

РАЗРАБОТКА ГИБРИДНОГО АЛГОРИТМА ОБУЧЕНИЯ СТРУКТУРЫ ДИНАМИЧЕСКОЙ БАЙЕСОВСКОЙ СЕТИ НА ОСНОВЕ МЕТОДА ЛЕВЕНБЕРГА – МАРКВАРДТА

Т.В. Азарнова¹, С.А. Баркалов², П.В. Полухин¹

¹ Воронежский государственный университет, г. Воронеж, Россия,

² Воронежский государственный технический университет, г. Воронеж, Россия

Для моделирования сложных стохастических процессов функционирования современных многопользовательских информационно-коммуникационных систем достаточно эффективно применяются динамические байесовские сети. Динамические байесовские сети представляют собой графические вероятностные модели, отражающие топологию и стохастические причинно-следственные связи между элементами моделируемых процессов обработки информации. Построение топологии динамических байесовских сетей, адекватно отражающей вероятностные и функциональные связи между элементами процессов, является ключевым фактором успеха при моделировании с помощью данного инструментального средства. Топология сетей, как правило, строится или экспертным путем, или на основании обучения. Механизмы обучения позволяют получить основную структуру сети, а также определить условные связи и их направленность между отдельными вершинами сети. В статье рассмотрены вопросы применения математического аппарата проверки статистических гипотез о наличии зависимости между случайными величинами, опирающегося на статистические критерии Пирсона, Шварца, Акаике и Байеса – Дирихле. В отличие от статических байесовских сетей при определении структуры динамических байесовских сетей необходимо определять переменные и связи между ними не только внутри одного среза, но и между переменными различных срезов, которые реализуют транзитивные связи между временными этапами функционирования некоторого процесса или объекта. Построение структуры транзитивных связей между срезами является достаточно сложным и проблемным этапом почти всех существующих алгоритмов. В данной статье в рамках оптимизации алгоритмов построения динамических байесовских сетей с транзитивными связями между срезами представлен алгоритм обучения структуры динамической байесовской сети, базирующийся на методе Левенберга – Марквардта.

Ключевые слова: динамические байесовские сети, обучение структуры, статистические критерии проверки независимости между случайными величинами, метод Левенберга – Марквардта.

Введение

Процедура обучения динамических байесовских сетей представляет собой сложный и многоаспектный процесс, требующий детального понимания вероятностных и временных связей между узлами сетей. Одним из ключевых факторов обучения является наличие априорной информации, полученной с помощью различных эвристических средств в процессе ретроспективного или текущего анализа прикладной области. Построение сети, как правило, сводится к построению некоторой базовой топологии, и настройки вероятностных связей между вершинами сети с помощью проверки статистических гипотез. Применение различных статистических критериев позволяет определить, какая из вершин является родительской или дочерней по отношению к другой вершине. Анализ существующих алгоритмов обучения структуры байесовских сетей показывает, что нет единого универсального алгоритма, который при моделировании процессов из различных предметных областей, имеющих различную специфику вероятностного распределения взаимосвязанных элементов процессов, обладал бы более высокой эффективностью по сравнению с другими алгоритмами. Все основные алгоритмы имеют свои достоинства и недостатки в решении задач определения топологических связей между узлами сети и направленности данных связей. Апробация различных алгоритмов показывает, что более качественное описание сложных процессов удается получить с помощью гибридных алгоритмов обучения, позволяющих комбинировать несколько различных математических подходов в единый алгоритм. Проблемными

зонами большинства алгоритмов обучения структуры динамических байесовских сетей является определение вершин, которые реализуют транзитивные связи между временными срезами и определение направленности данных связей. Наличие временных связей между вершинами соседних временных срезов свойственно не только вершинам, занимающим одинаковые позиции в структуре среза, но и вершинам с различной функциональностью в структуре среза. В статье рассмотрены вопросы, связанные с повышением эффективности решения описанных проблемных точек алгоритмов, и предложен оригинальный гибридный алгоритм обучения структуры динамической байесовской сети.

Критерий условной независимости при определении семантики динамических байесовских сетей

Байесовские сети представляют собой разновидность графических вероятностных моделей. Статическая байесовская сеть строится как ациклический ориентированный граф, состоящий из множества вершин $X = \{X_1, X_2, \dots, X_n\}$, соединенных дугами. Направление дуги от вершины X_1 к вершине X_2 говорит о том, что переменная X_1 является родительской по отношению к переменной X_2 . Вершины байесовской сети рассматриваются как случайные величины, в соответствие каждой из них ставится некоторое распределение вероятностей $P(X)$, при этом если вершина имеет одну и более родительских вершин, то такое распределение будет условным $P(X|parents(X))$ [1]. Понятие динамической байесовской сети является некоторым расширением понятия статической байесовской сети, динамическая байесовская сеть может быть представлена в виде нескольких байесовских сетей, взятых в определенной хронологической последовательности. Время может рассматриваться как дискретная или непрерывная характеристика анализируемых процессов; в рамках данного исследования предполагается, что время носит дискретный характер, и для системы связей между временными срезами выполняется свойство Марковских цепей.

Обучение динамических байесовских сетей функционально разделяется на обучение структуры графа и параметров вершин. Обучение параметров направлено на получение начального распределения вероятностей $P(X_0)$ для всех вершин первого временного среза сети. В данной статье основное внимание уделяется процедуре обучения структуры сети. Процедура обучения структуры производится на основе статистического анализа обучающей выборки, содержащей множества значений, принимаемых вершинами (узлами) сети [2]. Хорошие результаты обучения в процессе апробации показывают гибридные алгоритмы на основе ограничений условной независимости и оценки качества. Остановимся на математическом аппарате алгоритмов на основе выполнения гипотезы условной независимости. Сущность понятия условной независимости в терминах байесовских сетей тесно связана с понятием Марковского покрытия [3]. Для проверки гипотез об условной независимости вершин байесовской сети используется критерий χ^2 Пирсона. Основная сущность данного критерия заключается в оценке принадлежности выборки X вероятностному распределению $P(X)$. Если выборка $X = (X_1, X_2, \dots, X_n)$ принадлежит распределению $F(x)$, то каждый интервал значений элементов выборки X_i может быть разбит на m интервалов $I = (I_1, I_2, \dots, I_m)$ и вероятность попадания в данные интервалы будет определяться на основе следующего выражения [4]

$$P(X_i) = \int_I F(X) dX = \sum_I F(X). \quad (1)$$

В рассмотрении вводится следующая статистика критерия:

$$\chi^2(X, Y|Z) = \sum_{a,b,c} \frac{(N_{a,b,c} - E_{a,b,c})^2}{E_{a,b,c}} = n_{ac} \frac{(P(b|c) - P(b|a,c))^2}{P(b|c)}, \quad (2)$$

где $N_{a,b,c}$ – частота появления значений $X = a, Y = b, Z = c$, $E_{a,b,c}$ – ожидаемое число вхождений значений $X = a, Y = b, Z = c$, которое определяется следующим образом:

$$E_{a,b,c} = \frac{N_{ac} N_{bc}}{N_c}. \quad (3)$$

Гипотеза об условной независимости отвергается, если величина χ^2 принимает неправдоподобно большие значения.

В рамках разработанного в данном исследовании гибридного алгоритма критерий Пирсона используется как мера оценки устойчивости связи $A(X, Y|Z)$ между переменными X, Y . В резуль-

тате проведения тестов на условную независимость формируются множества родительских и дочерних вершин, связанных с каждой вершиной байесовской сети. На основании этих множеств формируется ненаправленная структура байесовской сети. Для определения направленности связей между узлами байесовской сети необходимо произвести поиск направленного графа с максимальным значением выбранной индикативной оценки. Изменение оценки напрямую связано с операциями добавления, удаления и изменения направленности связей между узлами байесовской сети. В качестве подобных оценок используются различные критерии: логарифм правдоподобия, критерий Шварца, Акаике и Байеса-Дирихле [5].

Значение логарифма правдоподобия для оценки направленности байесовской сети определяется выражением

$$L(G, \theta^G, D) = \sum_{i=1}^n \sum_{j=1}^{q_i} \sum_{k=1}^{r_i} N_{i,j,k} \ln \frac{N_{i,j,k}}{N_{i,j}}. \quad (4)$$

Критерий Шварца и Акаике строятся на основе вычисления логарифма правдоподобия (4). Обобщенное представление данных критериев имеет следующий вид:

$$Q(M) = L(G, \theta^G, D) - MF(N), \quad (5)$$

где N – размер обучающей выборки, M – число параметров байесовской сети Q^G :

$$M = \sum_{i=1}^{r_i} (r_i - 1) q_i. \quad (6)$$

Используя выражения (4) и (6), обобщенное представление критериев Шварца и Акаике (5) можно записать в следующем виде:

$$Q(M) = \sum_{i=1}^n \sum_{j=1}^{q_i} \sum_{k=1}^{r_i} N_{i,j,k} \ln \frac{N_{i,j,k}}{N_{i,j}} - \sum_{i=1}^{r_i} (r_i - 1) q_i F(N). \quad (7)$$

Множитель $F(N)$ является вариативным и может принимать значения $F(N) = 1$, если рассматривается критерий Акаике и $F(N) = \frac{\ln N}{2}$ для критерия Шварца. Критерий Байеса – Дирихле определяется через распределение Дирихле, а значения условных вероятностей вычисляются на основе локальной и глобальной независимости вершин байесовской сети. Выражения для критерия Байеса – Дирихле имеет следующий вид [6]:

$$P(D|G) = \prod_{i=1}^n \prod_{j=1}^{q_i} \frac{\Gamma(\sum_{k=1}^{r_i} \alpha_{i,j,s})}{\Gamma(\sum_{s=1}^{r_i} N_{i,j,s} + \alpha_{i,j,s})} \prod_{s=1}^{r_i} \frac{\Gamma(N_{i,j,s} + \alpha_{i,j,s})}{\Gamma(\alpha_{i,j,s})}, \quad (8)$$

где $D = \{D_1 = \{X_1^G = 1, X_2^G = 2, \dots, X_3^G = n\}, D_2 = \{X_1^G = 2, X_2^G = 3, \dots, X_3^G = n - 1\}, \dots, D_m = \{X_m^G = 3, X_m^G = 1, \dots, X_m^G = 2\}\}$ – совокупность обучающих выборок для θ^G , входящих в состав байесовской сети.

Разработка гибридного алгоритма обучения на основе метода Левенберга – Марквардта

Применение классических гибридных алгоритмов для обучения структуры динамической байесовской сети становится достаточно сложным, из-за присутствия связей между временными срезами сети. Для обхода данных ограничений в статье разработан и апробирован гибридный алгоритм обучения динамической байесовской сети на основе применения инструментов цепей Маркова и метода Левенберга – Марквардта. Алгоритм структурно разделяется на два основных шага решения задач обучения сети.

На первом шаге происходит определение базовой структуры динамической байесовской сети за счет формирования Марковского покрытия для каждого из узлов. Основой для построения Марковского покрытия является обработка тестов на условную независимость в соответствии с критерием Пирсона (2). Для работы с узлами сети, имеющими временные связи, используется математический аппарат цепей Маркова. При этом предполагается, что Марковская цепь является однородной. В процессе выполнения алгоритма происходит определение узлов-кандидатов, которые могут быть добавлены в Марковское покрытие M' для переменной Z . Для удаления ошибочно добавленных вершин предполагается повторное выполнения статистических тестов для каждого подмножества $M \subseteq M'$. Если переменная Z является временным узлом сети, то Z_t и Z_{t+1} описывают состояния данной переменной в моменты времени t и $t + 1$. Марковское покрытие для динамической байесовской сети до момента $t + 1$ включительно можно представить в виде следующего выражения

$$M_{t:t+1} = M_t \cup M_{t+1}. \quad (9)$$

С учетом предположения о Марковости рассматриваемого перехода, Марковское покрытие M_{t+1} будет соответствовать множеству дочерних вершин C_{t+1} , имеющих непосредственную связь с переменными из состояния t [7]. Процедура поиска ненаправленной структуры динамической байесовской сети состоит из следующих этапов.

На начальном этапе происходит определение входных параметров алгоритма: обучающая выборка D , текущая переменная Z и Марковское покрытие $M_{t:t+1} = \emptyset$.

На следующем этапе в цикле производится анализ устойчивости связи между текущей переменной цикла X_i и переменной Z при наличии всех подмножеств $M^* \subset M_{t:t+1}$ на основе вычисления критерия Пирсона. Если значения критерия превышает некоторое критическое значение с уровнем значимости $1 - \alpha$, то гипотеза об условной независимости отвергается, а переменная добавляется в состав Марковского покрытия.

В результате выполнения первого этапа алгоритма формируется результирующее Марковское покрытие, соответствующее искомой ненаправленной структуре динамической байесовской сети.

Для определения направленности связей между узлами динамической байесовской сети используется алгоритм локального поиска на основе метода Левенберга – Марквардта. Метод обладает достаточной высокой эффективностью и сводит к минимуму вероятность попадания оценочной функции в локальный оптимум. В общем случае метода Левенберга – Марквардта представляет комбинацию градиентного метода и метода Гаусса – Ньютона [8]. В основе метода Гаусса – Ньютона лежит метод наименьших квадратов, позволяющий произвести минимизацию искомой оценочной функции $\varphi(w)$ (критерии (4), (7) и (8)):

$$\varphi(w) = \min \|\Phi(w)\|^2 = \sum_{k=1}^n (g_k(w) - a_k)^2, \quad (10)$$

где $[g_k(w) - a_k]_{k=1}^n$ – разностное выражение для градиента $g_k(w)$.

Для работы с функцией (10) вводится матрица Якоби:

$$J(w) = \begin{bmatrix} \frac{dg_1(w)}{dw_1} & \dots & \frac{dg_1(w)}{dw_n} \\ \frac{dg_2(w)}{dw_1} & \dots & \frac{dg_2(w)}{dw_n} \\ \dots & \dots & \dots \\ \frac{dg_m(w)}{dw_1} & \dots & \frac{dg_m(w)}{dw_n} \end{bmatrix}, w \in R^m. \quad (11)$$

Транспонированная матрица по отношению к Якобиану (11) является градиентом $F'(w)$:

$$F'(w) \in R^{m \times n}, F'(w)_{ij} = \frac{dF_i}{dw_j}(w), F'(w) = J(w) = \nabla F(w)^T. \quad (12)$$

Используя Якобиан (11), можно определить значение w_{j+1} , если дано начальное значение w_j :

$$w_{j+1} = w_j - \frac{J^T(w_j)F(w_j)}{J^T(w_j)J(w_j)}. \quad (13)$$

Для формулировки метода Гаусса – Ньютона необходимо определить следующее неравенство

$$[J^T(w)J(w)] > Q(w). \quad (14)$$

С учетом того, что неравенство (14) допустимо только в тех случаях, когда $Q(w) \rightarrow 0$, выражение для метода Гаусса – Ньютона в точке w_{j+1} имеет следующий вид

$$w_{j+1} = w_j - [J^T(w)J(w)]^{-1}J^T(w)F(w). \quad (15)$$

Основным недостатком метода Гаусса – Ньютона является плохая сходимость в тех случаях, когда значения $Q(w) \rightarrow \infty$.

В основе метода Левенберга – Марквардта в качестве критерия оптимизации используются модели данных на основе обучающей выборки, а сущность авторского подхода заключается в поэтапной аппроксимации определенных начальных значений параметров к локальному оптимуму. Основой для построения метода Левенберга – Марквардта является регрессионная выборка $D = \{(x_k, y_k)\}_{k=1}^N$, $y = f(x)$ [9]. Сама же регрессионная модель задается в виде непрерывной и дифференцируемой функции $f(x_k, w)$. Второй параметр функции w определяет степень близости ожидаемых значений y_k и действительных значений y'_k . В таком случае задача обучения на основе критериев (4), (7) и (8) может быть сведена к минимизации следующей функции

$$E(w) = \sum_{k=1}^N (f(x_k, w) - y_k)^2, \quad (16)$$

$$w' = \arg \min_{w \in R^m} E(w).$$

Исходя из выражения (16), значение градиента и Гессиана, соответствующие функции $E(w)$, принимают следующий вид

$$\frac{dE}{dw_k} = 2 \sum_{k=1}^N (f(x_k, w) - y_k) \frac{df(x_k, w)}{dw_i}, \quad (17)$$

$$\frac{d^2E}{dw_k dw_j} = \left(\frac{df(x_k, w)}{dw_i} \right)^T \cdot \frac{df(x_k, w)}{dw_j} + \sum_{k=1}^N (f(x_k, w) - y_k) \frac{d^2 f(x_k, w)}{dw_i dw_j}.$$

С учетом того, что первое слагаемое представляет собой произведения матриц Якоби, обобщенное выражение для Гессиана можно переписать в следующем виде

$$H = J^T(w)J(w) + R(w), \quad (18)$$

где $R(w)$ – второе слагаемое выражения (18), описывающее вторые производные для функции $f(x_k, w)$.

С учетом того, что матрица $J^T(w)J(w)$ является положительно определенной, Левенбергом был введен параметр регуляризации $\lambda \geq 0$. Формула решения уравнения относительно Δw имеет следующий вид [10]

$$\Delta w = (J^T(w)J(w) + \lambda I(w))^{-1} J^T(w), \quad (19)$$

где $I(w)$ – единичная матрица.

Корректировка параметра регуляризации λ происходит на каждом шаге метода. Искомым значением Δw является значение, полученное на последней итерации. В свою очередь, классический метод Левенберга был преобразован Марквардтом, так как при увеличении λ , скорость аппроксимации $R(w)$ существенно снижается. Для этого единичная матрица $I(w)$ может быть заменена на диагональ Гессиана:

$$\Delta w = (J^T(w)J(w) + \lambda \text{diag}(J^T(w)J(w)))^{-1} J^T(w), \quad (20)$$

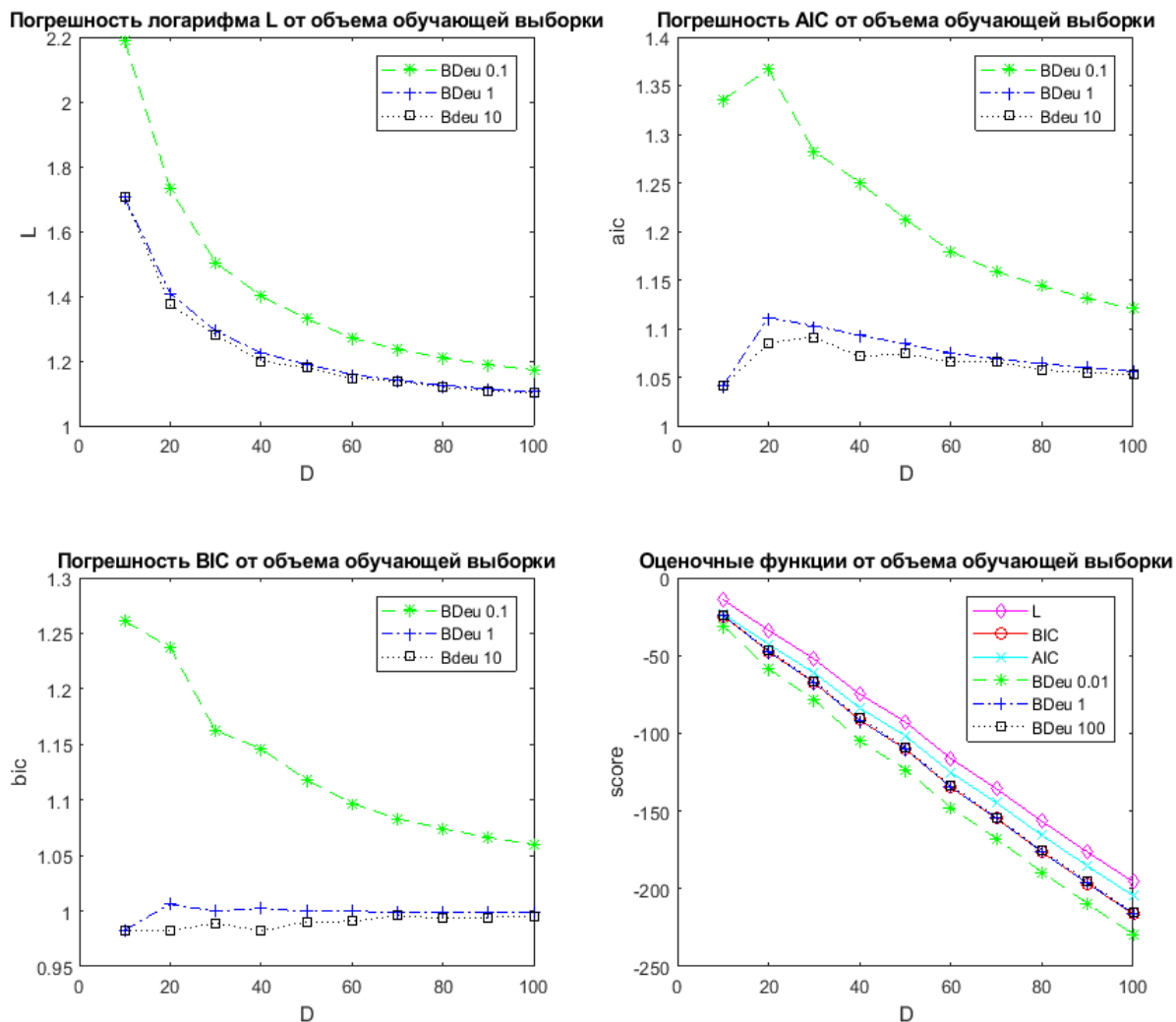
Выражение (20) описывает классический метод Левенберга – Марквардта. Исходя из того, что у метода Левенберга – Марквардта Гессиан пропорционален кривизне функция $f(w)$, то данное выражение будет приводить к малым итерациям при большой кривизне и большим при малой кривизне. Рассматривая данный метод применительно к процедуре обучения, видно, что вероятность попадания значений в локальные оптимумы достаточно мала, по сравнению с другими методами благодаря использованию параметра регуляризации λ . Это позволяет корректировать поведение метода в процессе проведения обучения. На начальных этапах выполнения метода, когда значения функции $f(w)$ не являются точными, используется метод Коши ($\lambda \gg 0$). В дальнейшем по мере означивания параметров w целесообразно брать значения $\lambda = 0$ и использовать метод Ньютона – Гаусса.

Проанализировав метод Левенберга – Марквардта, можно сказать, что его применение в процедуре обучения структуры динамической байесовской сети является обоснованным. В связи с этим разработанный гибридный алгоритм обучения структуры байесовской сети на основе объединения процедуры выполнения статистических тестов и метода Левенберга – Марквардта может быть разделен на два основных этапа.

На первом этапе происходит заполнение множеств $M_{t:t+1}$ на основе узлов-кандидатов в состав Марковского покрытия и исключения ошибочно добавленных переменных за счет выполнения χ^2 тестов. Тем самым происходит определение базовой (ненаправленной) структуры байесовской сети.

На втором этапе происходит определение направленности связей за счет вычисления оценок на основе метода Левенберга – Марквардта. Данный процесс характеризуется операциями добавления, удаления и изменения направленности связей, которые, в свою очередь, приводят к изменению оценок, назначаемых каждому из узлов сети.

Далее приведем зависимость критериев Шварца, Байеса – Дирихле, логарифма правдоподобия Акаике от размера обучающей выборки. Также оценим погрешность каждого из критериев методом Левенберга – Марквардта относительно критерия Байеса – Дирихле (см. рисунок).



Зависимость значений оценочных функций от объема обучающей выборки
методом Левенберга – Марквардта

Для проведения эксперимента по построению структуры байесовской сети были взяты сети: Alarm (37 узлов, 46 дуг, 509 параметров), Water (32 узла, 66 дуг, 10 082 параметра) и Mildew (34 узла, 46 дуг, 540 150 параметров). Таким образом, применение предложенного алгоритма позволит оптимизировать процедуру построения структуры динамической байесовской сети, а также адаптировать алгоритм к анализу временных связей между вершинами на основе математического аппарата Марковских цепей.

Заключение

Математические алгоритмы обучения структуры динамических байесовских сетей являются важным аспектом построения сложных разветвленных сетей обработки информации в современных информационно-коммуникационных системах. От реализации данных алгоритмов напрямую зависит эффективность применения процедур прогнозирования параметров сети методами вероятностного вывода. Применение процедур обучения параметров и структуры сети позволяет адаптировать и провести настройку сети на решения различного рода задач. В связи с этим разработка новых алгоритмов обучения структуры динамической байесовской сети является одним из приоритетных направлений развития инструментов моделирования с помощью динамических байесовских сетей. В рамках исследования, описанного в данной статье, разработан оригинальный алгоритм обучения структуры на основе метода Левенберга – Марквардта. Апробация применения данного алгоритма дала достаточно хорошие результаты. Алгоритм позволяет создать адаптивную самообучающуюся динамическую байесовскую сеть на основе построения прямых и

перекрестных связей между разнесенными во времени статическими байесовскими сетями. В статье представлена оценка качественных и количественных показателей алгоритма в процессе построения ряда динамических байесовских сетей из различных предметных областей. Эмпирические данные, полученные с использованием представленного алгоритма, доказывают эффективность выбранного математического аппарата и корректность выполнения основных процедур обучения структуры динамической байесовской сети.

Литература

1. Darwiche, A. *Modelling and Reasoning with Bayesian Networks* / A. Darwiche. – New York: Cambridge University Press, 2009. – 548 p.
2. Friedman, N. *Learning the structure of dynamic probabilistic networks* / N. Friedman, K. Murphy, S. Russel // *Proceedings of the Fourteenth conference of Uncertainty in artificial intelligence*. – San Francisco: Morgan Kaufman, 1998. – P. 139–147.
3. Тулупьев, А. *Байесовские сети, логико-вероятностный подход* / А. Тулупьев, С. Николенко, А. Сироткин. – СПб.: Наука, 2006. – 728 с.
4. Кельберт, М.Я. *Вероятность и статистика в примерах и задачах. Т. 1: Основные понятия теории вероятности и математической статистики* / М.Я. Кельберт, Ю.М. Сухов. – М.: МЦНМО, 2007. – 456 с.
5. Schwarz, G. *Estimation dimension of a Model* / G. Schwarz // *The Annals of Statistics*. – 1978. – Vol. 6, no. 2 – P. 461–464.
6. Рассел, С. *Искусственный интеллект: современный подход* / С. Рассел, Р. Норвиг. – М.: Вильямс, 2006. – 1408 с.
7. Азарнова, Т.В. *Разработка динамических байесовских моделей управления процессами тестирования веб-приложений* / Т.В. Азарнова, П.В. Полухин // *Актуальные проблемы прикладной математики, информатики и механики: материалы Междоунар. науч. конф.* – Воронеж: Научно-исследовательские публикации, 2017. – С. 490–498.
8. Стрижов, В.В. *Методы индуктивного порождения регрессионных моделей* / В.В. Стрижов. – М.: Вычислительный центр им. А.А. Дородницына РАН, 2008. – 62 с.
9. Вержбицкий, В.М. *Численные методы. Линейная алгебра и нелинейные уравнения* / В.М. Вержбицкий. – М.: Издат. дом «Оникс 21 век», 2005. – 432 с.
10. Васин, В.В. *Метод Левенберга – Марквардта и его модифицированные варианты для решения нелинейных уравнений с приложением к обратной задаче гравиметрии* / В.В. Васин, Г.Я. Пересторонина // *Труды института математики и механики УрО РАН*. – 2011. – Т. 17, № 2. – С. 53–61.

Азарнова Татьяна Васильевна, д-р техн. наук, зав. кафедрой математических методов исследования операций, Воронежский государственный университет, г. Воронеж; ivdas_92@mail.ru.

Баркалов Сергей Алексеевич, д-р техн. наук, декан факультета экономики, менеджмента и информационных технологий, зав. кафедрой управления строительством, Воронежский государственный технический университет, г. Воронеж; sbarkalov@nm.ru.

Полухин Павел Валерьевич, канд. техн. наук, преподаватель, Воронежский государственный университет, г. Воронеж; alfa_force@bk.ru.

Поступила в редакцию 28 августа 2018 г.

DEVELOPMENT OF THE HYBRID ALGORITHM OF TUTORING OF STRUCTURE OF DYNAMIC BAYESIAN NETWORK ON THE BASIS OF THE LEVENBERG-MARKVARDT METHOD

T.V. Asarnova¹, *ivdas_92@mail.ru*,
S.A. Barkalov², *sbarkalov@nm.ru*,
P.V. Polukhin¹, *alfa_force@bk.ru*

¹ Voronezh State University, Voronezh, Russian Federation,

² Voronezh State Technical University, Voronezh, Russian Federation

Dynamic Bayesian networks are used quite effectively for modeling complex stochastic processes of modern multi-user information and communication systems. Dynamic Bayesian networks are graphical probabilistic models that reflect topology and stochastic cause and effect relationships between elements of the handled simulated information processes. The construction of topologies of dynamic Bayesian networks that appropriately reflect the probabilistic and functional relationships between the elements of such processes is a main factor in the simulation using this tool. Network topology usually built either by expert means or be means of training. Training mechanisms allow to get spanning tree of the network, as well as to determine the conditional connections and their direction between the individual vertices of the network. In this article regard the usage of mathematical apparatus for testing statistical hypotheses based on conditional independency tests between random variables with the Pearson criteria, Schwartz, Akaike and Bayes-Dirichlet metrics. Unlike static Bayesian networks, when determining the structure of dynamic Bayesian networks, it is necessary to determine variables and relations between them not only within one slice, but also between variables of different slices, which implement transitive connections between the time slices that reflect functioning of a certain process or object. The construction of structure of transitive links between slices is a rather complex and problematic step in almost all existing algorithms. This article presents an algorithm for learning the structure of a dynamic Bayesian network based on the Levenberg-Marquardt method within the optimization of algorithms for constructing dynamic Bayesian networks with transitive links between slices.

Keywords: dynamic Bayesian networks, structure learning, stochastic values conditional independency statistical criteria's, Levenberg-Marquardt method.

References

1. Darwiche A. *Modelling and Reasoning with Bayesian Networks*. New York, Cambridge University Press, 2009. 548 p.
2. Friedman N., Murphy K., Russel S. Learning the Structure of Dynamic Probabilistic Networks. *Proc. of the Fourteenth Conference of Uncertainty in Artificial Intelligence*. SanFrancisco, Morgan Kaufman, 1998, pp. 139–147.
3. Tulup'yev A., Nikolenko S., Sirotkin A. *Bayesovskie seti, logiko-veroyatnostnyy podkhod* [Bayesian Networks, Logic-Probabilistic Approach]. St. Petersburg, Nauka Publ., 2006. 728 p.
4. Kel'bert M.Ya., Sukhov Yu.M. *Veroyatnost' i statistika v primerakh i zadachakh. T. 1: Osnovnye ponyatiya teorii veroyatnosti i matematicheskoy statistiki* [Probability and Statistics in Examples and Tasks. Vol. 1: Basic Concepts of Probability Theory and Mathematical Statistics]. Moscow, MTsNMO Publ., 2007. 456 p.
5. Schwarz G. Estimation Dimension of a Model. *The Annals of Statistics*, 1978, vol. 6, no. 2, pp. 461–464.
6. Russel S., Norvig R. *Iskustvennyy intellekt: sovremennyy podkhod* [Artificial Intelligence: Modern Approach]. Moscow, Vil'yams Publ., 2006. 1408 p.
7. Azarnova T.V., Polukhin P.V. [Development of Dynamic Bayesian Models of Management of Processes of Testing of Web Applications]. *Razrabotka dinamicheskikh bayesovskikh modeley upravleniya protsessami testirovaniya veb-prilozheniy. Aktual'nye problemy prikladnoy matematiki, informatiki*

i mekhaniki: materialy Mezhdunar. nauchn. konf. [Current Problems of Applied Mathematics, Informatics and Mechanics: Proc. of Int. Scient. Conf.]. Voronezh, Research Publications, 2017, pp. 490–498. (in Russ.)

8. Strizhov V.V. *Metody induktivnogo porozhdeniya regressionnykh modeley* [Methods of Inductive Generation of Regression Models]. Moscow, Computer Center of A.A. Dorodnitsyn of RAS Publ., 2008. 62 p.

9. Verzhbitsky V.M. *Chislennye metody. Lineynaya algebra i nelineynye uravneniya* [Numerical Methods. Linear Algebra and Nonlinear Equations]. Moscow, “Oniks 21 Vek” Publ., 2005. 432 p.

10. Vasin V.V., Perestoronina G.Ya. [The Levenberga-Markvardta Method and Its Modified Options for the Solution of the Nonlinear Equations with the Annex to the Return Problem of Gravitation Measurements]. *Works of Institute of Mathematics and Mechanics UB RAS*, 2011, vol. 17, no. 2, pp. 53–61. (in Russ.)

Received 28 August 2018

ОБРАЗЕЦ ЦИТИРОВАНИЯ

Азарнова, Т.В. Разработка гибридного алгоритма обучения структуры динамической байесовской сети на основе метода Левенберга – Марквардта / Т.В. Азарнова, С.А. Баркалов, П.В. Полухин // Вестник ЮУрГУ. Серия «Компьютерные технологии, управление, радиоэлектроника». – 2018. – Т. 18, № 4. – С. 16–24. DOI: 10.14529/ctcr180402

FOR CITATION

Asarnova T.V., Barkalov S.A., Polukhin P.V. Development of the Hybrid Algorithm of Tutoring of Structure of Dynamic Bayesian Network on the Basis of the Levenberg-Markvardt Method. *Bulletin of the South Ural State University. Ser. Computer Technologies, Automatic Control, Radio Electronics*, 2018, vol. 18, no. 4, pp. 16–24. (in Russ.) DOI: 10.14529/ctcr180402