

ВЫЯВЛЕНИЕ ПЛАГИАТА В ГРАФИЧЕСКИХ КОНСТРУКТОРСКИХ ДОКУМЕНТАХ В ХОДЕ УЧЕБНОГО ПРОЕКТИРОВАНИЯ

В. В. Миронов

Аннотация. В статье рассматривается не исследованная ранее задача выявления плагиата в графических конструкторских документах в процессе учебного проектирования разнообразных схем. Отмечается специфика этой задачи в сравнении с задачей обнаружения плагиата в одномерных текстовых документах, обусловленная размещением и соединением на схеме в двухмерном пространстве условных графических обозначений. Учебное проектирование, основанное на типовых заданиях, расширяет возможности для плагиата, поскольку задания могли неоднократно выполняться ранее. Предлагаемый подход к решению задачи основан на применении персонализированных шаблонов-заготовок, генерируемых для исполнителя конструкторского документа таким образом, что это позволяет выявить в готовом документе заимствования из других документов. В частности, предлагается использовать цифровые водяные знаки, которые внедряются в условные графические обозначения схемы для идентификации исполнителя. Описывается реализация этого подхода для конструкторских документов в графических форматах редактора Visio. Обсуждается применение разработанной информационной технологии в учебном процессе.

Ключевые слова: плагиат; графический конструкторский документ; цифровой водяной знак; учебное проектирование.

ВВЕДЕНИЕ

Плагиат – «умышленно совершаемое физическим лицом незаконное использование или распоряжение охраняемыми результатами чужого творческого труда, которое сопровождается доведением до других лиц ложных сведений о себе как о действительном авторе» (<https://ru.wikipedia.org/wiki/Плагиат>) – печальное и достаточно частое явление среди студентов в учебном процессе. Борьба с плагиатом является одной из задач эффективного управления образовательной деятельностью, а создание работоспособных методов и средств такой борьбы – достаточно актуальная задача с высокой социальной значимостью.

Учебный процесс неизбежно включает в себя отработку типовых задач и приемов профессиональной деятельности, что порождает соблазн у недобросовестных студентов воспользоваться результатами предыдущих решений этих или подобных задач для выполнения учебных заданий в короткий срок с меньшими умственными усилиями. Такая возможность особенно упрощается в современных условиях глобальной информатизации, когда результаты интеллектуального труда повсеместно представлены в цифровой форме. Студенты давно не пишут работы от руки, не вычерчивают схемы и чертежи, а создают их в виде электронных компьютерных документов. Информация легко размещается в глобальной Сети, после чего она становится легкодоступной для неограниченного числа людей. Повсеместно применяемая технология «сору-and-paste» – «скопируй и вставь» – позволяет легко позаимствовать интеллектуальный результат. В этих условиях преподаватели нуждаются в эффективных методах и инструментах для выявления плагиата в студенческих работах.

Процесс учебного проектирования иллюстрируется на рисунке 1, где в укрупненном виде показано взаимодействие участников процесса – исполнителей (студентов) и консультантов (преподавателей). В задачу консультанта входит проверка конструкторских документов (КД), подготовленных исполнителем на различных этапах проектирования. Эта проверка включает в себя плагиат-контроль – выявления плагиата.

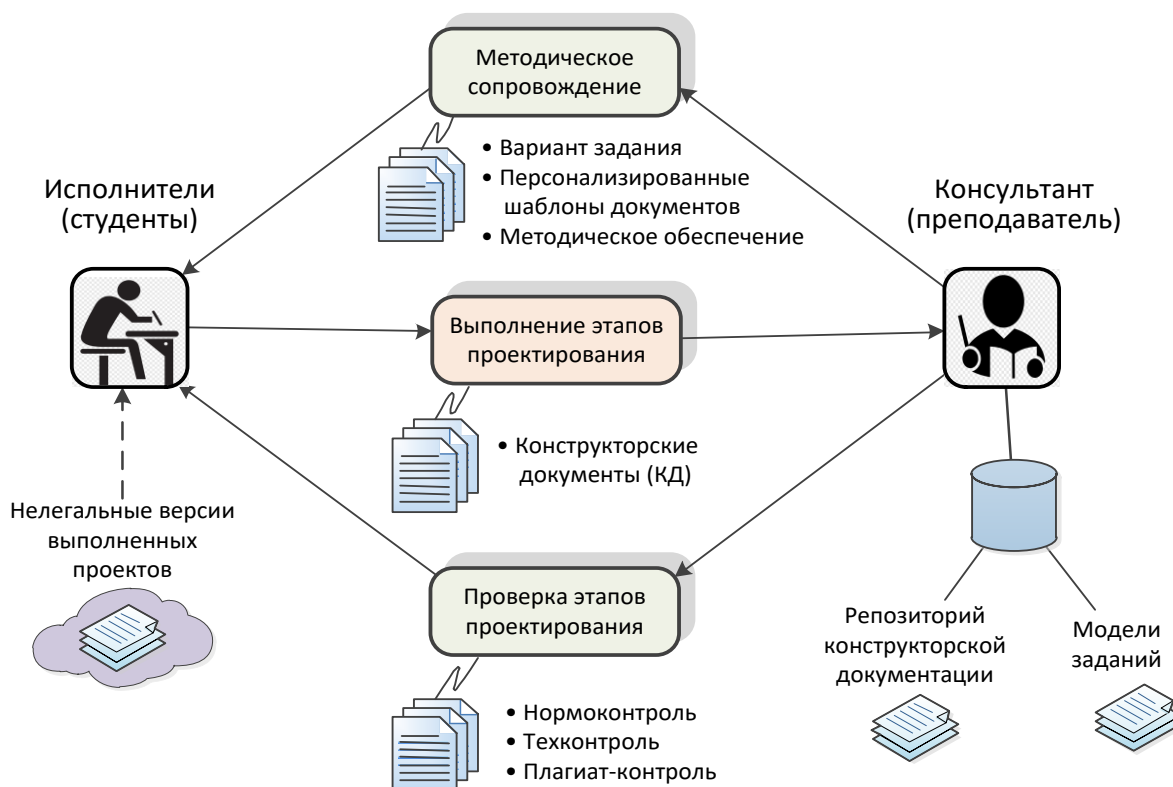


Рис. 1 Укрупненная схема процесса учебного проектирования.

В данной статье рассматривается не исследованная ранее задача выявления плагиата в учебных графических конструкторских документах. Специфические сложности при выявлении плагиата возникают на инженерных специальностях, предусматривающих разработку конструкторской документации, в том числе графической – схем и чертежей. Если вопросы выявления плагиата в текстовых документах хорошо исследованы и достаточно проработаны в информационно-технологическом плане, то графические конструкторские документы имеют специфику, не позволяющую напрямую использовать известные подходы.

Специфика этой задачи состоит в двухмерности пространства графического документа, в котором размещаются и соединяются между собой объекты – условные графические обозначения, что отличает ее от хорошо исследованной одномерной задачи обнаружения плагиата в текстовых документах. Графические конструкторские документы представляют собой схемы, составленные из условных графических обозначений (УГО), и чертежи, изображающие особенности реальных объектов в масштабе. И чертежи, и схемы могут, кроме того, содержать текстовые данные, характеризующие или поясняющие графическую информацию.

Современные открытые форматы векторной графики, применяемые для создания графических конструкторских документов, допускают программную обработку и анализ содержимого. Это в принципе позволяет автоматизировать процесс выявления плагиата. Возможны различные идеи и направления, ниже они обсуждаются на концептуальном уровне, и детально прорабатывается подход, основанный на сочетании цифровых водяных знаков и стеганографии для графических документов в форматах графического редактора Visio. При этом мы фокусируемся на задаче выявления плагиата в схемах, не затрагивая специфику чертежей.

СТЕПЕНЬ ПРОРАБОТАННОСТИ ТЕМЫ

В процессе анализа известных работ не удалось выявить каких-либо публикаций, посвященных плагиату в графических конструкторских документах, тем более учитывающих специфику учебного процесса. Из известных работ в этой области необходимо отметить обширные исследования по выявлению плагиата в текстах. Интерес представляют работы в области

цифровых водяных знаков, компьютерной и цифровой стеганографии, а также методы программной обработки графических документов. Эти подходы и результаты могут использоваться при решении рассматриваемой задачи.

Выявление плагиата в текстовых данных. Как уже отмечалось, эти вопросы хорошо проработаны как теоретически, так и практически (ru.wikipedia.org/wiki/Выявление_плагиата). Методы, основанные на оценке локального сходства («дактилоскопия», анализ встречаемости терминов) и оценке сходства в целом (основанной на цитировании, основанной на стилометрии) [1–4], достаточно успешно выявляют простой copy-paste-плагиат (C&P Plagiarism), а также преодолевают «маленькие хитрости» «замаскированного плагиата» (Disguised Plagiarism) (в виде незначительной переработки текста, такой как замена и перестановка словосочетаний в предложении и др.) [5–7]. Сложнее выявляется «плагиат перефразирования» (Paraphrase Plagiarism) и перевода на другой язык (Translation Plagiarism). Особенно сложно выявить автоматизированным путем «плагиат идей» (Idea Plagiarism).

Отметим также примеры применения этих подходов в учебном процессе для поиска плагиата в исходном коде компьютерных программ [8, 9]. Эти методы вполне применимы к другим видам учебных текстовых КД.

Цифровые водяные знаки и стеганография. Цифровой водяной знак (ЦВЗ, Digital watermark) – технология защиты авторских прав цифровых файлов на основе внедрения в них текста или изображения, идентифицирующего автора (ru.wikipedia.org/wiki/Цифровой_водяной_знак). Этот подход нередко совмещают со стеганографией (Steganography) (ru.wikipedia.org/wiki/Стеганография) – методами тайнописи, когда скрывается сам факт внедрения ЦВЗ [10–18]. Различают компьютерную стеганографию, использующую особенности компьютерных платформ (незадействованные области форматов файлов, подмена символов в названиях файлов и т. д.), и цифровую, основанную на внедрении сообщений в мультимедийные объекты (растрово-графические, аудио-, видеофайлы) с незаметными для глаза или уха искажениями.

Методы ЦВЗ и стеганографии представляют интерес для использования в графических КД в целом или его отдельных частей для идентификации исполнителя. Эти вопросы требуют дополнительного исследования.

Программная обработка графических документов. Графические КД создаются, как правило, компьютерными методами в среде графических редакторов в форматах векторной графики. Для автоматизированного выявления плагиата требуется доступ к внутренней структуре объектов графического документа в соответствии с используемым графическим форматом. Для этого необходимо, чтобы использовались открытые форматы для документов. Нужно знать, как задаются внутри документа те или иные элементы схемы или чертежа, чтобы компьютерная программа могла найти и извлечь нужную информацию. При использовании закрытых форматов требуется соответствующий программный интерфейс.

В настоящее время при создании учебных КД, как правило, используются графические редакторы (Visio, OpenOffice, LibreOffice Draw и т. п.), которые поддерживают открытые графические форматы. Это позволяет программам выявления плагиата непосредственно работать с внутренней структурой документов. Обсуждение связанных с этим вопросов имеется в работе [19].

ОБСУЖДЕНИЕ РЕШАЕМОЙ ЗАДАЧИ И ПРЕДЛАГАЕМОГО ПОДХОДА

Итак, имеющиеся наработки выявления плагиата в текстовых документах, подходы и методы цифровых водяных знаков и стеганографии, возможности программной обработки графических документов служат основой для решения новой задачи – выявления плагиата в графических документах в ходе учебного проектирования.

Подчеркнем, что решаемая задача относится к учебному, а не к реальному проектированию, поскольку при ее решении существенными являются некоторые специфические особенности учебного процесса, такие как типизация вариантов заданий учебного проектирования и повторяемость процесса во времени от семестра к семестру.

Нас прежде всего интересует С&Р-плагиат, а также плагиат «перерисовывания». С&Р-плагиат, при котором плагиатор копирует схему или ее фрагменты из чужого КД и вставляет их в свой, – недопустимый и вредный проступок, бесполезный в учебном плане. Плагиат «перерисовывания», при котором схема «перерисовывается» исполнителем по образцу схемы из чужого КД, – тоже недопустимый, хотя и менее вредный проступок (мы считаем плагиат такого рода менее вредным, поскольку в процессе «перерисовывания» у исполнителя больше возможностей почерпнуть что-то полезное).

Модель управления процессом учебного проектирования

На рисунке 2 представлена модель в виде основных контуров управления процессом учебного проектирования.

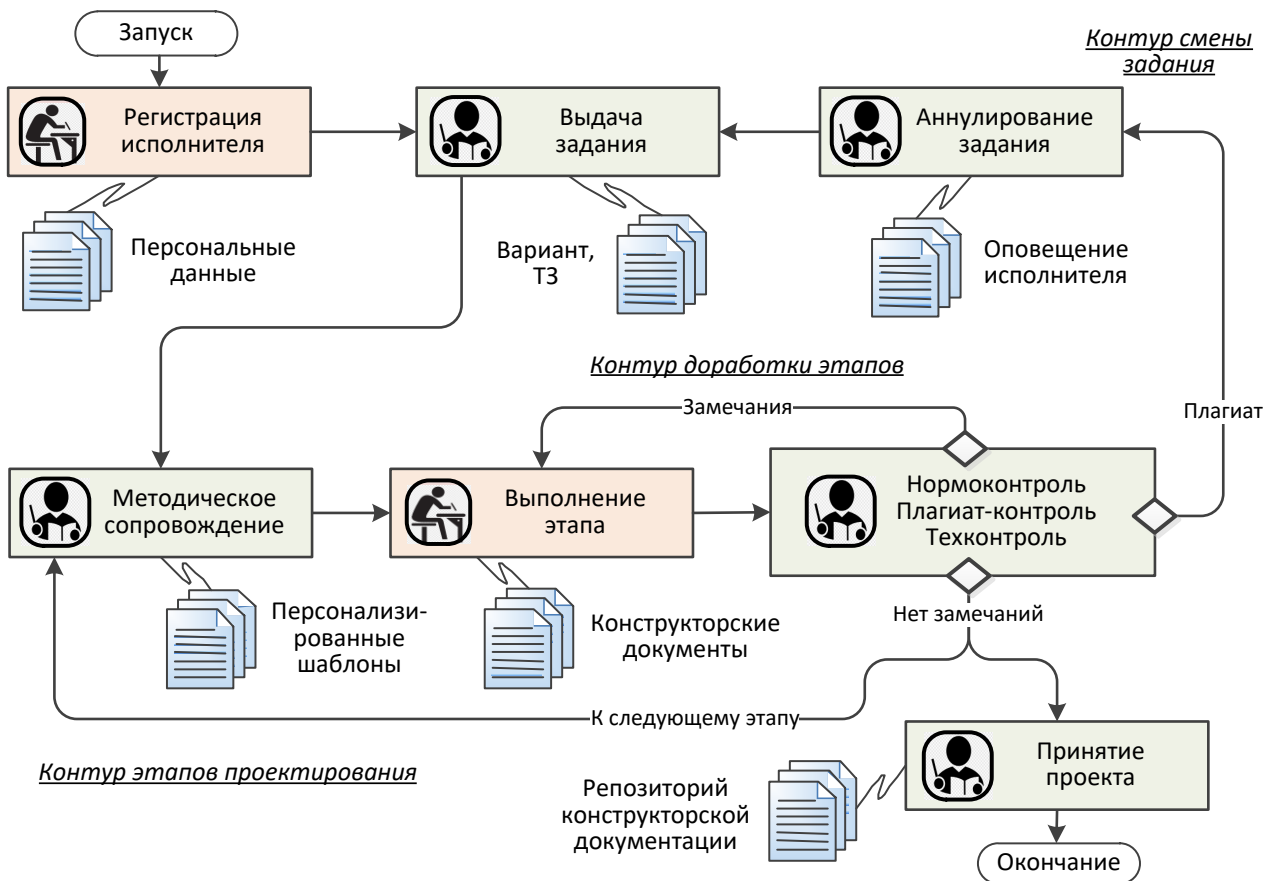


Рис. 2 Основные контуры управления процессом учебного проектирования

Процесс в целом включает регистрацию исполнителя, выдачу ему задания консультантом, последовательное выполнение этапов проектирования, принятие проекта консультантом. Процесс выполнения этапов содержит ряд вложенных контуров обратной связи. Контур этапов проектирования включает методическое сопровождение этапа со стороны консультанта, выполнение этапа исполнителем (результатом чего являются конструкторские документы), проверку конструкторских документов этапа консультантом. Этот контур содержит в себе контур доработки этапа, активизирующийся в случае замечаний по результатам нормо- и техконтроля. Также имеется контур смены (отмены) задания, активизирующийся в случае обнаружения плагиата.

Таким образом, процесс проектирования управляется по результатам проверки разработанных КД: 1) нормоконтроля, т. е. проверки формальных требований к документу; 2) техконтроля – проверки документа по существу содержащихся технических решений; 3) плагиат-контроля – проверки документа на наличие плагиата.

Отметим существенное обстоятельство: разработка конструкторских документов ведется на основе персонализированных шаблонов-заготовок, которые генерируются с учетом персональных данных исполнителя и особенностей технического задания проекта [20–23]. Это позволяет соответствующим образом подготовиться к последующему автоматизированному выявлению плагиата.

Существо предлагаемого подхода

Принципы, лежащие в основе предлагаемого подхода к выявлению плагиата, иллюстрируются на рисунке 3.

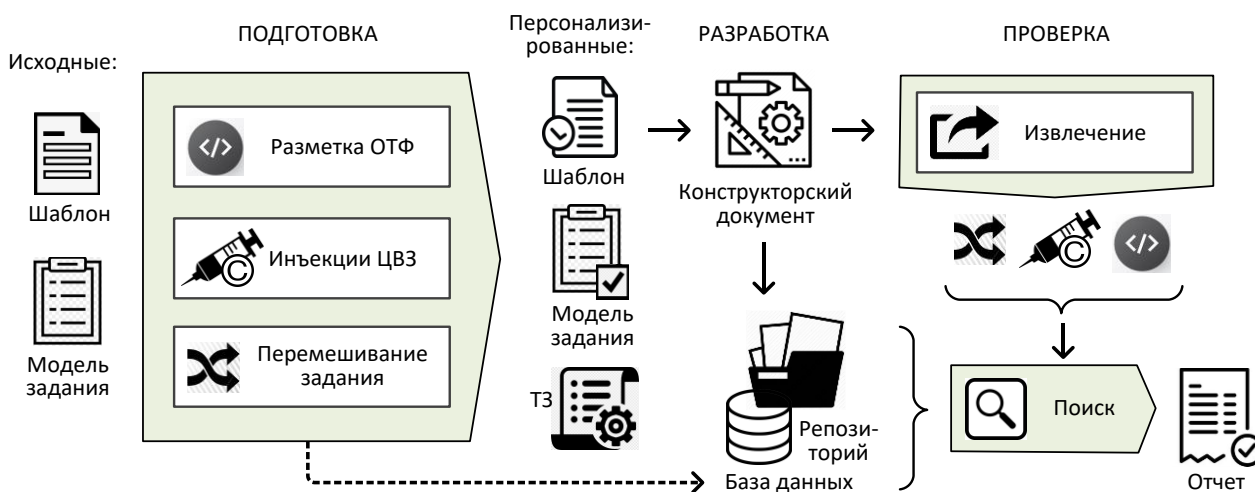


Рис. 3 Предлагаемый подход к выявлению плагиата

В общем виде они сводятся к следующему:

1. Проектирование КД ведется на основе персонализированных шаблонов-заготовок конструкторского документа (ПШ) и персонализированных технических заданий (ТЗ), которые генерируются на этапе подготовки так, чтобы в последствии на этапе проверки можно было идентифицировать исполнителя документа.

2. ПШ и ТЗ генерируются на основе общего шаблона КД и формализованной модели задания проектирования с учетом персональных данных исполнителя. При этом в ПШ и ТЗ вносятся не сообщаемые исполнителю идентификационные трансформации (ИТ).

3. Сигнатуры внесенных ИТ сохраняются в базе данных с привязкой к исполнителю для последующей идентификации по ним исполнителя на этапе проверки.

4. На этапе разработки КД действия исполнителя ограничиваются так, чтобы не нарушались сигнатуры внесенных ИТ, тем самым сохранялась возможность последующей идентификации исполнителя.

5. На этапе проверки КД из него извлекаются (вычисляются) сигнатуры ИТ, которые сравниваются с сигнатурами исполнителя, сохраненными в базе данных. В случае несовпадения фиксируется факт плагиата. Поиск в базе данных по сигнатурам других исполнителей позволит идентифицировать истинных исполнителей КД (или его частей).

Виды идентификационных трансформаций и их гранулярность

Обсудим вопрос о том, какие ИТ можно применять к КД и на каких уровнях КД это может быть сделано. На рисунке 3 указаны три достаточно очевидные трансформации: разметка ОТФ

(оригинальных текстовых фрагментов); инъекции ЦВЗ; перемешивание задания. Начнем с инъекций ЦВЗ.

Инъекции ЦВЗ. Предусматривается вставка в КД ЦВЗ, однозначно идентифицирующих исполнителя, для которого создается ПШ. Если на этапе проверки извлеченный ЦВЗ соответствует другому исполнителю, то это достаточно надежно свидетельствует о плагиате. Возможен случай, когда ЦВЗ не соответствует проверяемому исполнителю, но не найдено соответствие в базе других исполнителей. Или требуемый ЦВЗ вообще отсутствует. Это может быть следствием неполной базы данных либо нарушения исполнителем каких-то ограничений по сохранению сигнатур при разработке КД (см. п. 4 в списке выше).

Возможна различная степень гранулярности ЦВЗ:

- на уровне целого КД, когда ЦВД идентифицирует исполнителя документа целиком;
- на уровне условных графических обозначений в схеме, когда ЦВЗ идентифицирует исполнителей объектов, соответствующих данному условному обозначению;
- на уровне отдельных элементов условного графического обозначения, когда ЦВЗ идентифицирует отдельные фигуры, составляющие условное обозначение.

Таким образом, инъекция ЦВЗ позволяет эффективно обнаружить С&Р-плагиат. Плагиат «перерисовывание» этим способом не обнаруживается.

Разметка ОТФ. Оригинальные текстовые фрагменты – это такие фрагменты текстовых данных, которые создаются исполнителем (например, имена новых объектов на схеме, комментарии и т. п.). Содержимое ОТФ может служить отличительной особенностью данного конкретного КД. Разметка ОТФ нужна для того, чтобы отличить их от неоригинальных фрагментов, содержимое которых стандартное, определяется типом КД или прописано в ТЗ и не может служить отличительной особенностью данного конкретного КД.

Опыт организации учебного проектирования показывает, что в ходе «перерисовывания» схемы студенты-плагиаторы склонны дословно воспроизводить ОТФ. Поэтому к содержимому ОТФ можно применить методы выявления плагиата в текстовых данных. Таким образом, разметка ОТФ может помочь в выявлении плагиата типа «перерисовывания».

Сигнатура, сохраняемая в базе данных в данном случае, – это само текстовое содержимое фрагментов с привязкой к объектам документа. Гранулярность ОТФ может быть на уровне всего КД, т. е. объединять все фрагменты с привязкой к документу в целом. Другой вариант – множество ОТФ, привязанных к отдельным объектам КД. Данный вопрос требует отдельного исследования.

Перемешивание задания. Этот вид трансформации состоит во внесении персональных изменений в упорядочивание списков в техническом задании, выдаваемом исполнителю. Опыт автора по руководству учебным проектированием показывает, что последовательность размещения объектов в КД во многом коррелируется с последовательностью упоминания этих объектов в ТЗ. Отклонение от этой последовательности в проверяемом КД может быть косвенным признаком того, что КД выполнен на основе другого КД.

Идея перемешивания задания состоит в том, чтобы на этапе персонализации ТЗ вносить изменения в последовательность размещения объектов в определенных списках требований, параметров и т. п. и запоминать эту последовательность как сигнатуру в базе данных. На этапе проверки вычислять сигнатуру, соответствующую реальной последовательности этих объектов в КД. Сравнивая с базой данных, можно подобрать возможных кандидатов на КД-первоисточник. Таким образом, перемешивание задания может помочь в выявлении плагиата.

РЕАЛИЗАЦИЯ ПРЕДЛОЖЕННОГО ПОДХОДА

В настоящее время предложенный подход частично реализован в виде информационной технологии и опробован на практике в ходе учебного проектирования концептуально-логических моделей базы данных бизнес-процесса по дисциплине «Базы данных». В рамках курсового проектирования предусмотрена поэтапная разработка 12 КД, из них 8 схем моделей базы

данных разного уровня абстракции и охвата. Схемы выполняются студентами-исполнителями в среде векторного графического редактора Microsoft Visio. Процесс проектирования поддерживается веб-порталом (hsm.ugatu.su/artem/dbproj), созданным для этого на базе ситуационно-ориентированного подхода [24, 25]. Проектант-исполнитель регистрируется на портале, получает одно из типовых заданий. На каждом этапе для исполнителя генерируется персонализированный шаблон документа Visio, на основе которого исполнитель рисует результирующую схему модели. Готовый документ загружается на портал и проверяется сканером (нормоконтроль) и консультантом (техконтроль). После устранения ошибок, выявленных сканером, и замечаний, сделанных консультантом, происходит переход к следующему этапу.

Особенности применения ЦВЗ. Подход реализован в части выявления Р&С-плагиата на основе ЦВЗ. Применяемые ЦВЗ представляют собой цифровые идентификаторы исполнителя, встраиваемые в каждую фигуру, соответствующую условному графическому обозначению (т. е. в данном случае реализована гранулярность на уровне условных графических обозначений). Персонализированный шаблон-заготовка КД содержит набор образцов УГО-фигур, предназначенных для построения данного вида схемы. ЦВЗ встраиваются в каждый образец при генерации шаблона-заготовки.

В ходе построения схемы исполнителю запрещено использовать УГО-фигуры, кроме тех, которые приведены в полученном шаблоне-заготовке (нельзя самому создавать «с нуля» УГО-фигуры, допустимо только копировать и настраивать имеющиеся в шаблоне образцы). Нарушение этого требования автоматически выявляется сканером-нормоконтролёром. Данное ограничение гарантирует, что каждая УГО-фигура в корректно построенной схеме содержит ЦВЗ.

В ходе плагиат-контроля – проверки готовой схемы на плагиат – выполняется циклическое сканирование УГО-фигур в документе:

- отыскивается ЦВЗ УГО-фигуры. При отсутствии ЦВЗ фиксируется соответствующая ошибка;
- из ЦВЗ извлекается идентификатор исполнителя и сравнивается с идентификатором проверяемого исполнителя. При несовпадении значений идентификаторов фиксируется соответствующая ошибка.

По результатам сканирования формируется отчет об обнаруженных ошибках. По выявленном в ЦВЗ несовпадающим идентификаторам исполнителей в репозитории отыскивается документ, из которого было произведено незаконное копирование фигур.

Особенности размещения ЦВЗ в фигурах Visio. Техническая реализация описанной выше схемы требует ответа на вопрос о способе размещения ЦВЗ на внутреннем уровне файла Visio с учетом особенностей графического формата. Для первого варианта реализации был выбран самый простой путь – использовать определяемые пользователем ячейки (ОП-ячейки), предусмотренные графическими форматами Visio.

Графический редактор Visio использует два открытых формата, которые основаны на языке разметки XML и позволяют обрабатывать графические документы внешними программами – форматы VDX и VSDX. Первый формат представляет собой единый XML-файл, а второй – ZIP-архив, содержащий иерархию папок с XML-файлами, задающими отдельные части графического документа. Оба формата для каждой фигуры изображения предусматривают возможность задания ОП-ячеек, содержащих некоторые значения, относящиеся к этой фигуре. На рисунке 4 представлены модели, поясняющие место ОПЯ в общей структуре этих графических форматов (подробное описание приведено в [19]). Каждая фигура в обоих форматах задается соответствующим XML-элементом Shape. В формате VDX каждой ОП-ячейке соответствует вложенный XML-элемент User, имя ячейки задается XML-атрибутом NameU, а значение – вложенным XML-элементом Value. В формате VSDX множеству ОП-ячеек соответствует вложенный XML-элемент Section с атрибутом N='User', отдельная ячейка задается вложенным XML-элементом Row, имя ячейки задается атрибутом N, а значение – вложенным

XML-элементом Cell с атрибутом N='Value'. Представленные на рисунке 4 модели служат основой для извлечения значений ЦВЗ при сканировании фигур КД.

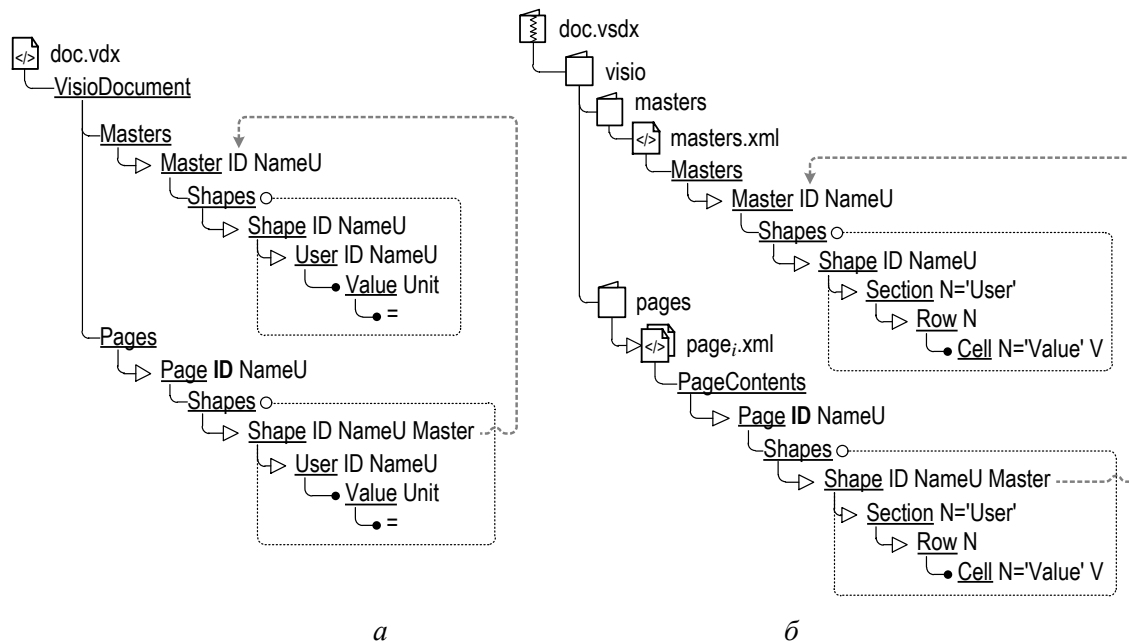


Рис. 4 Модели доступа к ЦВЗ в определяемых пользователем ячейках, документа Visio:
а – формат VDX; б – формат VSDX

На рисунке 5 иллюстрируется отображение ЦВЗ редактором Visio. Вверху рисунка приведен фрагмент КД «Модель локальная иерархическая». Основное окно отображает КД с результирующей схемой модели, а слева расположено окно образцов фигур, соответствующих УГО, из которых составлена схема. Внизу рисунка показан фрагмент предоставляемой редактором таблицы свойств выбранной фигуры схемы. Все ОПЯ этой фигуры представлены в разделе User-defined Cells, одна из строк этого раздела соответствует внедренному в фигуру ЦВЗ.

Отметим еще одну существенную техническую особенность применяемого способа внедрения ЦВЗ, связанную с концепцией образцов фигур в документах Visio. Конкретный экземпляр фигуры в графическом документе может быть создан тремя путями: 1) нарисован вручную, используя инструменты графического редактора (запрещено в данном случае); 2) скопирован из меню образцов фигур, прикрепленных к документу, и затем модифицирован; 3) скопирован из уже имеющегося изображения в документе и затем модифицирован. Поэтому при генерации персонализированного шаблона ЦВЗ внедряются как в образцы УГО-фигур, прикрепленные к шаблону, так и в УГО-фигуры, уже присутствующие на листах шаблона.

Обсуждение результатов практического использования

Разработанная технология использовалась автором для плагиат-контроля курсовых проектов, расчетно-графических работ в течение двух семестров. Всего через учебное проектирование с применением автоматизированного плагиат-контроля прошло более 250 студентов.

За этот период было выявлено около 30 случаев нарушений ЦВЗ. Среди них только 5 нарушений являются достоверно установленным плагиатом, когда исполнитель копировал фигуры из ранее выполненных тех же заданий. Из оставшихся примерно половина нарушений – это «псевдоплагиат» – студенты-исполнители копировали схемы, соответствующие другим вариантам задания, а потом модифицировали их для своего задания. Вторая половина случаев нарушения – это случаи отсутствия ЦВЗ. Эти случаи соответствуют копированию схемы из ранних конструкторских документов, выполненных до введения ЦВЗ. Отсутствие идентификатора исполнителя не позволяет установить, был ли это плагиат или «псевдоплагиат».

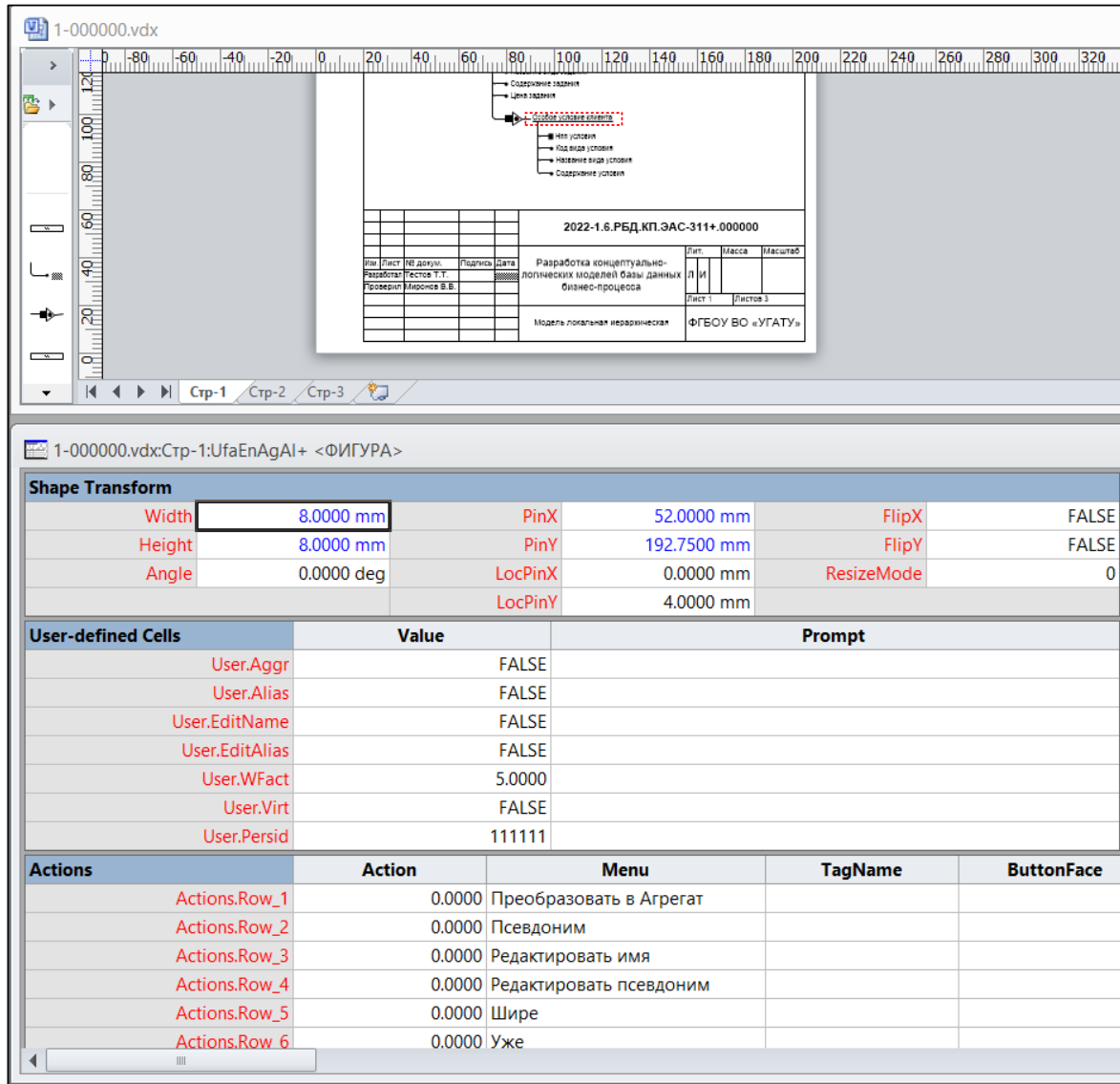


Рис. 5 Ячейки, определяемые пользователем, в документе Visio

Другая особенность: случаи плагиата в студенческих группах резко уменьшались после первых обнаружений и принятия соответствующих организационных мер. По-видимому, это объясняется тем, что факты выявления плагиата быстро становятся известны в студенческом сообществе, что уменьшает соблазн нарушения правил.

В целом автоматизация этого процесса заметно снизила трудоемкость плагиат-контроля, практически полностью избавив консультанта от необходимости трудоемкого ручного сличения «подозрительной» схемы с другими схемами в репозитории для этого варианта задания.

Отметим, что использованный способ задания ЦВЗ в ОП-ячейках не слишком надежен в плане скрытности, поскольку ОП-ячейки доступны для анализа и модификации стандартными средствами графического редактора. Этот способ применим первое время, пока студенты не разобрались, что к чему. В дальнейшем предполагается ввести ряд мер, повышающих скрытность ЦВЗ на основе стеганографии. Пока нереализованными остаются трансформации разметка ОТФ и перемешивания ТЗ, обсуждавшиеся выше, – это задача на будущее.

ЗАКЛЮЧЕНИЕ

Итак, рассмотрена не исследованная ранее задача выявления плагиата в графических конструкторских документах в процессе учебного проектирования. Специфика этой задачи со-

стоит в двухмерности пространства графического документа, в котором размещаются и соединяются между собой объекты – условные графические обозначения, что отличает ее от хорошо исследованной одномерной задачи обнаружения плагиата в текстовых документах. Другая особенность рассматриваемой задачи состоит в том, что учебное проектирование, основанное на типовых заданиях, сопровождается утечкой в Сеть конструкторских документов прошлых лет, что расширяет возможности и соблазны для плагиата.

Предлагаемый подход к решению задачи основан на применении персонализированных шаблонов-заготовок, содержащих набор фигур условных графических обозначений, которые генерируются для исполнителя конструкторского документа таким образом, что это позволяет выявить в итоговом документе заимствования из других документов. В частности, предложено использовать цифровые водяные знаки, внедряемые в условные графические обозначения схемы для идентификации исполнителя. Несоответствие ЦВЗ фигуры схемы исполнителю документа свидетельствует о том, что данная фигура была скопирована из другой схемы.

Этот подход реализован для конструкторских документов, выполненных в открытых графических XML-форматах редактора Visio. ЦВЗ реализованы в виде идентификаторов исполнителя, значения которых размещаются в определяемых пользователем ячейках, предусмотренных форматами Visio.

Разработанная информационная технология успешно применяется в учебном процессе проектирования графических моделей базы данных. В результате заметно снизилась трудоемкость плагиат-контроля по отношению к неавтоматизированной процедуре сличения «подозрительных» схем с другими схемами в репозитории конструкторских документов.

СПИСОК ЛИТЕРАТУРЫ / REFERENCES

1. Wahle Jan Philip, Ruas Terry, Foltýnek Tomáš, Meuschke Norman, Gipp Bela, Smits Malte (ed.). Identifying Machine-Paraphrased Plagiarism. Information for a Better World: Shaping the Global Future, Cham: Springer International Publishing, 2022. Vol. 13192. Pp. 393–413. arXiv:2103.11909, doi:10.1007/978-3-030-96957-8_34.
2. Reimers Nils, Gurevych Iryna. Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks. 2019. arXiv:1908.10084.
3. Bensalem Imene, Rosso Paolo, Chikhi Salim. On the use of character n-grams as the only intrinsic evidence of plagiarism// Language Resources and Evaluation. 2019. 53 (3): 363–396. doi:10.1007/s10579-019-09444-w.
4. Gipp Bela. Citation-based Plagiarism Detection, Springer Vieweg Research, 2014. ISBN 978-3-658-06393-1.
5. Bensalem Imene. Intrinsic Plagiarism Detection: A Survey//Plagiarism Detection: A focus on the Intrinsic Approach and the Evaluation in the Arabic Language (PhD thesis). 2020.
6. Foltýnek Tomáš, Meuschke Norman, Gipp Bela. Academic Plagiarism Detection: A Systematic Literature Review. 16 October 2019. ACM Computing Surveys. 52 (6): 1–42. doi:10.1145/3345317.
7. Седов А. В., Рогов А. А. Анализ неоднородностей в тексте на основе последовательностей частей речи // Современные проблемы науки и образования. 2013. Вып. 1. URL: <https://science-education.ru/ru/article/view?id=8339> (дата обращения: 20.04.2023). [[Sedov A. V., Rogov A. A. In: Sovremennyye problemy nauki i obrazovaniya, 2013, Issue 1. URL: <https://science-education.ru/ru/article/view?id=8339>. (in Russian)]]
8. Морозов П. А. Алгоритм выявления плагиата в исходном коде программ // Информатика: проблемы, методы, технологии: Мат-лы XXI Международной научно-методической конференции. Воронеж, 2021. С. 1457–1464. [[Morozov P. A. In: Informatika: problemy, metody, tekhnologii. Proc. XXI Int. Sci. Conf. Voronezh, Russia, 2021, pp. 1457-1464. (in Russian)]]
9. Ефимчик Е. А., Цибин А. И. Анализ и визуализация плагиата исходного кода в практических курсах по программированию // Компьютерные инструменты в образовании. 2020. №. 4. С. 79–92 [[Efimchik E. A., Tsibin A. I. In: Komp'yuternye instrumenty v obrazovanii, 2020, no. 4, pp. 79-92. (In Russian)]]
10. Halder Raju, Pal Shantanu, Cortesi Agostino. Watermarking Techniques for Relational Databases: Survey, Classification and Comparison//Journal of Universal Computer Science. 2010. 16 (21): 3164–3190. CiteSeerX 10.1.1.368.1075.
11. Zigmithos Athanasios, Papageorgiou Achilleas, Patsakis Constantinos. Social network content management through watermarking//2012 IEEE 11th International Conference on Trust, Security and Privacy in Computing and Communications. IEEE. Pp. 1381–1386. doi:10.1109/TrustCom.2012.264. ISBN 978-1-4673-2172-3. S2CID 17845019.
12. Shih F. Y. Digital watermarking and steganography: fundamentals and techniques. CRC press, 2017.
13. Byrnes O. et al. Data hiding with deep learning: A survey unifying digital watermarking and steganography //arXiv preprint arXiv:2107.09287. 2021.
14. Zhang C. et al. Udh: Universal deep hiding for steganography, watermarking, and light field messaging //Advances in Neural Information Processing Systems. 2020. T. 33. C. 10223–10234.
15. Hossen M. S. et al. A new approach to hiding data in the images using steganography techniques based on AES and RC5 algorithm cryptosystem // 2020 International Conference on Smart Electronics and Communication (ICOSEC). IEEE. 2020. C. 676–681.

16. Грибунин В., Оков И., Туринцев И. Цифровая стеганография. Litres, 2022. [[Gribunin V., Okov I., Turintsev I. Tsifrovaya steganografiya. Litres, 2022. (in Russian)]]
17. Федосеев В. А. и др. Цифровые водяные знаки и стеганография. Самара: Изд-во Самар. ун-та, 2019. [[Fedoseev V. A. et al. Tsifrovye vodyanye znaki i steganografiya. Samara, Russia: Samara Univ. Publ., 2019. (in Russian)]]
18. Коробкин А. И. Встраивание цифровых водяных знаков в изображения стеганографическим методом // Интеллектуальные, сенсорные и мехатронные системы – 2022. Минск: БНТУ, 2022. С. 14-19. [[Korobkin A. I. In: Intellekтуal'nye, sensornye i mekhatronnye sistemy — 2022. Minsk: BNTU, 2022, pp. 14-19. (in Russian)]]
19. Миронов В. В., Гусаренко А. С., Тугузбаев Г. А. Извлечение семантической информации из графических схем // Информатика и автоматизация. 2021. Т. 20. № 4. С. 940–970. [[Mironov V. V., Gusarenko A. S., Tuguzbaev G. A. In: Informatika i Avtomatizatsiya, 2021, vol. 20, no. 4, pp. 940-970. (in Russian)]]
20. Миронов В. В., Гусаренко А. С., Тугузбаев Г. А. Ситуационно-ориентированные базы данных: формирование персонализированных графических документов для поддержки учебного проектирования // Моделирование, оптимизация и информационные технологии. 2020. Т. 8. № 2 (29). [[Mironov V. V., Gusarenko A. S., Tuguzbaev G. A. In: Modelirovanie, Optimizatsiya i Informatsionnye Tekhnologii = Modeling, Optimization and Information Technology, 2021, vol. 20, no. 4, pp.]]
21. Миронов В. В., Гусаренко А. С., Диметриев Р. Р., Сарваров М. Р. Создание персонализированных документов на основе ситуационно-ориентированной базы данных // Вестник Уфимского государственного авиационного технического университета. 2014. Т. 18. № 4 (65). С. 191–197. [[Mironov V. V., Gusarenko A. S. et al. In: Vestnik UGATU, 2014, vol. 18, no. 4 (65), pp. 191-197. (in Russian)]]
22. Миронов В. В., Шакирова Г. Р., Яфаев В. Э. Иерархическая модель персонализированных документов и ее XML-реализация // Вестник Уфимского государственного авиационного технического университета. 2008. Т. 11. № 1. С. 164–174. [[Mironov V. V., Shakirova G. R. et al. In: Vestnik UGATU, 2008, vol. 11, no. 1, pp. 164-174. (in Russian)]]
23. Миронов В. В., Шакирова Г. Р. Программно-инструментальное средство для создания и ведения динамических xml-документов // Вестник Уфимского государственного авиационного технического университета. 2007. Т. 9. № 5. С. 54–63. [[Mironov V. V., Shakirova G. R. In: Vestnik UGATU, 2007, vol. 9, no. 5, pp. 54-63. (in Russian)]]
24. Миронов В. В., Юсупова Н. И., Гусаренко А. С. Ситуационно-ориентированные базы данных: современное состояние и перспективы исследования // Вестник Уфимского государственного авиационного технического университета. 2015. Т. 19. № 2 (68). С. 188–199. [[Mironov V. V., Yusupova N. I., Gusarenko A. S. In: Vestnik UGATU, 2015, vol. 19, no. 2 (68), pp. 188-199. (in Russian)]]
25. Миронов В. В., Гусаренко А. С., Юсупова Н. И. Структурирование виртуальных мультимедиа документов в ситуационно ориентированных базах данных с помощью entry-элементов // Труды СПИИРАН. 2017. № 4 (53). С. 225–240. [[Mironov V. V., Yusupova N. I., Gusarenko A. S. In: Trudy SPIIRAN, 2017, no. 4 (53), pp. 255-240. (in Russian)]]

Поступила в редакцию 25 апреля 2023 г.

МЕТАДАННЫЕ / METADATA

Title: Identify Plagiarism in Graphic Design Documents During Educational Design.

Abstract: The article deals with the previously unexplored problem of identifying plagiarism in graphic design documents in the process of educational design of various schemes. The specificity of this problem is noted in comparison with the problem of detecting plagiarism in one-dimensional text documents, due to the placement and connection of conventional graphic symbols on a diagram in a two-dimensional space. Instructional design based on generic assignments increases the scope for plagiarism, as the assignments may have been done repeatedly before. The proposed approach to solving the problem is based on the use of personalized blank templates generated for the developer of the design document in such a way that it makes it possible to identify borrowings from other documents in the finished document. It is proposed to use digital watermarks, which are embedded in the conventional graphic designations of the scheme to identify the performer. The implementation of this approach for design documents in graphic formats of the Visio editor is described. The application of the developed information technology in the educational process is discussed.

Key words: plagiarism; graphic design document; digital watermark; educational design.

Язык статьи / Language: русский / Russian.

Поддержка/Support: Уфимский университет науки и технологий / Ufa University of Science and Technologies.

Об авторах / About authors:

МИРОНОВ Валерий Викторович

ФГБОУ ВО «Уфимский университет науки и технологий», Россия. Проф. каф. автоматизированных систем управления. Дипл. радиофизик (Воронежск. гос. ун-т, 1975). Д-р техн. наук по упр. в техн. системах (Уфимск. гос. авиац. техн. ун-т, 1995). Иссл. в обл. иерархич. моделей и ситуац. управления. E-mail: mironov@list.ru
ORCID: <https://orcid.org/0000-0002-0550-4676>
URL: elibrary.ru/author_profile.asp?authorid=691759

MIRONOV Valeriy Viktorovich

Ufa University of Science and Technologies, Russia. Prof., Dept. of Automated Control Systems. Dipl. Radio physicist (Voronezh State Univ., 1975). Dr. of Tech. Sci. (Ufa State Aviation Technical University, 1995). Research in hierarchical models and situational management. E-mail: mironov@list.ru
ORCID: <https://orcid.org/0000-0002-0550-4676>
URL: elibrary.ru/author_profile.asp?authorid=691759