УДК 62-50

АЛГОРИТМ ВЕРОЯТНОСТНОГО ВЫВОДА В БАЙЕСОВСКИХ СЕТЯХ

А.Н. ТЕРЕНТЬЕВ, П.И. БИДЮК, Л.А. КОРШЕВНЮК

Предлагается новый, более простой и точный алгоритм вероятностного вывода в байесовских сетях на основе обучающих данных.

ВСТУПЛЕНИЕ

Применение технологий интеллектуального анализа данных становится все более актуальным благодаря совершенствованию методов анализа и вычислительных процедур, которые их реализуют. Это связано также с тем, что в мире постепенно накапливаются большие объемы информации, требующие надлежащей обработки и принятия решений. Успешное развитие любого предприятия напрямую зависит от его способности адекватно и оперативно реагировать на изменение внешней среды, а также умение прогнозировать результаты тех или иных воздействий. Так, в отчете Ассоциации американских банкиров отмечается, что 45 из 100 крупнейших банков США уже внедрили у себя системы интеллектуального анализа данных, и еще около 50 банков начали реализацию подобных проектов или планируют это сделать в ближайшее время.

Среди разнообразных методов интеллектуального анализа байесовские методы занимают одно из ведущих направлений. Благодаря универсальности по отношению к используемым типам данных и решаемых практических задач их можно использовать в различных областях человеческой жизнедеятельности — технике, экономике, финансах, информационных технологиях, медицине, биологии и других направлениях [1,2]. Классы решаемых задач также самые разнообразные — классификация, прогнозирование, диагностика, оценивание, управление и т.п. Одним из самых перспективных современных байесовских методов является байесовская сеть (БС).

При построении БС необходимо решать задачи обучения, оптимизации структуры сети и формирования вероятностного вывода. Задача вероятностного вывода в БС является важной и сложной задачей, относящейся к классу задач принятия решений. Один из методов формирования вероятностного вывода в БС предложен в работе [3]. Однако для реализации метода необходимо привести структуру БС к виду объединенного дерева (junction tree). Для этого нужно выполнить следующее.

1. Морализировать граф (структуру БС) — добавить дуги между несвязанными между собой предками каждой вершины сети.
2. Дезориентировать граф — направленные дуги заменить неориентированными ребрами.
3. Триангулировать граф — привести структуру графа к такому виду, чтобы не было циклов длиной больше трех.
4. Морализированный триангулированный граф привести к виду графа клик, т.е. к графу, состоящему из подграфов.
5. Граф клик привести к виду объединенного дерева.

И только после этого можно использовать алгоритм вероятностного вывода в объединенном дереве, который основывается на прохождении λ и π сообщений по дереву и последовательном пересчете таблиц условных вероятностей.

В работе [4] предложен метод поглощающего исключения (bucket elimination). Проблемой, возникающей при использовании этого метода, является обязательное наличие упорядоченного множества вершин, получение которого является сложной вычислительной задачей. Более подробно анализ проблемы и обзор методов формирования вероятностного вывода в БС рассмотрен в работах [5–7].

Очевидно, что существующие методы формирования вывода, предложенные в работах [3, 4], требуют серьезного преобразования структуры БС и трудоемких подготовительных вычислений. Поэтому разработка методов, позволяющих уменьшить вычислительную сложность, является актуальной и востребованной при моделировании процессов различной природы сетями Байеса.

ПОСТАНОВКА ЗАДАЧИ

Разработка метода вероятностного вывода в БС, состоящего из двух шагов. На первом шаге выполняется вычисление матрицы эмпирических значений совместного распределения вероятностей всей сети, на втором — вероятностей всех возможных состояний не инстанциированных вершин.

АЛГОРИТМ ВЕРОЯТНОСТНОГО ВЫВОДА В БС НА ОСНОВЕ ОБУЧАЮЩИХ ДАННЫХ

Входные данные, необходимые для алгоритма формирования вывода

1. **Множество обучающих данных** $D=\{d_1,...,d_n\}$, $d_i=\{x_i^{(1)}x_i^{(2)},\ldots$ $\ldots,x_i^{(N)}\}$ (нижний индекс — номер наблюдения, верхний — переменной), n — количество наблюдений, состоящих из N ($N \ge 2$) переменных $X^{(1)},X^{(2)},...,X^{(N)}$. Каждая j-я переменная ($j=1,...,N$) имеет $A^{(j)}=$ $=\{0,1,...,\alpha^{(j)}-1\}$ ($\alpha^{(j)} \ge 2$) состояний.

2. **Структура байесовской сети** g, представленная множеством из N предков ($\Pi^{(1)},...,\Pi^{(N)}$). Для каждой вершины $j = 1,...,N$, $\Pi^{(j)}$ — это множество родительских вершин такое, что $\Pi^{(j)} \subseteq \{X^{(1)},...,X^{(N)}\} \setminus \{X^{(j)}\}$ (вершина не может быть предком для самой себя — петли в графе должны отсутствовать).

3. **Множество инстанциированных вершин** $\{X^{(P_1)} = x^{(P_1)},\dots,X^{(P_v)} = = x^{(P_v)}\}$, т.е. вершин, находящихся в некотором определенном состоянии с единичной вероятностью. Если множество инстанциированных вершин пустое, то нужно использовать классический вероятностный вывод.

Алгоритм формирования вывода

Шаг 1.

По множеству обучающих данных вычисляется матрица эмпирических значений совместного распределения вероятностей всей сети $P(X^{(1)},...,X^{(N)})$. По формуле

$$P_{\text{matrix}}(X^{(1)} = x^{(1)},...,X^{(N)} = x^{(N)}) = \frac{n[X^{(1)} = x^{(1)},...,X^{(N)} = x^{(N)}]}{n},$$

где n — количество обучающих наблюдений; $x^{(j)} \in A^{(j)}$, а числитель вычисляется так:

$$n[X^{(1)} = x^{(1)},...,X^{(N)} = x^{(N)}] = \sum_{j=1}^{n} I(X^{(1)} = x_j^{(1)},...,X^{(N)} = x_j^{(N)}),$$

где функция $I(E) = 1$, когда предикат $E = \text{true}$, в противном случае $I(E) = 0$.

Шаг 2.

Перебираем последовательно все вершины БС. Если вершина не является инстанциированной, то нужно вычислить значения вероятностей возможных состояний этой вершины. Для этого делается последовательный перебор всех строк матрицы эмпирических значений совместного распределения вероятностей сети. Если значения вершин строки совпадают со значениями инстанциированных вершин и состоянием анализируемой вершины, то $P_{\text{matrix}}(X^{(1)},...,X^{(N)})$ прибавляется к значению вероятности соответствующего состояния вершины. После этого выполняется нормирование значений вероятностей ее состояний.

Алгоритм вычисления значений вероятностей всех возможных состояний неинстанциированных вершин

for $j = 1$ to N if $X^j \notin \{X^{(P_1)},...,X^{(P_v)}\}$ then
  begin
    $\text{sum} = 0$;

$\forall x^{(j)} \in A^{(j)}$ do

begin

for $k=1$ to $last\_string\_matrix$ do

begin

if $(X_{\text{matrix}}^{(P_1)} = x^{(P_1)})$ and…and $(X_{\text{matrix}}^{(P_v)} = x^{(P_v)})$ and $(X_{\text{matrix}}^{(j)} = x^{(j)})$ then

begin

$P(X^{(j)} = x^{(j)}) = P(X^{(j)} = x^{(j)}) + P_{\text{matrix}}(X_{\text{matrix}}^{(1)},\ldots,X_{\text{matrix}}^{(N)})$;

end;

end;

$\text{sum} = \text{sum} + P(X^{(j)} = x^{(j)})$;

end;

$\forall x^{(j)} \in A^{(j)}$ do

begin

$$P(X^{(j)} = x^{(j)}) = \frac{P(X^{(j)} = x^{(j)})}{\text{sum}}$$

end;

end;

Выходные данные

Выходными данными являются значения вероятностей всех возможных состояний всех неинстанциированных вершин.

ВЫВОДЫ

Моделирование процессов различной природы и сложности при помощи БС является одним из перспективных современных направлений в области интеллектуального анализа данных. В работе предложен метод формирования вероятностного вывода на основе БС с использованием обучающих данных. Этот метод более простой с вычислительной точки зрения по сравнению с такими методами, как вероятностный вывод в объединенном дереве [3] и поглощающим исключением [4]. Благодаря использованию информации об инстанциированных узлах достигается более высокая точность результата по сравнению с классическим вероятностным выводом, который основывается на таблицах условных вероятностей. В дальнейшем необходима разработка более совершенных методов вероятностного вывода для БС с непрерывными вершинами, т.е. переменными, подчиняющимися стандартным законам распределения, а также более интересного случая БС с неполными обучающими данными. Описанная научно-исследовательская работа выполнена в рамках гранта НТУУ «КПИ» 3/5-ГР.

ЛИТЕРАТУРА

1. *Tuzel O., Porikli F., Meer P.* A Bayesian approach to background modeling // TR2005-033. — Mitsubishi electric research laboratories, December 2005. — 8 p.
2. *Yedidia J., Freeman W. and Weiss Y.* Understanding belief propagation and its generalizations // TR-2001-22. — Mitsubishi electric research laboratories, January 2002. — 35 p.
3. *Lokeswarappa K.G*. Junction trees: motivation // Seminar CSE 714 on advanced topics in machine learning, March 2005. — 57 p.
4. *Dechter R.* Bucket elimination: a unifying framework for reasoning // ACM Press. — 1996. — **28**, № 61. — P. 1–51.
5. *Huang C., Darwiche A.* Inference in belief networks: g procedural guide // International journal of approximate reasoning. — 1994. — **11. —** P. 1–45.
6. *Lepar V., Shenoy P.* A Comparison of Lauritzen-Spiegelhalter, Hugin, and Shenoy-Shafer Architectures for computing marginals of Probability Distributions // Uncertainty in artificial intelligence, San Francisco, CA. — **14**. — 1999. — P. 328–337.
7. *Murphy K.P.* Dynamic Bayesian networks: representation, inference and learning: A PhD dissertation. — University of California, Berkeley, 2002. — 225 p.

Поступила 30.05.2007