

Лабораторная работа № 7

Методы анализа и исследования информационных систем

1 Цель работы:

- Изучить методы анализа и исследования информационных систем.
- Построить информационную иерархию информационной системы.
- Провести анализ информационной связности процессов.
- Оформить и защитить отчет у преподавателя.

2 Краткая теория

2.1. Взаимосвязь информационных процессов, систем и технологий

Система – есть сущность, которая в результате взаимодействия ее частей (компонентов) может поддерживать свое существование и функционировать как единое целое. Из определения следует, что поведение системы зависит не от природы свойств ее частей, а от того, как эти части соединены между собой.

Таким образом, системы функционируют как целое, что порождает у них свойства, отличающиеся от свойств составляющих ее частей. Эти свойства известны как эмерджентные. Разбиение системы на составляющие части для познания принципов ее функционирования называется анализом. Дополнением анализа является синтез – создание целого из частей. С помощью синтеза обретаётся понимание.

Любая информационная система создается под заранее известную цель, которая обычно является субъективной, поскольку она предлагается разработчиком, но эта цель должна исходить из объективных потребностей объекта автоматизации.

В настоящее время сформировалась идеология и практика применения различных средств извлечения, передачи, хранения, обработки и представления информации. Однако разрозненное их применение или использование в ограниченной совокупности не позволяло получить значительный системный эффект. Необходим подход к информационным

технологиям как к системе. Необходимо в состав системы включать те элементы, которые своим функционированием обеспечат реализацию заданной цели.

Соответственно, определение системы состоит в следующем: система – конечное множество функциональных элементов и отношений между ними, которые выделяются из окружающей среды в соответствии с поставленной целью в рамках определенного временного интервала ее реализации. Все то, что не вошло в состав системы, относят к окружающей среде. Окружающая среда может включать в себя другие системы, которые реализуют свои цели функционирования. Входы и выходы системы связаны с внешней средой.

Внешней средой для информационной системы может выступать производство, научное исследование, проектирование и т.д. В процессе взаимодействия с внешней средой реализуются основные функции информационной системы. Функции как проявление свойств системы тесно связаны с ее структурой.

Структура системы описывается на концептуальном, логическом и физическом уровне.

Концептуальный уровень позволяет качественно определить основные подсистемы, элементы и связи между ними.

На логическом уровне могут быть сформированы модели, описывающие структуру отдельных подсистем и взаимодействия между ними.

Физический уровень означает реализацию структуры на известных программно-аппаратных средствах.

2.2. Методы анализа и исследования информационных систем

Объектом исследования теории систем является модель, заданная разными способами. Для теории информационных систем наиболее приемлемым является задание модели на основе кибернетического подхода.

Кибернетика рассматривает поведение систем во взаимодействии с другими системами и окружающей средой. Коренными понятиями кибернетики являются система и информация.

Кибернетические системы рассматриваются как системы управления, а процессы управления как процессы переработки информации.

Кибернетическая система – система, для которой принято допущение об относительной изолированности в информационном значении и абсолютной проницаемости в материально-аналитическом значении. Это предполагает, что количество информации в системе конечно, что всякое поступление информации в систему и поступление информации из системы в среду, контролируемы и наблюдаемы.

Управление – функция системы, ориентированная либо на сохранение ее основного качества в условиях изменения среды, либо на выполнение некоторой программы, которая должна обеспечивать устойчивость функционирования, достижение определенной цели.

Система управления – система, в которой реализуются функции управления (рисунок 1).

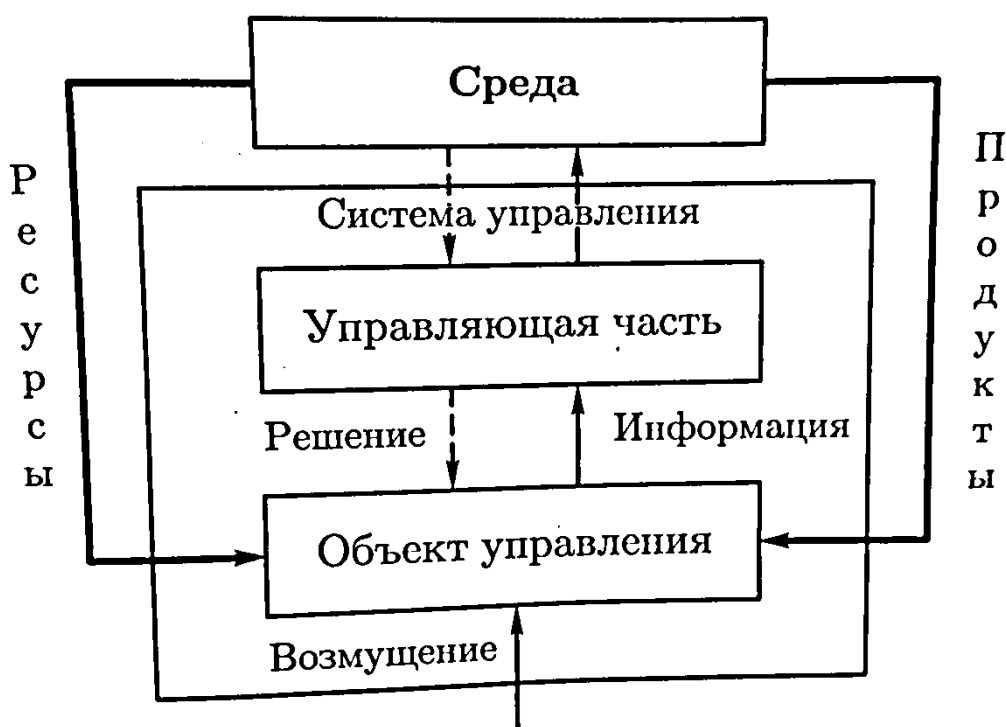


Рисунок 1 – Простейшая структура системы

Объект управления (ОУ) – динамическая система любой природы, преобразующая ресурсы в продукты и находящаяся под действием управляющих и возмущающих воздействий.

Возмущение - воздействие, выводящее систему в нежелательное состояние.

Решение – воздействие, выбранное из множества возможных на основе поставленной цели и принятого критерия.

Критерий – оценка вариантов решения.

Цель – состояние, к которому стремится система.

Управляющая часть (УЧ) – часть системы, вырабатывающая решения и передающая их на объект управления.

Среда – метасистема, в которую рассматриваемая система управления входит составной частью.

Функционирование системы – работа системы в рамках заданной структуры.

Развитие системы – работа системы в условиях острых противоречий, которые могут вызвать изменение структуры.

При более подробном рассмотрении объекта управления его можно представить как систему преобразования ресурсов в продукты (рисунок 2).

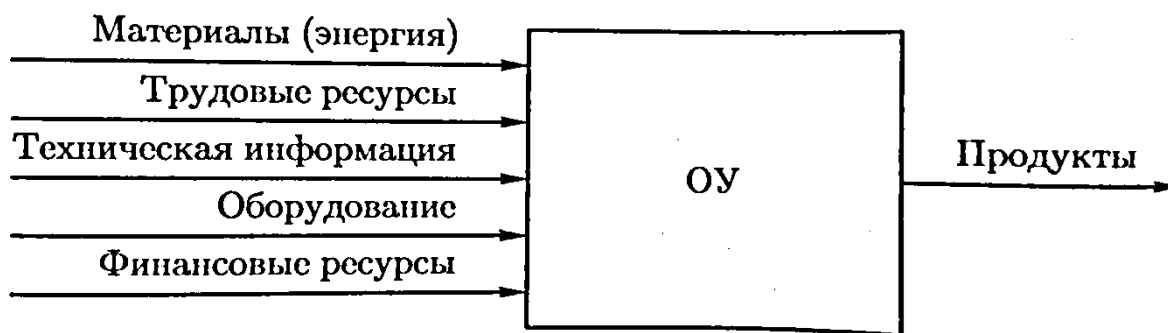


Рисунок 2 – Укрупненная схема объекта управления

При более подробном рассмотрении процессов в управляющей части можно получить совокупность этапов (рисунок 3).

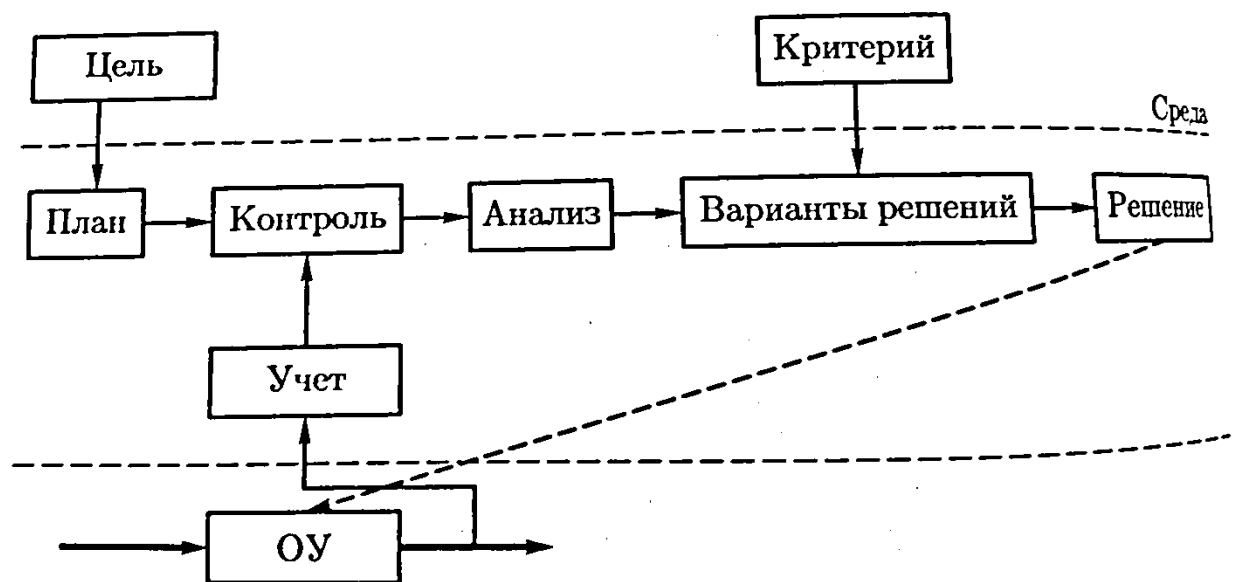


Рисунок 3 – Цикл управления

При анализе процесса управления из-за сложности объекта производят разделение процесса на части по различным признакам. Одним из главных признаков является вид иерархии: временная, пространственная, функциональная, ситуационная и информационная. Деление системы на части не может быть однозначным, т.к. границы между частями являются всегда в какой-то мере субъективными. Выбор того или иного принципа выделения составных частей должен удовлетворять следующим основным условиям: обеспечивать их максимальную автономность; учитывать необходимость координации их действий для достижения общей цели функционирования, а также совмещать отдельные части.

Временная иерархия. Признаком деления здесь является интервал времени от момента поступления информации о состоянии объекта управления до выдачи управляющего воздействия. Чем больше интервал, тем выше уровень (ранг) элемента. Управление может осуществляться в реальном времени, с интервалом в сутки, декаду, месяц, квартал и т.д. Причем управляющий интервал выбирается не произвольно, а исходя из критериев, определяющих устойчивость и эффективность функционирования всей системы.

Пространственная иерархия. Признаком деления является площадь, занимаемая объектом управления. Чем больше площадь объекта, тем выше его ранг. Данный признак является субъективным, так как не всегда площадь, занимаемая объектом, соответствует его значимости, и его можно использовать в случае аналогичности параметров элементов одного уровня.

Функциональная иерархия. В основе лежит функциональная зависимость (подчиненность) элементов системы. Такое разделение является субъективным, так как в этом случае трудно выделить границы между элементами системы.

Ситуационная иерархия. Деление на уровни производится в зависимости от эффекта, вызываемого той или иной ситуацией, например, от ущерба, возникающего в результате аварии или выхода из строя оборудования.

Информационная иерархия. В основе деления лежит оперативность и обновляемость информации. Именно через эти характеристики прослеживается иерархия информации по уровням управления предприятием.

На первом уровне хранится и обрабатывается повторяющаяся, часто обновляющаяся информация, необходимая для повседневной деятельности, т.е. для оперативного управления. Следующий уровень составляет информация более обобщенная, чем оперативная, и используемая не так часто. Информация группируется по функциональным областям и применяется для поддержки принятия решения по управлению производством. На верхнем уровне иерархии хранится и обрабатывается стратегическая информация для долгосрочного планирования. Для нее характерны высокая степень обобщенности, неповторяемость, непредсказуемость и редкое использования.

В общем виде функциональная модель процесса управления представлена на рисунке 4. Учет информации об объекте управления состоит в регистрации, классификации и идентификации. На основе разнообразных математических моделей, описывающих реальное и требуемое состояние объекта, и критериев оптимальности анализируют информацию о состоянии

объекта управления. Окончательная модель прогнозируемого состояния объекта управления формируется в виде плана. Возникающие за счет внешних воздействий отклонения от плана корректируются путем сравнения учетной и плановой информации, нового анализа и формирования управляющих воздействий (регулирования).

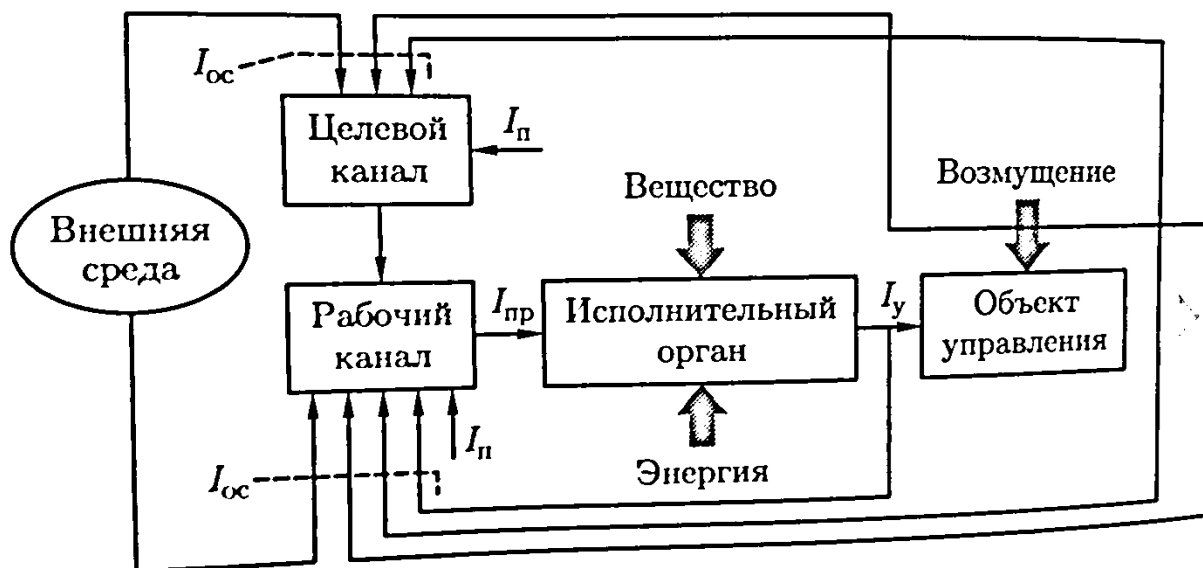


Рисунок 4 – Информационная структура системы управления

В большинстве случаев при информационном анализе процесса управления обычно рассматривают пассивную форму проявления информации, отражающую свойства внешней среды, объекта управления и самой управляющей системы. Однако не менее важное значение имеет и активная форма информации, являющаяся причиной изменения состояния управляемого объекта.

Принято выделять следующие качественно различимые формы проявления информации: осведомляющую – $I_{ос}$, преобразующую – $I_{п}$, принятия решения – $I_{пр}$ и управляющую – $I_{у}$.

К осведомляющей информации относят информацию о состоянии внешней среды, объекта управления и управляющей системы. Преобразующая информация включает информацию, содержащуюся в алгоритмах управления. Информация принятия решения является отражением образов и целей на

конечное множество принимаемых решений. Управляющей информацией является информация, вызывающая целенаправленное изменение состояния объекта управления.

В любой системе управления можно выделить два информационных канала: целевой и рабочий. В целевом канале на основе информационных процессов происходит выбор цели и принятие решения по выбору управляющего воздействия. В рабочем канале формируется информация, реализуемая исполнительным органом, осуществляющим целенаправленное изменение состояние объекта управления через вещественно-энергетические характеристики. Целевой канал может находиться как на одном уровне иерархии с рабочим, так и на более высоком.

Задача теории систем – создание общей методологии упрощения абстрактных систем и тем самым общей теории моделирования.

При кибернетическом подходе к исследованию системы рассматривается динамика ее функционирования во времени.

Представление системы на основе кибернетического подхода представлено на рисунке 5.

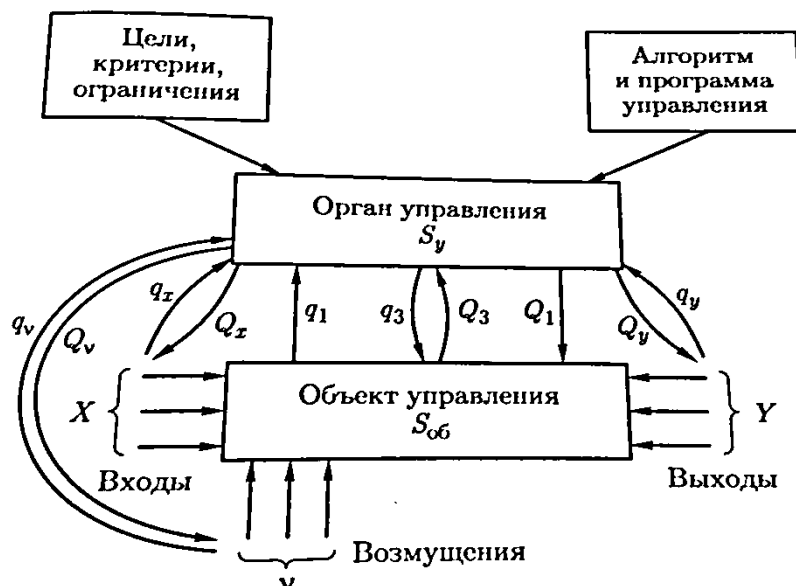


Рисунок 5 – Схема кибернетической системы

Признаки кибернетической системы – наличие объекта и органа управления, связанных обратными и прямыми информационными каналами,

образующими замкнутый контур; целей, критериев эффективности и ограничений; стратегии, плана, алгоритма и программ управления.

При отклонении объекта управления от заданной программы информация по каналам обратной связи поступает от объекта в орган управления. Поступившая информация разрабатывается и сопоставляется с информацией, характеризующей программу достижения целей, определяется рассогласование соответствующих параметров. В управляющем органе разрабатывается и принимается управленческое решение по устранению рассогласований, которое в виде управляющих воздействий подается на объект управления. Наличие всех необходимых признаков кибернетической системы обеспечивает устойчивость ее функционирования.

Для более детального изучения систем в кибернетическом аспекте можно произвести декомпозицию «черных ящиков» первичной модели.

Этапы построения кибернетической системы:

1. Выявление функций системы, реализующих поставленную задачу.
2. Выбор вариантов выполнения функций, т.е. выбор структурных элементов системы.
3. Синтез системы: выбор сочетаний компонентов системы.

В рамках построения информационной системы последовательно рассмотрим его составляющие: типизированные множества и отношения; основные операции; прочность и сцепление компонентов ИС; анализ информационной связности процессов; анализ функциональной связности систем; анализ функциональной связности данных; анализ информационной связности систем.

2.3. Типизированные множества и отношения. Основные операции.

Введем следующие обозначения:

A – множество элементов типа A ;

α – элемент α из множества A (элемент α типа A);

r – тип отношения r ;

ArB – отношение типа r между элементами типов A и B соответственно;

$s[i,j]$ – элемент матрицы S , или $\alpha \text{ r } b[i,j]$, или $A \text{ r } B[i,j]$, или $\alpha_i \text{ r } b_j$;
 $A \text{ r } B[* ,j]$ – j -й столбец;
 $A \text{ r } B[i, *]$ – i -я строка;
 type – базовый тип отношения «тип» ($\alpha \text{ type } A$);
 $:$ – базовый тип отношения «тип» ($\alpha:A$) вида 1-1;
 \rightarrow – базовый тип отношения «тип» ($\alpha \rightarrow A$) произвольного вида (1-1, 1-01, 1-M, 1-0M);
 A_i – i -е подмножество множества A ;
 from – базовый тип отношения «из» ($F_i \text{ from } A$);
 class – базовый тип отношения «обобщения/конкретизации»;
 det – базовый тип отношения «агрегирования/детализации»;
 in – базовый тип отношения «вход»;
 out – базовый тип отношения «выход»;
 and – логическая операция «И»;
 or – логическая операция «ИЛИ»;
 not – логическая операция «НЕ»;
 \otimes – операция логического перемножения матриц ($A \text{ r } 1 \text{ B } \otimes B \text{ r } 2 \text{ C}$);
 \times – операция арифметического перемножения матриц ($A \text{ r } 1 \text{ B } \times B \text{ r } 2 \text{ C}$);
 $(A \text{ r } 2 \text{ A})^t$ – транспонирование;
 fB – строка весовых характеристик f элементов из B ;
 $A \text{ r } B * fB$ – операция поэлементного арифметического умножения каждой строки $A \text{ r } B$ на элемент строки fB ;
 $W(B \text{ r } B):int$ – вес элементов матрицы отношения типа r , заданного на множестве элементов типа B ;
 $F(E \text{ r } A):int$ – функция целого типа, заданная на модели данных.

2.4. Прочность и сцепление компонентов ИС.

Вес взаимосвязи – приведенный к диапазону 0.0...1.0 вес связи между двумя элементами одного вида заданного отношения.

Матрица взаимосвязи – матрица весов взаимосвязи заданного отношения между элементами одного вида.

Абсолютный коэффициент связности – нормированное среднее значение соответствующих взаимосвязей с учетом внутренней связности элементов $\kappa = \text{Sum}(\text{ArA})/N^{**}2$, где ArA – матрица взаимосвязи элементов A по виду отношения r ; $\text{Sum}(\text{ArA})$ – сумма элементов ArA ; N – размерность ArA .

Основные свойства отношения детализации.

1. Отсутствие циклов:

$$\text{AcircleA} = \text{AndetA} \text{ and } (\text{AndetA})^t = I.$$

2. Если задано некоторое отношение ArB , и AdetA , BdetB , то

$$\text{ArB} \subseteq \text{AdetA} \otimes \text{ArB} \text{ and } \text{ArB} \subseteq \text{ArB} \otimes (\text{BdetB})^t.$$

Основные свойства отношения конкретизации:

1. Отсутствие циклов:

$$\text{AcircleA} = \text{AnclassA} \text{ and } (\text{AnclassA})^t = I.$$

2. Если задано некоторое отношение ArB , и AclassA , то $\text{ArB} \subseteq \text{AnclassA} \otimes \text{ArB}$.

Прочность элемента – степень взаимосвязи его внутренних элементов в соответствии с заданной матрицей взаимосвязи.

Сцепление элемента – степень взаимосвязи этого элемента с элементами окружения.

Интегральная оценка связности. Концепция независимости и прочности направлена на выделение и локализацию в проектируемой системе сильно связанных групп элементов.

Группа должна быть прочно или слабо связана с окружением. Один из механизмов группирования – агрегирование (детализация), приводящее к иерархической структуре рассматриваемого множества элементов. Желательно иметь численную оценку, характеризующую качество иерархической структуры. Оценка должна ухудшаться, если сильно связанные элементы оказываются в разных агрегированных группах. Естественно стремиться, чтобы структура каждого агрегированного элемента могла быть оценена аналогичным образом.

Будем оценивать качество некоторой структуры на основе матрицы коэффициентов связности $S=A_{fun}A$, рассчитываемых для различных типов элементов и различной природы связности. Частными случаями являются:

– полносвязная система $S=11$, где 1 – единичная матрица, все элементы которой равны 1 ;

– несвязанная система $S=I$, где I – единичная диагональная матрица.

Диагональные элементы $S[i,j]$ являются интегральной оценкой связности внутренней структуры и характеризуют прочность элементов A . Прочие элементы S характеризуют сцепление элементов A между собой.

Представим A как множество элементов модели первого уровня детализации. Нулевой уровень представим как композиционную модель в виде одного агрегированного элемента путем объединения всех элементов из A . Добавим новый элемент в A :

$$A_{det}A = \begin{vmatrix} 0 & & & \\ 0 & & I & \\ 0 & & & \\ 1 & 1 & 1 & 1 \end{vmatrix},$$

где I – единичная диагональная подматрица; $A[n+1]$ – агрегированный элемент.

Расширим матрицу S путем добавление первой нулевой строки и нулевого столбца:

$$Q = \begin{vmatrix} & & 0 \\ & S & 0 \\ 0 & 0 & 0 \end{vmatrix}.$$

Оценим прочность агрегированного элемента $A[n+1]$. При наличии N объединяемых элементов при максимальной их связности ($S=1$) число связей равно N^*2 . Таким образом, если все элементы характеризуются прочностью,

равной 1, а также каждый из них максимально связан с остальными ($S[i,j]=S[j,i]=1$), то сумма всех оценок связности будет равна $N^{**}2$.

Определим прочность агрегированного элемента:

$$K_s = S[n + 1, n + 1] = Sum(S)/N^{**}2$$

Деление на $N^{**}2$ обеспечивает условие нормировки оценки. Для полносвязной системы

$$S[n + 1, n + 1] = Sum(S)/N^{**}2 = 1$$

Для несвязной системы прочность агрегированного элемента определяется только прочностью его несвязных составляющих элементов:

$$S[n + 1, n + 1] = Sum(S)/N^{**}2 = 1/N$$

Таким образом, получена интегральная оценка, которая характеризует качество агрегирования группы взаимосвязных элементов. Оценка учитывает не только парные коэффициенты взаимности, но и прочность объединяемых элементов. Показатель прочности агрегированного элемента уменьшается при снижении прочности какого-либо элемента. Значит, полученный показатель оказывается чувствительным к изменениям структуры на всю глубину детализации. Коэффициент связности можно представить в виде суммы двух составляющих:

$$K_s = K_{in} + K_{out}$$

где K_{in} – интегральная оценка внутренней связности объединяемых элементов; K_{out} – интегральная оценка связности объединяемых элементов между собой.

Коэффициент внутренней связности равен:

$$K_{in} = Sum(S[i, i], i = 1, \dots, N)/N^{**}2$$

Коэффициент внешней связности определяется следующим образом:

$$K_{out} = Sum(S[i, j], i = 1, \dots, N, j = 1, \dots, N, i \neq j)/N^{**}2$$

Относительный коэффициент связности характеризует долю связности объединяемых элементов между собой по отношению к общей связности:

$$K_{so} = K_{in}/K_s$$

Ранее рассматривалась оценка связности на одном уровне декомпозиции структуры. При этом не учитывались внутренние взаимосвязи элементов и взаимосвязь объемлющего элемента с окружением. Применим описанный выше подход для оценки связности иерархических структур.

Пусть внутренняя связность элемента $MfunM[i,j]$ равна $S[i,j]$. На рисунке 6 приведен пример иерархической структуры и связности элементов.

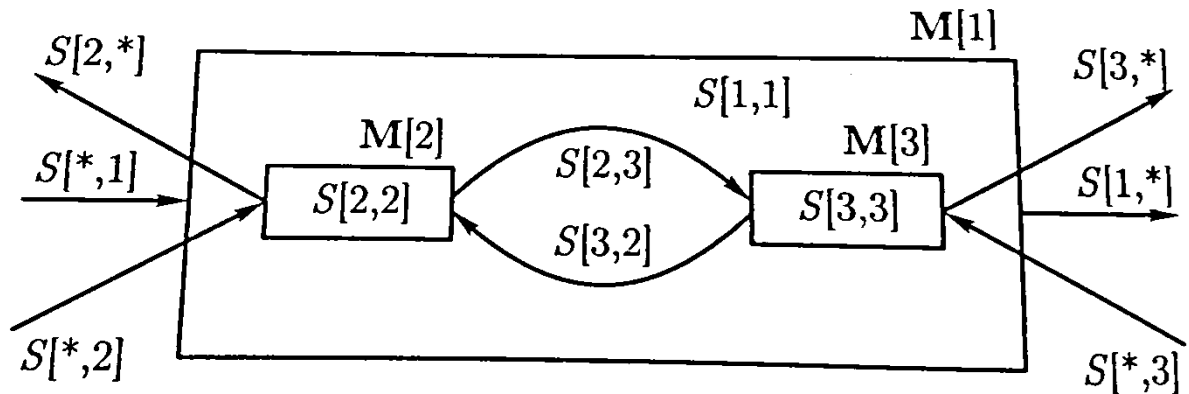


Рисунок 6 – Пример иерархической структуры

Задано отношение детализации $MdetM = \begin{vmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix}$.

Известны внутренние связности 2-го и 3-го элементов $S[2,2]$, $S[3,3]$, а также их взаимосвязи между собой $S[2,3]$, $S[3,2]$ и окружением $s[2,*]$, $S[3,*]$, $S[* ,2]$, $S[* ,3]$.

Требуется найти для агрегированного элемента $M[1]$ его прочность $S[1,1]$ и его связность окружением $S[1,*]$ и $[* ,1]$.

Прочность $M[1]$ оцениваем как интегральный коэффициент связности элементов его детализации $M[2]$ и $M[3]$:

$$S[1,1] = (S[2,2] + S[3,3]) / 2^{**}2 + (S[2,3] + S[3,2]) / 2^{**}2$$

Для определения связности агрегированного элемента $M[1,1]$ с окружением рассмотрим ряд частных случаев.

Вариант 1. Полная связность внутренних элементов $M[2]$ и $M[3]$ с окружением (элементом $M[4]$). В этом случае

$$S[2,4] = S[4,2] = S[3,4] = S[4,3] = 1$$

Естественно предположить, что в этом случае $S[1,4] = S[4,1] = 1$. Поэтому по условиям нормирования связность агрегированного элемента с окружением определяется следующим образом:

$$S[1,4] = (S[2,4] + S[3,4])/2$$

$$S[4,1] = (S[4,2] + S[4,3])/2$$

Это выражение остается справедливым и в случае, когда связность какого-либо из внутренних элементов с окружением равна 0.

В общем случае связь агрегированного элемента с окружением будет оцениваться как среднее по множеству связей с окружением его составляющих элементов на первом уровне детализации:

$$S[m, j] = \text{Sum}(S * \text{Mdet}M[m, *],) / \text{count}(\text{Mdet}M[m, *],)$$

где i – входит в состав агрегированного элемента $M[m, *]$.

Для перерасчета связности при выполнении операции агрегирования введем специальную операцию Agr:

$S_{\text{agr}} = \text{Agr}(S, \text{Mdet}M)$ – операция расчета связности и прочности агрегированных элементов M на основе исходной матрицы связности и отношения детализации $\text{Mdet}M$.

Операция Agr на основе матрицы связности S N элементов и матрицы группирования этих элементов $G_{\text{det}}S$ рассчитывает матрицу связности S_{agr} , в которой каждой группе соответствуют новые строки и столбцы.

Общий случай операции агрегирования имеет вид:

$$Q_{\text{ffun}}Q = \begin{vmatrix} M_{\text{ffun}}M & M_{\text{ffun}}F & 0 \\ F_{\text{ffun}}M & F_{\text{ffun}}F & F_{\text{ffun}}G \\ G_{\text{det}}M & G_{\text{ffun}}F & G_{\text{ffun}}G \end{vmatrix},$$

где M – множество агрегируемых элементов для данной операции;

F – множество элементов, которые не агрегируются в данной операции;

G – множество групп (новых агрегированных элементов, создаваемых в данной операции);

$E=M \cup F$ – исходное множество рассматриваемых элементов;

$$S = \left| \begin{array}{c|c} MffunM & MffunF \\ \hline FffunM & FffunF \end{array} \right|.$$

- исходная матрица связности элементов;

$GdetM$ – матрица группирования элементов из M в соответствующие группы;

$GffunG$ – матрица связности агрегированных элементов между собой ($GffunG[i,i]$ – прочность группы);

$GffunF, FffunG$ – подматрицы связности агрегированных элементов из G с элементами, которые не включены ни в одну группу;

$$S_{agr} = \left| \begin{array}{c|c} FffunF & FffunG \\ \hline GffunF & GffunG \end{array} \right|.$$

- матрица связности после выполнения операции агрегирования.

2.5. Анализ информационной связности процессов

Информационная связность процессов характеризует наличие общих данных, используемых соответствующей парой процессов (как по входу так, и по выходу). Абсолютный коэффициент информационной связности двух процессов равен 1, если эти процессы используют одно множество данных, и равен 0, если общих данных у процессов нет.

В качестве исходных данных должны быть заданы система, связность процессов которой оценивается; глубина детализации процессов по схеме требований действий; веса данных (поток).

Обозначим множество данных через E , множество процессов – через P .

Связи процессов и данных по входу и выходу зададим в виде соответствующих отношений:

$E_{in}P$ – входные данные процессов;

$E_{out}P$ – выходные данные процессов.

Пусть известна частота активизации процессов, заданная вектором fP .

Связь процессов с данными по входу или выходу определится следующим образом:

$$E_{inout}P = E_{in}P \text{ or } E_{out}P.$$

Совместное использование данных хотя бы одним процессом определится следующим образом:

$$L = E_{com}E = E_{inout}P \otimes (E_{inout}P)^t,$$

где $L[i,j] = 1$, если существует хотя бы один процесс, использующий совместно данные $E[i]$ и $E[j]$.

Для того, чтобы учесть количество процессов, совместно использующих пару данных, необходимо операцию логического умножения заменить на операцию арифметического умножения:

$$C = E_{count}E = E_{inout}P \times (E_{inout}P)^t,$$

где $C[i,j] > 0$ – количество процессов, совместно использующих данные $E[i]$ и $E[j]$.

Для анализа связности часто необходимо учитывать частоту активации процессов fP . В этом случае связность данных по использующим их процессам определится следующим образом:

$$fC = E_{fcount}E = (E_{inout}P * fP) \times (E_{inout}P)^t,$$

где $fC[i,j] > 0$, частота совместного использования данных $E[i]$ и $E[j]$ с учетом частоты активизации процессов; \times - операция арифметического умножения матриц.

Каждый диагональный элемент $fC[i,i] \geq fC[i,j]$ и $fC[i,i] \geq fC[j,i]$. Проведем нормирование соответствующих элементов матрицы весов взаимосвязей путем деления каждого из них на диагональный элемент:

$$S[i,j] = fC[i,j] / fC[i,i], S[j,i] = fC[j,i] / fC[i,i].$$

Матрица S определяет относительную частоту совместного использования данных, а каждый ее элемент можно назвать коэффициентом

совместного использования. При $S[i,j]=0$ соответствующие данные совместно не используются. При $S[i,j]=1$ соответствующие данные на рассматриваемом множестве процессов всегда используются совместно. Введем понятие коэффициента связности.

Коэффициент связности $S[i,j]$ – количественная мера связи заданного вида элемента i с элементом j , каждый из которых относится к одному типу.

Матрица S содержит полезную информацию для анализа модели предметной области и проектных решений по разработке ИС. Число элементов такой модели может составлять несколько тысяч, количество взаимосвязей, по крайней мере, на два порядка больше.

Множество данных можно рассматривать как набор сущностей модели данных, разрабатываемой в рамках конкретного проекта. Сущности в различной степени функционально зависимы между собой, что проявляется в наличии связей «FOREIGN KEY». Однако величина зависимости сущностей из проекта модели данных получена быть не может. Необходима процедурная модель обработки данных для количественной оценки связности сущностей. Описанный подход позволяет количественно оценить функциональные зависимости сущностей.

Модель данных представляет собой двухуровневое описание взаимосвязей двух множеств классов: атрибутов (A) и сущностей (E).

Структура сущностей описывается через состав атрибутов с указанием элементов ключевого первичного ключа ($primary_key$) и внешних связей сущностей ($foreign_key$).

$EstrA$ – структура сущностей;

$Eprimary_keyA$ – первичные ключи сущностей;

$Aforeign_keyE$ – внешние ключи сущностей.

Зависимости сущностей FOREIGN KEY можно описать следующим образом:

$$Eforeign_keyE = EstrA \otimes Aforeign_keyE,$$

где $E_{foreign_keyE[i,j]}$ равно 1, если существует хотя бы один атрибут из состава $E[i]$, который является ссылкой на $E[j]$.

Более сильная зависимость, когда связь является идентифицирующей:

$$E_{primary_keyE} = E_{primary_keyA} \otimes A_{foreign_keyE},$$

где $E_{primary_keyE[i,j]}$ равно 1, если существует хотя бы один ключевой атрибут из состава $E[i]$, который является ссылкой на $E[j]$.

Оценка функциональной зависимости данных основывается на анализе внешних связей сущностей ($foreign_key$). Наличие таких зависимостей определяется выражением:

$$E_{funE} = E_{strA} \otimes A_{foreign_keyE},$$

где $E_{funE[i,j]} = 1$, если существует хотя бы один атрибут из состава $E[i]$, который является ссылкой на сущность $E[j]$.

Пусть известна частота активизации связи, заданной атрибутом $A[j]$ из состава этой сущности $E[i]$. Вектор строка fA задает эту меру. Сущность $E[i]$ активизируется, когда активизируется или атрибут из ее состава или атрибут ссылка на данную сущность. Поэтому для оценки взаимосвязи атрибутов предварительно получим матрицу зависимости сущностей и атрибутов:

$$E_{funA} = (E_{strA} \text{ or } (A_{foreign_keyE})^t),$$

где $E_{funA[i,j]} = 1$, если атрибут $A[j]$ входит в состав $E[i]$ или является ссылкой на $E[i]$.

Тогда количественная оценка взаимосвязей сущностей по внешним связям с учетом их частоты активизации определится следующим образом:

$$fC = E_{ffunE} = (E_{funA} * fA) \times (E_{funA})^t,$$

где $fC[i,j] > 0$ – частота совместного использования данных $E[i]$ и $E[j]$ с учетом частоты активизации ссылочных связей;

* - операция поэлементного арифметического умножения;

\times - операция арифметического умножения матриц.

Далее необходимо нормировать E_{ffunE} по элементам главной диагонали: $S = \text{Norm}(E_{ffunE})$.

Требуется оценить связность сущностей для фрагмента модели данных, представленной на рисунке 7.

Последовательно строится матричная модель для фрагмента в соответствии с выражениями, приведенными выше.

$$EstrA = \begin{vmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{vmatrix} - \text{состав атрибутов сущностей}$$

$$Aforeign_keyE = \begin{vmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{vmatrix} - \text{внешние ключи сущностей}$$

$$fA = |0,0 \ 1,0 \ 0,0 \ 1,0 \ 0,0 \ 0,5 \ 0,5 \ 0,5 \ 0,5 \ 0,0| - \text{частота активизации атрибутов.}$$

$$EfunA = \begin{vmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 1 \end{vmatrix} - \text{функциональные зависимости сущностей}$$

и атрибутов;

$$EfunA * fA = \begin{vmatrix} 0,0 & 1,0 & 0,0 & 0,0 & 0,0 & 0,0 & 0,0 & 0,0 & 0,0 & 0,0 \\ 0,0 & 1,0 & 0,0 & 1,0 & 0,0 & 0,5 & 0,0 & 0,0 & 0,0 & 0,0 \\ 0,0 & 0,0 & 0,0 & 0,0 & 0,0 & 0,5 & 0,5 & 0,5 & 0,5 & 0,0 \\ 0,0 & 0,0 & 0,0 & 1,0 & 0,0 & 0,0 & 0,5 & 0,5 & 0,5 & 0,0 \end{vmatrix}$$

$$EffunE = \begin{vmatrix} 1,0 & 1,0 & 0,5 & 0,5 \\ 1,0 & 2,5 & 0,5 & 1,0 \\ 0,0 & 0,5 & 2,0 & 1,5 \\ 0,0 & 1,0 & 1,0 & 2,0 \end{vmatrix}$$

$$S = \text{Norm}(\text{Effun}E) = \begin{vmatrix} 1,0 & 0,4 & 0,25 & 0,25 \\ 1,0 & 1,0 & 0,25 & 0,5 \\ 0,0 & 0,2 & 1,0 & 0,75 \\ 0,0 & 0,4 & 0,5 & 1,0 \end{vmatrix}$$

$$K_s = \text{Sum}(S) / N^{**}2 = 8,3/16 = 0,53$$

Формируем GstrE путем введения двух групп, каждая из которых является агрегированием сущностей E[1],E[2] и E[3],E[4], соответственно:

$$\text{Gstr}E = \begin{vmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{vmatrix}$$

Используя матрицу функциональной зависимости сущностей, приведенную выше, и матрицу группирования сущностей, выполним операцию Agr. В результате получим матрицу функциональной зависимости групп GffunG:

$$\text{Gffun}G = \begin{vmatrix} 1,0 & 0,4 & 0,25 & 0,25 \\ 1,0 & 1,0 & 0,25 & 0,5 \\ 0,0 & 0,2 & 1,0 & 0,75 \\ 0,0 & 0,4 & 0,5 & 1,0 \end{vmatrix} \text{Agr} \begin{vmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{vmatrix}$$

$$\text{Gffun}G = \text{Agr}(S, \text{Gstr}E) = \begin{vmatrix} 0,85 & 0,31 \\ 0,15 & 0,81 \end{vmatrix}$$

На основе GffunG рассчитаем интегральные показатели полученной структуры:

$$K_z = \text{Sum}(\text{Gffun}G) / N^{**}2 = 2,12/4 = 0,53; K_{zn} = \text{Sum}(S[i,i], i = 1..N) / N^{**}2 = 0,415; K_{oul} = \text{Sum}(S[i,j], i = 1..N, j = 1..N, i \neq j) / N^{**}2 = 0,115; KJK_s = 0,783; K_{\phi}JK_s = 0,217.$$

За основу формирования супергрупп принимается матрица функциональной зависимости групп данных GffunG. Необходимо принять решение о количестве супергрупп. Можно последовательно проверять гипотезы о количестве супергрупп, равном 1, 2, 3 и т.д. Сравнение вариантов

предлагается проводить на основе интегрированной оценки связности иерархической структуры.

Рассмотрим вариант формирования структуры супергрупп при известном их количестве, например, равном 2. Необходимо так разделить элементы из G , чтобы интегрированный показатель качества разделения был максимален.

Приближительное решение:

1. Выбирают две самые прочные группы, которые имеют слабые связи между собой;

2. Для каждой из оставшихся групп решение о принадлежности к одной из супергрупп принимаем на основе интегральной оценки ее связности с каждой из супергрупп.

В результате получим две матрицы: $SGstrG$ — структура супергрупп; $SGffunSG = Agr(GffunG, SGstrG)$ — матрица функциональной зависимости групп сущностей.

Функциональное представление модели данных рассматривает атрибут как функцию доступа. Введем следующие обозначения:

$Q \subset E$ - множество независимых сущностей;

$Q_{primary_key} A = I$ — независимая сущность, имеющая простой первичный ключ;

$Z \supset E$ - множество зависимых сущностей;

$Z_{primary_key} A \otimes A_{foreign_key} E$ — в каждой строке количество ненулевых элементов > 1 .

Строгая типизация атрибутов задается отношением $A \rightarrow D$. Строго говоря, $A:B$ — на физическом уровне только базовые типы. Однако на логическом уровне $A_{foreign_key} E$ определяет способ реализации функциональной зависимости сущностями путем включения в состав одной сущности атрибута ссылки на элементы другой сущности.

Рассмотрим зависимости между элементами Q . Двойственность атрибутов следует из отношения $E[j]A[i]E$.

Атрибут $A[i] \rightarrow D$ задает отображение $V:D$ на множество параметров.

Параметры $VB:B$ — параметры простых (базовых) типов; $VQ:Q$ — параметры, ссылки на независимые сущности.

В частном случае $V:E$ заменяется на $V_{foreign_key}E$. С одной стороны, атрибут выступает как свойство конкретной сущности. Этот аспект отражается в отношении $EstrA$. С другой стороны,

$$t D[j]A[i]D=(A[i] \rightarrow D) \otimes A[i]inV \otimes V : D \text{ и } f(V:E) \rightarrow E[j]$$

Объектно-ориентированный подход требует выделения сильно связанных компонентов при формировании структуры классов. Потенциально каждая сущность является кандидатом для формирования соответствующих классов обработки и/или представления. Однако возможность обоснованно группировать сущности в иерархические структуры, несомненно, поможет конструкторам ИС более обоснованно формировать структуру классов приложения.

Для формирования групп сущностей используют методику оценки связности, на основе которой строится матрица $E/согт \mathbb{E}$.

Анализ функциональной связности систем. Функциональная связность систем характеризует наличие общих требуемых процессов для процессов, входящих в разные системы.

Коэффициент связности двух систем равен 0, если все процессы каждой из систем требуют для своего выполнения разных подмножеств процессов, и равен 1, если множества требуемых процессов одинаковы.

В качестве исходных данных должны быть заданы:

- система, связность компонентов которой оценивается;
- глубина детализации систем по отношению состава;
- веса процессов.

Анализ функциональной связности данных. Функциональная связность данных характеризует наличие общих процессов, которые используют по входу-выходу оцениваемые данные.

Коэффициент функциональной связности двух данных равен 1, если множество процессов, использующих эти данные по входу-выходу одинаково, и равен 0, если общих процессов у данных нет.

Анализ информационной связности систем. Анализ характеризует наличие общих данных у этих систем.

3. Задание для самостоятельного выполнения:

- Для информационной системы 2 из лабораторной работы №1 (согласно вашего варианта) разработать схему информационной иерархии.
- Для информационной системы из лабораторной работы № 1 (согласно вашего варианта) провести анализ информационной связности процессов.
- Составить отчет и защитить лабораторную работу у преподавателя.
 - Отчет включает в себя:
 - Титульный лист;
 - Название и цель работы;
 - Информационную иерархию системы с подробным описанием;
 - Подробный анализ информационной связности процессов.

Дополнительная информация:

Ниже приведена таблица из лабораторной №1 с вариантами индивидуальных заданий (смотри таблицу 1.1).

В данной работе рассмотрению подлежат социально-экономические системы.

Если выполнение задания вызывает серьезные затруднения, рекомендуется вначале рассмотреть описание технической системы, согласно своего индивидуального варианта.

Таблица 1.1 – Примеры систем для индивидуального выполнения

<i>№ п.п.</i>	<i>Техническая система</i>	<i>Социально-экономическая система</i>
1	САПР	Супермаркет
2	Грузовик	Банк
3	Кондиционер	Комитет по налогам и сборам
4	Телевизор	Гостиница
5	Телефон	Торговое предприятие
6	Фотоаппарат	Производственное предприятие
7	Трамвай	Оптовая фирма
8	Кофемолка	Магазин розничной торговли
9	Микрофон	Табачная фабрика
10	Кухонный комбайн	Кинотеатр
11	Самолет	Ресторан
12	Огнетушитель	Мебельная фабрика
13	Сигнализация	Трикотажная фабрика
14	Часы	Салон красоты
15	Холодильник	Студия звукозаписи
16	Магнитофон	Агентство по трудоустройству
17	Утюг	Лотерея
18	Микроскоп	Рекламное агентство
19	Копиров.аппарат	Ювелирная мастерская
20	Спутник	Строительная фирма
21	Робот	Агентство недвижимости
22	Микроволновая печь	Биржа