

УДК 519.95

СИСТЕМА ПРОГНОЗИРОВАНИЯ, ОСНОВАННАЯ НА МЕТОДАХ  
ПОЛУЧЕНИЯ И ОБРАБОТКИ ЗНАНИЙ

Н.Г.Загоруйко, С.Н.Бушуев, В.В.Григорьев, Ю.А.Устюгов

В в е д е н и е

Повышение точности прогнозирования развития технических средств и систем (ТСС) продолжает оставаться актуальной научно-технической проблемой, сложность решения которой определяется рядом особенностей:

- развитие ТСС сопровождается их постоянным усложнением и удорожанием;

- современные образцы ТСС обладают большим количеством характеристик, как правило, являющихся взаимосвязанными;

- развитие ТСС представляет собой управляемый процесс, целью которого являются своевременное обоснование рациональных значений характеристик образцов ТСС и их реализация в новых образцах;

- развитию ТСС принципиально присущи скачки (переход на новые физические принципы действия, новую элементную базу и т.д.), которые приводят к созданию более совершенных образцов;

- развитие современных ТСС характеризуется быстрым моральным старением образцов и, как следствие, необходимостью их своевременного обновления;

- имеет место существенное различие технических и стоимостных характеристик образцов ТСС одного целевого назначения, а в ряде случаев - и физических принципов их действия.

Отмеченные особенности существенно затрудняют решение задачи повышения точности прогнозирования развития ТСС и определяют необходимость совершенствования существующих методов и разработки новых подходов к прогнозированию.

Задача прогнозирования заключается в определении характеристик объекта прогнозирования в условиях заданной внешней среды (или наоборот) и решается процедурой, содержащей, как правило, следующие этапы:

а) описание предметной области (рассматриваемых характеристик объекта прогнозирования и окружающей среды) и фактов (конкретных значений характеристик);

б) обнаружение закономерностей на множестве фактов, описывающих ситуацию;

в) предсказание новых фактов на основе обнаруженных закономерностей;

г) оценка ошибок прогноза.

Основой для прогнозирования является информация о развитии объекта прогнозирования в прошлом и настоящем. Иными словами процесс прогнозирования можно представить, как процесс постепенного накопления знаний (фактов и закономерностей) об объекте прогнозирования и получения перечня наиболее вероятных ситуаций, которые могут сложиться в будущем, с помощью гипотез, выдвигаемых на основе имеющихся знаний.

## 1. Постановка задачи прогнозирования характеристик ТСС

Рассмотрим множество объектов  $A = \{a_1, a_2, \dots, a_n\}$ , представляющее собой множество ТСС определенного назначения и типа (например, множества различных по уровню развития средств связи, вычислительных средств и др.).

Каждому объекту  $a_i \in A$ ,  $i = \overline{1, n}$ , можно поставить в соответствие вектор характеристик, отражающих его важнейшие свойства  $\bar{a}_i = \langle x_{i1}, x_{i2}, \dots, x_{im} \rangle$ , где  $x_{ij} \in X_j$ ,  $x_{ij}$  - значение характери -

стики  $j$ -го типа для объекта  $a_i$ ,  $X_j$  - множество значений характеристики  $j$ -го типа,  $i = \overline{1, n}$ ,  $j = \overline{1, m}$ . Пусть  $\bar{A}$  - множество указанных векторов, соответствующее множеству  $A$ . Причем для каждого  $X_j$ ,  $j = \overline{1, m}$ , определен тип шкалы, в которой измеряется данная характеристика (шкала интервалов, шкала отношений, шкала порядка, абсолютная) [1].

Обозначим через  $X = X_1 \times X_2 \times \dots \times X_m$  множество всех наборов значений характеристик, или, по-иному, пространство возможных ситуаций предметной области. Элементами множества являются вектора  $\langle x_1, x_2, \dots, x_m \rangle$ , содержащие фиксированные значения выделенных характеристик, причем  $\bar{A} \subset X$ .

Состав выделенных объектов и существенных свойств зависит от вида объекта прогнозирования и окружающей среды, а также от решаемых пользователем задач. В частности, выбор множества  $A$  предполагает выделение характеристик объектов, значения которых отличают объекты этого множества от других по функциональному назначению (оперативные, надежностные и др.), по внешним параметрам (габариты, вес и др.) и другим свойствам (конструктивные особенности, интегральные по совокупности свойств, отражающие организационную структуру, состав элементной базы и др.).

Цель исследования (прогнозирование развития объектов типа  $A$ ) обязывает работать с временными характеристиками, отражающими определенные моменты жизненного цикла объекта (разработка, производство и эксплуатация), и стоимостными, отражающими затраты на этапы жизненного цикла. Таким образом, множество выделенных характеристик можно разбить по смысловому содержанию на несколько непересекающихся подмножеств: функциональные, технические, конструктивные, организационные и стоимостные характеристики.

Зададим временной интервал  $[T_H, T_K]$ , где  $T_H$  - время начала жизненного цикла первого объекта множества  $A$  (предполагается, что объекты множества  $A$  упорядочены по времени начала жизненного цикла),  $T_K$  - момент времени в будущем, для которого делается прогноз. Обозначим через  $T_0$  текущий момент времени развития объектов типа  $A$ . Таким образом,  $[T_H, T_0]$  - известный период развития (интервал наблюдения),  $[T_0, T_K]$  - прогнозируемый период развития (интервал упреждения).

Предполагается, что на момент времени  $T_0$  для объектов  $a_i \in A$ ,  $i = \overline{1, n}$ , в основном известны значения характеристик  $x_{ij}$ ,  $j = \overline{1, m}$ . То есть, известна таблица  $\{x_{ij}\}$ ,  $i = \overline{1, n}$ ,  $j = \overline{1, m}$ , называемая обучающей таблицей, которая отражает ретроспективный (прошедший) период развития объектов типа  $A$ .

Необходимо по частично заданному облику перспективного объекта  $a_{n+1}$  (на основе предполагаемых новых требований окружающей среды априори заданы значения некоторых характеристик, например, временных, стоимостных, части функциональных  $x_{n+1,1}, \dots, x_{n+1,s}$ ,  $s < m$ ) спрогнозировать  $(m-s)$  незаданных характеристик. Заданные значения характеристик перспективного объекта  $a_{n+1}$ , как правило, определяются экспертным путем.

В результате решения задачи прогнозирования должен быть определен целиком вектор  $\bar{a}_{n+1} = \langle x_{n+1,1}, \dots, x_{n+1,m} \rangle$ , описывающий перспективный объект  $a_{n+1}$ . Причем, в зависимости от целей прогнозирования заказчик - пользователь системы прогнозирования должен иметь возможность варьировать заданные и прогнозируемые наборы характеристик.

## 2. Определение подхода к решению задачи прогнозирования

В настоящее время известны более 100 способов и методов, относящихся ко всему комплексу проблем прогнозирования научно-технического развития. Обзоры существующих методов приведе-

ны в [2-4 и др.]. При этом трудно провести резкое различие между их качественными особенностями.

Основным требованием к методам прогнозирования является повышение обоснованности прогнозов, что в первую очередь предполагает повышение их точности. Это вызвано большой ценой ошибок прогнозирования, которые могут привести к принятию неоптимальных решений, и, как следствие, к потерям ресурсов и времени.

Каждый класс методов имеет свои особенности и ограничения и направлен на решение определенных задач прогнозирования. Выбор конкретного метода зависит от специфики объекта прогнозирования и окружающей среды, а также потребностей пользователя и его доверия к способам обнаружения и обработки закономерностей. Сравнительный анализ возможностей классов методов с точки зрения прогнозирования развития ТСС показывает, что каждый из них обладает рядом достоинств и недостатков. Отметим основные из них.

Прогноз, осуществляемый на основе моделирования, предполагает построение математической, а иногда и физической модели объекта прогнозирования, которая может, благодаря сходству с объектом прогнозирования, имитировать оригинал. Для процесса развития ТСС создание подобных моделей является трудной задачей, так как сложно представить существо и внутренний механизм прогнозируемого процесса.

Среди методов прогнозирования, основанных на моделировании, при решении поставленной задачи отдадим предпочтение статистическому моделированию на основе данных о ходе прошедшего развития, т.е. статистическому прогнозированию. Заметим однако, что прогноз, получаемый с помощью подобной модели, основывается на серьезном допущении: прогнозируемый процесс будет развиваться по тем же законам, по которым он развивался в прошлом и развивается в настоящем (явно не учитывается возможность

появления скачков развития в будущем). С другой стороны, способы статистического прогнозирования обладают высокой степенью точности в случае выполнения вышеуказанного условия и удобны для реализации на ЭВМ.

Способы эвристического прогнозирования достаточно сложны для реализации на вычислительной технике ввиду того, что процесс получения знаний от экспертов с трудом поддается формализации, хотя последние достижения в инженерии знаний [5-7] и практической разработке экспертных систем [8-10] значительно расширили эти возможности. Кроме того, при эвристическом прогнозировании суждения и предположения каждого эксперта являются субъективными. Но в то же время несомненное достоинство указанных методов - возможность предвидения скачкообразных изменений прогнозируемой характеристики.

Анализ отмеченных достоинств и недостатков показывает, что способы статистического и эвристического прогнозирования дополняют друг друга. Естественнo предположить, что способ, сочетающий в себе оба этих подхода, будет иметь меньше недостатков и давать более точный прогноз, но очевидно, что при этом трудоемкость и сроки получения прогноза увеличатся. В известной степени компенсировать это можно повышением степени автоматизации процесса прогнозирования за счет использования последних достижений в области новых информационных технологий.

Поставленная задача прогнозирования характеризуется малым объемом обучающей выборки и большой размерностью пространства признаков, так как предыстория развития ТСС обычно незначительна, а сложность образцов ТСС, как изделий на основе последних достижений науки и техники, велика. Кроме того, указанная задача имеет ряд особенностей:

а) вид и состав выделенных характеристик и обучающей выборки могут варьироваться от мнений экспертов и потребностей конкретного пользователя;

б) отдельные значения характеристик у некоторых объектов из обучающей выборки могут отсутствовать (по разным причинам - неизвестны, неточны, не могут быть получены, не имеют смысла);

в) характеристики могут измеряться в разных шкалах (абсолютная, порядка, отношений и др.);

г) помимо сведений, содержащихся в обучающей выборке, при прогнозе необходимо учитывать также экспертные суждения и оценки (например, значимость характеристик, зависимости между ними, исходя из их смыслового содержания, предельные значения характеристик и др.);

д) результат, а по возможности, и процесс его получения должны быть объяснены пользователю на понятном ему языке для повышения степени доверия к прогнозу.

Существующие методы (например, алгоритм заполнения пробелов ZET [11], методы предсказания разнотипных признаков в классе линейно-логических функций [12]), основанные на поиске эмпирических закономерностей в обучающей выборке и использовании для прогнозирования способа рассуждения по аналогии, позволяют в основном решить поставленную задачу. Но они не используют реальный смысл характеристик и взаимосвязей между ними, не могут учесть отмеченные ранее особенности "г" и "д" поставленной задачи прогнозирования и дать прогноз скачкообразного изменения исследуемой характеристики.

С другой стороны, особенности поставленной задачи прогнозирования в значительной мере соответствуют ключевым проблемам задачи предсказания в классификации экспертных задач [5]. Адекватным и удобным инструментом решения задачи автоматизации процесса прогнозирования с учетом всех ее особенностей может быть специализированная экспертная система прогнозирования, осуществляющая прогноз на основе знаний о закономерностях ретроспективного (прошлого) развития ТСС и суждениях (оценках) экспертов об их возможном развитии в будущем. По своим параметрам

она должна быть близка к партнерской системе, основные признаки которой излагаются в [13].

Предлагаемый подход к решению задачи автоматизации процесса прогнозирования характеристик перспективных ТСС, который по сути является процессом получения и обработки знаний об их развитии, предполагает:

а) определение состава и модели представления знаний, необходимых для осуществления прогноза;

б) разработку методики автоматизированного получения знаний, включающей процедуры:

- приобретения первоначальных знаний о предметной области (множество объектов определенного класса; множество характеристик, отражающих основные свойства этих объектов; диапазоны возможных значений характеристик, присущих данному классу; значимость характеристик; зависимости характеристик друг от друга);

- формирования знаний из данных, содержащих сведения о ретроспективном развитии ТСС, т.е. закономерностей изменения одних характеристик в зависимости от изменения других;

- получения оценок (суждений) экспертов о возможных исходах будущего развития ТСС;

в) разработку технологии обработки знаний, реализующую:

- логический вывод прогноза из закономерностей ретроспективного развития;

- интеграцию оценок экспертов о значениях характеристик перспективного образца прогнозируемого класса ТСС;

- согласование прогнозов, полученных из данных ретроспективного развития и от экспертов;

г) разработку структуры и алгоритма функционирования программного комплекса, обеспечивающего получение и обработку знаний с целью анализа и прогнозирования характеристик перспективных ТСС и предложений по его применению.



### 3. Структура и возможности системы прогнозирования

Разрабатываемый программный комплекс, реализующий предложенный подход анализа и обработки знаний о развитии ТСС с целью прогнозирования характеристик перспективных образцов, в соответствии со своим функциональным назначением должен обеспечивать:

- а) ввод и хранение данных о ретроспективном развитии ТСС, а также знаний (суждений и оценок) экспертов об их будущем развитии;
- б) автоматизированное извлечение знаний из данных;
- в) проверку непротиворечивости знаний и данных;
- г) прогнозирование облика (значений характеристик) перспективного образца ТСС на основе имеющихся знаний.

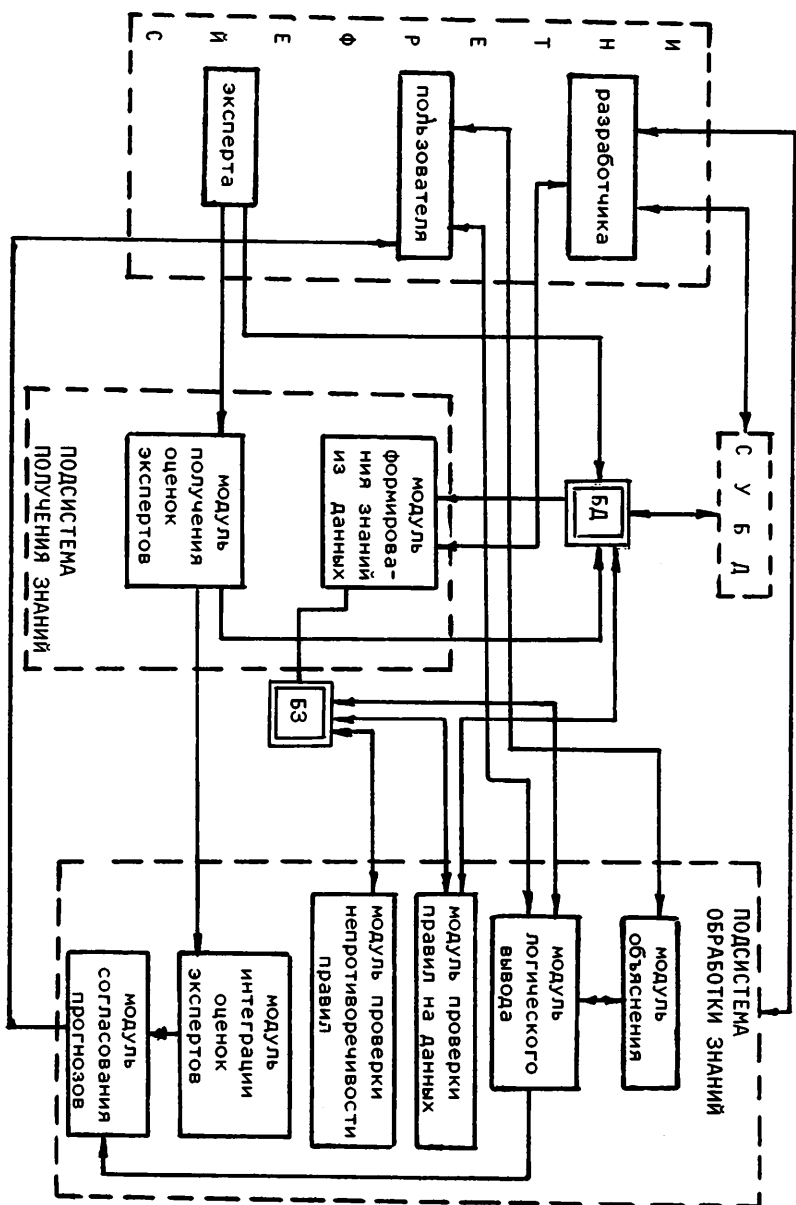
Кроме того, естественно, он должен иметь возможности для выполнения сервисных и обслуживающих функций таких, как:

- а) получение информации (визуальной и твердой копии) о содержимом базы данных (БД) и базы знаний (БЗ);
- б) копирование и восстановление БЗ и БД;
- в) получение объяснения полученного прогноза и др.

На рисунке приведена структурная схема разрабатываемой системы прогнозирования, реализующей вышеперечисленные функции.

Основными компонентами такой системы являются:

- интерфейс, позволяющий общаться с системой трем типам специалистов: разработчику системы, эксперту и пользователю (включает в себя модуль интерактивного опроса экспертов);
- БД, содержащая данные о ретроспективном развитии и оценки экспертов по характеристикам перспективных образцов ТСС, и соответствующая СУБД, реализующая программный доступ к ней;
- БЗ, предназначенная для хранения правил производственного вида, отражающих закономерности ретроспективного развития ТСС;
- модуль автоматизированного формирования знаний из данных;



Общая схема системы прогнозирования

- модуль логического вывода (получения гипотез), позволяющий на основе закономерностей ретроспективного развития осуществлять прогнозирование характеристик перспективных ТСС;
- модуль согласования правил БЗ с новыми данными;
- модуль обработки (интеграции) оценок экспертов;
- модуль проверки непротиворечивости знаний;
- модуль согласования прогнозов, полученных разными способами.

Работа с системой прогнозирования предполагает выполнение следующих основных этапов:

а) настройка системы прогнозирования на конкретную предметную область: определение объекта прогнозирования - класса ТСС, процесс развития которых будет анализироваться и прогнозироваться; выделение существенных характеристик, присущих указанным системам; определение множеств возможных значений выделенных характеристик; ввод данных о ретроспективном развитии объекта прогнозирования; получение и обработка (интеграция с целью получения единого согласованного мнения) экспертных суждений о выделенных характеристиках, т.е. установление взаимовлияния характеристик и их значимости относительно основного функционального назначения; получение и обработка экспертных оценок о значениях характеристик перспективных ТСС; создание БЗ, т.е. формирование правил БЗ, отражающих закономерности ретроспективного развития ТСС;

б) использование системы с целью получения прогноза: задание в режиме диалога собственных потребностей и предположений по значениям характеристик, которые не подлежат прогнозу, но могут быть использованы при определении значений прогнозируемых характеристик; задание перечня прогнозируемых характеристик; запуск процедуры обработки знаний и получение варианта прогноза значений характеристик перспективного образца ТСС; получение при необходимости объяснения полученного прогноза и

анализ его; повторение процедуры с начала с учетом полученных результатов, если результаты прогноза и его объяснение не удовлетворяют пользователя;

в) обслуживание системы, т.е. поддержание ее в работоспособном и актуальном состоянии: получение информации о содержимом БД и БЗ (на экране и в виде твердой копии); создание резервных копий и восстановление БД и БЗ; своевременная корректировка БД и БЗ в случае появления новых или уточнения имеющихся данных (фактов) и экспертных оценок о выделенных характеристиках; проверка непротиворечивости правил БЗ. (Действия по обслуживанию системы выполняются в процессе ее настройки и использования по мере необходимости.)

Реализация представленной системы прогнозирования выполняется на ПЭВМ, совместимой с IBM PC, в операционной среде MS DOS. Для реализации используются (после соответствующей адаптации) программы инструментальной системы ЭКСНА [14,15], разработанной в ИМ СО РАН совместно с международной лабораторией "Синтел". Для проверки возможности реализации представленного подхода к прогнозированию характеристик перспективных ТСС и предварительной оценки его эффективности выполнена разработка программного обеспечения, поддерживающего основные этапы получения и обработки знаний, основанные на подходе, изложенном в [16].

Методика получения, технология обработки знаний и поддерживающие их программные средства прошли предварительную экспериментальную проверку. В качестве ретроспективных данных были использованы данные о развитии двух типов средств управления. Размерности обучающих таблиц составили 8x9 и 9x15 соответственно.

В ходе проверки было выполнено два цикла получения и обработки знаний. При этом:

а) в автоматическом режиме сформирована БЗ из данных ретроспективного развития;

б) получены и согласованы мнения 10 экспертов по выделенным характеристикам;

в) проведена инверсная верификация методики прогнозирования, основанной на логическом выводе из закономерностей ретроспективного развития;

г) рассчитана средняя относительная ошибка прогнозов, полученных из закономерностей ретроспективного развития анализируемых средств управления (она составила 7,9% и 13,6% для первой и второй таблиц соответственно);

д) сформирован комбинированный прогноз путем согласования интегральной оценки экспертов с прогнозом, полученным из закономерностей ретроспективного развития; доверительный интервал комбинированного прогноза уменьшился в среднем на 59% относительно доверительных интервалов комбинируемых прогнозов, что по существу [17] означает повышение точности прогнозирования не менее чем в два раза.

Проверка показала правильность выбранного пути повышения точности и обоснованности прогнозов и техническую реализуемость предложенного подхода, целесообразность использования его именно для прогнозирования развития сложных технических систем, имеющих малую предысторию развития и большое число взаимосвязанных характеристик.

Вместе с тем, техническая реализация разрабатываемой системы прогнозирования, основанной на описанном подходе, требует решения ряда вопросов, связанных с оптимальной организацией БЗ и обеспечением приемлемых для пользователя временных характеристик системы по получению и обработке знаний.

Авторы выражают признательность Ёлкиной В.Н. за участие в обсуждении проблем разрабатываемого подхода к решению задачи

прогнозирования и коллективу Международной лаборатории "Синтел" за помощь в проведении экспериментальной проверки его реализуемости.

### Л и т е р а т у р а

1. ПФАНЦАГЛЬ И. Теория измерений. - М.: Мир, 1976.
2. ЯМПОЛЬСКИЙ С.М., ЛИСИЧКИН В.А. Прогнозирование научно-технического прогресса. - М.: Экономика, 1974.
3. ДОБРОВ Г.М. Прогнозирование в науке и технике. - М.: Наука, 1969.
4. ЗАГОРУЙКО Н.Г. Классификация задач прогнозирования на таблицах "объект-свойство" //Машинные методы обнаружения закономерностей. - Новосибирск, 1981. - Вып. 88: Вычислительные системы. - С. 3-8.
5. ХЕЙЕС-РОТ Ф., УОТЕРМАН Д., ЛЕНАТ Д. Построение экспертных систем. - М.: Мир, 1987.
6. ГАВРИЛОВА Т.А., ЧЕРВИНСКАЯ К.Р. Извлечение и структурирование знаний для экспертных систем. - М.: Радио и связь, 1992.
7. Приобретение знаний /Под ред. С.Осуги, Ю.Сазэки. - М.: Мир, 1990.
8. Экспертные системы. Принципы работы и примеры /Под ред. Р.Форсайта. - М.: Радио и связь, 1987.
9. Искусственный интеллект: В 3-х кн. Кн.1 Системы общения и экспертные системы: Справочник /Под ред. Э.В.Попова - М.: Радио и связь, 1990.
10. Материалы III конференции по искусственному интеллекту: КИИ-92. Сборник научных трудов в 2-х томах. - Тверь, 1992.
11. ЗАГОРУЙКО Н.Г., ЕЛКИНА В.Н., ТИМЕРКАЕВ В.С. Алгоритм ZET-75 заполнения пробелов в эмпирических таблицах и его применение //Машинные методы обнаружения закономерностей. -Новосибирск: Наука, 1976. - С. 57-63.
12. ЛБОВ Г.С. Логические функции в задачах эмпирического предсказания //Эмпирическое предсказание и распознавание образов. - Новосибирск, 1978. - Вып. 76: Вычислительные системы. - С. 34-64.
13. ЗАГОРУЙКО Н.Г. Партнерские системы //Анализ данных и знаний в экспертных системах. - Новосибирск, 1990. - Вып. 134: Вычислительные системы. - С. 3-18.

14. ЕЛКИНА В.Н., ЗАГОРУЙКО Н.Г. Блок анализа данных в экспертной системе ЭКСНА. //Экспертные системы и анализ данных. - Новосибирск, 1991. - Вып. 144: Вычислительные системы. - С. 54-175.

15. БУШУЕВ М.В., ЕЛКИНА В.Н., ЗАГОРУЙКО Н.Г., ШЕЛЯКИНА Е.Н. Блок анализа знаний в инструментальной экспертной системе ЭКСНА //Методы и системы искусственного интеллекта.-Новосибирск,1992. - Вып. 145: Вычислительные системы. - С. 29-78.

16. БУШУЕВ С.Н., ГРИГОРЬЕВ В.В. Комплексная методика получения знаний и их обработка с целью прогнозирования развития. - Настоящий сборник. - С.32-44.

17. ЧУЕВ Ю.В., МИХАЙЛОВ Ю.Б., КУЗЬМИН В.И. Прогнозирование количественных характеристик процессов. - М.: Сов.радио, 1975.

Поступила в ред.-изд.отд.

11 мая 1994 года