

МОДЕЛЬ СИНХРОНИЗАЦИИ БАЗ ДАННЫХ НА ОСНОВЕ АКТУАЛЬНОСТИ ИНФОРМАЦИИ

Удальцов А. В.¹

Ключевые слова: синхронизация баз данных, актуальность информации, система поддержки принятия решений, информационные ресурсы, имитационное моделирование, AnyLogic.

Цель исследования. Разработка подхода к управлению синхронизацией баз данных, ориентированного на поддержание актуальности информационных элементов в системе поддержки принятия решений.

Метод исследования. Аналитический с привлечением математического аппарата для определения актуальности информационных элементов.

Результат исследования в области систем поддержки принятия решений при выполнении задач планирования операций. В статье рассматриваются вопросы построения модели синхронизации баз данных в информационной системе специального назначения. Приведена аналитическая модель оценки вероятности сохранения актуальности информации в системе поддержки принятия решений на момент ее использования и построена имитационная модель синхронизации баз данных на основе актуальности информации в среде разработки AnyLogic. Данная модель позволяет оценить состояния актуальности информационных ресурсов и получить оптимальное время для синхронизации информационных ресурсов, хранящихся в базах данных.

Научная новизна заключается в создании модели синхронизации баз данных, которая позволяет оценивать актуальность информации и определять оптимальное время для ее обновления, что улучшает управление данными в системах специального назначения.

Введение

Автоматизированные системы управления (АСУ) широко используются для решения задач, связанных с управлением сложными объектами и процессами. Однако рост сложности этих задач и увеличение объемов обрабатываемой информации требует внедрения систем поддержки принятия решений (СППР), способных эффективно справляться с возросшей нагрузкой и помогать в принятии оптимальных решений.

Система поддержки принятия решений — это человеко-машинная система, которая помогает пользователю, опираясь на данные, математические модели и экспертные знания, анализировать различные варианты решения слабоструктурированных и неструктурированных задач и находить оптимальное или допустимое решение [1]. При этом основную роль в процессе принятия решения играет человек (лицо, принимающее решение, ЛПР). СППР лишь поддерживает и облегчает процесс выбора, но окончательное решение всегда остается за ЛПР. Выводы и рекомендации системы могут быть полностью или частично отклонены или скорректированы ЛПР на основе дополнительных знаний и субъективных суждений, не учитываемых системой.

Для принятия эффективных управленческих решений ЛПР необходимо обладать актуальными информационными ресурсами, которые хранятся в базах данных (БД). Эти ресурсы состоят из информационных элементов (ИЭ), которые могут находиться как в

¹ Удальцов Александр Владимирович, адъюнкт Военной академии связи, Санкт-Петербург.
E.mail: axil2003@yandex.ru.

локальных, так и в удаленных хранилищах. Для обеспечения актуальности этих элементов требуется регулярная синхронизация баз данных. Цель данного исследования — разработка подхода к управлению синхронизацией БД, ориентированного на поддержание актуальности ИЭ.

Оценка вероятности актуальности информации в системе поддержки принятия решений

Задача синхронизации БД возлагается на систему управления базами данных (СУБД). Процесс синхронизации может быть как синхронный, так и асинхронный. При синхронном процессе данные фиксируются сразу после изменения ИЭ, а при асинхронном процессе – через строго определенный промежуток времени Δt :

$S_1(t) \rightarrow S_2(t)$ -синхронный процесс,

$S_1(t) \rightarrow S_2(t + \Delta t)$ - асинхронный процесс,

где S_1, S_2 – удаленные сервера.

Поддержание актуальности информационных элементов (ИЭ) синхронным способом приводит к значительной дополнительной нагрузке на сеть передачи данных, а также на серверное оборудование и программное обеспечение, что может снижать доступность этих элементов. В то же время для ЛПР отсутствие доступа к ИЭ при выполнении оперативных задач является недопустимым [2]. При использовании асинхронного метода синхронизации возникает риск того, что ИЭ могут содержать устаревшие значения, что может привести к принятию неадекватных решений. Поэтому возникает важный вопрос: как определить оптимальную частоту синхронизации ИЭ

Для решения этого вопроса введем понятие коэффициента актуальности ИЭ $A(t)$, который изменяется в интервале от $[0;1]$. Пример зависимости коэффициента актуальности от времени показан на рис. 1.



Рис. 1. Пример зависимости актуальности от времени

Пусть Pr_1, Pr_2, \dots, Pr_m – задачи, которые решает ЛПР, а $In = \{In_1, In_2, \dots, In_n\}$ – множество информационных элементов, хранящихся в БД. Будем считать, что In_i , которые хранятся на том же ресурсе, к которому ЛПР имеет постоянный локальный доступ, всегда будут актуальными, т.е. $\forall t > 0, A_{In_i}(t) = 1$, и, соответственно, они не требуют синхронизации.

Если существует удаленный In_i , то для него справедливо следующее условие:

$$\begin{cases} Y = 1, \text{ если } A(In_i) < A_i^{mp} \\ Y = 0, \text{ если } A(In_i) \geq A_i^{mp} \end{cases} \quad (1)$$

где $A(In_i)$ – актуальность i -го ИЭ; A_i^{mp} – требуемая пороговая актуальность i -го ИЭ; Y – бинарная индикаторная функция, указывающая на необходимость проведения

синхронизации ($Y = 1$) или отсутствие такой синхронизации ($Y = 0$).

Предположим, задача Pr_j требует для своего решения подмножество ресурсов $\mathbf{In}_j = \{In_1, In_2, \dots, In_{M_j}\}$, $\mathbf{In}_j = \bigcup_{k=1}^{M_j} In_k$. Тогда для расчета $A(\mathbf{In}_j)$ можно предложить следующую формулу:

$$A(\mathbf{In}_j) = \frac{1}{M} \sum_{k=1}^{M_j} A(In_k), \quad (2)$$

Для расчета оценки вероятности актуальности информации в СППР предлагается использовать модель оценки вероятности сохранения актуальности информации в СППР на момент ее использования.

После первоначального ввода данных в систему о реально существующих информационных ресурсах эти данные, как было сказано ранее, устаревают со временем, т.е. теряют свою актуальность для выполнения СППР своих функций [3]. Требуемая пороговая актуальность синхронизируемых данных обеспечивается на основе своевременного выявления изменений, реализации эффективных технологий синхронизаций данных в СППР, а также за счет достаточно частого обновления применяемых данных в БД [4].

Вероятность сохранения актуальности информации в СППР $P(A_{In})$ на момент ее использования вычисляют по формулам:

– при изменении состояния ИЭ:

$$P(A_{In}) = \frac{\varepsilon}{\varepsilon + T_{\delta\delta}^{Pr_m}}, \quad (3)$$

– при обновления информации в СППР вне зависимости от наличия или отсутствия изменения текущего состояния ИЭ:

$$P(A_{In}) = \frac{\varepsilon^2}{(\varepsilon + T_{\delta\delta}^{Pr_m})(\varepsilon + q)}, \quad (4)$$

В случае, когда обновление информации в СППР осуществляется строго через постоянный интервал времени q , вычисляется по формуле

$$P(A_{In}) = \frac{\varepsilon^2}{q(\varepsilon + T_{\delta\delta}^{Pr_m})} \left(1 - e^{-\frac{q}{\varepsilon}} \right), \quad (5)$$

где ε – время между изменениями ИЭ реальных объектов относительно информации, хранимой в БД, используемой в СППР;

$T_{\delta\delta}^{Pr_m}$ – время подготовки, передачи и ввода в БД данных от источников информации;

q – среднее время между соседними обновлениями.

Необходимые для моделирования границы исходных значений ε задают в исходных данных решаемых задач в СППР.

Для вычисления времени подготовки, передачи и ввода в БД данных от источников информации, воспользуемся формулой:

$$T_{\delta\delta}^{Pr_m} = T_{сч}^{Pr_m} + T_{зн}^{Pr_m}, \quad (6)$$

$T_{сч}^{Pr_m}$ – время считывания исходных данных Pr_m -ой задачи;

$T_{зн}^{Pr_m}$ – время записи обновленных данных при решении Pr_m -ой задачи.

Время считывания исходных данных складывается из времен выполнения запросов к БД.

$$T_{сч}^m = \sum_{s=1}^{S_{Pr_m}^{сч}} T_{запр}^{Pr_m s}, \quad (7)$$

$T_{запр}^{Pr_m s}$ – время выполнения s -го запроса Pr_m -ой задачи;

$S_{Pr_m}^{сч}$ – количество запросов Pr_m -ой задачи.

Количество запросов и их параметры определяются составом и структурой считываемых ИЭ при решении задач, которые решает ЛПР [5]. Каждый запрос состоит их набора операций реляционной алгебры, операции поиска по сети с других узлов. Таким образом, операции над данными можно разделить на четыре типа: соединение; проекция; выбор; обмен данными между узлами по сети.

Значение времени выполнения каждого типа операций над ИЭ одного объема для однотипной вычислительной техники близки к среднему значению. Известные значения количества операций каждого типа, выполняемых на ИЭ, и объемах обрабатываемых ИЭ позволяют получить количественную оценку времени считывания ИЭ при решении задач ЛПР [6].

Учитывая аддитивность времени выполнения различных операций над ИЭ на одном узле, время считывания исходных данных при решении задач, можно рассчитать как сумму времен операций указанных четырех типов, образующих выполняемые при решении задачи запросы, по формуле:

$$T_{сч}^{Pr_m} = \sum_{s=1}^{S_{Pr_m}^{сч}} \sum_{l=1}^4 (N_l^{Pr_m s} \cdot t_l(V_s^{сч Pr_m})), \quad (8)$$

$S_{Pr_m}^{сч}$ – количество запросов, выполняемых Pr_m -ой задачей;

$N_l^{Pr_m s}$ – количество операций l -го типа, выполняемых s -м запросом Pr_m -ой задачи;

$t_l(V_s^{сч Pr_m})$ – время выполнения операции l -го типа над данными объема $V_s^{сч Pr_m}$;

$V_s^{сч Pr_m}$ – объем обрабатываемых данных s -м запросом Pr_m задачи в Кбайтах.

Время записи обновления данных при решении задачи ЛПР включает в себя операции обмена данными по сети и можно его представить как сумму операций обмена данными по сети:

$$T_{зн}^{Pr_m} = \sum_{k=1}^{S_{Pr_m}^{зн}} D_{Pr_m k}^{зн} \cdot t_4(V_{Pr_m k}^{зн}), \quad (9)$$

$D_{Pr_m}^{зн}$ – количество обновляемых таблиц;

$S_{Pr_m}^{зн}$ – количество таблиц, обновляемых при решении Pr_m -ой задачи;

$t_4(V_{Pr_m k}^{зн})$ – время передачи k -ой таблицы по сети;

$V_{Pr_m k}^{3n}$ – объем k -ой таблицы в Кбайтах.

Сложив выражения получим среднее время подготовки, передачи и ввода в БД данных от источников информации $T_{\delta\delta}^{Pr_m}$:

$$T_{\delta\delta}^{Pr_m} = \sum_{s=1}^{S_{Pr_m}^{cu}} \sum_{l=1}^4 (N_l^{Pr_m s} \cdot t_l(V_s^{cu Pr_m})) + \sum_{s=1}^{S_{Pr_m}^{3n}} D_{Pr_m k}^{3n} \cdot t_4(V_{Pr_m k}^{3n}), \quad (10)$$

Таким образом полученное выражение позволяет определить количественное значение времени подготовки, передачи и ввода в БД данных от источников информации при решении задача ЛПР.

Экспериментальное исследование для оценки актуальности информационных элементов

Проведем экспериментальное исследование для оценки актуальности ИЭ, обновляющие БД за счет процесса синхронизации, обеспечивающих полезность данных в режиме реального функционирования СППР [7].

Важной задачей является определение обоснованного периода обновления данных о состоянии реальных объектов принимаемых в решении задач ЛПР. Возникает диалектическое противоречие, с одной стороны, обновление ИЭ по мере изменения их состояния необходимо для обеспечения достоверной информации с последующим ее применением по назначению. С другой стороны, слушком частое обновление ИЭ необоснованно перегружает каналы связи и компьютерную память, приводит к программным сбоям, создает недопустимые временные задержки, может рассинхронизировать информационные процессы в СППР, нарушая тем самым режим реального времени функционирования самой СППР и лишая необходимой информационно-аналитической поддержки ЛПР в процессе управления информацией на рабочем месте [8].

Допустим полнота отражения оперативной информации об обстановке будет с вероятностью не ниже требуемой, тогда необходимо определить такой оптимальный период обновления информации в СППР, при котором актуальность ИЭ будет $A(In_i) \geq A_i^{mp}, A_i^{mp} = 0,95$.

Анализ совокупности обновляемой информации при круглосуточной работе позволил выявить три варианта условий:

- обычные условия, характеризующиеся частотой изменения состояния системы не реже раз в неделю;
- повышенные условия, возникающие в системе характеризующиеся частотой изменения состояния не реже 1 раза в сутки;
- высокие условия, характеризующиеся частотой изменения состояния системы не реже 4 раз в час.

Время подготовки, передачи и ввода в БД данных от источников информации съема, передачи и ввода в БД вычисляется по формуле (10) в среднем 5,3 с, 10,7 с, 20,4 с соответственно.

Модель оценки вероятности сохранения актуальности информации в СППР на момент ее использования рассчитываются по формулам (1), (5).

Моделирование для определения искомого периода обновления информации в

информационной системе осуществлено по этим исходным данным с использованием модели оценки вероятности сохранения актуальности информации в СППР на момент ее использования.

Сравнительные результаты расчетов приведены на рисунках 2-4.

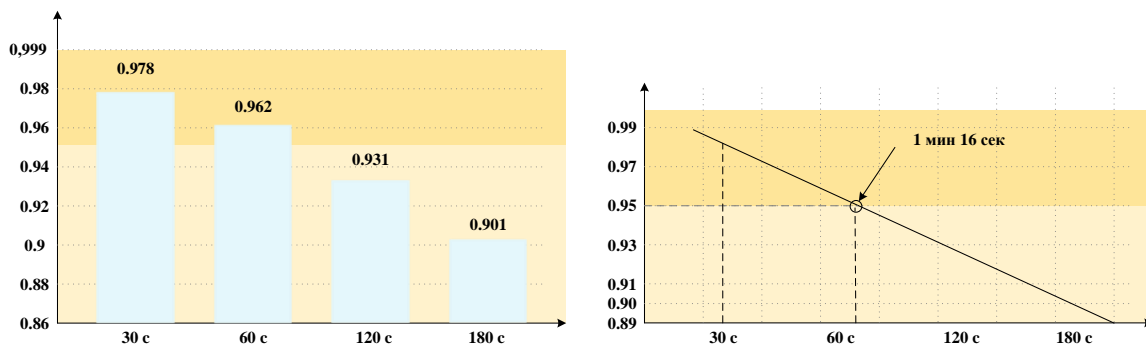


Рис. 2. Вероятность сохранения актуальности информации для высоких условий загрузки оборудования $T_{\delta\delta}^1 = 5,3$ с и зависимость вероятности сохранения актуальности информации от периода обновления (в секундах)

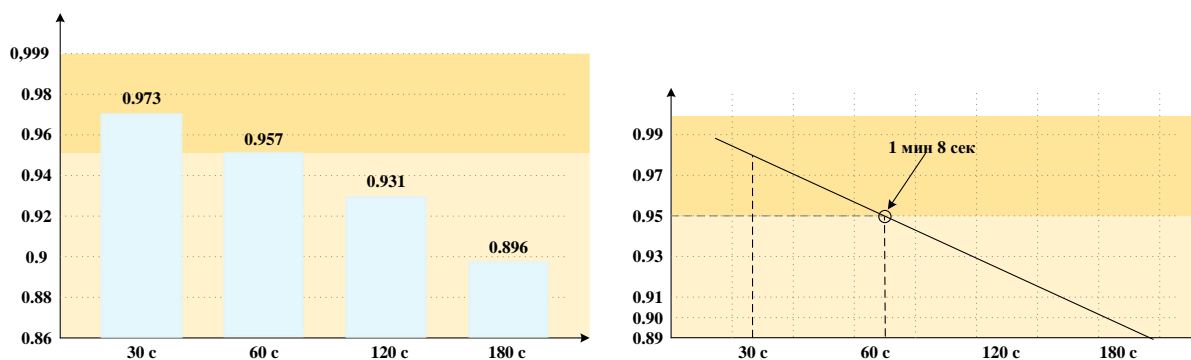


Рис. 3. Вероятность сохранения актуальности информации для высоких условий загрузки оборудования $T_{\delta\delta}^1 = 10,7$ с и зависимость вероятности сохранения актуальности информации от периода обновления (в секундах)

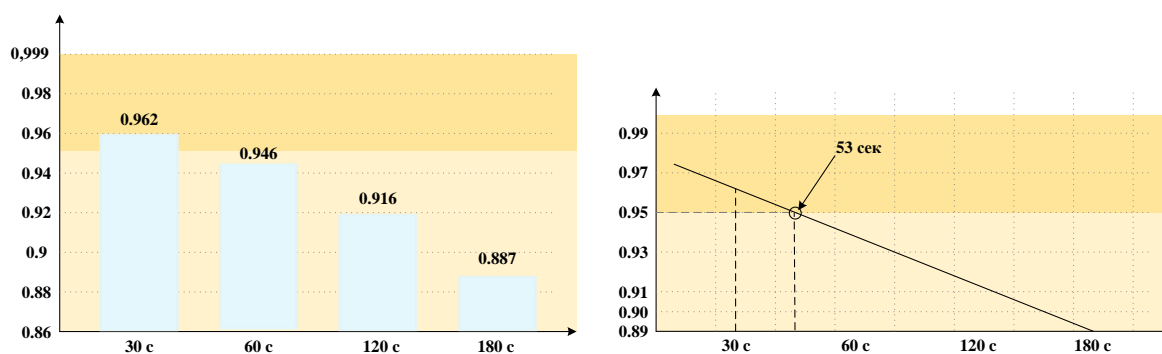


Рис. 4. Вероятность сохранения актуальности информации для высоких условий загрузки оборудования $T_{\delta\delta}^1 = 20,4$ с и зависимость вероятности сохранения актуальности информации от периода обновления (в секундах)

Анализ результатов расчетов показывает, что для обеспечения актуальности информации в СППР для ЛПР с вероятностью не ниже 0,95, период обновления может быть выбран следующим образом:

- для высоких условий синхронизируемой информации и времени подготовки,

передачи и ввода в БД данных от источников информации 5,3 сек – до 1 мин 18с;

– для высоких условий обновляемой информации и времени подготовки, передачи и ввода в БД данных от источников информации 10,7 сек – до 1 мин 8с;

– для высоких условий обновляемой информации и времени подготовки, передачи и ввода в БД данных от источников информации 20,4 сек – до 53 с.

Разработаем имитационную модель синхронизации баз данных на основе оценки вероятности сохранения актуальности информации в СППР на момент ее использования.

В качестве среды моделирования была выбрана среда имитационного моделирования AnyLogic 7.3.6. AnyLogic – современная среда разработки моделей на языке Java с русскоязычным графическим интерфейсом и тщательно продуманной контекстной справочной системой [9]. AnyLogic содержит большую библиотеку визуальных компонентов. Разработчик может также создавать и добавлять в среду собственные компоненты. Модели сохраняются как Java-апплеты. AnyLogic-модели обладают хорошими средствами 2D-3D симуляции, интерактивности и развитыми возможностями проведения экспериментов (в том числе оптимизационных). Поэтому модель позволяет пользоваться всеми преимуществами объектно-ориентированного моделирования. Схема имитационной модели представлена на рисунке 5.

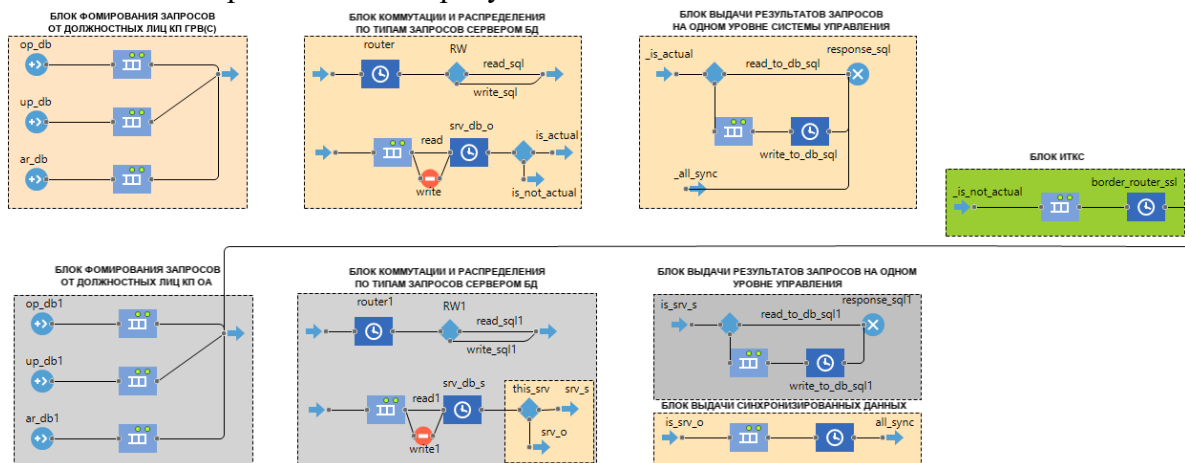


Рис. 5. Схема имитационной модели

На основе построенной имитационной модели была собрана статистика и проведен анализ системы при изменении времен выполнения различных операций над ИЭ на одном узле, считывания исходных данных при решении задач времени записи обновления данных при решении задачи ЛПР, времени между изменениями ИЭ реальных объектов относительно информации, хранимой в БД [10]. Симуляция проводится в режиме реального времени с выполнением 3 информационно-расчетных задач, длительность моделирования – 10 суток. Присутствует возможность произвести моделирование в режиме виртуального времени.

Выводы

В статье рассмотрена модель синхронизации баз данных на основе актуальности информационных элементов в системе поддержки принятия решений. Результатом моделирования является вывод оптимального времени синхронизации для каждого информационного элемента с заданной требуемой актуальностью информации.

Литература

1. Допира Р.В., Потапов А.Н., Брежнев Д.Ю., Гетманчук А.В., Семин М.В. Оценка эффективности методического и алгоритмического обеспечения системы поддержки принятия решений специального назначения // Программные продукты и системы. 2019. № 2. С. 273–282.
2. Саенко И.Б. Анализ проблемы синхронизации локальных баз данных в распределенной информационной системе / И.Б. Саенко, А.В. Удальцов, А.В. Ермаков // Труды Научно-исследовательского института радио. – 2022. – №4. – С. 37–41. – DOI 10.34832/NIPR.2022.11.4.004.
3. ГОСТ Р 59341-2021. Системная инженерия. Защита информации в процессе управления информацией системы. М.: Изд-во стандартов, 2021.
4. Фабияновский И.Н. Обеспечение своевременности обмена информационными ресурсами на основе технологии распределенного реестра // Известия Тульского государственного университета. Технические науки. 2021. Вып. 9. С. 184–194.
5. Храмов В.Ю. методы и средства проектирования баз данных /В.Ю. Храмов, А.И. Кустов, Э.Б. Ханов – Воронеж: Воронежский ЦНТИ – филиал ФГБУ «РЭА» Минэнерго России, 2015. – 188 с.
6. Хемди А.Таха. Исследование операций. М.: ООО «И.Д. Вильямс», 2016. 912 с.
7. Бирюков М.А., Брунилин А.А., Саенко И.Б. Имитационный подход к моделированию системы разграничения доступа к единому информационному пространству // Актуальные проблемы инфотелекоммуникаций в науке и образовании. Сборник научных статей: в 4-х томах. Под редакцией С.В. Бачевского. 2017. С. 78–83.
8. Анфилатов В.С., Авраменко В.С., Пантюхин О.И. Теоретические основы автоматизации управления войсками и связью. Часть 1. Системные основы автоматизации управления войсками и связью: учеб. Пособие. СПб.: ВАС, 2014. 312 с.
9. Карпов Ю. Имитационное моделирование систем. Введение в моделирование с AnyLogic 5. [Текст] / Ю. Карпов. – СПб.: БХВ-Петербург, 2005.–400 с.
10. Шаталова Ю.Г. Исследование модели обработки запросов к базе данных в среде AnyLogic // Восточно-Европейский журнал передовых технологий 4/2 (58) С. 48–50.

A DATA BASE SYNCHRONIZATION MODEL BASED ON INFORMATION RELEVANCE

Udaltsov A. V.²,

Keywords: *database synchronization, information relevance, decision support system, information resources, simulation modeling, AnyLogic.*

Objective. *Development of an approach to managing database synchronization aimed at maintaining the relevance of information elements in a decision support system.*

Research method. *Analytical with the involvement of mathematical tools for determining the relevance of informational elements.*

The result *of the research in the field of decision support systems when performing operation planning tasks. The article discusses issues related to building a model for synchronizing databases in a special-purpose information system. An analytical model is presented for assessing the probability of preserving the relevance of information in a decision support system at the time of its use, and a simulation model for database synchronization based on the relevance of information has been built in the AnyLogic development environment. This model allows us to assess the state of relevance of information resources and obtain the optimal time for synchronizing information resources stored in databases.*

The scientific novelty lies *in creating a database synchronization model that allows assessing the relevance of information and determining the optimal time for updating it, which improves data management in special-purpose systems.*

² Alexander V. Udaltsov, adjunct of Military Academy of Communications named after Marshal of the Soviet Union S.M. Budyonny, Saint Petersburg, E-mail: axil2003@yandex.ru.

References

1. Dopira R.V., Potapov A.N., Brezhnev D.Yu., Getmanchuk A.V., Semin M.V. Evaluation of the effectiveness of methodological and algorithmic support for a special-purpose decision support system // Software products and systems. 2019. No. 2. P. 273-282.
2. I.B. Saenko, Analysis of the problem of synchronization of local databases in a distributed information system / I.B. Saenko, A.V. Udaltsov, A.V. Ermakov // Proceedings of the Radio Research Institute. - 2022. - №4. – P. 37-41. – DOI 10.34832/NIIR.2022.11.4.004.
3. GOST R 59341-2021. System engineering. Information protection in the process of information management of the system. Moscow: Publishing House of Standards, 2021.
4. Fabiyanovsky I.N. Ensuring timeliness of exchange of information resources based on distributed ledger technology // Izvestiya Tula State University. Technical Sciences. 2021. Issue 9. P. 184-194.
5. Khramov V.Yu. Methods and means of database design / V.Yu. Khramov, A.I. Kustov, E.B. Khanov - Voronezh: Voronezh TsNTI - branch of FGBU "REA" of the Ministry of Energy of Russia, 2015. - 188 p.
6. Khemdi A. Takha. Operations research. Moscow: LLC "I.D. Williams", 2016. 912 p.
7. Biryukov M.A., Brunilin A.A., Saenko I.B. Simulation approach to modeling the access control system for a single information space // Actual problems of infotelecommunications in science and education. Collection of scientific articles: in 4 volumes. Edited by S.V. Bachevsky. 2017. P. 7
8. V.S. Anfilatov, V.S. Avramenko, O.I. Pantyukhin. Theoretical Foundations of Automation of Troops and Communications Control. Part 1. Systemic Foundations of Troops and Communications Control Automation: Textbook. St. Petersburg: Military Academy of Communications, 2014. 312 p.
9. Y. Karpov. Simulation Modeling of Systems. Introduction to Modeling with AnyLogic 5. St. Petersburg: BHV-Petersburg, 2005. 400 p.
10. Yu.G. Shatalova. Study of the Model of Processing Requests to the Database in the AnyLogic Environment. Eastern European Journal of Advanced Technologies 4/2 (58). Pp. 48–50.