

А. А. Фурса¹

аспирант кафедры экономической кибернетики Санкт-Петербургского государственного университета

ИМИТАЦИОННАЯ МОДЕЛЬ ОЦЕНКИ ДОСТАТОЧНОСТИ ЧИСЛЕННОСТИ СОТРУДНИКОВ ОТДЕЛЕНИЯ БАНКА

В филиалах и дополнительных офисах коммерческого банка (далее — ДО КБ) периодически возникает задача оценки адекватности численности работающего персонала объему и сложности выполняемой работы и при необходимости дальнейшей корректировки данной численности. Особенно актуальной становится подобная проблема в период кризиса и выхода из него, когда любое предприятие стремится к разумному сокращению издержек, но при этом, сильно уменьшив численность сотрудников, банк может потерять клиентов из-за ухудшения качества обслуживания. Значительная часть руководителей решает эти задачи на основании собственных экспертных оценок либо опираясь на некие нормативы, устанавливаемые Институтом труда, Министерством труда и социального развития и другими учреждениями. Количественные методы (см., напр., Колосова, 2008, Одегов, 2005) также находят свое применение, однако существуют трудности при оценке параметров, необходимых для принятия решения. Большинство упомянутых методов не относится к экономико-математическим методам. В то же время для обоснования решения о сохранении или изменении численности персонала возможно применение математических методов — как дополнительного или даже основного аргумента в пользу того или иного решения.

Часто при анализе деятельности систем, подобных ДО КБ, используется аппарат теории массового обслуживания (ТМО). Однако при видимых достоинствах (простота расчетов) для получения корректных результатов при применении моделей, базирующихся на использовании ТМО, необходимо выполнение ряда предпосылок (в частности, о том, что поток клиентов — простейший и время между приходом клиентов распределено по показательному закону) (см., напр., Вентцель, 2003, Чернов, 2000). В действительности экспериментальными данными указанные предпосылки могут не подтверждаться (см., напр., Васин, 2004). В таком случае возможно применение к рассматриваемой ситуации аппарата имитационного моделирования. В данной статье мы опишем разработанную нами имитационную модель деятельности универсального ДО КБ, результатом применения которой является расчет показателей, позволяющих сделать вывод об адекватности численности сотрудников конкретных категорий уровню их текущей загрузки, на основании которых может быть принято решение об изменении численности персонала.

¹ Эл. адрес: annafursa@bk.ru

Опишем исследуемую систему в терминах производственного менеджмента: клиенты ДО КБ выступают в качестве деталей (предметов труда), которым требуется та или иная «обработка» (в зависимости от совершаемой операции), а работники разных категорий выступают в роли станков разных видов, причем способных выполнять виды работы и не требующие «переналадки» (переход от выполнения одной функции при обслуживании одного клиента к выполнению иной функции при обслуживании следующего клиента осуществляется мгновенно). Присвоим различным типам работников индексы $j = 1, \dots, 9$ (табл. 1).

Таблица 1

**Перечень категорий работников универсального ДО,
деятельность которых описывается в рамках модели**

Индекс категории сотрудников (j)	Наименование категории сотрудников
1	Менеджер по обслуживанию физических лиц (МОФЛ)
2	Кассир (КАС)
3	Контролер (КОН)
4	Специалист по работе с клиентами (СРК)
5	Кредитный работник (КР)
6	Администратор кредитных операций (АКО)
7	Операционный работник (ОР)
8	Бухгалтер последующего контроля (БПК)
9	Менеджер по обслуживанию юридических лиц (МОЮЛ)

После уточнения перечня категорий сотрудников можно формализовать множество операций выполняемых работниками каждой категории (множества $M^1, \dots, M^9, M^j = \{m_{ij}\}, i_j = 1, \dots, I_j$ соответственно). Каждой цели обращения клиентов в ДО можно сопоставить строго определенную последовательность обработки обращения работниками разных типов. В соответствии с целями клиенты разбиваются на группы, в рамках которых клиенты неразличимы. Обозначим число целей обращения клиентов — физических лиц как $K^{\text{ФЛ}}$, а клиентов — юридических лиц как $K^{\text{ЮЛ}}$.

Для того чтобы произвести моделирование деятельности универсального ДО, нам требуется знание следующих параметров:

- интенсивности входящих потоков клиентов всех разновидностей (как юридических, так и физических лиц) ($\gamma_k^{\text{ФЛ}}, k = 1, \dots, K^{\text{ФЛ}}, \gamma_k^{\text{ЮЛ}}, k = 1, \dots, K^{\text{ЮЛ}}$);
- плотности распределения соответствующих длительностей обслуживания при обращении к различным группам специалистов для выполнения различных операций (см. табл. 1) ($f_{ij}^i, j = 1, \dots, 9, i_j = 1, \dots, I_j$);
- количество работников каждой категории ($n_j, j = 1, \dots, 9$);
- ограничение на длину очереди к работникам, непосредственно работающим с клиентами (МОФЛ, СРК, МОЮЛ) ($m_j, j = 1, 4, 9$);
- среднее количество клиентов каждого типа на момент начала функционирования системы ($\Lambda_0^{\text{ФЛ}}, \Lambda_0^{\text{ЮЛ}}$).

Для начала моделирования также нужно задать параметры моделирования, а именно: число имитируемых рабочих дней (D), продолжительность рабочего дня для обслуживания юридических и физических лиц ($T^{\text{ЮЛ}}, T^{\text{ФЛ}}$, как правило, $T^{\text{ЮЛ}} < T^{\text{ФЛ}}$). Опишем в общем виде схему моделирования работы системы (рис. 1—4).

В блоке 1 осуществляется переход к расчетам по очередному имитируемому рабочему дню. Для первого дня устанавливается номер дня $d_l = 1$, при переходе от предыдущего имитируемого дня к следующему устанавливается $d_l = d_l + 1$, для последнего дня $d_l = D$.

В блоке 2 для каждого вида клиентов (юридические и физические лица) в исходной очереди $(\Lambda_0^{\text{ФЛ}}, \Lambda_0^{\text{ЮЛ}})$ определяется количество клиентов каждого типа в соответствии с долями клиентов каждой разновидности, поступающих на об-

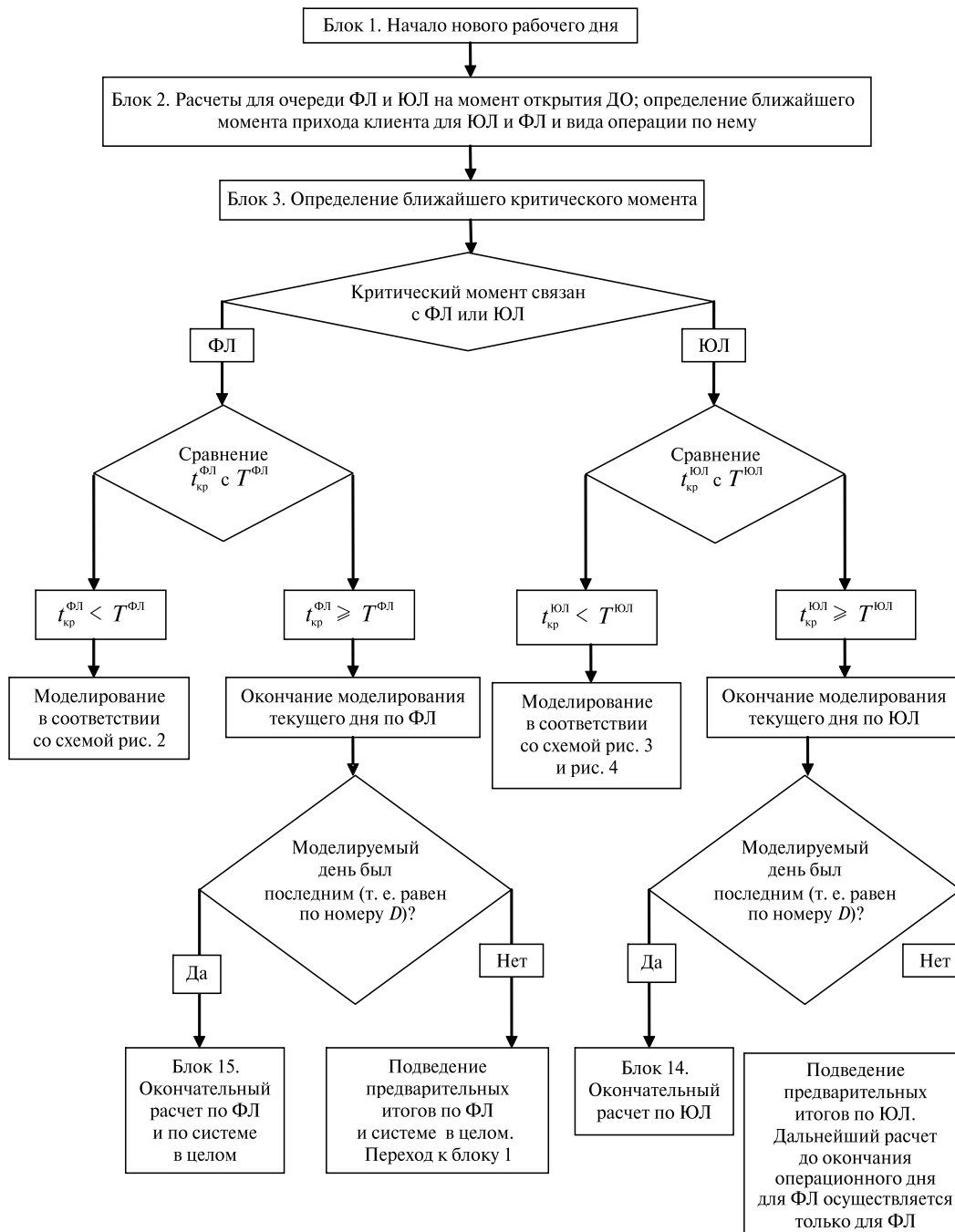


Рис. 1. Общая схема моделирования процесса работы универсального ДО КБ

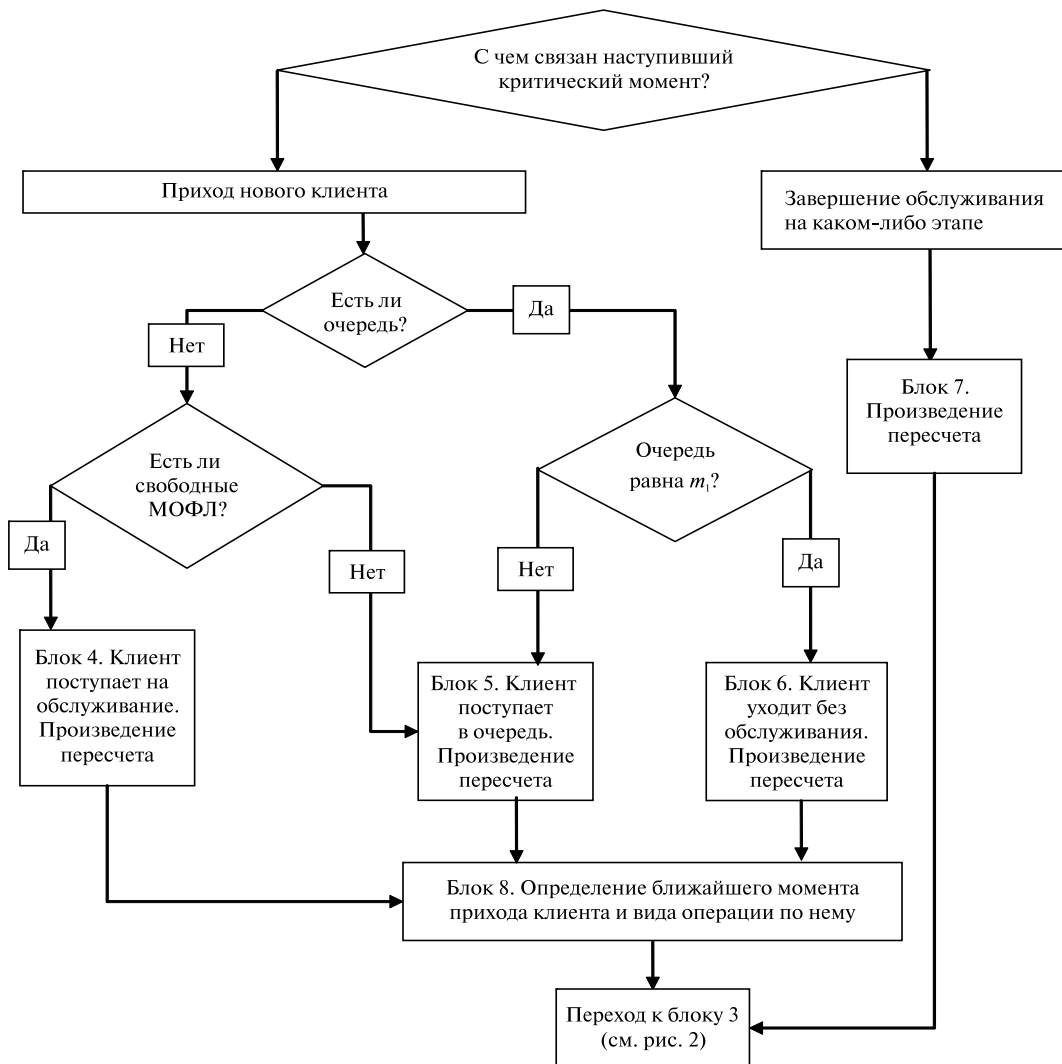


Рис. 2. Схема моделирования процесса работы универсального ДО КБ, связанной с обслуживанием физических лиц

служивание к МОФЛ ($p_k^{\text{ФЛ}}$, $k = 1, \dots, K^{\text{ФЛ}}$), МОЮЛ ($p_k^{\text{ЮЛ}}$, $k \in K_1 = \{1, \dots, K^{\text{ЮЛ}}\}$) и СРК ($p_k^{\text{ЮЛ}}$, $k \in K_2 = \{1, \dots, K^{\text{ЮЛ}}\}$, $K_1 \cup K_2 = \{1, \dots, K^{\text{ЮЛ}}\}$, $K_1 \cap K_2 = \emptyset$):

$$p_k^{\text{ФЛ}} = \frac{\gamma_k^{\text{ФЛ}}}{\sum_{k=1}^{K^{\text{ФЛ}}} \gamma_k^{\text{ФЛ}}}, \quad (1)$$

$$p_k^{\text{ЮЛ}} = \frac{\gamma_k^{\text{ЮЛ}}}{\sum_{k=1}^{K^{\text{ЮЛ}}} \gamma_k^{\text{ЮЛ}}}. \quad (2)$$

Клиенты — юридические лица делятся на тех, кто поступает на обслуживание к МОЮЛ, и тех, кто поступает на обслуживание к СРК:

$$\Lambda_0^{\text{СРК}} = \Lambda_0^{\text{ЮЛ}} \sum_{j \in K_2} p_k^{\text{ЮЛ}}, \quad (3)$$

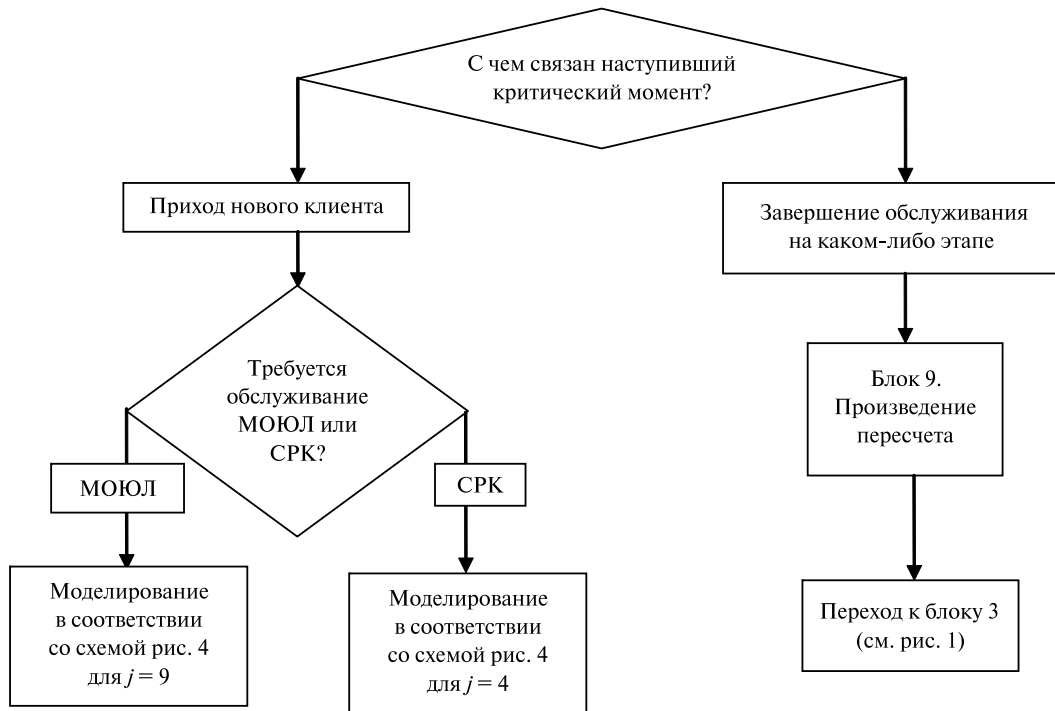


Рис. 3. Схема моделирования процесса работы универсального ДО КБ, связанной с обслуживанием юридических лиц

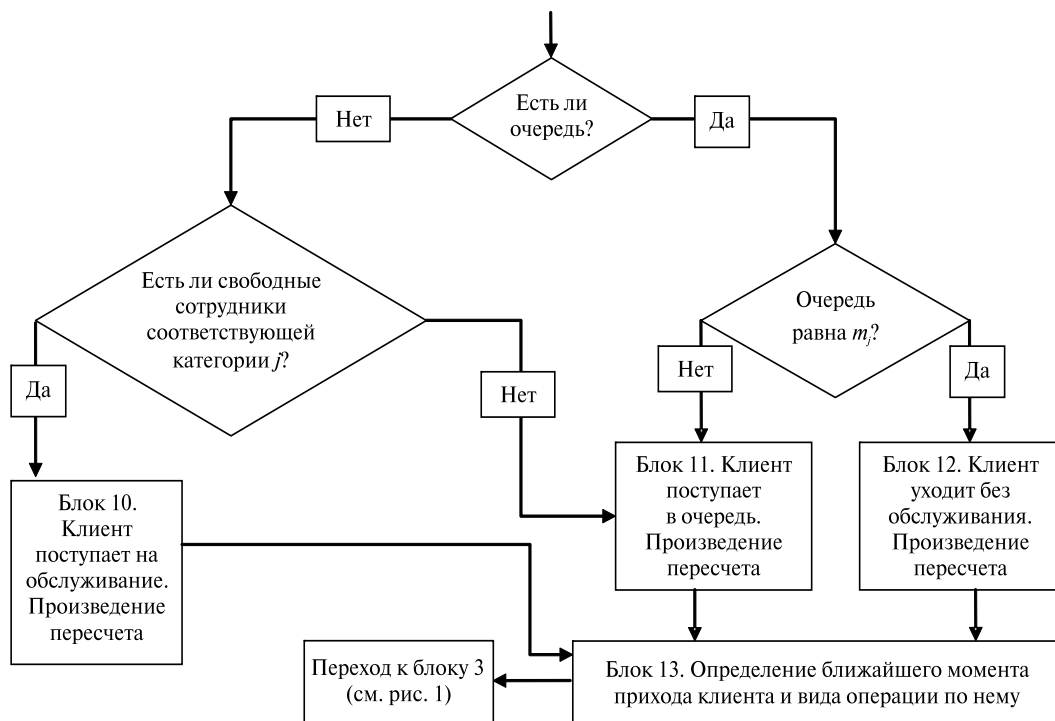


Рис. 4. Схема моделирования процесса работы универсального ДО КБ, связанной с обработкой вновь поступивших заявок от клиентов — юридических лиц (моделируется деятельность СРК либо МОЮЛ)

$$\Lambda_0^{\text{МОЮЛ}} = \Lambda_0^{\text{ЮЛ}} \sum_{j \in K_1} p_k^{\text{ЮЛ}}. \quad (4)$$

Устанавливается стартовое для текущего дня количество вызовов в системе ($N_l^{\text{ФЛ}} = \Lambda_0^{\text{ФЛ}}$, $N_l^{\text{ЮЛ}} = \Lambda_0^{\text{ЮЛ}}$). Входящие заявки могут быть адресованы трем категориям работников — МОФЛ, СРК, МОЮЛ ($j = 1, 4, 9$). Для каждой категории j осуществляется следующая процедура: определяется, куда попадают поступившие вызовы:

- Если $\Lambda_0^{\text{МОЮЛ}} \leq n_1$ ($\Lambda_0^{\text{СРК}} \leq n_4$, $\Lambda_0^{\text{МОЮЛ}} \leq n_9$), все вызовы сразу попадают на обслуживание; в соответствии с плотностями распределения соответствующих длительностей обслуживания рассчитывается время их обслуживания; для свободных МОФЛ (СРК, МОЮЛ) фиксируется начало времени их простоя; отказов в системе нет: $N_{\text{отк}_j}^{\text{ФЛ}} = 0$ ($N_{\text{отк}_j}^{\text{ЮЛ}} = 0$).

- Если $n_1 < \Lambda_0^{\text{МОЮЛ}} \leq n_1 + m_1$ ($n_4 < \Lambda_0^{\text{СРК}} \leq n_4 + m_4$, $n_9 < \Lambda_0^{\text{МОЮЛ}} \leq n_9 + m_9$), то n_1 (n_4 , n_9) вызовов сразу попадают на обслуживание, рассчитывается время их обслуживания; $(\Lambda_0^{\text{ФЛ}} - n_1)$ ($(\Lambda_0^{\text{СРК}} - n_4)$, $(\Lambda_0^{\text{МОЮЛ}} - n_9)$) вызовов попадают в очередь, фиксируются моменты начала их пребывания в очереди; отказов в системе нет: $N_{\text{отк}_j}^{\text{ФЛ}} = 0$ ($N_{\text{отк}_j}^{\text{ЮЛ}} = 0$).

- Если $\Lambda_0^{\text{МОФЛ}} > n_1 + m_1$ ($\Lambda_0^{\text{СРК}} > n_4 + m_4$, $\Lambda_0^{\text{МОЮЛ}} > n_9 + m_9$), n_1 (n_4 , n_9) вызовов сразу попадают на обслуживание, рассчитывается время их обслуживания, m_1 (m_4 , m_9) вызовов попадают в очередь, фиксируются моменты начала их пребывания в очереди, $(\Lambda_0^{\text{МОФЛ}} - n_1 - m_1)$ ($(\Lambda_0^{\text{СРК}} - n_4 - m_4)$, $(\Lambda_0^{\text{МОЮЛ}} - n_9 - m_9)$) вызовов получают отказ: $N_{\text{отк}_j}^{\text{ФЛ}} = \Lambda_0^{\text{МОФЛ}} - n_1 - m_1$ ($N_{\text{отк}_j}^{\text{ЮЛ}} = (\Lambda_0^{\text{СРК}} - n_4 - m_4) + (\Lambda_0^{\text{МОЮЛ}} - n_9 - m_9)$).

Далее генерируется момент следующего прихода клиента (в соответствии с установленными интенсивностями входящих потоков клиентов всех разновидностей ($\gamma_k^{\text{ФЛ}}$, $k = 1, \dots, K^{\text{ФЛ}}$, $\gamma_k^{\text{ЮЛ}}$, $k = 1, \dots, K^{\text{ЮЛ}}$) и его тип (в соответствии с (1) и (2)).

В блоке 3 определяется очередной критический момент (т.е. ближайший момент, в который система изменит свое состояние). Это может быть момент поступления нового клиента, сгенерированный в блоках 2, 8, 13, либо момент окончания обслуживания очередной заявки у одного из сотрудников (который был установлен в блоках 2, 4, 7, 9, 10). Определяется, связан ли этот момент с обслуживанием юридического или физического лица и как он соотносится с продолжительностью моделируемого дня d_l .

Если ближайший критический момент связан с обслуживанием юридического лица и $t_{\text{кр}}^{\text{ЮЛ}} \geq T^{\text{ЮЛ}}$, типовое (описанное в схеме) моделирование обслуживания юридических лиц в рамках текущего рабочего дня заканчивается. Дальнейший расчет производится максимально приближенно к действительности: осуществляется обслуживание только тех заявок, которые уже находятся в системе, время прихода новых клиентов — юридических лиц не генерируется, т.е. новые заявки в систему не поступают. По окончании обслуживания всех заявок юридических лиц, находившихся в системе на момент $T^{\text{ЮЛ}}$, фиксируется фактическое время окончания функционирования системы: $T_{\text{max}}^{\text{ЮЛ}}$. Если моделируемый рабочий день был последним ($d_l = D$), то осуществляется переход к блоку 14 (подведение итогов моделирования в части юридических лиц), если нет — подводятся предварительные итоги. В качестве показателей конкретного моделируемого дня d_l ($l = 1, \dots, D$) выступают:

- количество вызовов юридических лиц всех видов, поступивших в систему ($N_l^{\text{ЮЛ}}$) — оно было определено в момент прихода последнего посетителя;

- количество обслуженных вызовов юридических лиц всех видов ($N_{\text{обсл}_l}^{\text{ЮЛ}}$) — оно было определено в момент окончания обслуживания последней заявки;
- количество вызовов юридических лиц, получивших отказ ($N_{\text{отк}_l}^{\text{ЮЛ}}$) — оно было определено в момент получения последнего за текущий день отказа;
- время фактического окончания работы ДО в части обслуживания юридических лиц ($T_{\text{max}_l}^{\text{ЮЛ}}$);
- время простоя работников каждого типа (для работников, занимающихся обслуживанием только юридических лиц) — в ходе моделирования фиксируются моменты начала и окончания простоя по каждому рабочему месту, в результате возможно суммировать простои по каждому типу рабочих мест $j = 4, 9$;
- время ожидания вызовов перед обслуживанием сотрудников каждого вида (для работников, занимающихся обслуживанием только юридических лиц) — в ходе моделирования фиксируются моменты начала и окончания пребывания в очереди по каждой заявке, в результате возможно определить среднее время пребывания в очереди к каждому типу рабочих мест $j = 4, 9$).

Как было сказано ранее, $T^{\text{ЮЛ}} < T^{\text{ФЛ}}$. В таком случае после подведения предварительных итогов дня l (либо завершения блока 14, если $l = D$) моделирование производится только для физических лиц (полностью аналогично тому, как оно осуществляется для юридических лиц, с той разницей, что вновь поступающая заявка может поступить только к одному типу сотрудников — МОФЛ), до того момента, пока не станет верным неравенство: $t_{\text{кр}}^{\text{ФЛ}} \geq T^{\text{ФЛ}}$. После этого в зависимости от номера моделируемого дня либо подводятся предварительные итоги работы системы в части обслуживания физических лиц, либо реализуется блок 15. В качестве рассчитываемых промежуточных показателей могут выступать:

- количество вызовов физических лиц всех видов, поступивших в систему ($N_l^{\text{ФЛ}}$) — оно было определено в момент прихода последнего посетителя;
- количество обслуженных вызовов физических лиц всех видов ($N_{\text{обсл}_l}^{\text{ФЛ}}$) — оно было определено в момент окончания обслуживания последней заявки;
- количество вызовов физических лиц, получивших отказ ($N_{\text{отк}_l}^{\text{ФЛ}}$) — оно было определено в момент получения последнего за текущий день отказа;
- время фактического окончания работы ДО в части обслуживания физических лиц ($T_{\text{max}_l}^{\text{ФЛ}}$);
- время простоя работников каждого типа (для работников, занимающихся обслуживанием не только юридических лиц) — в ходе моделирования фиксируются моменты начала и окончания простоя по каждому рабочему месту, в результате можно суммировать простои по каждому типу рабочих мест $j = 1, 2, 3, 5, 6, 7, 8$;
- время ожидания вызовов перед обслуживанием сотрудников каждого вида (для работников, занимающихся обслуживанием не только юридических лиц) — в ходе моделирования фиксируются моменты начала и окончания пребывания в очереди по каждой заявке, в результате можно определить среднее время пребывания в очереди к каждому типу рабочих мест $j = 1, 2, 3, 5, 6, 7, 8$).

На основании данных показателей можно также рассчитать их производные, например, процент вызовов, получивших отказ, коэффициент загрузки рабочих мест типа j ($j = 1, \dots, 9$) и т. д.

Если вновь наступивший критический момент $t_{\text{кр}}^{\text{ЮЛ}}$ оказывается меньше $T^{\text{ЮЛ}}$, то моделирование зависит от того, чем он был обусловлен. Если он связан с приходом нового клиента, то пересчитывается число клиентов, поступивших в систему за день:

$$N_l^{\text{ЮЛ}} = N_l^{\text{ЮЛ}} + 1. \quad (5)$$

Далее обработка заявки осуществляется однотипно для СРК и МОЮЛ: следует определить наличие очереди к тому типу сотрудников, к которому должен обратиться вновь прибывший клиент:

- если очереди нет (т. е. число людей в очереди $k_{\text{оч}}^j = 0$, где $j = 4, 9$), то:
 - если есть свободные СРК (МОЮЛ), клиент сразу поступает на обслуживание (блок 10). Отмечается момент окончания простоя начавшего обслуживания рабочего места. Тип операции, которую он хочет осуществить ($i_1 = 1, \dots, I_1$), был определен в блоках 2, 8, 13. Генерируется время его обслуживания в соответствии с плотностью распределения длительности обслуживания $f_{i_1}^j$;
 - если все СРК (МОЮЛ) заняты обслуживанием клиентов, осуществляется переход к блоку 11 (описанному ниже);
 - если очередь есть, то определяется ее текущее состояние ($k_{\text{оч}}^j$) по сравнению с максимально возможной длиной очереди (m_j):
 - если $k_{\text{оч}}^j < m_j$, вызов поступает в очередь (блок 11), длина очереди пересчитывается по формуле:

$$k_{\text{оч}}^j = k_{\text{оч}}^j + 1. \quad (6)$$

Отмечается время начала пребывания в очереди для данной заявки;

- если $k_{\text{оч}}^j = m_j$, вызову отказывается в обслуживании (блок 12), на единицу увеличивается число отказов:

$$N_{\text{отк}_l}^{\text{ЮЛ}} = N_{\text{отк}_l}^{\text{ЮЛ}} + 1. \quad (7)$$

В случае прохождения блоков 10, 11 или 12 (т. е. блоков, в рамках которых начинается обслуживание нового клиента), осуществляется переход к блоку 13, где вновь определяются время прихода следующего клиента и его тип (аналогично тому, как это делалось в блоке 2).

Если критический момент был обусловлен завершением обслуживания заявки на одном из этапов (в данном случае обслуживающим сотрудником может быть не обязательно СРК либо МОЮЛ), то осуществляется переход к блоку 9, где определяется:

- Какое далее требуется обслуживание для данной заявки? Заявка поступает в очередь к соответствующему типу сотрудников. Если есть свободные сотрудники, то генерируется ее время обслуживания и она сразу же поступает в обработку. Отмечается момент окончания простоя начавшего обслуживания рабочего места. В ряде случаев клиенту требуется последовательное выполнение разных видов операций у сотрудника одной категории. В данном случае они осуществляются одним и тем же сотрудником без перерыва между ними (т. е. клиент как бы оказывается первым следующим в очереди к данному сотруднику). Например, при подаче клиентом — юридическим лицом кредитной заявки клиенту требуется три операции, выполненные СРК подряд. В данном случае они выполняются подряд одним и тем же СРК. Для заявок, попавших в очередь при переходе к очередному этапу обслуживания, отмечается время начала пребывания в очереди к сотруднику конкретного типа. Если заявка была полностью обслужена на завершившемся этапе, то она добавляется к числу заявок, обслуженных системой:

$$N_{\text{обсл}_l}^{\text{ЮЛ}} = N_{\text{обсл}_l}^{\text{ЮЛ}} + 1. \quad (8)$$

- Есть ли очередь к сотрудникам того типа ($j = 2, \dots, 9$), у которого только что завершилось обслуживание? Если да, то из очереди на обслуживание поступает

новый вызов (очередь уменьшается на 1, генерируется время его обслуживания). Для данного вызова отмечается момент окончания ожидания в очереди к текущему типу сотрудников:

$$k_{\text{оч}}^j = k_{\text{оч}}^j - 1, j = 2, \dots, 9. \quad (9)$$

Если очереди к сотрудникам нет, фиксируется момент начала простоя рабочего места.

Далее (после реализации блоков 7, 8, 9 либо 13) осуществляется повторный переход к блоку 3 — определяется следующий критический момент.

По окончании моделирования для юридических лиц (для дня $l = D$) осуществляется блок 14, в рамках которого производится расчет окончательных показателей работы системы в части обслуживания юридических лиц. В качестве таких показателей выступают:

- *среднее количество вызовов юридических лиц всех видов, поступивших в систему в течение дня ($\bar{N}^{\text{ЮЛ}}$);*
- *среднее количество обслуженных в течение дня вызовов юридических лиц всех видов ($\bar{N}_{\text{обсл}}^{\text{ЮЛ}}$);*
- *среднее количество вызовов юридических лиц, получивших отказ в течение дня ($\bar{N}_{\text{отк}}^{\text{ЮЛ}}$);*
- *среднее время фактического окончания работы ДО в части обслуживания юридических лиц ($\bar{T}_{\text{max}}^{\text{ЮЛ}}$);*
- *среднее время простоя работников каждого типа (для работников, занимающихся обслуживанием только юридических лиц);*
- *среднее время ожидания вызовов перед обслуживанием сотрудников каждого вида (для работников, занимающихся обслуживанием только юридических лиц).*

После того как для последнего моделируемого дня ($d = D$) завершается моделирование для физических лиц, в рамках блока 15 осуществляется окончательный расчет параметров работы системы, связанных с обслуживанием физических лиц, а также юридических и физических лиц. В качестве рассчитываемых показателей выступают:

- *среднее количество вызовов физических лиц всех видов, поступивших в систему в течение дня ($\bar{N}^{\text{ФЛ}}$);*
- *среднее количество обслуженных в течение дня вызовов физических лиц всех видов ($\bar{N}_{\text{обсл}}^{\text{ФЛ}}$);*
- *среднее количество вызовов физических лиц, получивших отказ в течение дня ($\bar{N}_{\text{отк}}^{\text{ФЛ}}$);*
- *среднее время фактического окончания работы ДО в части обслуживания физических лиц ($\bar{T}_{\text{max}}^{\text{ФЛ}}$);*
- *среднее время простоя работников каждого типа (для работников, занимающихся обслуживанием не только юридических лиц);*
- *среднее время ожидания вызовов перед обслуживанием сотрудников каждого вида (для работников, занимающихся обслуживанием не только юридических лиц).*

На основании данных показателей можно также рассчитать их производные. В результате, получив оценку работы системы для заданного количества работников различных категорий, можно принять решение об увеличении либо сокращении их численности, что приведет к улучшению показателей эффективности работы системы.

В результате применение имитационного моделирования позволяет избежать зависимости от выполнения предпосылок ТМО, гибко учитывать особенности конкретной моделируемой системы, принять во внимание не только на-

личие в ней специалистов разных категорий, но и характер обращений клиентов. Также данный подход применим при наличии нестационарных (зависимых от времени) процессов: при моделировании времени прихода следующего клиента можно проверять, в какой интервал времени осуществляется моделирование и какой вид имеет интенсивность входящего потока клиентов на данном временном отрезке. Возможным является перенос полученных результатов из банковской сферы в любую ситуацию, где идет процесс обслуживания структурированных (по необходимым операциям) требований клиентов и имеются статистические данные, позволяющие оценить входные параметры системы.

Источники

- Васин А. С.* Моделирование и оценка параметров распределения длительности обслуживания клиентов банка // *Финансы и кредит*. 2004. № 12 (150). С. 19—21.
- Вентцель Е. С.* Теория вероятностей: учебник. М., 2003.
- Колосова М.* Методы оптимизации численности персонала: четыре подхода // *Управление персоналом*. 2008. № 16. С. 70—74.
- Одегов Ю. Г.* Банковский менеджмент: управление персоналом: учеб. пособие. М., 2005.
- Чернов В. П.* Математика для экономистов / под ред. А. Ф. Тарасюка. М., 2000.