

МАТЕМАТИКА И МАТЕМАТИЧЕСКО ОБРАЗОВАНИЕ, 2017
MATHEMATICS AND EDUCATION IN MATHEMATICS, 2017
*Proceedings of the Forty-sixth Spring Conference
of the Union of Bulgarian Mathematicians
Borovets, April 9–13, 2017*

**МЕТОДИ ЗА СПРАВЯНЕ С МЕХАНИЧНОТО
ПРЕПИСВАНЕ НА ДОКЛАДИ ОТ СТУДЕНТИ
ПО ЕДИН ИТ-УЧЕБЕН ПРЕДМЕТ**

Филип Петров

Представени са опити за справяне с проблема с механичното преписване при защита на доклади в сравнително нов и много динамичен откъм учебно съдържание учебен предмет в Техническия университет – София. Изследването е проведено в рамките на 7 последователни учебни години.

Въведение. Предметът „Програмни технологии за сигурен код“ (ПТСК) се въвежда за първи път в учебната програма на специалност КСТ [3] (сега КСИ [4]) през учебната 2009/2010 г. в Техническия университет – София. Курсът се провежда в четвърти курс, втори семестър за бакалавърска степен в рамките на 10 седмици. Учебните групи традиционно са три, като се разделят на две подгрупи. Провеждат се общо 5 лабораторни упражнения, като едната подгрупа идва в нечетните седмици, а другата – в четните. Числеността на подгрупите обикновено е в порядъка от 12 до максимум 15 студенти. Допълнително има между 5 и 12 магистри от изравнително обучение, за които понякога се сформира отделна група (при 8 или повече), а друг път се разпределят по групите с бакалаврите. Така общият брой на студентите, които участват в курса по ПТСК, е между 80 и 100 през всяка учебна година. В съдържателната част курсът има значителен обем учебен материал и се работи с конспект от над 60 въпроса. За провеждането на лабораторните упражнения е прието, че няма технологично време да бъдат разгледани всички въпроси и поради тази причина там е поставен фокус изключително върху теми, свързани със сигурното програмиране в интернет. Изпитът се провежда в рамките на два астрономически часа и включва развиване на два теоретични въпроса от конспекта и решаване на една задача, която е свързана с материала, изучаван по време на упражнения. За успешно преминаване на курса студентът трябва да има минимум оценка „среден 3“ на всеки от двата въпроса и на задачата. Крайната оценка се формира като средноаритметично на получените три.

Още от първата учебна година е въведена практика за даване на възможност на студентите да се освободят от изпит чрез изготвяне на доклад на тема по техен личен избор. Темите могат да са свързани с въпроси от конспекта, които не са били разглеждани по време на упражнения, с доразвиване и допълване на материал от лабораторните упражнения или теми, които въобще не са предвидени в курса, но са от предметната област на „сигурно програмиране“. Важно условие за

всички доклади е да има демонстрация с практически пример. Успешното написване и предаване на доклад, който отговаря на поставените изисквания, гарантира освобождаване само от задачата на изпита. Успешната „публична защита“ на този доклад може да освободи студента и от развиването на теоретичните въпроси. Освобождаване е допустимо само с получени оценки „много добър 5“ или „отличен 6“. Публичната защита се състои в изнасяне на кратък практически-ориентиран урок пред състудентите от съответната подгрупа и отговаряне на свързаните с него въпроси, които могат да са зададени от всички участници в залата. За всеки доклад асистентът предварително подготвя критични бележки и въпроси, които не са известни на студентите преди публичната защита.

Основните цели на преподавателите, свързани с въвеждането на тези допълнителни възможности за успешно преминаване на курса, са:

- студентите да добият значително по-задълбочени познания в някои от областите на сигурното програмиране;
- да се стимулира творческото мислене у студентите, като се подтикнат да търсят нови, актуални и интересни практически примери за своите доклади;
- да се подобрят презентационните умения на студентите, което е една добра тренировка за предстоящите защиты на техните дипломни работи;
- студентите да бъдат подготвени за някои елементи от нормативните изисквания към текстовата част на предстоящите им дипломни работи.

Тематичното разпределение за упражненията е следното:

- Упражнение 1: Управление на потребителски сесии. Открадване и фиксиране на потребителски сесии;
- Упражнение 2: Хеширане и съхранение на пароли в база от данни, употреба на salt за защита срещу rainbow таблици и на key stretching срещу brute force атаки;
- Упражнение 3: SQL инжекции, XSS, XSSI и XSRF атаки;
- Упражнение 4: Защита от автоматизирани атаки с ботове, филтриране на спам при електронна поща и уебсайтове;
- Упражнение 5: Публична защита на студентски доклади.

До началото на упражнение 3 студентите имат възможност да заявяват темите, по които ще пишат доклади (асистентът ги одобрява или отхвърля, ако са неподходящи/дублиращи). Една седмица след края на упражнение 4 е крайният срок за предаване на завършените доклади. Преди началото на упражнение 5 асистентът трябва да прочете докладите и да подготви въпроси по тях.

Проблемът с механичното преписване. Трябва да се отбележи, че докладите за освобождаване от изпит по ПТСК се очаква да имат образователна, а не научно-изследователска стойност. Очакванията са студентите да се специализират в конкретна област, да добият разширени знания в нея и да ги предадат на своите колеги по време на защитата. В този смисъл докладите се различават от класическите определения за „курсов проект“ [2], въпреки че по обем на свършената работа и трудност на изпълнение би трябвало да са сходни. Изискването за авторски труд е свързано с изготвянето на собствен практически пример, който да демонстрира разгледаната в теоретичен план софтуерна уязвимост. Когато условията са поставени по такъв начин, е некоректно да се каже, че „преписването“ е сериозен

проблем. Съвсем нормално и очаквано е студентите да използват готови материали по избраните от тях теми и да включват големи количества от чужди текстове в своите доклади. Единственото важно условие е тези текстове да бъдат адаптирани по адекватен начин и да бъдат цитирани коректно. Голям проблем обаче би било „механичното преписване“ – свалянето и предаването на готов цялостен доклад, – защото при него ще липсва задълбоченото разбиране на предадените текстове и код, а оттам идва и голямата вероятност да се стигне до некачествена публична защита. Последното не е желано, защото, както беше споменато, тези защиты освен за оценяване се използват и като компонент от обучението на останалите студенти.

В първата година изложеният модел се приложи безпроблемно, без да се налагат специални методи за справяне с механичното преписване, защото курсът е нов не само за Техническия университет – София, а и за България като цяло. Самият процес на адекватно превеждане на текст от чужд език до известна степен гарантира и разбиране. Бяха срещнати редица технически и организационни трудности, породени по-скоро от неопитност на преподавателя, в резултат на които качеството на публичните защиты беше ниско. Още същата учебна година обаче готови доклади се разпростригнаха по сайтове за споделяне на информация и групи в социални мрежи. Прие се, че моделът няма да бъде устойчив във времето и ще е нужно да се въведат мерки срещу преписване.

Методи за справяне с механичното преписване по ПТСК. Има множество възможни решения за справяне с механичното преписване, които включват софтуерно подпомагане [8]. Още от третата учебна година авторът започна да използва софтуер за откриване на преписване и плагиатство – Turnitin в Moodle [1], – с който да бъдат спрени поне най-грубите опити за плагиатство. Мярката даде мигновен ефект, който се изразява в много по-малко доклади през следващата учебна година (таблицы 1 и 2). Отчетлив е спадът на броя заявени теми и незадоволителните публични защиты и през учебната 2012/2013 г. Ефектът обаче беше краткосрочен

Таблица 1. Резултати преди въвеждане на софтуер за следене за механично преписване

Учебна година	2009/2010	2010/2011
Заявили тема	75	74
Предали доклад	60	59
Хванати в механично преписване	0	1
Освободени от задача	60	58
Освободени от изпит	13	31
Незадоволителна публична защита	47	27

Таблица 2. Резултати след въвеждане на софтуер за следене за механично преписване

Учебна година	2011/2012	2012/2013
Заявили тема	84	54
Предали доклад	69	50
Хванати в механично преписване	14	6
Освободени от задача	55	44
Освободени от изпит	21	33
Незадоволителна публична защита	34	11

– през 2013/2014 г. се забеляза отново повишаване на броя заявени теми за доклади. Недостатъчен беше и спадът на незадоволителните публични защити през учебната 2011/2012, когато софтуерът вече беше въведен, а това е ясна индикация, че проблемът не е разрешен в своята същност:

- Незадоволителните публичните защити продължават да са сериозен проблем, а именно те са очевидната индикация, че студент е предал текст, който не разбира;
- Има ясно видима „наследственост“ – студентите от предишни години предават опита си на тези след тях. Започват да се търсят нови начини за заобикаляне на предприетите мерки;
- Преподавателите не желаят да превърнат практиката с докладите в състезание по надлъгване между софтуерни продукти и студенти, защото това няма да изпълни основната цел на защитите им – да служат като взаимообучителен метод.

Курсът по ПТСК има една характерна особеност, която го отличава от много други учебни предмети – учебното съдържание се обновява и актуализира изключително динамично. Практически примери, с които могат да се демонстрират различните типове уязвимости в реално работещи уебсайтове, са актуални буквално в рамките на няколко дни. Затова беше взето решение тази свръхдинамика на учебното съдържание да бъде употребена, като се включи изискване за пример за актуална уязвимост като компонент на написаните доклади. Хипотезата е, че това би довело до по-устойчиво справяне с механичното преписване и по този начин ще се спести допълнително натоварване на преподавателя, защото предадените доклади (не непременно заявените) би следвало да намалее количествено. Намаленото натоварване за преподавателя ще бъде употребено, като бъде въведена т.нар. „**предварителна защита**“ за избрани „съмнителни“ доклади (такива с голям процент съвпадение на текст и код от доклади, които са предавани през предишни години), в която преподавателят ще задава няколко контролни въпроса по докладите на студентите и по този начин ще се уверява, че до известна степен на сигурност ще бъдат адекватни за публична защита. Следвайки тази идея, моделът за предлагане, утвърждаване, предаване и защита на доклади за освобождаване от изпит по ПТСК беше модифициран по следния начин:

- Етап 1: До края на учебната седмица на упражнение 3, студент, който желае да пише доклад, предлага избрана от него тема при следните ограничения:
 - Темата трябва да е свързана с тематиката на курса и да не дублира материал от лабораторните упражнения;
 - Не е позволено двама студенти от една и съща учебна подгрупа да пишат по еднакви теми;
- Етап 2: В рамките на няколко дни (обикновено 1–2), асистентът одобрява темата, като внася към нея допълнителен специфичен акцент, който да я направи уникална спрямо предавани през предишни учебни години доклади;
- Етап 3: До една седмица след упражнение 4, докладите трябва да бъдат подготвени и предадени;
- Етап 4: В срок от 7 дни асистентът одобрява или отхвърля предадените доклади, като с помощта на Turnitin за избрани от тях провежда предварителна защита;

- **Етап 5:** Провежда се публична защита чрез изнесен урок в упражнение 5.

Вижда се ясно, че ролята на преподавателя в етап 2 с утвърждаването на темите на докладите е ключовият момент за успешното прилагане на този модел, защото именно то е своеобразен гарант, че темата на студента ще притежава елемент на уникалност и няма да бъде тривиална за механично преписване. Все пак е нормално да се очаква припокриване на голяма част от текстовете на доклади от минали години, особено що се отнася до литературния преглед. Затова при внасянето на уникалния елемент за демонстрация на актуална уязвимост той трябва да бъде поднесен така, че да осигури нужда от цялостно ревизиране и обновяване на старите (преписаните) текстове. В помощ на предложения модел се прие да се използва базата от данни CVE [5] на организацията с нестопанска цел Mitre Corporation [6]. Базата CVE е публичен регистър със свободен достъп, в който се индексират и описват известни в публичното пространство уязвимости на софтуерни продукти. Всяка уязвимост се описва подробно и се индексира със собствен уникален идентификатор. Налична е търсачка чрез [5] и [7]. Когато преподавателят знае, че дадена тема е била използвана в предишни години, той може най-лесно да добави в изискванията към тази тема „да се представи практически пример с CVE-201...“ според намерена от него свързана с темата актуална или сравнително актуална уязвимост. Ефектът би бил още по-добър, ако такава уязвимост се зададе директно като тема на доклада.

Обновеният модел беше частично приложен „в движение“ през учебната 2013/2014 г., след като се отчете отново бум на числеността на заявените доклади. Ефектът от предварителните защиты беше значителен, защото 14 студенти, за които софтуерът за разпознаване на плагиатство не беше дал категорична индикация, не бяха допуснати до публична защита (таблица 3). Това може да се дължи или на изобретателност в механичното преписване, или (по-вероятно) някой друг да е написал докладите на тези студенти, а те да не са положили достатъчно усилия да се запознаят в детайли с тях.

Истинският ефект от въведения модел се забеляза през 2014/2015 г. и 2015/2016 г. Броят не само на предадените, но и на заявените теми за доклади намаля буквално наполовина спрямо предишните учебни години. Опитите за механично преписване почти изчезнаха, а броят на незадоволителните публични защиты намаля драстично. Субективното усещане на преподавателите е, че студентите значително повишават качеството на докладите си. Не може да не се спомене и един негативен ефект, а именно, че броят на отличните доклади (на освободени от изпит) също намаля средно с около 25% спрямо минали учебни години (таблица 3).

Таблица 3. Резултати след въвеждане на предварителна защита

Учебна година	2013/2014	2014/2015	2015/2016
Заявили тема	73	33	39
Предали доклад	65	33	35
Хванати в механично преписване	8	0	2
<i>Отпаднали при предварителна защита</i>	14	8	12
Освободени от задача	43	25	23
Освободени от изпит	32	21	18
Незадоволителна публична защита	11	4	5

Заклучение. Софтуерните продукти са добър помощник за преподавателите както за бързо отстраняване на най-грубите опити за плагиатство, така и за формиране на предварително очакване за идентифициране на по-съмнителни доклади. Когато докладите имат образователна, а не научно-изследователска стойност, много по-устойчив ефект имат предварителната защита и внасянето на елементи на уникалност към изискванията на докладите. При по-малко учебни групи или повече на брой асистенти вероятно най-ефективно за този модел би било да се проведе предварителна защита за всеки студент.

ЛИТЕРАТУРА

- [1] Ю. ПЕНЕВА, С. ИВАНОВ. Студентското плагиатство в дигиталната епоха. Пета национална конференция „Образованието в информационното общество“, Пловдив, 31 май – 01 юни 2012 г.
- [2] Ф. ПЕТРОВ, Д. АНДРЕЕВ. Ползи от научно-изследователската дейност на студенти и докторанти по информатика и информационни технологии извън университета. Сборник с доклади от научна конференция на младите изследователи 13 май 2016 г., Велико Търново, Издателство Фабер, 7–13, ISBN 978-619-00-0476-9, 2016
- [3] Учебен план на специалност “Компютърни системи и технологии”, Професионално направление: 5.3. Комуникационна и компютърна техника, Технически университет – София, приет на ФС на ФКСУ на 12.04.2011 г.
- [4] Учебен план на специалност “Компютърно и софтуерно инженерство”, Професионално направление: 5.3. Комуникационна и компютърна техника, Технически университет – София, приет от ФС на ФКСУ на 18.12.2012 г.
- [5] Common Vulnerabilities and Exposures, Master Copy, <https://cve.mitre.org/cve/cve.html>.
- [6] Mitre Corporation, <https://www.mitre.org>.
- [7] National Vulnerability Database, <https://nvd.nist.gov>.
- [8] D. WEBER-WULFF. Plagiarism Detection Software: Promises, Pitfalls, and Practices. Handbook of Academic Integrity, Springer Singapore, 2016, 625–638, ISBN 978-981-287-099-5.

Филип Петров Петров
Факултет по математика и информатика
Софийски университет „Св. Климент Охридски“
бул. „Джеймс Баучър“ 5
1164 София, България e-mail: philip@abv.bg

PLAGIARISM DETECTION METHODS AGAINST STUDENT REPORTS FOR AN IT SUBJECT IN UNIVERSITY

Philip Petrov

The paper presents methods for plagiarism detection for student reports in a relatively new IT subject in Technical University – Sofia. The research is made in 7 consecutive years with different measures against plagiarism – both with and without software.