectively captures this distribution offset can significantly contribute to distinguishing failed disks from healthy ones.

IV. Method

This study aims to predict whether a disk will be damaged in the next few days based on the disk data composed of multiple SMART attributes recorded during a consecutive time. As revealed by the analysis in Section III, a notable observation is that the distribution offset in healthy disks tends to be small, while in failed disks, it exhibits to be large. Motivated by the observation, we propose a framework **SiaDFP** to capture the distribution offset of the disk for predicting the impending failure. The disk is predicted to fail in the next days if the offset exceeds the pre-set threshold, and on the contrary, it is predicted to remain healthy. Fig. 10 shows a comprehensive illustration of **SiaDFP**. The framework mainly consists of a CP-MAP component and two sub-networks. The CP-MAP component is to capture change points within SMART attributes, which is a significant feature in failed disks. The sub-networks have parallel structures sharing parameters, while each sub-network contains a 2D-Attention component and a feature integration component. The 2D-Attention component is to capture the important abnormal features within the SMART attributes. The feature integration component is to synthesize the final feature representation of the disk by combining the outputs of the CP-MAP component and the 2D-Attention component. In addition, We implement the notation $\boldsymbol{D} \in \mathbb{R}^{n \times 2m}$ to denote the disk with n SMART attributes recorded in consecutive $2m$ days. $\boldsymbol{D}^{h-}$ denotes the disk data for the first half of the time period within $\boldsymbol{D}$, while $\boldsymbol{D}^{h+}$ denotes the disk data for the second half of the time period within $\boldsymbol{D}$. In this section, we will introduce the framework **SiaDFP** in detail.

A. CP-MAP Component

As noted in data investigation and analysis III, the presence of change points within the SMART attributes of failed disks is indicative of abnormal disk behavior. To capture the change points in SMART attributes, we proposed the CP-MAP mechanism, which leverages the Bayesian Change Point Detection Algorithm [34]. In detail, for each time series of SMART attribute $\boldsymbol{D}_{i\cdot} = (\boldsymbol{D}_{i\cdot 1}, \boldsymbol{D}_{i\cdot 2}, \ldots, \boldsymbol{D}_{i\cdot 2m})$ in $\boldsymbol{D}$, the algorithm is employed to assess the probability of specific days when change points, denoted as $\boldsymbol{D}i \cdot \tau*$, appear within the SMART attributes. A probability exceeding 0.5 is regarded as an indication that a change point appears on that day.

Afterward, a position mark-map $\boldsymbol{M} \in \mathbb{R}^{n \times 2m}$, where the dimension equals to $\boldsymbol{D}$, is generated to mark the position of change points in $\boldsymbol{D}$. The mark-map $\boldsymbol{M}$ is a discrete binary matrix, where 1 in $\boldsymbol{M}$ denotes that the corresponding position in $\boldsymbol{D}$ is a change point, while 0 in $\boldsymbol{M}$ denotes that the corresponding position in $\boldsymbol{D}$ is not a change point. Furthermore, the position mark-map $\boldsymbol{M} \in \mathbb{R}^{n \times 2m}$ is split into $\boldsymbol{M}^{h-} \in \mathbb{R}^{m \times n}$ and $\boldsymbol{M}^{h+} \in \mathbb{R}^{m \times n}$ as a part of the input to the sub-networks. The information of change points within mark-maps $\boldsymbol{M}^{h-}$ and $\boldsymbol{M}^{h+}$ could be

Authorized licensed use limited to: Nanjing Univ of Post & Telecommunications. Downloaded on March 09,2025 at 02:06:28 UTC from IEEE Xplore. Restrictions apply.

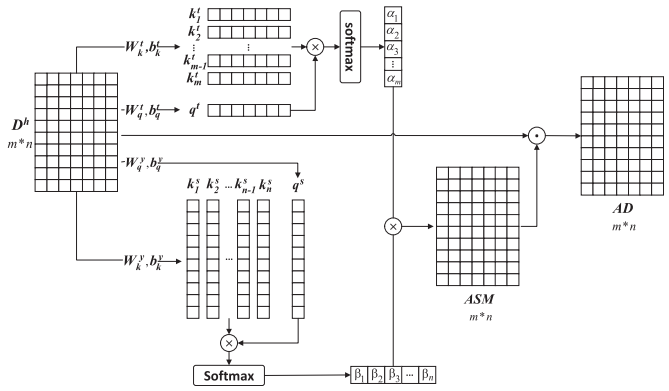


Fig. 11. Principle of 2D-Attention.

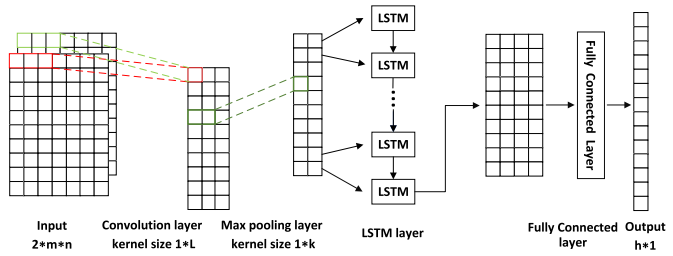


Fig. 12. Structure of feature integration component.

further embedded into final feature represents of disk via the sub-networks.

B. 2D-Attention Component

As mentioned in III, the significance of features varies across different time for a disk, as well as across different SMART attributes of a disk. However, the attention mechanism employed in [19] only focuses on capturing the importance of features at different time (i.e., the feature in different days). To effectively capture the significance of features across different time and different SMART attributes, the 2D-Attention mechanism is proposed. Fig. 11 illustrates the architecture of the 2D-Attention mechanism. In the time domain, we implement the attention mechanism [35] to distinguish the importance of SMART attributes in different days:

$$\begin{aligned} \boldsymbol{q^t} &= tanh(\boldsymbol{D}^h \cdot \boldsymbol{W_q^t} + \boldsymbol{b_q^t}), \\ \boldsymbol{k_j^t} &= tanh(\boldsymbol{W_k^t} \cdot \boldsymbol{D}_{:j}^h + \boldsymbol{b_k^t}), \\ \alpha_j &= \frac{exp(\boldsymbol{q^{t\mathrm{T}}} \cdot \boldsymbol{k_j^t})}{\sum_j^m exp(\boldsymbol{q^{t\mathrm{T}}} \cdot \boldsymbol{k_j^t})}, \end{aligned} \tag{3}$$

where $\boldsymbol{W_q^t} \in \mathbb{R}^{m\times 1}$, $\boldsymbol{b_q^t} \in \mathbb{R}^{n\times 1}$, $\boldsymbol{W_k^t} \in \mathbb{R}^{n\times n}$, $\boldsymbol{b_q^t} \in \mathbb{R}^{n\times 1}$ are the model parameters and $j \in (0, m]$. $\boldsymbol{D}^h$, the SMART attributes of a disk during the half of the time, is first transformed into a high-level context vector $\boldsymbol{q^t}$, and $\boldsymbol{D}_{:j}^h$ is transformed into vector representation $\boldsymbol{k_j^t}$ in the time domain. Then, we calculate the relation between $\boldsymbol{k_j^t}$ and the context vector $\boldsymbol{q^t}$ and obtain the importance of $\boldsymbol{k_j^t}$ based on a softmax function. The attention mechanism generates a weight distribution $\boldsymbol{\alpha}$ for days. Intuitively, α_j denotes the importance of a day for the disk failure prediction. In the SMART attributes domain, we implement the other attention mechanism to distinguish the importance of each SMART attribute for the disk failure prediction:

$$\begin{aligned} \boldsymbol{q^s} &= tanh(\boldsymbol{W_q^s} \cdot \boldsymbol{D}^h + \boldsymbol{b_q^s}), \\ \boldsymbol{k_i^s} &= tanh(\boldsymbol{D}_{i:}^h \cdot \boldsymbol{W_k^s} + \boldsymbol{b_k^s}), \\ \beta_i &= \frac{exp(\boldsymbol{k_i^s} \cdot \boldsymbol{q^{s\mathrm{T}}})}{\sum_i^n exp(\boldsymbol{k_i^s} \cdot \boldsymbol{q^{s\mathrm{T}}})}, \end{aligned} \tag{4}$$

where $\boldsymbol{W_q^s} \in \mathbb{R}^{1\times n}$, $\boldsymbol{b_q^s} \in \mathbb{R}^{1\times m}$, $\boldsymbol{W_k^s} \in \mathbb{R}^{m\times m}$, $\boldsymbol{b_q^s} \in \mathbb{R}^{1\times m}$ are model parameters to learn and $i \in (0, n]$. $\boldsymbol{q^s}$ is a context vector of SMART data, and $\boldsymbol{k_i^s}$ is a vector representation of each SMART attribute in SMART attributes domain. We calculate the relation of $\boldsymbol{k_i^s}$ with $\boldsymbol{q^s}$, then obtain the weight of the $\boldsymbol{k_i^s}$ through a softmax function. The attention mechanism generates a weight distribution $\boldsymbol{\beta}$ for each SMART attribute, and β_i denotes the importance of a SMART attribute. Finally, we represent the feature of the disk data with the feature map $\boldsymbol{AD}^h$ which the abnormal feature is enhanced through the (5):

$$\begin{aligned} \boldsymbol{ASM}^h &= \boldsymbol{\beta} \cdot \boldsymbol{\alpha}, \\ \boldsymbol{AD}^h &= \boldsymbol{ASM}^h \odot \boldsymbol{D}^h, \end{aligned} \tag{5}$$

where $\boldsymbol{ASM}^h$ decided by the weight vectors in the time domain and SMART attributes domain (e.g., $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$), and denotes the distinct importance of a SMART attributes in a day.

C. Feature Integration Component

The occurrence of change points within the SMART attributes can indicate the abnormal status of a disk. Additionally, the temporal SMART data can reflect changes in SMART attributes over time. Consequently, both change points and temporal information of SMART attributes are crucial features for disk failure prediction. To embedding the feature of change points while capturing the temporal features, we propose the feature integration component, which is combined with the network of CNN and LSTM. Fig. 12 illustrates the architecture of the component. CNN is used to embedding the position information of change points into feature map based on the position mark-map generated by CP-MAP component while extracting important feature of the SMART attributes for the next LSTM. In detail, we first concatenate the position mark-map $\boldsymbol{M}^h$ with $\boldsymbol{AD}^h$ into a two-channel map $\boldsymbol{MAD}^h \in \mathbb{R}^{2\times m\times n}$. Since $\boldsymbol{M}^h$ could mark the position of change points in $\boldsymbol{AD}^h$, the convolutional layer of CNN is able to embedding the position information of change points into feature map with the convolution calculation for the two-channel map $\boldsymbol{MAD}^h$. Subsequently, a max-pooling layer is applied to select important feature and reduce computational complexity for subsequent layers. In addition, LSTM is implemented to capture changes in the SMART attributes by learning temporal features. Finally, a fully connected layer generates an integrated feature representation of the disk.

Authorized licensed use limited to: Nanjing Univ of Post & Telecommunications. Downloaded on March 09,2025 at 02:06:28 UTC from IEEE Xplore. Restrictions apply.

Algorithm 1: Pseudocode of SiaDFP.

Input: A disk $D \in \mathbb{R}^{n \times 2m}$
Initialization:
$\boldsymbol{W_q^t} \in \mathbb{R}^{m \times 1}$, $\boldsymbol{b_q^t} \in \mathbb{R}^{n \times 1}$, $\boldsymbol{W_k^t} \in \mathbb{R}^{n \times n}$, $\boldsymbol{b_q^t} \in \mathbb{R}^{n \times 1}$,
$\boldsymbol{W_q^s} \in \mathbb{R}^{1 \times n}$, $\boldsymbol{b_q^s} \in \mathbb{R}^{1 \times m}$, $\boldsymbol{W_k^s} \in \mathbb{R}^{m \times m}$, $\boldsymbol{b_q^s} \in \mathbb{R}^{1 \times m}$,
$pre\text{-}set\ threshold$
Define:
CPD(x): Detect the change points position of x;
Ecuclidean(x_1, x_2): Caculate the distance between x_1 and x_2;
Output: The status of the disk
function CP-MAP($\boldsymbol{D}$, h)
  Pos = CPD($\boldsymbol{D}$) // Pos is the position of the change points
  Transfer Pos to mark-map // mark-map is a binary matrix
  Split mark-map into $\boldsymbol{M}^{h-}$ and $\boldsymbol{M}^{h+}$ based on time
  return $\boldsymbol{M}^h$
  end function
function 2D-Attention($\boldsymbol{D}^h$)
  Caculate $\boldsymbol{AD}^h$ by (3), (4), and (5)
  return $\boldsymbol{AD}^h$
end function
split $\boldsymbol{D}$ into $\boldsymbol{D}^{h-}$ and $\boldsymbol{D}^{h+}$ based on time
for each $h \in (h-, h+)$ **do**
  $\boldsymbol{M}^h$ = CP-MAP($\boldsymbol{D}$, h);
  $\boldsymbol{AD}^h$ = 2D-Attenion($\boldsymbol{D}^h$);
  $\boldsymbol{MAD}^h$ = Concatenate $\boldsymbol{AD}^h$ and $\boldsymbol{M}^h$;
  $\boldsymbol{output}^h$ = Integrate feature of $\boldsymbol{MAD}^h$
end for
Distance = euclidean($\boldsymbol{output}^{h-}$, $\boldsymbol{output}^{h+}$)
if $Distance > pre\text{-}set\ threshold$ **then**
  return Failure
else
  return Healthy
end if

D. Model Training and Prediction

The prediction process of **SiaDFP** is shown in Algorithm 1. We implement the contrastive loss function described in (6) to magnify the distribution offset of failed disks while mitigating that of healthy disks.

$$\begin{aligned} D\text{-}O &= ||Sub\text{-}Network\text{-}II(\boldsymbol{D}^{h+}) \\ &\quad - Sub\text{-}Network\text{-}I(\boldsymbol{D}^{h-})||, \\ Loss(\boldsymbol{D}^{h-}, \boldsymbol{D}^{h+}, Y) &= \frac{1}{2}(1-Y)D\text{-}O^2 \\ &\quad + \frac{1}{2}Y\left\{max(margin - D\text{-}O, 0)\right\}^2. \end{aligned} \tag{6}$$

Y is the label of the disk, where 0 represents a failed disk and 1 represents a healthy disk. $margin$ is a constant to limit the range of offset. In addition, The parameters of **SiaDFP** are updated through the back-propagation algorithm during the training process.

TABLE I
OVERVIEW OF DATASET

Dataset	Model	Original		Post-Processing	
		#F	#H	#P	#N
Dataset-1	Seagate ST4000DM000	940	34738	6128	7258
Dataset-2	Seagate ST8000DM002	103	9888	504	794
Dataset-3	Seagate ST31000524NS	433	22962	4337	5826

V. EXPERIMENTS

In this section, experiment evaluation are conducted on the real-world dataset Backblaze and Baidu to verify the effectiveness of the **SiaDFP**,[1] and mainly discuss the following issues:

1) *Performance of Disk Failure Prediction:* How does the performance of **SiaDFP** compare to the performance of baselines?
2) *Component Exploration:* Whether the 2D-Attention and CP-MAP are effective for the disk failure prediction?
3) *Case Study:* Why **SiaDFP** is effective for the disk failure predicion. Why does SiaDFP not predict the failure of minority samples?

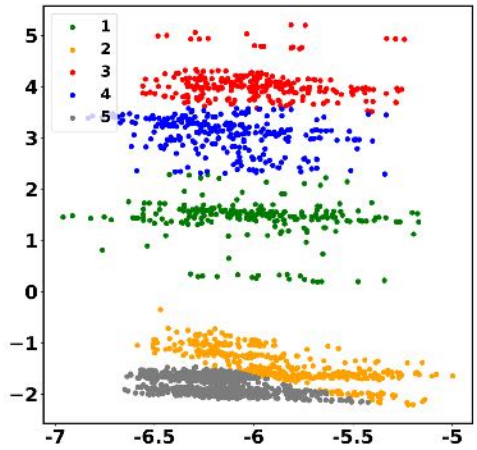
A. Experimental Setting

1) Data Processing: The experimental evaluation is conducted based on the Backblaze dataset and Baidu dataset, including the disk models Seagate ST4000DM000, Seagate ST8000DM002 and Seagate ST31000524NS.[2] Table I shows the overview of the datasets. As observed from Table I, the number of healthy disks (denoted as '**#H**') is greater than that of the failed disks (denoted as '**#F**'), which causes the severe data imbalance. To address this issue, we employ an upsampling method TPS [20] to augment the number of failed samples from the failed disks, using a sliding time window within the fixed size. The *leading time* in TPS is a period between the occurrence time of prediction action and the occurrence time of the disk failure. The failed samples are generated as the time window sliding across the *leading time* of failed disks. Additionally, we obtain the healthy samples by the time window sliding across the lifetime of the healthy disks.

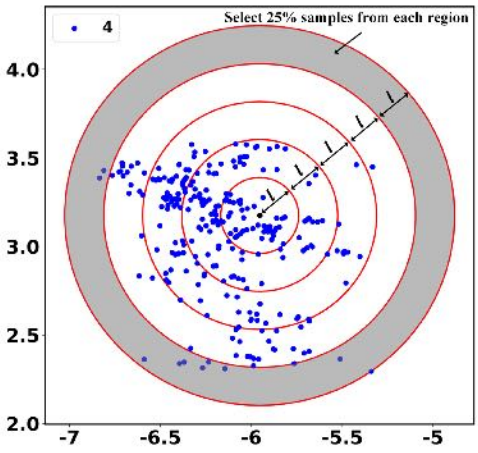
To collect the representative healthy samples for further balancing the number of healthy samples and failed samples, a subsampling method called Region Balanced Sampling (RBS) is proposed. Fig. 13 illustrates the design of RBS. Initially, we employ the K-Means clustering algorithm to partition the healthy samples into five clusters. To select the representative samples, each cluster is further divided into five regions using the contour line. The contour lines are generated by the distance l, where l represents one-fifth of the distance from the centroid to the furthest point as shown in Fig. 13. From each region, 25% of the samples are selected to construct the set of healthy

[1] https://github.com/coderxor/SiaDFP.git
[2] Comparative experiments involving additional disk models can be found in Appendix B, available online

Authorized licensed use limited to: Nanjing Univ of Post & Telecommunications. Downloaded on March 09,2025 at 02:06:28 UTC from IEEE Xplore. Restrictions apply.



(a) Distribution of healthy disk samples in Dataset-1, and samples are split into 5 clusters based on the K-Means clustering algorithm.



(b) Processing of sample selection. Each cluster is split into 5 regions, and 25% samples are selected from each region.

Fig. 13. Processing of the RBS.

samples. After the process of upsampling and subsampling, Dataset-1 contains 6128 failed samples (denoted as '**#P**') and 7258 healthy samples (denoted as '**#N**'). Dataset-2 contains 504 failed samples and 794 healthy samples. Dataset-3 contains 4337 failed samples and 5826 healthy samples. The label of **#P** is 0, and that of **#N** is 1. In addition, we randomly selected 80% of the failed samples and healthy samples as the training samples and the remaining samples as the testing samples.[3] Moreover, not all SMART attributes contribute significantly to model training for disk failure prediction. For Dataset-1 and Dataset-2, we employed a selection criterion based on the presence of distribution offset in SMART attributes leading up to disk failure. For Dataset-3, which comprises only 12 SMART attributes, all attributes were included in model training due to the limited number available. This strategic selection process ensures that the model focuses on the most relevant attributes, optimizing its predictive capabilities. Furthermore, since the range of values in different SMART attributes varies widely, we normalize the value of SMART attributes to avoid bias according to (7):

$$x = \frac{x - x_{min}}{x_{max} - x_{min}}, \tag{7}$$

where x is the original value of a SMART attribute. x_{max} is the maximum value of the SMART attribute among the training samples, and x_{min} is the minimum value of the SMART attribute among training samples.

2) Baseline: We compared the proposed model with several state-of-the-art baselines in disk failure prediction to evaluate the effectiveness of the **SiaDFP**. These baselines can be divided into two categories: **SVM** [36], **DT** [13], **RF** [14], **GBDT** [37], **RGF** [21] are based on traditional machine learning-based methods, and **LSTM** [17], **CNN+LSTM** [18], **AMANDA** [19], **NTAM** [20] are based on deep learning methods.

B. Performance of Disk Failure Prediction

Table II shows the performance of baselines and **SiaDFP** on Dataset-1, Dataset-2 and Dataset-3. As observed from Table II that the performance of **GBDT** exceeds other traditional machine learning based models on Dataset-1, while the **RF** performs well on Dataset-2. In addition, deep learning models (i.e., **LSTM**, **CNN+LSTM**, **NTAM**, **AMANDA**, **SiaDFP**) significantly outperform those of traditional machine learning-based models on Dataset-1. However, except for the **SiaDFP**, the other deep learning models perform worse than those traditional machine learning-based methods on Dataset-2. The better performance of **SiaDFP** on Dataset-1 and Dataset-2 demonstrates that the mechanism of extracting features from disk series with 2D-Attention and CP-MAP in the model could leverage the temporal information of disks and effectively capture the characteristics of disk failure, even on a small-scale dataset.

To efficiently apply the models in real data center scenarios, we explored the overhead of the models about model size and inference time. The number of model parameters is implemented to denote the model size, and the time taken to predict a sample is implemented to denote the inference time. The exploration was conducted on a server with 32 GB RAM and two 10 GB NVIDIA GeForce GTX 1080 Ti GPUs. Table III shows the overhead of the baselines and the proposed model in Dataset-1. In contrast to traditional machine learning models, deep learning models demonstrate superior performance but often come with larger model sizes and longer inference times. Notably, **SiaDFP** excels in performance across various datasets, despite its larger model size and longer inference times compared to other models.

C. Component Exploration

1) CP-MAP vs 2D-Attention: Both CP-MAP and 2D-Attention component are applied to mine the abnormal information from disk data. CP-MAP utilizes expert knowledge-based extraction methods, capturing change points using Bayesian Change Point Detection Algorithm. Different from CP-MAP, 2D-Attention relies on learning-based extraction, utilizing well-trained deep neural networks to capture abnormal information. In this section, a series of experiments are conducted to explore the effectiveness of CP-MAP and 2D-Attention.

[3]The link to the training samples and test samples used in this paper

Authorized licensed use limited to: Nanjing Univ of Post & Telecommunications. Downloaded on March 09,2025 at 02:06:28 UTC from IEEE Xplore. Restrictions apply.

TABLE II
PERFORMANCE OF BASELINES AND **SIADFP**

CATEGORY	MODEL	Dataset-1				Dataset-2				Dataset-3			
		Accuracy	Recall	Precision	F1	Accuracy	Recall	Precision	F1	Accuracy	Recall	Precision	F1
Machine Learning	SVM	0.5560	**0.9900**	0.5304	0.6908	0.6779	**0.8724**	0.6018	0.7123	0.9674	0.9639	0.9660	0.9649
	DT	0.4420	0.4270	0.4130	0.4070	0.7331	0.7852	0.6802	0.7289	**0.9758**	**0.9672**	0.9805	**0.9738**
	RF	0.6770	0.5608	**0.7317**	0.6350	**0.8527**	0.6778	**1.0000**	**0.8080**	0.9737	0.9503	**0.9929**	0.9712
	GBDT	**0.6990**	0.6846	0.6990	**0.6950**	0.8190	0.6174	0.9787	0.7572	0.9695	0.9526	0.9814	0.9668
	RGF	0.6790	0.6606	0.6867	0.6734	0.7116	0.5033	0.7894	0.6147	0.9690	0.9402	0.9928	0.9658
Deep Learning	LSTM	0.7902	0.5143	0.8626	0.6444	0.8296	0.0955	**1.0000**	0.1744	0.4650	**1.0000**	0.4650	0.6348
	LSTM+1D-Attention	0.9047	0.8943	0.8885	0.8680	**0.8476**	0.2058	0.9333	0.3373	0.4640	0.9966	0.4644	0.6336
	LSTM+2D-Attention	**0.9585**	**0.9648**	**0.9261**	**0.9450**	0.8116	**0.4191**	0.9047	**0.5728**	**0.8477**	0.5921	**0.9931**	**0.74191**
	CNN+LSTM	**0.9794**	0.9574	0.9865	0.9717	0.8587	0.2573	0.9722	0.4069	0.6429	0.8939	0.5107	0.6500
	CNN+LSTM+CP-MAP	0.9574	**0.9864**	0.9717	**0.9794**	0.8421	0.1691	0.9583	0.2875	**0.8127**	0.5154	**0.9619**	**0.6712**
	CNN+LSTM+1D-Attention	0.9607	0.8943	0.9990	0.9438	0.8739	0.3676	0.9090	0.5235	0.4661	**0.9988**	0.4655	0.6350
	CNN+LSTM+2D-Attention	0.9533	0.9633	**1.0000**	0.9761	**0.8836**	**0.3897**	**0.9814**	**0.5578**	0.4640	0.9966	0.4644	0.6336
	BaseNTAM	0.7206	0.5094	0.5765	0.4698	0.8988	0.4632	**1.0000**	0.6331	0.8995	0.7346	0.9911	0.8438
	NTAM	0.9630	0.9287	0.9700	0.9489	0.9127	0.5514	0.9740	0.7042	0.9185	0.7821	**0.9968**	0.8765
	BaseNTAM+2D-Attention	**0.9858**	**0.9631**	**0.9983**	**0.9804**	**0.9210**	**0.5808**	**1.0000**	**0.7348**	**0.9915**	**0.9561**	0.9606	**0.9232**
	BaseAMANDA	0.7206	0.3350	0.7865	0.4698	0.8545	0.2352	0.9696	0.3786	0.9825	0.8982	0.9128	0.8273
	AMANDA	**0.9930**	**0.9934**	0.9877	**0.9906**	0.9030	0.5000	**0.9714**	0.6601	0.9897	**0.9901**	0.9821	0.9861
	BaseAMANDA+2D-Attention	0.9915	0.9828	**0.9942**	0.9885	**0.9196**	**0.6838**	0.8611	**0.7622**	**0.9936**	0.9844	**0.9983**	**0.9913**
	BaseSiaDFP	0.7542	0.9885	0.6020	0.7483	0.9376	0.6764	**0.9892**	0.8034	0.9811	0.9616	0.9976	0.9793
	BaseSiaDFP+CP-MAP	0.7633	0.8272	0.7863	0.8062	0.9376	0.6911	0.9690	0.8068	0.9811	0.9706	0.9885	0.9794
	BaseSiaDFP+1D-Attention	0.9631	0.9140	0.9850	0.9482	0.9584	0.7941	0.9818	0.8780	0.9826	0.9729	0.9896	0.9812
	BaseSiaDFP+2D-Attention	0.9851	0.9799	0.9970	0.9884	0.9626	0.8161	0.9823	0.8915	0.9837	0.9762	0.9885	0.9823
	SiaDFP	**0.9954**	**0.9901**	**0.9975**	**0.9938**	**0.9695**	**0.8529**	0.9830	**0.9133**	**0.9963**	**0.9954**	**0.9966**	**0.9960**

TABLE III
OVERHEAD OF MODELS

MODEL	MODEL SIZE (/mb)	INFERENCE TIME (/s)
SVM	3.73e-4	0.38e-4
DT	1.81e-5	0.54e-4
RF	0.37e-1	10.03e-4
GBDT	0.17e-1	0.40e-4
RGF	0.15e-1	9.04e-4
LSTM	0.58e-1	1.72e-4
CNN+LSTM	1.81e-1	1.29e-4
NTAM	0.47e-1	1.13e-4
AMANDA	4.27e-1	2.56e-4
SPA	7.60e-1	3.34e-4
SiaDFP	5.12e-1	2.97e-4

Table II shows the performance of baselines and **BaseSiaDFP** combined with CP-MAP and 2D-Attention. **BaseSiaDFP** is the model which only contains the feature integration component where the 2D-Attention component and the CP-MAP component are removed. **BaseSiaDFP+2D-Attention** is the model which combines the feature integration component with the 2D-Attention component. **BaseSiaDFP+CP-MAP** is the model which combines the feature integration component with the CP-MAP component.

As observed from Table II, both of **BaseSiaDFP+CP-MAP** and **BaseSiaDFP+2D-Attention** outperform **BaseSiaDFP** on all three datasets. The phenomenon indicates that the abnormal information captured by CP-MAP and 2D-Attention effectively promotes the model to learn the characteristics of disk failure. In addition, **BaseSiaDFP+2D-Attention** performs better than the **BaseSiaDFP+CP-MAP** on all three datasets. As shown in Fig. 14, there are some intersections of the important feature captured by 2D-Attention and the change point captured by CP-MAP, where 2D-Attention is capable of capturing more crucial information (i.e., the position of deep yellow) which stably improves the performance of models. Consequently, the effectiveness of the learning-based extraction method, 2D-Attention, surpasses that of the expert knowledge-based extraction method, CP-MAP, when mining abnormal information for disk failure prediction.

To further compare the effectiveness of CP-MAP and 2D-Attention, we combine the baselines with the components to obtain the integrated models. As CNN is essential for CP-MAP, we conduct the comparative experiment based on **CNN+LSTM**. Specifically, **CNN+LSTM+CP-MAP** is the model which combined **CNN+LSTM** with CP-MAP, while **CNN+LSTM+2D-Attention** is the model which combined the **CNN+LSTM** with 2D-Attention. As observed from Table II, the performance of **CNN+LSTM+2D-Attention** exceeds the performance of **CNN+LSTM+CP-MAP**. The phenomenon further illustrates

Authorized licensed use limited to: Nanjing Univ of Post & Telecommunications. Downloaded on March 09,2025 at 02:06:28 UTC from IEEE Xplore. Restrictions apply.

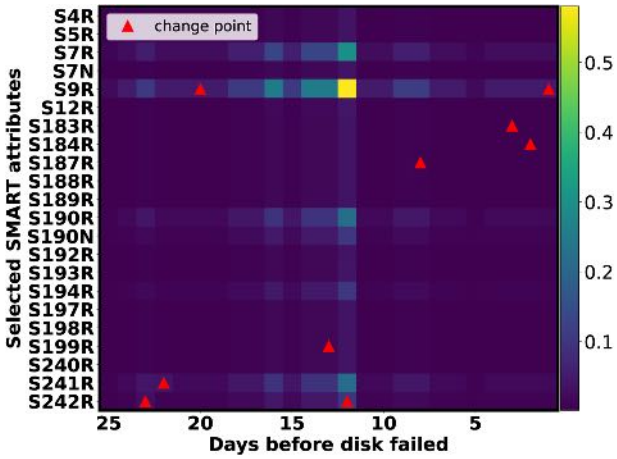
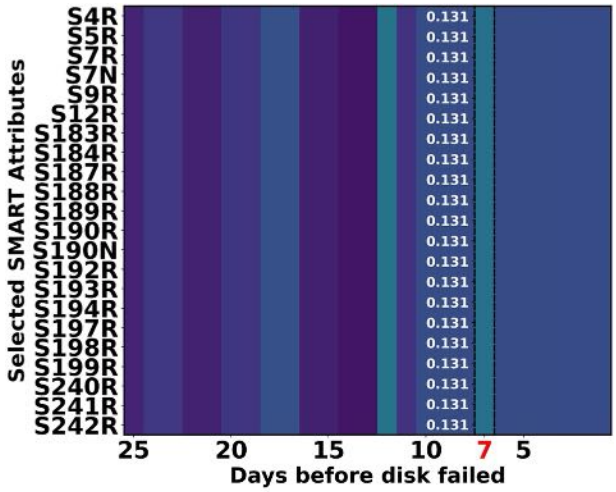


Fig. 14. Score map generated through the 2D-Attention component of **BaseSiaDFP+2D-Attention** trained by Dataset-1 (i.e., disk Z300GZ6V). The position with deep yellow represents that the corresponding SMART attribute in disk Z300GZ6V is essential. The red triangles in the score map mark the positions of change points which are captured by the CP-MAP component in **BaseSiaDFP+CP-MAP** on Dataset-1.

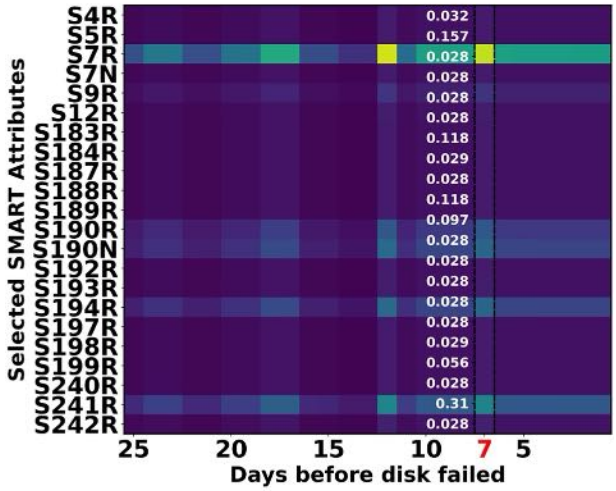
that 2D-Attention is superior to CP-MAP in promoting performance of the model on disk failure prediction. Moreover, as observed from Table II, **SiaDFP**, which consists of the CP-MAP component and the 2D-Attention component, outperforms those models that include only one of them. The phenomenon indicates that the change point captured by CP-MAP complement supplement the important features extracted by 2D-Attention for mining the abnormal information from disk data. Therefore, the combination of the CP-MAP mechanism and the 2D-Attention mechanism could mine the more comprehensive abnormal information of disks and exhibit excellent performance in disk failure prediction.

2) 2D-Attention vs 1D-Attention: To compare the effectiveness of 2D-Attention and 1D-Attention (i.e., traditional attention mechanism [35]), we combine the baselines and **BaseSiaDFP** with the 2D-Attention and 1D-Attention components. Table II shows the performance of these integrated models. For **LSTM** and **CNN+LSTM**, the 2D-Attention and 1D-Attention components are integrated to the original models. In addition, 1D-Attention is utilized in the original **NTAM** and **AMANDA**. Therefore, for **NTAM** and **AMANDA**, we first remove 1D-Attention from them to obtain the basic models (i.e., **BaseNTAM** and **BaseAMANDA**) and then integrate 2D-Attention with the basic models to obtain the combined models.

As observed from Table II, the models combined with 2D-Attention outperform the corresponding models combined with 1D-Attention and the corresponding basic models. 1D-Attention exclusively focuses on capturing the feature importance in the time dimension. However, besides the time dimension, 2D-Attention could capture the feature importance for various SMART attributes. As shown in Fig. 15, 2D-Attention assigns different importance score to various SMART attributes on a specific day, while 1D-Attention assigns equal importance weight to all SMART attributes on that day. Compared to 1D-Attention, the feature of each SMART attribution extracted by



(a) Score map generated by 1D-Attention.



(b) Score map generated by 2D-Attention.

Fig. 15. Score map generated by 1D-Attention and 2D-Attention in Dataset-1.

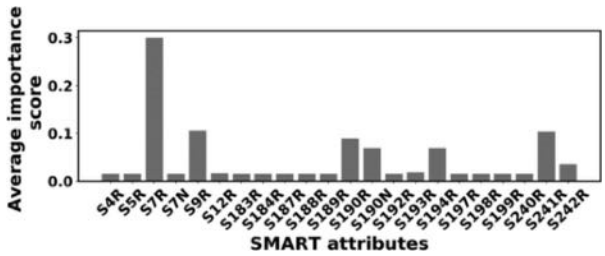


Fig. 16. Average importance score generated by 2D-Attention for SMART attributes in Dataset-1.

2D-Attention are more refined and precise for enhancing the performance of the model on disk failure prediction. Furthermore, to identify the most crucial SMART attribute for predicting disk failures, which can assist data center administrators in diagnosing disk health, we count the average importance score generated by 2D-Attention for each SMART attributes. As shown in Fig. 16, SMART_7_RAW (Seek Error Rate) emerges as the most critical attribute concerning disk failure prediction. This finding

Authorized licensed use limited to: Nanjing Univ of Post & Telecommunications. Downloaded on March 09,2025 at 02:06:28 UTC from IEEE Xplore. Restrictions apply.

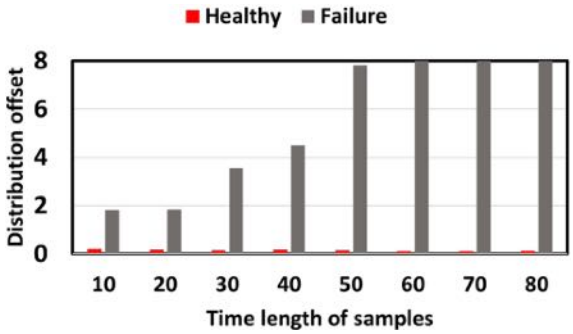


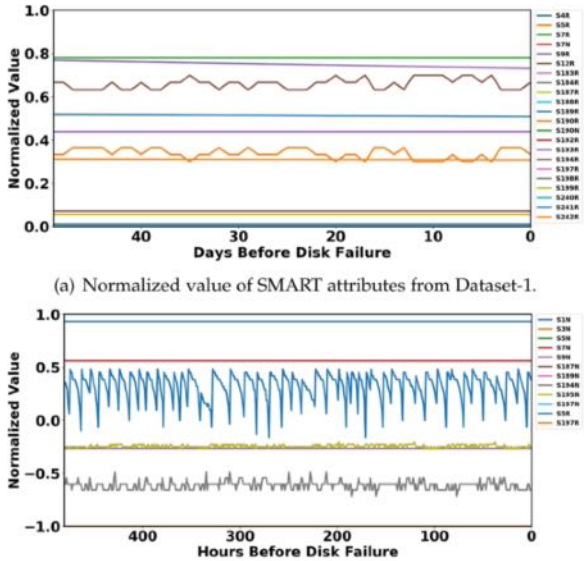
Fig. 17. Distribution offset of healthy samples and failed samples.

provides valuable guidance for administrators when monitoring disk health and making predictions regarding potential failures.

D. Case Study

In the comparative experiments and the component exploration experiments, **SiaDFP** exhibit a standout performance for the disk failure prediction when compared to other baselines. In this section, we conduct series of analysis to explore the reason for **SiaDFP** effectiveness and the limitation of **SiaDFP**. As demonstrated in Fig. 9 within the Section III (i.e., Data Investigation & Analysis), the distinction between the distribution offset of healthy samples and failed samples is present but not readily discernible (i.e., the rate of distribution offset in failed samples compared to that in healthy samples is slightly greater than 1). However, after the healthy samples and failed samples being extracted feature by **SiaDFP**, the distribution offset of the two samples exhibit obvious disparity as shown in Fig. 17. This disparity allows for the clear differentiation of the two samples. The result demonstrates that distribution offset is crucial abnormal feature for the disk failure prediction, while **SiaDFP** could effectively capture the abnormal feature based on the Siamese neural network. Additionally, the component exploration experiments in Section V-C underscores the effectiveness of both the 2D-Attention mechanism and the CP-MAP mechanism on the other abnormal features for disk failure prediction.

However, it is worth noting that the performance of **SiaDFP** falls short of achieving perfection across various metrics (such as Accuracy, Recall, Precision, and F1) in the three datasets. The best levels of **SiaDFP** in these metrics are attained at 99% in Dataset-1 and Dataset-3, still with a 1% gap remaining from achieving a perfect 100%. To understand the reasons behind this limitation, we conducted an analysis of the data from the disks in the datasets that hinder the performance of **SiaDFP**. Fig. 18 illustrates the normalized value of SMART attributes in the failed disks from Dataset-1 and Dataset-3 that **SiaDFP** cannot correctly distinguish. We can learn that the value of SMART attributes remain stable without any changes before the disks failed. In other words, the disks experience abrupt failure without exhibiting any discernible signs, rendering **SiaDFP** unable to effectively capture any distribution offset from these disks. Consequently, one limitation of **SiaDFP** is its inability



(a) Normalized value of SMART attributes from Dataset-1.

(b) Normalized value of SMART attributes from Dataset-3.

Fig. 18. Normalized value of SMART attributes.

to distinguish disks that experience sudden and unanticipated failures.

VI. Conclusion

In this paper, we dive into the research of disk failure prediction and propose a novel prediction framework **SiaDFP** based on the Siamese neural network, which contains a CP-MAP component and two parallel sub-networks. Each sub-network consists of a 2D-Attention component and a feature integration component. We conduct the experiments on the real-world dataset Backblaze and Baidu to evaluate **SiaDFP**. The experimental results demonstrate that: (1)CP-MAP and 2D-Attention are capable of mining the abnormal information from time series of disks for the disk failure prediction; (2)Compared to the traditional attention mechanism (i.e., 1D-Attention mechanism), 2D-Attention could capture more representative features based on the information from two dimensions (i.e., the time domain and the SMART attributes domain); (3) The performance of **SiaDFP** which combined with CP-MAP and 2D-Attention could exhibit the standout performance on disk failure prediction. (4) The limitation of **SiaDFP** is that it can not distinguish the failed disks experienced a sudden damage.

References

[1] E. Pinheiro, W.-D. Weber, and L. A. Barroso, "Failure trends in a large disk drive population," in *Proc. 5th USENIX Conf. File Storage Technol.*, 2007, pp. 17–28.

[2] J. Sikora, "Disk failures in the real world: What does an MTTF of 1,000,000 hours mean to you?" in *Proc. 5th USENIX Conf. File Storage Technol.*, 2007, pp. 1–16.

[3] K. V. Vishwanath and N. Nagappan, "Characterizing cloud computing hardware reliability," in *Proc. 1st ACM Symp. Cloud Comput.*, 2010, pp. 193–204.

[4] Y. Wang, Q. Miao, E. W. Ma, K.-L. Tsui, and M. G. Pecht, "Online anomaly detection for hard disk drives based on mahalanobis distance," *IEEE Trans. Rel.*, vol. 62, no. 1, pp. 136–145, Mar. 2013.

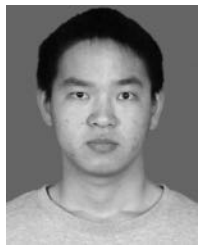
Authorized licensed use limited to: Nanjing Univ of Post & Telecommunications. Downloaded on March 09,2025 at 02:06:28 UTC from IEEE Xplore. Restrictions apply.

[5] W. Jiang, C. Hu, Y. Zhou, and A. Kanevsky, "Are disks the dominant contributor for storage failures? A comprehensive study of storage subsystem failure characteristics," *ACM Trans. Storage*, vol. 4, no. 3, pp. 1–25, 2008.
[6] J. Zhao, S. Xiong, D. B. Bogy, K. Sun, and L. Fang, "Hard-particle-induced physical damage and demagnetization in the head-disk interface," *Microsystem Technol.*, vol. 19, pp. 1313–1317, 2013.
[7] C. Huang et al., "Erasure coding in windows azure storage," in *Proc. USENIX Annu. Tech. Conf.*, 2012, pp. 15–26.
[8] Z. Huang, H. Jiang, K. Zhou, C. Wang, and Y. Zhao, "XI-code: A family of practical lowest density MDS array codes of distance 4," *IEEE Trans. Commun.*, vol. 64, no. 7, pp. 2707–2718, Jul. 2016.
[9] S. Ghemawat, H. Gobioff, and S.-T. Leung, "The google file system," in *Proc. 19th ACM Symp. Operating Syst. Princ.*, 2003, pp. 29–43.
[10] B. Allen, "Monitoring hard disks with smart," *Linux J.*, no. 117, pp. 74–77, 2004.
[11] J. Xiao, Z. Xiong, S. Wu, Y. Yi, H. Jin, and K. Hu, "Disk failure prediction in data centers via online learning," in *Proc. 47th Int. Conf. Parallel Process.*, 2018, pp. 1–10.
[12] J. F. Murray, G. F. Hughes, K. Kreutz-Delgado, and D. Schuurmans, "Machine learning methods for predicting failures in hard drives: A multiple-instance application," *J. Mach. Learn. Res.*, vol. 6, no. 5, pp. 783–816, 2005.
[13] J. Li et al., "Hard drive failure prediction using classification and regression trees," in *Proc. IEEE/IFIP 44th Annu. Int. Conf. Dependable Syst. Netw.*, 2014, pp. 383–394.
[14] J. Shen, J. Wan, S.-J. Lim, and L. Yu, "Random-forest-based failure prediction for hard disk drives," *Int. J. Distrib. Sensor Netw.*, vol. 14, no. 11, 2018, Art. no. 1550147718806480.
[15] I. C. Chaves, M. R. P. De Paula, L. G. Leite, J. P. P. Gomes, and J. C. Machado, "Hard disk drive failure prediction method based on a Bayesian network," in *Proc. IEEE Int. Joint Conf. Neural Netw.*, 2018, pp. 1–7.
[16] C. Xu, G. Wang, X. Liu, D. Guo, and T.-Y. Liu, "Health status assessment and failure prediction for hard drives with recurrent neural networks," *IEEE Trans. Comput.*, vol. 65, no. 11, pp. 3502–3508, Nov. 2016.
[17] J. Zhang, J. Wang, L. He, Z. Li, and S. Y. Philip, "Layerwise perturbation-based adversarial training for hard drive health degree prediction," in *Proc. IEEE Int. Conf. Data Mining*, 2018, pp. 1428–1433.
[18] S. Lu, B. Luo, T. Patel, Y. Yao, D. Tiwari, and W. Shi, "Making disk failure predictions {SMARTer}!," in *Proc. 18th USENIX Conf. File Storage Technol.*, 2020, pp. 151–167.
[19] J. Wang, W. Bao, L. Zheng, X. Zhu, and P. S. Yu, "An attention-augmented deep architecture for hard drive status monitoring in large-scale storage systems," *ACM Trans. Storage*, vol. 15, no. 3, pp. 1–26, 2019.
[20] C. Luo et al., "NTAM: Neighborhood-temporal attention model for disk failure prediction in cloud platforms," in *Proc. Web Conf.*, 2021, pp. 1181–1191.
[21] M. M. Botezatu, I. Giurgiu, J. Bogojeska, and D. Wiesmann, "Predicting disk replacement towards reliable data centers," in *Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discov. Data Mining*, 2016, pp. 39–48.
[22] H. Zhou et al., "Proactive drive failure prediction for cloud storage system through semi-supervised learning," *IEEE Trans. Dependable Secure Comput.*, to be published, doi: 10.1109/TDSC.2023.3286093.
[23] S. Xu and X. Xu, "Convtrans-TPS: A convolutional transformer model for disk failure prediction in large-scale network storage systems," in *Proc. IEEE 26th Int. Conf. Comput. Supported Cooperative Work Des.*, 2023, pp. 1318–1323.
[24] M. Züfle, C. Krupitzer, F. Erhard, J. Grohmann, and S. Kounev, "To fail or not to fail: Predicting hard disk drive failure time windows," in *Proc. Int. Conf. Meas. Modelling Eval. Comput. Syst.*, 2020, pp. 19–36.
[25] R. Lu et al., "Perseus: A {Fail-Slow} detection framework for cloud storage systems," in *Proc. 21st USENIX Conf. File Storage Technol.*, 2023, pp. 49–64.
[26] J. Bromley, I. Guyon, Y. LeCun, E. Säckinger, and R. Shah, "Signature verification using a "siamese" time delay neural network," in *Proc. Adv. Neural Inf. Process. Syst.*, 1993, pp. 737–744.
[27] S. Chopra, R. Hadsell, and Y. LeCun, "Learning a similarity metric discriminatively, with application to face verification," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, 2005, pp. 539–546.
[28] M. Khalil-Hani and L. S. Sung, "A convolutional neural network approach for face verification," in *Proc. IEEE Int. Conf. High Perform. Comput. Simul.*, 2014, pp. 707–714.
[29] J. Mueller and A. Thyagarajan, "Siamese recurrent architectures for learning sentence similarity," in *Proc. AAAI Conf. Artif. Intell.*, 2016, pp. 2786–2792.
[30] N. Bölücü, B. Can, and H. Artuner, "A siamese neural network for learning semantically-informed sentence embeddings," *Expert Syst. Appl.*, vol. 214, 2023, Art. no. 119103.
[31] L. Bertinetto, J. Valmadre, J. F. Henriques, A. Vedaldi, and P. H. Torr, "Fully-convolutional siamese networks for object tracking," in *Proc. Eur. Conf. Comput. Vis.*, 2016, pp. 850–865.
[32] B. Li, J. Yan, W. Wu, Z. Zhu, and X. Hu, "High performance visual tracking with siamese region proposal network," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2018, pp. 8971–8980.
[33] R. P. Adams and D. J. MacKay, "Bayesian online changepoint detection," 2007, *arXiv:0710.3742*.
[34] P. Fearnhead, "Exact and efficient Bayesian inference for multiple changepoint problems," *Statist. Comput.*, vol. 16, no. 2, pp. 203–213, 2006.
[35] A. Vaswani et al., "Attention is all you need," in *Proc. Adv. Neural Inf. Process. Syst.*, 2017, pp. 5998–6008.
[36] B. Zhu, G. Wang, X. Liu, D. Hu, S. Lin, and J. Ma, "Proactive drive failure prediction for large scale storage systems," in *Proc. IEEE 29th Symp. Mass Storage Syst. Technol.*, 2013, pp. 1–5.
[37] X. Huang, "Hard drive failure prediction for large scale storage system," Ph.D. dissertation, UCLA, 2017.

Xiaoyu Fang received BS degree from the Department of Software Engineering, Nanjing University of Posts and Telecommunications, Nanjing, China, in 2020. He is currently working toward the MS degree with the Department of Computer Science, Nanjing University of Posts and Telecommunications, Nanjing, China. His research interests include Deep Learning and AI for IT Operations.

Wenbai Guan received BS degree from the Department of Software Engineering, Nanjing University of Posts and Telecommunications, Nanjing, China, in 2021. He is currently working toward the MS degree with the Department of Computer Science, Nanjing University of Posts and Telecommunications, Nanjing, China. His research interests include Reinforcement learning and AI for IT Operations.

Jiawen Li received BS degree from the Department of Software Engineering, Nanjing University of Posts and Telecommunications, Nanjing, China, in 2022. She is currently working toward the MS degree with the Department of Computer Science, Nanjing University of Posts and Telecommunications, Nanjing, China. Her research interests include AI for IT Operations.



Chenhan Cao received BS degree from the Department of Economics and Management, Nanjing University of Posts and Telecommunications, Nanjing, China, in 2020. He is currently working toward the MS degree with the Department of Computer Science, Nanjing University of Posts and Telecommunications, Nanjing, China. His research interests include Recommender System and AI for IT Operations.

Bin Xia received the PhD degree with the School of Computer Science and Engineering, Nanjing University of Science and Technology, Nanjing, in 2018. He is currently an Associate Professor with the School of Computer Science, Nanjing University of Posts and Telecommunications. He was in the Knowledge Discovery Research Group (KDRG) that is led by Prof. Tao Li. He is leading the X-Lab of AI Operations and Recommender System (XOR). His research area includes Deep Learning, Recommender System, and AI for IT Operations.

Authorized licensed use limited to: Nanjing Univ of Post & Telecommunications. Downloaded on March 09,2025 at 02:06:28 UTC from IEEE Xplore. Restrictions apply.