



HAL
open science

Évaluation des algorithmes d'apprentissage de structure pour les réseaux bayésiens dynamiques.

Ghada Trabelsi, Philippe Leray, Mounir Ben Ayed, Adel Alimi

► **To cite this version:**

Ghada Trabelsi, Philippe Leray, Mounir Ben Ayed, Adel Alimi. Évaluation des algorithmes d'apprentissage de structure pour les réseaux bayésiens dynamiques.. 7èmes journées francophones sur les réseaux bayésiens (JFRB 2014), 2014, Paris, France. hal-01005775

HAL Id: hal-01005775

<https://hal.science/hal-01005775v1>

Submitted on 17 Apr 2020

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.

Evaluation des algorithmes d'apprentissage de structure pour les réseaux Bayésiens dynamiques

Ghada Trabelsi

National School of Engineers of Sfax,
(REGIM), University of Sfax, Tunisia
Laboratoire d'Informatique de Nantes
Atlantique (LINA), Nantes, France
Email: ghada.trabelsi@gmail.com

Philippe Leray

Laboratoire d'Informatique de Nantes
Atlantique (LINA), Nantes, France
Email: philippe.leray@univ-nantes.fr

Mounir Ben Ayed

Adel M. Alimi
National School of Engineers of Sfax,
(REGIM), University of Sfax, Tunisia
Email:mounir.benayed@ieeeg.org
adel.alimi@ieeeg.org

Les réseaux bayésiens (RB) sont connus comme l'un des formalismes les plus complets et cohérents pour l'acquisition et la représentation des connaissances et du raisonnement à partir de données incomplètes et/ou incertaines. L'apprentissage de la partie graphique de ces modèles à partir de données est un problème NP-difficile [?]. De nombreuses méthodes ont été proposées sur ce sujet [?]. L'évaluation de ces algorithmes passe classiquement par l'utilisation de métriques connues comparant un modèle appris et un modèle de référence ayant servi à générer les données d'apprentissage.

Les réseaux bayésiens dynamiques (RBD) sont une classe de modèle générale et flexible permettant de représenter des processus temporels stochastiques complexes [?]. Certains algorithmes d'apprentissage de structure ont été proposés pour ces modèles, reprenant les principes des méthodes proposées dans le cas statique. Assez souvent, ces méthodes se concentrent sur un cas particulier de RBD, les 2-TBN, réseaux à deux tranches de temps. Par contre, la comparaison de ces algorithmes est une tâche difficile car les métriques et/ou les réseaux de référence utilisés changent d'un article à un autre. De plus, l'évaluation de ces algorithmes est souvent limitée à des réseaux ayant un nombre limité de variables.

Nous proposons dans ce travail deux contributions ayant pour objectif l'analyse comparative de n'importe quel algorithme d'apprentissage de structure pour les 2-TBN : (1) un algorithme de génération de grands 2-TBN s'appuyant sur une approche de *tiling* proposée dans le cas statique, et (2) une métrique pour évaluer la qualité de l'apprentissage adaptant la distance structurelle de Hamming (SHD) au cas dynamique, en prenant bien en compte toutes les spécificités liées à ces modèles.

Un réseau bayésien dynamique (RBD) définit la distribution de probabilité d'une collection de variables aléatoires $\mathbf{X}[t]$ où $\mathbf{X} = \{X_1 \dots X_n\}$ est l'ensemble des variables observées au cours du temps discret t . Dans ce travail, nous considérons une classe spéciale de RBD, appelé réseaux bayésiens à deux tranches de temps (2-TBN). Un 2-TBN est un RBD qui satisfait la propriété de Markov d'ordre 1 $\mathbf{X}[t-1] \perp \mathbf{X}[t+1] \mid \mathbf{X}[t]$. En conséquence, un 2-TBN est décrit par la paire (M_0, M_{\rightarrow}) .

Les modèles de référence souvent utilisés dans l'évaluation

des algorithmes d'apprentissage ont souvent un nombre limité de variables. Deux solutions sont possibles pour contrôler cette dimension. Il est tout d'abord possible de générer aléatoirement des modèles graphiques probabilistes de n'importe quelle taille [?]. Tsamardinos et al. [?] proposent de leur côté un algorithme pour la génération de grands RB par empilement (*tiling*) d'un premier modèle de référence.

Une fois l'apprentissage réalisé, il faut ensuite appliquer une métrique comparant le modèle de référence et le modèle appris. Plusieurs mesures ont ainsi proposées dans la littérature. Tsamardinos et al. [?] ont proposé une adaptation de la distance de Hamming entre les graphes tenant compte d'un problème d'identifiabilité statistique des modèles de dépendances. En effet, certains arcs d'un réseau bayésien peuvent être inversés sans changer les propriétés de dépendances/indépendances décrites par le modèle. Les réseaux bayésiens sont donc identifiables à une classe d'équivalence près, représentée par un graphe partiellement dirigé (PDAG) aussi appelé graphe essentiel. La métrique proposée par Tsamardinos et al. compare alors les PDAG du modèle de référence et du modèle appris afin de ne comparer que les orientations d'arcs qui sont réellement statistiquement distinguables.

L'apprentissage de structure étant une tâche difficile, certains travaux proposent de réduire l'espace des solutions à l'aide de connaissances a priori, par exemple en déclarant le fait que certains arcs sont obligatoirement présents (ou absents) dans le graphe final. Les classes d'équivalence décrites précédemment se doivent alors de prendre aussi cette information en compte. Meek [?] a ainsi proposé un algorithme permettant d'identifier un graphe essentiel compatible avec ce type de connaissances a priori.

Nous décrivons ici des nouveaux outils pour l'évaluation des algorithmes d'apprentissage de structure pour les RBD. La première contribution décrit un algorithme de génération de 2-TBN. La seconde contribution présente une nouvelle métrique pour comparer deux modèles de type 2-TBN.

A. Génération de 2-TBN

Dans le même état d'esprit que l'approche de *tiling* de Tsamardinos et al. nous proposons de générer les deux com-

posantes d'un 2-TBN (M_0 et M_{\rightarrow}) à partir d'un réseau bayésien statique de référence M . Ainsi M_0 peut être obtenu par *tiling* de n copies de M , nous permettant de contrôler le nombre de variables du 2-TBN. M_{\rightarrow} est ensuite obtenu par *tiling* de 2 copies de M_0 . Les paramètres de l'algorithme de *tiling* permettent aussi de contrôler le degré de connectivité entre les deux tranches de temps (arcs "temporels").

L'exécution de cette procédure sur plusieurs RB de référence usuels nous a permis de générer plusieurs 2-TBN de référence que nous mettons à disposition en ligne ¹ pour tous les chercheurs intéressés par l'apprentissage de structure de 2-TBN.

B. Métrique de comparaison de 2-TBN

La composante graphique d'un 2-TBN est définie par deux graphes G_0 (graphe initial) et G_{\rightarrow} (graphe de transition). Nous proposons d'évaluer la différence entre deux 2-TBN par une paire de distance de Hamming, celle entre les deux graphes initiaux, et celle entre les deux graphes de transition. Comme nous l'avons vu auparavant, tenir en compte des classes d'équivalence en comparant les PDAG, mais ce n'est pas suffisant pour la comparaison des graphes de transition. En effet, certaines directions d'arcs sont maintenant imposées de part la nature temporelle du modèle, même ces relations ne sont pas identifiables statistiquement. Les algorithmes d'apprentissage des 2-TBN tiennent compte de ces informations temporelles avec l'orientation automatique des arcs de l'instant t vers l'instant $t + 1$ et pas le contraire. Il est alors important de ne pas perdre cette information lors de la transformation du graphe en son PDAG.

Nous avons donc proposé une nouvelle adaptation de la distance de Hamming prenant à la fois en compte les problèmes d'identifiabilité soulevés par Tsamardinos et al. [?] et des contraintes "temporelles", en adaptant l'approche de Meek [?]. La figure 1 montre un exemple d'utilisation de cette distance structurelle. La distance de Hamming entre PDAG compte 4 erreurs entourées dans la deuxième ligne. L'application de notre correction temporelle (sur la dernière ligne) mène à une métrique plus précise, ne détectant qu'une seule erreur.

Nous nous sommes concentré dans ce travail sur la proposition d'outils pour l'évaluation des algorithmes d'apprentissage de la structure de 2-TBN. Ces outils ont été utilisés pour l'évaluation d'un algorithme d'apprentissage hybride capable de travailler avec des grands 2-TBN [?]. Des algorithmes d'apprentissage récent traitent le problème du passage à l'échelle en simplifiant l'espace de recherche par l'interdiction de dépendances intra-temporelles [?], [?]. Ainsi, ils sont capables de traiter de plus grands modèles et/ou de travailler avec des modèles à k tranches de temps avec $k > 2$. Nous travaillons actuellement à l'extension de notre procédure de génération de modèles dans ce même contexte.

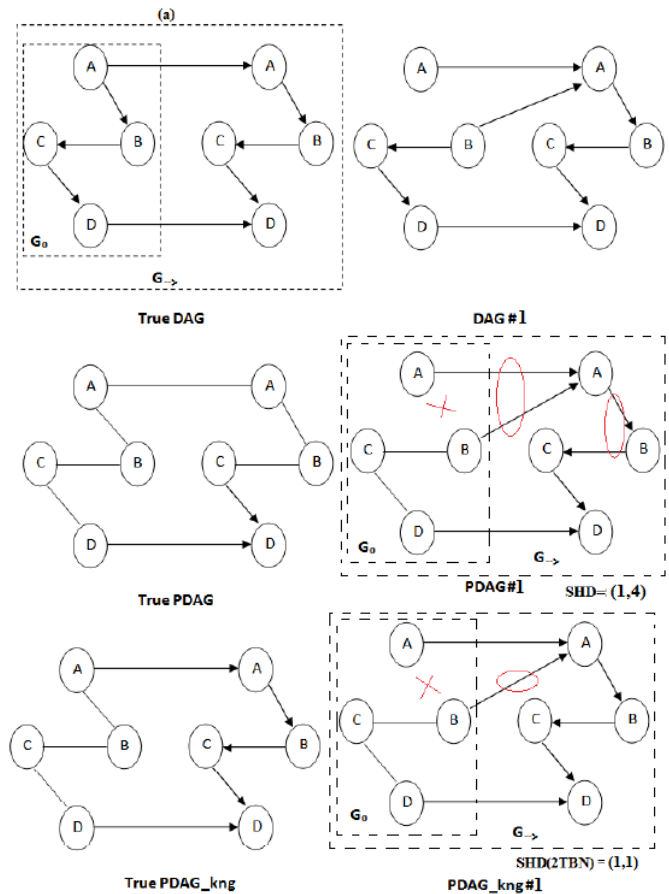


Figure 1. Comparaison de deux 2-TBN (TrueDAG et DAG#1). La première ligne donne les deux modèles graphiques. La seconde ligne donne les graphes essentiels associés à ces deux graphes. La dernière ligne donne les graphes essentiels après correction temporelle.

¹<https://sites.google.com/site/dynamicbenchmarking/>