

Анализ экспертной информации при технологических прогнозах

На основе разработанного обобщенного алгоритма экспертного опроса проводится отбор перспективных, разрабатываемых и осваиваемых технологий. Предлагаемая методика используется для получения более обоснованных результатов за счет применения одновременно нескольких различных методов анализа экспертных оценок. Применяются две группы методов обработки данных: численные и турнирные. Сопоставление результатов технологического прогнозирования с возможностями их реализации приводит к выбору предпочтительного перечня технологий и стратегии развития экономики.

Прогнозирование перспективных, разрабатываемых и осваиваемых технологий* включает анализ технологического потенциала отраслей, оценку рейтинга технологий в отраслях, и оценку рейтинга технологий по трем основным критериям: а) значимость и актуальность; б) инвестиционная привлекательность; в) коммерческий потенциал.

Сопоставление результатов технологического прогнозирования с возможностями их реализации приводит к выбору перечня технологий и к выработке соответствующей стратегии развития экономики.

Методика экспертного оценивания включает четыре основных этапа:

1. Сбор предложений по решению научных проблем и формирование исходных списков по различным направлениям.
2. Формирование экспертных комиссий по различным направлениям.
3. Формирование критериев оценки проектов (прогнозных решений).
4. Оценивание проектов по выбранным критериям и их ранжирование

для последующего обсуждения и принятия окончательного решения.

После определения целей и этапов экспертизы формируется рабочая группа по проведению опросов, состоящая из руководителя работ и аналитиков, которые непосредственно занимаются обработкой экспертных суждений.

• Под перспективными, разрабатываемыми и осваиваемыми технологиями далее будем понимать прогнозные решения или проекты.

Формирование исходных списков проектов

Рабочая группа организует сбор научных предложений по представлению специалистов институтов РАН, вузов и других учреждений. Полученные предложения распределяются по определенным направлениям.

Информация о проектах должна отражать следующие позиции: 1) наименование проекта; 2) разработчики проекта; 3) решаемые задачи;

4) краткая аннотация с ретроспективным анализом проблемы; 5) особенности методики и основные научные характеристики; 6) планируемый период реализации.

Отбор экспертов

При формировании состава экспертных комиссий по определенным направлениям необходимо учитывать: а) профессиональный уровень и авторитет в научной области; б) опыт научной деятельности; в) авторитет в научной среде. Кроме того, группа не должна состоять из представителей одной узкой специальности, одной научной школы или одного учреждения.

Большую роль в оценке компетентности экспертов может сыграть *анкета эксперта*, содержащая персональные, профессионально-квалификационные данные и информацию об области научных интересов.

Для оценки компетентности экспертов могут быть использованы различные косвенные методы. Они подразделяются на две группы: 1) методы, основанные на оценке объективных качеств, самооценки, взаимной оценки, профессиональных признаков (таких как стаж работы, должность, ученая степень, и т.д.); 2) методы, основанные на оценке степени влияния данного эксперта на согласованность группы. Последняя группа методов эффективна лишь при достаточной однородности группы экспертов.

Практика показывает, что на этапе предварительного формирования экспертной комиссии можно использовать метод самооценки. Этот метод позволяет при минимальных затратах времени и без проведения трудоемких экспертиз на стадии отбора произвести достаточно надёжную оценку компетентности. Суть процедуры заключается в подготовке специальных анкет, заполняя которые эксперт показывает свою эрудицию и аналитические способности. Важно, чтобы эксперт, заполнил анкету в короткий срок (5-10 мин.).

Примерное обращение к кандидату в эксперты может выглядеть следующим образом: «Вас просят оценить вопросы соответственно объёму Ваших знаний. Прочтите внимательно все вопросы и выберите тот, который Вы знаете лучше других. Дайте этому вопросу оценку 5. Затем найдите вопрос, который Вы знаете хуже других и дайте ему оценку 1. После этого оцените оставшиеся вопросы относительно этих двух, используя баллы 1, 2, 3, 4, 5. Таким образом, вопрос, о котором Вы знаете почти столько же, как и по первому, также получит оценку 5. Вопрос, расположенный примерно в «середине» следует оценить в 3 балла и т.д.

Заметим, что оценка является относительной и зависит от объёма Ваших знаний по данной проблематике. Не старайтесь улучшить оценки – руководствуйтесь Вашим личным впечатлением».

Получив эти данные, можно рассчитать суммарный численный индекс, как индивидуальной компетентности эксперта, так и групповой по каждому вопросу.

Подготовка экспертизы

Подготовка экспертизы начинается с создания рабочей группы. Руководителем группы, ответственным за решение проблемы является начальник соответствующего департамента компании (организации). Далее в рабочую группу должны входить ведущие специалисты компании и отраслевых институтов по данной проблематике. Рабочая группа формирует вспомогательную группу экспертизы, куда включаются специалисты, занимающиеся непосредственно подготовкой экспертизы, сбором и обработкой экспертных суждений.

Рабочая группа формирует экспертную комиссию для проведения экспертизы. В состав экспертной комиссии включаются научные работники, компетентные в соответствующей области, производственники, хорошо знающие практические и конкретные вопросы внедрения НТП на отраслевых производствах, а также представители администрации, имеющие большой опыт организации разработки и внедрения результатов НТП.

Чтобы избежать снижения точности и надежности экспертизы, в экспертную группу нужно ввести не менее 5 человек. Известный опыт проведения экспертных опросов подобного рода показывает, что оптимальным следует считать состав комиссии из 8 – 10 экспертов.

При составлении списка экспертов следует учитывать также и представительность группы – желательно, чтобы экспертная группа состояла из представителей нескольких научных направлений по предметной области экспертизы.

Один из экспертов назначается (или выбирается) руководителем экспертной комиссии. Он является также членом рабочей группы. Через него осуществляется связь и взаимодействие между рабочей группой и экспертной комиссией.

Рабочая группа составляет подлежащий экспертизе исходный список оцениваемых объектов, который раздается экспертам для анализа с описанием целей экспертизы.

Работа экспертов над исходным списком состоит в том, что они исключают ненужные или добавляют недостающие, по их мнению, объекты в списке, то есть устанавливают список базовых объектов (базовый список).

Базовые списки анализируются рабочей группой и затем составляется окончательный список объектов по следующему правилу: если более половины всех экспертов исключили (добавили) объект, то он не входит (входит) в окончательный список.

Формирование критериев оценки проектов

В основе методического подхода к решению этой задачи лежит метод так называемого аналитического иерархического вида с приданием *весов значимости* критериям на различных уровнях

иерархии. Подобная структура позволяет на каждом уровне и в каждой группе близких по смыслу критериев производить отбор, ранжирование и оценивать их приоритетность.

Применительно к рассматриваемой задаче структура системы критериев представляет собой двухуровневое иерархическое дерево, на верхнем уровне которого находятся группы однотипных критериев, отвечающих различным целям оценивания. Нижний уровень дерева определяют непосредственно сами критерии, по которым эксперты оценивают проекты. Критериально-целевая система оценивания формируется рабочей группой.

По каждому критерию разработана качественная балльная шкала из небольшого числа градаций. Это, с одной стороны, облегчает работу эксперта, а с другой, - повышает надежность и обоснованность получаемых оценок.

Для лучшей дифференциации при отборе проектов желательно определить относительную важность или *приоритетность* критериев, как на уровне групп, так и отдельно внутри каждой группы. С этой целью применяется разработанная информационная технология экспертных исследований, опирающаяся на математический аппарат анализа и обработки суждений эксперта, который будет рассмотрен ниже.

Необходимым этапом формализации процесса выработки оптимальной стратегии научно-технического развития является идентификация критериально-целевой структуры управления. Такая идентификация возможна только в рамках её важнейшего раздела – теории целенаправленных иерархических систем. Согласно этой теории, системный подход предполагает качественную структуризацию оптимизационной задачи путем построения так называемого «дерева целей». Для этого необходимо выяснить, что понимает ЛПР под глобальными целями, а также какие подцели и критерии обеспечат требуемую формализацию конечных целей. В рассматриваемой проблематике конечные цели должны предполагать создание и последующее внедрение новшеств, а это, в свою очередь, определяет промежуточные цели по направлениям научно-технического развития, соответствующим организациям и предприятиям.

Поскольку цели дают качественное описание предпочтительности (оптимальности), нужно найти соответствующие им критерии, которые выражали бы это описание в количественном виде, то есть были измеримы в физических шкалах, либо оценены экспертами. В теории систем цели и критерии объединяются в систему на основе задания их иерархии. Для этого используется понятие конкретизации – подразделения, разбиения цели на подцели более низкого уровня, что способствует разъяснению смысла целей вышестоящего уровня [3]. Так строится иерархия «сверху вниз» - агломеративное дерево целей. Сначала определяются новые требования к деятельности отрасли (или региона и т.п.) в перспективном периоде с перечнем критериев оценки деятельности отрасли. Далее устанавливается уровень детализации целей научно-технического развития. Например: отрасль в целом; основные комплексные технологии для производства однотипной продукции; элементы комплексных технологий; составные части этих элементов, включая технологию, оборудование и кадры [1].

Важно, чтобы при разделении целей на подцели каждый из аспектов цели более высокого уровня был учтен в одной из подцелей. При этом надо предотвратить разрастание дерева, как по горизонтали, так и по вертикали.

Процесс построения иерархической системы может происходить не единственным образом. Однако в любом случае, следуя принципам иерархичности и конкретизации, система должна быть построена так, чтобы полученный на нижнем уровне набор критериев отвечал следующим условиям:

- полнота – охват всех важных сторон и аспектов проблемы;
- информативность – полезность заключенной в них информации для анализа;
- разложимость–возможность упрощения процесса оценивания результатов, то есть независимого рассмотрения отдельных целей;
- избыточность – отсутствие дублирования при учете целей.

После того, как структурированы цели и критерии оценки научно-технического развития, необходимо установить их приоритетность, что является вторым этапом рассматриваемой задачи идентификации. При этом на каждом уровне иерархии приоритетность или сравнительная важность элементов относительно элементов более высокого уровня должна быть выражена в количественной форме – в виде весовых коэффициентов или весов.

Таким образом, назначение методики состоит в использовании экспертных данных для построения критериально-целевой структуры и вычисления весов приоритетности критериев. Полученные веса в дальнейшем используются при моделировании оптимального выбора перспективных комплексных технологий.

Идентификация критериально-целевой структуры научно-технического развития социально-экономических систем предполагает существенное использование информации, получаемой от высококвалифицированных специалистов – экспертов.

Более неоднозначная ситуация имеет место при рассмотрении методов экспертного опроса для определения весов приоритетности целей и критериев в задачах принятия решений, что является предметом исследований теории экспертных оценок.

Математическая теория экспертных оценок является довольно молодым научным направлением. За последние два десятилетия экспертные методы получили бурное развитие, как в теоретическом, так и в прикладном аспектах. Происходит разработка понятийного аппарата, формирование стройной логической схемы получения, анализа и обработки экспертной информации.

Если абстрагироваться от конкретного содержания задачи определения весов приоритетности целей и критериев, то возникает известная задача упорядочения однотипных объектов. По данной проблематике имеется огромное число публикаций, ввиду ее актуальности и отсутствия точных методов решения. Дело в том, что объекты, подлежащие упорядочению, не могут быть измерены в физической шкале сравнения или, как принято говорить, имеют нечисловую природу. Поэтому в основе предлагаемых методов лежит экспертное оценивание. Однако, несмотря на множество имеющихся методов, проблема получения весов приоритетности для корректного использования в задачах

многокритериального принятия решений остается далекой от достаточно полного решения. Дело в том, что подавляющему большинству предлагаемых процедур присущ ряд серьезных недостатков, существенно ограничивающих возможности их применения на практике. В первую очередь это касается тех способов получения экспертных оценок, где предлагается, что эксперт может дать необходимую информацию почти в любом виде.

В ряде методов считается, что эксперт может непосредственно сравнить объекты и выставлять им численные оценки или, например, может оценить интервал между уровнями предпочтительности объектов и затем попарно сравнивать эти интервалы. В интерактивной процедуре эксперт в процессе диалога должен определить, удовлетворяют ли его вычисленные на ЭВМ веса и, если нет, скорректировать управляющий параметр модели, представляющий собой отношение максимального и минимального весов.

Существует ряд распространенных методов, в которых необоснованно используются арифметические операции над экспертными оценками относительно шкал, в которых они измерены [13]. Примером может служить так называемый метод средних рангов, в котором ранги объектов, назначенные группой экспертов, усредняются и полученные величины считаются искомыми весами, хотя известно, что в порядковых шкалах операция осреднения некорректна.

Существуют подходы, в которых весовые коэффициенты определяются на основе информации о парных сравнениях многомерных объектов и их критериальных оценках [6,15,16]. Однако, во-первых, такие процедуры требуют решения сложных оптимизационных задач и, во-вторых, полученные веса могут быть использованы только для данного конкретного множества объектов, что снижает их практическую ценность.

Большую группу составляют методы, основанные на вероятностной трактовке экспертных суждений. Здесь предполагается, что суждения могут изменяться под действием случайных факторов [17]. Отсюда следует, что оценка важности объекта также является случайной величиной, а эксперт оценивает каждую ее реализацию. Поскольку от одного эксперта нельзя получить набор суждений при независимых повторных испытаниях, то требуемые частоты находятся в результате опроса группы экспертов, которую считают однородной [17]. Таким образом, статистические модели требуют, как минимум, наличия больших экспертных групп и проверки их однородности.

Для рассматриваемой здесь задачи наиболее известными и разработанными являются процедуры, в основе которых лежит модель Терстоуна [5, 6], а также некоторые другие методы, использующие, например, в качестве исходной информации матрицы частот предпочтений экспертов [18].

Анализ существующих методов экспертных оценок приводит к необходимости выработки иного концептуального взгляда на характер экспертной информации. Следует исходить из того, что компетентный эксперт проводит оценку возможных вариантов предпочтений в достаточной степени субъективно и выражает свои суждения продуманно. Факторы же, способные повлиять на его мнение носят главным образом не случайный характер, а имеют нестатистическую природу, связанную с размытостью понятий и категорий человеческого мышления и многозначностью языка.

Поэтому говорить о распределении вероятностей индивидуальных суждений нет достаточных оснований, ибо нет возможности оценить воспроизводимость экспертных оценок и их статистическую устойчивость. Следовательно, в условиях, не обеспечивающих предпосылки вероятностной концепции, можно применять более адекватный математический аппарат, опирающийся на идеи теории нечетких множеств [19]. Понятия нечеткого множества и лингвистической переменной позволяет исследователю моделировать неопределенности, не имеющие стохастической природы (связанные, скажем, с ошибками измерения). В частности, для описания высказываний эксперта удобно использовать термы соответствующих лингвистических переменных, которые в свою очередь формализуются через функции принадлежности нечетких множеств [19]. Таким образом, обработка и анализ экспертных оценок сводится к оперированию с функциями принадлежности, для чего имеется развитый математический аппарат.

Такой подход согласуется с трудно оспариваемой точкой зрения, согласно которой человеку значительно легче давать качественную вербальную информацию, нежели количественные оценки. Кроме того, качественные суждения, по мнению большинства психологов, являются устойчивыми к модификациям структуры целей эксперта. Таким образом, качественные оценки значительно более надежны, нежели количественные, хотя и менее информативны. Отсюда следует важный вывод: если получаемая от экспертов информация имеет качественный характер, то и найденные веса приоритетности также следует считать приближенными, так как после обработки нечеткой вербальной информации нельзя получить точные количественные результаты.

Такой подход к решению обсуждаемой проблемы идентификации целевой структуры лежит в основе разработки методики экспертного спроса и специальных экспертных процедур анализа и обработки экспертных оценок.

Методика построения критериально-целевой структуры

Первым этапом идентификации критериально-целевой структуры научно-технического развития социально-экономической системы является построение дерева целей – агломеративной иерархической структуры целей, подцелей и критериев деятельности системы.

Анализ ситуации на основе обобщения известных подходов позволяет предложить достаточно простую процедуру решения этой задачи. Приведем описание методики работы с экспертами.

1. Реализуем подготовительный этап экспертизы. С этой целью образуется инициативная группа по проведению экспертного опроса, в которую могут входить ЛПР – руководитель работ, ответственный за решение проблемы, несколько экспертов, выбранных по рекомендации ЛПР, и один или несколько человек, которых условно назовем аналитиками. В задачу последних входят – подготовка и проведение экспертизы, а также ее математическое обеспечение (анализ и обработка информации).

2. Инициативная группа определяет глобальные цели исследования и основные факторы, характеризующие степень достижения этой цели. Далее группа вырабатывает, так называемые базовые

или первоначальные списки элементов на каждом уровне иерархии создаваемого дерева целей. Естественно при этом выявляются связи «сверху-вниз» между элементами различных уровней.

3. Поскольку базовые списки могут быть где-то неполными, а где-то избыточными, следующий этап экспертизы состоит в формировании расширенной экспертной комиссии, в которую в индивидуальном порядке передаются подготовленные списки.

4. Работа эксперта со списками заключается в том, что он по своему усмотрению может добавить или исключить какие-то элементы в списках, после чего скорректированные списки передаются аналитикам для обобщения и анализа.

5. Ввиду того, что несогласованность суждений экспертов неизбежна, для выработки и принятия окончательного решения целесообразно собрать всех экспертов за «Круглым столом», чтобы обсудить все спорные моменты.

Здесь может возникнуть ситуация, когда за один тур обсуждения не удастся прийти к единому мнению по каким-то вопросам. Поэтому следует провести повторные туры по методу «Дельфи» [6], пока не будет найдено взаимоприемлемое решение. Как правило, определенные трудности по согласованию мнений возникают на нижних уровнях дерева (критерии и показатели), в связи, с чем здесь обычно имеет место избыточность элементов. Поэтому требуется провести исследование списков на избыточность с помощью специальных математических методов визуализации данных и снижения размерности. Эти исследования необходимы еще и потому, что в основе реализации второго этапа идентификации критериально-целевой структуры – определении весов приоритетности целей и критериев – лежит метод парных сравнений. Трудозатраты экспертов при исследовании данного метода существенно зависят от количества оцениваемых объектов.

Если собрать экспертов за «круглым столом» не представляется возможным, то для формирования окончательных списков следует использовать мажоритарное правило отбора, например, правило простого большинства: элемент остается или исключается из списка, если за это «проголосовало» более половины экспертов, входящих в комиссию. В сложных случаях, как и выше, следует оставить спорные элементы в списке и применить методы снижения размерности.

Применение данной группы математических методов является желательным (хотя и необязательным) этапом анализа критериально-целевой структуры. Несмотря на то, что применение методов снижения размерности выходит за рамки темы настоящего исследования и является самостоятельной сложной проблемой, для полноты картины рассмотрим краткое представление об этой группе методов.

Цель процедур снижения размерности – сокращение избыточности числа критериев за счет выделения наиболее информативных критериев на базе статистического анализа их структурных связей.

Методы снижения размерности пространства критериев основаны на переходе от большого числа k исходных критериев $f_1, \dots, f_k$ к существенно меньшему числу информативных переменных $y_1, \dots, y_m$, $m < k$. В качестве исходной информации принимается матрица измерений критериев в экспериментальных точках.

В общем случае задача снижения размерности ставится как экстремальная: необходимо определить такой набор переменных $\vec{y} = (y_1, \dots, y_m)$ из класса допустимых преобразований критериев $\vec{f} = (f_1, \dots, f_m)$, при котором достигает максимума выбранная мера информативности $I_m(\vec{y})$.

Конкретный набор меры I_m и класса допустимых преобразований приводит к различным методам снижения размерности методом факторного анализа, главных компонент, многомерного шкалирования и дискриминантного анализа [21]. Однако снижение размерности критериального пространства в этих методах происходит не за счет исключения части критериев из исходного набора $\vec{f}$, а за счет использования меньшего числа новых латентных переменных $\vec{y}$, которые являются функциями от $\vec{f}$ и плохо поддаются интерпретации.

Поэтому целесообразно использовать метод, позволяющий снизить размерность, оставаясь в той же системе критериев, т.е. перейти к системе $f_{i1}, \dots, f_{im}$, наилучшим образом прогнозирующих оставшиеся $k - m$ критериев. Такая система критериев получается в результате решения оптимизационной задачи $I_m \rightarrow \max$.

Здесь мерой информации служит функционал:

$$I_m = \sum_{q=m+1}^k R_{iq},$$

где R_{iq} – множественный коэффициент корреляции, характеризующий тесноту связи i_q -го критерия с совокупностью $f_{i1}, \dots, f_{im}$.

Таким образом, процедура выбора информативных критериев позволяет снизить размерность дерева целей и применять на втором этапе идентификации метод парных сравнений.

Обобщенный алгоритм экспертного опроса

Рассматриваемая задача определения приоритетности целей и критериев в виде весов в теории экспертных оценок относится к классу задач количественного оценивания вариантов. Анализ литературы по этому вопросу позволил выделить четыре основных этапа реализации экспертизы подобного рода. Каждый из этапов представляет собой самостоятельную задачу. Решение этих задач возможно

только во взаимосвязи и неразрывно связано с тем методом и теми предпосылками, которые исследователь выбирает для решения основной задачи экспертизы.

Приведем обобщенный алгоритм проведения экспертизы:

1. Получение исходной информации от экспертов.
2. Анализ приемлемости (непротиворечивости) суждений эксперта.
3. Агрегирование экспертных оценок и получение результирующих весов приоритетности.

Получение экспертной информации

По способам получения экспертной информации можно выделить две основные группы методов – это методы непосредственного оценивания приоритетности объектов в некоторой априорно заданной числовой или балльной шкале и методы, использующие информацию о предпочтительности в виде парных сравнений.

Заметим, что между «выходными» данными экспертизы, т.е. видом окончательных оценок, и «входными» - теми, которые непосредственно запрашиваются у эксперта, нет взаимнооднозначного соответствия. Так, основываясь на частотах ответов типа «лучше-хуже», в которых отражается лишь порядок предпочтения, можно строить количественные оценки предпочтительности. Наоборот, нередко опросы, когда проставляемые экспертами баллы, т.е. некоторые числа непосредственно годятся лишь для ранжирования объектов. В экспертизе используются также вербальные шкалы, парные и множественные сравнения, но все они, в конечном счете, направлены на оценивание одного из перечисленных типов.

Если шкала заранее не задана, единственные вопросы, которые не связаны с сильными предположениями о способности экспертов к самостоятельной квантификации своих субъективных представлений, - это вопросы о порядке предпочтения. Порядок может устанавливаться как на различных подмножествах исходного множества объектов (парные и множественные сравнения, ранжировки), так и на комбинации объектов (сравнение «разностей», «сумм», лотерей). В зависимости от выбранного варианта получаем конкретный тип данных, которые запрашиваются у эксперта. Оценки предпочтительности, рассчитанные по этим данным, могут быть порядковыми (ранги), количественными или чем-то промежуточными (например, приближенно-количественными). С формальной точки зрения они образуют некую апостериорную шкалу, в которой измерены объекты.

Для получения количественных оценок на основе «порядковой» информации, т.е. по ответам типа «лучше-хуже» есть два принципиально разных способа: запрашивать как можно более полные сведения о предпочтениях каждого эксперта и их квалифицировать или анализировать относительно простые данные, но по группе в целом, считая, что за частотами «побед» и «поражений» объектов стоят неизвестные значения их предпочтительности.

Для определения интенсивности предпочтения чаще всего используют баллы. При отсутствии априорной шкалы это может иметь два объяснения. Первое: каждый эксперт имеет свою субъектив-

ную шкалу, в которой и оцениваются объекты. Такое предположение, правда, выглядит несколько фантастично. Более правдоподобно второе – баллы рассматриваются как простой способ получения углубленной информации о предпочтениях. В этом случае мы опять сталкиваемся с проблемой проверки «достаточности» соответствующей информации для приближенно-количественного представления предпочтений. Но так как в явном виде она нам недоступна – она «спрятана» в назначенные баллы, то решить эту проблему невозможно. Кроме того, по сравнению с методами предыдущей группы здесь появляется такой сильный источник искажений как «самоквантификация»: эксперт должен сам подобрать числа, отражающие его представления о качестве объектов, и вряд ли он бывает точным. Можно указать и на деформации, связанные с недостаточным числом градаций в предложенной шкале. Таким образом, хотя баллы всегда несут более богатую информацию о предпочтениях, чем «порядковые» данные, они же могут содержать очень существенные искажения. Баллы не оставляют возможности проверять степень «количественности» предпочтений, и в тех задачах оценивания, о которых идет речь, лишь с определенной натяжкой могут обрабатываться как числа.

Проблема баллов отягощается также необходимостью совмещения оценок разных экспертов. Если два эксперта имеют полностью совпадающие предпочтения, это еще не означает, что они используют для их выражения одинаковыми оценками шкалы и каждому объекту будут приписаны одни и те же баллы. Существующие здесь технические приемы (фиксация оценки лучшего объекта и т.д.) не решают проблемы в целом.

Представляется, что более надежные, контролируемые результаты дает подход, когда для упорядочения (выбора) и для количественной оценки используются данные только о порядке предпочтения. Как уже отмечалось, для практических экспертиз больше подходят относительно простые формы: индивидуальный выбор «лучших» объектов, ранжировки, парные и множественные сравнения. «Платой» за объединение информации является то, что в отдельных случаях количественные оценки, рассчитанные с помощью «частотных» методов, лишь приблизительно отражают субъективную интенсивность предпочтения.

Индивидуальный выбор одного или нескольких лучших объектов, наоборот, прост для эксперта, но дает бедную информацию о его предпочтениях. Это снижает надежность результата.

Наиболее подходящими для рассматриваемой задачи экспертного оценивания являются парные сравнения. Преимущество парных сравнений – в большей свободе для выражения предпочтений, простоте самой процедуры оценивания с точки зрения эксперта, особенно в ситуации плохой «различимости» объектов. Очевидно, что эксперту легче сравнивать пары объектов, нежели «разместить» на числовой шкале сразу все множество объектов. Отметим еще два важных достоинства этого способа. Во-первых, определенная «избыточность» получаемой информации позволяет провести анализ суждений эксперта и тем самым эффективно решить задачу экспертизы. Недостатком этой формы оценивания является сравнительная трудоемкость процедуры, требующей для упорядочения n объектов проведения $n(n - 1)/2$ сравнений. Реально рекомендуется применять парные сравнения, когда число объектов не превышает 10–15.

В связи с этим, методы, основанные на парных сравнениях, получили большое распространение и довольно часто используются на практике для нахождения весов приоритетности. При этом от эксперта фактически требуется количественная информация о предпочтительности – качественные оценки типа «лучше-хуже», как отмечалось, не позволяет с точки зрения теории измерений конкретно произвести упорядочение объектов на числовой шкале.

Проверка состоятельности индивидуальных предпочтений

Экспертное оценивание проводится тогда, когда объективное измерение невозможно или нецелесообразно. Тем не менее, в традиционном представлении проверка состоятельности индивидуального ответа означает его сравнение с некоторым эталоном – «истинной» оценкой или ее приближением. Предлагается, например, метод сравнения с групповой средней, которая в этом случае принимается за эталон. Понятно, для того, чтобы усредненное мнение соответствовало «истинному», основная часть экспертов должна давать правильные ответы. Проверить, так ли это на самом деле в конкретной экспертизе, практически невозможно, а провести предварительное тестирование удастся далеко не всегда. Вообще, надежда на использование «истинной» или близкой к ней оценки в анализе экспертных данных, который должен выявить эту оценку, вряд ли оправданы. Скорее упомянутая и аналогичные ей идеи применимы для незначительной корректировки уже известного результата. Более того, в многочисленных экспертизах, предназначенных для нахождения разумного компромисса между заведомо несовпадающими точками зрения, само понятие истинности или эталона теряет смысл. Как же тогда судить о приемлемости ответа. Рассмотрим характерный пример.

Пусть в каком-то министерстве решили выяснить влияние социальных факторов, в частности, уровня образования на качество работы по некой группе рабочих профессий в целом. Так как результат нужно получить в виде упорядочения факторов по их сравнительной важности, вместо трудоемкого выборочного обследования с использованием методов математической статистики привлечены эксперты-люди, хорошо знающие реальные условия отраслевого производства. Перед экспертами – довольно сложная задача оценки связей между комплексными, неоднозначными признаками. Предположим, что она отчасти конкретизирована, и каждый признак (фактор) замещен простым измеримым показателем: уровень образования – числом лет обучения, качество работы – процентом сдачи продукции с первого предъявления и т.д. Однако и в такой постановке задача остается в значительной степени неопределенной, так как не ясно, каким образом давать единую оценку по профессиям, для которых значимость факторов, вообще говоря, различна.

При неопределенности вопроса не определено и понятие истинности ответа. Каждый эксперт вправе ориентироваться на ту или иную профессию (группу профессий), считая ее основной для отрасли, например, из-за наибольшего удельного веса в общей численности занятых или из-за ее прогрессивности, значимости и т.п. В этой ситуации цель каждого ответа и всего опроса – установление некой общей закономерности (если таковая имеется) как ненулевой равнодействующей множества конкретных истин (контекстов). Если принять за факт, что задача, в принципе, разрешима, то

информативными естественно считать те ответы, которые дают нужную информацию, в данном случае – упорядочение (ранжирование) социальных факторов по их влиянию на качество работы по группе в целом. Точнее, представления эксперта считаем информативными, если они могут быть представлены оценками шкалы заданного типа (в частности, порядковой). Таким образом, проблема «истинности» индивидуального основания замещается проблемой его возможности.

Данный подход актуален не только для нечетко поставленных экспертных задач, которые составляют большинство в социологии и экономике, но и для задач согласования интересов, где он позволяет выделить незрелые, «хаотические» предпочтения. Впервые он был предложен С.Кендаллом и Б.Смитом для проверки возможности строгого упорядочения объектов, затем распространен на задачи группировки и нестрогого ранжирования [7]. В методе, предложенном С.Кендаллом и Б.Смитом, для определения внутренней согласованности парных сравнений в качестве меры непротиворечивости суждений эксперта используется число циклических (нетранзитивных) троек элементов в матрице.

В общем случае схема построения критерия индивидуальной согласованности эксперта имеет следующий вид:

1. выбор меры согласованности экспертных оценок;
2. задание «эталонной» модели образования противоречивых суждений;
3. расчет распределений меры для принятой модели при некоторых предположениях о распределении модельных параметров.

В подавляющем большинстве работ в качестве «эталонной» модели рассматриваются случайные суждения эксперта. Тогда анализ согласованности состоит в проверке гипотезы о случайном формировании экспертных оценок. В качестве соответствующей статистики, для которой строятся распределения, выступает выбранная мера согласованности.

Проверка групповой согласованности

Главная цель экспертиз – получение некой «объективной» информации о вариантах, причем такая информация не рассматривается как директивная. Исходное соглашение состоит в том, чтобы в качестве «объективной» использовать оценку, которая получена от группы экспертов как равнодействующая их индивидуальных мнений. При этом предполагается, что разумный компромисс между различными точками зрения достижим. В ситуации согласования различных интересов экспертные методы дают возможность построить «проект» соглашения, который, в известном смысле, наиболее близок к исходным позициям всех участников.

Необходимость получения результата, обобщающего ответы нескольких теоретически равноправных экспертов, делает задачу согласования мнений центральной как для каждого конкретного опроса, так и для направления в целом.

Существуют две возможности согласования: путем последовательного сближения исходных точек зрения или применения некоего формального правила, дающего «справедливый» компромисс,

когда мнения различны и фиксированы. Так как в реальности полное единство взглядов – скорее исключение, чем закономерность, использование формальных методов часто является вынужденной мерой, на которую идут, когда обсуждение вопроса заходит в тупик. Простейший пример – принятие решения большинством голосов. В то же время во многих экспертизах обмен мнениями между участниками не предусматривается (дорого, бесполезно и т.д.), и формальные методы становятся основным инструментом поиска решений. При этом нагрузка на экспертов снижается, но возрастает трудоемкость сбора и обработки данных.

Промежуточное положение между неформальным обсуждением и формальным согласованием (агрегированием) занимают процедуры, которые можно назвать формализованными. Они рассчитаны на постепенное сближение мнений, но при этом регламентируются контакты между экспертами и тип экспертной информации. Краткая характеристика этих процедур дана ниже. Здесь отметим основные проблемы, связанные с агрегированием предпочтений.

Основной недостаток агрегирования состоит в использовании относительно бедной информации о вариантах. Более сложные соображения и мотивы, которыми руководствуется эксперт, остаются «за кадром». В неформальном обсуждении они, наоборот, влияют на общее решение. Другая проблема, связанная с формальным согласованием – большое многообразие имеющихся методов и, как следствие, не единственность результата. Обычно рекомендуется подтверждать полученное решение проведением альтернативных расчетов, однако в том случае, когда решение изменяется вместе с методом, выбор остается за организаторами экспертизы или самими экспертами. Эти сложности можно рассматривать как плату за «волевое» согласование.

С содержательной точки зрения существенно, чтобы любая формально построенная равнодействующая выражала то общее, что содержится в полученных ответах. Понятно, что этого можно добиться только при условии достаточной согласованности (близости) мнений в группе. Если они противоречивы, то агрегирование превращается в бессмысленное усреднение. Следовательно, в каждом конкретном опросе должен решаться вопрос о правомерности использования формальной результирующей как «представителя» группы, то есть должна проверяться согласованность ответов.

Рассмотрим методы опроса, в которых эксперт не только указывает свои оценки, но и может корректировать их во взаимодействии с другими экспертами. Выделяются два основных типа таких процедур опроса [7]:

1. процедуры с личными контактами между экспертами,
2. процедуры без личных контактов, но с обратной связью.

К первому типу относится традиционная «дискуссия за круглым столом». В ходе дискуссии эксперт может неоднократно вызывать суждения, учитывая точки зрения других участников опроса. За последние 40 лет появилось лишь несколько существенно новых предложений, развивающих исходную идею дискуссии. К ним относятся, в частности, метод мозговой атаки и процедура номинальной группы. Метод мозговой атаки состоит в проведении совместного заседания экспертов по определенным правилам, направленным на создание атмосферы свободного высказывания мнений. Принципи-

альным является отказ экспертов от критической оценки высказанных во время заседания суждений. Процедура, как правило, применяется в ситуациях, требующих поиска нестандартных решений. В методе номинальной группы участники независимо дают свои оценки, затем следует их общее обсуждение, после чего эксперты снова проставляют оценки и т.д.

Хотя в практике экспертиз, особенно отечественных, широко применяются процедуры с прямыми контактами между экспертами, приоритет в настоящее время принадлежит процедурам второго типа. Они глубже проработаны в научном плане и активно применяются за рубежом. Эти методы ведут свою историю от широко известной процедуры «Дельфи» [7]. Она состоит в следующем. Экспертам предъявляют оцениваемый объект (или объекты). В первом туре каждый эксперт дает объекту числовую оценку. После этого исследователь подсчитывает и сообщает всем экспертам среднюю оценку (или медиану) и показатель разброса оценок (например, интервал между крайними квартилями). Экспертов, давшие крайние оценки просят обосновать свое мнение. Эти обоснования сообщаются (при сохранении анонимности) всем экспертам, после чего аналогично проводится второй тур опроса. Итерации заканчиваются тогда, когда, по мнению исследователя, будет достигнута достаточная близость оценки экспертов. Этот исходный вариант процедуры («стандартный Дельфи») повлек за собой множество модификаций. Так, обособленность особенность метода «Дельфи II» состоит в том, что эксперт в каждом туре дает высшую, среднюю и низшую оценку: соответственно, по обратной связи передаются медианы этих трех оценок. Метод «упрощенный Дельфи» представляет компромисс между Дельфи и дискуссией. Эксперты сидят «за круглым столом», но устно выступают лишь те, которые представили крайние оценки. В методе Форда предлагается более изощренная схема, направленная на уменьшение конформизма. В первом туре эксперт сообщает верхнюю и нижнюю границы, между которыми, по его мнению, лежит точная оценка. Подсчитываются средние высшая и низшая оценки. Середина интервалов между ними сообщается экспертами, и их просят указать, выше или ниже ее лежит точная оценка. «Большинством голосов» выбирается одна из областей, например, верхняя. В следующем туре середина этой области предъявляется экспертам с аналогичным вопросом и т.д. Опрос заканчивается при достижении достаточно малого интервала.

Огромный опыт использования процедуры Дельфи, широко применяемой с конца 60-х годов, убеждает, что сообщение усредненной оценки часто оказывает неоправданное сильное давление на мнения экспертов. Ведь, по существу, эксперт оказывается в ситуации известных психологических опытов по выявлению конформизма. Данное обстоятельство привело к критике «Дельфи» и к разработке новых процедур опроса. В этих процедурах участник опроса на каждом шагу узнает не среднюю оценку, а мнение одного из своих коллег. Сравниваются два равноправных суждения, а не частное мнение с мнением коллектива. Перейдем к описанию таких методов.

Л.А.Панковой и М.В. Шнейдерманом предложена так называемая последовательная процедура опроса. На первом туре случайно выбирается один из экспертов. Он дает оценку объекту и отдельно от нее текст, содержащий известные эксперту сведения и мнения о различных особенностях оцениваемого объекта, которые эксперт принимает во внимание при вынесении оценки. Во втором туре та

же информация поступает от второго случайно выбранного эксперта. Эти два эксперта знакомятся с текстами (но не оценками) друг друга, после чего сообщают исследователю свои скорректированные оценки. В третьем туре первые два эксперта получают текст третьего эксперта, а он – их тексты, и оценки снова корректируются и т.д. Этот метод имеет характерное правило остановки, принципиально отличное от используемого в «Дельфи». Критерием окончания опроса является не сближение, а стабилизация мнений экспертов.

В процедуре «качественной обратной связи» С.Пресса в первом туре все эксперты одновременно дают оценки и подробные текстовые обоснования своих оценок. Затем составляется и передается каждому эксперту свободная информация, суммирующая обоснования всех экспертов. Во втором туре снова каждый эксперт дает оценку и ее обоснование. Критерием остановки является стабилизация индивидуальных оценок.

С.Б.Котляром предложена процедура, согласно которой в каждом туре эксперты случайно разбиваются на пары, и обмениваются оценками. Правило остановки не указано, но, по-видимому, опрос заканчивается при достаточном сближении оценок.

Приведенный неполный перечень процедур проведения «динамических» экспертных опросов позволяет выделить основные признаки, по которым эти процедуры различаются. К таким признакам относятся способ организации обратной связи, вид передаваемой по ней информации и правило остановки процедуры.

Приведем те выводы, которые нашли подтверждение в большинстве исследований:

1. В целом в описанных процедурах итерации повышаются точность ответов, причем в этом отношении между процедурами нет существенного различия. Однако большую точность дает обратная связь, предусматривающая одновременную передачу и экспертных оценок и их обоснований.
2. Эффект сближения оценок сильнее всего проявляется в «Дельфи», однако достижение согласия в этом случае не гарантирует точности итоговых оценок.
3. Стабилизация оценок является более надежным критерием остановки, чем их сближение.
4. Субъективная удовлетворенность эксперта тем выше, чем интенсивнее его взаимодействие с другими экспертами в процессе опроса. В то же время, точность итоговой оценки не всегда положительно коррелируется с удовлетворенностью экспертов.
5. По трудоемкости и затратам времени лучшей является процедура индивидуальной обратной связи.

Известные в литературе формальные методы определения групповой согласованности наиболее разработаны для проверки экспертных суждений, заданных в виде неметризованных парных сравнений и ранжирований. В первую очередь к ним относится коэффициент конкордации С.Кендалла, свойства которого достаточно полно исследованы в [18]. Имеется также целый ряд аналогичных мер согласования, основанных, главным образом, на вычислении результирующей матрицы предпочтений $P = (P_{ij})$, где $P_{ij} = m_{ij}/m$, m_{ij} - число экспертов, считающих i -й объект предпочтительнее j -го (или приписавших i -му объекту ранг j), m – общее число экспертов.

Однако для рассмотрения задачи определения весов приоритетности такие меры групповой согласованности оказываются малопригодными, так как при их использовании теряется содержащая в парных сравнениях информация об интенсивности предпочтений, что приводит к слишком грубым оценкам согласованности.

Схема анализа групповой согласованности аналогична схеме проверки непротиворечивости ответов эксперта, которая описана выше. В качестве «эталонной» модели несогласованности экспертов используется модель случайного поля высказанных точек зрения экспертов с равномерным распределением. Проверка согласованности экспертной группы состоит в принятии или отклонении статистической гипотезы о случайном наборе полученных суждений. С этой целью необходимо произвести статистическое моделирование выбранной меры и согласованности и построить соответствующее распределение этой меры.

Агрегирование экспертных предпочтений

Рассматриваем последний этап обобщенного алгоритма проведения экспертизы – задача упорядочения или количественной оценки объектов по данным в виде набора индивидуальных ранжировок или матриц парных сравнений. Мы обсудим также некоторые элементы теории группового выбора, предмет которой – анализ и синтез различных формальных правил агрегирования.

Предположим, что имеется конечное множество объектов $X = \{x_1, \dots, x_n\}$; каждый их экспертов попарно сравнивает все объекты и упорядочивает их по предпочтительности, и на основании его ответов заполняется матрица парных сравнений $A^k = (a_{ij}^k)$, k – номер эксперта, элементы которой определяются следующим образом: $a_{ij}^k = 0$ при $i = j$, в случае $i \neq j$.

$$a_{ij}^k = \begin{cases} 1, & \text{если эксперт предпочитает объект } x_i \text{ объекту } x_j; \\ 0, & \text{если эксперт предпочитает объект } x_j \text{ объекту } x_i; \\ 1/2, & \text{если эксперт считает } x_i \text{ и } x_j \text{ равноценными.} \end{cases}$$

Парные сравнения являются более общим типом экспертной информации, чем упорядочения. Действительно, если задано упорядочение, то определен результат парного сравнения любых двух

объектов. С другой стороны для множества парных сравнений, определяемого упорядочением, всегда выполняется следующее условие: если объект x_i предпочитается объекту x_j , а x_j предпочитается x_l , то x_i предпочитается x_l (это свойство называется транзитивностью, а если, напротив, x_l предпочитается x_i , то говоря, что x_i , x_j и x_l образуют цикл предпочтения). Для парных сравнений общего вида выполнение этого условия, вообще говоря не предполагается. Матрицы парных сравнений будем рассматривать как исходные данные для построения итогового упорядочения объектов.

Введем матрицу $A = (a_{ij})_n$:

$$A = \sum_{k=1}^n A^k.$$

Элемент матрицы A:

$$a_{ij} = \sum_{k=1}^n a_{ij}^k$$

равен количеству экспертов, предпочитающих объект x_i объекту x_j , плюс половина количества экспертов, считающих x_i и x_j равноценными. Пользуясь спортивной аналогией, часто полезной при анализе экспертных оценок, можно сказать, что a_{ij} есть общее количество «очков», набранных объектом x_i против объекта x_j .

Если в упорядоченности по предпочтительности нет равноценных объектов (такое упорядочивание называется строгим), будем иногда представлять его ранжировкой т.е. функцией $r = r(x)$, которая самому предпочтительному объекту ставит в соответствие число (ранг) 1, следующему – ранг 2 и т.д. Если $r(x_i) < r(x_j)$, будем говорить, что в ранжировке (и в упорядочении, которое она представляет) объект x_i стоит выше, чем x_j (или x_i превосходит x_j в r).

Рассмотрим несколько распространенных методов упорядочения объектов по результатам парных сравнений. Простейший метод упорядочения состоит в следующем. Для каждого объекта x_i вычисляется сумма результатов всех его сравнений с другими объектами:

$$S = \sum_{j=1}^n a_{ij} = \sum_{j=1}^n \sum_{k=1}^m a_{ij}^k, \quad i = \overline{1, n}, \quad \text{и объекты упорядочиваются по убыванию величины } S_i.$$

Этот метод называется методом строчных сумм.

Если мнения экспертов заменить кругами спортивного турнира, то метод сумм совпадает со стандартной процедурой распределения мест в игровых видах спорта.

Алгоритм строчных сумм имеет свою аксиоматику. Для некоторых моделей парных сравнений он может быть получен методом максимального правдоподобия или методом наименьших квадратов и обладает многими другими свойствами.

В методе строчных сумм показатель S_i , определяющий место объекта x_i в итоговом упорядочении, зависит от результатов всех сравнений x_i . Ж.А.Кондросье в работе 1785г. предложил определять взаимное положение любых двух объектов x_i и x_j в коллективном упорядочении, в зависимости от результатов их сравнений друг с другом: x_i ставится выше, чем x_j , если большинство индивидуумов предпочитают x_i объекту x_j . В наших обозначениях:

$$a_{ij} > a_{ji} \rightarrow r(x_i) < r(x_j).$$

Однако, такой метод противоречив даже в случае, когда мнения экспертов являются упорядочениями объектов.

Из-за наличия «циклов коллективного предпочтения» правила большинства, вообще говоря, недостаточно для построения итогового коллективного предпочтения.

Интересен вопрос: существуют ли вообще «разумные» методы построения итогового упорядочения. Для одного из возможных пониманий «разумности» ответ на этот вопрос дает анализ, проведенный К.Дж.Эрроу. В работах [4, 9] исследуется совместимость четырех требований к методу агрегирования, которые можно записать в следующем виде:

1. Универсальность. Пусть $h > 2$ - количество объектов, $m > 1$ - число индивидуумов. Итоговое упорядочение определено для любой совокупности индивидуальных упорядочений множества объектов.
2. Монотонность. Если объект x превосходит y в итоговом упорядочении, построенном для набора индивидуальных упорядочений, R отличается от R^1 лишь тем, что один из участников, сравнивая x с некоторым z , изменил свое мнение в пользу x , то x превосходит y в итоговом упорядочении, построенном для R^1 .

3. Бинарная независимость. В итоговом упорядочении для любых X, Y превосходство X над Y или Y над X зависит лишь от результатов индивидуальных сравнений объектов X и Y друг с другом.
4. Ненавязанность решения. Для любых объектов X, Y может быть построен набор индивидуальных упорядочений, такой, что X превосходит Y в соответствующем итоговом упорядочении.

Теорема Эрроу состоит в том, что четырем перечисленным условиям удовлетворяют только «диктаторские» правила, т.е. такие методы, в которых итоговое упорядочение всегда совпадает с упорядочением одного из участников, причем этот участник выбирается заранее.

Безусловно, есть ситуации, в которых «диктаторское» правило разумно, но несомненно и то, что во многих случаях решение следует принимать демократическим путем. При этом, как следует из теоремы Эрроу, нужно отказаться хотя бы от одного из четырех перечисленных ограничений.

Исследуем теперь свойства бинарно-независимых «демократических» правил агрегирования, допуская при этом возможность нарушения условий универсальности (условия монотонности и ненавязанности решения, как правило, не вызывают возражений, и мы будем по-прежнему требовать их выполнения). Иными словами, предположим, что процедура агрегирования строит итоговое предпочтение на парах объектов, которое, вообще говоря, не является упорядоченным. «Демократичность» правила будем принимать здесь как равноправие всех экспертов (избирателей). Допустим также, что правило агрегирования нейтрально к вариантам, т.е. итоговое предпочтение не меняется при любом переименовании объектов. Предположим, наконец, для простоты, что индивидуальные упорядочения являются строгими.

Можно показать, что всем этим ограничениям удовлетворяют только правила « k -большинство».

Согласно правилу « k -большинству» для любых $x, y \in X$ коллектив предпочитает объект x объекту y тогда и только тогда, когда x стоит выше y не менее, чем в k индивидуальных упорядочениях, где $k \leq m$ - заранее выбранное число. Правило большинства, очевидно, является « k -большинством» при $k > \left\lfloor \frac{m}{2} \right\rfloor + 1$. Невыполнение условий универсальности для правил « k -большинство» означает, что полученное коллективное предпочтение может быть упорядоченным, и, в частности, при $k < m$ в нем возможны нарушения транзитивности.

В ряде случаев применение правил типа « k -большинство» может приводить к совсем неожиданным результатам. Для описания этих результатов обратимся к так называемой модели динамического голосования. В этой модели бесконечное множество объектов ("программ") отождествляется с метрическим пространством, а голосование представляет собой последовательность туров, в каждом из которых избиратели сравнивают программу, выбранную коллективом в предыдущем туре, с некоторой вновь предложенной, и решение принимается " k -большинством". Изучаются свойства последова-

тельности выбранных программ в предположении, что, сравнивая две программы, избиратель голосует за ту, которая ближе по метрике в его "идеале" ("идеал" - одна из программ того же пространства).

Согласно теореме Маккельви для любых двух программ X и Y существует последовательность выбираемых программ, "ведущая" от X к Y . В частности, программа Y может быть значительно дальше от всех "идеалов", т.е. значительно хуже для всех избирателей, чем программа X . При достаточной размерности пространства аналогичной результат может быть доказан для любого правила "к-большинство".

Таким образом, из приведенных рассуждений следует вывод о желательности учета интенсивности предпочтений. С другой стороны, типична ситуация, когда не существует объективного метода, позволяющего сравнивать степени предпочтения для разных экспертов. Единственный выход, на наш взгляд, состоит в косвенном учете степеней предпочтения посредством учета результатов сравнения X и Y с другими объектами (как это делается, например, в методе строчных сумм, удовлетворяющих всем условиям теоремы Эрроу, кроме бинарной независимости).

Несмотря на описанные недостатки правила большинства, имеется немало ситуаций, в которых применение его содержательно обосновано. В этих случаях часто возникает задача выбора наилучших объектов при циклических итоговых предпочтениях. Одним из широко известных алгоритмов ее решения является метод, предложенный Д.Блэком, согласно которому «лучшими» признаются объекты, не уступающие ни одному объекту по правилу большинства, а если таковых нет, лучшими объявляются объекты с максимальной строчной суммой.

Правило Блэка соединяет два совершенно различных принципа выбора. Это приводит к нарушению важного условия монотонности, если использовать данное правило для построения коллективного упорядочения, т.е. на каждом шаге выбирать из множества X «лучшие» объекты среди еще невыбранных.

Более корректным методом упорядочения объектов при нарушении транзитивности в правиле большинства, на наш взгляд, является метод Копленда, согласно которому объекты упорядочиваются по убыванию показателя, равного разности числа объектов «худших» к «лучшим» по большинству, чем данный объект.

Далее рассмотрим несколько методов, которые, как и метод строчных сумм, не удовлетворяют условию бинарной независимости.

Модель Цермело-Брэдли-Терри (ЦБТ) состоит в следующем.

Каждому объекту X_i соответствует его «сила», и вероятность того, что объект X_i предпочитается X_j ($i \neq j$) в индивидуальном парном сравнении, равна:

$$\pi_{ij} = \pi_i / (\pi_i + \pi_j), \quad i, j = \overline{1, n}, \quad i \neq j$$

(равноценность объектов в этой модели не допускается).

Данная модель характеризуется тем, что для любых объектов x_i и x_j отношение вероятностей предпочтения при их сравнении равно отношению «сил» π_i и π_j :

$$\pi_{ij} / \pi_{ji} = \pi_i / \pi_j$$

Для оценки параметров $\pi_1, \dots, \pi_n$, чаще всего используют метод максимального правдоподобия, который приводит к системе уравнений:

$$\sum_{i \neq j}^n m \frac{\hat{\pi}_i}{\hat{\pi}_i + \hat{\pi}_j} = S_i, \quad i, j = \overline{1, n}, \quad \sum_{i=1}^n \pi_i = 1,$$

Система решается численными методами, а решения $\hat{\pi}_1, \dots, \hat{\pi}_n$ являются оценками параметров $\pi_1, \dots, \pi_n$.

Тогда, решения системы [12], [13] $\hat{\pi}_1, \dots, \hat{\pi}_n$ упорядочены так же, как строчные суммы $S_1, \dots, S_n$.

Оценки максимального правдоподобия могут быть использованы для проверки соответствия модели ЦБТ полученным данным. Нулевая гипотеза H_0 состоит в том, что вероятности предпочтений задаются моделью ЦБТ:

$$H_0: \pi_{ij} = \frac{\pi_i}{\pi_i + \pi_j}, \quad i, j = \overline{1, n}, \quad i \neq j.$$

Для проверки нулевой гипотезы используется статистика:

$$\lambda^2 = \sum_{i, j \neq i} (a_{ij} - \frac{m \hat{\pi}_i}{\hat{\pi}_i + \hat{\pi}_j})^2 \frac{\hat{\pi}_i + \hat{\pi}_j}{m \hat{\pi}_i}$$

Гипотеза H_0 отвергается на уровне значимости α , если λ^2 больше α -процентной точки распределения λ^2 с $(n-1)(n-2)/2$ степенями свободы.

Оценки параметров $\pi_1, \dots, \pi_n$, полученные методом максимального правдоподобия, не всегда могут быть естественно интерпретированы на содержательном уровне.

В [18] описаны методы оценки параметров $\pi_1, \dots, \pi_n$ и проверки адекватности для модели ЦБТ. В этой же работе содержится обзор других статических моделей парных сравнений.

Метод упорядочения объектов, предложенный Т.Уэи и М.Кендаллом основан на интеративной процедуре вычисления оценок объектов. На первом шаге в качестве оценки S_i^1 каждого объекта x_i выбирается его строчная сумма:

$$S_i^1 = S_i, \quad i = \overline{1, n}.$$

На шаге $l+1$ объект x_i оценивается взвешенной суммой полученных на шаге l оценок S_i^l его «соперников». Весаами являются набранные против них «очки» a_{ij} :

$$S_i^{l+1} = \sum_{j=1}^n a_{ij} S_i^l, \quad i = \overline{1, n}, \quad l = 1, 2, \dots$$

Далее к оценкам применяется операция нормирования, сохраняющая их отношения, т.е. вектор S^l оценок, полученных на шаге l , заменяется вектором:

$$\tilde{S}^l = S^l / \sum_{i=1}^n S_i^l$$

Можно показать, что последовательность $\{S^l\}$ сходится к некоторому вектору $\tilde{S}$, компоненты которого и предлагается использовать в качестве окончательных оценок объектов. Из теоремы Перрона-Фробениуса следует, что вектор $\tilde{S}$ является правым собственным вектором матрицы A , соответствующем ее наибольшему собственному значению, что позволяет находить вектор $\tilde{S}$ методами линейной алгебры.

Недостатком метода Уэи-Кендалла состоит в том, что при оценивании объекта x_i главным образом учитываются результаты тех объектов, x_i которым x_j часто предпочитается (a_{ij} велико), и слабо учитываются результаты объектов, превосходящих x_i в большинстве парных сравнений. Такая несбалансированность проявляется в отсутствии у данного метода свойства «транспонируемости», согласно которому изменение всех индивидуальных парных предпочтений на противоположные должно приводить к инверсии итогового упорядочения.

При обсуждении метода строчных сумм рассматривается мера близости между ранжировкой r и матрицей A , равная взвешенной сумме «очков», набранных в A всеми объектами против объектов, стоящих выше их в r . В методе Кемени-Слейтера рассматривается мера близости, в которой та же сумма берется без весов, т.е. не учитывается, насколько выше в r стоит объект x_j , против которого объект x_i получил результат a_{ij} :

$$D(r, A) = \sum_{r(x_i) < r(x_j)} a_{ij}$$

Рассмотрим задачу наложения ранжировок r_0 , таких, что

$$D(r_0, A) = \min_r D(r, A)$$

Эта задача приводит к так называемой медиане Кемени [6].

Итоговая ранжировка r_0 , оптимальная по критерию Кемени-Слейтера, вообще говоря, не единственна. Для нахождения медианы Кемени используются методы дискретной оптимизации (ветвей и граней, динамического программирования и др.), сложность которых весьма высока и экспоненциально растет при увеличении размерности задачи, или эвристические алгоритмы, не гарантирующие нахождения всех оптимальных решений.

Для некоторых исходных данных ранжировки, оптимальные по критерию Кемени-Слейтера, весьма сильно отличаются друг от друга. В силу перечисленных свойств применение метода Кемени-Слейтера на практике требует осторожности.

В данном обсуждении мы коснулись лишь некоторых наиболее распространенных методов получения коллективных упорядочений (или численных оценок объектов) из большого числа алгоритмов, описанных в литературе.

Наиболее предпочтительным из изложенных методов, на наш взгляд, является метод строчных сумм и методы, основанные на нахождении собственных векторов матрицы предпочтений. Вместе с тем, необходимо отметить, что качество того или иного метода существенно зависит от специфики проводимого опроса.

Оценивание проектов

Оценивание проектов происходит с помощью анонимного анкетирования по всей системе критериев. В качестве оценок берутся балльные градации шкал соответствующих критериев. После опроса полученные данные представляют собой набор (по числу экспертов) таблиц типа “проекты-критерии”, которые подлежат математической обработке.

Целью обработки является получение некоей согласованной обобщенной оценки проекта по всей экспертной группе. Технология экспертиз предполагает анализ согласованности экспертной группы, так как результирующая оценка имеет смысл только при наличии достаточной близости мнений членов экспертной комиссии. В противном случае необходим более детальный анализ точек зрения экспертов на проблему путём определения характера расслоения группы. Так, внутри группы может оказаться представительная высокосо согласованная подгруппа экспертов, на мнение которой и следует опираться. Если такой подгруппы нет, то поставленная перед экспертизой задача не может быть решена. Это означает, что при подготовке опроса были допущены серьёзные просчёты, которые следует проанализировать и внести коррективы.

Предлагаемая методология оценивания для получения более обоснованных результатов предполагает использование одновременно нескольких различных методов моделирования экспертных оценок. При анализе информации применяются две группы методов обработки данных: численные и турнирные.

Первая группа содержит три процедуры. В двух из них - методе Спирмена и методе определения среднего интервального - ранжирования проектов строятся на основе статистических характеристик (среднего и уклонения). В методе обобщённых строчных сумм данные преобразуются в кососимметрические матрицы парных сравнений. Ранжирования здесь строятся по результатам решения системы линейных уравнений.

Вторая группа методов содержит четыре турнирные процедуры, которые используют соотношения оценок между парами сравниваемых объектов (в нашем случае между проектами). Применение турнирных показателей полезно, когда важны не значения оценок экспертов, а их отношения. По балльным оценкам каждого эксперта строится матрица парных сравнений по аналогии с таблицей спортивного турнира. В качестве результатов используются турнирные показатели по сумме очков (метод Коупленда), по квадратному корню из суммы квадратов очков, по минимальному выигрышу и по среднему геометрическому очков. Одновременно каждый метод позволяет проверять согласованность экспертной группы.

Эти же методы могут быть применены при вычислении весов приоритетности критериев. Экспертам необходимо произвести ранжирование критериев внутри групп, а также оценить важность самих групп на верхнем уровне (тоже в виде ранжирования), и по этим оценкам будут вычислены веса критериев. Эти веса учитываются уже в моделях оценивания проектов и их последующем упорядочении.

В результате по каждому методу получаем свои ранжирования, которые сводим в единую таблицу и сравниваем между собой для построения окончательной ранжировки.

Литература:

1. АЛЕКСАНДРОВ Н.И., КОМКОВ Н.И. *Моделирование организации и управления решением научно - технических проблем*. М.: Наука, 1988.
2. ЛАРИЧЕВ О.И. *Объективные модели и субъективные решения*. М.: Наука, 1987.
3. МЕСАРОВИЧ М., ТАКАХАРА Я. *Общая теория систем: математические основы*. М.: Мир, 1978.
4. КИНИ Р.Д., РАЙФА Х. *Принятие решений при многих критериях: предпочтения и замещения*. М.: Радио и связь, 1981.
5. БЕШЕЛЕВ С.Д., ГУРВИЧ Ф.Г. *Математико-статистические методы экспертных оценок*. М.: Статистика, 1980.
6. ЛИТВАК Б.Г. *Экспертная информация: методы получения и анализа*. М.: Радио и связь, 1982.
7. ПАНКОВА Л.А., ПЕТРОВСКИЙ А.М., ШНЕЦДЕРМАН М.В. *Организация экспертизы и анализ экспертной информации*. М.: Наука, 1984.

8. ПОДИНОВСКИЙ В.В. *Что такое коэффициенты важности и как их корректно применять?* // Нечисловая статистика, экспертные оценки и смежные вопросы. М.: ВИНТИ, 1984. – с.163-165.
9. ГЛАТОВ В.А., ПАВЕЛЬЕВ В.В. *Векторная стратификация*. М.: Наука, 1987.
10. АЛЕСКЕРОВ Ф.Т., БАУМАН Е.В., ВОЛЬСКИЙ В.И. *Методы обработки интервальных экспертных оценок* // Автоматика и телемеханика. – 1984. – 3. – с.127-133.
11. ЖУКОВИН В.Е. *Многокритериальные модели принятия решений с неопределённостью*. Тбилиси: Мецниереба, 1983.
12. ВЛЭДУЦ С.Г., ЛЕВНЕР Е.В. *Об одном экспертном методе оценивая важности критериев в задачах многокритериальной оптимизации* // Нечисловая статистика, экспертные оценки и смежные вопросы. М.: ВИНТИ, 1984. – с.65-66.
13. РАСТРИГИН Л.А., ЭЙДУК Я.Ю. *Адаптивные методы многокритериальной оптимизации* // Автоматика и телемеханика. – 1985. – 1. – с.5-26.
14. ОРЛОВ А.И. *Устойчивость в социально – экономических моделях*. М.: Наука, 1979.
15. РАЕВ А.Г. *Об одном способе определения весовых коэффициентов частных критериев при построении аддитивного интегрального показателя* // Автоматика и телемеханика. – 1984. – 5. – с.162-165.
16. БАТИЩЕВ Д.И., АНУЧИН В.Ф. *Построение коэффициентов важности частных критериев на основе нечисловой информации* // Нечисловая статистика, экспертные оценки и смежные вопросы. М.: ВИНТИ, 1984. – с.188-189.
17. ТЮРИН Ю.Н. *О математических задачах в экспертных оценках* // Вопросы кибернетики. – 1979. – 58. – с.7-16.
18. ШМЕРЛИНГ Д.С., ДУБРОВСКИЙ С.А. и др. *Экспертные оценки: методы и применение*. – Там же, с.290-382.
19. *Обработка нечёткой информации в системах принятия решений* / Под ред. БОРИСОВА А.Н. М.: Радио и связь, 1989.
20. КОЗЕЛЬСКИЙ Ю. *Психологическая статистика. Классификация и снижение размерности* / Под ред. АЙВАЗЯНА С.А. М.: Финансы и статистика, 1989.