

УДК 681.327.8

**С.В. Павлов**

Одесская государственная академия холода, г. Одесса  
кафедра информационно-коммуникационных технологий  
E-mail: [semyen@rambler.ru](mailto:semyen@rambler.ru)

## ПОШАГОВОЕ ВЫДЕЛЕНИЕ КЛАССОВ И КЛАССИФИКАЦИОННЫХ ФОРМУЛ ИНФОКОММУНИКАЦИОННЫХ УСЛУГ

### *Abstract*

*Pavlov S.V. Incremental selection of classes and classifications formulas of infocommunication services. Classification of infocommunication services by the method of discriminant analysis is considered. The classes of services and their classifications formulas are certain. Approach of forming of requirements to the networks and equipment for the grant of set of infocommunication services is certain.*

**Keywords:** *infocommunication services – discriminant analysis – classification of services.*

### *Анотація*

*Павлов С.В. Покрокове виділення класів і класифікаційних формул інфокомунікаційних послуг. Розглянута класифікація інфокомунікаційних послуг методом дискримінантного аналізу. Визначені класи послуг і їх класифікаційні формули. Визначений підхід формування вимог до мереж і устаткування для надання набору інфокомунікаційних послуг.*

**Ключові слова:** *інфокомунікаційні послуги – дискримінантний аналіз – класифікація послуг.*

### *Аннотация*

*Павлов С.В. Пошаговое выделение классов и классификационных формул инфокоммуникационных услуг. Рассмотрена классификация инфокоммуникационных услуг методом дискриминантного анализа. Определены классы услуг и их классификационные формулы. Определен подход формирования требований к сетям и оборудованию для предоставления набора инфокоммуникационных услуг.*

**Ключевые слова:** *инфокоммуникационные услуги - дискриминантный анализ - классификация услуг.*

**Общая постановка проблемы.** Актуальность классификации параметров предоставления инфокоммуникационных услуг объясняется непрерывным появлением новых услуг, и, соответственно, необходимостью их упорядочивания и группировки, потому что группами услуг оперировать проще, чем каждой из услуг отдельно, а также постоянной необходимостью общества к быстрому манипулированию информацией. Классы услуг формируются на основании существующих методов классификации и снижения размерности. Любая из существующих или прогнозируемых услуг сопоставляется одному из выделенных классов. В свою очередь класс услуги определяет требования к сети и оборудованию, необходимые для предоставления данной услуги. Таким образом, для определения требований к сетям и оборудованию новой услуги достаточно будет лишь определить ее класс. Учет требований инфокоммуникационных услуг предоставит возможность четко планировать развитие и оценить необходимые ресурсы для последующей модификации существующих сетей. Классификация услуг также полезна при планировании пакетов услуг оператором связи, так как последний будет наглядно представлять возможности конкретной телекоммуникационной сети по представлению инфокоммуникационных услуг, а также требования, которым данная сеть должна отвечать (требования к показателям предоставления услуг, таким как полоса пропускания, коэффициент ошибок, время доступа к услуге и другим показателям).

Цель работы состоит в классификации инфокоммуникационных услуг (далее — ИКУ) методом дискриминантного анализа.

Объект исследования — параметры предоставления инфокоммуникационных услуг [1].

Состояние вопроса. Необходимость в классификации ИКУ обусловлена постоянным увеличением числа этих услуг и обоснована в [1]. Для выполнения такой классификации потребовалось снижение размерности параметров предоставления услуг, выделенных в соответствии с рекомендациями Международного союза электросвязи [2] (таких как информационная скорость доступа, готовность сети, вид передаваемой информации и другие параметры). Постановка этой задачи приведена в [3]. Для решения указанной задачи задействован факторный и кластерный анализ параметров предоставления ИКУ [5].

Выбран метод для проведения классификации ИКУ — пошаговый дискриминантный анализ с включением [6]. В статье выполнена классификация параметров предоставления инфокоммуникационных услуг методом дискриминантного анализа. Выделены классы ИКУ, формирующие требования к сетям и оборудованию, и формулы классификации, определяющие класс любой услуги.

**Постановка задач исследования.** Для предоставления ИКУ необходимо сформировать требования к сетям и оборудованию (к пропускной способности, допустимому коэффициенту ошибок и другим параметрам). Эти требования выдвигаются на основании классов услуг, поэтому необходимо провести классификацию параметров предоставления ИКУ [1]. Для этого целесообразно применить пошаговый дискриминантный анализ [6] — выделить классы услуг и характерные для них классификационные формулы [4]. Классификационные формулы определяют класс услуги на основании параметров ее предоставления. Каждый найденный класс формирует требования к сетям и оборудованию для предоставления ИКУ [3]. На основании поученной модели классов инфокоммуникационных услуг можно определить класс любой из услуг и определить требования к сетям и оборудованию для ее предоставления.

Основная идея дискриминантного анализа заключается в определении отличия совокупности параметров ИКУ по среднему какой-либо переменной (или линейной комбинации переменных), и использовании этой переменной для предсказания новых членов и их принадлежность к той или иной группе. Общий вид функции дискриминации между двумя совокупностями параметров инфокоммуникационных услуг [4, 7]:

$$d_{ik} = a_{0k} + a_{1k} \cdot x_{i1k} + a_{2k} \cdot x_{i2k} + \dots + a_{jk} \cdot x_{ijk} + \dots + a_{mk} \cdot x_{imk}, \quad (1)$$

где:  $i=1, 2, \dots, n_k$ ,  $k=1, 2, \dots, p$ ;

$p$  — количество непересекающихся подмножеств строк, на которые разбита таблица наблюдений;

$a_{jk}$  — неизвестные коэффициенты;

$d_{ik}$  — расчетное значение функции для  $i$ -го объекта из группы  $k$ , состоящей из совокупности  $n_k$  измерений;

$x_{ijk}$  — значение  $j$ -ой дискриминантной переменной;

$j=1, 2, \dots, m$  — столбцы матрицы наблюдений.

Исходными данными для дискриминантного анализа приняты показатели, полученные в результате использования кластерного и факторного анализа [5], а именно:

- 1) «Время доступа к службе».
- 2) «Информационная скорость доступа».
- 3) «Вид передаваемой информации».
- 4) «Наименование услуги».
- 5) «Код службы, вид услуги».
- 6) «Готовность сети».

На данном этапе исключим показатель «Код службы, вид услуги», т.к. службы формируются для предоставления группы услуг, а дополнительные и интеллектуальные услуги не могут предоставляться без основной, а только расширяют ее функциональность. В результате чего данный показатель не может оказывать первостепенное влияние на формирование классов ИКУ.

Также исключим показатель «Наименование услуги», т.к. этот показатель лишь словесно определяет ту или иную услугу. Данный параметр неоднозначен и может быть изменен в не зависимости от модификации способа предоставления услуги (обратное утверждение также справедливо) либо по желанию поставщика услуг.

Таким образом, в качестве исходных данных для дискриминантного анализа в данной работе приняты следующие показатели предоставления инфокоммуникационных услуг:

- 1) «Время доступа к службе»;
- 2) «Информационная скорость доступа»;
- 3) «Вид передаваемой информации»;
- 4) «Готовность сети».

В качестве группирующей переменной выберем показатель «Вид передаваемой информации». Этот показатель однозначно характеризует типы ИКУ. Независимые переменные — «Время доступа к службе», «Информационная скорость доступа», «Готовность сети».

Для обозначенных переменных применим пошаговый дискриминантный анализ с включением [4]. В данном анализе дискриминантных функций модель дискриминации строится по шагам. На каждом шаге просматриваются все переменные, и находится та из них, которая вносит наибольший вклад в различие между совокупностями параметров ИКУ. Эта переменная должна быть включена в модель на данном шаге, после чего происходит переход к следующему шагу.

В нашем случае последний шаг будет третьим, число переменных в модели — три (так как независимых переменных — три). На следующем этапе необходимо рассчитать  $p$  линейных дискриминантных функций, равных количеству анализируемых параметров предоставления ИКУ, после чего с использованием коэффициентов  $a_{jk}$  и постоянной  $a_{0k}$  (для расчета коэффициентов дискриминантных функций нужен статистический критерий, оценивающий различия между группами) можно провести классификацию любой услуги. Для этого необходимо подставить значения параметров предоставления услуги в дискриминантные уравнения для каждой  $k$ -ой группы и рассчитать значения оценок отклика, при  $k = 1, 2, \dots, p$ . Услуга классифицируется как принадлежащая классу  $k$ , для которого величина  $d_k$  (дискриминантные функции  $d_k$  основаны на максимуме отношения межгрупповой вариации к внутригрупповой) имеет максимальное значение [4].

Результаты пошагового дискриминантного анализа для решения поставленной задачи:

- последняя включенная переменная — «Информационная скорость доступа», со значением статистики F-критерия:  $F(9.526) = 20.19004$ , уровень значимости:  $p < 0.0000$ ;
- значение лямбды Уилкса: 0.0740269, приближенное значение F-статистики, связанной с лямбдой Уилкса:  $F(27.1536) = 81.88042$ , уровень значимости:  $p < 0.0000$ .

Значения статистики лямбды Уилкса лежат в интервале  $[0;1]$ . Значения, лежащие около нуля, свидетельствуют о хорошей дискриминации (что и наблюдается в нашем случае), то есть мощность дискриминации близка к единице.

В результате вычислений получена классификационная матрица и сделан вывод, что не все объекты были правильно отнесены к выделенным группам. Анализируя правильность соотношения «вида передаваемой информации» к группам был уменьшен размер выборки (удалены противоречивые и единичные экстремумы). Эти действия повторялись до тех пор, пока процент правильности в матрице классификации не составил сто процентов (таблица 1).

В результате применения пошагового дискриминантного анализа с включением и с результатами изменения выборки, получим:

- последняя включенная переменная — «Информационная скорость доступа», соответствующее значение статистики F-критерия:  $F(5.327) = 615.6304$ , уровень значимости:  $p < 0.0000$ ;
- значение лямбды Уилкса: 0.0000008, приближенное значение F-статистики, связанной с лямбдой Уилкса:  $F(15.903) = 9559.37$ , уровень значимости:  $p < 0.0000$ .

В таблице 2 показаны переменные, включенные в модель.

Так как рассмотренные виды передаваемой информации не являются полной группой, выполнен переход к общему названию классов:

Таблица 1 — Окончательная матрица классификации

	Процент правильности	Текст, неподвижные изображения p=.08358	Речь p=.23881	Речь, подвижные изображения p=.17313	Данные p=.36418	Датаграммы p=.11940	Передача разнородной информации p=.02090
Текст, неподвижные изображения	100.00	28	0	0	0	0	0
Речь	100.00	0	80	0	0	0	0
Речь, подвижные изображения	100.00	0	0	58	0	0	0
Данные	100.00	0	0	0	122	0	0
Датаграммы	100.00	0	0	0	0	40	0
Передача разнородной информации	100.00	0	0	0	0	0	7
Всего	100.00	28	80	58	122	40	7

Таблица 2 — Анализ дискриминантных функций

Размер выборки: N=335	Лямбда Уилкса	Частная лямбда	F-критерий (9,526)	p-уровень	Толерантность	1- толерантность (R-Sqr.)
Готовность сети	0,044554	0,000019	3526746	0,00	0,028821	0,971179
Время доступа к службе	0,000056	0,014810	4351	0,00	0,028821	0,971179
Информационная скорость доступа	0,000009	0,096031	616	0,00	0,999996	0,000004

- «Текст, неподвижные изображения» — «Класс 1»;
- «Речь» — «Класс 2»;
- «Речь, подвижные изображения» — «Класс 3»;
- «Данные» — «Класс 4»;
- «Датаграммы» — «Класс 5»;
- «Передача разнородной информации» — «Класс 6»;

В результате применения дискриминантного анализа получены функции классификации, приведенные в таблице 3. Каждый класс характеризуется своей функцией классификации. Поэтому, любую ИКУ можно отнести к одному из классов, подставив параметры предоставления этой услуги в найденные классификационные формулы и определив максимум из полученных результатов (услуга принадлежит к тому классу, в котором результат расчета классификационной формулы максимален).

С помощью этих функций можно вычислить классификационные значения для новых услуг по формулам:

$$\begin{aligned}
 \text{Класс 1} &= 1009587.542 \cdot r + 7376.42011 \cdot t - 0.00068 \cdot v - 48010733.850; \\
 \text{Класс 2} &= 1061876.336 \cdot r + 7758.28017 \cdot t - 0.00072 \cdot v - 53112682.711; \\
 \text{Класс 3} &= 1009199.763 \cdot r + 7373.50287 \cdot t - 0.00068 \cdot v - 47973857.657; \\
 \text{Класс 4} &= 1058661.329 \cdot r + 7735.09141 \cdot t - 0.00071 \cdot v - 52791558.362; \\
 \text{Класс 5} &= 1008811.992 \cdot r + 7370.58570 \cdot t - 0.00068 \cdot v - 47936997.927; \\
 \text{Класс 6} &= 1058668.730 \cdot r + 7735.29855 \cdot t + 0.00002 \cdot v - 52792544.131;
 \end{aligned}
 \tag{2}$$

где:  $r$  — готовность сети;  
 $t$  — время доступа к службе;  
 $v$  — информационная скорость доступа.

На основании полученных формул классификации сформируем классы ИКУ (таблица 4), где  $r$  — готовность сети;  $t$  — время доступа к службе;  $v$  — информационная скорость доступа.

Таблица 3 — Функции классификации<sup>1</sup>

	Текст, неподвижные изображения $p=.08358$	Речь $p=.23881$	Речь, подвижные изображения $p=.17313$	Данные $p=.36418$	Датаграммы $p=.11940$	Передача разнородной информации $p=.02090$
Готовность сети	1009587.542	1061876.336	1009199.763	1058661.329	1008811.992	1058668.730
Время доступа к службе	7376.42011	7758.28017	7373.50287	7735.09141	7370.58570	7735.29855
Информационная скорость доступа	-0.00068	-0.00072	-0.00068	-0.00071	-0.00068	0.00002
Константа	-48010733.85	-53112682.71	-47973857.66	-52791558.36	-47936997.93	-52792544.13

Таблица 4 — Классы ИКУ

Класс 1	$1009587.542 \cdot r + 7376.42011 \cdot t - 0.00068 \cdot v - 48010733.850$
Класс 2	$1061876.336 \cdot r + 7758.28017 \cdot t - 0.00072 \cdot v - 53112682.711$
Класс 3	$1009199.763 \cdot r + 7373.50287 \cdot t - 0.00068 \cdot v - 47973857.657$
Класс 4	$1058661.329 \cdot r + 7735.09141 \cdot t - 0.00071 \cdot v - 52791558.362$
Класс 5	$1008811.992 \cdot r + 7370.58570 \cdot t - 0.00068 \cdot v - 47936997.927$
Класс 6	$1058668.730 \cdot r + 7735.29855 \cdot t + 0.00002 \cdot v - 52792544.131$

С помощью классификационных формул можно определить класс любой ИКУ. Каждый класс содержит требования к сетям и оборудованию [1]. Поэтому, на основании формул (2) можно определить, какие требования должны выполняться сетью для предоставления каждой услуги или групп услуг. Анализ полученных требований определит, способна ли рассматриваемая телекоммуникационная сеть предоставлять необходимый набор услуг, что позволит, например, операторам связи принять решение о предоставлении новых услуг или о необходимости модернизации существующей сети (сегмента сети). Зная требования услуг к сетям и оборудованию можно планировать предоставление услуг на заданной телекоммуникационной сети.

**Выводы.** В результате применения пошагового дискриминантного анализа к параметрам предоставления ИКУ получен набор классов, представленный в таблице 4.

Полученные результаты проверены пошаговым дискриминантным анализом со стандартным методом (соответствие полное). Для соотношения новой услуги к одному из обозначенных классов, необходимо подставить значения параметров предоставления в приведенные выше формулы. Новая услуга относится к тому классу, для которого классификационное значение максимально.

Рассмотрим пример. Определим класс услуги «Передача данных по телефонной сети общего пользования». Параметры этой услуги следующие: информационная скорость доступа — 33.6 кбит/с; готовность сети — 99.5%; время доступа к службе — 15 сек.

<sup>1</sup> Приведенные данные округлены до тысячных с целью экономии места.

Вычислим значения классификационных функций для данной услуги (расчет производим по формулам таблицы 4):

$$\begin{aligned} \text{Класс1} &= 1009587.542 \cdot 99.5 + 7376.42011 \cdot 15 - 0.00068 \cdot 33.6 - 48010733.850 = \\ &= 52553872.85780201 \end{aligned}$$

$$\begin{aligned} \text{Класс2} &= 1061876.336 \cdot 99.5 + 7758.28017 \cdot 15 - 0.00072 \cdot 33.6 - 53112682.711 = \\ &= 52660386.89935798 \end{aligned}$$

$$\begin{aligned} \text{Класс3} &= 1009199.763 \cdot 99.5 + 7373.50287 \cdot 15 - 0.000682 \cdot 33.6 - 47973857.657 = \\ &= 52552121.28170202 \end{aligned}$$

$$\begin{aligned} \text{Класс4} &= 1058661.329 \cdot 99.5 + 7735.09141 \cdot 15 - 0.00071 \cdot 33.6 - 52791558.362 = \\ &= 52661270.22079399 \end{aligned}$$

$$\begin{aligned} \text{Класс5} &= 1008811.992 \cdot 99.5 + 7370.58570 \cdot 15 - 0.00068 \cdot 33.6 - 47936997.927 = \\ &= 52550354.03965201 \end{aligned}$$

$$\begin{aligned} \text{Класс6} &= 1058668.730 \cdot 99.5 + 7735.29855 \cdot 15 + 0.00002 \cdot 33.6 - 52792544.131 = \\ &= 52661023.98157801 \end{aligned}$$

Расчеты показывают, что при подстановке значений в функцию «Класса 4», результат максимален. Следовательно, услуга «Передача данных по телефонной сети общего пользования» соотносится к «Классу 4» (действительно, для «Класса 4» характерна передача данных).

Таким образом, можно сопоставить с полученными классами любую существующую либо прогнозируемую услугу. Соотнесение к классам услуг, по формулам, приведенным в таблице 4, проверено программным путем. Для этого было разработано программное приложение, определяющее класс услуги по введенным данным на основании указанных формул. Результат соотнесения идентичен.

### Литература

1. Гайворонская, Г. С. Классификация инфокоммуникационных услуг и разработка требований к сети для их предоставления [Текст] / Г.С. Гайворонская, С.В. Павлов // Сб. науч. тр. IV семинара «Информационные системы и технологии». (Приложение к журналу Холодильна техніка і технологія). — Одесса : ОДАХ, 2006. — С. 12–13.
2. International Telecommunication Union [Electronic resource] / Интернет-ресурс. — Режим доступа: [www/ URL: http://www.itu.int](http://www.itu.int). — Загл. с экрана.
3. Павлов, С.В. О размерности пространства параметров предоставления инфокоммуникационных услуг [Текст] / С. В. Павлов // Зв'язок, №1. — 2008. — С. 59.
4. Прикладная статистика: Классификации и снижение размерности. Справ, изд. [Текст] / С.А. Айвазян, В.М. Бухштабер, И.С. Енюков и др.: под ред. С. А. Айвазяна. — М.: Финансы и статистика, 1989. — 607 с.
5. Павлов, С.В. Снижение размерности параметров предоставления инфокоммуникационных услуг методом факторного анализа [Текст]/С.В. Павлов//III Міжнародна науково-технічна конференція «Сучасні інформаційно-комунікаційні технології»: Тезиси. — Киев: ДУІКТ, 2007. — С. 63.
6. Павлов, С.В. Классификация инфокоммуникационных услуг методом дискриминантного анализа [Текст] / С.В. Павлов // Материалы Международной конференции «Проблемы телекоммуникаций – 2008»: Тезиси.— Киев: КПИ, 2008. — С. 340.
7. Электронный учебник StatSoft [Электронный ресурс] / Электронный учебник по статистике. — Internet-ресурс: Режим доступа: [www/ URL: http://www.statsoft.ru/home/textbook/](http://www.statsoft.ru/home/textbook/) — Загл. с экрана.

Здано в редакцію:  
23.02.2009р.

Рекомендовано до друку:  
д.т.н, проф. Чичикало Н.І.