

# Применение нечетких систем управления для нахождения скрытых закономерностей в спортивной аналитике

А.А. Писаревский, В.В. Холмогоров

*Аннотация - В данной работе рассматривается технология применения аппарата нечеткой логики для решения проблем анализа данных на примере процесса нахождения скрытых закономерностей в спортивной аналитике.*

*Настоящая статья описывает простой метод расчета нечетких отношений для выявления новых ассоциативных правил на основе заданной базы знаний из спортивной аналитики, что и позволит выявить скрытые закономерности.*

*В качестве основного примера приведены несколько эмпирических правил, описывающих аспекты карьеры футболиста. Полученная нечеткая база знаний используется для расчета новых правил, которыми в дальнейшем возможно оперировать для последующего анализа карьеры игрока.*

*После проведения интеллектуального анализа игровой карьеры и применения методов нечеткой логики, можно сделать вывод, что достаточно собрать показатели нескольких аспектов профессиональной жизни футболиста, а также заранее составить несколько аксиом, чтобы определить, сколько приблизительно он зарабатывает и в клубе какого статуса он играет.*

*Тема применения нечеткой логики в сфере аналитики изучена крайне мало, так как нечеткая логика используется в основном при создании экспериментальных экспертных систем. Но ее применение также может быть ориентировано на анализ слабоструктурированных данных, в которых не всегда возможно обнаружить закономерности между элементами. Поэтому необходимо проводить дальнейшие исследования на эту тему.*

**Ключевые слова – нечеткая логика, продукционные отношения, нечеткие системы управления, интеллектуальный анализ, спортивная аналитика.**

## ВВЕДЕНИЕ

Одним из аппаратов, описывающих теоретическую компоненту реализации продуктов искусственного интеллекта, является нечеткая логика.

Благодаря внедрению нечетких множеств и, в частности, показателя степени принадлежности элементов, кондиционеры, к примеру, могут принимать не только количественный показатель температуры («сделай 25°C»), но и качественный («сделай теплее»).

Можно согласиться, что в большинстве ситуаций человеку гораздо удобнее оперировать подобным описанием, а не конкретными (или четкими) числами, которые еще нужно правильно подобрать.

Но, конечно, практическое применение нечеткой логики неограничено созданием умных бытовых систем. Не может быть оставлен без внимания интеллектуальный анализ данных, или data mining. В частности, очень популярно проведение исследований с использованием правил (аксиом) из базы знаний предметной области, что является ядром любой нечеткой системы управления. И, в частности, той, которая встроена в умный кондиционер и преобразует нечеткое «сделай теплее» в температуру комфорта.

Данная статья описывает простой метод расчета нечетких отношений для выявления новых ассоциативных правил на основе заданной базы знаний из спортивной аналитики, что и позволит выявить скрытые закономерности. Разумеется, существующие алгоритмы глубинного анализа данных, такие как Apriori и FP-дерево, также могут быть применены для задач подобного рода. Но при значительном увеличении входных данных и исследуемых факторов поиск ассоциативных правил становится все более трудоемким и затратным, а вышеприведенные методы – менее эластичными и эффективными.

Для наглядности и простоты ознакомления с предлагаемым подходом будут рассмотрены несколько эмпирических правил, затрагивающих аспекты карьеры футболиста. После этого будут сравнены результаты применения предлагаемого метода нахождения новых правил с выводом, полученным с помощью программы FisPro – инструмента проектирования нечетких систем управления.

## 1. НЕЧЕТКОЕ ПРЕОБРАЗОВАНИЕ И АКСИОМЫ О КАРЬЕРЕ СПОРТСМЕНА

Для работы с правилами настоящей предметной области сначала необходимо определить входные и выходные переменные системы управления.

Пусть на вход подаются продолжительность карьеры футболиста и его возраст – оба показателя, отражающих опыт и умение игрока. На выходе надо получить информацию об уровне клуба, за который выступает игрок, и размере его заработной платы в клубе.

Статья получена 20.12.2022 г.

А.А. Писаревский, МИРЭА (e-mail: pisarevskiy1977@gmail.com).

В.В. Холмогоров, МИРЭА (e-mail: Hv13@mail.ru).

Так как будет использоваться аппарат нечеткой логики, необходимо провести *фазификацию* (трансформацию четких сигналов в нечеткие множества) вышеупомянутых множеств путем определения категориальных групп значений для каждого множества. Например, для переменной «возраст» можно определить четыре категории: «очень молодой» (17-24), «молодой» (25-28), «старый» (29-32) и «очень старый» (33-40), где категории и границы интервалов каждой категории определены на основе эмпирических познаний (данное распределение содержит субъективный характер).

Аналогично, указываются нечеткие распределения для остальных трех переменных. При этом стоит упомянуть, что для переменных «продолжительность карьеры», «возраст» и «зарплата» предполагаются нормальные (Гауссовы) распределения величин, в то время как для переменной «уровень клуба», термы которой есть названия клубов (по одному из каждого уровня), – треугольное.

Теперь, после прохождения, по сути, первого этапа классической архитектуры нечеткой системы управления, следует перейти к построению набора аксиом, или базы знаний. Задаются три правила:

1. Если футболист опытный и молодой, то он получает высокую зарплату.
2. Если футболист неопытный и старый, то он не играет в топ-клубе.
3. Если футболист опытный и старый, то он играет в топ-клубе.

Все удобство нечетких отношений заключается в том, что на основе этих правил мы можем вывести новые, тем самым получив скрытые закономерности. То есть, зная информацию об опыте и возрасте игрока, возможно определить, сколько он зарабатывает и за клуб какого уровня он выступает.

Необходимо продемонстрировать валидность одного правила: если футболист опытный и молодой, то он играет в топ-клубе.

Данный пример является довольно примитивным, так как нужно вывести только одно правило. В крупномасштабных интеллектуальных системах входные нечеткие базы знаний гораздо более насыщены аксиомами, из чего можно вывести гораздо больший набор правил.

После того, как была проведена фазификация четких входных и выходных множеств и задана база ассоциативных правил, можно переходить к этапу валидации нашего целевого правила путем вычисления продукционных отношений.

## 2. ПРОДУКЦИОННЫЕ ОТНОШЕНИЯ И ВЫВОД НОВЫХ ПРАВИЛ

Для того, чтобы исследовать новые правила, необходимо сначала воспользоваться формулой расчета *матрицы отношений* для двух нечетких множеств  $U$  и  $V$ :

$$R = \sum_{i=1}^m \sum_{j=1}^n (\mu_U(u_i) \wedge \mu_V(v_j)) / (u_i, v_j) \quad (1),$$

где

- $R$  – матрица нечетких отношений на  $U \times V$ ;
- $u_i$  – элемент под номером  $i$  множества  $U$ ;
- $v_j$  – элемент под номером  $j$  множества  $V$ ;
- $\mu_U$  – функция принадлежности множества  $U$ ;
- $\mu_V$  – функция принадлежности множества  $V$ ;
- $m$  – размер множества  $U$ ;
- $n$  – размер множества  $V$ .

С помощью такого расчета можно вычислить матрицу отношений для простейшего правила с одной посылкой (антецедентом) и одним заключением (консеквентом) [1]. Но исходная база знаний содержит двумерные посылки: некоторый «опыт» и некоторый «возраст» футболиста даются в качестве условия. Поэтому необходимо обобщить процедуру составления матрицы продукционных отношений для случаев, когда в правиле имеется конъюнкция двух посылок и/или двух заключений.

Данное обобщение эквивалентно вычислению продукционных отношений (1) между переменными посылки/заключения (с использованием декартова произведения) и замене посылки/заключения найденной матрицей отношений.

После того, как были получены матрицы отношений для наших правил, можно переходить к их *композиции* (или *свертке*) – операции расчета нового правила из антецедента (посылки)  $X$  одного правила и консеквента (заключения)  $Z$  другого (пусть консеквент первого правила и антецедент второго равны  $Y$ ). Это действие и является ключевым в решении задачи: на основе двух правил и предварительного расчета их матриц отношений  $R_1$  и  $R_2$  производится свертка [2]:

$$\mu_{R_1 \circ R_2}(x, z) = \bigvee_y (\mu_{R_1}(x, y) \wedge \mu_{R_2}(y, z)) \quad (2),$$

где

- $\mu_{R_1 \circ R_2}$  – функция принадлежности нечеткого отношения между множествами  $X$  и  $Z$ ;
- $\mu_{R_1}$  – функция принадлежности нечеткого отношения между множествами  $X$  и  $Y$ ;
- $\mu_{R_2}$  – функция принадлежности нечеткого отношения между множествами  $Y$  и  $Z$ ;
- $x$  – элемент множества  $X$ ;
- $y$  – элемент множества  $Y$ ;
- $z$  – элемент множества  $Z$ .

Теперь рассмотренный выше аппарат можно применить для вывода (или опровержения) правила «если футболист опытный и молодой, то он играет в топ-клубе».

## 3. АНАЛИЗ ЗАВИСИМОСТИ КАРЬЕРНОГО УСПЕХА СПОРТСМЕНА ОТ ОПЫТА

В данном разделе поэтапно проводится нечеткий продукционный анализ следующих двух правил из предоставленной базы знаний:

1. Если футболист опытный и молодой, то он получает высокую зарплату.

2. Если футболист опытный и старый, то он играет в топ-клубе.

Следует дать комментарий: для генерации корректных правил необходимо совпадение заключения первого правила и посылки второго (как предполагается в операции свертки, описанной в разделе 2), но для генерации всех правил, которые потом пройдут проверку на корреляцию ввода и вывода (как далее будет показано в этом разделе) это совпадение не является обязательным.

Выполняется фазификация входных и выходных переменных, выраженных через конкретные термины:

```
experienced = {'1-3': 0.3, '4-6': 0.5, '7-10': 0.9, '11-15': 0.7}
young = {'17-24': 1.0, '25-28': 0.7, '29-32': 0.4, '33-40': 0.1}
top_club = {'Victoria': 0.2, 'Arsenal': 0.6, 'Real Madrid': 0.9, 'Bayern Munich': 1.0}
high_salary = {'< 0.1': 0.4, '0.1-0.2': 0.5, '0.2-0.5': 0.7, '0.5-1.0': 1.0}

# complements
not_experienced = {k:1.0-v for k, v in talent.items()}
not_young = {k:1.0-v for k, v in young.items()}
not_top_club = {k:1.0-v for k, v in top_club.items()}
not_high_salary = {k:1.0-v for k, v in high_salary.items()}
```

Рисунок 1 – Фазификация входных и выходных переменных в среде программирования JupyterLab

Далее вручную проводится преобразование посылок для возможности обобщить продукционные отношения для условий с двумя лингвистическими переменными.

А затем вычисляются матрицы продукционных отношений для каждой аксиомы (приведены фрагменты полученных матриц):

...	...
[0.4, 0.5, 0.7, 0.9]	[0.0, 0.0, 0.0, 0.0]
[0.4, 0.5, 0.7, 0.7]	[0.2, 0.3, 0.3, 0.3]
[0.4, 0.4, 0.4, 0.4]	[0.2, 0.6, 0.6, 0.6]
[0.1, 0.1, 0.1, 0.1]]	[0.2, 0.6, 0.7, 0.7]]

Для вывода нового правила необходимо произвести композицию (свертку) отношений двух известных аксиом, у одной из которых есть нужная посылка, а у другой – нужное заключение (2).

В конечном итоге, эти отношения сворачиваются в матрицу отношений правила, которое мы хотим вывести:

...
[0.2 0.3 0.3 0.3]
[0.2 0.3 0.3 0.3]
[0.2 0.3 0.3 0.3]
[0.1 0.1 0.1 0.1]]

Здесь приведен уникальный фрагмент конечной матрицы отношений, который демонстрирует монотонность вправо. Это говорит о том, что корреляция между antecedentом ("опытный И молодой") и консеквентом ("играет в топ-клубе") рассматриваемого правила является положительной, что и подтверждает его валидность.

Таким образом, после применения методов нечеткой логики и отношений, была расширена исходная база

знаний. Осталось сравнить полученный результат с выводом программы нечеткого преобразователя FisPro.

#### 4. ИНТЕЛЛЕКТУАЛЬНЫЙ АНАЛИЗ С ПОМОЩЬЮ ПРОГРАММЫ FISPRO

После применения ручного счета для вывода новых правил о карьере футболиста, выполняется та же самая задача, но уже с использованием программы FisPro, функционал которой основан на классической архитектуре нечетких систем управления [3].



Рисунок 2 – Схема архитектуры нечеткой системы управления

После выполнения данных механизмов на выходе получаются новые продукционные правила.

Если предоставить программе исходную базу аксиом о футболисте, которые были продиктованы ранее, а также эмпирическим путем задать распределения для входных и выходных нечетких множеств, то выходит целевое правило:

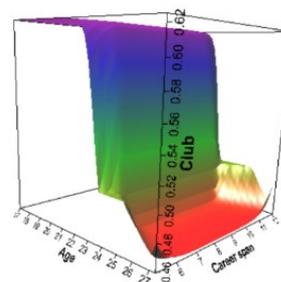


Рисунок 3 – График зависимости выходной переменной (Club) от входных (Age и Career span)

Как видно по 3D-графику зависимости уровня игры клуба от возраста и продолжительности карьеры игрока, если игрок молод (от 17 до 22 лет) и отыграл в профессиональный футбол более 5 лет, то клуб, за который он выступает, вероятно является топ-клубом.

#### ЗАКЛЮЧЕНИЕ

В данной статье был продемонстрирован несложный математический и вычислительный аппарат для расширения базы эмпирических знаний о карьере спортсмена, что может быть эффективным и полезным инструментом в спортивной аналитике.

После проведения интеллектуального анализа игровой карьеры и применения методов нечеткой логики, можно сделать вывод, что достаточно собрать показатели нескольких аспектов профессиональной жизни футболиста, а также заранее составить несколько аксиом, чтобы определить, сколько приблизительно он зарабатывает и в клубе какого статуса он играет.

Тема применения нечеткой логики в сфере аналитики изучена крайне мало и требует дальнейшего обсуждения и

исследования, так как использование предлагаемого метода может упростить анализ карьеры игроков, что является полезным, например, в период трансферных окон.

#### БИБЛИОГРАФИЯ

- [1] Круглов В.В., Дли М.И., Голунов Р.Ю. Нечеткая логика и искусственные нейронные сети. М.: Физматлит. — 2001. — 221 с.
- [2] Новак В., Перфильева И., Мочкрож И. Математические принципы нечёткой логики = Mathematical Principles of Fuzzy Logic. — Физматлит, 2006. — 352 с.
- [3] Зак Юрий Александрович. Принятие решений в условиях нечетких и размытых данных: Fuzzy-технологии. — М.: «ЛИБРОКОМ», 2013. — 352 с.
- [4] Усков А. А., Кузьмин А. В. Интеллектуальные технологии управления. Искусственные нейронные сети и нечеткая логика. — М.: Горячая Линия — Телеком, 2004. — 143 с.

# Applying fuzzy control systems to unveil hidden patterns in sports analytics

A.A. Pisarevsky, V.V. Kholmogorov

*Abstract - This paper examines the technology of applying fuzzy logic apparatus to solve problems of data analysis using the process of unveiling hidden patterns in sports analytics.*

*The article describes a simple method for computing fuzzy relations to identify new associative rules based on a given knowledge base from sports analytics, which will reveal hidden patterns.*

*As a basic example, several empirical rules describing aspects of a player's career are given. The resulting fuzzy knowledge base is used to calculate new rules, which can later be operated to further analyze a player's career.*

*After conducting an intellectual analysis of a player's career and applying fuzzy logic methods, we can conclude that it is enough to collect the indicators of several aspects of a football player's professional life, as well as to make some axioms in advance to determine how much he earns approximately and in which club he plays.*

*The topic of application of fuzzy logic in analytics is studied very little, because fuzzy logic is used mainly in creating experimental expert systems. But its application can also be focused on the analysis of weakly structured data, in which it is not always possible to detect patterns between elements. Therefore, it is necessary to conduct further research on this topic.*

**Key words - fuzzy logic, fuzzy relations, fuzzy control systems, intelligent analysis, sports analytics.**

## REFERENCES

- [1] Kruglov V. V., Dli M. I., Golunov R. Y. Nechyotkaya logika i iskusstvennye neyronnye seti. M.: Fizmatlit. — 2001. — 221 p.
- [2] Novak V., Perfil'eva I., Mochkrozhd I. Matematicheskie printsipy nechyotkoy logiki = Mathematical Principles of Fuzzy Logic. — Fizmatlit, 2006. — 352 p.
- [3] Zak Yuriy Aleksandrovich. Prinyatie resheniy v usloviyakh nechetkikh i razmytykh dannyykh: Fuzzy-tehnologii. — M.: «LIBROKOM», 2013. — 352 p.
- [4] Uskov A. A., Kuz'min A. V. Intellektual'nye tehnologii upravleniya. Iskusstvennye neyronnye seti i nechetkaya logika. — M.: Goryachaya Liniya — Telekom, 2004. — 143 p.