HAL
open science

Toward Behavioral Modeling of a Grid System: Mining the Logging and Bookkeeping files

Xiangliang Zhang, Michèle Sebag, Cecile Germain-Renaud

To cite this version:

Xiangliang Zhang, Michèle Sebag, Cecile Germain-Renaud. Toward Behavioral Modeling of a Grid System: Mining the Logging and Bookkeeping files. DSMM07, Oct 2007, Omaha, United States. inria-00174285

HAL Id: inria-00174285

https://inria.hal.science/inria-00174285v1

Submitted on 22 Sep 2007

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Toward Behavioral Modeling of a Grid System: Mining the Logging and Bookkeeping files

Xiangliang Zhang, Michèle Sebag, Cécile Germain
INRIA, CNRS, and Université Paris-Sud 11, 91405 Orsay, France
{xlzhang,sebag,cecile}@lri.fr

Abstract

Grid systems are complex heterogeneous systems, and their modeling constitutes a highly challenging goal. This paper is interested in modeling the jobs handled by the EGEE grid, by mining the Logging and Bookkeeping files. The goal is to discover meaningful job clusters, going beyond the coarse categories of "successfully terminated jobs" and "other jobs". The presented approach is a three-step process: i) Data slicing is used to alleviate the job heterogeneity and afford discriminant learning; ii) Constructive induction proceeds by learning discriminant hypotheses from each data slice; iii) Finally, double clustering is used on the representation built by constructive induction; the clusters are fully validated after the stability criteria proposed by Meila (2006). Lastly, the job clusters are submitted to the experts and some meaningful interpretations are found.

1. Introduction

The increasing complexity of computer science systems has brought on new demands for Autonomic Computing systems (*http://www.research.ibm.com/autonomic/*). Concretely, autonomic systems are expected to "manage themselves in accordance with high-level objectives from humans" [1]; they should be able to discover problems, send reports to the system administrator in real-time, and recover from functioning incidents. In order to do so, such systems should be self-aware, i.e. endowed with a model of their normal functioning mode to detect a deviation from its normal functioning mode and/or to anticipate upcoming problems. As behavioral models of complex systems can hardly be based on prior knowledge (see below), Autonomic Computing is becoming a source of challenging applications in Machine Learning and Data Mining (see e.g. [2]).

This paper is concerned with the behavioral modeling of the EGEE grid (*Enabling Grid for E-Science in Europe* [1]). EGEE involves about 20,000 CPUs, 5 Petabytes of storage, and it concurrently handles 20,000 jobs on average. Modeling such a system raises diverse difficulties. The EGEE topology and state is hardly known with certainty at any point in time and its load depends on the collective behavior of the users, reflecting the diverse communities sharing the grid and their agendas.

As a first step toward a behavioral model of the EGEE system, this paper focuses on modeling the jobs handled by EGEE. These jobs are described from the Logging and Bookkeeping (L&B) files recorded by the grid broker. The challenge is to extract clusters of jobs with meaningful interpretations from the rough partition of good ("successfully terminated") and bad (all other cases) jobs.

This goal raises several difficulties besides the data size (5 Gb). Firstly, job traces are represented using a structured representation (Job Description Language) and there is no natural metric on this representation space. Secondly, jobs are highly heterogeneous, with orders of magnitude variations depending on the period of the year and on the users.

The approach presented in this paper is a three-step process, combining constructive induction and clustering. The constructive induction step uses the training set and the clustering step operates on the test set. In the first step, the training set is partitioned into subsets, in order to both i) alleviate the jobs heterogeneity, and ii) afford the use of discriminant learning algorithms. The partitions are based on the users and on the weeks. Every subset is expected to provide a (more) consistent snapshot of the grid use, filtering out the user-related or the period-related aspects. In the second step, every subset undergoes discriminant learning; hypotheses discriminating the good from the bad jobs in the subset are extracted. A stochastic algorithm optimizing the area under the ROC curve called Roger [3] is used. In the third step, the hypotheses are used as new features to redescribe the jobs in the test set. Inspired from [4], the double clustering of the hypotheses and the jobs is achieved.

[1] http://www.eu-egee.org/

The significance of the job clusters is assessed by comparing the clusters independently extracted from the user-based and week-based hypotheses. The clusters are fully validated after the stability criteria proposed by Meila [5]. Finally, these clusters are submitted for interpretation to the experts, and some meaningful interpretations are provided.

The paper is organized as follows. Section 2 briefly reviews the state of the art in dimensionality reduction and clustering. Section 3 describes the data preparation and slicing and the constructive induction process. Section 4 describes the double clustering of the hypotheses and the jobs. Section 5 explains the experimental setting and reports on the results. The paper ends with a discussion and perspectives for further research.

2. State of the art

How to construct a meaningful and tractable representation for a problem domain, referred to as constructive induction, is known to be a key difficulty for Machine Learning. Some paradigms proposed for constructive induction are dimensionality reduction and clustering.

2.1. Dimensionality reduction

Let us make the standard distinction between linear and non linear dimensionality reduction approaches.

Classical approaches in linear dimensionality reduction are Principal Component Analysis (PCA) and Singular Value Decomposition (SVD). The main limitation of these well-established methods in our context is due to the highly heterogeneous distribution of the descriptors in the initial representation of the jobs. Typically, when the values for a numerical attribute differ by orders of magnitude, the mean value provides little information.

Since 2000, several approaches of unsupervised non linear dimensionality reduction have been presented, such as Isomap [6] and Locally Linear Embedding (LLE) [7]. The baseline Isomap and LLE approaches are limited in our context due to their complexity (at least quadratic in the number of examples).

Some extensions of non linear dimensionality reduction to the supervised setting have also been proposed. For instance, Sugiyama [8] proposed a supervised dimension reduction method called local Fisher discriminant analysis (LFDA) which performs well on multimodal problems, combining Fisher discriminant analysis and Locality Preserving Projection. Another approach for Semi-Supervised Nonlinear Dimensionality Reduction, proposed by Yang et al. [9], uses prior knowledge to improve the stability of the solution.

Closely related to dimensionality reduction is distance learning. For instance, Weinberger et al. [10] formalized distance learning as an optimization problem, maximizing (minimizing) the distance of every point to its K nearest neighbors pertaining to different (same) classes; a Mahalanobis distance is determined by solving this optimization problem using semi-definite programming. Again, this approach has quadratic complexity in the number of examples.

2.2. Clustering

The so-called K-means clustering approach relies on a distance function, which is used to partition the space into K subsets; a key question concerns the stability of the clusters, as these usually depend on random initialization. The relationship between PCA and K-means was studied by Ding and He [11], proving that the cluster centers are related to the PCA eigenvectors. Based on this result, under the well-separateness assumption, Meila proved that any good clustering can be approximated (up to a cluster permutation) from the principal components of the data [5]. Accordingly, a bound on the quality and stability of a clustering can be derived from its distance to the PCA eigenvectors.

The well-separateness assumption requires that the data do not live in a manifold of dimension less than $K-1$. Note that, if the assumption does not hold, this will not imply that the quality of K-means clusters will be low; rather, it implies that this quality cannot be assessed after the above bound.

3. Constructive induction with slicing

3.1. Data preparation

The L&B files used in the paper is collected from EGEE Resource Brokers which are being monitored to display the status of all jobs submitted to the grid. The logs are updated in real time and describe the job behavioral traces; the life cycle of every job is represented as a variable-length sequence of events, where the number of events ranges from 1 to 174. The sequences are generated in real-time according to the job executing statuses in the grid system. About 300,000 jobs and 3,300,000 events are represented in the dataset [2]. Events that describe some service achieved for the job are saved using three different tables, the structure of which is represented in Fig. 1.

As widely acknowledged, data preparation is a most time-consuming task. A software suite has been designed to identify functional dependencies, segment compound attributes, parse the strings and more generally filter out redundant information and format every job into a propositional vector.

Numerical attributes are normalized. Prior knowledge is used to parse the categorical attribute values (machine

[2] These will be available soon in public.

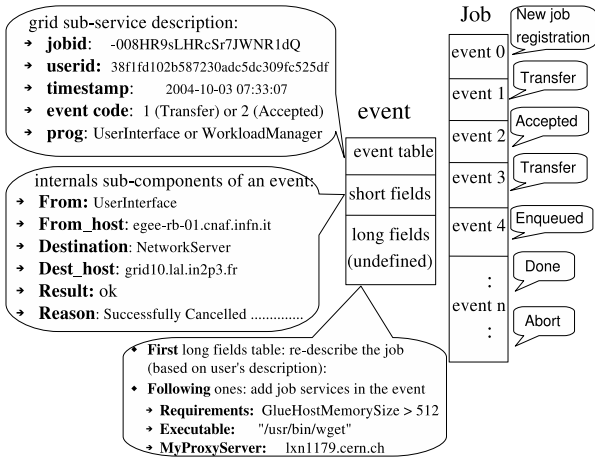


Figure 1. L&B Data structure

names, file names, user logins). Thereafter, every categorical attribute with m values is converted into m boolean attributes; more precisely, sufficiently frequent values derive boolean attributes; other values are merged together as an "other-values" boolean attribute. Eventually, every job is described through 408 numerical and boolean attributes.

3.2. Data slicing

As mentioned in the introduction, the dataset is composed of good jobs (30% of the jobs) and bad jobs. Both classes of jobs are very heterogeneous, for (at least) two reasons. The first type of heterogeneity is due to the users: they differ based on their level of expertise, and their scientific field. The second type of heterogeneity is due to the mutualization paradigm of the EGEE grid: the load of the grid hugely varies along time; typically some weeks are heavily loaded while other weeks are not.

In order to get rid of this heterogeneity, the proposed approach, inspired from [12], proceeds by aggressively subsampling the training dataset (90% of all data). More precisely, job subsets are defined by retrieving: i) all jobs submitted by a given user; ii) all jobs submitted during a given week. The subsets which do not include a sufficient number of jobs or percentage of good/bad jobs were removed.

3.3. Constructive induction

The second step of the presented approach is concerned with constructive induction. Each data subset constructed as above undergoes discriminant learning; the hypotheses learned to discriminate good from bad jobs in the given subset will be used as new features for describing all jobs.

In order to enforce the stability of the discriminant hypotheses, an AUC-based learning approach, optimizing the so-called Wilcoxon-Mann-Whitney statistics [13], was considered.

As shown by [13], the WMW criterion, equivalent to the optimization of the area under the ROC curve (AUC), constitutes a stable learning criterion (e.g. quadratic in the number of examples as opposed to the linear mis-classification rate). Further, AUC-based hypotheses directly provide an estimation of the classification probability.

To find "sufficiently" diverse hypotheses from every data subset for the same reason advocated in ensemble methods, the learning algorithm achieves the stochastic optimization of the AUC criterion, using a $(\lambda + \mu)$ Evolution Strategy termed ROGER (*ROC-based Genetic Learner*) to find linear hypotheses with complexity $\mathcal{O}(n \ln n)$ where n denotes the number of examples [3]. Finally, a set of ℓ hypotheses is obtained from each training subset by independently running the ROGER algorithm ℓ times.

Every such hypothesis, mapping every job onto a real value, defines a new attribute (feature) of the problem domain.

These new features together define a new representation of the problem domain. Finally two representations, referred to as U-representation and W-representation, are considered. They respectively include the set of features (hypotheses) learned from user-based and week-based data subsets.

Formally, let $h_{u,i}$ (respectively $h_{w,i}$) denote the i-th hypothesis extracted by ROGER from the job subset associated to user u (resp. to week w). The U-representation (W-representation) associates to every job $\mathbf{x}$ the real-valued vector $U(\mathbf{x})$ ($W(\mathbf{x})$) defined as $(h_{u,i}(\mathbf{x}))$ $((h_{w,i}(\mathbf{x}))$) for $i = 1 \ldots \ell$ and u (w) varying in the set of users (weeks); let $N'_u = N_u \times \ell$ ($N'_w = N_w \times \ell$) denote the dimensionality of vector $U(\mathbf{x})$ ($W(\mathbf{x})$).

4. Clustering stability and double clustering

4.1. Clustering stability

Following Meila [5], clustering $C = \{C_1, \ldots, C_K\}$ is represented as a $J \times K$ matrix $\widetilde{C}$, where $\widetilde{C}_{i,k}$ is 1 iff the i-th example belongs to C_k and 0 otherwise. Let $\widehat{C}$ be defined by normalizing $\widetilde{C}$; denoting n_k the size of cluster C_k,

$$\widehat{C}_{i,k} = \begin{cases} 1/\sqrt{n_k} & \text{if the } i\text{-th example belongs to } C_k \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

The similarity between two clusterings $\widehat{C}$ and $\widehat{C}'$ defined on the same dataset is computed from the scalar product of $\widehat{C}$ and $\widehat{C}'$:

$$S(\widehat{C}, \widehat{C}') = \|\widehat{C}^T \widehat{C}'\|^2_{Frobenius} = \sum_{i,j=1}^{K} n^2_{i,j} \frac{1}{n_i n'_j} \tag{2}$$

where $n_{i,j}$ is the number of jobs in $C_i \bigcap C'_j$, n_i and n'_j respectively are the size of C_i and C'_j.

Theorem

With above notations, similarity $S(\widehat{C}, \widehat{C}')$ admits a lower bound and an upper bound as follows:

$$K \geq S(\widehat{C}, \widehat{C}') \geq \frac{J}{(J-K+1)} \frac{1}{K} \tag{3}$$

After Meila [5], the upper bound is used to measure the clustering stability and hence its quality: the closer $S(\widehat{C}, \widehat{C}')$ to K, the better $\widehat{C}$ and $\widehat{C'}$ are. However, when the well separateness assumption does not hold, the upper bound is not necessarily reached; in such cases, the lower bound gives an indication about the quality of clusterings.

4.2. Double clustering

The new representations proposed for the jobs suffer from both their high dimensionality (some thousand features) and high redundancy (all features are based on the initial attributes and furthermore subsets of features are learned from the same data subset). Therefore, a double clustering approach inspired from [4] will be used. The double clustering mechanism proceeds as follows (considering independently the U- and the W-representation):

1. Features belonging to the U-representation (W-representation) are clustered using K-means;

2. The examples are clustered using K-means based on the average feature of each feature cluster obtained in the first step.

This procedure eventually defines two clusterings on the training set, respectively based on the U-clustered and W-clustered representation.

The validity of the approach is finally assessed from the stability of the clusters obtained from both representations, as detailed in next section.

5. Experiments

5.1. Experimental settings

The training set is first sliced into single-user and single-week subsets as detailed in Section 3.2. Finally the total number of single-user (respectively single-week) subsets is $N_u = 36$ (resp. $N_w = 47$).

On each subset, ROGER is launched $\ell = 50$ times, using the same parameters over all runs. The number of parents: $\mu = 10$; the number of offspring: $\lambda = 70$. Maximum number of evaluations is 1000.

Table 1. Data used

		N. total	N. training	N. testing
Good jobs		88,131	78,131	10,000
Bad jobs	NAR	117,369	100,000	17,369
	GNG	40,906	36,000	4,906
	ABU	2,561	2,000	561
Total		248,967	216,131	32,836

In this way, $N'_w = 47 \times 50 = 2350$ hypotheses are extracted from single-week datasets and $N'_u = 36 \times 50 = 1800$ hypotheses from single-user datasets.

Finally, the test set is rewritten using the U- and W-representations defined from the ROGER hypotheses, and it undergoes the double clustering(Section 4.2); the number T of feature clusters ranges from 6 to 36 with increment step 2. The number K of example clusters ranges from 4 to 32 with increment step 1.

5.2. Failure types

As already mentioned, about 70% jobs are not terminated successfully[14]. On grid system, there are various factors resulting in the job failures, such as Hardware problem, Misconfiguration, security attacks, mistakes of users. According to the causes of failure, the "bad" jobs are roughly divided into 3 classes:

1. No Adequate Resource(NAR): It happened at the mapping step on Resource Broker. Possible scenarios are "User error" or "True mismatch".

2. Generic and Non Generic errors(GNG): Generic errors which do not refer to specific job. There are many kinds of these errors, such as: "The user proxy expired", "Job size exceeds limits". There are also Non Generic errors which happen accompanied by generic errors, such as: IO problems reported as "Cannot download/upload/retriev".

3. Aborted By User (ABU): the user requests the job to be canceled. Two overlapping cases are "True user reconsideration" and "Grid problem".

The four job categories represented in the dataset are described in Table 1; remind that the learning algorithms are only provided with the "good" and "bad" information.

5.3. Experimental results

Fig. 2 shows the double clustering based on ROGER hypotheses when $T = 30$ and $K = 29$. Most clusters are pure with respect to the four categories. A few clusters present

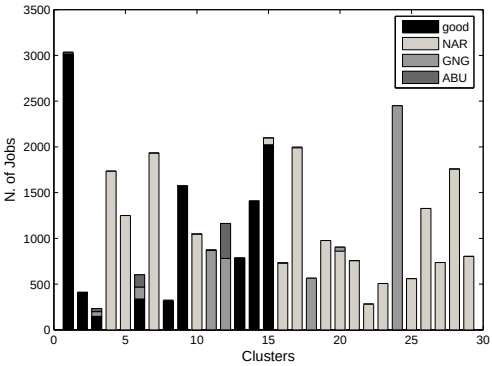


Figure 2. ROGER-based representation: job clusters for $T = 30$ and $K = 29$

a mixture of GNG and ABU jobs. Each cluster is labeled from the majority category in the cluster.

The quality of the clustering is first assessed from its error rate, that is the percentage of the examples which belong to a cluster with different label.

In order to show the influence of the number of example and feature clusters on the error rate, Fig. 3.(a) reports the error rates depending on K , i) averaged over all values of T, ii) obtained for the best T value, considering the U-clustered representation. The error rates obtained by excluding the ABU class are also given. Fig. 3.(b) symmetrically illustrates the error rates depending on T. The error rates of W-clustered representation are similar to those of U-clustered representation.

To avoid reporting too optimistic results (misclassifying all of the ABU jobs does not show on the global error rate), the y axis in Fig. 3 is the average over all four classes, of the percentage of misclassified examples in this class. When discarding the ABU class, the stability of the error is excellent wrt the number K of job clusters as well as the number T of feature clusters.

The second goal of experiments is to assess the clustering stability, measured as follows. Denoting $D(C, C') = \frac{1}{K} S(C, C')$ as defined in Section 4.1, we introduce the self-stability of the U-clustered and W-clustered representation, and the mutual stability between them.

- Self-stability: For each K, its stability index $\overline{D}_K = \frac{1}{|T_i|(|T_i|-1)} \sum_{T_i} \sum_{all T_j \neq T_i} D(C^{KT_i}, C^{KT_j})$, where $|T_i|$ is the number of different T_i. C^{KT_i} is the clustering obtained when the number of example clusters and feature clusters are K and T_i.

- Mutual stability:
  For each K, $\overline{D}'_K = \frac{1}{|T_i|} \sum_{T_i} D(C_w^{KT_i}, C_u^{KT_i})$.

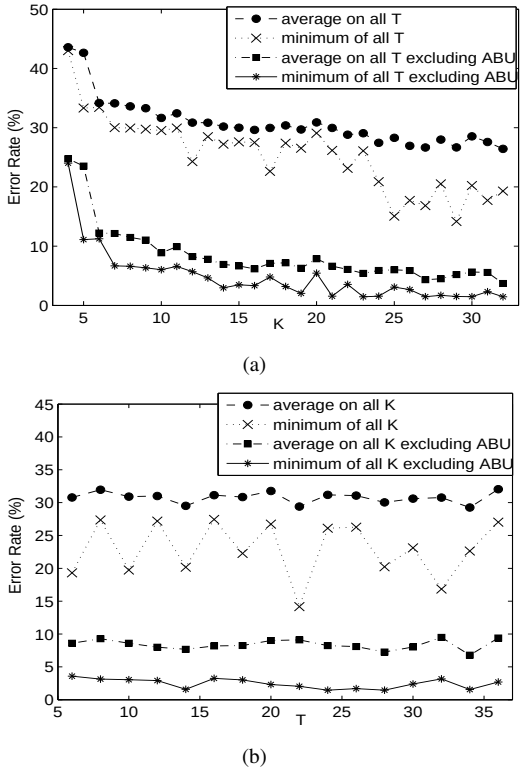


Figure 3. Error rate: influence of K (a) and T (b) considering U-clustered representation

For each T, $\overline{D}'_T = \frac{1}{|K_i|} \sum_{K_i} D(C_w^{K_i T}, C_u^{K_i T})$.

The above stability indices are displayed in Fig. 4.

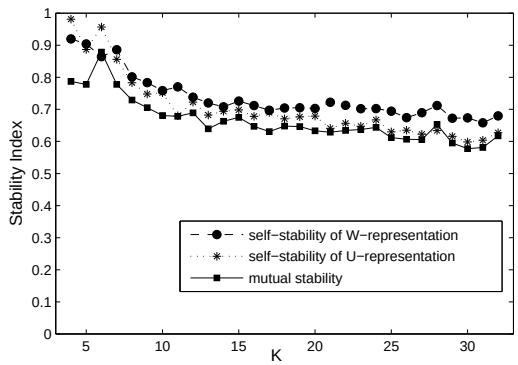
6. Concluding remarks

The above results raise the following remarks.

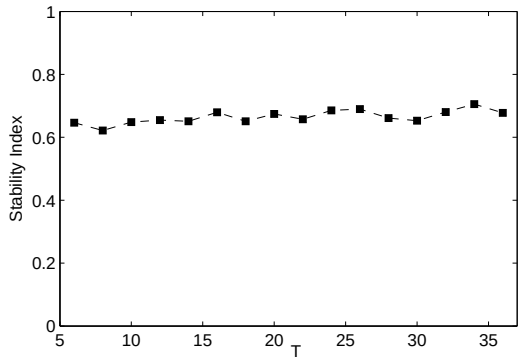
With respect to the classification accuracy, it is shown (Fig. 2 and Fig. 3) that the clusters obtained after the presented approach successfully identify the job categories in the test set, complying with the given types of failures (NAR, ABU and GNG) and even suggesting refinements of these categories.

Most generally (Fig.3 (a)), the error rate is very low in all categories, particularly so for a high number of clusters ($K > 20$), except for the ABU category, which appears to be rather polymorphic. Indeed there are many diverse reasons why the user might want to abort a job. In the case of the double clustering, the influence of T appears to be quite limited (Fig.3 (b)) in the considered range.

The cluster stability is excellent for small values of K ($K = 6$) and it gracefully decreases as K increases; the stability index observed for $K \simeq 30$ is still quite good (close



(a)



(b)

Figure 4. Self-stability and mutual stability of the clusters. (a) shows the self-stability of U- and W-representations and mutual stability between them on different K. (b) shows the mutual stability changing with T

to 0.6 Fig. 4 (a)). In the meanwhile, the error rate decreases as K increases. The number T of feature clusters has no impact on the error rate or on the cluster stability in the considered range (Fig.3 (b)).

In conclusion, the paper presents two contributions regarding constructive induction and stable clustering of heterogeneous datasets. Empirically, the approach was found to be successful to cluster the EGEE grid jobs, identifying classes which were unknown to the algorithm, and finding even finer-grained subclasses, which are under inspection by the user. The stability analysis provides guarantees about the quality of the clusters found after two independent protocols. The double clustering approach was shown effective in reducing the dimensionality of the redescription without degrading the error rate.

Our next step is to construct user and job profiles to enable some personalized support and prediction of job failures.

References

[1] J. O. Kephart, and D. M. Chess. *The vision of autonomic computing computer*. 36(1):41-50, 2003.

[2] I. Rish, M. Brodie, S. Ma, et al. Adaptive diagnosis in distributed dystems. *IEEE Transactions on Neural Networks (special issue on Adaptive Learning Systems in Communication Networks)*. 16:1088-1109, 2005.

[3] M. Sebag, N. Lucas, J. Azé. Impact studies and sensitivity analysis in medical data mining with ROC-based genetic learning. *ICDM*. 637-640, 2003.

[4] N. Slonim, and N. Tishby. Document clustering using word clusters via the information bottleneck method. *Research and Development in Information Retrieval*. 208-215, 2000.

[5] M. Meila. The uniqueness of a good optimum for K-means. *ICML*. 625-632,2006.

[6] J. B. Tenenbaum, V. D. Silva, and J. C. Langford. A global geometric framework for nonlinear dimensionality reduction. *Science*. 290:2319-2323, 2000.

[7] S. Roweis, L. Saul. Nonlinear dimensionality reduction by locally linear embedding. *Science*. 290:2323-2326, 2000.

[8] M. Sugiyama. Local fisher discriminantanalysis for supervised dimensionality reduction. *ICML*. 905-912, 2006.

[9] X. Yang, H. Fu, H. Zha, et al. Semi-supervised nonlinear dimensionality reduction. *ICML*. 1065-1072, 2006.

[10] K. Q. Weinberger, J. Blitzer, and L. K. Saul. Distance metric learning for large margin nearest neighbor classification. In *Proc. Neural Info. Processing Systems (NIPS)*. 1473-1480, 2005.

[11] C. Ding, and X. He. K-means clustering via principal component analysis. *ICML*. 225-232, 2004.

[12] M. Kearns, and M. Li. Learning in the Presence of Malicious Errors, in *SIAM J. Comput*. 22:807-837, 1993.

[13] S. Rosset. Model selection via the AUC. *ICML*. 89-96, 2004.

[14] K. Neocleous, M. Dikaiakos, P. Fragopoulou, et al. Failure management in grids: the case of the EGEE infrastructure. Technical report, Institute on System Architecture, CoreGRID - Network of Excellence. 2006.