

## СИСТЕМНЫЕ ТЕХНОЛОГИИ ГЕНЕРАЦИИ И АНАЛИЗА СЦЕНАРИЕВ

Гожий А.П., Коваленко И.И.

**Введение.** Основным показателем современного экономического развития выступает коренное изменение в технологическом способе производства, которое базируется на инновационном развитии общества. При этом большое внимание будет уделяться так называемым «сгенерированным» или «прорывным» технологиям, т.е. технологиям, которые еще не находятся на конкурентной стадии своего развития, но основываясь на самых прогрессивных идеях и новых эффектах, имеют все основания для государственного финансирования [1].

Для анализа принципиально новых проектов и прогноза фундаментальных и прикладных исследований разрабатывают сценарии, в основе которых лежит построение так называемого «дерева целей» или «прогнозного графа», которые отображают как инновационную потребность, так и инновационную альтернативу, в том числе оценку необходимых ресурсов и производственных возможностей.

Разработка сценариев – высококвалифицированная, творческая и слабоформализуемая работа, для которой нет исчерпывающего алгоритма поведения. При разработке сценариев желательна многовариантность описания будущей ситуации, что дает возможность анализировать и определять наиболее рациональную стратегию воздействия на ситуацию. Для разработки сценариев привлекаются различные группы высококвалифицированных специалистов, получивших название «аналитики»(А), «консультанты»(К), «эксперты»(Э), «лица, принимающие решения» (ЛПР), которые в своей работе выполняют определенные функции.

Понятно, что при разработке сценариев ни один из имеющихся методов изолированно, вне связи с другими не может дать исчерпывающих результатов. Поэтому стремление к использованию отдельных методов в определенной последовательности с установлением определенных взаимосвязей между ними, является очевидной необходимостью. В современной трактовке такой процесс получил название технологическое предвидение [1], которое и формируется на основе методологии системного анализа. Следует отметить, что вопросы формализации процедур построения сценариев на основе системного применения ряда инструментальных методов являются наименее разработанными, по ней практически отсутствует библиография, сослаться можно, пожалуй, лишь на работу [2]. В данной статье предложена системная методология построения сценариев, в рамках которой рассматриваются следующие вопросы:

- систематизированы и выделены основные группы так называемых «инструментальных методов», используемых для построения и анализа сценариев (методы генерации информации качественного характера; графовые модели, необходимые для построения деревьев целей или событий; вероятностно-статистические методы для получения оценок событий на графах и методы сворачивания сценариев; методы поддержки принятия решений по выбору лучшего из множества альтернативных сценариев).

- предложено с позиций системного подхода рассматривать определенную совокупность используемых инструментальных методов в виде системы методов (СМ), между элементами которой (в качестве элементов выступают отдельные методы) существуют определенные связи или формы взаимодействия;

- на основе анализа СМ рассмотрены процедуры формирования системных технологий выбора различных вариантов и последовательностей методов в соответствии с установленными критериями и принципами оптимальности их использования.

На начальном этапе изучения проблемы и её качественного анализа используются методы генерации информации качественного характера (мозговая атака, метод аналогий, метод контрольных вопросов, морфологический анализ, эвристические методы и др.), с помощью которой анализируют характерные особенности, определяют направления исследований, формулируют важнейшие критерии для выбранной проблемы (рис.1).



Рис.1 Системная методология построения сценариев

На этапе написания сценариев используются методы теории графов (построение дерева целей, дерева событий, прогнозный граф и др.), что дает возможность получить топологию цепочек событий, связанных причинно-следственными связями. Для количественного оценивания вероятностей событий, из которых формируется сценарий, и последующего его сворачивания применяются вероятностно-статистические методы (стратегия Байеса, многошаговый алгоритм Байеса, методы параметрической, робастной и непараметрической статистики).

После написания сценариев и оценивания их реалистичности последним важным этапом является их представление группе лиц, принимающих решения. В данной ситуации применяются методы поддержки принятия решений (однокритериальные методы при объективных моделях, многокритериальные методы при объективных моделях и многокритериальные методы при субъективных моделях).

**Постановка задачи.** Представим некоторую совокупность методов, используемую на различных этапах построения сценариев в виде системы методов, между элементами которой (в качестве элементов выступают отдельные методы), существуют определенные связи или формы взаимодействия (рис.2). Выявление таких связей и форм взаимодействия между методами с целью формирования процедуры по применению определенной их совокупности будем называть системной технологией выбора методов генерации и анализа сценариев.

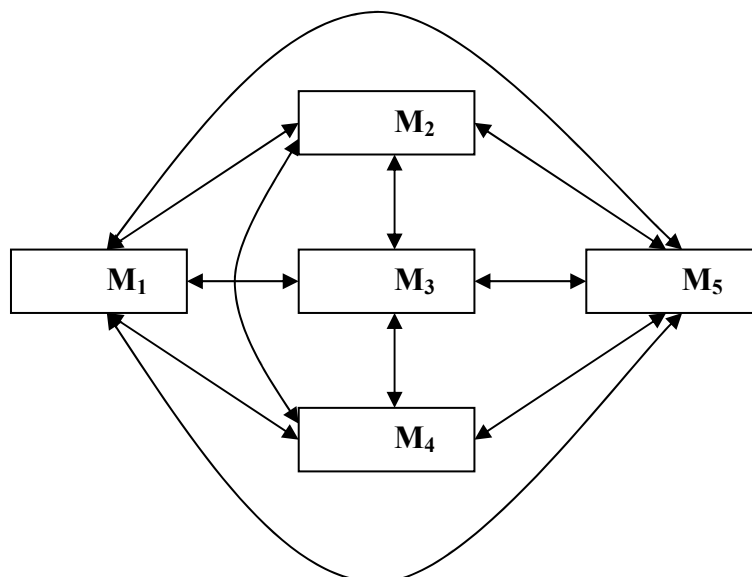


Рис. 2 Системное представление методов в виде ориентированного графа

Математически строго задачу выбора (В) можно описать парой:  $V = \langle M, O \rangle$ , где априори задано  $M = \{M_i\}$ ,  $i = 1, n$  – множество исходных методов генерации и анализа сценариев (M), O – принцип оптимальности, задаваемый лицом, принимающим решение (ЛПР) из совокупности требований по допустимости ( $O_{доп.}$ ) и критериальных требований (далее критериев), причем  $O = O_{доп.} \cup O_{кр.}$ ,  $O_{доп.} \cap O_{кр.} = \emptyset$ . Требования по допустимости  $O_{доп.}$  представляют собой требования к характеристикам объектов (методов) выбора (В) посредством отношений R-типа:  $R = \{=, \neq, <, >, \leq, \geq, \langle \rangle\}$ .

Критерии  $O_{кр.}$ , формируемые ЛПР, отображают его целевые устремления с учетом информационного обеспечения на данном этапе процедуры выбора. Решением задачи выбора (В) будем называть подмножество  $M_{opt} \in M$ , полученное с помощью принципа оптимальности O в следующей последовательности:

$$\begin{array}{ccc}
 M \rightarrow M_{доп.} \rightarrow M_{opt.} \\
 \uparrow \quad \quad \uparrow \\
 O_{доп.} \quad O_{кр.}
 \end{array}$$

то есть вначале выделяется множество допустимых методов  $M_{\text{доп}}$ , удовлетворяющих требованиям по допустимости, а затем уже на этом усеченном множестве проводится выбор методов по критериям  $O_{\text{кр}}$ . Рассмотренная процедура выбора методов была положена в основу построения ряда системных технологий некоторые из которых рассмотрим более подробно.

Системная технология выбора методов генерации информации на основе текстологического подхода. Методы генерации информации, используемые экспертами при построении сценариев в качестве инструментальных имеют как правило, качественный характер, что не позволяет формализовать взаимосвязи между ними аналитически. Поэтому одним из подходов, позволяющим сформировать системную технологию выбора последовательности их применения может быть подход на основе использования текстологических моделей извлечения знаний, основанных на изучении текстов.

Среди методов извлечения знаний этот подход является наименее разработанным, по нему отсутствует библиография поэтому для его краткого изложения воспользуемся положениями изложенными в [3]

Задачу извлечения знаний из текстов можно сформулировать как задачу понимания и выделения смысла текста, и представить следующей схемой:

$$M_1 \rightarrow V \rightarrow T \rightarrow I \rightarrow M_2,$$

где  $M_1$  – смысл который пытался заложить автор, его модель мира,  $M_2$  – смысл который постигает читатель в процессе интерпретации  $I$ , при этом  $T$  (текст) – словесное представление  $M_1$ , то есть результат вербализации  $V$ .

Рассмотрим подробнее какие источники питают модель  $M_1$  и создают текст  $T$ . В работе [3] такая модель представлена в виде:

$$M_1 = (\alpha, \beta, \gamma, \delta, \theta)_L,$$

где  $\alpha$  – первичный материал наблюдений;  $\beta$  – система научных понятий в момент создания текста;  $\gamma$  – субъективные взгляды автора, результат его личного опыта;  $\delta$  – некоторые “общие места” в тексте;  $\theta$  – заимствования из других источников (статей, монографий) и т.д.. При этом все перечисленные компоненты погружены в языковую среду  $L$ . Основным моментом понимания текста является формирование некоторой смысловой структуры текста за счет установления внутренних связей между отдельными важными (ключевыми) словами и фрагментами, а также за счет абстрактных понятий, обобщающих конкретные фрагменты знаний. При этом смысловая структура текста формируется посредством выделения «опорных», ключевых слов или «смысловых вех», которые связываются в единую семантическую структуру. Таким образом, процесс понимания (или интерпретации)  $I$  и модель  $M_2$  определяются следующими компонентами [3]:

$$M_2 = [(\alpha, \beta, \gamma, \theta), \omega, \varepsilon, \varphi],$$

где  $(\alpha, \beta, \gamma, \theta)$  – экстракт компонентов, почерпнутый из текста  $T$ ;  $\omega$  – предварительные знания аналитика о предметной области;  $\varepsilon$  – общенаучная эрудиция аналитика;  $\varphi$  – его личный опыт.

Процесс  $I$  – это сложный, не поддающийся формализации процесс, в основе которого лежит процедура разбивки текста на части («смысловые группы»), а затем компрессия (сжатие) текста в виде набора ключевых слов (НКС), передающих основное содержание текста. В качестве НКС могут служить любые части речи (существительное, прилагательное, глагол и т.д.) или их сочетаний.

На основании изложенного, рассмотрим одну из возможных технологий выбора методов качественного анализа, и представим её таблицей 1.

При этом в качестве критерия, на основе которого была сформирована последовательность применения данных методов являлось, было установление смысловых связей между понятиями в их описании.

Технологии выбора методов качественного анализа

Методы генерации информации качественного характера	Решаемые задачи	Дополнительные методы или информационные ресурсы
1. Аналогий	Поиск прототипа, принципов решения проблем	Информация из архивов, метод контрольных вопросов
2. Контрольных вопросов	Поиск путей решения проблем	-
3. Эвристических приемов	Преобразование прототипа, разрешение противоречия	Фонды эвристических приемов (индивидуальные, специализированные, межотраслевые)
4. Мозговой атаки (мозговой штурм, синектические методы)	Коллективный поиск идей, решений, предложений, проведение аналогий	Методы: аналогий, контрольных вопросов, эвристических приемов
5. Морфологического анализа	Анализ функциональных возможностей, варианты решений, их классификаторы и др.	Метод эвристических приемов
6. Написание сценария (дерево целей, прогнозный граф)	Системное прогнозирование всех аспектов генерации ОИМ	Метод морфологического анализа, информация, сгенерированная другими методами

**Системная технология формирования альтернативных стохастических графов.**

Развитие сценарного подхода привело к разработке двух основных взаимосвязанных между собой методов : прогнозного графа и «дерева целей», которые характеризуются альтернативным и стохастическим характером.

Это обстоятельство обуславливает адекватность моделирования этапов создания сценариев с помощью альтернативных стохастических графов.

Необходимо отметить, что преобразование, например, дерева целей в альтернативный граф не является алгоритмической проблемой, а выполняется на экспертном уровне. Тем не менее на основе анализа и последующего определения основных типов вершин дерева целей, может быть создан инструментарий, помогающий экспертам в автоматизированном режиме строить альтернативные графы, отображающие различные сценарии. В целом модель разработки альтернативных графов может быть организована как некоторая суперпозиция связанных между собой вершин различных типов.

Простейшими в данной модели являются вершины графа, на входе и выходе которых реализуются логические условия  $\wedge$  ( $\wedge$  - логическая операция «и»).

Для отображения различного рода альтернатив на входах и выходах вершин могут быть использованы логические условия  $\wedge$ ,  $\vee$ ,  $\bar{\vee}$  ( $\bar{\vee}$  - логическая операция, исключающая «или»). Причем любой тип входа может быть скомбинирован с любым типом выхода. Проведенный анализ рассматриваемых моделей показывает, что для отображения альтернативных ситуаций в реальном процессе среди всех типов вершин, которые образуются различной комбинацией входов и выходов, достаточно выбрать шесть следующих типов :

$$\wedge \in \wedge, \wedge \in \vee, \wedge \in \bar{\vee}, \vee \in \wedge, \vee \in \vee, \vee \in \bar{\vee}$$

Запись типов вершин в приведенном виде рассматривается таким образом. Для произвольной вершины  $e$  графа имеются логические условия на входе и выходе.

Например, тип  $\bigwedge_e \bigvee$  означает, что на входе  $e$  имеет место условие «и», т.е. вершина  $e$  считается свершенной после окончания всех работ, непосредственно предшествующих ей;

условие  $\bigvee_e$  на выходе вершины  $e$  означает, что будет реализовываться одна и только одна работа из всех работ, исходящих из неё.

При реализации разработки встречаются ситуации, когда дальнейшее осуществление процесса, т.е. выполнение исходящих из событий дуг-работ существенно зависит от реализации дуг на входе событий. Для отображения таких ситуаций вводится дополнительно два типа вершин:

$$\bigvee_e (\bigvee P_i), i \in \Gamma_e^-;$$

$$\bigvee_e \bigvee (P_i), i \in \Gamma_e^+,$$

где  $\Gamma_e^-$  - множество событий, из которых исходят работы, входящие в вершину  $e$ ;

$\bigvee_e \bigvee (P_i)$  – обозначает тип вершины, реализация которой на выходе зависит от реализации дуги  $i$  на входе события  $e$  с данным логическим условием;  $\Gamma_e^+$  - множество событий (вершин), которые исходят из вершины  $e$ .

Таблица 2.

Кодировка входа	Логические возможности на входе	Обозначения
1.	Логическая операция $\bigwedge$	$\bigwedge_e$
2.	Логическая операция $\bigvee$	$\bigvee_e$
3.	Логическая операция $\bigvee$ вводится только для приведения графа к одной конечной вершине	$\bigvee_e$
Логические возможности на выходе		
1.	Логическая операция $\bigwedge$	$e \bigwedge$
2.	Логическая операция $\bigvee$ , $\sum P_{e,j} = 1, j \in \Gamma_e^+$ Логическая операция $\bigvee$ , $0 < P_{e,j} \leq 1$ для всех $j \in \Gamma_e^+$	$e \bigvee$ $e \bigvee$
3.	Логическая операция $\bigvee$ , реализация на выходе зависит от реализации дуги $(i,e)$ на входе события $e$ , $0 < P_{e,j} \leq 1$ для всех $j \in \Gamma_e^+, i \in \Gamma_e^-$	$e \bigvee (P_i)$
4.	Логическая операция $\bigvee$ , реализация на выходе зависит от реализации дуги $(i,e)$ на входе события $e$ , $\sum P_{e,j} = 1$ для всех $j \in \Gamma_e^+, i \in \Gamma_e^-$	$e \bigvee (P_i)$
5.	Логическая операция $\bigvee$ , реализация на выходе зависит от реализации дуги $(i,e)$ на входе события $e$ , $\sum P_{e,j} = 1$ для всех $j \in \Gamma_e^+, i \in \Gamma_e^-$	$e \bigvee (P_i)$

Перечисленные типы вершин, вообще говоря, исчерпывают различные ситуации, возникающие на начальных стадиях процесса разработки сценариев, допускающие альтернативу, но модель позволяет использовать не только указанные типы вершин, но и любую комбинацию из входов и выходов, приведенных в таблице 2.

Таким образом, рассмотренный системный подход основан на представлении процесса создания сценария в виде многовариантного альтернативного стохастического

графа, в котором в той или иной комбинации применяются восемь типов вершин, отображающих различные ситуации в разрабатываемых сценариях.

### Системная технология выбора вероятностно-статистических методов

Рассмотрим один из возможных подходов системного использования статических методов. При этом в качестве критериев выбора будем рассматривать априорные информационные уровни задания функции распределения вероятностей, которая является полной характеристикой измеряемых случайных величин (табл. 3). Рассмотрим данные уровни более подробно [5]:

1. Задание строгих параметрических моделей распределения вероятностей  $F$  (нормального, равномерного, пуассоновского, экспоненциального и др. законов) обеспечивается объединением следующих информационных составляющих в символах алгебры логики:

$$F \sim I_1 \wedge I_2 \wedge I_3 \wedge I_4,$$

где  $F$  – функция распределения вероятностей;  $I_1$  – информация об общей схеме (механизме) формирования случайных величин данной природы;  $I_2$  – область значений случайной величины;  $I_3$  – примеры реальных признаков, подчиняющихся данному закону;  $I_4$  – аналитическое задание (модель) закона и его определяющие параметры.

В данной ситуации для обработки измерительных данных могут быть использованы методы параметрической статистики для решения задач (табл. 3).

2. Задание полной окрестности параметрических моделей распределения вероятностей базируется на том, что распределения реально наблюдаемых случайных величин в подавляющем большинстве отличны от строгих параметрических моделей (например, от нормальных или гауссовских).

Такие модели можно охарактеризовать как "искаженные" параметрические модели (например, от тяжелые "хвосты", многомодальность, резко выраженные асимметрия или эксцесс). Такие искажения обязаны своим появлением наличием, как правило, в исследуемых выборках данных определенной доли (обычно относительно небольшой) так называемых "засоряющих" значений. Формально такие модели представляются симметричными и несимметричными смесями распределения вероятностей соответственно [4]:

$$F(x) = (1 - \varepsilon)\varphi(x; a; \sigma_0^2) + \varepsilon\varphi(x; a; \sigma_1^2)$$

$$F(x) = (1 - \varepsilon)\psi(x; \theta; \sigma) + \varepsilon\hat{h}(x; a; \sigma_1^2),$$

где  $\varphi(x; a; \sigma_0^2)$  – плотность нормального распределения со средним значением  $a$  и дисперсией  $\sigma_0^2$ ;  $\varepsilon$  – доля "засоряющих" наблюдений;  $\sigma_0^2 < \sigma_1^2$ ;  $\theta$  – параметр сдвига;  $\hat{h}$  – плотность некоторого симметричного закона распределения.

Для рассмотренных моделей целесообразно использовать методы робастной (устойчивой) статистики с целью получения несмещенных (малосмещенных) и эффективных оценок параметров положения и масштаба (табл.3).

3. Задание только различий между распределениями вероятностей рассматривается в рамках непараметрической статистики, которая не делает предположений о том, что функция распределения результатов наблюдений принадлежит тем или иным параметрическим семействам распределений.

Поэтому самым характерным для непараметрической статистики является то, что в ее задачах распределение вероятностей считается полностью неизвестным, а сами задачи формулируются в терминах только различий между классами или внутри класса неизвестных распределений.

Таблица 3.

Уровень априорного информационного обеспечения	Задание строгих параметрических моделей распределения вероятностей	Задание полной окрестности параметрических моделей распределения вероятностей	Задание только различий между распределениями вероятностей
Характер статистических методов	Методы параметрической статистики	Методы робастной статистики	Методы непараметрической статистики
Основные решаемые задачи	Оценивание параметра положения. Оценивание параметра масштаба. Задачи оценивания других числовых характеристик законов распределения.	Получение несмещенных (малосмещенных) и эффективных оценок параметров положения и масштаба. Задачи устойчивого оценивания других числовых характеристик распределений вероятностей.	Проверка непараметрических гипотез: задача согласия; задача сдвига (расположения); задача расположения и симметрии; задача масштаба и др.

Одной из наиболее развитых и распространенных категорий непараметрических задач является проверка непараметрических гипотез. Рассмотрим одну из них, которая получила название задачи согласия. Пусть задано известное непрерывное распределение  $F(x)$ . Из неизвестного распределения  $G(x)$ , принадлежащего классу всех распределений, берется выборка  $x_1, x_2, \dots, x_n$ . Формируются гипотезы:

нулевая гипотеза  $H_0 : F = G$  – простая гипотеза;

альтернатива  $\left. \begin{array}{l} \text{а) } H_1^+ : F < G \\ \text{б) } H_1^- : F > G \end{array} \right\}$  – односторонние гипотезы;

в)  $H_1 : F \neq G$  – двусторонняя гипотеза

Для наглядности все приведенные рассуждения представлены в форме таблицы 3.

В заключении необходимо отметить, что изложенный материал не претендует на полноту изложения затронутой проблемы, тем не менее, предложенные подходы по мнению авторов могут представить интерес для специалистов различных сфер деятельности.

This paper is description system methodology of development screenplays. The task of purposes system methodology for development screenplays apply different instrumental methods.

1. М.З.Згуровський Системна методологія передбачення – Київ: "Політехніка", 2001. - 49 с
2. М.З.Згуровський , Н.Д. Панкратова Технологическое предвидение.-Київ: "Політехніка", 2005. - 154 с.
3. Т.А.Гаврилова, В.Ф.Хорошевский Базы знаний интеллектуальных систем.-СПб.:Питер, 2001.-384 с.
4. Устойчивые статистические методы оценки данных./ Под редакцией Н.Г.Волкова; Пер.с английского Ю.А.Малахова.-М.:Машиностроение, 1984.-232 с.
5. И.И.Коваленко, А.П. Гожий, Т.В. Пономаренко Системный подход к выбору методов анализа сложных физических систем// Материалы международной научн. Конференции ISDMIT-2005, Евпатория, 2005, с.75-78.