

## ИДЕНТИФИКАЦИЯ ВЕРОЯТНОСТНЫХ ХАРАКТЕРИСТИК ЗАКОНОВ РАСПРЕДЕЛЕНИЯ ДЛЯ ИМИТАЦИОННОЙ МОДЕЛИ СЛУЖБЫ SERVICE DESK

С. А. Крол, В. Д. Алёшин (Москва)

В настоящем докладе представлены результаты исследований круглосуточной службы технической поддержки пользователей банковских систем автоматизации (далее – служба Service Desk). Далее излагаются результаты идентификации как законов распределения (входного потока заявок, времён их исполнения), так и их численных значений. Исследования были начаты во Внешэкономбанке в период его преобразования в государственную корпорацию (июнь–август 2007 года). Заказчиком исследования выступал ИТ-директор банка.

Полученные результаты являлись частью работ по выработке рекомендаций по реорганизации службы Service Desk и нахождению оптимальных параметров службы в условиях прогнозируемого увеличения нагрузки на ИТ-инфраструктуру Внешэкономбанка. Рекомендации по реорганизации службы предполагалось получить на основе построения и исследования имитационной модели службы Service Desk. Для этой модели и решалась задача идентификации законов распределения и их параметров.

### Исходные данные

В качестве исходных данных выступали фиксировавшиеся службой Service Desk обращения в службу и времена исполнения обращений. Длина временного ряда – 3 года. Для идентификации законов распределения и времен обработки обращений использовался отрезок временного ряда обращений в службу Service Desk за 2006 год. Первичная верификация имитационной модели проводилась на отрезке временного ряда данных с 01.01.2007 г. по 01.04.2007 г.

### Выбор и методика оценки вероятностных характеристик моделей

В качестве основных вероятностных характеристик имитационной модели службы Service Desk, рассмотренной как система массового обслуживания, были выделены:

- **интенсивность входного потока заявок** (длины интервалов между обращениями пользователей ИТ-систем в Service Desk)
- **временные характеристики обработки заявок в подразделениях первой и второй линии поддержки.**

Подразделением первой линии поддержки во Внешэкономбанке является т.н. диспетчерская служба, подразделениями второй линии – шесть функциональных отделов ИТ-департамента. Исходная организация службы не запрещала обращения пользователей непосредственно в отделы второй линии поддержки; часть инцидентов устранялась диспетчерской службой. В этой связи идентификации при создании моделей подлежали семь параметров интенсивности поступления заявок и семь временных параметров обработки заявок.

При решении этой задачи для каждого из указанных четырнадцати параметров была применена методика, предусматривающая следующие шаги:

- предварительный выбор семейства теоретического распределения;
- оценку параметров (положения, масштаба и формы) этих распределений;
- проверку гипотезы о правильности сделанного выбора;
- верификацию полученных результатов с помощью сравнения расчетных и реальных данных.

### Оценка интенсивности входного потока заявок

Первый результат, полученный в ходе оценки интенсивности входного потока заявок, – необходимость выделения из исходного временного ряда неслучайной компоненты. На рис. 1. представлены длины интервалов между поступлением заявок в 2006 году. Можно заметить регулярно возникающие пики значительной длины. Указанные пики соответствовали интервалам между последней принятой заявкой рабочего дня банка (09:00 – 18:00) и первой поступившей заявкой следующего дня. При этом, однако, имелось ненулевое количество заявок, принятых в нерабочее время.

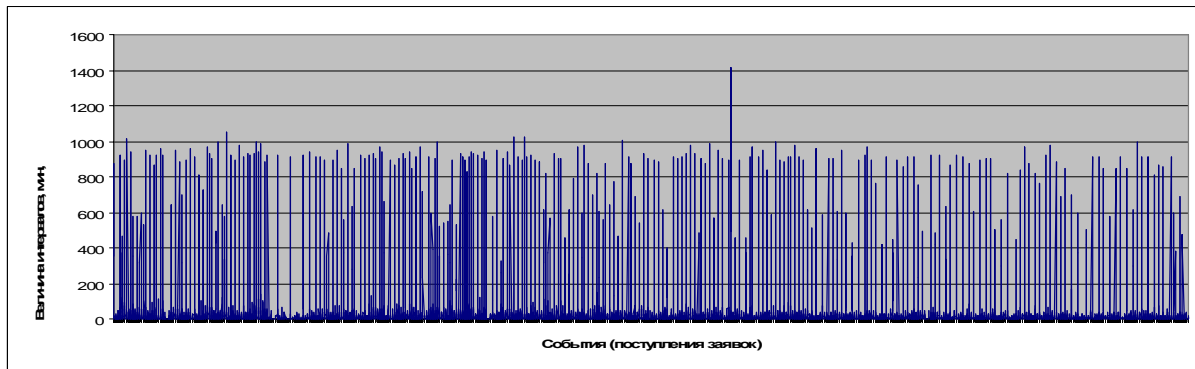


Рис. 1. Длины интервалов между поступлением заявок в 2006 г.

В этой связи исходный временной ряд был разбит на два. Первый временной ряд включал обращения в службу Service Desk в рабочее время (с 09:00 до 18:00), второй – обращения в службу в нерабочее время (с 18:01 до 08:59 следующего дня). Идентификация законов распределения осуществлялась отдельно для каждой последовательности данных. Таким образом, количество параметров модели, отражающих интенсивность поступления заявок, увеличилось с семи до четырнадцати, однако длительности интервалов между поступлением приобрели искомый случайный характер.

За первоначальную гипотезу о вероятностном распределении длин интервалов между поступлением в Service Desk заявок, в соответствии с теоретическими рекомендациями, было принято предположение об экспоненциальном характере этого распределения для каждого из адресатов поступления заявок, как в дневное, так и в ночное время.

Для проверки указанных четырнадцати гипотез на основании полученных за 2006 год данных в пакете STATISTICA строились и анализировались с помощью критерия Колмогорова-Смирнова гистограммы плотности распределения интервалов между поступлением заявок (рис. 2). Результаты проверки гипотез приведены в табл. 1, 2.

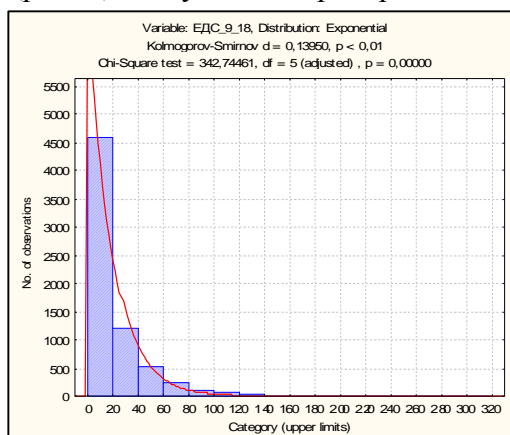


Рис.2. Результаты проверки гипотезы об экспоненциальном распределении поступивших заявок (диспетчерская служба) в рабочее время в 2006 г.

Таблица 1

**Результаты проверки гипотезы об экспоненциальном распределении заявок, поступивших в службу Service Desk в 2006 г. в рабочее время**

Функциональное подразделение	Результаты подбора закона распределения		
	Среднее значение длины интервала между поступлением заявок в 2006 г., мин	Значение критерия Колмогорова-Смирнова	Гипотеза об экспоненциальном характере распределения принята?
Диспетчерская служба	<b>19,21</b>	<b>0.13950</b>	Да
Отдел 1	<b>63,30</b>	<b>0.51992</b>	Да
Отдел 2	<b>123,30</b>	<b>0.20109</b>	Да
Отдел 3	<b>151,8</b>	<b>0.17755</b>	Да
Отдел 4	Заявок не было	–	–
Отдел 5	<b>49,68</b>	<b>0.20293</b>	Да
Отдел 6	Заявок не было	–	–

Таблица 2

**Результаты проверки гипотезы об экспоненциальном распределении заявок, поступивших в службу Service Desk в 2006 г. в нерабочее время**

Функциональное подразделение	Результаты подбора закона распределения		
	Среднее значение длины интервала между поступлением заявок в 2006 г., мин	Значение критерия Колмогорова-Смирнова	Гипотеза об экспоненциальном характере распределения принята?
Диспетчерская служба	Заявок не было	–	–
Отдел 1	<b>395,00</b>	<b>0.18618</b>	Да
Отдел 2	Заявок не было	–	–
Отдел 3	Заявок не было	–	–
Отдел 4	Заявок не было	–	–
Отдел 5	<b>241,91</b>	<b>0.2858</b>	Да
Отдел 6	Заявок не было	–	–

Как следует из приведенных таблиц, принятыми оказались все гипотезы об экспоненциальном характере распределения интервалов между поступлением заявок (в случае, если таковые заявки поступали).

**Оценка временных характеристик обработки заявок**

В соответствии с теорией массового обслуживания и практикой имитационного моделирования теоретическим законом, "подозрительным" с точки зрения успешной

применимости для оценки распределения вероятностной переменной вида "время выполнения какой-либо задачи", является *гамма-распределение*. Масштабный параметр  $\beta$  и параметр формы  $\alpha$  этого распределения связаны между собой двумя уравнениями:

$$\left\{ \begin{array}{l} \alpha\beta = \mu, \quad \text{где } \mu \text{ – среднее значение;} \\ \alpha\beta^2 = \sigma^2, \quad \text{где } \sigma^2 \text{ – дисперсия;} \end{array} \right. \quad (1)$$

Решение этой системы уравнений позволяет определить максимально правдоподобные значения параметров гамма-распределения.

Для проверки семи гипотез о применимости гамма-распределения для оценки времени обработки заявок в семи подразделениях, входящих в первую и вторую линию поддержки Service Desk, были использованы соответствующие данные за 2006 год и так же, как и в предыдущем случае, применен критерий согласия Колмогорова-Смирнова.

В табл. 3 и 4 приведены результаты проверки указанных гипотез и рассчитанные на основе формул (1) и (2) параметры распределений.

Таблица 3

**Результаты проверки гипотезы о гамма-распределении длительностей обработки заявок, поступивших в службу Service Desk в 2006 г.**

Функциональное подразделение	Результаты подбора закона распределения	
	Значение критерия Колмогорова-Смирнова	Гипотеза об гамма – распределении принята?
Диспетчерская служба	0,42774	Да
Отдел 1	0,43048	Да
Отдел 2	0,69746	Да
Отдел 3	0,63228	Да
Отдел 4	0,59954	Да
Отдел 5	0,71291	Да
Отдел 6	0,23075	Да

Таблица 4

**Результаты проверки гипотезы о гамма-распределении длительностей обработки заявок, поступивших в службу Service Desk в 2006 г.**

Функциональное подразделение	Параметры гамма – распределения			
	$\mu$	$\sigma^2$	$\alpha$	$\beta$
Диспетчерская служба	33	8978	0,12	272
Отдел 1	51	24023	0,11	471
Отдел 2	115	520286	0,025	4524
Отдел 3	128	249575	0,066	1950
Отдел 4	679	7549535	0,061	11119
Отдел 5	145	565567	0,037	3900
Отдел 6	2517	9410679	0,67	3739

### Использование полученных результатов при построении модели в системе AnyLogic

В имитационных моделях AnyLogic текущей и целевой реализации службы Service Desk приведенные выше результаты оценки вероятностных распределений случайных параметров были отражены:

- при установке атрибутов *interarrivalTime* для семи объектов класса *Source*, служащих для порождения входных транзактов. В качестве единственного параметра  $\beta$  использованной функции *exponential* ( $\beta$ ) взято среднее за 2006 год значение длины интервала между поступлением заявок;
- при установке атрибутов *delayTime* для семи блоков класса *ProcessQ*, реализующих в модели процесс обработки заявок в подразделениях. Значения найденных параметров сведены в табл. 5.

Таблица 5

#### Сводные результаты подбора теоретических вероятностных распределений для статистических параметров модели

Статистический параметр модели	Наименование и параметры распределения	
	Рабочее время	Нерабочее время
Диспетчерская служба. Входной поток	expo(1/19.21 )	–
Отдел 1. Входной поток	expo(1/63.30)	–
Отдел 2. Входной поток	expo(1/123.30)	expo(1/395)
Отдел 3. Входной поток	expo(1/151.8)	–
Отдел 4. Входной поток	–	–
Отдел 5. Входной поток	expo(1/49.68 )	expo(1/241.91)
Отдел 6. Входной поток	–	–
Диспетчерская служба. Время обработки	gamma(0.12, 272)	
Отдел 1. Время обработки	gamma(0.11, 471)	
Отдел 2. Время обработки	gamma(0.025, 4524)	
Отдел 3. Время обработки	gamma(0.066, 1950)	
Отдел 4. Время обработки	gamma(0.061, 11119)	
Отдел 5. Время обработки	gamma(0.037, 3900)	
Отдел 6. Время обработки	gamma(0.67, 3739)	

#### Проверка адекватности полученных оценок вероятностных характеристик модели

Для проверки полученных результатов были использованы статистические данные о приеме и обработке заявок, собранные службой Service Desk в первом квартале 2007 года, и созданная имитационная модель текущей реализации службы. Количественные результаты, полученные на модели и в реальности, были сравнены.

Результаты сравнения, приведенные в табл. 6, показали, что расхождения между модельными и реальными показателями не превысили 25%.

Таблица 6

## Сравнительная оценка реальных и полученных на модели результатов работы службы Service Desk в 1 квартале 2007 г.

Количественный результат работы Service Desk в 1 кв. 2007 г.	Зарегистрировано	Получено на модели	Расхождение, %
Всего заявок поступило	<b>4429</b>	<b>3836</b>	<b>-13.4 %</b>
из них в диспетчерскую службу	1919	1646	-14.2 %
в отдел 1	460	447	-2.8 %
в отдел 2	400	480	+20.0 %
в отдел 3	0	1	–
в отдел 4	190	190	0.0 %
в отдел 5	1186	1071	-9.7 %
в отдел 6	1	1	–
Всего заявок обработано	<b>4383</b>	<b>3820</b>	<b>-12.8 %</b>
из них в диспетчерской службе	260	204	-21.5 %
в отделе 1	2029	1853	-10.3 %
в отделе 2	372	405	-8.7 %
в отделе 3	103	79	-23.3 %
в отделе 4	249	217	-12.9 %
в отделе 5	1364	1055	-22.7 %
в отделе 6	6	7	+16.7 %

Приведенные результаты позволяют сделать следующие выводы:

- применимость способа выделения из исходного временного ряда неслучайной компоненты путем раздельного рассмотрения входных потоков заявок на дневном и ночном интервалах работы службы Service Desk подтверждена;
- идентифицированные законы распределения имеют место;
- подобранные параметры распределений позволяют обеспечить адекватность имитационной модели на уровне 25%.

Дальнейшее уточнение характеристик модели осуществлялось по мере получения статистических данных службы Service Desk за 2007–2008 гг.

## 7. Выводы

Проведенное исследование продемонстрировало, каким образом при построении имитационной модели службы Service Desk могут быть идентифицированы основные вероятностные характеристики. Предложенная методика (разделение дневных и ночных потоков входных заявок Service Desk, выдвижение и верификация гипотез о теоретических распределениях и оценка их параметров), будучи использована как инструмент при решении аналогичных задач, позволит сократить временные затраты на моделирование.