

Рубцова А.В.

Смоленский государственный университет, г. Смоленск, Россия

РАЗРАБОТКА СЕРВИСА ДЛЯ РЕШЕНИЯ ЗАДАЧ ПРОГНОЗИРОВАНИЯ МЕТОДАМИ НЕЧЕТКОЙ МАТЕМАТИКИ

АННОТАЦИЯ

В статье обсуждается теоретическая обоснованность выбора методов нечеткой математики при решении задач прогнозирования. Проводится краткий анализ популярных, распространенных методов решения и анализа решения задач прогнозирования в различных сферах повседневной жизни.

КЛЮЧЕВЫЕ СЛОВА

Прогнозирование; программирование; нечеткая математика; методы прогнозирования; разработка сервиса; решение задач прогнозирования; анализ данных.

Rubtsova A.V.

Smolensk State University, Smolensk, Russia

DEVELOPMENT OF SERVICE FOR THE SOLUTION OF TASKS OF FORECASTING BY METHODS OF INDISTINCT MATHEMATICS

ABSTRACT

In the article theoretical justification of the choice of methods of indistinct mathematics in case of the solution of tasks of forecasting a is discussed. The short analysis of popular, widespread methods of the decision and the analysis of the solution of tasks of forecasting in various spheres of everyday life is carried out.

KEYWORDS

Forecasting; programming; indistinct mathematics; forecasting methods; service development; solution of tasks of forecasting; analysis of data.

Сегодня перспективным является применение интеллектуальных технологий, методы теории нечетких множеств являются удобным средством моделирования сложных динамических процессов в условиях многокритериальности, ограниченности и неопределенности. Процесс развития жизнедеятельности того или иного явления есть процесс труднопредсказуемый, практически непрогнозируемым, который зависит от многих граничных условий и факторов, и в этих условиях целесообразно использовать нечеткие модели вывода. Вопросы практического применения нечетких моделей рассматривались в работах Х. Танаки, К. Сонга [1], однако универсальный подход к прогнозированию на данный момент не сформулирован. В зависимости от вида анализируемой деятельности существуют рекомендации, но определенного, общего подхода не существует.

Под «разработкой сервиса» понимается проектирование программного продукта, который будет удовлетворять требованиям: анализа, сбора и обработки информации. В нашем случае, с помощью такого сервиса требуется осуществить сервисное обслуживание, которое способно качественно повысить уровень жизни людей. Но достичь этого в действительности возможно только при условии, если обслуживаемый и обслуживающий понимают друг друга в процессе обслуживания. Вопрос адаптации сервиса к потребностям конечного пользователя стремится к вопросу о теоретическом фундаменте метода, на который разработчик склонен опираться-метод нечеткой математики. Данный аспект будет освещен в данной статье.

Следует рассмотреть исторический аспект рассматриваемого метода. Как известно, обобщениями теории множеств являются: математическая теория нечетких множеств (fuzzy sets) и нечеткая логика (fuzzy logic). Как известно, данные понятия были впервые предложены американским ученым Лотфи Заде (Lotfi Zadeh) в 1965 г. Основной причиной появления новой теории стало наличие нечетких и приближенных рассуждений при описании человеком процессов, систем, объектов. Первоначально, до того, как нечеткий подход к моделированию сложных систем

получил признание во всем мире, прошло много времени с момента отделения теории нечетких множеств, как отдельной науки. При этом, первый период (конец 60-х–начало 70 гг.) характеризуется развитием теоретического аппарата нечетких множеств ([2]). Второй период (70–80-е годы) интерпретируется появлением первых практических результатов в области нечеткого управления сложными техническими системами (парогенератор с нечетким управлением).

Одновременно стало уделяться внимание вопросам построения экспертных систем, построенных на нечеткой логике, в большей степени иллюстрируется на разработке нечетких контроллеров. Нечеткие экспертные системы для поддержки принятия решений находят широкое применение в таких отраслях как экономика, образование, медицина. Наконец, в третьем периоде, который длится с конца 80-х годов и продолжается в настоящее время, появляются пакеты программ для построения нечетких экспертных систем, а области применения нечеткой логики заметно углубляются. Она применяется в автомобильной, аэрокосмической и транспортной промышленности, в области финансов, анализа и принятия управленческих решений и многих других.

Безусловно, оба рассмотренных вида прогнозирования часто работают в совокупности, например, на основании сложного алгоритма вычисляются будущие значения временного ряда, а далее, эксперт проверяет эти цифры на адекватность. На этом этапе эксперт может внести ручные корректировки, которые при его высокой квалификации, способны положительно повлиять на качество прогноза. Работа нашего сервиса основана именно на сложном алгоритме вычисления будущих значений из совокупности данных.

Если требуется повысить точность экспертного прогнозирования, например, в задачах маркетинга (что непременно является яркой иллюстрацией), то нужно следовать ряду определенных рекомендаций. Если же перед разработчиком стоит задача повышения точности прогнозирования за счет сложных, быстрых, программно реализованных математических моделей, то стоит взглянуть в сторону консенсус-прогноза, то есть прогноза, составленного на основании обзора независимых прогнозов.

Относительно теоретической обоснованности выбора метода в разработке сервиса, следует отметить его недостатки, чтобы в будущем их нейтрализовать. Классические нечеткие системы обладают тем недостатком, что для формулирования правил и функций принадлежности необходимо привлекать экспертов той или иной предметной области, что не всегда удается обеспечить. Решением этого недостатка являются, как отмечают специалисты [3], адаптивные нечеткие системы (adaptive fuzzy systems). В таких системах подбор параметров нечеткой системы производится в процессе обучения на экспериментальных данных. Алгоритмы обучения адаптивных нечетких систем относительно трудоемки и сложны по сравнению с алгоритмами обучения нейронных сетей, и, как правило, состоят из двух стадий: 1. Генерация лингвистических правил; 2. Корректировка функций принадлежности. Первая задача относится к задаче переборного типа, вторая – к оптимизации в непрерывных пространствах. При этом возникает определенное противоречие: для генерации нечетких правил необходимы функции принадлежности, а для проведения нечеткого вывода – правила. Кроме того, при автоматической генерации нечетких правил необходимо обеспечить их полноту и непротиворечивость.

Значительная часть методов обучения нечетких систем использует генетические алгоритмы. В англоязычной литературе этому соответствует специальный термин – Genetic Fuzzy Systems [3]. Значительный вклад в развитие теории и практики нечетких систем с эволюционной адаптацией внесла группа испанских исследователей во главе с Ф. Херрера (F. Herrera) [4].

По нашему мнению, быстрым и удобным источником для сбора данных являются базы данных. Они известны простым обывателям, разнообразием своих возможностей и общей доступностью пользователям, которые владеют первоначальными навыками работы за компьютером. Нечеткие запросы к базам данных (fuzzy queries) – перспективное направление в современных системах обработки информации [2]. Данный инструмент дает возможность формулировать запросы на естественном языке, например: "Вывести список недорогих предложений о съеме жилья близко к центру города", что невозможно при использовании стандартного механизма запросов. Для этой цели разработана нечеткая реляционная алгебра и специальные расширения языков SQL для нечетких запросов. Большая часть исследований в этой области принадлежит западноевропейским ученым Д. Дюбуа и Г. Праде [6].

Нечеткие ассоциативные правила (fuzzy associative rules) – инструмент для извлечения из баз данных закономерностей, которые формулируются в виде лингвистических высказываний. [1] Здесь введены специальные понятия нечеткой транзакции, поддержки и достоверности нечеткого ассоциативного правила.

Нечеткие когнитивные карты (fuzzy cognitive maps) были предложены Б. Коско в 1986 г. [3]

и используются для моделирования причинных взаимосвязей, выявленных между концептами некоторой области. В отличие от простых когнитивных карт, нечеткие когнитивные карты представляют собой нечеткий ориентированный граф, узлы которого являются нечеткими множествами. Направленные ребра графа не только отражают причинно-следственные связи между концептами, но и определяют степень влияния (вес) связываемых концептов. Активное использование нечетких когнитивных карт в качестве средства моделирования систем обусловлено возможностью, во-первых, наглядного представления анализируемой системы, именно возможность иллюстрирования результатов работы является, на наш взгляд, самым привлекательным при получении итоговых результатов. Так, во-вторых легкостью интерпретации причинно-следственных связей между концептами. Иначе говоря, получение возможности быстрого и понятного синтеза и анализа полученного результата. Основные проблемы в данном случае связаны с процессом построения когнитивной карты, который не поддается формализации. Кроме того, необходимо доказать, что построенная когнитивная карта адекватна реальной моделируемой системе [7] Одним словом, что мы получим понятно и эффективно обработанную цепочку данных. Для решения данных проблем разработаны алгоритмы автоматического построения когнитивных карт на основе выборки данных.

К нечетким методам относятся колоссальное количество методов: нечеткие деревья решений, нечеткие сети Петри, нечеткие методы кластеризации нечеткая ассоциативная память, нечеткие самоорганизующиеся карты и другие гибридные методы.

Если подвергнуть сравнению привычную многим классическую логику и нечеткую в последней явно заметно отличие: вместо величин истина и ложь используется степень истинности, принимающая любые значения из бесконечного множества от 0 до 1 включительно. Следовательно, логические операции уже нельзя представить в виде таблицы. В нечеткой логике они задаются функциями.

Однако, надо помнить, что не существует универсальных методов прогнозирования на любые варианты зависимости чисел. Выбор метода прогнозирования и его эффективность зависят от многих условий, и в частности от требуемой длины или времени прогнозирования.

По времени прогнозирования различают краткосрочный, среднесрочный и долгосрочный прогноз. Краткосрочный прогноз характеризует собой прогноз «на завтра», то есть прогноз на несколько шагов вперед. Среднесрочный прогноз – это обычно прогноз на один или на половину сезонного цикла. А при построении долгосрочного прогноза стандартные статистические методы прогнозирования практически не используют, и требуется использование комплексных подходов. Например, использование нейронных сетей или регрессионных моделей.

Для построения прогноза, важно правильно понимать термины, используемые при построении моделей и хорошо ими оперировать. После построения любой модели важно проверять, насколько адекватно она построена, насколько верно она соответствует действительности. Для этого можно, во-первых, провести визуальный анализ со сдвигом прогноза на несколько шагов назад, иначе говоря, данный шаг можно охарактеризовать как приближенная оценка предполагаемого результата. А во-вторых, воспользоваться анализом остатков – стандартным методом проверки адекватности любой построенной модели. Таким образом, главная функция прогнозирования заключается в проведении научного анализа связанных с тематикой исследования процессов и тенденций, а также в предвидении новых ситуаций и выявлении характерных проблем. Основные функции прогнозирования состоят также в исследовании объективных связей явлений и процессов в конкретных условиях, результаты которого мы планируем обработать с помощью сконструированного сервиса. Хотелось бы также отметить следующие две характерные особенности: адекватность и рентабельность прогнозирования. Под адекватностью понимается максимальное приближение прогнозной модели к реальной действительности, тенденциям, закономерностям. Рентабельность прогнозирования заключается в том, чтобы эффект от разработанного прогноза должен превышать затраты на его разработку.

Числовые данные масштабируются в принятый для сети диапазон. Обычно данные масштабируются по линейной шкале. При обработке некоторых данных используются алгоритмы, которые автоматически находят масштабирующие параметры для преобразования числовых значений в нужный диапазон. [8] На первый план, в общем случае, выходит формулирование и применение алгоритма, который с помощью функции способен автоматически сгенерировать результат и обработать массив числовой информации.

Так, нам становится более понятно, как действовать с числовыми характеристиками. Велика вероятность, что нам придется пользоваться нечисловыми данными. Примеров данному явлению существует колоссальное количество. Например, в образовании можно сгенерировать

данные по численности педагогического и ученического составов. Для достаточно трудоемко запоминать имена, фамилии, место жительства с помощью чисел. По крайней мере, потому что для этого понадобится специалист, который запомнит все числа соответствующие данным и понадобится сложнейшая функция для нахождения характеристик статистики.

Чаще всего нечисловые данные представляют в виде номинальных переменных. Номинальные переменные могут быть двузначными (например, Пол = {Мужчина, Женщина}) или многозначными (то есть принимать больше двух значений). Двузначную номинальную переменную легко превратить в числовую (например, Мужчина = 1, Женщина = 0. Такой способ представления информации может иметь место и в других отраслях нашей повседневной жизни, например, в медицине (формирование электронной очереди, количество пациентов, медицинский состав учреждения). Данные и общую систематизацию данных мы встречаем повсеместно.

Обратимся к общей схеме работы сервиса (рис. 1). Конечный пользователь получает результат работы при помощи обработки выполненного им запроса некоторой базы данных, включающей в себя комплекс числовых характеристик некоторого явления. Отметим также, что задача нашего сервиса состоит в формировании некоторого прогноза, методом нечеткой математики, что выделяет его работу среди подобных. В свою очередь, подобная работа уникальна по своему проявлению, поскольку задача баз данных заключена не только в работе над сбором и анализом, но и над синтезом полученной информации. Значительная часть информации имеет нечеткую структуру и построение прогнозов на основе методов нечеткой математики нечетких временных рядов должно быть составной частью комплексной системы прогнозирования и принятия решений.

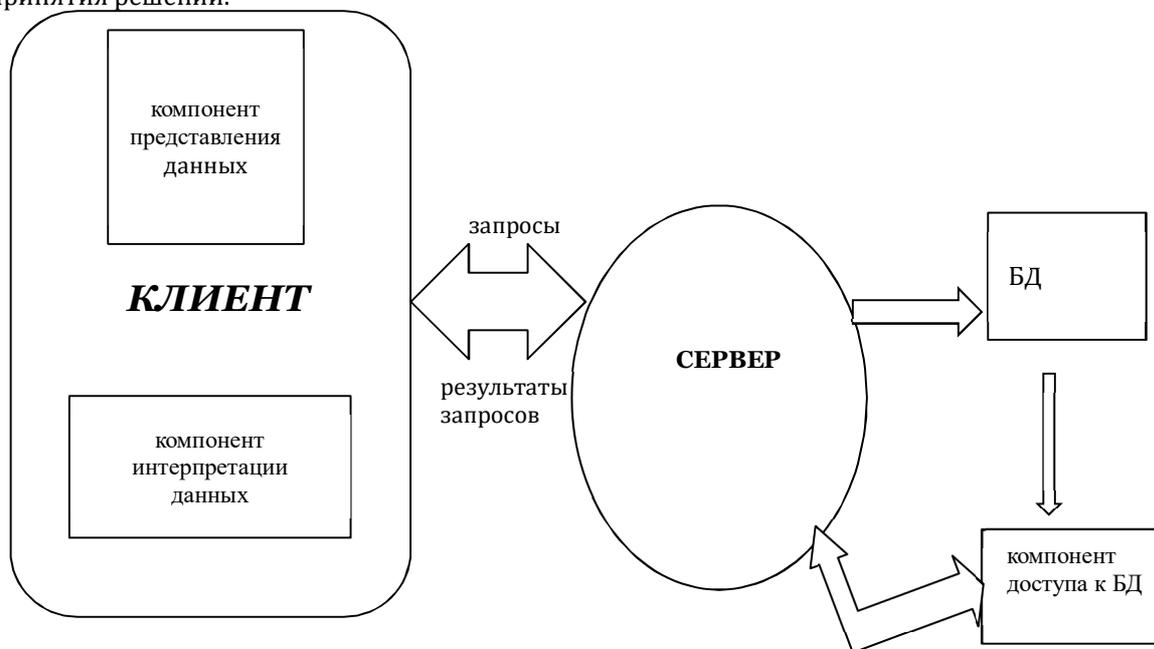


Рис. 1. Общая схема работы сервиса

В заключении рассмотрим такой аспект, как не совсем достоверные данные. Например, когда физик делает замеры сигнала - значения некоторых переменных могут быть испорчены шумом или частично отсутствовать. Такие переменные можно заменить на среднее значение этой переменной. Если данных немного, можно включить в рассмотрение случаи с пропущенными значениями.

Принципы прогнозирования обеспечивают методологическое единство разнообразных методов и моделей прогнозирования. После получения предвиденных значений при наличии правильных можно получить абсолютные и относительные отклонения на всем контрольном множестве, для каждого шага прогнозирования. При наличии удовлетворительных результатов прогнозирования можно считать, что полученный сервис будет максимально точно и цело отвечать общим требованиям. Далее рассмотрим методы нечеткой математики и их интерпретацию.

Говоря об основании выбора именно таких методов-нечеткой математики хотелось бы выделить следующую идею. Известно, что все операторы языков программирования требуют именно четких условий, так получается всегда приходится из нечеткой степени истинности получать четкий критерий срабатывания. Данное явление можно проиллюстрировать на таком примере: мы знаем, что живые организмы обладают свойством раздражимость. Получается если

взять три одинаковых организма, их жизненный цикл будет протекать в нормальном режиме, спокойное состояние. Добавив к одному организму-раздражитель, мы получим реакцию, а значит режим спокойствия нарушается. Такое явление имеет название дефаззификация, именно основываясь на нём будет работать разрабатываемый сервис. Работа выполнена под руководством к.ф.-м.н., доцента Кристалинского В.Р.

Литература

1. Cox I J., Lewis R. W., Rasing R.S. and etc. Application of neural computing in basic oxygen steelmaking // J. of Materials Processing Techn., 2002.V.120, №1-3.-p.310-315.
2. BaseGroup Labs URL:<https://basegroup.ru/community/articles/fuzzylogic-math>
3. Заде Л. Понятие лингвистической переменной и его применение к принятию приближенных решений. – М.: Мир, 1976.
4. Kosko B. Fuzzy systems as universal approximators // IEEE Transactions on Computers, vol. 43, No. 11, November 1994. – P. 1329-1333.
5. Cordon O., Herrera F., A General study on genetic fuzzy systems // Genetic Algorithms in engineering and computer science, 1995. – P. 33-57.
6. Дюбуа Д., Прад А. Теория возможностей. Приложения к представлению знаний в информатике.
7. Прикладные нечеткие системы / Под ред. Т. Тэрано, К. Асаи, М. Сугэно. – М.: Мир, 1993.- 368 с.
8. Chen S.M. Forecasting enrollments based on fuzzy time series // Fuzzy Sets Systems, 1996, vol. 81, no. 3, pp. 311-319.

References

1. Cox I J., Lewis R. W., Rasing R.S. and etc. Application of neural computing in basic oxygen steelmaking // J. of Materials Processing Techn., 2002.V.120, №1-3.-p.310-315.
2. BaseGroup Labs URL:<https://basegroup.ru/community/articles/fuzzylogic-math>.
3. Zade L. Ponyatie lingvisticheskoy peremennoy i ego primeneniye k prinyatiyu priblizhennykh resheniy. – М.: Mir, 1976.
4. Kosko B. Fuzzy systems as universal approximators // IEEE Transactions on Computers, vol. 43, No. 11, November 1994. – P. 1329-1333.
5. Cordon O., Herrera F., A General study on genetic fuzzy systems // Genetic Algorithms in engineering and computer science, 1995. – P. 33-57.
6. Dyubua D., Prad A. Teoriya vozmozhnostey. Prilozheniya k predstavleniyu znaniy v informatike
7. Prikladnye nechetkie sistemy / Pod red. T. Terano, K. Asai, M. Sugeno. – М.: Mir, 1993.- 368 s.
8. Chen S.M. Forecasting enrollments based on fuzzy time series // Fuzzy Sets Systems, 1996, vol. 81, no. 3, pp. 311-319.

Поступила 12.10.2016

Об авторах:

Рубцова Алиса Валерьевна, магистрант физико-математического факультета Смоленского государственного университета, 409223@gmail.com.