Инновации в информационно-аналитических системах: сб. научн. трудов. Вып. 2 – Курск: Науком, 2011. - 93 с., ил. ISBN 978-5-4297-0003-8

УДК 004.9

Сизов А.С., Халин Ю.А.

О ВОЗМОЖНОСТИЯ ПРОГНОЗИРОВАНИЯ СОСТОЯНИЯ КОНКУРИРУЮЩИХ ПРЕДПРИЯТИЙ НА ОСНОВЕ ДСМ-МЕТОДА АВТОМАТИЧЕСКОГО ПОРОЖДЕНИЯ ГИПОТЕЗ

Предложен метод прогнозирования состояний конкурирующих предприятий, основанный на обработке разнородных данных с помощью автоматического порождения гипотез

В условиях рыночной экономики промышленные предприятия функционируют в конкурентной среде. Для руководства предприятий важной задачей является не только прогнозирование собственного развития, но и дальнейших действий основных конкурентов.

Конкурентная среда постоянно изменяется, вследствие этого руководству предприятия необходимо принимать управленческие решения не имея абсолютных правил, т.е. система управления должна в любой момент адаптироваться при поступлении информации (сигналов) из внешней среды. Кроме того она должна быть устойчива к шумам и выбросам в данных. Среди множества методов обучения (нейронные сети, эволюционное моделирование, вероятностные автоматы) следует выделить ДСМ-метод. Главным преимуществом данного метода является быстрое обучение (требуется относительно небольшое количество обучающих примеров), кроме того полученные правила ДСМ-метод выдаёт в явном виде, что облегчает интерпретацию результатов. Основным недостатком метода отмечают неприменимость в условиях противоречий в обучающих примерах. Таким образом, система управления предприятием должна быть интеллектуальной.

Согласно [1] интеллектуальная система определяется через иерархию (когнитивную технологическую магистраль):

1 уровень – интеллектуальная система, данные вводятся и интерпретируются экспертом;

2 уровень – когнитивная система, которая обладает перцепцией, т.е. может получать информацию из окружающей среды;

3 уровень - интеллектуальная система управления, которая обладает эффекторами, т.е. может влиять на окружающую среду.

Система третьего уровня должна обладать механизмом целеполагания, она должна выводить следствия из имеющейся ситуации.

ДСМ-метод формализует схему правдоподобного и достоверного вывода, называемую ДСМ-рассуждением. ДСМ-рассуждение является синтезом познавательных процедур: индукции, аналогии и абдукции. ДСМ-метод был создан как средство автоматизированного построения формализации знаний о предметной области средствами так называемых квазиаксиоматических теорий (КАТ).

ДСМ-метод оперирует сущностями трёх видов: объекты предметной области, свойства этих объектов, возможные причины свойств:

- **O** — множество объектов (конкурирующих предприятий);
- **P** — множество свойств объектов (целевых признаков),
- **C** — множество возможных причин свойств объектов (элементы структуры объектов),
- **V** — множество оценок. **V** = {+1, –1, 0, τ}.

Предполагается, что объекты имеют структуру и причинами свойств объектов являются фрагменты этой структуры.

На вход системы, реализующей ДСМ-метод подаётся некоторое множество изучаемых объектов и сведения об их структуре, о наличии или отсутствии у них определенных свойств, а также, в некоторых случаях, о связи между структурой объектов и их свойств. Кроме того имеется ряд целевых признаков, каждый из которых разбивает исходное множество объектов на четыре непересекающихся подмножества:

1. Объекты, про которые известно, что они обладают данным признаком (+1),
2. Объекты, про которые известно, что они не обладают данным признаком (-1),
3. Объекты, для которых существуют аргументы как за, так и против того, что они обладают данным признаком (0),
4. Объекты, о которых неизвестно, обладают они этим признаком или нет (**τ**).

Результатом применения ДСМ-метода являются гипотезы двух типов:

1. гипотезы о связи определенных структурных фрагментов изучаемых объектов со свойствами, которыми они обладают,
2. гипотезы о наличии или отсутствии целевых признаков у объектов, для которых изначально это было неизвестно, формируемые на основании установленной взаимосвязи между свойствами объектов и их структурными компонентами.

Для решения прогнозирования действий конкурирующих предприятий основными свойствами (целевыми признаками) будут являться показатели конкурентоспособности:

p_1 – уровень конкурентоспособности вырастет,

p_2 – уровень конкурентоспособности упадет,

p_3 – уровень конкурентоспособности не изменится.

Множество причин $C = \{c_i\}$ выступают факторы, влияющие на уровень конкурентоспособности предприятия. Эти факторы были рассмотрены в работе

[3]. Следует отметить также, что данные факторы основываются на данных конкурентной разведки.

В ходе конкурентной разведки добываются разнородные данные. ДСМ-метод позволяет обрабатывать разнородные данные, в том числе и плохо формализуемые. Главным условием применения является хорошее структурирование данных.

В общем случае ДСМ-метод включает в себя следующие этапы.

Введём функцию F: O×P→V, представляющую начальную ситуацию.

F(o, p) = +1 — известно, что объект *o* обладает свойством *p*;

F(o, p) = –1 — известно, что объект *o* не обладает свойством *p*;

F(o, p) = 0 — есть аргументы как за, так и против того, что объект *o* обладает свойством *p*;

F(o, p)=τ — неизвестно, обладает ли объект *o* свойством *p.*

Функция **F** может быть представлена в виде матрицы:

$$\begin{bmatrix} f_{11} & \cdots & f_{1j} \\ \vdots & \ddots & \vdots \\ f_{i1} & \cdots & f_{ij} \end{bmatrix}$$

Если $f_{ij} = \tau$, то говорят, что для пары (o_i, p_i) функция $F(o_i, p_i)$ недоопределена. Задача ДСМ-метода состоит в том, чтобы с помощью формирования гипотез доопределить исходную матрицу.

В общем случае ДСМ-метод включает в себя следующие этапы. На первом этапе формируются гипотезы о возможных причинах свойств. В результате получим функцию H: C×P→V.

H(C, P) = +1 — c является возможной причиной наличия свойства p или (+)-гипотезой;

H(C, P) = –1 — c является возможной причиной отсутствия свойства p или (-)-гипотезой;

H(C, P) = 0 — есть аргументы как за то, что c является причиной наличия свойства p, так и за то, что c есть причина отсутствия этого свойства или (+)-гипотезой (противоречивой гипотезой);

H(C, P) = τ — неизвестно, является ли c причиной наличия p или причиной отсутствия этого свойства.

Значения функции H для каждой пары (C, P) находятся с помощью правил правдоподобного вывода. Эти правила называются правилами первого рода. Сокращенное обозначение — PIR_1 (от Plausible Inference Rules). Правила первого рода можно рассматривать как функцию, использующую матрицу F для получения матрицы H, т.е.$H = PIR_1(F)$. Правила формулируются с помощью специальных предикатов M^+ и M^-. Предикаты M^+ используются для нахождения сходства для положительных примеров, а предикаты M^- для нахождения сходства для отрицательных примеров. Согласно этим правилам:

Через $M^+(F, C, P)$ будем обозначать тот факт, что с удовлетворяет (+)-условию для p относительно F .

Через $M^-(F, C, P)$ – тот факт, что с удовлетворяет (-)-условию для p относительно F .

Через $M^0(F, C, P)$ – тот факт, что с удовлетворяет (0)-условию для p относительно F .

$$H(C,P)=\begin{cases}+1,\ если\ M^+(F,C,P)\ \&\ \overline{M^-}(F,C,P)\ \&\ \overline{M^0}(F,C,P),\\ -1,если\ M^-(F,C,P)\ \&\ \overline{M^+}(F,C,P)\ \&\ \overline{M^0}(F,C,P),\\ 0,если\left(M^+(F,C,P)\ \&\ M^-(F,C,P)\right)\cup M^0(F,C,P),\\ \tau,если\ \overline{M^+}(F,C,P)\ \&\ \overline{M^-}(F,C,P)\ \&\ \overline{M^0}(F,C,P)\end{cases}$$

Гипотезы о причинах формируют новый фрагмент базы знаний, который на следующем этапе используется для порождения прогнозов.

Используя матрицу гипотез о возможных причинах, можно сформировать гипотезы о наличии или отсутствии свойства *P* у тех объектов *O*, для которых изначально не было известно, обладают они этим свойством или нет, т.е. для которых $F(O, P) = \tau$.

В результате мы получим функцию F': $O \times P \to V$. $F'(o, p) = F(o, p)$, если $F(O, P) \neq \tau$.

Значения функции F' находятся с помощью правил правдоподобного вывода. Эти правила называются правилами второго рода. Сокращенное обозначение — PIR_2. Правила второго рода можно рассматривать как функцию, использующую матрицы *F* и *H* для получения матрицы *F'*, т.е. $F' = PIR_2(F, H)$. Согласно этим правилам через $П^+(H, o, p)$, $П^-(H, o, p)$, $П^0(H, o, p)$ будем обозначать тот факт, что объект *o* для свойства *p* относительно *H* удовлетворяет (+)-условию, (–)-условию и 0-условию, соответственно.

Определим функцию $F'(S, P) = F(S, P)$, если $F(S, P) \neq \tau$. в противном случае:

$$F'(S,P)=\begin{cases}+1,\ если\ \Pi^+(H,S,P)\ \&\ \overline{\Pi^-}(H,S,P)\ \&\ \overline{\Pi^0}(H,S,P),\\ -1,если\ \Pi^-(H,S,P)\ \&\ \overline{\Pi^+}(H,S,P)\ \&\ \overline{\Pi^0}(H,S,P),\\ 0,если\left(\Pi^+(H,S,P)\ \&\ \Pi^-(H,S,P)\right)\cup\Pi^0(H,S,P),\\ \tau,если\ \overline{\Pi^+}(H,S,P)\ \&\ \overline{\Pi^-}(H,S,P)\ \&\ \overline{\Pi^0}(H,S,P)\end{cases}$$

Правила первого рода (процедура индукции) и правила второго рода (процедура аналогии) последовательно применяются до тех пор, пока в результате их работы порождается хотя бы одна новая гипотеза, т.е. применение правил первого рода приводит к изменению матрицы гипотез о возможных причинах свойств объектов, а применение правил второго рода – к изменению матрицы гипотез о возможном наличии или отсутствии свойства *p* у объектов. При этом номер шага является показателем правдоподобия рассуждений.

Заключительным этапом работы ДСМ-метода является проверка условия каузальной полноты. Проверка этого условия интерпретируется, как рассуждение по абдукции – условие выполняется, если полученные гипотезы объясняют исходные данные. Цель проверки условия состоит в том, чтобы определить, можно ли принимать полученные в результате работы метода гипотезы. Если условие каузальной полноты не выполняется, необходимо изменить применяемую когнитивную технику (например, выбрать другой способ кодирования структуры объектов) или входной набор объектов (как правило, набор расширяется).

Таким образом, после применения ДСМ-метода может быть три исхода:

1. Все факторы начального состояния базы фактов объясняются полученными гипотезами. Тогда гипотезы и предсказания принимаются, т.к. удовлетворяют критерию достаточного основания. Для решения нашей задачи определяются наиболее вероятные состояния конкурирующих предприятий.

2. Некоторые факты начального состояния базы фактов не объясняются, но большинство из них принимаются с учётом их недообоснованности. В нашем случае состояния конкурирующих предприятий определяются с учётом вероятности их ложности и возникает необходимость статистической проверки гипотез.

3. После расширения базы фактов появляются необъяснённые факты. В этом случае полученные гипотезы нельзя считать обоснованными.

Библиографический список

1. Автоматическое порождение гипотез в интеллектуальных системах [Текст]. Сост Е.С. Панкратова, В.К. Финн; Под общ. ред. В.К.Финна. Предисл. Ю.М. Арского – М.: Книжный дом «Либроком», 2009.

2. Фатхутдинов Р.А. Управление конкурентоспособностью организации [Текст] / Р.А. Фатхутдинов. – М.: Изд-во Эксмо, 2004. – 544с.

3. Халин, Ю. А. Математическая модель оценки конкурентоспособности предприятия на основе теории нечётких множеств. // Математические и компьютерные методы в технических, гуманитарных и общественных науках [Текст]. / Кол. монография по ред. проф. В.И. Левина. – Пенза: Приволжский Дом Знаний; МИЭМП, 2011. – 188с. (с. 7 – 19).