

СХЕМА ПРЕОБРАЗОВАНИЯ ПОКАЗАТЕЛЕЙ В ИСОД

А. Д. БАРСКИЙ

(Москва)

Методологической основой комплекса программ и инструкций по обработке данных в ИСОД является детальная схема преобразования показателей, построение которой затруднено отсутствием единого их списка. Несмотря на кажущуюся простоту, подготовка такого списка требует труда высококвалифицированных специалистов. Особую сложность при идентификации показателей представляет наделение их всеми необходимыми признаками: определениями, обстоятельствами, дополнениями и т. п., которые зачастую бывают заложены в шапке документа либо неявно зафиксированы временем, местом или условиями появления документа. Поэтому все способы идентификации показателей на основе лингвистического анализа их полных наименований требуют тщательной проверки при занесении их в список. Пользование таким списком при шифровке вызывает также ряд трудностей, связанных с несопоставимостью полных наименований показателей списка и наименований в документе. Теоретически возможны два пути подготовки исходных данных для построения схемы преобразования показателей, причем и в том и в другом случае первым этапом работы является построение и шифровка единого списка показателей: а) описание алгоритмов формирования каждого элемента единого списка показателей без привязки к документам; б) описание алгоритмов формирования показателей по каждому документу с использованием шифров, полученных на первом этапе.

Попытка реализации первого пути показала, что практически не всегда возможно поставить в соответствие каждому показателю однозначный алгоритм его формирования на основе материалов обследования, так как зачастую одноименные показатели рассчитываются различными способами в отдельных подразделениях и в различных документах при решении аналогичных планово-экономических задач.

Основная трудность, возникающая при реализации второго пути, — сложность идентификации конкретных показателей каждого из документов с показателями, зашифрованными в едином списке. При этом необходимы многократные уточнения списка из-за часто встречающихся неточностей в формулировке призначной части, что превращает процесс подготовки исходных данных в итерационный (каждое изменение и уточнение наименований показателей в шифраторе требует уточнения шифров уже подготовленных элементов). «Сходимость» процесса подготовки исходных данных в основном зависит от уровня квалификации разработчиков. Кроме этого, отрицательную роль играет неоднозначность алгоритмов формирования показателей, имеющих одинаковые наименования. В результате возникает проблема выбора одного алгоритма-представителя. В противном случае в схеме происходит наложение различных алгоритмов формирования, привязанных к одному и тому же показателю. В связи с необходимостью точного определения алгоритмов формирования каждого показателя при построении алгоритмической сети ИСОД, целесообразно считать алгоритм формирования показателя единственным признаком, по которому можно идентифицировать и различать показатели. При этом вводятся следующие основные определения:

1) показатель — конкретное операционное выражение сущности экономической категории. Каждому наименованию показателя может соответствовать только одна операция формирования и наоборот;

2) операция — совокупность показателей, участвующих в формировании заданного показателя и оператора, преобразующего их. Под оператором формирования показателя понимаются арифметические действия, логические преобразования — сортировка, группировка и т. д., а также элементарные процессы обработки данных, например переписка, регистрация.

Из приведенных определений следует, что если двум показателям, одинаково названным в различных документах, соответствуют различные операции формирования, то в едином списке они должны быть названы по-разному. Иными словами, с точки зрения организации систем обработки данных было бы правильно считать, что каждой операции формирования должен однозначно соответствовать определенный показатель.

В качестве примера можно привести показатель: рост производительности труда, который может рассчитываться двумя способами: а) по факторам роста производительности труда, б) прямым счетом.

На основе операционного определения показателя разработан метод выделения единого списка показателей, а также построения алгоритмической сети системы управления. Построение алгоритмической сети такого рода можно осуществить в три этапа.

1. Упорядочение технологической сети формирования показателей. Упорядоченной называем сеть, элементы которой распределены по уровням формирования в зави-

симости от заданных для каждого элемента отношений подчиненности. Технологическая сеть формирования показателей отражает последовательность выполнения арифметических, логических и технических процедур по преобразованию показателей. Техническими процедурами преобразования показателей называются операции по регистрации и переписыванию показателей. Таким образом, производится разбиение показателей по уровням на основе технологии их формирования. На стадии сбора исходных данных об операциях формирования показателей и при построении технологической сети принимается условное допущение, что в системе информации отсутствует дублирование показателей. Это позволяет кодировать показатели любым способом, обеспечивающим их однозначное опознание. Наиболее удобен составной код, который строится из номера документа и номера показателя внутри документа. В результате принятых допущений упорядоченная технологическая сеть представит достаточно полную модель преобразования показателей в системе и даст возможность выделить исходный для построения алгоритмической сети нулевой уровень формирования показателей, т. е. перечень первичных показателей системы.

II. Неформальный анализ списка первичных показателей. На этом этапе унифицируется полученный в результате построения технологической сети список показателей нулевого уровня, т. е. первичных. Цель унификации — установление смыслового соответствия между показателями нулевого уровня, иными словами, устранение дублирования в списке первичных показателей. В настоящее время невозможно предложить какие-либо эффективные методы формального смыслового анализа первичных показателей, поэтому на данном этапе работ использовался метод неформального анализа показателей.

Большую роль для анализа первичных показателей должны сыграть интенсивно разрабатываемые методы семантико-лингвистического анализа. Однако уже сейчас сам факт ограничения применения неформального анализа уровнем первичных показателей позволяет существенно снизить трудоемкость проведения работ по выделению единого списка показателей и построению алгоритмической сети системы управления.

III. Выделение и упорядочение алгоритмической сети формирования показателей. Алгоритмическая сеть формирования показателей отражает последовательность выполнения математических и логических процедур, т. е. процедур, приводящих к изменению смысла и значения показателей, причем сама технология построения такого рода сети основана на правиле, по которому невозможно наличие в ней хотя бы двух одинаковых (по способу формирования) показателей. Построение алгоритмической сети ведется по уровням преобразования показателей, начиная с нулевого. Необходимы и достаточным условием идентификации двух показателей является совпадение операций их формирования (определение 2)). Проиллюстрируем это примером. Допустим, на первом уровне технологической сети имеются показатели «001» и «002»; заданы операции их формирования: показатель «001» формируется из показателей «003» и «004» оператором f ; показатель «002» — из показателей «005», «006» тем же оператором f . В результате анализа показателей нулевого уровня обнаружено, что показатель «003» идентичен показателю «005», а «004» идентичен «006». Тогда можно записать

$$\langle 003 \rangle \equiv \langle 005 \rangle$$

$$\langle 004 \rangle \equiv \langle 006 \rangle$$

$$f \equiv f.$$

Исходя из определения 1), можно утверждать, что показатель «001» идентичен показателю «002» и им присваивается общий шифр. Показателям же с разными операциями формирования присваиваются различные шифры. Аналогично анализируются показатели второго и каждого следующего уровня вплоть до конечного. На заключительном этапе показатели, получаемые простой перепиской, переименовываются с присвоением им шифров показателей, с которых они воспроизводятся.

Реализация предложенного метода требует некоторых изменений в организации обследования потоков информации. В первую очередь это потребует повышения роли координирующей группы обследования.

Обследование потоков информации ведется по подразделениям органа управления в два этапа. На I этапе укрупненно изучаются потоки информации на уровне документов. Бланки документов передаются в координирующую группу для составления общего списка документов и шифровки. Особое внимание уделяется документам, зафиксированным более чем в одном отделе. Тщательно сверяется форма документов и уточняется перечень показателей, встречающихся в них. Это связано с необходимостью единообразной шифровки показателей в документах, зафиксированных разными группами исследователей в различных подразделениях.

Закачивается I этап подготовкой исходных данных для обследования движения и преобразования показателей. Для этого шифруются реквизиты в собранных

бланках документов. Шифровка реквизитов может производиться децентрализованно для документов только одного отдела. Для документов, зафиксированных в двух и более отделах, шифровка показателей должна производиться централизованно и единообразно во всех копиях бланков для различных групп исследователей.

II этап — изучение и сбор данных о взаимосвязи показателей и процедурах их формирования. Результаты обследования фиксируются в специальных разработочных таблицах с использованием шифров показателей и документов, разработанных на I этапе. Возможны случаи, когда группировочные признаки, необходимые для расчета того или иного показателя, явно не выделены в бланках документов. Тогда признак должен быть записан в ведомость-приложение для последующего уточнения шифра в координирующей группе. Ведомость-приложение для записи дополнительных показателей имеет следующую форму

Номер документа	Условный шифр признака	Наименование признака	Уточненный шифр признака

Макет разработочной таблицы строится в предположении, что каждый производный показатель состоит максимум из двух формирующих. Когда число формирующих показателей больше двух, вводится понятие фиктивного показателя.

Для различения фиктивных и реальных показателей вводятся также понятия, число показателей, участвующих в формировании данного, и индекс порядка. Реальным показателям устанавливается индекс порядка, равный нулю. Фиктивными показателями присваиваются индексы порядка, отличные от нуля.

Общее правило выделения фиктивных показателей рассмотрим на примере. Так, если показатель g_n рассчитывается с использованием ряда других: $g_1, g_2, \dots, g_{n-1}$, то способ расчета показателя g_n может быть представлен двухуровневым графом (рис. 1).

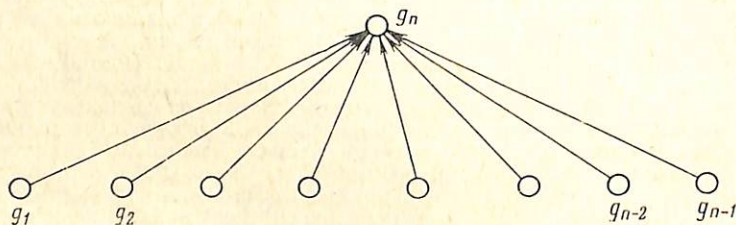


Рис. 1

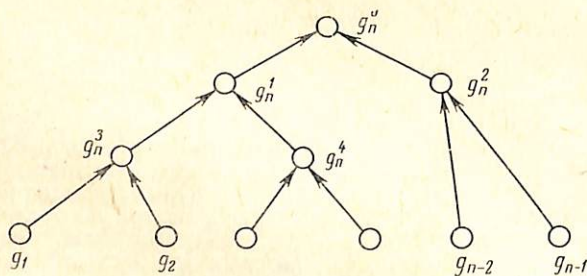


Рис. 2

Для записи по предложенному макету граф формирования показателя должен быть перестроен (см. рис. 2), где показатели $g_1 - g_{n-1}$ представляются элементами входа, попарно соединяющимися в фиктивные элементы.

В этом графе g_n — искомый реальный показатель, $g_n^1, g_n^2, g_n^3, g_n^4$ — фиктивные с различными порядковыми индексами. Для облегчения поиска фиктивным показателям устанавливается тот же номер, что и реальному, но с порядковым индексом, отличным от нуля.

Для каждого описываемого показателя устанавливается количество показателей, участвующих в расчете. Для первичных количество участников расчета будет равно 0, для переписываемых — 1, а для формируемых — 2 и более

Предположим, необходимо описать процесс формирования показателя: «Расходы по содержанию и эксплуатации оборудования на единицу данного вида сравнимой продукции по плану». При этом используются показатели: 1) итого расходов по содержанию и эксплуатации оборудования по плану; 2) сумма основной заработной платы производственных рабочих на выпуск данного вида сравнимой продукции по плану; 3) сумма основной заработной платы производственных рабочих на товарную продукцию по плану; 4) сумма основной заработной платы производственных ра-

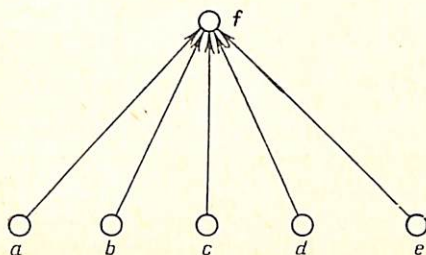


Рис. 3

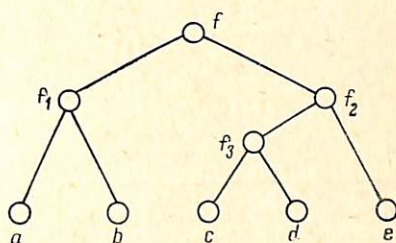


Рис. 4

бочих на незавершенное производство по плану; 5) план выпуска данного вида продукции в физических единицах.

Обозначим формируемый показатель через f , формирующие показатели соответственно через a, b, c, d, e , тогда формирование f может быть представлено

$$f = \frac{ab}{(c+d)e}.$$

Элементарный граф формирования f показан на рис. 3.

По предложенной методике граф формирования должен быть перестроен в соответствии с рис. 4, или $f = f_1 / f_2$, $f_1 = ab$, $f_2 = f_3 e$, $f_3 = c + d$.

В разработочной таблице операция формирования показателя расписывается по строкам (см. табл. 1).

Таблица 1

№ п/п	Формируемый показатель	Индекс порядка	Количество участников	I формирующий показатель	Индекс порядка	Операция	II формирующий показатель	Индекс порядка
1	f	0	5	f	1	(:)	f	2
2	f	1	5	a	0	(×)	b	0
3	f	2	5	f	3	(×)	l	0
4	f	3	5	c	0	(+)	d	0

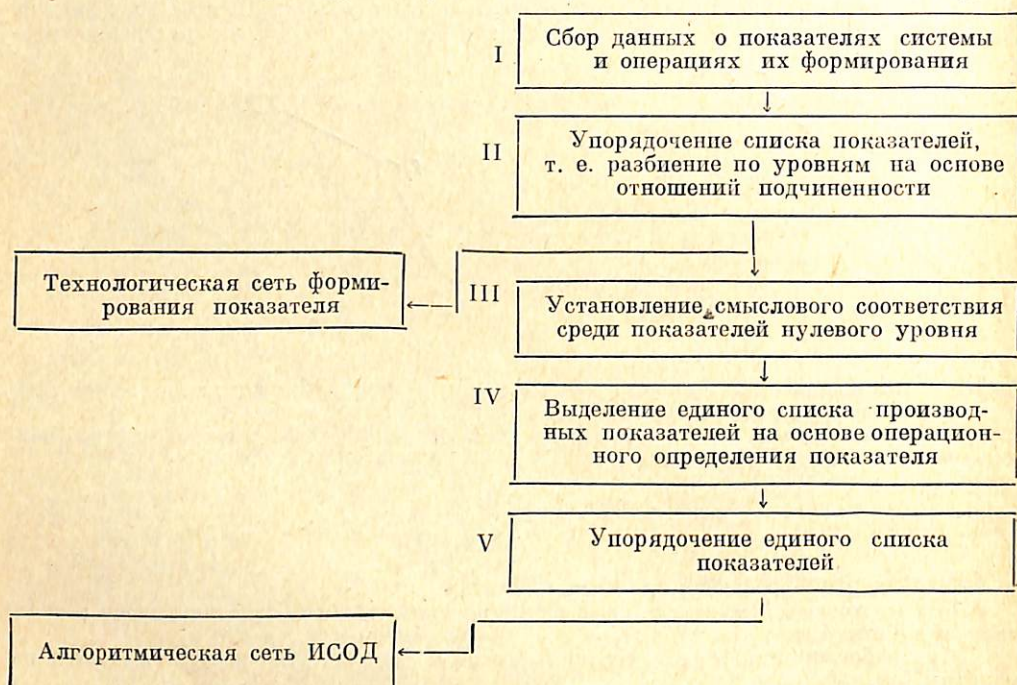
Таблица 2

Графы	Количество знаков	Графы	Количество знаков
№ по порядку	5	Шифр I формирующего документа	4
Шифр формируемого показателя	3	Шифр подразделения	2
Шифр формируемого документа	4	Индекс порядка	1
Шифр подразделения	2	Шифр операции	2
Индекс порядка	1	Шифр II формирующего показателя	3
Вид показателя	1	Шифр II формирующего документа	4
Количество показателей, участвующих в формировании данного	1	Шифр подразделения	2
Шифр I формирующего показателя	3	Индекс порядка	1

В целом макет разработочной таблицы представляется следующим образом (см. табл. 2).

Заполнением разработочных таблиц заканчивается этап обследования потоков информации. Анализ материалов обследования проводится на ЭВМ.

Таким образом, процедура построения алгоритмической сети ИСОД имеет вид



Программа выделения единого списка показателей на ЭВМ разработана С. П. Забаринской (ЦЭМИ АН СССР). В связи с тем, что программирование производилось на языке «АЛГОЛ-60», все характеристики показателя, используемые при построении алгоритмической сети, были разбиты на четыре элемента. Соответственно этому исходные данные составляют четыре массива: A — шифры формируемых показателей; B — операции формирования; C — первые формирующие показатели; D — вторые формирующие показатели.

Эти массивы разбиваются на подмассивы соответственно уровням технологической сети обработки данных. Кроме того, вводятся два массива, обеспечивающие перевод и расшифровку новых номеров показателей: F — шифры всех показателей, введенные в машину (по возрастающим номерам); G — новые шифры всех показателей (строится в соответствии с массивом F).

Алгоритм выделения единого списка показателей содержит три части.

I. Упорядочение исходных данных (массивы F и G). Необходимость этого вызвана тем, что унификация показателей нулевого уровня производится попарно. Поэтому возможен случай, когда установлено смысловое соответствие между f_i и f_j , а также между показателями f_j и f_k . Отсюда можно заключить, что существует также соответствие между показателями f_i и f_k . При большом объеме исходных данных такая работа без применения ЭВМ затруднена. В связи с этим в качестве составного раздела в программу составления единого списка показателей введена эта часть.

II. Перенумерация показателей нулевого уровня. Введение этого раздела программы диктуется необходимостью перенумерации показателей нулевого уровня на основе установленных смысловых соответствий.

III. Перенумерация производных показателей. Основная часть программы, где производится перенумерация производных показателей на основе определений 1) и 2). При этом показатели, формируемые с использованием арифметических операций, подчиненных коммутативному закону, проверяются на соответствие дважды: 1) сравниваются соответственно элементы s и d I показателя с элементами s и d II показателя; 2) сравниваются соответственно элементы s и d I показателя с элементами d и s II показателя. Перенумерация производных показателей производится по уровням формирования.

Поступила в редакцию
8 XII 1970