

**Естественные
и
технические
науки[®]**

№ 4 (142) 2020 г.

ISSN 1684-2626

**Журнал «Естественные и технические науки» входит
в Перечень ВАК РФ и Международную базу данных
и систему цитирования Chemical Abstracts.**

Журнал публикует основные научные результаты диссертаций на соискание ученой степени кандидата наук и на соискание ученой степени доктора наук по естественным и техническим наукам. В соответствии с пунктом 5 Правил издания, текущие номера которых или их переводные версии входят хотя бы в одну из международных реферативных баз данных и систем цитирования, считаются включенными в Перечень по отраслям науки, соответствующим их профилю. При рассмотрении вопроса о присвоении ученого звания публикации соискателя ученого звания в данных изданиях засчитываются в качестве научных трудов, опубликованных в рецензируемых научных изданиях, включенных в Перечень (на основании пункта 5 Правил формирования перечня рецензируемых научных изданий, утвержденных приказом Минобрнауки России от 12.12.2016 № 1586, и письма Минобрнауки МН-06.2/1059 от 08.11.2018).

Учредитель – Издательство «Спутник+»

Компьютерный набор и верстка – О. Щуклин

Адрес редакции: Россия, 109428, Москва, Рязанский проспект, д. 8А

Телефон: (495) 730-47-74, 778-45-60, 730-48-71 (с 9 до 18, обед с 14 до 15)

E-mail: print@sputnikplus.ru

**Издание зарегистрировано
Министерством Российской Федерации по делам печати,
телерадиовещания и средств массовых коммуникаций**

**Свидетельство о регистрации
ПИ № ФС77-39983 от 20 мая 2010 г.**

Объем 19,38 печ. л.
Тираж 1000 экз. Заказ № 109.
Подписано в печать 28.05.2020.

Отпечатано в ООО «Издательство «Спутник+»

Порядок направления, рецензирования и опубликования научных статей

На основании решения редакционной коллегии журнала
«Естественные и технические науки» № 03 от 01.12.08 утвержден следующий
Порядок рецензирования статей, поступивших в редакцию журнала:

1. Статью необходимо направлять на электронный или почтовый адрес редакции. Поступающая в редакцию статья рассматривается на предмет соответствия профилю журнала, требованиям к оформлению и регистрируется с присвоением ей индивидуального номера. Редакция в течение 3-х дней уведомляет авторов о получении рукописи статьи. Рукописи, оформленные не должным образом, не рассматриваются.

2. Рукописи всех статей, поступивших в редакцию журнала, подлежат обязательному рецензированию. К рецензированию привлекаются ученые, доктора наук, обладающие неоспоримым авторитетом в сфере научных знаний, которой соответствует рукопись статьи, имеющие в течение последних 3-х лет публикации по тематике рецензируемой статьи. Рецензии хранятся в редакции в течение 5 лет. Рецензентом не имеет права быть автор (или один из авторов) рецензируемой статьи. Рецензенты информируются о том, что рукописи статей являются частной собственностью их авторов и представляют собой сведения, исключающие их разглашение и копирование.

3. В случаях, когда редакция журнала не располагает возможностью привлечь к рецензированию эксперта подходящего уровня в сфере знаний, к которой имеет отношение рукопись, редакция обращается к автору с просьбой предоставить внешнюю рецензию. Внешняя рецензия предоставляется при подаче статьи (что, тем не менее, не исключает принятый порядок рецензирования). Рецензии обсуждаются редколлегией, являясь причиной для принятия или отклонения рукописей. Рукопись, адресуемая в редакцию, также может сопровождаться письмом из направляющей организации за подписью ее руководителя.

4. Рецензия должна беспристрастно давать оценку рукописи статьи и заключать в себе исчерпывающий разбор ее научных достоинств и недостатков. Рецензия составляется по предлагаемой редакцией форме или в произвольном виде и должна освещать следующие моменты: научную ценность результатов исследования, актуальность методов исследования и статистической обработки данных, уровень изучения научных источников по теме, соответствие объема рукописи статьи в целом и отдельных ее элементов в частности, т.е. текста, таблиц, иллюстраций, библиографических ссылок. В завершающей части рецензии необходимо изложить аргументированные и конструктивные выводы о рукописи и дать ясную рекомендацию о необходимости либо публикации в журнале, либо переработки статьи (с перечислением допущенных автором неточностей и ошибок).

5. Если в рецензии на статью сделан вывод о необходимости ее доработки, то она направляется автору на доработку вместе с копией рецензии. При несогласии автора с выводами рецензента, автор вправе обратиться в редакцию с просьбой о повторном рецензировании или отозвать статью (в этом случае делается запись в журнале регистрации). Тогда новой датой поступления в редакцию журнала доработанной статьи считается дата ее возвращения. Доработанная статья направляется на повторное рецензирование тому же рецензенту. Редакция журнала оставляет за собой право отклонения рукописи статьи в случае неспособности или нежелания автора учесть пожелания рецензента.

6. Срок рецензирования между датами поступления рукописи статьи в редакцию и вынесения редколлегией решения в каждом отдельном случае определяется ответственным секретарем с учетом создания условий для максимально оперативной публикации статьи, но не более 2-х месяцев со дня получения рукописи.

7. Рецензии на статьи предоставляются редакцией экспертным советам в ВАК по их запросам.

8. Редакция журнала не сохраняет рукописи статей, не принятые к публикации. Рукописи статей, принятые к публикации, не возвращаются.

9. Редакция издания направляет авторам представленных материалов копии рецензий или мотивированный отказ, а также обязуется направлять копии рецензий в Министерство образования и науки Российской Федерации при поступлении в редакцию издания соответствующего запроса.

Главный редактор: Хавкин А.Я.

Редакционная коллегия журнала:

А.Я. Хавкин

главный редактор, доктор технических наук, профессор кафедры нефтегазовой и подземной гидромеханики Российского государственного университета нефти и газа (национального исследовательского университета) им. И.М. Губкина, сопредседатель бюро секции «Нанотехнологии для нефтегазового комплекса» Нанотехнологического общества России, Почетный нефтяник РФ, лауреат премии Миннефтепрома СССР, лауреат премии им. академика И.М. Губкина, лауреат премии им. Н.К. Байбакова, кавалер медали ЮНЕСКО «За вклад в развитие нанонауки и нанотехнологий»

Т.П. Анцупова

доктор биологических наук, профессор кафедры неорганической и аналитической химии Восточно-Сибирского государственного университета технологий и управления, Почетный работник высшего профессионального образования РФ, Почетный работник Восточно-Сибирского государственного университета технологий и управления

А.И. Белолобцев

доктор сельскохозяйственных наук, профессор, заведующий кафедрой метеорологии и климатологии Российского государственного аграрного университета – МСХА им. К.А. Тимирязева

С.С. Валеев

доктор технических наук, профессор, заведующий кафедрой информатики Уфимского государственного авиационного технического университета

И.А. Гарагаиш

доктор физико-математических наук, профессор, главный научный сотрудник, заведующий лабораторией геомеханики Института физики Земли им. О.Ю. Шмидта Российской академии наук, член Российского национального комитета по теоретической и прикладной механике

О.А. Графский

доктор технических наук, профессор вычислительной техники и компьютерной графики Дальневосточного государственного университета путей сообщения

А.В. Дерюгина

доктор биологических наук, доцент, заведующая кафедрой физиологии и анатомии, заместитель директора по научной работе Института биологии и биомедицины, ведущий научный сотрудник лаборатории по разработке методов нейропротекции Национального исследовательского Нижегородского государственного университета им. Н.И. Лобачевского

В.А. Завьялов

доктор технических наук, профессор кафедры автоматизации и электроснабжения Национального исследовательского Московского государственного строительного университета

С.Н. Золотухин

доктор биологических наук, профессор кафедры микробиологии, вирусологии, эпизоотологии и ветеринарно-санитарной экспертизы, декан факультета ветеринарной медицины и биотехнологии Ульяновского государственного аграрного университета им. П.А. Столыпина, Заслуженный деятель науки и техники Ульяновской области

И.И. Иванов

доктор биологических наук, профессор, главный научный сотрудник лаборатории общей биофизики кафедры биофизики биологического факультета Московского государственного университета им. М.В. Ломоносова, лауреат Государственной премии СССР (1983)

<i>Е.А. Калашикова</i>	доктор биологических наук, профессор кафедры генетики, биотехнологии, селекции и семеноводства Российского государственного аграрного университета-МСХА им. К.А. Тимирязева
<i>Ю.Г. Калпин</i>	доктор технических наук, профессор кафедры кузовостроения и обработки давлением Московского государственного технического университета «МАМИ»
<i>В.Ф. Касьянов</i>	доктор технических наук, профессор, зав. кафедрой технической эксплуатации зданий Московского государственного строительного университета, Заслуженный работник высшей школы РФ, Почетный работник профессионального высшего образования РФ, Почетный строитель России и г. Москвы, Почетный работник ЖКХ России
<i>Л.Г. Константинова</i>	доктор биологических наук, профессор, заведующая лабораторией экологии микроорганизмов Института биоэкологии Каракалпакского отделения АН Республики Узбекистан
<i>Т.А. Краснова</i>	доктор технических наук, профессор, заведующая кафедрой аналитической химии и экологии Кемеровского технологического института пищевой промышленности, Почетный работник Высшей школы, Заслуженный эколог РФ
<i>Т.В. Мальцева</i>	доктор физико-математических наук, профессор, проректор по научной и инновационной работе Тюменского государственного архитектурно-строительного университета
<i>Л.Г. Моисейкина</i>	доктор биологических наук, профессор Калмыцкого государственного университета им. Б.Б. Городовикова, Почетный работник высшего образования РФ
<i>А.Н. Николаев</i>	доктор технических наук, профессор, заведующий кафедрой оборудования пищевых производств Казанского государственного технологического университета, профессор кафедры теоретических основ теплотехники Казанского государственного технического университета
<i>Ю.Р. Осипов</i>	доктор технических наук, профессор кафедры теории и проектирования машин и механизмов Вологодского государственного технического университета, Почетный работник высшего образования РФ
<i>Н.Д. Поляхов</i>	доктор технических наук, профессор Санкт-Петербургского государственного электротехнического университета «ЛЭТИ»
<i>О.А. Решетник</i>	доктор технических наук, профессор, заведующая кафедрой технологии пищевых производств Казанского государственного технологического университета, член Общества биотехнологов России им. Ю.А. Овчинникова, член Российского отделения Общества микробиологов, Заслуженный деятель науки и техники Республики Татарстан
<i>О.И. Ручкина</i>	доктор технических наук, профессор кафедры теплогазоснабжения, вентиляции и водоснабжения Пермского национального исследовательского политехнического университета
<i>Ф.Н. Сарапулов</i>	доктор технических наук, профессор, заведующий кафедрой электротехники и электротехнологических систем Уральского государственного технического университета

Уважаемые подписчики!

Вы можете подписаться на любой из наших журналов. Подписка производится как в России, так и за ее пределами.

Подписные индексы наших журналов:

1. «Актуальные проблемы современной науки» – № 41774
2. «Аспирант и соискатель» – № 41535
3. «Вопросы гуманитарных наук» – № 42954
4. «Естественные и технические науки» – № 42943
5. «Вопросы экономических наук» – № 25784
6. «Педагогические науки» – № 26028
7. «Современные гуманитарные исследования» – № 83645

Чтобы подписаться, Вам необходимо вырезать расположенный ниже купон и вписать в него название журнала и индекс. Подписка по Объединенному зеленому каталогу.

Ф. СП-1

Министерство связи РФ
УФПС «Моспочтамт»

АБОНЕМЕНТ на газету
журнал

(индекс названия)

Количество
комплектов

на 2012 год по месяцам:

1	2	3	4	5	6	7	8	9	10	11	12

Куда

(почтовый индекс)

(адрес)

Кому

(фамилия, инициалы)

доставочная карточка

на газету
журнал

(индекс названия)

ПВ

место

литер

Стои-
мость

по
каталогу
за
доставку

руб. коп.
руб. коп.

Количество
комплектов

на 2012 год по месяцам:

1	2	3	4	5	6	7	8	9	10	11	12

Куда

(почтовый индекс)

(адрес)

Кому

(фамилия, инициалы)

СОДЕРЖАНИЕ

ЕСТЕСТВЕННЫЕ НАУКИ

Физико-математические науки

Механика

Динамика, прочность машин, приборов и аппаратуры

Аксенов Е.В. (Самарский национальный исследовательский университет имени академика С.П. Королева)

Формирование системы прогнозирования вибрационного состояния газотурбинного двигателя 14

Химия

Неорганическая химия

Байсангурова А.А., Магомадова М.А., Ильхаева З.С. (Чеченский государственный университет)

Результаты фазового равновесия системы $Na_2O(Na_2CO_3) - Na_2CrO_4$ 16

Биологические науки

Физико-химическая биология

Физиология и биохимия растений

Мусин Х.Г. (Институт биохимии и генетики – обособленное структурное подразделение Уфимского федерального исследовательского центра Российской академии наук), Кулуев Б.Р. (Институт биохимии и генетики – обособленное структурное подразделение Уфимского федерального исследовательского центра Российской академии наук; Башкирский государственный университет), Якупова А.Б. (Башкирский государственный университет), Сагитов А.М. (ООО «Бионикум», г. Уфа), Чемерис А.В. (Институт биохимии и генетики – обособленное структурное подразделение Уфимского федерального исследовательского центра Российской академии наук)

Рост культур бородачатых корней в колбах и прототипах биореакторов дождевального типа 19

Биотехнология

Доан Тху Тхуи (Вьетнамский национальный университет сельского хозяйства, Вьетнам), Калашникова Е.А. (Российский государственный аграрный университет – МСХА имени К.А. Тимирязева), Зайцева С.М. (Московская государственная академия ветеринарной медицины и биотехнологии – МВА имени К.И. Скрябина; Российский государственный аграрный университет – МСХА имени К.А. Тимирязева), Киракосян Р.Н. (Российский государственный аграрный университет – МСХА имени К.А. Тимирязева)

*Влияние препарата Мицефит на клональное микроразмножение *Dioscorea caucasica* Lipsky 32*

Саркисова М.В., Чередниченко М.Ю. (Российский государственный аграрный университет – МСХА имени К.А. Тимирязева)

*Оценка анатомо-морфологических особенностей растений *Marsilea hirsuta* R.Br. в культуре in vitro и ex vitro 35*

Общая биология

Ботаника

Ибрагимова А., Ализаде В. (Институт ботаники Национальной академии наук Азербайджана, Азербайджан)

Diversity and Distribution of Woody Plants in Qakh district (Azerbaijan) 39

Сагындыкова М.С., Иманбаева А.А. (Мангышлакский экспериментальный ботанический сад, Актау, Казахстан), Ишмуратова М.Ю. (Карагандинский государственный университет им. Е.А. Букетова, Казахстан)

*Исследование динамики накопления эфирных масел в сырье *Ferula foetida* в зависимости от возраста и места произрастания* 42

Микробиология

Соломашенко Н.И., Кириллова О.Г., Поландова В.Н. (Центр гигиены и эпидемиологии в Ставропольском крае), Гандрабура Н.И., Харина Е.И. (Северо-Кавказский федеральный университет)

Анализ уровня иммунной прослойки населения Ставропольского края к возбудителям иксодового клещевого боррелиоза, моноцитарного эрлихиоза и гранулоцитарного анаплазмоза человека методом ИФА за 2017–2019 годы 46

Экология

Белов А.Н., Репш Н.В., Ключников Д.А., Кольцова Л.А., Хилькова М.К. (Филиал Дальневосточного федерального университета в г. Уссурийске (Школа Педагогики))

Растительность морского побережья Хасанского района в зоне интенсивных рекреационных нагрузок 49

Паразитология

Сидорова К.А. (Государственный аграрный университет Северного Зауралья), Драгич О.А. (Государственный аграрный университет Северного Зауралья; Тюменский индустриальный университет), Пантелеева Е.А. (Государственный аграрный университет Северного Зауралья), Татарникова Н.А. (Пермский государственный аграрно-технологический университет имени академика Д.Н. Прянишникова)

Лечебно-профилактические мероприятия при варроатозе пчел в условиях Северного Зауралья 52

Биологические ресурсы

Иванова Л.А., Чурмасова Л.А. (Московский государственный университет пищевых производств), Фоменко И.А. (Московский государственный университет пищевых производств; ОАО «Биохиммаш», г. Москва); Фоменко В.В. (ОАО «Биохиммаш», г. Москва), Кузьмичева Т.П. (Московский государственный университет пищевых производств), Константинова Н.А. (ОАО «Биохиммаш», г. Москва)

*Противораковые, антиоксидантные и антидиабетические свойства растений рода *Zingiber* sp.* 57

Маловичко Л.В., Голованова Н.А., Сашина Л.М. (Российский государственный аграрный университет – МСХА имени К.А. Тимирязева)

*Пищевой спектр золотистой цурки (*Merops apiaster* L., 1758) в аридной зоне Ставрополя* 61

Физиология

Физиология

Ананьев В.Н. (Государственный научный центр Российской Федерации – Институт медико-биологических проблем Российской академии наук), Прокопьев Н.Я. (Тюменский государственный университет), Бычков В.Г., Ананьева О.В. (Тюменский государственный медицинский университет Министерства здравоохранения Российской Федерации), Гуртовой Е.С. (Тюменский государственный университет), Ананьев Г.В. (Московский педагогический государственный университет)

Влияние инвазии описторхозом на изменение реактивности сердечно-сосудистой системы к норадреналину и ацетилхолину при холодовой адаптации 69

Сельскохозяйственные науки

Агрономия

Общее земледелие, растениеводство

Орлова Е.Е., Зубик И.Н., Иванова И.В., Козлова Е.А., Шарафутдинов Х.В., Аниськина Т.А. (Российский государственный аграрный университет – МСХА имени К.А. Тимирязева)

Влияние схемы посадки на сортовые особенности некоторых гибридов Тагетеса прямостоячего (Tagetes erecta L.)..... 75

Луговое хозяйство и лекарственные, эфирно-масличные культуры

Ланин Д.О., Маринич М.Н., Виноходов А.С. (Белгородский государственный национальный исследовательский университет), Коцарева Н.В. (Белгородский государственный национальный исследовательский университет; Белгородский государственный аграрный университет имени В.Я. Горина)

Проективное покрытие искусственных растительных сообществ с различным составом в агроэкосистемах с развитием линейной эрозии..... 81

Ветеринария и Зоотехния

Кормопроизводство, кормление сельскохозяйственных животных и технология кормов

Киселев В.Л. (Научно-исследовательский институт пушного звероводства и кролиководства имени В.А. Афанасьева), Киселев А.Л. (ООО «ВерумБио», г. Москва), Яхин А.Я. (Научно-исследовательский институт пушного звероводства и кролиководства имени В.А. Афанасьева), Тинаева Е.А. (Московская государственная академия ветеринарной медицины и биотехнологии – МВА имени К.И. Скрябина)

Применение АНТИСЕПТИКА-СТИМУЛЯТОРА Д2-фракция для увеличения рентабельности производства при выращивании бройлеров..... 84

Науки о Земле

Технология бурения и освоения скважин

Иванова Т.Н. (Чайковский филиал Пермского национального исследовательского политехнического университета; Институт механики Удмуртского федерального исследовательского центра Уральского отделения Российской академии наук)

Возможности роторных управляемых систем..... 87

Разработка и эксплуатация нефтяных и газовых месторождений

Хавкин А.Я. (Российский государственный университет нефти и газа (национальный исследовательский университет) имени И.М. Губкина)

Особенности повышения дебитов скважин в баженовской свите..... 91

Увеличение нефтеизвлечения путем гармонизации интересов государства и недропользователей.... 94

Метеорология, климатология, агрометеорология

Любомудров А.А. (Национальный исследовательский ядерный университет «МИФИ»)

О возможном наличии дополнительного источника тепла в недрах планеты Земля..... 104

Геодезия

Лабузнов А.В. (Национальный исследовательский Московский государственный строительный университет)

Использование сплайн-функций при наблюдении за деформациями зданий..... 110

ТЕХНИЧЕСКИЕ НАУКИ

Транспортное, горное и строительное машиностроение

Колесные и гусеничные машины

Поливаев О.И. (Воронежский государственный аграрный университет имени императора Петра I), Винокуров В.Д., Можейко А.В., Золотых Е.Д. (Военный учебно-научный центр Военно-воздушных сил «Военно-воздушная академия имени профессора Н.Е. Жуковского и Ю.А. Гагарина» (г. Воронеж) Министерства обороны Российской Федерации)

Износостойкость фрикционных накладок сцепления тракторов с демпферными пружинами..... 113

Дорожные, строительные и подъемно-транспортные машины

Пашков А.В. (Российский университет транспорта)

Колебания стрелы башенного крана при наличии груза 117

Информатика,

вычислительная техника и управление

Автоматизация и управление технологическими процессами и производствами

Кан С.В., Титова Ж.И., Сергазыкызы А. (Карагандинский государственный индустриальный университет, Казахстан)

Структура базовой системы регистрации несчастных случаев, происшедших на предприятии ... 121

Лукашенко Д.В. (Научно-исследовательский институт Федеральной службы исполнения наказаний)

Особенности автоматизированной системы управления библиотечной деятельностью в информационно-образовательной системе..... 124

Управление в социальных и экономических системах

Лукашенко Д.В. (Научно-исследовательский институт Федеральной службы исполнения наказаний)

Проблемы и подходы к информатизации образовательной деятельности как социального процесса 126

Математическое и программное обеспечение вычислительных машин, комплексов и компьютерных сетей

Андрианов И.А., Ржеуцкая С.Ю., Григорьева А.Н., Сорокин А.Н., Полянский А.М. (Вологодский государственный университет)

Подсистема выявления плагиата в программном коде для дистанционного практикума по программированию 128

Теоретические основы информатики

Устин А.М. (Национальный исследовательский Мордовский государственный университет им. Н.П. Огарёва)

Алгоритм оценки пространственной схожести изображений..... 134

Технология продовольственных продуктов

Процессы и аппараты пищевых производств

Шогенов М.Ю., Солталиев М.С., Адамаев М.С., Алоев Р.Б., Диданов М.Ц. (Кабардино-Балкарский государственный университет им. Х.М. Бербекова), Щеренко А.П. (Московский институт энергобезопасности и энергосбережения)

Влияние параметров конструкции и режимов работы шелушильной машины на энергоемкость процесса и качество шелушения риса 138

Процессы и машины агроинженерных систем

Электротехнологии и электрооборудование в сельском хозяйстве

Васильев Д.А., Пантелеева Л.А. (Ижевская государственная сельскохозяйственная академия)

Энергоэффективное управление асинхронным электродвигателем 143

Строительство и архитектура

**Теплоснабжение, вентиляция, кондиционирование воздуха,
газоснабжение и освещение**

Кашуркин А.Ю. (Национальный исследовательский Московский государственный строительный университет; Научно-исследовательский институт строительной физики Российской академии архитектуры и строительных наук)

Регулирующие устройства систем вентиляции и кондиционирования воздуха 148

Безопасность деятельности человека

Охрана труда

Самчук-Хабарова Н.Я., Гапонов В.Л., Хорошун Э.Г. (Донской государственный технический университет)

Повышение объективности экспертной оценки управления рисками 154

**Математическое и программное обеспечение
вычислительных машин, комплексов
и компьютерных сетей**

Андрианов И.А., кандидат технических наук, доцент

Ржеуцкая С.Ю., кандидат технических наук, доцент

Григорьева А.Н., специалист по учебно-методической работе

Сорокин А.Н., кандидат технических наук, доцент

Полянский А.М., кандидат технических наук, доцент

(Вологодский государственный университет)

**ПОДСИСТЕМА ВЫЯВЛЕНИЯ ПЛАГИАТА В ПРОГРАММНОМ КОДЕ
ДЛЯ ДИСТАНЦИОННОГО ПРАКТИКУМА ПО ПРОГРАММИРОВАНИЮ ***

В статье рассмотрена реализация подсистемы выявления плагиата в исходном программном коде для системы дистанционного практикума по программированию. Целью работы является эффективное обнаружение плагиата в студенческих программах, разрабатываемых на различных языках программирования. Предложен оригинальный подход, позволяющий упростить данную проблему за счёт специфических свойств современных оптимизирующих компиляторов. В итоге задача сводится к более простой задаче поиска совпадающих фрагментов в строках. Для её решения применена структура данных «суффиксное дерево». Сделаны модификации внутренней структуры дерева и алгоритма его построения, что позволило снизить потребление памяти.

Ключевые слова: поиск по сходству, электронное обучение, суффиксные деревья.

* Работа выполнена при поддержке РФФИ грант №19-01-00103 А

Andrianov I.A.

Rzheutskaya S. Yu.

Grigor'eva A.N.

Sorokin A.N.

Polianskii A.M.

**ANALYSIS AND DEVELOPMENT OF A SUBSYSTEM FOR SIMILARITY
SEARCH IN PROGRAM CODE FOR THE REMOTE PROGRAMMING
WORKSHOP SYSTEM**

The article describes the implementation of the plagiarism detection subsystem in the source code for the remote programming workshop system. The aim of this work is effective detection of plagiarism in student programs written in various programming languages. We have proposed an original approach that allows simplify this problem using the specific properties of modern optimizing compilers. As a result, the task is reduced to a simpler task of finding matching fragments in text data. To solve it, the suffix tree data structure is used. Modifications were made to the internal structure of the tree and its construction algorithm, which allowed to reduce memory consumption.

Keywords: similarity search, e-learning, suffix trees.

I. Введение

Дистанционный практикум по программированию в Вологодском государственном университете используется уже несколько лет при преподавании ряда учебных курсов. Данный ресурс и руководство по его использованию доступны по адресу [1]. В настоящий момент в системе содержится около двух тысяч задач по нескольким учебным курсам и обширный архив решений. Задачи имеют различную сложность – от простейших до олимпиадных. Продолжается развитие и совершенствование набора задач для обеспечения соответствия их требованиям образовательных стандартов нового поколения [2], рекомендациям SWEBOK [3], профессиональных и отраслевых стандартов программной инженерии. Развиваются возможности построения индивидуальной траектории обучения, формирования и контроля достижения обучающимся необходимого уровня профессиональных компетенций, которые формируются в соответствии с рекомендациями [4] на основе системы оценочных индикаторов.

Важной задачей является повышение объективности оценок уровня приобретённых обучаемым компетенций, в частности – контроль самостоятельности решения студентами предлагаемых задач. Эта задача определяет актуальность выполненного авторами исследования.

II. Анализ проблемы, постановка задачи исследования

Работа учащихся с дистанционным практикумом [1] производится следующим образом. Решением каждой задачи является программа на одном из допустимых в системе языков программирования. Программа рассматривается в виде черного ящика. Система выполняет компиляцию её текста, после чего запускает программу с различными заранее подготовленными входными данными и проверяет правильность ответов. В зависимости от пройденных тестов определяется набранная сумма баллов.

Одна из проблем, которая выявилась в процессе эксплуатации системы, состояла в следующем. К сожалению, периодически находятся недобросовестные студенты, пытающиеся выдать чужие программы за свои либо использовать элементы чужого кода. При этом могут меняться названия идентификаторов, добавляться лишние переменные и функции, меняться форматирование исходного текста, и так далее. Простое сравнение файлов не позволяет выявлять такие случаи – для этого необходимо специализированное программное обеспечение. Выделим основные требования к нему.

Во-первых, оно должно справляться с как можно большим количеством различных способов модификации кода, не влияющих на конечный результат его работы. Во-вторых, оно не должно требовать существенных вычислительных ресурсов, поскольку нередки случаи, когда с системой одновременно работает несколько десятков и даже сотен пользователей (например, при проведении турниров по программированию). В-третьих, необходимо обеспечить возможность сравнительно легко добавлять в систему поддержку новых языков программирования.

III. Анализ методов решения поставленной задачи

Изучение работ в данной области, а также сравнение существующих инструментов для поиска плагиата в программном коде позволило нам выделить несколько подходов к решению рассмотренной задачи.

Первый подход носит название «подсчёт атрибутов» [5]. Он основан на сравнении таких характеристик, как число операторов, операндов, служебных символов, частоты ссылок на переменные и функции и т.п. В некоторых работах предлагается учитывать даже такие стилистические особенности кода, как частота использования пробелов, заглавных букв в именах и др.

Следующий подход основан на поиске общих подстрок сравниваемых файлов, он хорошо работает при незначительных модификациях исходного кода. В работе [6] предложено улучшить данный способ с помощью параметризованного сравнения, то есть сравнения строк в предположении, что они могут отличаться лишь изменением имен идентификаторов. Для эффективного выполнения такого сравнения разработана специальная структура данных – параметризованное суффиксное дерево.

Ещё одна большая группа подходов базируется на синтаксическом и, иногда, даже семантическом анализе исходных текстов. Примером является система Plague [7], в которой строятся и затем сравниваются так называемые профили исходных текстов, отражающих последовательности используемых в программе операторов. В работе [8] предлагается использовать для сравнения программ их графы вызовов, то есть ориентированные графы, изображающие вызовы функций в программе. В [9] на основе синтаксического анализа исходного кода строятся кластеры древовидных структур, которые в дальнейшем и используются для определения степени сходства. В [10] предложено использовать метод статической трассировки: на этапе синтаксического анализа определяется порядок выполнения функций, согласно ему из текста извлекаются ключевые слова и используются при сравнении.

Встречаются и другие интересные подходы, не относящиеся напрямую к рассмотренным выше. Например, в [11] предлагается использовать методы машинного обучения. Построенная модель обучается с помощью экспертов, что со временем повышает точность её ответов.

Анализ рассмотренных подходов позволяет утверждать, что они обладают, по крайней мере, одним из двух недостатков: либо обеспечивается выявление лишь сравнительно простых случаев умышленных изменений кода, либо имеется сильная привязка к конкретным языкам программирования, что делает сложным и трудоёмким добавление поддержки новых языков в систему.

IV. Упрощение задачи за счёт возможностей современных компиляторов

Заметим, что большинство современных оптимизирующих компиляторов обладают следующими особенностями.

- 1) Скомпилированный код обычно не зависит от имён локальных идентификаторов и особенностей форматирования текста.
- 2) Альтернативные конструкции языка программирования во многих случаях компилируются в одинаковый или очень похожий код.
- 3) Лишний код (не влияющий на результат) не компилируется, если транслятору удаётся его распознать.
- 4) При изменении типов данных на похожие типы скомпилированный код меняется незначительно.

Для подтверждения данных тезисов был проведён ряд экспериментов, в которых исследовались различные случаи плагиата в исходном коде программ, отправленных студентами в проверяющую систему. Проанализировав выявленные случаи плагиата, мы обнаружили целый ряд примеров, подтверждающих вышеприведённые тезисы.

Например, эксперимент, подтверждающий тезис 2, приведён в табл. 1. Здесь выполнена замена цикла *for* на альтернативный цикл *while*. Для удобства результат компиляции показан не в машинном коде, а на языке ассемблера. Как можно видеть, последовательность машинных команд не изменилась.

Таблица 1

Пример замены оператора на альтернативный вариант

Входной файл 1	Выходной файл 1	Входной файл 2	Выходной файл 2
for(int i=0; i<20; i++) { c+=a; a*=2; }	xor edx,edx @3: add ecx,eax mov ebx,eax add ebx,ebx mov eax,ebx @5: inc edx cmp edx,20 jl short @3	int i=0; while(i<20) { c+=a; a*=2; i++; }	xor edx,edx @2: add ecx,eax mov ebx,eax add ebx,ebx mov eax,ebx inc edx cmp edx,20 jl short @2

На основе полученных экспериментальных результатов был предложен следующий подход, позволяющий упростить задачу определения уровня сходства двух заданных программ. На первом шаге выполняется компиляция программ. На следующем шаге из скомпилированного кода извлекаются последовательности кодов операций, которые рассматриваются как входные строки для дальнейшего сравнения. Тем самым мы сводим исходную задачу к более простой задаче определения степени сходства двух строк.

В качестве замечания заметим, что существует затруднение при использовании данного подхода для интерпретируемых языков. Однако, это затруднение вполне преодолимо. Во-первых, часть интерпретаторов строит некоторое внутреннее представление программы (промежуточный код, байт-код и т.п.), и его можно использовать. Но даже если такой возможности нет, можно разработать модуль для лексического и синтаксического анализа. Хотя это сравнительно трудоёмко, но всё же такую работу нужно будет проводить не для всех поддерживаемых в системе языков, а для единичных случаев.

V. Использование суффиксных деревьев для сравнения строк

Согласно предыдущему пункту мы перешли от сравнения программных кодов к более простой задаче – вычислению уровня сходства двух строк. При этом возникает закономерный вопрос: какую именно метрику использовать при сравнении? Наиболее известной метрикой для сравнения строк является расстояние Левенштейна ([12], глава 11). Однако алгоритм её вычисления имеет квадратичную сложность, и потому для наших целей не подходит.

Мы предлагаем использовать в качестве метрики для сравнения двух заданных строк t и p следующую величину: количество общих подстрок t и p , имеющих длину не менее l (значение константы l подбирается эмпирически). Такая метрика даёт вполне адекватные результаты сравнения, и при этом она может вычисляться за линейное время с помощью структуры данных «суффиксное дерево» (далее – СД). Примечание: альтернативный вариант для поиска по сходству с помощью подстрок переменной длины («мультиграмм») рассматривается в нашей статье [13].

СД представляет собой дерево, в котором в компактном представлении хранятся все подстроки входной строки. Как время построения дерева, так и объём памяти линейно зависят от размера входных данных. Подробное описание данной структуры и алгоритмов работы с ней можно найти в [12].

В основе вычисления предложенной выше метрики лежит построение структуры данных, называемой статистикой совпадений. Пусть t и p – две исходные строки. Тогда статистикой совпадений для t относительно p будет массив $ms[1..|t|]$, где $ms[i]$ – длина наибольшей общей подстроки t , начинающейся в позиции i , которая встречается также где-то в p . Алгоритм вычисления выполняет построение СД для строки p и проход по его вершинам согласно символам строки t . Подробно алгоритм описан в [12]. Несложно видеть, что, построив массив ms , мы можем легко вычислить искомую метрику.

При работе с СД мы столкнулись со следующей проблемой. Специфика входных данных в нашей задаче предполагает относительно большой размер алфавита. Действительно, каждый символ алфавита – это одна операция в откомпилированном коде, а количество различных видов операций может достигать нескольких сотен. Возникает проблема эффективного представления дерева в памяти. Если хранить в каждом узле дерева массив всех возможных переходов, то требования к памяти будут абсолютно неприемлемыми. Для более разумного расхода памяти необходимо хранить выходящие из узлов дуги более эффективно – например, с помощью сбалансированных деревьев. Однако, даже в этом случае требования к памяти остаются довольно высокими – до 100 – 200 байт на один узел дерева (в зависимости от реализации).

Для решения данной проблемы мы предлагаем метод, полученный на основе комбинации идей из [14] и [15]. В статье [14] показана возможность преобразования уже построенного дерева к бинарному виду путём замены символов на дугах дерева их бинарными кодами. В новом дереве будет столько же вершин, что и в исходном. Однако, из каждой вершины будет вы-

ходить не более двух дуг (начинающихся с символов '0' и '1'), что устраняет необходимость в сложных способах их представления. Пример такого СД показан на рисунке 1. Каждая дуга дерева нагружена подстрокой исходной строки, записанной в двоичном коде. Для экономии памяти реально на дугах хранятся лишь пары чисел – позиции этих подстрок: например, $\langle 2, 6 \rangle$ означает, что подстрока "11000" начинается в позиции 2 и заканчивается в позиции 6.

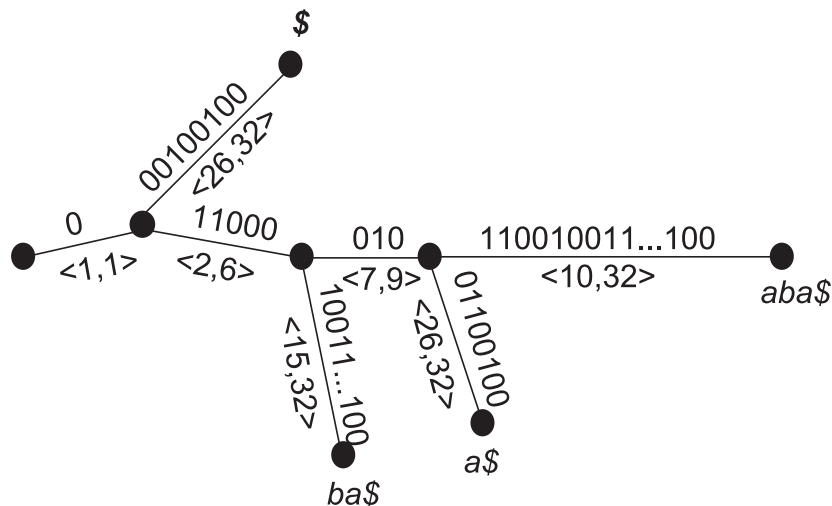


Рис. 1. Бинарное суффиксное дерево над строкой "aba\$"

Наша идея состоит в том, что такое дерево можно не конвертировать из обычного СД, а изначально строить в таком виде. Для этого заметим, что позиции начала соседних суффиксов, входящих в дерево, в двоичном представлении исходной строки находятся на одинаковом расстоянии друг от друга (оно равно количеству битов, которое занимает код одного символа). Это означает, что мы имеем так называемое равномерно разреженное СД. Известен модифицированный алгоритм Укконена для построения такой структуры за линейное время [15]. Применяв его для входной строки, в которой каждый символ заменён на его двоичное представление, мы получим искомое бинарное суффиксное дерево.

Нами проведено экспериментальное сравнение реализации предложенного алгоритма построения дерева с существующей реализацией алгоритма Укконена, взятой с известного справочного ресурса по алгоритмам Е-Махх (e-maxx.ru/algo/ukkonen). Эксперимент показал, что наша версия требует в среднем в 2,5 раза меньше памяти, при этом проигрывая в скорости работы всего лишь около 30%.

Заключение

В данной статье предложен простой и легко реализуемый подход к обнаружению плагиата в программном коде. Использование постобработанного скомпилированного кода в качестве входных данных для сравнения позволяет автоматически игнорировать многие умышленно внесённые изменения в код, не влияющие на результат его работы. В итоге задача упрощается и сводится к поиску общих фрагментов в текстовых данных.

Для решения данной задачи были использованы суффиксные деревья. На основе идей из [14] и [15] предложены изменения внутренней структуры и алгоритма построения суффиксного дерева для более эффективной работы с большими алфавитами. В итоге удалось добиться более чем двукратного снижения расхода памяти при незначительном ухудшении производительности.

На основе полученных результатов разработан модуль контроля плагиата для системы дистанционного практикума по программированию Вологодского государственного университета, который позволяет обнаруживать подозрительные студенческие решения и сообщать об этом преподавателям.

ЛИТЕРАТУРА

1. Дистанционный практикум по программированию ВоГУ [Электронный ресурс]. URL: <http://atpp.vstu.edu.ru/acm>
2. Федеральный государственный образовательный стандарт высшего образования – бакалавриат по направлению подготовки 09.03.04 Программная инженерия [Электронный ресурс]. URL: http://fgosvo.ru/uploadfiles/FGOS%20VO%203++/Bak/090304_B_3_17102017.pdf
3. ISO/IEC TR 19759-2015 Software Engineering – Guide to the software engineering body of knowledge (SWEBOK) – Second Edition [Электронный ресурс]. URL: <https://www.iso.org/standard/67604.html>
4. Полянский А.М., Смирнова Е.А. Проектирование рабочей программы дисциплины на основе элементов компетенций. / Открытое образование. 2018, № 22 (3). – С. 35–51.
5. Grier S. A Tool that Detects Plagiarism in PASCAL Programs. / SIGSCE Bulletin. 1981, Vol. 13, No. 1.
6. Baker B. A theory of parameterized pattern matching: algorithms and applications. / Proc. of the 25th ACM Symp. on the Theory of Computing. 1993. – Pp. 71–80.
7. Clough P. Plagiarism in natural and programming languages: an overview of current tools and technologies. [Электронный ресурс] / Department of Computer Science, Univ. of Sheffield. 2000. – URL: <http://www.dcs.shef.ac.uk/~cloughie/papers/Plagiarism.pdf>
8. Cholakov T., Birov D. Duplicate code detection algorithm. / CompSysTech '15: Proc. of the 16th Int. Conf. on Computer Systems and Technologies, June 2015. – Pp. 104–111.
9. Hyo-Sub Lee, Kyung-Goo Doh. Tree-pattern-based duplicate code detection. / DSMM '09: Proc. of the ACM first international workshop on Data-intensive software management and mining, November 2009. – Pp. 7–12.
10. Jeong-Hoon Ji, Gyun Woo, Hwan-Gue Cho. A source code linearization technique for detecting plagiarized programs. / ITiCSE '07: Proc. of the 12th annual SIGCSE conference on Innovation and technology in computer science education", June 2007. – Pp. 73–77.
11. Huang Qiubo, Tang Jingdong, Fang Guozheng. Research on Code Plagiarism Detection Model Based on Random Forest and Gradient Boosting Decision Tree. / ICDMML 2019: Proceedings of the 2019 International Conference on Data Mining and Machine Learning, April 2019. – Pp. 97–102.
12. Гасфилд Д. Строки, деревья и последовательности в алгоритмах: Информатика и вычислительная биология. / Пер. с англ. И.В. Романовского – СПб.: Невский Диалект; БХВ-Петербург, 2003. – 654 с.
13. Андрианов И.А., Григорьева А.Н. Разработка обновляемого индекса для поиска по регулярным выражениям и сходству. / Информационные системы и технологии. 2018, № 6 (110). – С. 74–82.
14. Anderson A., Nilsson S. Efficient implementations of suffix trees. / Software – Practice and Experience. 1995, Vol. 25. – P. 129–141.
15. Kärkkäinen Ju., Ukkonen E. Sparse suffix trees. / Proc. of the 2nd Annual International Conf. on Comput. and Combinatorics. Lecture Notes in Computer Science. 199, Vol. 1090. – Pp. 219–230.