

УДК 519.95:681.3.06:001.18:303.444

ТЕХНОЛОГИЯ ПРОГНОЗИРОВАНИЯ РАЗВИТИЯ
ТЕХНИЧЕСКИХ СРЕДСТВ И СИСТЕМ НА ДЛИТЕЛЬНУЮ ПЕРСПЕКТИВУ
ПРИ СУЩЕСТВЕННО ОГРАНИЧЕННОЙ ОБУЧАЮЩЕЙ ВЫБОРКЕ

Ю.А. Устюгов

В в е д е н и е

Разработка программ развития технических средств и систем (далее - ТС) на длительную перспективу (например, на 1995-2015 гг. и более длительный период времени) предполагает решение целого ряда в общем случае взаимосвязанных задач [1]:

а) построение наиболее рационального плана НИОКР, серийного производства и эксплуатации ТС на период 1995-2015 гг.;

б) выбор минимальной номенклатуры ТС, обеспечивающих на интервале времени 1995-2015 гг. выполнение требуемого объема работ при ограниченных ресурсах;

в) выделение приоритетного ряда ТС и др.

Реально известна информация о серийно выпускаемых и находящихся на стадии НИОКР технических средств и систем. Эта информация, как правило, высвечивает перспективу на 5-7 лет, т.е. до 2000-2002 гг. Вместе с тем для успешного решения сформулированного выше класса задач требуется информация о возможном облике рассматриваемых ТС на промежутке времени 2002-2015 гг. Получить такую информацию можно в ходе прогнозирования развития ТС на указанном периоде времени.

В статье излагается технология прогнозирования развития ТС на длительную перспективу, апробированная в двух разных предметных областях. По существу, в ней обобщаются результаты методологических, научно-методических и экспериментальных пяти-летних исследований, направленных на разработку надежной и эффективной технологии долгосрочного прогнозирования развития ТС в условиях априорной ограниченности (дефицита) исходных данных о них.

Существенным в обоих приложениях являлось то, что объем обучающей выборки был ограничен как по числу ТС, так и по количеству характеристик ТС, определяющих пространство признаков. Кроме того, от 10% до 40% элементов исходной базы данных были заполнены пробелами.

Определена взаимосвязь излагаемой технологии с созданным экспериментальным программным обеспечением, инструментально поддерживающим автоматизированный режим выполнения различных этапов технологии.

1. Постановка задачи на прогнозирование

На период времени Δt составляется сводная перспективная программа НИОКР, серийного производства и эксплуатации ТС. Пусть $\Delta t = T_n - T_k$, где T_n и T_k - годы соответственно начала и конца планового промежутка; T - текущий год планового промежутка, т.е. $T \in [T_n, T_k]$.

На год T_n ретроспективные, серийно выпускаемые, находящиеся в стадии ОКР и НИР технические средства и системы составляют заданный ряд $\{.\}^3$. Пусть ТС заданного ряда пронумерованы следующим образом:

$$\{n\}^3 = \{1, n_{lm}^3\}, \quad l = \overline{1, L}; \quad m = \overline{1, M}, \quad (1)$$

где n_{1m}^3 - заданное число ТС 1-го вида ψ -типа.

Для n -го ТС, $n = \overline{1, N}$, введем следующие обозначения:

α_n^1 - объем работы, выполняемый n -м ТС для решения 1-й задачи;

t_n^H - год начала НИОКР по созданию n -го ТС;

t_n^C - год начала серийного производства n -го ТС;

$t_n^Э$ - год снятия с серийного производства n -го ТС, при этом эксплуатация n -го ТС продолжается;

t_n^B - год снятия с эксплуатации n -го ТС;

C_n^H - стоимость НИОКР по созданию n -го ТС;

C_n^C - стоимость серийного производства одного экземпляра n -го ТС;

$C_n^Э$ - стоимость эксплуатации в течение года одного экземпляра n -го ТС;

U_n - число лиц, задействованных для обеспечения эксплуатации n -го ТС;

X_{qn} - q -я характеристика n -го ТС, ($q = \overline{1, Q}$, где Q - общее число характеристик n -го ТС).

1.1. Исходные данные для прогнозирования. Облик заданных рядов ТС $\{.\}^3$ полностью известен. Другими словами, известны характеристики: $\alpha_n^1, t_n^H, t_n^C, C_n^H, C_n^C, C_n^Э, X_{qn}$ при $n = \overline{1, n_{1m}^3}$; $l = \overline{1, L}$; $m = \overline{1, M}$; $q = \overline{1, Q}$, которые определяют обучающую выборку для прогнозирования. Для повышения эффективности и управляемости процедуры прогноза облик прогнозируемых рядов $\{.\}^n$ перспективных ТСО частично задается экспертами. В частности, априори известны характеристики t_n^H, t_n^C, X_{qn} при $n = \overline{1, n_{1m}^n}$; $l = \overline{1, L}$; $m = \overline{1, M}$; $q = \overline{1, Q_1}$, $Q_1 < Q$, где n_{1m}^n - число прогнозируемых ТС 1-го вида, ψ -го типа; Q_1 - число характеристик ТС прогнозируемого ряда, определяемых накануне прогноза.

Таким образом, в исходных данных на прогнозирование будут определены значения части временных и иных характеристик перспективных комплексов.

1.2. Конечные результаты прогнозирования. На этапе прогнозирования требуется доопределить облик прогнозируемых рядов перспективных ТС $\{.\}^n$, т.е. найти значения следующих характеристик: $\alpha_n^1, C_n^H, C_n^C, C_n^Э, X_{qn}$ при $n = \overline{1, n_{1m}^n}$; $l = \overline{1, L}; m = \overline{1, M}$; $q = \overline{Q_1 + 1, Q}$.

Таким образом, в ходе прогнозирования перспектив развития ТС на основе данных о N^3 ТС,

$$N^3 = \sum_{l=1}^L \sum_{m=1}^M n_{1m}^3, \quad (2)$$

формируются данные о N^n ТС,

$$N^n = \sum_{l=1}^L \sum_{m=1}^M n_{1m}^n. \quad (3)$$

2. Основные этапы технологии прогнозирования

В общем случае можно определить следующие основные этапы технологии прогнозирования.

1. Анализ предметной области.
2. Формулировка требований к процедуре прогнозирования.
3. Постановка задачи на прогнозирование.
4. Формализация и классификация задач прогнозирования.
5. Выработка требований к составу и содержанию исходных данных для решения задач прогнозирования.
6. Сбор, анализ и систематизация исходных данных для процедуры прогнозирования.
7. Формирование баз данных по ретроспективным, серийно выпускаемым и находящимся в стадии НИР и ОКР ТС.
8. Подготовка баз данных к процедуре прогнозирования.

9. Разработка сценария проведения эксперимента.
10. Подготовка и редактирование входных файлов.
11. Реализация сценария прогноза.
12. Анализ и редактирование баз данных результатов прогноза.
13. Завершение процедуры прогнозирования.

На первом этапе определяется признаковое пространство анализа ТС, т.е. величины Q , L , M , N^3 , $N^П$, а также проводится классификация ТС.

Формирование требований к процедуре прогнозирования на втором этапе осуществляется с точки зрения тех задач, которые решает потенциальный пользователь и заказчик прогнозных исследований. На этом этапе, в частности, уточняются требования к надежности прогноза (надежность оценивается процентом ожидаемой погрешности прогноза), определяются значения величин Δt , T_n , T_k .

На третьем этапе осуществляется формализация исходных данных и конечных результатов (см. п. 1). Заметим, что именно в тезаурусе и обозначениях, введенных на этом этапе, формулируются результаты исследований, проводимых на последующих этапах.

Формализация и классификация задач прогнозирования, выполняемая на четвертом этапе, по сути позволяет определить функциональные возможности реализуемого научно-технического подхода.

В ходе выполнения этапов 1-5 по сути решаются системные вопросы, в частности, определяются структура и размерность баз данных, наполняемых и корректируемых на последующих этапах (6-12), определяются требования к экспериментальному программному комплексу, инструментально поддерживающему выполнение, например, этапа 12.

Опыт показывает, что выполнение этапа 6 связано с целым рядом трудностей. Наиболее ответственным и сложным является

выявление характеристик α_n^1 , $n = \overline{1, n^3}$, $l = \overline{I, L}$, и определение их значений. В выбранном пространстве признаков часть ТС либо плохо описаны, либо часть характеристик для отдельных ТС имеют уникальный характер.

Обучение механизма процедуры прогнозирования осуществляется на ретроспективе анализируемых ТС, причем чем глубже взгляд в прошлое этих изделий (чем больше удален год T_H от текущего года T , например, $T - T_H \geq 20$), тем труднее восстановить значения таких важных характеристик ТС как t_n^H и t_n^C , $n = \overline{1, n^3}$, определяющих временную ось исследований.

Сформулированные выше утверждения относительно сложности определения характеристик α_n^1 , t_n^H , t_n^C справедливы и в отношении отыскания значений характеристик C_n^H , C_n^C , C_n^3 , $n = \overline{1, n^3}$. Во-первых, это объясняется меняющимся темпом инфляции, во-вторых, стоимость эксплуатации C_n^3 n -го ТС в год $T \in [T_H, T_K]$ часто определяют как 5-7% от C_n^C , что не всегда справедливо, так как величина C_n^3 существенно зависит в первую очередь от характеристик надежности ТС X_{qn} , $q \in Q$.

Чаще всего экспертным путем на этапе 6 определяется строка предельно допустимых значений всех характеристик на момент времени $T = T_K$.

На этапе 7 осуществляется наполнение баз данных по ретроспективным, серийно выпускаемым и находящимся в стадиях НИР и ОКР техническим средствам и системам. Именно на этом этапе появляется множество элементов баз данных, заполненных пробелами ("дырками"). Именно в "дырках" публикуется несовершенство собранных исходных данных для прогнозирования. Причина несовершенства - в существенной ограниченности обучающей выборки.

Основная цель этапа 8 - устранить "дырки" в исходных данных. На этом этапе осуществляется, с одной стороны, формирование и настройка механизма процедуры прогнозирования развития

ТС, с другой стороны, так как заполнение пробелов в исходных базах данных - суть прогнозирование значений характеристик в местах расположения "дырок", созданный на этапе 8 механизм прогнозирования на этом же этапе впервые и применяется. Создание механизма процедуры прогнозирования на этапе 8, адаптированного и настроенного на выбранную предметную область, составляет цель процесса "лечения" баз исходных данных. Это весьма сложный, а главное трудоемкий, хотя и автоматизированный процесс. Именно на этом этапе формируется механизм гарантии надежности прогнозирования. Опыт проведения прогнозирования перспектив развития технических средств и систем, показывает, с одной стороны, трудоемкость и сложность процедуры прогнозирования, с другой стороны, плохую описанность ТС в выбранном признаковом пространстве. В силу этих обстоятельств в ходе "лечения" баз данных, с одной стороны, часть недостающих исходных данных вычисляется по определенным методикам вручную. С другой стороны, с помощью специальных методов из баз данных устраняется часть строк и столбцов, часть строк из одних баз данных искусственно вводится в "чужие" базы данных. Наконец, систематически при "лечении" баз данных на этапе подготовки к процедуре прогнозирования выполняется известная [2] процедура заполнения пробелов. В отличие от ранее практиковавшихся и успешно работающих в других предметных областях подходов развиваемый в данной статье принципиально основан на том, что центр тяжести всех усилий по обеспечению требуемого значения ожидаемой погрешности прогноза находится на этапах подготовки баз данных к процедуре прогнозирования (этап 8), а не на этапе собственно процедуры прогноза (в этом случае погрешность метода снижается за счет введения специальных фильтров на выходе процедуры прогнозирования, устраняющих ложные прогнозы). Другими словами, добившись модификацией и редактированием баз данных перед этапом прогнозирования

ожидаемой погрешности при заполнении пробелов в интересующем столбце одной строки (прогнозного варианта) 15-20%, можно далее построчно предлагать для заполнения новые неизвестные значения прогнозируемой характеристики и всякий раз получать во всем массиве прогнозных вариантов данную погрешность, не прибегая к дополнительным фильтрам. Сказанное выше позволяет заключить: технология при достаточной описанности ТС в выбранном признаковом пространстве обеспечивает ожидаемую ошибку прогноза 15-20%. Таким образом, по своим результатам этап 8 весьма прост: в исходных базах данных устраняется часть "собственных" строк и столбцов и вносится часть строк из "чужих" баз данных с тем, чтобы обеспечить потребную надежность прогнозирования. Этап 8 - стержневой этап технологии: до 50-60% ресурсов времени работы с ПЭВМ при реализации технологии уходит на его выполнение.

При выполнении этапов 1-7 исполнители (эксперты) по-своему классифицируют ТС, т.е. определяют влияние различных ТС друг на друга в базах данных (по строкам) и корреляцию характеристик ТС (по столбцам). В ходе выполнения этапа 8 модифицируется, редактируется автоматизированным способом классификация ТС, перечень рассматриваемых характеристик, т.е. по существу - признаковое пространство.

При выполнении этапа 9 прежде всего выявляются вопросы, на которые нужно ответить в ходе экспериментального прогнозирования. Созданное экспериментальное программное обеспечение позволяет ответить на любые вопросы пользователя и заказчика прогнозных исследований. В каждом конкретном случае будет в конечном счете свой сценарий проведения эксперимента, по-своему подготовленные и оформленные во входных файлах (этап 10) исходные данные. В зависимости от версии экспериментального программного обеспечения последнее может инструментально поддерживать автоматизированный режим выполнения различных этапов технологии.

Этапы 1-6, 9, 12, 13 технологии в настоящее время выполняются вручную в силу объективной необходимости участия человека при выполнении этих этапов. Выполнение этапов 7, 8, 10, 11 с помощью ЭВМ классом не ниже АТ-286 с сопроцессором не предполагает каких-то специальных требований к памяти.

В целом более старые версии экспериментального программного обеспечения, работа над которыми планомерно осуществляется, предоставляют потенциальному пользователю более широкое поле органов управления для экспериментального прогнозирования перспектив развития ТС.

В младших версиях экспериментального программного обеспечения вариантность исследований ограничивается вводимыми константами во входные файлы данных. В связи с этим пользователю при выполнении этапа 10 предоставлено право менять, например, значения любых характеристик N^3 ТС, находящихся в базе данных для прогнозирования. В связи с этим выполнение этапа 10 сводится по существу к редактированию базы исходных данных для прогнозирования, в которой отсутствуют "дырки".

Заметим, что у пользователя всегда существует возможность редактирования исходных данных на этапах 7 и 8. При этом на этапе 8 пользователь может по своему усмотрению "вылечить" исходные данные: заполнить часть "дырок". Остальные "дырки" заполнит инструмент. При более детальном выполнении этапа 8 может проводиться анализ исходных данных на предмет противоречий в них.

Выполнение этапа 11 длится в зависимости от объема исходных данных и числа просматриваемых вариантов десятки минут (например, 20 минут при $N^1 = 10^4$).

Анализ и редактирование результатов прогноза на этапе 12 сводится к исключению из N^1 технических средств и систем, неприемлемых с точки зрения пользователя (экспертов), вводу своих прогнозных вариантов ТС (их требуется описать в том же при-

знаковом пространстве), изменению части значений характеристик в $N^П$ ТС (число $N^П$ при этом не меняется). На этапе 12 осуществляется выбор потребного числа прогнозных вариантов анализируемого класса ТС для процедуры планирования [1].

В целом процедура прогнозирования носит итерационный характер. На этапе 13 принимается решение о достаточности результатов прогнозирования для ответа на вопросы, сформулированные при постановке эксперимента (на этапе 9). Если данных не достаточно и есть возможность их расширить, то осуществляется возврат к этапам 6-9. При необходимости уточняется даже признаковое пространство (уход на этап 1). Кроме того, сценарий предусматривает, как правило, проведение серий экспериментов. В связи с этим, по завершении выполнения очередного эксперимента естественен переход к началу выполнения следующего эксперимента сценария. При положительном исходе на этапе 13 консервируются входные и выходные файлы данных. На этом экспериментальное прогнозирование завершается.

Важным достоинством развиваемого подхода является осознанная управляемость основными параметрами процесса прогноза (выбранные характеристики, требуемое число прогнозных вариантов, принудительный просмотр в требуемых точках признакового пространства ...). Это становится понятным, когда нужно мотивировать все решения выбора при реализации технологии.

3. Анализ предметной области

Существенной особенностью и сложностью создания механизма прогнозирования является то, что в период разработки последний однажды настраивается на данные, близкие к реальным, и в период эксплуатации реализует свои функции без перестройки. Другими словами, перестройка механизма прогнозирования при, например, существенном изменении структуры и содержания баз исходных данных может осуществляться только с участием разработчи -

ков экспериментального программного обеспечения. Отмеченная особенность технологии прогнозирования связана с тем, что в период настройки механизма прогнозирования на конкретную предметную область (конкретные ряды заданных и прогнозируемых перспективных ТС, с помощью которых решаются выбранные L задач) осуществляется генерация множества математических моделей для прогнозирования.

Выбор адекватных моделей из сгенерированного множества осуществляется на этапе 8 в интерактивном режиме с использованием знаний экспертов-разработчиков.

Выбранная совокупность математических моделей по существу составляет базу моделей - математический инструмент, настроенный на решение задач прогнозирования при заданных структуре и содержании баз исходных данных для прогнозирования, отражающих особенности предметной области.

Отмеченное выше в первую очередь объясняет почему при создании инструментальных средств прогнозирования необходимы исходные данные, максимально близкие к реальным. Это положение должно выполняться уже при реализации первого этапа технологии.

4. Требования к процедуре прогнозирования

В общем случае комплексы ТС (КТС) состоят из различного набора технических средств и систем. Прогнозирование развития КТС, в связи с этим, сводится в конечном счете к прогнозированию развития различных ТС, составляющих КТС. Процедура прогнозирования должна позволять прогнозировать развитие любых ТС как компонент КТС. Существует отдельный случай, когда КТС рассматривается как неделимая на ТС система со своей предисторией, своими характеристиками. В этом частном случае процедура прогнозирования должна обеспечить прогноз развития КТС по той же технологии, по которой выполняется прогноз отдельных ТС. Другими словами, процедура прогнозирования должна быть универсальной.

Ретроспектива развития во времени как отечественных так и зарубежных ТС должна составлять обучающую выборку для процедуры прогнозирования. При подготовке исходных данных не должно быть отличий для ТС различного народно-хозяйственного назначения.

Пользователь в режиме диалога должен иметь возможность редактирования исходных данных для прогноза и заказывать требуемое число прогнозных вариантов ТС.

В интеллектуальной технологии прогнозирования могут быть заложены различные методы прогнозирования. Пользователю должно быть недоступно их различие. Процедура должна быть настроена на предметную область таким образом, чтобы в технологии работали наиболее эффективные для данной предметной области методы прогнозирования.

Источниками исходной информации для процедуры прогнозирования должны быть исходные данные (ряды) о ретроспективных, серийно выпускаемых, находящихся в стадиях НИР и ОКР технических средств и систем. В перспективе дополнительным источником для разрабатываемой процедуры будут знания экспертов, извлекаемые с помощью специально разработанных инструментальных средств.

Таким образом, по заданным рядам ТС и заказываемому числу прогнозных вариантов разрабатываемая процедура должна обеспечить формирование прогнозного ряда ТС, например, до 2020 г. включительно.

Пользователь при анализе результатов прогноза должен иметь возможность их редактировать следующим образом:

а) исключать из прогнозного ряда ТС отдельные изделия, характеристики которых по мнению пользователя, как эксперта, противоречивы;

б) вводить в прогнозный ряд ТС свои варианты прогнозных изделий;

в) редактировать отдельные характеристики оставленных в прогнозном ряде ТС изделий.

Число прогнозных вариантов ТС может либо определяться диапазонами и шагом изменения варьируемых при прогнозировании характеристик ТС, либо задаваться пользователем. В первом случае преследуется цель просмотреть тенденции развития в наперед заданном наборе точек пространства признаков. Во втором случае - в произвольном наборе заданного числа точек.

Опыт проведения прогнозных исследований в разных предметных областях показал, что чаще всего возможности и потребности заказчиков удовлетворялись при $T_H = 1975$ г., $T_K = 2015-2020$ гг., $\Delta t = 40-45$ лет. Вместе с тем ретроспективу целесообразно просматривать настолько глубоко, насколько она прописана. То есть T_H можно брать раньше 1975 г., например, 1945 г. В этом случае при создании механизма процедуры прогнозирования появится возможность в большей степени учесть революционные всплески в прошлом.

5. Функциональные задачи прогнозирования

После создания и настройки механизма процедуры прогнозирования на заданную предметную область пользователь должен иметь возможность решать функциональные задачи прогнозирования, пример формулировки части из которых приведен ниже.

ПРИМЕР. Исследовать влияние на прогнозируемые характеристики перспективных ТС $\{.\}^n$ (см.стр. 74):

а) изменение возможностей ТС заданного ряда $\{.\}^3$ выполнять 1-ю работу ($l = \overline{1, L}$), т.е. α_n^1 , $n \in N^3$;

б) переноса времени начала НИОКР t_n^H , направленной на создание n -го ТС, $n \in N^3$;

в) переноса времени начала серийного производства t_n^C n -го ТС, $n \in N^3$;

г) изменения стоимости НИОКР C_n^H , направленной на создание n -го ТС, $n \in N^3$;

д) изменения стоимости серийного производства C_n^C n -го ТС, $n \in N^3$;

е) изменения стоимости эксплуатации $C_n^Э$ n -го ТС, $n \in N^3$;

ж) изменения значения q -й характеристики n -го ТС, $q = \overline{1, Q}$; $n \in N^3$;

з) изменения числа Q_1 (или состава) заданных при прогнозе характеристик X_{qn} , $q = \overline{1, Q_1}$; $n \in N^3$;

и) изменения величины n_{1m}^3 при $l = \overline{1, L}$; $m = \overline{1, M}$.

6. Требования к составу и содержанию исходных данных для прогнозирования

Характеристики α_n^1 , t_n^H , t_n^C , C_n^H , C_n^C , $C_n^Э$, X_{qn} при $n = \overline{1, n_{1m}^3}$;

$l = \overline{1, L}$; $m = \overline{1, M}$; $q = \overline{1, Q}$, а также t_n^H , t_n^C , X_{qn} при $n = \overline{1, n_{1m}^p}$;

$l = \overline{1, L}$; $m = \overline{1, M}$; $q = \overline{1, Q_1}$; $Q_1 < Q$, определяют состав исходных данных для прогнозирования. Возможности n -го ТС, $n \in N^3$, выполнять задачу $l \in [1, L]$ определяет величина α_n^1 . Принципиально важно, чтобы единицы измерения α_n^1 в пределах одной задачи l были одинаковы для всех $n \in N^3$.

Величины n_{1m}^3 должны быть не менее 15, $l = \overline{1, L}$, $m = \overline{1, M}$. Точное значение потребного числа реализаций ТС из ретроспективы можно определить при тщательном исследовании заданного ряда ТС N^3 , степени взаимосвязи и взаимообусловленности характеристик ТС заданного ряда.

Стоимость НИОКР, серийного производства и эксплуатации ТС заданного ряда должны быть выражены в ценах одного года (инфляция учитывается на этапе планирования [1]).

Значения характеристик X_{qn} , $q = \overline{1, Q}$, $n \in N^3$, в ряде случаев могут восстанавливаться при их пропуске, но это не должно носить систематический характер.

Для настройки процедуры прогноза на выбранную предметную область важно указать при подготовке исходных данных предельно допустимые значения всех указанных характеристик на период T_k . Эти данные должны быть выработаны при групповом экспертном опросе. Для генерации сетки изменений от \min к \max по каждой из варьируемых характеристик необходимо указать шаг изменения этих характеристик. Для процедуры прогнозирования нужно определять примерное потребное число прогнозных вариантов ТС, т.е. длину прогнозного ряда N^n . Наконец, пользователь перед прогнозом должен определить число Q_1 и перечень априори заданных характеристик X_{qn} , $q = \overline{1, Q_1}$, $n \in N^n$.

При подготовке исходных данных целесообразно использовать ТС, уже реально применяемые и планируемые к применению в выбранной предметной области, а также рассматривать ТС, применяемые в других отраслях народного хозяйства и выпускаемые за рубежом. Экономический анализ ситуации при планировании все расставит на свои места. Важно, чтобы выбираемые ТС имели свой набор исходных данных, начинающихся с α_n^1 . Процедура прогноза будет работать тем надежнее и эффективнее, чем точнее, и, главное, реальнее будут исходные данные.

7. Сбор, анализ и систематизация исходных данных для процедуры прогнозирования

Методические рекомендации. Фундаментальным вопросом при подготовке исходных данных для прогнозирования является правильный выбор интегральной характеристики ТС α_n^1 , $1 = \overline{1, L}$, и определение ее значений для всех $n \in N^3$.

Общая рекомендация при подготовке исходных данных одна - они должны максимально отражать истинные значения параметров и характеристик исследуемых ТС. Специфика разработки и настройки технологии прогнозирования такова, что на этапе создания ее делают адекватной исходным данным. По сути идет процесс обуче-

ния технологии именно на исходных данных. Другими словами, отклонение истинных данных от подготовленных для процедуры разработки неминусею сказывается на точности прогноза.

Опыт исследований в направлении создания интеллектуальных систем прогнозирования показывает, что подготовка исходных данных - наиболее ответственный и достаточно трудоемкий этап. Вместе с тем технология прогнозирования предусматривает подключение заказчика, выдающего исходные данные, дважды: первый раз - при выдаче исходных данных, второй раз - при анализе результатов прогноза и их экспертном редактировании.

Из множества потенциально измеримых характеристик ТС заданного ряда целесообразно брать в состав исходных данных лишь те, которые, во-первых, с течением времени (например, в развитии m_1 -го типа ТС, $l = \overline{1, L}$, $m_1 = \overline{1, M_1}$) изменяются (постоянные характеристики не должны входить в множество характеристик $q \in [1, Q]$), во-вторых, те характеристики, которые приоритетно влияют на основные характеристики ТС.

8. Подготовка баз данных к процедуре прогнозирования

Опыт показывает, что даже в тщательно отобранных исходных данных осталась масса пробелов. Основная задача этапа 8 технологии прогнозирования перспектив развития ТС состоит в том, чтобы так модифицировать и отредактировать исходные базы данных, чтобы в них были заполнены максимально точно все пробелы.

Можно считать, что базы данных тогда готовы к процедуре прогнозирования, когда по ним можно построчно заполнять пробелы (прогнозировать) в значениях прогнозируемых характеристик с допустимой ожидаемой погрешностью (например, меньше 15-20%).

Процесс подготовки баз данных носит итерационный характер и включает в себя таксономическое исследование баз данных. В конечном счете важно узнать: какие столбцы и строки из баз данных можно безболезненно выбросить, чтобы по остатку (обучающей

выборке) процесс прогнозирования проводился с приемлемой погрешностью. Важным приемом (средством) "лечения" баз данных является отыскание в "чужих" базах данных аналогов для базы данных, по которой осуществляется прогноз.

Таким образом, процесс подготовки баз данных к процедуре прогнозирования в конечном счете сводится, во-первых, к устранению в базах данных части "своих" строк и столбцов, во-вторых, к добавлению в базы данных части "чужих" строк (переход на псевдореальную классификацию), в-третьих, к заполнению всех пробелов в базах данных перед осуществлением процедуры прогноза. Ясно, что всем указанным операциям должно предшествовать изучение влияния столбцов друг на друга, поиск строк-аналогов, определение приоритетного ряда характеристик под каждый изучаемый столбец.

Указанные выше операции выполнялись итерационным способом автоматизированно с помощью программных средства SIM, ACQUIS, ZET [2]. Программа SIM позволяет осуществлять поиск объектов-аналогов для каждого ТС базы данных. Программа ACQUIS позволяет выявлять множество признаков, по которым можно различать классы (надежно классифицировать в заданном признаковом пространстве). С помощью программы ZET осуществляется восстановление (прогнозирование) значений в "дырке" и редактирование баз данных.

9. Сценарий проведения эксперимента

Излагаемая в статье технология прогнозирования универсальна в том смысле, что подготовка ответа на любые возникающие при прогнозных исследованиях вопросы не потребует изменения технологии, при этом лишь по-иному нужно собирать и компоновать исходные данные для прогнозирования. Это же утверждение касается применения экспериментального программного обеспече -

ния, инструментально поддерживающего автоматизированный режим выполнения этапов 11-13 технологии.

Другими словами, под возникающие при прогнозных исследованиях вопросы всякий раз необходимо разрабатывать свой сценарий эксперимента, определяющий набор входных и выходных данных и алгоритм применения экспериментального программного обеспечения.

Набор входных данных эксперимента делится на три группы: прогнозируемые, варьируемые и постоянные величины. Подчеркнем: в каждом эксперименте перечень прогнозируемых характеристик, перечень, диапазон и шаг изменения варьируемых характеристик, перечень и значения постоянных характеристик уникальны и определяются задачами эксперимента.

10. Подготовка входных файлов данных

Выполнение этапов 10-12 технологии прогнозирования в настоящее время осуществляется автоматизированно с помощью экспериментального программного обеспечения, созданного на основе программы ZET [2].

По мере совершенствования процедур выполнения этапов технологии прогнозирования вплоть до их полной автоматизации, предусматривающей участие пользователя в ключевых фазах выполнения этих этапов, перечень параметров, доступных для вариации пользователем, будет расширяться. Это означает, что и подготовка входных файлов усложнится, станет более трудоемкой.

11. Реализация сценария прогноза

В настоящее время для двух предметных областей на базе программ ZET [2] в Международной лаборатории СИНТЕЛ при участии автора данной статьи создано и успешно апробировано две версии программного обеспечения для экспериментального прогнозирования перспектив развития ТС.

При создании экспериментального программного обеспечения реализовано два разных подхода. В первой редакции обеспечен режим генерации заданного числа прогнозных вариантов, при этом задаются диапазоны и шаг изменения варьируемых характеристик. Процедура формирования прогнозных вариантов включает два этапа. На первом этапе по заданным диапазонам интервалов вариации формируется сетка возможных значений переменных. Затем осуществляется генерация возможного числа прогнозных вариантов. Далее, на втором этапе, с помощью интеллектуальных фильтров из сгенерированного множества выбирается та часть прогнозных вариантов, которая отвечает системе правил, заложенных в базу знаний.

Если в первой версии надежность прогнозирования в большей степени обеспечивалась за счет использования интеллектуальных фильтров, то во второй версии экспериментального программного обеспечения эта задача решается на этапе подготовки баз данных к процедуре прогнозирования. Во второй версии пользователь опосредованно выбором набора варьируемых характеристик и множества их значений с неравномерным шагом изменения задает в конечном счете требуемое число прогнозных вариантов.

Экспериментальное программное обеспечение требует применения ПЭВМ не ниже АТ-286 с сопроцессором со стандартной памятью в среде ДОС не ниже версии 3.3. Таблицы, предъявляемые для обработки, по объему не должны превышать 15000 знаков.

12. Анализ и редактирование баз данных результатов прогноза

Результаты выполнения экспериментального прогнозирования ТС помещаются в выходной файл данных.

Заключительный этап целесообразнее всего проводить специалистам-экспертам исследуемой предметной области, т.е. пользователям и заказчикам прогноза.

Редактирование баз данных результатов прогноза сводится к тому, что пользователь по своему усмотрению может устранять из баз данных отдельные строки, вставлять новые строки (прогноз эксперта), изменять отдельные элементы строки и столбцов.

Анализ результатов прогноза должен дать возможность ответить на вопрос о том, какой набор варьируемых характеристик ТС дает наиболее близкий к истине результат.

Важной целью проведения анализа и редактирования результатов прогнозирования перспектив развития ТС является осуществление направленного отбора приемлемых прогнозных вариантов ТС для формирования исходного ряда для планирования, в который наряду с заданными (известными) образцами ТС могут входить и прогнозные [1].

13. Завершение процедуры прогнозирования

В общем случае процесс прогнозирования носит итерационный характер. Так после анализа полученных результатов прогноза пользователь может принять решение о возобновлении работы по прогнозированию с начала, с середины, либо с предыдущего этапа. Другими словами, можно менять ожидаемую ошибку прогнозирования кардинально, путем изменения признакового пространства, либо дополнением базы данных новыми данными. На этом этапе можно либо принять решение о применении другого метода прогнозирования, либо его специальной разработке.

Целесообразно создавать архив данных прогнозирования, записывая в него положительные результаты прогнозирования (входные и выходные файлы варианта экспериментального исследования).

З а к л ю ч е н и е

Изложенная технология прогнозирования развития технических средств и систем на длительную перспективу охватывает весь спектр вопросов от выбора и постановки задачи на прогно-

зирование до редактирования результатов прогноза и их применения, например, при планировании развития этих средств и систем. Технология успешно апробирована в двух принципиально разных предметных областях. Центральное место в технологии занимают этапы подготовки входных файлов, реализации сценария прогноза, анализа и редактирования баз данных результатов прогноза. Созданные две версии экспериментального программного обеспечения инструментально поддерживают автоматизированный режим выполнения этих этапов. Выполнение остальных этапов технологии осуществляется вручную.

В настоящее время ведется работа над созданием инструментальных средств, позволяющих осуществлять в автоматизированном режиме сбор, анализ и систематизацию исходных данных, подготовку баз данных к процедуре прогнозирования, разработку сценария проведения эксперимента, завершение процедуры прогнозирования.

Например, программа ACQUIS, применяемая при подготовке баз данных к процедуре прогнозирования (выбор компетентных и выявление "лишних" столбцов в базе данных), используется на этапе разработки сценария проведения эксперимента при выборе целесообразного набора варьируемых характеристик, наиболее существенно влияющих на прогнозируемые характеристики.

Принципиально первые восемь (из тринадцати) этапов технологии (от анализа предметной области до подготовки баз данных к процедуре прогнозирования) выполняет разработчик инструментальных средств, используемых пользователями при выполнении последующих пяти этапов технологии (от разработки сценария проведения эксперимента до завершения процедуры прогнозирования).

Изложенная технология допускает различный уровень интеллектуализации. Одна из версий экспериментального программного обеспечения, используемого для автоматизированного прогнозирования перспектив развития технических средств и систем, включа-

ла программные средства, реализующие интеллектуальные фильтры данных, построенных на базах знаний, извлеченных, в частности, из баз данных. Вместе с тем возможна такая конфигурация инструментальных средств, поддерживающая автоматизированный режим прогнозирования, при которой наряду с созданными версиями экспериментального программного обеспечения как функциональный резерв будут использованы инструментальные средства, позволяющие решать задачу прогнозирования на основе использования только базы знаний [3]. В этом случае потребуются специальные инструментальные средства, интегрирующие результаты прогнозирования, выполненные разными методами.

Среди важных преимуществ излагаемой технологии является то, что она, с одной стороны, использует все преимущества ZET-метода [2] (локальный и индивидуальный подход при заполнении "дырок" не зависит от размерности баз данных и более точен) и, с другой стороны, имеет регулируемую и измеримую надежность прогнозирования.

Технология поддерживает решение любых вопросов, возникающих как при подготовке и проведении прогнозирования, так и при применении результатов прогноза. Появление новых аспектов будет отражаться на этапах разработки сценария эксперимента, подготовки исходных данных и не потребует вносить изменения в созданное экспериментальное программное обеспечение. Изложенная технология наряду с решением задач прогнозирования перспектив развития технических средств и систем может найти применение при прогнозировании развития любых процессов, допускающих описание, изложенное в статье. Например, технологию можно использовать при прогнозе развития ситуаций в сфере управления, минимизации измерений и объемов обрабатываемой информации, при разработке методов ускорения вычислений в системах управления реального времени, при решении задач технической диагностики сложных систем и комплексов, в сфере обучения (при подготовке

и тестировании обучаемых), при совершенствовании существующих и создании новых методов и средств радиолокации.

Автор искренне признателен Н.Г.Загоруйко, В.Н. Ёлкиной, Т.П.Киприяновой, Е.Н.Шемакиной и Н.В.Ефановой за плодотворное участие в обсуждении излагаемой технологии, проведении экспериментального прогнозирования перспектив развития технических средств и систем в двух предметных областях.

Л и т е р а т у р а

1. УСТЮГОВ Ю.А. Технология долгосрочного планирования НИОКР, серийного производства и эксплуатации технических систем на основе прогнозирования их развития с учетом ресурсных ограничений и меняющихся во времени требований //Настоящий сборник. - С.45-70.

2. ЁЛКИНА В.Н., ЗАГОРУЙКО Н.Г. Блок анализа данных в экспериментальных системах //Экспертные системы и анализ данных. - Новосибирск, 1991. - Вып. 144: Вычислительные системы. - С. 54-175.

3. ЗАГОРУЙКО Н.Г., БУШУЕВ С.Н., ГРИГОРЬЕВ В.В., УСТЮГОВ Ю.А. Система прогнозирования основанная на методах получения и обработки знаний //Настоящий сборник. - С.

Поступила в ред.-изд.отд.

18 октября 1994 года