

# КВАЛИМЕТРИЯ МОДЕЛЕЙ. ПОКАЗАТЕЛИ КАЧЕСТВА И МЕТОДЫ ОЦЕНИВАНИЯ

*Микони С.В., Соколов Б.В., Юсупов Р.М. (Санкт-Петербург)*

## **Введение**

В работах [ ] рассматривались проблемы квалиметрии моделей. В работе [Казань] были предложены показатели, по которым можно сопоставлять качество моделей. В настоящей работе эти показатели конкретизируются применительно к различным задачам применения моделей. Рассматриваются методы оценивания качества моделей и приводится пример применения этих методов для решения конкретной задачи выбора модели.

## **1. Жизненный цикл модели**

Качество продукта потребления определяется как «способность удовлетворять установленные или предполагаемые потребности пользователя». Применительно к качеству модели необходимо определить потребности её пользователя. Учитывая, что качество характеризуется не единственным свойством модели, а их совокупностью, оценивание качества необходимо выполнять по нескольким критериям. Они должны отражать как этапы формирования свойств модели, так и этапы их реализации.

Качество модели формируется на начальных этапах её жизненного цикла, а реализуется при её конкретном использовании. Жизненный цикл модели любого объекта состоит из следующих этапов:

1. Разработка;
2. Модификация;
3. Настройка на задачу;
4. Решение задачи;
5. Анализ результатов и переход на 3 или 2 или 1.

### **1.1. Разработка модели**

Разработка модели делится на следующие этапы: проектирование формальной модели объекта, разработку на её основе компьютерной модели, проверку компьютерной модели.

На первом этапе изучаются свойства моделируемого объекта, и выбирается отражающий их математический аппарат. Основной целью этапа создания математической модели является обеспечение *полноты* моделируемых свойств объекта и *точности* их отражения в модели. По этим требованиям и определяется *соответствие* (адекватность) модели объекту.

Для представления свойства моделируемого объекта может применяться различный математический аппарат. Например, для распознавания символов алфавита могут применяться методы теории графов (скелетные графы эталонных символов), теории фильтрации (волновые алгоритмы) и нейронные сети. Каждый из этих методов имеет свои преимущества и недостатки, влияющие на показатели качества модели.

Компьютерная модель объекта создаётся двумя способами: параметризацией типовой модели, либо программированием оригинальной модели. На этом этапе решаются проблемы *размерности модели* и *сложности вычислений*. Размерность модели

оценивается объемом требуемой памяти, а сложность вычислений – временем решения задачи. Если параметризация типовой модели не позволяет добиться нужных результатов, выбирается другая типовая модель. Например, выбирается другая архитектура нейросетевой модели.

Проверка спроектированной модели осуществляется с применением методов верификации и тестирования. Устранение обнаруженных ошибок осуществляется в процессе отладки модели. На этом этапе качество модели характеризуется числом *пропущенных ошибок* проектирования.

### 1.2. **Модификация модели**

Модификация модели может потребоваться при изменении условий задачи, существенно не влияющих на применяемый математический аппарат. Базовые соотношения остаются неизменными. Меняется состав переменных. Например, изменился состав распознаваемых символов. Эти изменения влекут изменение контрольных, обучающих и тестовых примеров. Изменение состава переменных влияет на *размерность* модели.

Если нужный результат не получается модификацией компьютерной модели, меняется формальная модель вплоть до привлечения другого математического аппарата.

### 1.3. **Настройка на задачу**

Настройка на задачу представляет собой преобразование модели-прототипа в модель-экземпляр. Переменные модели заменяются константами, характеризующими конкретную задачу. В практическом плане настройка на задачу сводится к сбору и вводу в модель-прототип исходных данных. Этот этап оценивается *трудоёмкостью* подготовки модели к решению задачи.

### 1.4. **Решение задачи**

Все подготовительные этапы характеризуются *трудоёмкостью* в человеко-часах и *временем* выполнения этапа. Этап решения задачи оценивается привлекаемыми *вычислительными ресурсами* и *временем* решения задачи.

### 1.5. **Анализ результатов**

Результаты решения задачи анализируются на предмет достоверности. Достоверность вычисляемых величин определяется их *точностью*.

Таким образом, полная оценка качества модели должна отражать все этапы её жизненного цикла. Рассмотрим подробнее показатели качества моделей.

## 2. **Адекватность**

Степень соответствия свойств модели и объекта-оригинала является одним из важнейших показателей её качества. Одним из очевидных признаков соответствия модели объекту-оригиналу является подобие их свойств. Например, периодичность (колебательный процесс) адекватно выражается тригонометрическими функциями, скорость в задачах динамики – первой производной непрерывной функции, а ускорение – её второй производной.

Принцип прямого подобия реализуется на аналоговой вычислительной технике. Появление компьютеров нарушило принцип прямого подобия в силу самой дискретности их функционирования. Для решения задач функционального анализа были разработаны численные методы. Для вычисления значений неэлементарных функций применяются, как правило, итеративные алгоритмы.

Имитационные и многоагентные модели сложных систем ещё реализуют принцип структурного и функционального подобия с системой оригиналом так же, как и модели более примитивных объектов. Однако этот принцип уже не реализуем в нейросетевых моделях вычислительного интеллекта. Причина этого очевидна. Умственная деятельность, выполняемая нейронами человеческого мозга, не имеет прямого соответствия моделируемым ею системам. В этом и заключается основная проблема конструирования нейросетевых моделей. Поскольку процесс решения конкретной сложной задачи человеческим мозгом в настоящее время неизвестен, учёные избрали эмпирический (опытный) подход к созданию архитектуры нейросетевых моделей.

Таким образом, подобие в способах реализации функций объекта не является универсальным показателем адекватности модели. Отсюда логичен переход от подобия в реализации функций к подобию самих функций, т. е. переход к модели чёрного ящика, оперирующего только вход-выходными последовательностями. Его называют также гештальт-моделью в силу целостного взгляда на объект моделирования.

Рассмотрим возможность измерения адекватности модели поставленной задаче на примере распознавания рукописных символов. Пусть модели  $M_1$  и  $M_2$  распознают  $|X_1|$  и  $|X_2|$  символов соответственно. Очевидно, что качество модели 1 выше, если  $|X_1|/|X| > |X_2|/|X|$ , где  $|X|$  – число рукописных символов, подлежащих распознаванию. Если при этом  $|X_1| < |X|$ , модель 1 *частично* адекватна поставленной задаче.

Но оценивать следует не только число, но и составы  $X_1$  и  $X_2$  распознаваемых символов. Если  $X_1 \subseteq X$ ,  $X_2 \subseteq X$  и  $X_1 \cap X_2 = X_2$ , то  $X_2 \subseteq X_1$ , из чего следует, что модель  $M_1$  в большей степени соответствует заданным требованиям. Если же  $X_1 \cap X_2 \neq X_1$  и  $X_1 \cap X_2 \neq X_2$ , то модели *дополняют* друг друга. Очевидно, что если потребуется распознать символы из множества  $X_1 \setminus X_2$ , то следует использовать модель  $M_1$ , а для распознавания символов из множества  $X_2 \setminus X_1$  применить модель  $M_2$ .

### 3. Достоверность результата

Достоверность результата конкретизируется относительно типа решаемой задачи: точность для численных задач; отношение к глобальному оптимуму для оптимизационных задач; вероятность при решении статистических задач; коэффициент уверенности при решении логических задач.

Точность решения численной задачи определяется числом знаков после запятой. На точность результата влияет число итераций. Достоверность результата распознавания определяется вероятностью правильного распознавания анализируемой сущности. На вероятность правильного распознавания символа влияет состав и мощность обучающей выборки, а также количество итераций (эпох) обучения модели.

Таким образом, достоверность результата связана с трудоёмкостью его получения.

В задачах с высокой степенью неопределённости отсутствует база объективного сравнения. Иными словами, в момент получения результат просто не с чем сравнивать. Применяемые методы формализации неопределённости позволяют только вычислить коэффициент уверенности или степень риска. В этих условиях достоверность результата оценивается путём получения его разными методами.

### 4. Показатели оценивания качества оптимизационной модели

Возникает вопрос: «А можно ли сравнить качество нейросетевой модели с качеством модели другой природы?». На этот вопрос можно ответить утвердительно, если представить сопоставляемые модели как оптимизационные. Действительно, достоверность результата, получаемого с помощью нейросетевой модели, зависит от

качества её обучения. А задача обучения нейросетевой модели относится к классу задач оптимизации. Она формулируется как задача минимизации интегральной ошибки при обучении с учителем. Эта ошибка, как известно, определяется объёмом и качеством обучающей выборки.

С другой стороны, алгоритмы решения задач классическими методами также можно рассматривать как задачи оптимизации получаемого результата. Его достоверность (точность решения задачи) определяется объёмом итераций.

Для оценивания качества задач оптимизации привлекаются следующие критерии:

1. **Точность поиска** – значение окрестности локального оптимума, в которую приводит алгоритм после выполнения заданного числа итераций.

2. **Скорость сходимости** – число итераций, необходимое для достижения заданной точности.

3. **Время счёта** – время поиска на ЭВМ локального оптимума с заданной точностью, отнесенное к коэффициенту сложности задачи (или к быстродействию ЭВМ).

4. **Стабильность** – свойство алгоритма незначительно увеличивать число итераций при малых возмущениях выбора начальных точек, а также вследствие погрешности вычислений.

5. **Надёжность** – свойство алгоритма приводить к оптимуму при многократном повторении поиска из разных начальных точек.

#### 5. Задачи оценивания качества модели

В предыдущих разделах были рассмотрены различные свойства, характеризующие качество моделей. Состав свойств, подлежащих оцениванию, зависит от проблемы, которой озабочено лицо, принимающее решение (ЛПР). Приведём примеры постановки задач по оцениванию качества моделей.

1. Решается проблема разработки модели под новую задачу. Предложено два (или более) проекта моделей. Составлены технические задания на их разработку. Функциональность моделей одинакова, т.е. они в равной мере адекватны решаемой задаче. Доказана соизмеримая достоверность результатов решения задачи. Каждый разработчик расхваливает свой проект. Перед лицом, принимающим решение (ЛПР), возникла проблема выбора наиболее приемлемой модели. Его интересуют, прежде всего, временные затраты на проектирование и применение модели.

При такой постановке задачи модель оценивается по этапам жизненного цикла. Каждому этапу предъявляется частная цель и задаётся важность. При многократном решении типовой задачи к наиболее важным этапам относятся ввод исходных данных и время решения задачи. Если задача уникальная, то к наиболее важным относят первые два этапа жизненного цикла модели.

При ограниченности ресурсов ЛПР может оценивать этапы жизненного цикла моделей по трудоёмкости в человеко-часах или человеко-днях. Может принять комплексный вариант оценивания, охватывающий как трудоёмкость, так и время выполнения этапов.

2. Модели, разработанные и представленные для оценивания, отнесены к классу оптимизационных. В этом случае оцениванию подвергаются свойства, изложенные в разделе 5. Важность свойств определяется в зависимости от сферы использования моделей: исследование, ограниченное или многократное решение типовой задачи и пр.

3. Разработанные и представленные для оценивания модели различаются функциональностью. В этом случае для их оценивания привлекаются такие признаки как адекватность, сложность по памяти, вычислительная сложность, достоверность результата. Важность признаков ЛПР определяет в зависимости от имеющихся вычислительных ресурсов, требуемой функциональности модели и достоверности результата.

Перечисленными примерами не исчерпываются всевозможные постановки задач по оцениванию качества моделей. При детализации рассмотренных свойств модели оцениваются иерархией признаков. Важной особенностью этих задач является многоаспектность оценивания. Для решения таких задач используются методы критериального и функционального выбора.

#### **6. Методы оценивания качества моделей**

Методы критериального выбора оценивают объекты по значениям критериев. Они ограничиваются использованием порядковой шкалы и исчислением предикатов.

Методы функционального выбора оценивают объекты по значениям функций, созданных на основе критериев или по предпочтениям экспертов. Они используют численные шкалы и вычислительные операции.

Выбор метода осуществляется в зависимости от особенностей решаемой задачи, ресурсов ЛПР и требования к достоверности результатов. Под ресурсами ЛПР понимается время, выделенное на решение проблемы, и финансы, требуемые на создание модели выбора.

Методы критериального выбора обычно применяются для моделей выбора, содержащих малое число признаков с малым числом качественных значений. Увеличение числа признаков влечёт увеличение мощности множества Парето. Получение линейного порядка на множестве альтернатив связано с увеличением объёма экспертных оценок.

Методы функционального выбора ориентированы на большое число признаков с численными значениями. Они гарантируют получение линейного порядка на множестве альтернатив, но требуют дополнительной информации для реализации. К ней относится выбор шкал признаков, задание их важности, выбор функций, отображающих значения признаков в абсолютную шкалу  $[0, 1]$  и выбор вида обобщающей функции.

В направлении усложнения различают: функции достижения идеальной и реальной цели, функции отклонения от цели и функции полезности [ ]. Линейная функция достижения идеальной цели представляет собой нормирующую функцию целевого критерия. Кусочно-линейные функции достижения реальной цели и отклонения от цели формируются на основе ограничительного критерия. Нелинейная функция полезности отражает предпочтения эксперта о полезности делений на шкале признака и требует наиболее глубокого проникновения в суть задачи выбора.

Применительно к задаче оценивания качества моделей методы критериального выбора представляются неперспективными, поскольку эта задача требует установления линейного порядка на множестве и характеризуется не менее чем четырьмя признаками, измеренными в различных числовых шкалах. Выбор подходящего метода функционального выбора определяется глубиной проникновения в решаемую задачу.

## 7. Пример оценивания качества моделей

В качестве примера решим проблему разработки модели под новую задачу. Пусть представлены две модели-претендента с временными характеристиками, приведёнными табл. 1.

Модель	Разработка, месяц	Модификация, неделя	Настройка, часов	Решение, мин	Анализ, дней
№ 1	2	2	5	8	1
№ 2	3	1	3	6	2

ЛПР заинтересовано в минимизации времени на любом этапе жизненного цикла модели. Поэтому каждый признак подлежит минимизации. Из сравнения векторных оценок двух моделей следует, что они несравнимы и требуется дополнительная информация для выявления лучшего варианта. Учитывая численный характер векторных оценок, воспользуемся методами функционального выбора. Любой из методов требует задания важности признаков, по которым оцениваются модели.

Вычислим весовые коэффициенты этапов жизненного цикла моделей с помощью матрицы парных сравнений (МПС). Для этого достаточно задать  $N - 1$  предпочтений, соответствующих гамильтонову пути в матрице размерностью  $N \times N$ . Остальные клетки матрицы достраиваются автоматически с сохранением коэффициента согласованности [ ].

Воспользуемся МПС кратности предпочтений. Сформулируем следующие предпочтения, для простоты ограничившись лингвистическим значением умеренного превосходства:

Умеренное превосходство (Разработка, Модификация);

Умеренное превосходство (Настройка, Модификация);

Умеренное превосходство (Решение, Настройка);

Умеренное превосходство (Решение, Анализ).

Умеренное превосходство по Т. Саати кодируется как отношение 3:1 [ ]. Ниже представлена матрица, соответствующая введённым предпочтениям. Введённые в 4 клетки предпочтения выделены полужирным шрифтом. На их основе вычислены значения в остальных клетках матрицы.

Этап	Разработка	Модификация	Настройка	Решение	Анализ	Вес
Разработка		<b>3,00</b>	1,00	0,33	1,00	0,16
Модификация	0,33		0,33	0,11	0,33	0,05
Настройка	1,00	<b>3,00</b>		0,33	1,00	0,16
Решение	2,99	8,98	<b>3,00</b>		<b>3,00</b>	0,47
Анализ	1,00	2,99	1,00	0,33		0,16

На основе содержимого матрицы вычислен вектор весовых коэффициентов, показанный в столбце «Вес»:  $w = (0,16; 0,05; 0,16; 0,47; 0,16)$ .

В табл. 1 этапы жизненного цикла моделей представлены в различных единицах измерения. В то время как разработка модели измеряется месяцами, решение задачи измеряется минутами и т. д. Поэтому первоочередной задачей является задание области определения функций, создаваемых на шкалах признаков. Все функции должны отображать свои области определения в абсолютную шкалу [0, 1]. Наиболее простыми являются функции достижения идеальной цели (ДИЦ), вычисляемые нормированием значений критерия диапазоном шкалы.

Шкалы и предпочтения ЛПР

Таблица 2

Признак	Мин.зн.	Макс.зн.	Вес	Предпочтение	Нижн. граница	Верхн. граница
Разработка	2	3	0,16	Мин	0	5
Модификация	1	2	0,05	Мин	0	3
Настройка	3	5	0,16	Мин	0	8
Решение	6	8	0,47	Мин	0	10
Анализ	1	2	0,16	Мин	0	3

В левой части таблицы приведены границы выборочной шкалы для каждого признака. В правой части таблицы они расширены экспертным способом. Решение задачи методом оптимизации по достижению идеальной цели (ДИЦ) с аддитивной обобщающей функцией приведено в табл. 3.

Решение задачи методом ДИЦ

Таблица 3

Модель	Разработка	Модификация	Настройка	Решение	Анализ	Оценка	Ранг
№ 1	0,60	0,33	0,38	0,20	0,67	0,37	2
№ 2	0,40	0,67	0,63	0,40	0,33	0,44	1

Каждому признаку сопоставлены значения функций ДИЦ, а в столбце «Оценка» приведены значения обобщающей функции. Согласно этим значениям предпочтение отдаётся модели № 2.

Если ЛПР задаёт временные ограничения по минимальным значениям выборочных шкал, то метод оптимизации по отклонениям от цели даёт следующие результаты (см. табл. 4).

Решение задачи методом отклонений от цели

Таблица 4

Модель	Разработка	Модификация	Настройка	Решение	Анализ	Отклонение	Ранг
№ 1	0,00	0,50	0,40	0,50	0,00	0,324	2
№ 2	0,33	0,00	0,00	0,00	0,50	0,133	1

Каждому признаку сопоставлены нормированные значения штрафов при отклонении от цели. Модель № 1 и № 2 имеют соответственно 2 и 3 нулевых штрафа по тем критериям, чьи значения совпадают с минимальными границами выборочных шкал. В столбце отклонение (от цели) приведены средневзвешенные штрафы по всем критериям. Поскольку у модели № 2 он меньше, она признаётся лучшей.

Применим метод оптимизации по достижению реальной цели (РИЦ) с аддитивной обобщающей функцией. Совместим его как частный случай с методом оптимизации по полезности. Для этого построим нисходящую логистическую функцию полезности для признака «Решение». Функции РИЦ для остальных признаков строятся автоматически, полагая 50-ти процентную полезность в точках достижения реальной цели. Графики функций всех признаков в порядке их перечисления в табл. 2 приведены на рис. 1. На них показаны значения полезности признаков для модели № 2.

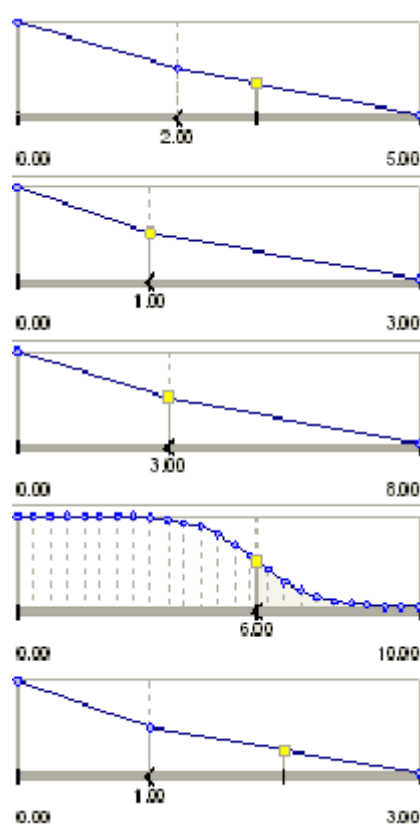


Рис. 1. Графики функций полезности пяти признаков

Решение задачи методом оптимизации по полезности с аддитивной обобщающей функцией приведено в табл. 5.

Решение задачи методом многомерной теории полезности

Таблица 5

Модель	Разработка	Модификация	Настройка	Решение	Анализ	Оценка	Ранг
№ 1	0,67	0,50	0,50	0,06	0,75	0,36	2
№ 2	0,50	0,75	0,70	0,50	0,50	0,54	1

Решение задачи разными методами, невзирая на различные численные оценки, дало одинаковые результаты. Такой результат ожидаем в связи с тем, что модель № 2 имеет большее число нулевых штрафов, а признаку «Решение» дана почти 50-ти процентная важность.

### Заключение

Предложены методы многокритериального оценивания моделей и приведены примеры их применения.

Данная работа была выполнена при финансовой поддержке РФФИ (гранты №№ 11-08-01016, 11-08-00767, 12-07-13119-офи-м-РЖД, 12-07-00302, 13-07-00279, 13-08-00702, 13-08-01250, 13-01-00912), Программы фундаментальных исследований ОНИТ РАН (проект №2.11), проекта ESTLATRUS 2.1/ELRI –184/2011/14 «Integrated Intelligent Platform for Monitoring the Cross-Border Natural-Technological Systems» (2012–2013 гг.), проекта ESTLATRUS/1.2./ELRI-121/2011/13 «Baltic ICT Platform».



## Литература

1. Калинин В.Н., Резников Б.А. Теория систем и управления (структурно-математический подход). – Л.: ВИКИ, 1987.
2. Юсупов Р.М., Иванищев В.В., Костельцев В.И., Суворов А.И. Принципы квалиметрии моделей // IV СПб Международная конференция «Региональная информатика-95», тез. докладов. – СПб, 1995.
3. Ростовцев Ю.Г., Юсупов Р.М. Проблема обеспечения адекватности субъектно-объектного моделирования// Известия ВУЗов. Приборостроение. - № 7, 1991.
4. Андрианов Ю.М., Суббето А. И. Квалиметрия в приборостроении. – Л. Машиностроение, 1990. – 216.
5. Технология системного моделирования / Е.Ф. Аврамчук, А.А. Вавилов, С.В. Емельянов и др.; Под общ. ред. С.В. Емельянова и др. – М.: Машиностроение; Берлин: Техника, 1988.
6. Соколов Б.В., Юсупов Р.М. Концептуальные основы оценивания и анализа качества моделей и полимодальных комплексов // Известия РАН. Теория и системы управления, 2004, №6, стр. 5-16.
7. Охтилев М.Ю., Соколов Б.В., Юсупов Р.М. Интеллектуальные технологии мониторинга и управления структурной динамикой сложных технических объектов. М.: Наука, 2006. 410 с.
8. Карпов Ю. . Имитационное моделирование систем. Введение в моделирование с AnyLogic 5. –СПб.: БХВ-Петербург, 2005, –400с.
9. Соболева Т.С., Чечкин А.В. Дискретная математика. – М.: Академия, 2006, –255 стр.
10. Микони С.В. Многокритериальный выбор на конечном множестве альтернатив. Учебное пособие. – СПб.: Лань, 2009, 272 с.
11. Микони С.В. Дискретная математика для бакалавра: множества, отношения, функции, графы: Учебное пособие. –СПб.: Издательство «Лань», 2012. –192 с.

**Конференция «Информационные технологии в управлении» ИТУ-2014, – СПб.: ОАО "Концерн "ЦНИИ "Электроприбор", 2014. – с. 44-53.**