

ПОСТАНОВКА ЗАДАЧ АВТОМАТИЗИРОВАННОГО АНАЛИЗА ДАННЫХ В ПОДСИСТЕМЕ АУДИТА ПРЕДПОСЫЛКИ СРЕДНЕГО УРОВНЯ ИТ СППР

В статье рассматривается проблема постановки задач автоматизированного анализа данных в подсистеме аудита предпосылки Положений (стандартов) бухгалтерского учета в ИТ СППР, как составляющих методики обобщенно-множественного отображения информации. Определены функциональные особенности учета отображений множеств данных элементарных задач двух видов (по функциональному назначению) за период проверки в подсистеме среднего уровня и наборов данных задач за период проверки. Выполнена постановка задач автоматизированного анализа данных за период проверки в подсистеме аудита предпосылки среднего уровня. Данные постановки задач инвариантны относительно особенностей предприятия и являются основанием для создания методики автоматизации решения задач анализа данных в ИТ СППР.

Ключевые слова: постановка задачи, автоматизированный анализ данных, информационная технология, СППР аудита, методика обобщенно-множественного отображения информации, средний уровень.

T. NESKORODIEVA

Donetsk National University named Vasyly Stus

TASKS STATEMENT OF AUTOMATED DATA ANALYSIS OF THE PREREQUISITES AUDIT SUBSYSTEM OF THE AVERAGE LEVEL DSS IT

The article discusses the problem of setting tasks for automated data analysis in the audit subsystem of the prerequisites for Accounting Regulations (standards) in IT DSS, as components of the generalized-multiple display of information. The functional features of accounting for the mappings of the sets of these tasks for the verification period in the mid-level subsystem are determined. This made it possible to determine that the direct mapping of these tasks is ambiguous, and the reverse is unambiguous, and also that the background check is not automatically decomposed into solving problems for each set of accounting objects and operations. Based on the methodology of generalized-multiple display of information, two types of tasks for automated data analysis for the period of verification in the audit subsystem of the medium level are formulated. Two types of tasks of automated data analysis for the period of verification in the audit subsystem of the medium level have been formulated. The first local sub-task is data analysis to determine equivalent and nonequivalent subsets when displaying data of an average level of sets of parallel consecutive operations during the verification period. The second local subtask is data analysis to determine equivalent and nonequivalent subsets when displaying data of an average level of accounting for sets of objects that are transformed as a result of many parallel sequential operations during the verification period. The problem statement data is invariant with respect to the enterprise features and is the basis for creating a solution automation technique of data analysis tasks in IT DSS. These statement of tasks are invariant with respect to the characteristics of the enterprise and are the basis for creating methods for automating the solution of problems of data analysis in IT DSS.

Key words: task statement, automated data analysis, information technology, audit DSS, generalized-plural presentation of information methodology, middle level.

Постановка проблемы. В настоящее время актуальной научно-технической проблемой информационных технологий финансово-экономической сферы является автоматизация анализа больших объемов данных финансово-экономической информации предприятий хранящихся и поступающих в режиме онлайн в базы данных локальных и глобальных компьютерных систем с целью формирования рекомендаций принятия решений при аудите. Так, например, эффективность ИТ бизнеса в сокращении коррупции обоснована в [1]. В [2] рассмотрены теоретические и практические вопросы связанные с применением ИТ в аудите. Определены изменения, которые вносит в методологию аудита его комплексная компьютеризация.

Анализ последних источников. Применение современных систем учета и управления, на предприятии которые интегрированы в глобальные, многоуровневые системы информационных компьютерных сетей дает потенциальные возможности обработки больших объемов данных, которые используются не в полной мере [3]. В [4] представлена архитектура подсистем внешнего и внутреннего аудита системы Audit 4.0, внедрение и функционирование которых сталкивается с проблемой обработки больших объемов данных. Технические проблемы внедрения ИТ в контрольно-ревизионной деятельности выделены в [5].

Создание современных ИТ аудита предполагает формирование рекомендуемых решений на основании результатов автоматизированного применения интеллектуальных технологий анализа данных. Говоря об автоматизации процедур анализа как составляющих при проектировании ИТ аудита, следует отметить, что в литературе представлены методы обнаружения мошенничества, основанные на применении моделей нейронных (вероятностных) сетей, логистических регрессий, деревьях решений, экспертных оценок, статистических методах анализа данных [6–10]. Недостатком перечисленных методов является тот факт, что их применение дает результат, если выполняются определенные условия, которые зависят в частности от свойств показателей. Следовательно они не являются универсальными, что затрудняет автоматизацию анализа на их основании и их использование при решении задач автоматизации на всех этапах аудита.

Проектирование ИТ СППР аудита согласно методике обобщенно-множественного отображения информации основано на формализации функциональной структуры преобразований проверяемых данных в виде последовательности отображений множеств. Первый этап создания методики – постановка задач. Первый вид задач – определение эквивалентных и неэквивалентных подмножеств при отображении множеств данных параллельно последовательных операций. Второй вид задач по функциональному

назначению – выделение эквивалентных и неэквивалентных подмножеств при отображении данных множеств объектов двух видов, которые преобразуются в результате множества параллельно последовательных операций. Методика постановки задач проиллюстрирована на примере подсистемы аудита предпосылки «Полнота» материальных расходов [11]. Данные элементарные задачи образуют набор взаимосвязанных задач по отношению ко всем множествам объектов учета и операций при учете материальных расходов за период проверки на среднем уровне.

Цель статьи: выполнить формальную постановку задач автоматизированного анализа данных элементарных задач подсистемы аудита предпосылки за период проверки инвариантную относительно особенностей предприятия и как составляющих методики обобщенного множественного отображения данных.

Для этого необходимо решить следующие **задачи:**

- определить функциональные особенности учета данных элементарных задач аудита за период проверки;
- выполнить формальную постановку задач автоматизированного анализа данных задач за период проверки.

В порядке общности методика рассматривается для подсистемы аудита предпосылки «Полнота» данных среднего уровня экономико-производственной (основной) деятельности предприятия так как в силу количества видов подмножеств и видов взаимосвязей между ними, количества элементарных задач – это наиболее общая и трудоемкая задача. Для наглядности и связи с предыдущими результатами автора [11] составляющие методики рассматриваются на примере подсистемы аудита предпосылки «Полнота» материальных расходов.

Изложение основного материала. Функциональные особенности учета данных элементарных задач за период проверки. Обозначим элементы предметной области аудита на среднем уровне:

E – множество объектов,

X – множество операций,

$\overline{e^{(c)}}$ – вектор обобщенных показателей среднего уровня учета множества объектов E ,

$\overline{x^{(c)}}$ – вектор обобщенных показателей среднего уровня учета множества операций X ,

$n^{(c)}$ – номер документа среднего уровня учета,

s и S – вид множества и множество видов объектов учета,

u и U – подвид и множество подвидов объектов учета,

t или T – период учета (проверки).

Совокупность задач аудита в [11] представлена в виде наборов задач двух видов по функциональному назначению. Задача первого вида – определение эквивалентных и неэквивалентных подмножеств значений обобщенных показателей среднего уровня $\overline{e^{(c)}}$ и $\overline{e^{(c)}}$ множеств объектов E_s и E_{s+1} двух видов (или подвидов), которые преобразуются в результате множества параллельных операций $X_s^{s+1}(N_s)$ за период t . Задача второго вида – определение эквивалентных и неэквивалентных подмножеств при отображении значений обобщенных показателей среднего уровня $\overline{x^{(c)}}$ и $\overline{x^{(c)}}$ множеств параллельно последовательных операций $X_{s-1}^s(N_s)$ и $X_s^{s+1}(N_{s+1})$ за период t . Для определения возможных режимов анализа данных (параллельный (независимый), одновременный (взаимосвязанный), последовательный) наборов задач сформулированных выше, определим функциональную структуру их учета за период T .

Значения обобщенных характеристик множеств объектов E_s накапливаются в хронологическом порядке по каждому виду s за периоды квантования периода t_l , обобщаются за весь период T и образуют множества данных среднего уровня учета объектов каждого вида:

$$e_s^{(c)} = \left\{ \overline{e^{(c)}(n_s^{(c)}, t_l)}, \overline{e^{(c)}(n_s^{(c)}, T)}, t_l \in T, l = \overline{1, L} \right\}, s \in S, \quad (1)$$

где t_l – период квантования периода учета T .

Множество данных (1) каждого вида формируется на основании подмножеств данных по подвидам u объектов за соответствующие периоды:

$$e_{s,u}^{(c)} = \left\{ \overline{e^{(c)}(n_{s,u}^{(c)}, t_l)}, \overline{e^{(c)}(n_{s,u}^{(c)}, T)}, t_l \in T, l = \overline{1, L} \right\}, u \in U, s \in S. \quad (2)$$

Тогда множество данных среднего уровня учета объектов $e^{(c)}$ за период t , $t \in \{t_l \in T, l = \overline{1, L}\}$ можно представить в виде объединения непересекающихся подмножеств данных объектов каждого вида (3), которые в свою очередь можно представить в виде непересекающихся подмножеств данных подвидов (4):

$$e^{(c)} = \bigcup_{s=1}^S e_s^{(c)}, e_s^{(c)} \cap e_g^{(c)} = \emptyset, s \neq g, \tag{3}$$

$$e_s^{(c)} = \bigcup_{u \in U} e_{s,u}^{(c)}, e_{s,u}^{(c)} \cap e_{s,g}^{(c)} = \emptyset, u \neq g, s \in S, \tag{4}$$

где $e^{(c)}$ – набор данных среднего уровня, характеризующий множество объектов учета.

Значения обобщенных характеристик множеств параллельно последовательных операций за период $t, t \in \{t_l \in T, l = \overline{1, L}\}$ образуют множество данных среднего уровня учета этапов экономико-производственной деятельности (например, осуществления материальных расходов):

$$x_{s-1}^{s(c)} = \left\{ \overline{x(n_s^{(c)})} : m_{s,s-1} = 1, s \in S, (s-1) \in S \right\}, \tag{5}$$

$$x_{s-1,u}^{s,p(c)} = \left\{ \overline{x(n_{s,u}^{(c)})} : m_{u,p} = 1, u \in U, p \in P \right\}, s \in S, \tag{6}$$

где $x_{s-1}^{s(c)}$ – набор данных множества параллельно последовательных операций,

$x_{s-1,u}^{s,p(c)}$ – набор данных подмножества параллельно последовательных операций,

$m_{s,s-1}$ – элемент матрицы смежности видов объектов учета (характеризует виды объектов взаимосвязанные операций, в системе счетов определяется корреспонденцией счетов),

$m_{u,p}$ – элемент матрицы смежности подвидов объектов учета (характеризует подвиды объектов взаимосвязанные операций, в системе счетов определяется корреспонденцией субсчетов),

n – номер документа учета.

Тогда множество данных среднего уровня учета $x^{(c)}$ множества операций X за период $t, t \in \{t_l \in T, l = \overline{1, L}\}$ можно представить в виде объединения непересекающихся множеств и подмножеств данных множеств параллельно последовательных операций:

$$x^{(c)} = \bigcup_{s=1}^S x_s^{(s+1)(c)} = \bigcup_{l=3}^5 x_l^{(c)}, x_l^{(c)} \cap x_m^{(c)} = \emptyset, l \neq m, \tag{7}$$

$$x_s^{s+1(c)} = \bigcup_{u \in U} \bigcup_{p \in P} x_u^{p(c)}, x_{u_1}^{p_1(c)} \cap x_{u_2}^{p_2(c)} = \emptyset, (u_1 \neq u_2) \vee (p_1 \neq p_2). \tag{8}$$

Для определения свойств взаимосвязей данных среднего уровня множеств объектов $e^{(c)}$ (4) и множеств операций $x^{(c)}$ (8) за период $t, t \in \{t_l \in T, l = \overline{1, L}\}$, определим их свойства для каждой элементарной задачи.

Φ_1 – документ с номером $n_{s,u}^{(c)}$ учета обобщенных показателей множества операций $X(N_{s,u})$, которые находятся в отношении увеличения $r^{+(c)}$ к множеству объектов учета $E(N_{s,u})$ подвида u единственный;

Φ_2 – документы с номерами $n_{s+1,p}^{(c)}$ учета обобщенных показателей множеств операций $X(N_{s+1,p})$, которые находятся в отношении $r^{-(c)}$ уменьшения к множеству объектов учета $E(N_{s,u})$ в общем случае больше или равно 1:

$$\forall E_{s,u} \left(\exists! n_{s,u}^{(c)} : x(n_{s,u}^{(c)}) r_s^{+(c)} e(n_{s,u}^{(c)}) \right) \wedge \left(\exists n_{s+1,p}^{(c)} : x(n_{s+1,p}^{(c)}) r_{s+1,p}^{-(c)} e(n_{s,u}^{(c)}) \right). \tag{9}$$

Следовательно, отношения соответствия данных множеств объектов учета (4) и операций (8) за период T одно-мнозначные. Формализуем данные отношения соответствия в виде графов $G_1^{(c)}$ (рис. 1) и $G_2^{(c)}$ (рис. 2) соответственно.

$$G_1^{(c)}(s,u) = (V_1^{(c)}(s,u), R_1^{(c)}(s,u)), V_1^{(c)}(s,u) = (e_{s,u}^{(c)}, x_{s,u}^{(c)}, x_{s+1,1}^{(c)}, \dots, x_{s+1,P}^{(c)}), \\ R_1^{(c)}(s,u) = (r_s^+, r_s^-(n_{s,1}^{(c)}), \dots, r_s^-(n_{s,P}^{(c)})), u \in U, s = \overline{1, S}, \tag{10}$$

$$G_2^{(c)}(s,u) = (V_2^{(c)}(s,u), R_2^{(c)}(s,u)), V_2^{(c)}(s,u) = (e_{s,u}^{(c)}, x_{s,u}^{(c)}, x_{s+1,1}^{(c)}, \dots, x_{s+1,P}^{(c)}, e_{s+1,1}^{(c)}, \dots, e_{s+1,P}^{(c)}),$$

$$R_2^{(c)}(s,u) = (r_s^-(n_{s,1}^{(c)}), \dots, r_s^-(n_{s,P}^{(c)}), r_{s+1}^-(n_{s+1,1}^{(c)}), \dots, r_{s+1}^-(n_{s+1,P}^{(c)})), u \in U, s = \overline{1, S}, \quad (11)$$

где $G_1^{(c)}$ – граф отношений соответствий данных задач первого вида,

$G_2^{(c)}$ – граф отношений соответствий данных задач второго вида,

V – множество вершин графа G , R – множество ребер графа G ,

e – вершина, характеризующая данные среднего уровня множества объектов учета,

s и $s+1$ – номера, характеризующие виды множеств объектов учета,

x – вершина, характеризующая данные среднего уровня множества операций,

$r^{+(c)}$ ($r^{-(c)}$) – ребро, характеризующее отношение соответствия данных множества операций к

увеличению (уменьшению) множества объектов учета,

p – количество подвидов операций и объектов учета,

$n^{(c)}$ – номер документа (подсистемы) среднего уровня, содержащего обобщенные показатели по документам первого уровня множества N .

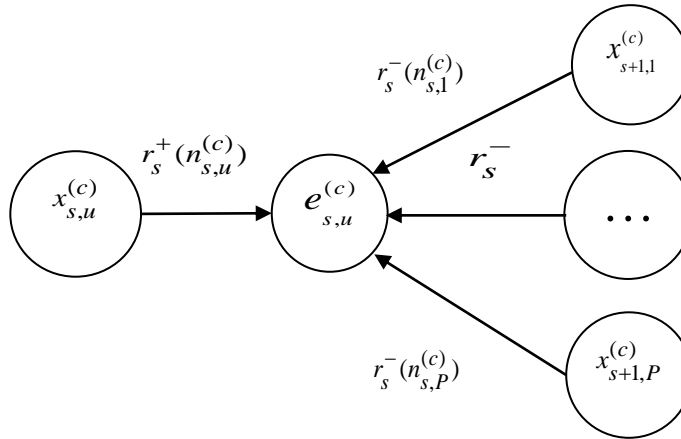


Рис. 1. Граф $G_1^{(c)}$ отношений соответствий данных задач первого вида для одного множества объектов учета

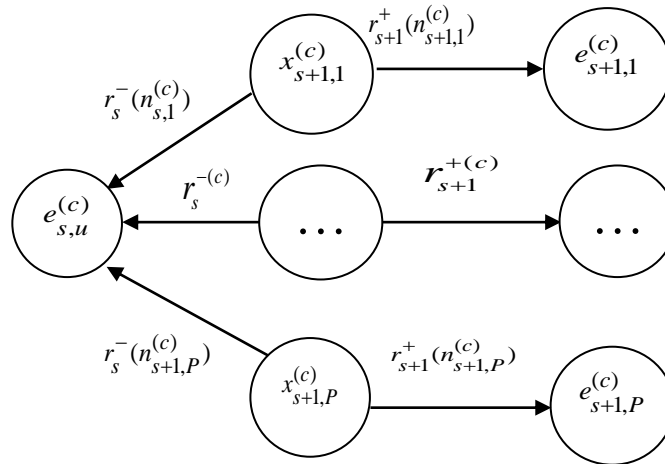


Рис. 2. Граф $G_2^{(c)}$ отношений соответствий данных задач второго вида для одного множества операций

Выделим функциональные особенности системы учета наборов задач соотношения между элементами, которые формализованы в виде графов $G_1^{(c)}$ и $G_2^{(c)}$ (рис. 1 и рис. 2) за период проверки.

$\Phi_1^{(c)}$ – учет данных каждой задачи за период T не выделяется в отдельную подсистему;

$\Phi_2^{(c)}$ – учет отношений соответствий $r^{(c)}$ множеств операций в документах среднего уровня в общем случае (для разных методик учета отпуска сырья (инвентарный, обратной калькуляции) и для всех видов сырья основных и вспомогательных материалов) осуществляется в ручном режиме или обобщаются за период T на основании значений нижнего уровня, которые учитываются в ручном режиме (из за разрыва во времени учета);

$\Phi_3^{(c)}$ – взаимосвязи между наборами переменных документов, между которыми существуют отношения соответствия $G_1^{(c)}$ и $G_2^{(c)}$ в процессе учета в общем случае автоматически не проверяются.

Из $\Phi_1^{(c)}$ следует, что задача проверки наборов задач не декомпозируется в автоматическом режиме на элементарные задачи.

Из $\Phi_2^{(c)}$ и $\Phi_3^{(c)}$ следует, что для выявления элементарных задач, в которых нарушены взаимосвязи переменных необходимо проводить анализ данных за весь период проверки

Постановка задач. Первая локальная подзадача – анализ данных с целью определения эквивалентных и неэквивалентных подмножеств при отображении данных среднего уровня множеств параллельно последовательных операций за период проверки T . Вторая локальная подзадача – анализ данных с целью определения эквивалентных и неэквивалентных подмножеств при отображении данных среднего уровня учета множеств объектов, которые преобразуются в результате множества параллельно последовательных операций за период проверки T .

Методика 1 локальной подзадачи реализуется в подсистемах аудита данных материальных расходов: «Производство полупродукта (цех 1)», «Производство (цех 2)», «Готовая продукция». Методика 2 локальной подзадачи реализуется в подсистемах аудита взаимосвязей между перечисленными подсистемами учета материальных расходов.

Выводы. Определены функциональные особенности учета данных задач среднего уровня, которые инвариантны относительно особенностей предприятия. Это позволило определить, что прямое отображение данных задач среднего уровня многозначно, а обратное однозначно. Установлено, что проверка предпосылки не декомпозируется в автоматическом режиме на решение элементарных задач по каждому множеству объектов и операций.

Выполнена постановка задач автоматизированного анализа данных в подсистеме аудита предпосылки, которые инвариантны относительно особенностей предприятия. Это позволяет перейти к созданию инструментария анализа данных, результаты которого станут основой для формирования рекомендуемых решений в ИТ СППР аудита.

Литература

1. Xinli Hu. Effectiveness of information technology in reducing corruption in China Electronic Library, 2015 Vol. 33. Issue 1. P. 52–64, doi: 10.1108 / el-11-2012-0148.
2. Ивахненко С. В. Информационные технологии аудита и внутрихозяйственного контроля в контексте мировой интеграции : монография / Ивахненко С. В. – Житомир : ЧП «Рута», 2010. – 432 с.
3. Kirkos E., Spathis C., Manolopoulos Y. Data mining techniques for the detection of fraudulent financial statement. *Expert Syst. Appl.*, 2007. Vol. 32. Issue 4. p. 995–1003.9. doi:10.1016/j.eswa.2006.02.016.
4. Dai J., Vasarhelyi M. A. Imagineering Audit 4.0. *Journal of Emerging Technologies in Accounting*, 2016. № 13(1). P. 1–15.
5. Яремко С.А. Дослідження проблем впровадження сучасних інформаційних систем аудиту в контрольно-ревізійній діяльності / С.А. Яремко, В.В. Коваленко // Комп'ютерно-інтегровані технології освіта наука виробництво. – Луцьк, 2014. – № 14 – С. 179–182.
6. West J., Bhattacharya M. and Islam R. Intelligent Financial Fraud Detection Practices: a comprehensive review. *Computers & Security*, 2016. Vol. 57. P. 47–66. doi: 10.1007 / 978-3-319-23802-9_16
7. Lin Chi-Chen, Chiu An-An, Yan Huang Shaio, C. Yen. David. Detecting the financial statement fraud: The analysis of the differences between data mining techniques and experts' judgments. *Knowledge-Based Systems*, 2015. № 89. P. 459–470: doi : 10.1016 / j.knosys.2015.08.011.
8. Бююль А. SPSS: Искусство обработки информации. Анализ статистических данных и восстановление скрытых закономерностей / Бююль А., Цёфель П. ; пер. с нем. – СПб : ДиаСофтЮП, 2005. – 608 с.
9. Zgurovsky M.Z., Zaychenko Y.P.: The fundamentals of computational intelligence: system approach. Springer International Publishing Switzerland. 2017. 356 p.
10. Hu Z., Bodyanskiy Y. Tyshchenko O. Self-Learning and adaptive algorithms for business applications. Emerald Publishing Limited. 2019. 382 p. doi: 10.1108/978-1-83867-171-620191001
11. Нескорородева Т.В. Постановка задач подсистемы аудита предпосылки среднего уровня ИТ СППР / Тетяна Нескорородева // Вісник Черкаського державного технологічного університету. – 2020. – № 2. – С. 38–48. doi: 10.24025/2306-4412.2.2020.194985.

References

1. Xinli Hu. Effectiveness of information technology in reducing corruption in China Electronic Library, 2015 Vol. 33. Issue 1. P. 52–64, doi: 10.1108 / el-11-2012-0148.
2. Ivahnenkov S. V. Informacionnye tehnologii audita i vnutrihozyajstvennogo kontrolya v kontekste mirovoj integracii : monografiya / Ivahnenkov S. V. – Zhitomir : ChP «Ruta», 2010. – 432 s.
3. Kirkos E., Spathis C., Manolopoulos Y. Data mining techniques for the detection of fraudulent financial statement. *Expert Syst. Appl.*, 2007. Vol. 32. Issue 4. p. 995–1003.9. doi:10.1016/j.eswa.2006.02.016.
4. Dai J., Vasarhelyi M. A. Imagineering Audit 4.0. *Journal of Emerging Technologies in Accounting*, 2016. № 13(1). P. 1–15.
5. Iaremko S.A. Doslidzhennia problem vprovadzhennia suchasnykh informatsiinykh system audytu v kontrolno-reviziiinii diialnosti / S.A. Yaremko, V.V. Kovalenko // Kompiuterno-intehrovani tekhnolohii osvita nauka vyrobnytstvo. – Lutsk, 2014. – № 14 – S. 179–182.
6. West J., Bhattacharya M. and Islam R. Intelligent Financial Fraud Detection Practices: a comprehensive review. *Computers & Security*, 2016. Vol. 57. P. 47–66. doi: 10.1007 / 978-3-319-23802-9_16

-
7. Lin Chi-Chen, Chiu An-An, Yan Huang Shaio, C. Yen. David. Detecting the financial statement fraud: The analysis of the differences between data mining techniques and experts' judgments. Knowledge-Based Systems, 2015. № 89. P. 459–470: doi : 10.1016/j.knosys.2015.08.011.
 8. Byuyul A. SPSS: Iskusstvo obrabotki informacii. Analiz statisticheskikh dannyh i vosstanovlenie skrytyh zakonomernostej / Byuyul A., Cyofel P. ; per. s nem. – SPb : DiaSoftYuP, 2005. – 608 s.
 9. Zgurovsky M.Z., Zaychenko Y.P.: The fundamentals of computational intelligence: system approach. Springer International Publishing Switzerland. 2017. 356 p.
 10. Hu Z., Bodyanskiy Y. Tyshchenko O. Self-Learning and adaptive algorithms for business applications. Emerald Publishing Limited. 2019. 382 p. doi: 10.1108/978-1-83867-171-620191001
 11. Neskorođeva T.B. Postanovka zadach podsistemy audita predposylki srednego urovnya IT SPPR / Tetyana Neskorođeva // Visnik Cherkaskogo derzhavnogo tehnologichnogo universitetu. – 2020. – № 2. – S. 38–48. doi: 10.24025/2306-4412.2.2020.194985.

Надійшла / Paper received: 17.04.2020

Надрукована / Paper Printed : 02.06.2020