

АЛГОРИТМЫ ЭКСПЕРТНО-СТАТИСТИЧЕСКОЙ ОБРАБОТКИ ИНФОРМАЦИИ В ЗАДАЧАХ ПРИНЯТИЯ РЕШЕНИЙ

Мандель А.С.

Институт проблем управления им. В.А. Трапезникова РАН

almandel@yandex.ru

Аннотация. Рассматривается проблема создания систем поддержки принятия решений (СППР) в условиях неопределенности. Отмечено, что во многих прикладных задачах наиболее адекватным математическим средством анализа действий СППР в условиях неопределенности являются экспертно-статистические методы обработки информации. Основное внимание уделяется методам повышения уровня достоверности работы СППР. Описывается разработанный для этого испытательный стенд.

Ключевые слова: системы поддержки принятия решений, условия неопределенности, экспертно-статистическая обработка информации, испытательный стенд.

Введение

Исследуется задача создания системы поддержки принятия решений (СППР) [1, 2] в условиях неопределенности. Во многих случаях использование термина «неопределенность» означает, что процесс принятия решений осуществляется в отсутствие исходной математической модели ситуации принятия решений. Более того, нередко даже в результате накопления достаточной для построения более или менее точной модели статистической информации выясняется, что выработанные по такой модели рекомендации по принимаемым решениям отвергаются лицами, принимающими решения (ЛПР). Это означает, что в формируемой в результате обработки собранной информации математической модели отсутствуют некие важные для процесса принятия решений атрибуты, которые на интуитивном уровне известны экспертам. При этом в роли ЛПР, как правило, выступают высококвалифицированные эксперты по той предметной области, к которой относится создаваемая СППР.

Очерченный выше класс систем требует для своего исследования, а также проектирования соответствующих СППР математического инструмента, развитого в работах автора настоящего доклада [3, 4] и получившего название экспертно-статистической обработки информации. Одним из важнейших средств экспертно-статистической обработки является так называемый метод аналогов [5 – 7], применение которого позволяет организовать конструктивный диалог между экспертом-пользователем и СППР. На основании такого диалога осуществляется уточнение модели и повышается степень адекватности вырабатываемых СППР вариантов решений. Чтобы оценить

эти точность и степень адекватности необходимо ставить серьезные и дорогостоящие эксперименты с СППР, цена которых в значительной степени обусловлена тем, что к участию в них привлекаются высокооплачиваемые специалисты-эксперты.

Настоящая работа развивает результаты исследований [6] и [8] по созданию вспомогательной по отношению к СППР системы, которую в дальнейшем будет называть испытательным стендом.

1 Испытательный стенд, как средство анализа эффективности СППР

Функция описываемого испытательного стенда (имитационной системы) состоит в моделировании действий экспертов (в дальнейшем, сокращенно Э) при прогнозировании временных рядов с использованием метода аналогов. Система моделирует такие действия Э, как выбор аналогов, принятие решений о корректировке списка аналогов, а также решений о непосредственной корректировке значений прогнозного ряда. С помощью различных параметров система имитирует особенности поведения Э, обладающих различными психологическими характеристиками и уровнями профессионализма, включая такие, как степень профессионализма, склонность к завышению («оптимист») или занижению («пессимист») прогнозов и т. д. Имеется два режима работы с системой, представляющих два варианта ее применений. Первый вариант – «интерфейсный» – позволяет отслеживать каждый шаг процедуры моделирования, просматривать все промежуточные решения системы и выводить окончательные результаты прогнозирования в виде таблиц и графиков. Второй вариант системы – автоматический – сводится к выполнению большого числа сеансов моделирования, в результате чего в базе данных (БД) системы накапливается значительный объем статистической информации, на основании которой можно формировать более или менее определенные выводы о качественных характеристиках исследуемого способа формирования прогнозов.

Так как основной функцией стенда является моделирование действий реальных Э, то очевидно, что достоверность результатов работы системы напрямую зависит от того, насколько точно в системе отражены особенности поведения Э при формировании прогнозов с помощью метода аналогов. Ниже перечислены типовые ситуации, возникающие при прогнозировании по методу аналогов, а также возможные при этом действия и решения Э. Приведенные описания основаны на наблюдении за работой Э при прогнозировании временных рядов с помощью метода аналогов. В данном параграфе внимание сосредоточено на особенностях поведения Э, на попытках объяснить и описать причины тех или иных действий Э и на выборе вытекающих из поведения Э значений параметров имитационной среды и СППР.

Количество выбираемых Э объектов-аналогов. При выборе объектов-аналогов возникает вопрос о том, какое число аналогов является, с одной стороны, минимально необходимым, а, с другой, достаточным для описания объекта прогнозирования (ОП). Практика показала, что, как правило, Э старается выбрать не менее двух аналогов, так как выбор одного аналога подразумевает практическую однозначность прогноза значений временного ряда. Кроме того, при выборе только одного аналога возникает ощущение, что Э просто не захотел тратить время на поиск дополнительных объектов-аналогов, которые позволили бы уточнить прогноз для ОП. С другой стороны, выбор большого числа аналогов также не получил широкого распространения, так как, во-первых, большое количество сильно различающихся между собой аналогов говорит о неуверенности Э при формировании прогнозов, а выбор значительного числа похожих друг на друга аналогов является явно нецелесообразным, так как в этом случае добавление дополнительного аналога практически не отражается на результате прогнозирования.

Реакция Э на малое расхождение между прогнозом и фактическими значениями временных рядов. При поступлении первых статистических данных об ОП эксперт, руководствуясь тем, насколько эти данные похожи на сформированный прогноз, может внести корректировки в список аналогов. Впрочем, эксперименты показали, что вполне вероятны ситуации, когда прогноз оказывается достаточно близким к поступающим статистическим данным, но Э все равно принимает решение осуществить корректировку списка аналогов. Подобное поведение может иметь несколько объяснений. Во-первых, Э может вспомнить о каком-либо объекте из данной предметной области, который должен присутствовать в списке аналогов, но не был включен в него ранее по некоторым субъективным причинам (скажем, Э о нем просто забыл). Эксперт также может добавить в список аналогов объект, который не внесет значительных корректировок в результирующий прогноз, но способен как бы «подтвердить» прогноз Э, сформированный на основе ранее выбранного списка аналогов. Или, напротив, в подобной же ситуации эксперт может удалить из списка аналогов объект, который не вносит каких-либо корректировок в прогноз, и

дублирует другой объект из списка аналогов. Кроме того, несмотря на близость поступающих статистических данных к сформированному прогнозу, Э может принять решение о том, что ОП может изменить свое поведение в будущем, поэтому необходимо скорректировать список аналогов, а, следовательно, и прогноз для данного ОП. В любом случае, корректировка списка аналогов при близости поступающих значений временных рядов к значениям прогноза является пусть маловероятным, но вполне допустимым действием эксперта.

Реакция Э на значительное расхождение между прогнозом и фактическими значениями временных рядов. Рассмотрим ситуацию, когда поступающие значения временного ряда заметно отличаются от значений прогнозного ряда. Очевидно, что в этом случае Э может пожелать изменить значения прогноза для следующих точек ряда с достаточно большой вероятностью, так как поступающие фактические данные указывают на значительную ошибку в прогнозе. Реакция Э на подобную ситуацию определяется несколькими факторами. Во-первых, эксперименты показали, что вплоть до определенного уровня расхождение между статистикой и прогнозом не является для Э существенным. Причем речь идет не только о малых величинах ошибок. Иногда даже относительно большие отклонения могут восприниматься Э как не слишком существенные. Подобное поведение может быть связано с особенностями конкретной предметной области и ситуацией, в которой осуществляется прогнозирование. Имеется корреляция и с психологическим типом того индивидуума, который является экспертом («холерик» реагирует почти всегда, «флегматик» никогда не спешит). Границы, в пределах которых расхождение между статистикой и прогнозом не является для Э существенным, называются *областью безразличия*.

В том случае, когда ошибка прогноза выходит за рамки области безразличия, вероятность того, что Э среагирует на выявленное расхождение и внесет в прогноз соответствующие корректировки, немала. Причем минимальное значение этой вероятности, то есть вероятности внесения изменений в список аналогов, достаточно велико, но почти никогда не равно единице. Для того, чтобы данная вероятность равнялась единице, расхождение между прогнозом и статистикой должно быть значительным (что бывает нечасто, если Э, действительно, профессионален), так как в противном случае Э может проигнорировать ошибку прогноза. Причины подобного пренебрежения могут быть разными. Например, Э может счесть, что выявленное расхождение относится только к уже ранее зарегистрированным значениям, и что, несмотря на это расхождение, будущие значения, все равно, будут близки к сформированному прогнозу. Таким образом, существует функциональная зависимость между величиной расхождения и вероятностью внесения изменений в список аналогов. На границе области безразличия данная вероятность принимает свое минимальное значение, а при некотором достаточно большом уровне расхождения эта вероятность достигает своего максимально возможного значения, то есть равна единице.

Ручная корректировка прогнозов, построенных на основе аналогов. После того, как Э при поддержке ЭСС сформировал список аналогов и ввел необходимые для дальнейшей работы коэффициенты, начинается процедура построения экспертно-статистической системой прогноза временного ряда на основе полученной от Э информации. После того, как такой прогноз просчитан, выясняется, что Э достаточно охотно (то есть с большой вероятностью) соглашается (или хочет) скорректировать значения ближайших прогнозных точек временного ряда. Это объясняется двумя обстоятельствами.

Во-первых, насколько бы, по мнению Э, аналоги ни были близки к ОП, каждый новый объект может обладать своими уникальными особенностями, которые отсутствуют у других, похожих на него объектов. Подобное особенно актуально в предметных областях, которые связаны с большим числом определяющих ситуацию случайных факторов. Вторая особенность носит, скорее, психологический характер: как правило, Э хочет, чтобы «последнее слово», окончательное решение оставалось за ним, а не за программой. Все это приводит к тому, что сформированный на основе аналогов прогноз с некоторой немалой вероятностью корректируется (хотя бы незначительно) Э. При этом подправляются, как правило, только некоторые ближайшие к моменту корректировки точки временного ряда, так как, во-первых, указать точные значения отдаленных в будущее точек временного ряда сложнее, и, во-вторых, в момент корректировки значения удаленных точек могут представлять для Э меньший интерес, по сравнению с ближайшими точками ряда.

Величина осуществляемых Э корректировок зависит от профессионализма Э. Как правило, профессиональный Э с помощью корректировок добивается уточнения прогноза, то есть старается приблизить значение точки прогнозного ряда к ожидаемому тем же экспертом фактическому значению временного ряда. Значительные корректировки значений прогнозного ряда

маловероятны, поскольку таким образом Э как бы расписывается в том, что выбранные им в качестве аналогов объекты таковыми не являются.

Длина прогнозного ряда. Решение о том, на какой отрезок времени необходимо сформировать прогноз для ОП, принимается Э самостоятельно или совместно с ЛПР. В любом случае, продолжительность данного отрезка определяется двумя факторами: актуальностью прогноза и возможностями Э. Актуальность в данном случае означает, что прогноз может быть интересен для Э и/или для ЛПР только в течение определенного, ограниченного отрезка времени. Это, как правило, объясняется особенностями конкретной предметной области и спецификой рассматриваемых временных рядов. Например, в некоторых случаях подлинный интерес для Э могут представлять только первые несколько точек или десятков точек временного ряда, так как значения последующих точек будут малыми и не внесут каких-либо изменений в стратегию деятельности, основанную на формируемом Э прогнозе. Кроме того, особенности конкретной предметной области могут оказаться таковы, что позволят разбивать временной ряд на отрезки, и осуществлять прогнозирование для каждого отрезка в отдельности независимо от других отрезков. Кроме того, Э зачастую просто не в состоянии сделать прогноз для точек «отдаленного будущего», поскольку точность прогноза подобных точек может оказаться крайне невысока. Таким образом, прогноз, как правило, осуществляется на некоторый ограниченный, фиксированный отрезок времени.

Момент последней корректировки прогноза. Предположим, что Э прогнозирует значения временного ряда, в котором определяющими являются только первые точки, а остальные играют значительно меньшую роль, и практически не влияют на решения, принимаемые на основе формируемых прогнозов. Очевидно, что в этом случае принципиально важным оказывается прогноз самых первых точек ряда. Жизнь показывает, что во многих случаях Э может утратить интерес к корректировке прогноза после того, как уже поступили фактические значения первых точек, поскольку эффект (например, экономический) от уточнения прогнозных значений последующих точек резко снижается, и, даже несмотря на возможные значительные ошибки в прогнозе будущих значений, прогноз может быть оставлен без изменений.

Группы похожести. При выборе аналогов из БД системы Э не только опирается на собственную память, но и руководствуется рекомендациями системы. Сделав предположения о будущем поведении ОП, эксперт начинает отбирать объекты, временные ряды которых, по его мнению, близки к предполагаемому временному ряду ОП. При этом Э первоначально пытается отнести ОП к той или иной группе объектов, имеющих определенные атрибуты или представленные отрезками временных рядов, которые отличают объекты этой группы от других объектов из БД системы. При этом различия между объектами в рамках конкретной группы могут являться для Э менее существенными или вообще несущественными. В частности, отнесенные к одной группы объекты могут оказаться настолько похожи друг на друга, что конкретизация аналогов в рамках группы может стать для Э задачей куда более сложной, чем выбор самой группы.

Делая предположения о принадлежности ОП к конкретной группе, Э может допустить ошибку, обусловленную недостаточностью или недостоверностью его знаний, отсутствием опыта, или какой-либо иной причиной. Данная ошибка, как правило, становится очевидной только после того, как для данного ОП поступили первые измеренные (наблюденные) значения временного ряда, и эти значения доказывают, что, на самом деле, данный ОП расклассифицирован Э неправильно, то есть относится к другой группе, и для него необходимо подбирать новые аналоги из соответствующей группы.

Количество групп, на которые можно разбить объекты, зависит от конкретной предметной области. Разброс может быть достаточно широким. С одной стороны, чем больше используется групп, тем точнее получаются прогнозы, формируемые с помощью метода аналогов. Однако, с другой стороны, чем больше групп, тем больше вероятность отнести ОП к неправильной группе.

2 Профессиональные и психологические характеристики экспертов

Очевидно, что результаты прогнозирования зависят от профессионализма Э. При этом ЭСС может рекомендовать Э внести те или иные коррективы, изменить списки аналогов или выполнить какие-то иные операции. Однако в ситуации, когда для данного ОП полностью отсутствует какая бы то ни было статистическая информация, ЭСС не может оценить, в частности, то, насколько верно Э выбрал группу, к которой он отнес ОП. На основе наблюдений за работой экспертов можно утверждать, что уровень профессионализма Э при использовании метода аналогов проявляется в двух основных операциях: при работе со списком аналогов и при ручной корректировке значений

сформированного на основе аналогов прогноза. При работе со списком аналогов – в процессе добавления или удаления объектов из списка – профессионализм Э определяется тем, насколько точно он выбирает группу, к которой, по его мнению, относится ОП, а, следовательно, насколько точно он подбирает аналоги для ОП. Непрофессиональный Э может сильно ошибаться при отнесении ОП к той или иной группе или – в крайнем, но вполне вероятном случае – вообще не разбираться в объектах данной предметной области.

При ручной корректировке прогноза профессионализм Э определяется тем, насколько точно он «угадывает» будущие значения ряда. Возможны ситуации, когда для конкретного ОП в БД отсутствуют близкие к нему, объекту прогнозирования, объекты-аналоги. Однако даже, когда такие аналоги имеются, Э, тем не менее, все-таки может захотеть подкорректировать сформированный на их основе прогноз. В этом случае Э проводит ручную корректировку значений прогнозного ряда, и чем профессиональнее поведение Э, тем ближе скорректированные им значения оказываются к будущим фактическим значениям ряда.

Возможны различные комбинации описанных выше проявлений профессионализма. Например, Э может хорошо ориентироваться в группах, на которые разбиваются объекты из БД, но при этом он не будет готовым к внесению уточняющих корректировок в прогнозный ряд. Возможна и обратная ситуация: Э не слишком свободно ориентируется в объектах из БД, но при этом способен достаточно точно прогнозировать будущие значения характеризующего ОП временного ряда. Наиболее профессиональные эксперты способны не только точно классифицировать объект и относить его к правильной группе, но и корректировать значения сформированного на основе аналогов прогноза так, чтобы эти значения оказывались еще более близкими к будущим фактическим значениям временного ряда ОП.

Помимо уровня профессионализма Э, как и все люди, может относиться к определенному психологическому типу, являться пессимистом или оптимистом и обладать рядом других особенностей. В результате сформированные Э прогнозы могут содержать систематические ошибки (завышения или занижения). Скажем, профессиональный эксперт-оптимист, несмотря на способность к точной классификации ОП (то есть относя его к правильной группе), может с помощью ручных корректировок постоянно завышать значения прогноза прибыли в силу одного лишь своего природного оптимизма.

3 Настраиваемые параметры имитационного стенда

В данном разделе приводится описание основных параметров, заложенных в имитационный стенд. Все настраиваемые параметры стенда являются отражением особенностей поведения реальных Э при прогнозировании временных рядов с использованием метода аналогов.

Минимальное и максимальное число аналогов. Эти параметры обозначают, соответственно, минимальное и максимальное число аналогов, которое может быть отобрано в системе на каждом шаге процесса моделирования. По умолчанию, минимальное количество аналогов равно двум, максимальное – пяти.

Вероятность «случайной» корректировки экспертом списка аналогов. Параметр обозначает вероятность «случайной» корректировки системой списка аналогов, то есть вероятность корректировки списка аналогов в том случае, когда полученные статистические данные близки к прогнозу и находятся в области безразличия. По умолчанию, вероятность случайной корректировки равна 0,05.

Реакция системы на несоответствие между статистическими данными и прогнозом. Данный параметр обозначает вероятность того, что система примет решение о корректировке списка аналогов в том случае, когда наблюдается расхождение между статистическими данными и прогнозом, которое выходит за пределы области безразличия. При принятии решений о корректировке списка аналогов данный параметр используется совместно с величиной расхождения между прогнозом и статистическими данными. По умолчанию, «исходное» значение параметра равно 0,3.

Область безразличия. Данный параметр характеризует такое расхождение между статистическими данными и прогнозом, которое является допустимым с точки зрения точности прогноза. Границы области безразличия выражаются в процентах. Формула расчета расхождений описана в соответствующем разделе. По умолчанию, значение параметра равно 15%.

Вероятность изменения значений прогноза. На каждом шаге процесса моделирования система строит прогноз на базе выбранных аналогов. Затем значения некоторых точек полученного прогноза могут быть «случайным» образом скорректированы. Данный параметр обозначает

вероятность корректировки значения каждой прогнозной точки. По умолчанию, значение параметра равно 0,3.

Число прогнозных точек, значения которых могут быть скорректированы. Данный параметр обозначает число точек, следующих за той последней по времени точкой, фактическое значение которой уже известно, и обладающих той особенностью, что моделируемый эксперт (то есть испытательный стенд) склонен к тому, чтобы корректировать их значения. По умолчанию, значение данного параметра равно 3.

Число точек, которые учитываются при выборе аналогов. Данный параметр обозначает число точек временного ряда, используемых при вычислении объектов, близких по своим временным рядам к объекту ОП. По умолчанию, значение параметра равно 10.

Длина прогнозного периода. Данный параметр обозначает максимальное число шагов (точек, начиная с первой от текущего момента времени), на которое осуществляется прогнозирование временного ряда. По умолчанию, значение параметра равно 30.

Последняя корректировка списка аналогов. Данный параметр обозначает номер итерации, на которой система в последний раз принимает решение о корректировке списка аналогов. Подразумевается, что после этого момента список аналогов становится оптимальным. По умолчанию, значение параметра равно 5.

Число групп похожести. Для формирования списков объектов, похожих на ОП, все объекты из БД разбиваются на группы похожести относительно ОП. Каждая группа похожести характеризуется нижней и верхней границей расстояния до ОП. Данный параметр обозначает число групп похожести, на которые разбиваются объекты БД. По умолчанию, значение данного параметра равно 5.

Границы групп похожести. Для каждой группы похожести задаются верхняя и нижняя границы расстояний до ОП.

Вероятности удаления объектов, входящих в группы похожести, из списка аналогов. На каждой итерации процесса моделирования система принимает решение о корректировке списка аналогов. Корректировка списка подразумевает удаление из списка аналогов одних объектов и добавление в него других. Данные параметры характеризуют вероятности того, что объекты, входящие в соответствующие группы похожести, будут удаляться из списка аналогов. Для групп похожести, близких к ОП, эта вероятность мала; для групп, удаленных от ОП, эта вероятность велика.

Профессионализм эксперта. Данный параметр является двухмерным, характеризует степень профессионализма эксперта и влияет на точность всех выполняемых системой вычислений. Первая составляющая параметра показывает, насколько точно моделируемый Э умеет «угадывать», к какой группе или классу относится ОП, или, иными словами, насколько точно он умеет подбирать объекты, аналогичные ОП. Данная составляющая параметра «профессионализм эксперта» может принимать целочисленные значения на отрезке [1, 5], причем, чем больше значение параметра, тем выше профессионализм Э. Вторая составляющая параметра показывает, насколько точно моделируемый эксперт умеет «угадывать» сами значения прогнозируемого ряда. Данная составляющая также принимает целочисленные значения в интервале [1, 5], и снова, чем больше ее значение, тем профессиональнее действует моделируемый Э.

Склонность эксперта к завышению или занижению прогноза. Данный параметр характеризует склонность Э к завышению или занижению прогнозов, или, иными словами, степень его оптимизма или пессимизма. Значение этого параметра выбирается независимо от значения параметра, характеризующего степень профессионализм Э, и принимает одну из трех «градаций»: «оптимист», «пессимист» и «нейтральный эксперт».

Параметр, использующийся при вычислении среднеквадратического отклонения скорректированного значения прогноза. В том случае, когда на соответствующем шаге моделирования система принимает решение о корректировке значения прогноза, сформированного на основе аналогов, новое значение рассматриваемой прогнозной точки считается имеющим нормальное распределение со средним в виде исходного значения. Значение данного параметра определяет среднеквадратическое отклонение случайной величины, которая принимается в качестве нового значения данной прогнозной точки. По умолчанию, значение параметра равно 6.

Заключение

Описан испытательный стенд (система имитационного моделирования), предназначенная для экспериментальной оценки степени адекватности математических моделей, используемых в

системах поддержки принятия решений (СППР). Это позволяет своевременно формировать оценки уровня достоверности рекомендуемых СППР вариантов принимаемых решений.

Литература

1. *Трахтенгерц Э.А.* Компьютерная поддержка принятия решений. М.: СИНТЕГ, 1998.
2. *Трахтенгерц Э.А.* Компьютерная поддержка переговоров при согласовании управленческих решений. М.: СИНТЕГ, 2003.
3. *Мандель А.С.* Экспертно-статистические системы в задачах управления и обработки информации: часть I. // Приборы и системы управления, 1996, N^o12.
4. *Мандель А.С.* Экспертно-статистические системы в задачах управления и обработки информации: часть II. // Приборы и системы управления, 1997, N^o1.
5. *Беляков А. Г., Мандель А.С.* Прогнозирование временных рядов на основе метода аналогов (элементы теории экспертно-статистических систем). Препринт. М.: ИПУ, 2002.
6. *Мандель А.С.* Метод аналогов в прогнозировании коротких временных рядов: экспертно-статистический подход. // Автоматика и телемеханика. 2004. №4.
7. *Mandel A.* Expert-Statistical Data Processing Using the Method of Analogs / Proceedings of the 11th IEEE International Conference on Application of Information and Communication Technologies (AICT2017, Moscow). М.: IEEE, 2017. Vol. 1. P. 147-151.
8. *Мандель А.С.* Моделирование действий экспертов в процессе принятия ими прогностических решений. // Автоматизация в промышленности, 2004, № 7 – С.50–54.