

Approche entropique pour l'analyse de modèle de chroniques

Nabil Benayadi*, Marc Le Goc*, Philippe Bouché*.

*Laboratoire des Sciences de l'Information et des Systèmes - LSIS

UMR CNRS 6168 - Université Paul Cézanne

Avenue Escadrille Normandie Niemen 13397 Marseille Cedex 20 – France

{nabil.benayadi, marc.legoc, philippe.bouche}@lsis.org

Résumé. Cet article propose d'utiliser l'entropie informationnelle pour analyser des modèles de chroniques découverts selon une approche stochastique (Bouché et Le Goc, 2005). Il décrit une adaptation de l'algorithme TemporalID3 (Console et Picardi, 2003) permettant de découvrir des modèles de chroniques à partir d'un ensemble d'apprentissage contenant des séquences d'occurrences d'événements discrets. Ces séquences représentent des suites d'alarmes générées par un système à base de connaissance de monitoring et de diagnostic de systèmes dynamiques. On montre sur un exemple que l'approche entropique complète l'approche stochastique en identifiant les classes d'événements qui contribuent le plus significativement à la prédiction d'une occurrence d'une classe particulière.

1 Introduction

La découverte de connaissances temporelles est un enjeu majeur pour le diagnostic de systèmes dynamiques (Das et al., 1998), (Dousson et Vu Duong, 1999), (Keogh et Smyth, 1997), (Agrawal et al., 1995), (Faloutsos et al., 1994). Récemment, Bouché P. et Le Goc M. (2005) ont proposés une approche stochastique pour découvrir des modèles de chroniques à partir d'une séquence d'événements discrets. Nos travaux visent à compléter cette approche pour identifier les classes d'événements contribuant le plus significativement à la prédiction de l'occurrence d'une classe particulière.

Les arbres de décisions (Breiman, 1984), (Murthy, 1998), sont largement utilisés pour classer des séquences de données (Kadous, 1999), (Geurts, 2001), (Drucker et Hubner, 2002), (Rodriguez et Alonso, 2004). Récemment, l'algorithme ID3 (Quinlan, 1986) a été adapté pour construire des arbres temporels de décision (Console et al., 2003) à partir d'un ensemble de situations. Cette adaptation montre que l'entropie informationnelle permet d'identifier les variables contribuant le plus significativement à une prise de décision.

Nous proposons donc d'utiliser un critère entropique pour analyser des modèles de chroniques. Après un bref rappel sur les arbres temporels de décision, cet article présente une adaptation de l'algorithme proposée par Console pour la déduction de modèles de chroniques à partir d'un ensemble de séquences d'occurrences d'événements discrets et montre sur un exemple comment l'approche entropique peut être utilisée pour compléter l'approche stochastique.