

Modèles comportementaux de la grille : enjeux et exemples

Cecile Germain-Renaud

▶ **To cite this version:**

Cecile Germain-Renaud. Modèles comportementaux de la grille : enjeux et exemples. Premières journées scientifiques France-Grilles, Sep 2011, Lyon, France. inria-00632438

HAL Id: inria-00632438
https://inria.hal.science/inria-00632438

Submitted on 14 Oct 2011

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.

Modèles comportementaux de la grille : enjeux et exemples

Cécile Germain-Renaud, pour la collaboration Grid Observatory (1)
(1) cecile.germain@lri.fr, Laboratoire de Recherche en Informatique, U. Paris Sud, CNRS, INRIA

Overview

Grid research and engineering require behavioral models of its use and middleware. These models must be inferred from the data, through Machine Learning. Technology and usages of Grids and Clouds change rapidly, challenging analysis and evaluation. We summarize two of these challenges and present new methods for addressing them.

Des modèles comportementaux pour l'opérationnel et la recherche

Des modèles descriptifs, génératifs (descriptifs paramétriques), ou prédictifs sont nécessaires pour opérer, pour dimensionner, et pour faire évoluer, la grille. Quelques exemples sont les suivants. Pour opérer : le scheduling de glite repose sur un prédicteur. Pour dimensionner : comment adapter la fréquence d'acquisition d'information pour améliorer l'extensibilité du système d'information (BDII) [Fie] ? Pour faire évoluer : l'adoption des techniques autonomiques demande des démonstrations convaincantes [Hub]. Les modèles sont indispensables lorsque les possibilités d'expérimentations en vraie grandeur sont limitées.
Ces modèles doivent être inférés à partir des données de fonctionnement, ce qui soulève alors les défis du déluge de données. Au premier degré, il s'agit du traitement de données massives, qui demande des ontologies encore à créer. Du point de vue de l'analyse, ces données relèvent de deux défis récents en apprentissage statistique : l'absence d'interprétation de référence, les données n'étant en général pas étiquetées sémantiquement ; le caractère extrêmement bruité de ces données, découlant des limitations inhérentes à tout système d'acquisition lui-même réparti, et intégrant en général une hiérarchie de caches.

L'initiative Grid Observatory

L'initiative Grid Observatory, qui participe à EGI-Inspire, adresse à la fois l'acquisition des données comportementales de la grille et leur analyse. Nous renvoyons à [Ger] pour une présentation de l'architecture d'acquisition. Elle se situe à la fois dans un objectif de Data Curation (indexation, exhaustivité, validation), et dans le contexte Open Data (mise à disposition des utilisateurs scientifiques sans restriction).

Non-stationarité et valeurs extrêmes

Dans le domaine des modèles descriptifs ou génératifs, les grilles sont généralement décrites comme des systèmes complexes, font l'hypothèse de distributions statistiques stationnaires et exhibent de la dépendance à longue portée [Li, Lin, Min]. Mais les mécanismes concrets, sociologiques et économiques, qui régissent le comportement et l'évolution de la grille ne motivent pas naturellement l'hypothèse de stationnarité. La difficulté est alors que l'ensemble des outils statistiques, des plus classiques (calcul de moments, ajustement de modèles), à d'autres plus sophistiqués, comme le clustering, ne sont plus justifiés. Une alternative consiste à considérer une autre classe de modèles : une séquence de processus stationnaires séparés par des points de rupture. Outre une meilleure vraisemblance par rapport aux mécanismes concrets, étudier cette classe est justifié par la confusion possible entre processus à ruptures et dépendance à longue portée, bien documentée théoriquement et empiriquement [Gra, Bha, Fea].
L'évaluation se heurte à une autre difficulté : la fréquence de valeurs extrêmes. Les indicateurs élémentaires comme, pour une distribution statistique, l'erreur quadratique moyenne (RMSE, Root Mean Squared Error), qui mesure la distance entre l'estimateur et la valeur mesurée, ou pour le cas spécifique des séries chronologiques, ou comme la corrélation entre série prédite et mesurée, sont profondément perturbés. Dans le domaine du scheduling, cette difficulté a été soulignée par les travaux autour du projet BQP (Batch Queue Predictor) [Mut], à partir du besoin utilisateur : fournir un intervalle de confiance réaliste pour les prédictions.
Les études présentées dans la section suivante illustrent ces deux questions ; [Ger] et [Elt] en donnent une présentation plus complète, ainsi que l'application à la détection des phénomènes

de sites attirant des jobs pour les détruire (blackholes).

Vers des modèles comportementaux réalistes

Modèles auto-régressifs locaux

Les données analysées sont les séries chronologiques (time series) de la charge de quatre Computing Elements (CE) particulièrement chargés. Le problème de sélection de modèle consiste alors à découvrir les points de rupture, et à ajuster un modèle paramétrique dans chaque segment. Les principaux résultats de cette étude sont les suivants.

- Un modèle non-stationnaire peut être identifié à partir des données empiriques.
- La limitation de la gamme des modèles locaux à affine par morceaux (autorégressif -AR) est suffisamment puissante.
- Une stratégie de validation par bootstrap permet de suppléer à l'absence de résultats théoriques concernant la qualité de l'estimation des paramètres du modèle.

Davis a proposé dans [Dav] AutoPARM, une méthode de sélection de modèles AR par morceaux basé sur le principe de Description de Longueur Minimale (MDL) de type 2-parties : le codage du modèle, qui décrit simultanément les ruptures et la paramétrisation AR d'une part, et celui de l'écart du modèle aux données (innovation) d'autre part. La sélection de modèle devient alors un problème d'optimisation : minimiser la fonction de codage. Dans ce travail, l'optimisation a été effectuée par un algorithme génétique (de type N-islands, pour prévenir une convergence prématurée dans un paysage d'optimisation difficile). La principale limitation de la méthode est son manque d'extensibilité, la complexité pour chaque évaluation de la fonction d'objectif étant proportionnelle à la taille des données. Un travail en cours adresse cette limitation, en exploitant un algorithme évolutionnaire plus avancé, CMA-ES [Han].

La validation, en l'absence d'interprétation de référence, est évidemment cruciale. Nous avons défini la méthodologie suivante.

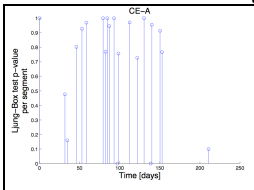


Figure 1. Independence of the residuals /Indépendence des résidus : le test

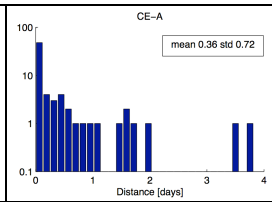


Figure 2. Distribution of the distances of the bootstrapped samples. / Distribution des distances des échantillons générés par bootstrap.

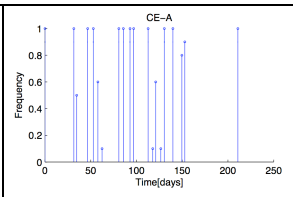


Figure 3. The clustered model: Frequency of the breakpoints. /Fréquence des ruptures dans le modèle clusterisé

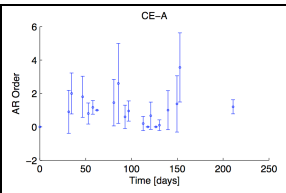


Figure 4. The clustered model: AR order ± one standard deviation. / Modèle clusterisé : ordres des modèles AR ± un écart-type

Evaluation de la qualité intrinsèque de l'ajustement. Il s'agit de vérifier que, sur chaque segment, les innovations sont effectivement un bruit blanc (indépendance des résidus, sans hypothèse de gaussianité). La figure 1 donne les p-values du test de Ljung-Box, qui autorisent à ne pas rejeter cette hypothèse, à de rares exceptions près sur des segments très courts.

Evaluation de la stabilité du modèle. Le modèle doit être robuste à de petites perturbations, qui correspondent à l'indéterminisme de la génération des données et de la mesure. Comme d'autres données (pour la même période et le même CE) ne sont évidemment pas disponibles, la technique classique du bootstrap a été employée. Des séries chronologiques suivant le modèle, et incluant une permutation aléatoire des innovations à l'intérieur de chaque segment, sont générées, puis segmentées et ajustées. On peut alors évaluer finement le degré de similarité entre les modèles résultants par la distribution des distances entre les points de ruptures voisins. Ceci fournit une meilleure évaluation qu'un critère numérique unique, comme l'approximation par le critère AIC_C de la distance de Kullback-Leibler [Hur], pour les mêmes raisons que la non-pertinence du RMSE évoquées ci-dessus. La figure 2 montre cette distribution ; elle est fortement dominée par les petites valeurs de la distance. Lorsque l'unité est la période d'échantillonnage, plus de 97% des distances sont à moins de 10 points, ce qui est en fait la résolution de la méthode, et la moyenne est proche de 1, avec un écart-type inférieur à 10

dans tous les cas. La segmentation peut donc être considéré comme stable.
Enfin, pour bâtir un modèle finalisé, les points de rupture issus du bootstrap sont groupés par clustering (clustering hiérarchique par centroïde) (figures 3). Les limites de la robustesse pour le paramètre d'ordre du modèle AR (les intervalles de confiance dans la figure 4) s'expliquent par la stagnation relative de la fonction cible de l'optimisation dans cette direction. Afin de construire un ordre plus robuste, les modèles AR doivent être réconciliés. Le bagging [Lee] donne des fondements théoriques à la réconciliation de modèles, par moyenne ou par vote. Ici, le nombre de modèles restant faible en raison du temps de calcul, le vote doit être préféré.

Evaluation des performances du scheduling gLite

Dans gLite, l'allocation des tâches aux sites est partiellement basée sur une prédiction du temps d'attente (Estimated Response Time, ERT). Le GO a crée un processus de monitoring de cette information, et un ensemble d'outils pour interpoler et sous-échantillonner les ERT et les temps d'attente réels (Actual Response Time, ART). La trace analysée correspond à l'activité du LAL d'octobre 2008 à février 2009. Nous avons examiné les quatre indicateurs suivants: RMSE; test du chi-2; corrélation de séries chronologiques ; enfin, la courbe ROC (Receiver Operating Characteristic). Ce dernier indicateur estime le rapport coût/bénéfice d'un prédicteur.

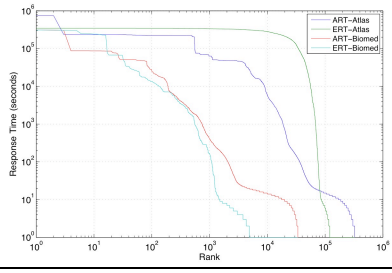


Figure 5. Distribution of response time / Distribution des temps de réponse

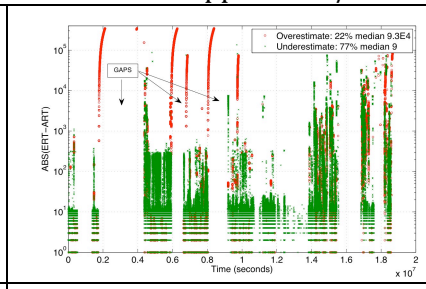


Figure 6. Absolute value of the errors /valeurs absolue des erreurs

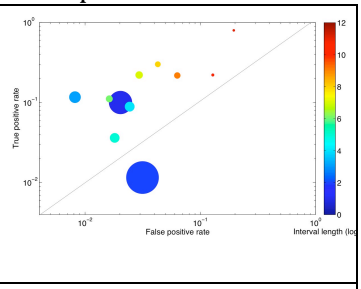


Figure 7. ROC curve - Atlas Queue /courbe ROC pour la queue Atlas

Le RMSE est au moins un ordre de grandeur supérieur à la moyenne ou la médiane de l'ART, indiquant donc une très mauvaise qualité de prédiction. L'inspection des distributions (figure 5) explique ce mauvais résultat par la domination de très grandes valeurs rares (qui ne sont pas des aberrations) de l'ERT. Comme on peut s'y attendre, le chi-2 rejette l'hypothèse de similitude: 4 degrés de liberté, valeur p = 0 pour la queue Atlas, 1 degré de liberté, valeur p = 0,2112 pour la queue Biomed. L'inspection de la dynamique du système (figure 6) montre que ces anomalies correspondent à des pics d'ERT très localisés, et résultant probablement d'un dysfonctionnement du prédicteur. La corrélation donne également de mauvais résultats, avec plusieurs dizaines d'heures de décalage.
Les indicateurs devraient intégrer à la fois la fiabilité de l'estimation (à la BQP), et la difficulté de la tâche, ce qui est essentiel vu des profils de charge très irréguliers: l'estimation exacte d'une ERT nul pour une queue vide n'est pas très instructive.
Lorsque l'objectif est de comparer des classifieurs, une stratégie alternative est consiste à mesurer la précision en fonction de la sensibilité, pour chaque classificateur considéré. Le meilleur cas est le point (0, 1), où toutes les prévisions sont exactes, le pire est (1,0), où toutes les prévisions sont erronées, la diagonale séparant la région où la prédiction est meilleure qu'aléatoire, de celle où la prédiction est pire que aléatoire.
L'idée est alors de considérer que l'ERT définit un ensemble de classifieurs associés à des intervalles disjoints: pour un intervalle de temps donné, l'ERT prédit que le temps d'attente tombera ou non dans cet intervalle. Afin de ne pas être trop dépendant d'un ensemble de données particulier, nous avons défini les limites des intervalles comme une progression géométrique, ce qui donne une croissance exponentielle de la taille de l'intervalle. La figure 7 montre la courbe ROC; la surface des points est proportionnelle à la population réelle de l'intervalle, et la barre latérale donne la taille de l'intervalle. L'ERT peut être alors être considéré comme un bon prédicteur dans la plupart des cas, puisque dans la région supérieure. Dans le cas de la queue cas Atlas, le grand intervalle mal prédit provient de l'effet de pic précité.

Perspectives

Le but de cet article était de montrer sur deux exemples que les objectifs de modélisation soulèvent des questions scientifiques difficiles, ouvrant la voie à des interactions entre l'apprentissage statistique et les utilisateurs opérationnels et scientifiques des modèles de la grille. D'autres études [Zha] traitent du diagnostic de défaillances par des techniques de clustering sous deux contraintes : streaming (les données doivent être traitées en temps réel car très volumineuses), et détection des ruptures pour des modèles non paramétriques.
L'exploitation avancée des données du Grid Observatory ouvre également d'autres perspectives de recherche, dans le domaine des ontologies : l'intégration de données hétérogènes pourrait dans une certaine mesure suppléer à l'absence d'interprétation de référence.
La constitution d'une communauté interdisciplinaire nouvelle n'est jamais simple, nous souhaitons avoir montré qu'elle peut être fructueuse. Contribuera aussi à cette démarche la publication de données de consommation énergétique, via le projet Green Computing Observatory, dont une démonstration est proposée dans cette conférence.

Remerciements

Le Grid Observatory est partiellement financé par : le programme PEPS-CNRS 2008-2010, le programme Logiciels et Systemes Complexes de la Région Ile de France opéré par la fondation DIGITEO, le programme Moyens de Recherche Mutualisés de l'Université Paris-Sud, le GIS France-Grilles, et EGI-Inspire.

Références

[Hub] M. C. Huebscher and J. A. McCann. A survey of autonomic computing: degrees, models, and applications. ACM Comput. Surv., 40:7:1–7:28, 2008.
[Fie] L. Field and R. Sakellariou. How Dynamic is the Grid? Towards a Quality Metric for Grid Information Systems. 11th IEEE/ACM Int. Conf. on Grid Computing, pp. 113-120, 2010.
[Ger] Cécile Germain-Renaud, Alain Cady, Philippe Gauron, Michel Jouvin, Charles Loomis, Janusz Martyniak, Julien Nauroy, Michèle Sebag. The Grid Observatory. 11th IEEE/ACM International Symposium on Cluster, Cloud, and Grid Computing (CCGrid 2011).
[Li] H. Li and M. Muskulus. Analysis and modeling of job arrivals in a production grid. SIGMETRICS Perform. Eval. Rev., 34(4):59–70, 2007.
[Lin] D. Lingrand, T. Glatard, and J. Montagnat. Modeling the latency on production grids with respect to the execution context. Parallel Computing, 35(2009):493–511, 2009.
[Min] T. N. Minh, L. Wolters, and D. Epema. A Realistic Integrated Model of Parallel System Workloads. 10th IEEE/ACM Int. Symp. on Cluster, Cloud and Grid Computing, pp 464–473, 2010.
[Gra] C. W. J. Granger and N. Hyung. Occasional structural breaks and long memory with an application to the S&P 500 absolute stock returns. J. Empirical Finance, 11(3):399–421, 2004.
[Bha] R. N. Bhattacharya, V. K. Gupta, and E. Waymire. The Hurst effect under trends. J. Appl. Probab., 20(3):649–662, 1983.
[Fea] P. Fearnhead. Exact Bayesian curve fitting and signal segmentation. IEEE Trans. Signal Process., 53(6):2160–2166, 2005.
[Mut] Andrew Mutz, Rich Wolski, and John Brevik. Eliciting honest value information in a batch-queue environment. 8th IEEE/ACM Int. Conf. on Grid Computing, pp 291-297. 2007
[Elt] T. Elteto, C. Germain-Renaud, P. Bondon, et M. Sebag. Discovering Piecewise Linear Models of Grid Workload. 10th IEEE/ACM Int. Symp on Cluster, Cloud and Grid Computing, pp 474-484. 2010
[Dav] R. A. Davis, T. Lee, and G. Rodriguez-Yam. Structural break estimation for nonstationary time series models. J. American Statist. Assoc., 101:229–239, 2006.
[Han] N. Hansen et al. Reducing the Time Complexity of the Derandomized Evolution Strategy with Covariance Matrix Adaptation (CMA-ES). Evolutionary Computation, 11(1) :1-18, 2003.
[Hur] C. M. Hurvich; C.L. Tsai. Regression and time series model selection in small samples. Biometrika 76(2):pp. 297-307, 1989.
[Lee] T-W Lee and Y. Yang. Bagging binary and quantile predictors for times series. J. of Econometrics, 135(1-2):465–497, 2006.
[Zha] X. Zhang, C. Furtlehner, J. Perez, C. Germain-Renaud, M. Sebag. Toward Autonomic Grids: Analyzing the Job Flow with Affinity Streaming. 15th ACM Conf. on Knowledge Discovery and Data Mining, pp. 987-996, 2009