

DETECTION OF CHEATING IN ONLINE CONTESTS

Václav ŠIMANDL

Abstract: This paper deals with the detection of contestants' unfair behaviour in online contests. We discuss the possibilities of detecting cheating using automated data processing. We defined some indicators referring to abnormalities in the contestants' work. We created software which reveals the presence of these indicators in the work of any particular contestant in the Bebras contest. We inputted the last year's data into it. The software generated a list of contestants who may have cheated during the contest. We compared this list with the list of contestants who were disqualified by the school co-ordinators. We then contacted selected school co-ordinators in order to check whether they share our suspicion that some pupils had cheated during the contest.

Key words: Pupils' cheating, detection of cheating, data mining, Bebras contest

ODHALOVÁNÍ PODVÁDĚNÍ V ONLINE SOUTĚŽÍCH

Resumé: Příspěvek se zabývá problematikou odhalování nežádoucího jednání soutěžících v rámci online soutěží. Diskutujeme zde možnosti, jak odhalovat podvádění soutěžících pomocí automatizovaného zpracování dat, která vznikají během soutěžení. Definovali jsme několik jevů, které mohou poukazovat na abnormality ve snažení soutěžících. Vytvořili jsme software pro zjišťování přítomnosti těchto jevů u jednotlivých soutěžících v soutěži Bobřík informatiky a naplnili jej daty loňského ročníku soutěže. Pro ověření našeho přístupu jsme seznam soutěžících, u kterých jsme měli podezření na podvádění, porovnali se seznamem soutěžících, kteří byli ze soutěže vyřazeni školními koordinátory soutěže. Kontaktovali jsme vybrané školní koordinátory soutěže s cílem zjistit, zda u konkrétních žáků školy sdílí naše podezření na podvádění.

Klíčová slova: Podvádění žáků, zjištění podvádění, dolování dat, Bobřík informatiky

1 Podvádění jako problém českého školství

Přetrvávajícím problémem komplikujícím hodnocení znalostí a dovedností žáků ve školním prostředí je podvádění žáků, přičemž s tímto jevem se setkávají i organizátoři online soutěží. Podvádění ve škole je takové jednání žáka, které porušuje stanovená školní pravidla, přináší mu nezasloužené výhody oproti spolužákům a snižuje spolehlivost a přesnost pedagogických závěrů, které lze z jeho výkonů vyvozovat [1]. Jak uvádí Mareš [2], aktuální společenské klima podvádění ve škole přehlíží nebo toleruje. Podle něho tak vzniká obecné mínění, že podvádění je vlastně běžné, zatímco nepodvádět je cosi výjimečného. Mareš se také domnívá, že podvod je pro řadu lidí podvodem, pouze když je odhalen a jeho tvůrce identifikován a je-li podvodný čin jednoznačně a prokazatelně doložen a tvůrce podvodu je nezpochybnitelně prokázán úmysl podvádět.

2 Formy podvádění a způsoby jeho odhalování *Formy podvádění*

Způsobů podvádění ve škole existuje poměrně velké množství, přičemž většinu z nich lze označit jako žákovské podvádění. Mareš tento druh podvádění dále rozděluje na podvádění tradiční a elektronické [2]. Mezi tradiční formy podvádění patří napovídání, opisování, zjišťování obsahu písemných zkoušek dopředu tak, aby se na ně jedinci mohli lépe připravit, prozrazování obsahu písemných zkoušek dalším třídám a studijním skupinám a „zastupování“ prospěchově slabších studentů při ústním či písemném zkoušení studenty schopnějšími [1]. Dále je možné do této formy zařadit například plagiátorství, opakované využití jedné seminární práce v různých předmětech a úpravu naměřených dat podle požadovaného výsledku [2].

Mezi formy elektronického podvádění je řazeno nedovolené používání mobilních telefonů [1], použití elektronického taháku [2], vytvoření kopie obrazovky nebo pořízení fotografie písemného testu pomocí mobilního telefonu a její předání dalším žákům [2, 3], spolupráce několika studentů prostřednictvím emailu nebo diskuzních

skupin [3], vypracování elektronického testu jinou osobou nebo s nedovolenou pomocí jiné osoby [3], vypracování písemné práce dopředu, její nakopírování do školního počítače a následně vydávání za práci vytvořenou ve škole [2] a plagiátorství [1]. Rowe upozorňuje na specifický typ podvádění při elektronickém testování, kterým je vícenásobné vyplnění a odeslání testu v případě, že s předchozím výkonem není žák spokojen [3]. Poměrně nebezpečné jsou metody založené na prolomení ochrany počítačového systému s cílem získat přístup k datům pro žáka nepřístupným [3].

Problematická může být při podvádění role učitele, kdy učitel může podvádění nejen tolerovat, ale i sám iniciovat. Učitelova aktivní účast při podvádění může mít podobu učitelova napovídání, které se objevuje zejména při snaze porovnat mezi sebou jednotlivé třídy nebo školy [4]. Učitel tehdy nerespektuje požadavek jednotného zadávání otázek, učebních úloh a určitým způsobem napovídá žákům správný postup, přičemž se řídí obavou, že za případné nedostatky ponese zodpovědnost jen on, bez ohledu na složení třídy, zvláštnosti žáků, kvalitu osnov a učebnic [4].

V rámci online soutěží, ve kterých dochází ke srovnání výkonů žáků z různých škol, lze předpokládat oba výše uvedené jevy – podvádění žáků i nežádoucí pomoc učitele. Jak uvádí Vaníček, větším problémem než podvádění jednotlivých žáků jsou situace, kdy svým žákům v soutěži nedovoleně napomáhá sama škola, nebo jejich podvádění alespoň kryje s cílem dosáhnout celostátního úspěchu [5]. Tuto druhou situaci dokumentuje ve slovenském prostředí Gujberová příkladem, kdy oslovení žáci zdůvodnili zvolení určité odpovědi v online soutěži tím, že jim to poradila učitelka [6]. Ačkoliv by bylo možné se domnívat, že podvádění žáků v soutěžích je ve velké míře rozšířeno po celém světě, diskuze na loňském mezinárodním workshopu soutěže Bebras ukazuje, že s podváděním žáků mají ze všech 30 pořadajících zemí problémy především Česká republika a Slovensko [7]. Tento stav zřejmě vystihuje i názor Hausnera, že pojmy jako čestnost, přímost a smysl pro fair play jaksi vymizely z našich škol [8].

Odhalování podvádění

Odhalování podvádění je důležitou činností, která má kromě mravního rozměru i významnou diagnostickou funkci. Navrací totiž přesnost do hodnocení žákovských výkonů, jehož objektivita byla podváděním narušena [4].

Během času bylo vypracováno mnoho postupů, jak odhalovat případné podvádění žáků ve škole. Podle Mareše je nejdostupnější metodou pro odhalování tradičního podvádění pozorování [2]. Již před započítím písemného zkoušení je možné pozorovat, zda nechybí kopie testů, někteří studenti nesedí v učebně jinde než obvykle a zda na lavicích nebo na tabuli nejsou nově nějaké nápisy [2]. Během zkoušení si může učitel všimnout, že žák má na lavici nebo v jejím okolí věci, které tam nepatří, má na sobě věci, které obvykle nenosí, vyměňuje si s jiným žákem určité materiály, nahlíží do spolužákova textu, zjevně nepracuje naplno a na cosi čeká nebo že dokončil práci v neobvykle krátké době [2]. Při opravě písemné práce může zjistit, že žák napsal práci výrazně lépe, než byly jeho dosavadní výkony, že písemné práce dvou nebo více žáků jsou nápadně shodné a vyskytují se zde i stejné chyby nebo že znění práce se nápadně shoduje s textem v učebnici [2].

Pro odhalování elektronického podvádění je podle našeho názoru možné použít obdobné metody. Během zkoušení na počítačích je možné si všimnout, zda žák nemá spuštěné programy, které k práci standardně nepotřebuje. Při práci v online prostředí může učitel kontrolovat, zda žáci nemají v internetovém prohlížeči otevřeno více záložek a v některých z nich nejsou zobrazeny nepovolené stránky (včetně rozhraní emailové schránky, chatovacích aplikací nebo Facebooku). Toto pozorování žáků při práci může být ulehčeno pomocí vhodného monitorovacího software, který zobrazuje náhledy obrazovek jednotlivých počítačů v učebně.

Odhalování elektronického podvádění v písemných pracích, zejména plagiátorství, je podle Mareše možné dvěma způsoby. První způsob si všímá zvláštností jednotlivých složek odevzdané práce – autora, obsahu textu, stylu textu, slovní zásoby, bibliografie a formálních znaků textu. Toto zkoumání provádí sám učitel, přičemž může například zjistit, že text neodpovídá věku, znalostem, zkušenostem a obvyklému způsobu vyjadřování daného studenta, že je styl práce nekonzistentní nebo že formátování textu není jednotné [2]. Druhý způsob si všímá odevzdané práce jako celku, přičemž kontrola je realizována pomocí vhodných softwarových nástrojů. Je zkoumáno, zda jde o originální práci studenta nebo jde o kompilaci textů z necitovaných zdrojů [2].

Jak uvádí Cizek, pro odhalování podvádění v didaktických testech je možné použít také statistické zjišťování, které je založeno na hledání

neobvykle podobných odpovědí [9]. Souhrn statistických metod, které se věnují odhalování podvádění žáků, ve své práci poskytuje Bliss [10]. Popisované metody se zaměřují na odhalování kopírování odpovědí mezi žáky, znalosti znění otázek před započítáním testu a snahy zapamatovat si přesné znění otázek. Uvedené metody jsou založeny na vzájemném porovnávání odpovědí žáků, kteří skládali test v určitý čas na určitém místě, hledání odchylky výkonu žáka od jeho očekávaného výkonu, identifikaci žáků s velmi krátkým časem vyplňování testu nebo odhalení žáků, kteří při testu nereflektovali očekávanou časovou náročnost jednotlivých otázek. Jak však uvádí Rowe, počet správných odpovědí žáka v testu a (ne)stálost jeho známek mezi několika testy nejsou spolehlivým ukazatelem podvádění [3].

Při zkoumání podvádění žáků je třeba pamatovat také na názor Hansona, že žádná statistická metoda zkoumání podvádění nemůže poskytnout nezvratné důkazy, že k podvádění skutečně došlo [11].

3 Postupy odhalování podvádění v online soutěžích

Pro organizátory online soutěží je většina metod, založených na pozorování žáků před soutěží a během soutěže, jen těžko použitelná, využít je lze pouze na úrovni pověřených zástupců z jednotlivých škol. Záleží pak jen na důslednosti každého takového zástupce, jak bude reagovat v případě zjištěného porušení pravidel soutěže. Je tak potřeba se zaměřit na metody založené na dolování dat, které čerpají potřebné informace ze strukturovaných dat generovaných při soutěži. Díky nim lze zjistit významnou odchylku ve výkonu některých žáků nebo skupin žáků oproti ostatním soutěžícím.

Metody odhalování podvádění, které jsou uvedeny výše, jsme se rozhodli zpřesnit a rozšířit tak, aby byly použitelné v online soutěžích. Pro potřeby stanovení jevů, které lze při online soutěžení zkoumat, jsme definovali následující modelové uspořádání soutěže:

- V soutěži existuje více věkových kategorií.
- Žáci soutěží online pod dozorem učitele.
- Různé skupiny soutěžících soutěží v různý čas.
- Není třeba hromadná registrace žáků před soutěží.
- V každé věkové kategorii jsou vyhlašováni tzv. úspěšní řešitelé, tedy soutěžící, kteří dosáhli určeného minimálního počtu bodů.

Zkoumané rizikové jevy

Na základě tohoto uspořádání soutěže jsme určili sadu rizikových jevů, které mohou ukazovat na podvádění žáků v takové soutěži. Každému jevu jsme přiřadili písmenné označení, na které se budeme v dalším textu odkazovat. V závorce vysvětlujeme, proč je vhodné daný jev zkoumat:

- **S:** Žáci, kteří soutěžili současně, odpovídali zcela shodně, ať jejich odpovědi byly správné či nikoliv (je pravděpodobné, že žáci spolupracovali nebo si napovídali).
- **R:** Žák pracoval velice rychle a je zároveň úspěšným řešitelem soutěže (je pravděpodobné, že žák soutěžil opakovaně nebo znal znění otázek od spolužáků, kteří soutěžili dříve).
- **M:** Žák pracoval samostatně nebo ve velmi malé skupince (je možné, že žák soutěžil mimo výuku či ve skupince, která nebyla řádně dozorována).
- **P:** Žák vykázal velkou časovou prodlevu mezi registrací a spuštěním testu (je možné, že žák nejprve spolupracoval s kamarádem a teprve poté spustil svůj test, aby získal více času).
- **K:** Žáci školy pracovali v průměru neobvykle rychle oproti průměru celé věkové kategorie a přitom byli úspěšní (je možné, že značná část žáků spolupracovala).
- **L:** Žáci školy dosáhli v jedné věkové kategorii významně lepších výsledků než jejich spolužáci v jiných kategoriích (je možné, že značná část žáků v dané věkové kategorii spolupracovala).
- **Z:** V průběhu soutěžení se na škole výrazně zvyšovala úspěšnost žáků v dané věkové kategorii (je možné, že si žáci sdělovali znění otázek mezi třídami).

Tyto rizikové jevy (dále je budeme označovat jako indikátory) jsme rozdělili do dvou základních skupin – na indikátory individuální a hromadné. Jako individuální indikátory označujeme jevy, které ukazují na abnormalitu přímo u konkrétního soutěžícího (tj. první čtyři uvedené indikátory); jako hromadné indikátory označujeme jevy, které ukazují na abnormalitu v celé skupině soutěžících (tj. poslední tři uvedené indikátory).

Ačkoliv je možné většinu výše uvedených indikátorů využít pro odhalení podvádění u všech účastníků soutěže, zajímavou skupinou, na níž je možné ověřit použité metody, je skupina úspěšných řešitelů. Domníváme se, že u nich je případné podvádění závažnější než u ostatních soutěžících, protože nese riziko získání výhod

z významného výsledku, jako je zohlednění při přijímacím řízení či ministerský program Excellence, který přináší školám finance na základě úspěšnosti jejich žáků v předmětových soutěžích.

Míra pravděpodobnosti podvádění soutěžících

Míra pravděpodobnosti podvádění je číselně vyjádřený souběh výskytu různých jevů, detekovaných při dolování dat jednotlivých soutěžících a popsanych výše uvedenými indikátory.

Každý indikátor nese určitý stupeň pravděpodobnosti, že opravdu došlo k podvodu. Jednotlivé indikátory pak mají různou důležitost při stanovení pravděpodobnosti podvádění. Proto jsme indikátorům přiřadili rozdílnou váhu: každému individuálnímu indikátoru jsme přiřadili váhu 1 a každému hromadnému indikátoru váhu 0,3. Čím vyšší je vypočítaný součet vah u soutěžícího, tím je vyšší pravděpodobnost, že se daný soutěžící v rámci soutěže dopustil podvádění.

Úspěšné řešitele pak můžeme rozdělit do tří skupin podle součtu vah indikátorů, které se jej týkají:

- Úspěšní řešitelé, u nichž považujeme za téměř vyloučené, že se dopustili podvádění. Součet vah indikátorů u těchto soutěžících je menší než 1.
- Úspěšní řešitelé, u nichž nepovažujeme za pravděpodobné, že se dopustili podvádění. Součet vah indikátorů u těchto soutěžících je roven 1.
- Úspěšní řešitelé, u nichž považujeme za možné až pravděpodobné, že se dopustili podvádění. Součet vah indikátorů u těchto soutěžících je větší než 1.

4 Případová studie: podvádění v soutěži Bobřík informatiky

Možnosti odhalování podvádění žáků v online soutěžích jsme se rozhodli ověřit na příkladu národního kola soutěže Bobřík informatiky. V této pasáži stručně nastíníme, jak je soutěž organizována a data soutěžících zaznamenána, protože to má význam pro pochopení implementace metod zjišťování podvádění.

Tato soutěž probíhá jako jednorázový online test, který je organizován přímo na školách v počítačových učebnách [12]. Soutěžit je možno v celkem pěti věkových kategoriích, přičemž každé kategorii je vyhrazen jeden soutěžní den [13]. Test, na jehož vyplnění mají soutěžící obvykle 40 minut (maximální čas vyplňování

testu se u jednotlivých kategorií mírně liší), se skládá z 12 nebo 15 úloh, nejčastěji uzavřených výběrových otázek s jednou správnou odpovědí a úloh interaktivních. Otázky v testu jsou soutěžícím vygenerovány v náhodném pořadí, přičemž i pořadí odpovědí je u každé otázky náhodné.

Testové otázky jsou bodovány podle obtížnosti, za správně zodpovězenou těžší otázku získá soutěžící více bodů než za zodpovězení lehké otázky, ovšem za nesprávnou odpověď se body odečítají. Soutěžící má možnost na otázku neodpovědět, pak nezískává ani neztrácí žádné body [12]. Úspěšným řešitelem je soutěžící, který získal více než polovinu maximálního počtu bodů po odečtení vstupní bodové dotace [12].

Pro průběh soutěže je důležitá role tzv. školního koordinátora soutěže, který se stará o registraci školy do soutěže a průběh soutěže na škole [12]. Registrace jednotlivých žáků před soutěží není nutná, žáci se registrují sami těsně před spuštěním testu, kdy musí mimo jiné zadat tajný kód školy [13], který jim sdělil školní koordinátor. Po skončení soutěžního dne by měl školní koordinátor soutěže zkontrolovat seznam soutěžících za svou školu, potvrdit správnost seznamu a případně vyřadit neplatné soutěžící (žáky nedodržující pravidla soutěže, soutěžící v nižší věkové kategorii, vyplňující test opakovaně, případně kolegy z řad učitelského sboru a podobně) [13].

Data vzniklá činností soutěžících jsou ukládána do databáze, která tvoří základ celé soutěžní aplikace. Mezi ukládaná data patří údaje, jež je každý účastník povinen uvést při registraci: jméno, pohlaví, třída a označení školy, které je určeno na základě zadaného kódu školy. Do databáze se také ukládá čas registrace účastníka, čas spuštění testu a čas ukončení testu. V databázi jsou dále uchovávány odpovědi soutěžícího na jednotlivé otázky, přičemž je uloženo vždy číslo soutěžícího, číslo otázky, označení zvolené odpovědi a čas zvolení této odpovědi. Jestliže soutěžící změnil odpověď na některou otázku, nevytvoří se v databázi další záznam, ale vyhledá se příslušný záznam obsahující jeho předchozí odpověď a ten se aktualizuje. Po ukončení testu jsou odpovědi soutěžícího vyhodnoceny a je vypočítán jeho celkový bodový zisk, který je zaznamenán. Do databáze se též promítají informace o tom, zda byl daný soutěžící vyřazen školním koordinátorem soutěže.

5 Implementace metod hledání indikátorů v soutěži Bobřík informatiky

Ačkoliv by bylo možno zkoumat jednotlivé indikátory v rámci soutěže Bobřík informatiky jen za pomoci vhodných SQL dotazů do databáze, v níž jsou uložena data soutěže, vytvořili jsme specializovaný software, který umožňuje efektivnější a pohodlnější obsluhu. Tato aplikace umožňuje na základě spolupráce s databází soutěže tvořit sestavy soutěžících, u nichž se vyskytují uvedené indikátory. Část procesu zpracování dat se odehrává mimo vytvořenou aplikaci v tabulkovém editoru, přičemž pro export i import dat používáme CSV soubory. Pro každý z výše uvedených indikátorů jsme vypracovali jednu metodu hledání úspěšných řešitelů soutěže, u nichž se daný indikátor vyskytuje.

ad S: Základem metody bylo systematické prohledávání množiny všech úspěšných řešitelů soutěže školy v kategorii a zjišťování, které dvojice soutěžících měly zcela stejné odpovědi na všechny otázky. U těchto dvojic jsme dále ověřovali, zda soutěžili současně (časový rozestup začátku vyplňování nejvýše 5 minut nebo rozestup konce soutěžení nejvýše 10 minut). Pokud ano, zaznamenali jsme oba soutěžící pro další zpracování.

ad R: Identifikace soutěžících probíhala dvěma způsoby. Kromě celkového času soutěžení (tj. doby mezi spuštěním testu a jeho ukončením) jsme sledovali ještě „čistý“ čas soutěžení, což je doba mezi spuštěním testu a zodpovězením poslední otázky (ačkoliv časová prodleva mezi zodpovězením poslední otázky a ukončením testu byla obvykle velmi krátká, u některých soutěžících činila až tři čtvrtiny celkového času soutěžení). V případě celkového i „čistého“ času soutěžení jsme hledali úspěšné řešitele, jejichž čas soutěžení byl nejmeně o pět minut kratší, než byl časový průměr úspěšných řešitelů dané kategorie.

ad M: Základem metody byla systematická kontrola všech soutěžících školy v kategorii. Zajímali nás soutěžící, u kterých jsme zjistili, že v časovém intervalu 10 minut před začátkem soutěžení až 10 minut po něm začali soutěžit nejvýše 3 další soutěžící z dané školy. Jestliže nalezení soutěžící byli zároveň úspěšnými řešiteli soutěže, zaznamenali jsme je pro další zpracování.

ad P: Zaměřili jsme se na úspěšné řešitele soutěže, u nichž časová prodleva mezi registrací a spuštěním testu byla delší než 10 minut. Tyto soutěžící jsme zaznamenali pro další zpracování.

ad K: Identifikace soutěžících probíhala ve dvou krocích. Nejprve jsme vypočítali průměrný čas soutěžení úspěšných řešitelů dané školy v kategorii a porovnali jej s průměrným časem všech úspěšných řešitelů v této kategorii. Jestliže byl průměrný čas úspěšných řešitelů školy nejmeně o pět minut kratší než průměrný čas všech úspěšných řešitelů, poznamenali jsme si číslo školy a název kategorie. Ve druhém kroku jsme vytvořili seznam všech úspěšných řešitelů z této školy a kategorie.

ad L: Identifikace soutěžících probíhala opět ve dvou krocích. Nejprve jsme pro každou kategorii, v níž žáci školy soutěžili, vypočítali průměrný bodový zisk všech soutěžících a posléze jsme vyhledali nejvyšší z těchto průměrů. Jestliže byl tento bodový průměr nejmeně o 25 bodů vyšší než bodové průměry v ostatních kategoriích, poznamenali jsme si číslo školy a název kategorie. Ve druhém kroku jsme vytvořili seznam všech úspěšných řešitelů z dané školy a kategorie.

ad Z: Zaměřili jsme se na školy, z nichž v kategorii soutěžilo alespoň 12 soutěžících. Soutěžící této skupiny jsme vzestupně seřadili podle času začátku soutěžení a skupinu posléze rozdělili na dvě poloviny. První polovina skupiny byla kontrolní a druhou polovinu jsme dále rozdělili na třetiny (tj. šestiny původní skupiny), čímž vznikly tři prověřované skupiny. Pro kontrolní skupinu i pro každou z prověřovaných skupin jsme vypočítali bodový průměr zde zařazených soutěžících. Pokud byl bodový průměr některé z prověřovaných skupin alespoň o 30 % vyšší než bodový průměr kontrolní skupiny, zaznamenali jsme úspěšné řešitele z prověřované skupiny pro další zpracování.

6 Pilotní výzkum a zjištěné výsledky

Výše uvedené postupy odhalování podvádění jsme aplikovali na soutěžící v 6. ročníku soutěže Bobřík informatiky, který se konal v listopadu 2013. Tohoto ročníku se zúčastnilo celkem 34454 soutěžících z 373 škol, přičemž 8501 soutěžících se stalo úspěšnými řešiteli soutěže. 341 úspěšných řešitelů však bylo vyřazeno koordinátory soutěže, řádných úspěšných řešitelů tedy bylo 8160.

Podle námi vypracované metodiky vylučujeme podvádění u 6349 (tj. 78 %) řádných úspěšných řešitelů, za nepravděpodobné jej považujeme u 1262 (tj. 15 %) řádných úspěšných řešitelů a za možné až pravděpodobné jej považujeme u 549 (tj. 7 %) řádných úspěšných řešitelů. Podrobné výsledky uvádíme v tabulce 1.

Taktéž jsme se zaměřili na úspěšné řešitele, kteří byli vyřazeni školními koordinátory. Pro ověření naší metodiky jsme zjišťovali, zda jsme daného vyřazeného úspěšného řešitele označili jako podvádějícího. Podvádění jsme určili jako možné až pravděpodobné u 149 (tj. 43 %) vyřazených úspěšných řešitelů, jako nepravděpodobné u 68 (tj. 21 %) vyřazených úspěšných řešitelů a jako téměř vyloučené u 124 (tj. 36 %) vyřazených úspěšných řešitelů. Podrobné výsledky jsou uvedené v tabulce 1.

Součet vah	Počet úspěšných řešitelů		
	Řádných	Vyřazených	Celkem
0	5412	92	5504
0,3	903	19	922
0,6	34	13	47
0,9	0	0	0
1	1262	68	1330
1,3	349	37	386
1,6	60	12	72
1,9	0	0	0
2	85	45	130
2,3	45	45	90
2,6	8	8	16
2,9	0	0	0
3	1	0	1
3,3	1	1	2
3,6	0	1	1

Tabulka 1: Počty úspěšných řešitelů podle součtu vah

7 Diskuze výsledků pilotního výzkumu

Ověření na školách

Abychom zjistili, zda námi vytvořený postup neoznačuje jako podvádějící žáky, kteří

se podvádění nedopustili, oslovili jsme 13 školních koordinátorů soutěže s dotazem, zda si během soutěžení vyjmenovaných žáků nevšimli něčeho neobvyklého. Do emailu jsme vždy vložili seznam několika vytipovaných řádných úspěšných řešitelů z dané školy, u nichž jsme označili podvádění za možné až pravděpodobné, a u každého žáka jsme zdůvodnili, proč si u něj nejsme jisti dodržováním pravidel. Dotazovali jsme se na soutěžící z kategorií Kadet, Junior nebo Senior (8. ročník základní školy a výše). Školy, které jsme oslovili, jsme vybírali podle počtu úspěšných řešitelů, které náš systém označil jako možná až pravděpodobně podvádějící, a také jsme do tohoto výběru

zařadili školy, za něž soutěžil žák s vysokým součtem vah indikátorů (2,3 a více).

Školní koordinátoři vesměs podvádění svých žáků odmítali, dva z nich nicméně přislíbili prošetření celé situace. Ačkoliv jsme dobré výsledky žáků nezpochybňovali, koordinátoři nás poměrně často upozorňovali na úspěchy žáka v jiných soutěžích a žákův zájem o informatiku. Krátký čas soutěžení některých žáků vysvětlovali jejich nadáním pro informatiku a případně dobrou přípravou na soutěž. Samostatné soutěžení žáka bylo podle jejich slov obvykle způsobeno tím, že z celé třídy soutěžilo pouze několik nadaných žáků, zatímco ostatní žáci dostali jiné úkoly. Několik školních koordinátorů pak přiznalo, že sdělili kód školy žákovi, který byl v době konání soutěže doma z důvodu nemoci. Žádný školní koordinátor nám však nedokázal uspokojivě objasnit situaci, kdy někteří žáci měli zcela shodné odpovědi na všechny otázky.

Možné příčiny vyřazování řešitelů školními koordinátory

U vyřazených úspěšných řešitelů, u nichž jsme podvádění označili za téměř vyloučené, jsme se pokusili odhadnout, z jakého důvodu došlo k jejich vyřazení školními koordinátory. Podrobným zkoumáním seznamu soutěžících jsme zjistili, že ze 124 takových řešitelů:

- 17 soutěžících byli učitelé,
- 14 soutěžících zadalo nesmyslné jméno nebo přezdívku (tito soutěžící mohli být také učitelé, žáci spadající do jiné kategorie a podobně),
- 5 osob soutěžilo ve více kategoriích,
- 55 soutěžících bylo školními koordinátory vyřazeno hromadně, kdy byla vyřazena celá směna, třída či kategorie (z databáze soutěže však není patrné, zda byly tyto skupiny soutěžících koordinátorem cíleně vyřazeny, nebo zda byly z nějakého důvodu pouze nepotvrzeny – např. koordinátor na potvrzení soutěžících zapomněl),
- u zbylých 33 osob jsme nebyli schopni pravděpodobný důvod vyřazení nalézt.

Případy, kdy soutěží učitel nebo žák uveďte nesmyslné jméno, nelze podle našeho názoru automatizovaně odhalit. Stejně tak je problematické odhalování vícenásobného startu jednoho soutěžícího, pokud například pozmění své jméno a třídu. Věříme však, že odhalování právě těchto situací by měl zvládat školní koordinátor soutěže.

8 Diskuze vhodnosti použitých indikátorů a metod

Ačkoliv metody, které jsme použili během zkoumání, byly z hlediska implementace dobře realizovatelné, považujeme za důležité diskutovat jejich přínos k odhalování podvádění.

Metoda hledání dvojic soutěžících, kteří měli zcela stejné odpovědi na všechny otázky (indikátor S), se ukázala být přínosná a během podrobnějšího zkoumání jsme navíc zjistili, že vytipované dvojice mají v mnohých případech i podobné pořadí odpovědí na soutěžní otázky (ačkoliv jsou otázky soutěžícím generovány v náhodném pořadí). To nás utvrzuje v názoru, že tato metoda je velmi vhodná pro hledání podvádění žáků. Zároveň se domníváme, že by bylo možno ji ještě rozšířit o hledání dvojic, které zodpovídaly testové otázky ve stejném nebo velice podobném pořadí.

Metoda hledání soutěžících, kteří pracovali samostatně (indikátor M), se ukázala být nevhodná pro kontrolu těch škol, za které soutěží pouze několik vybraných žáků v celé kategorii. Tito soutěžící byli metodou označeni za potencionálně podvádějící, ačkoliv ve skutečnosti pracovali pod řádným dozorem učitele. Metoda nicméně odhalila praxi některých škol, které žákům, ze zdravotních důvodů se neúčastnícím výuky, zasílají soutěžní kód školy, aby mohli soutěžit z domova. Tento postup pravidla soutěže Bobřík informatiky sice nevylučují, ale žák by měl být posléze vyřazen školním koordinátorem soutěže, protože nesoutěžil pod dozorem vyučujícího.

Metoda hledání soutěžících, kteří vykazali velkou časovou prodlevu mezi registrací a samotným startem soutěže (indikátor P), se ukázala být nepříliš vhodná. Je totiž založena na odhalování chyby podvádějícího soutěžícího, který se registruje, ale nespustí již samotný test, čímž na sebe upoutá pozornost. S tím souvisí fakt, že s její pomocí se nám podařilo oproti ostatním metodám identifikovat pouze velice málo soutěžících.

Metoda založená na odhalování tzv. rychlých soutěžících (indikátor R) je podle našeho názoru přínosná, avšak její podobu pro použití v budoucnu bude nutné modifikovat. Především bude potřeba upravit výpočet mezního času, který tvoří předěl mezi normálním a abnormálně krátkým soutěžním časem, přičemž se nabízí například výpočet pomocí směrodatné odchylky času soutěžení úspěšných řešitelů v dané soutěžní kategorii. Za vhodnou považujeme také kontrolu čistého času soutěžení, která pomáhá odhalit

žáky, kteří pracují velice rychle, avšak tuto rychlost soutěžení maskují tím, že odloží ukončení testu, ačkoliv již na jeho řešení nepracují.

Metody založené na hledání abnormalit v celé skupině soutěžících s sebou přinášejí riziko, že vytipovaná skupina je odlišná od skupiny referenční přirozeně a zjištěné odchylky se zde vyskytují, aniž došlo k podvádění. Tyto metody jsou nevhodné zejména pro kontrolu škol, ve kterých není jednotný postup při výběru soutěžících. Pokud například v jedné kategorii soutěží celé třídy a ve druhé kategorii pouze vybraní žáci, je pravděpodobné, že průměrný bodový zisk soutěžících ve druhé kategorii bude výrazně vyšší než v kategorii první. Analogicky, jestliže bude v určité kategorii na škole během dne soutěžit několik tříd a po výuce budou v této kategorii soutěžit vybraní žáci ze zbylých tříd, je opět pravděpodobné, že skupina těchto vybraných žáků bude mít výrazně vyšší průměrný bodový zisk než třídy soutěžící na počátku dne. I přes zmíněná negativa se domníváme, že tyto metody lze použít jako podpůrné postupy umožňující komplexnější pohled na určitou skupinu soutěžících než metody zkoumající každého soutěžícího jednotlivě.

9 Závěr

Podvádění žáků v online soutěžích je aktuálním problémem, kterému musí organizátoři těchto soutěží čelit. Navrhli jsme sadu indikátorů, jež mohou ukazovat na podvádění soutěžících, a metody umožňující sledování těchto indikátorů u soutěžících. Vytvořili jsme softwarové řešení, které umožňuje sledovat vybrané indikátory v rámci online soutěže, a toto řešení jsme otestovali na soutěžících loňského ročníku soutěže Bobřík informatiky. Výsledky jsme ověřovali u vybraných školních koordinátorů, kde bylo podvádění indikováno významně, s dotazem, zda sdílí naše podezření na podvádění u vytipovaných žáků. Odpovědi škol vyzněly vesměs v tom smyslu, že se daní žáci podvádění nedopustili. Některé z oslovených škol navíc reagovaly podrážděně.

I přes nepříliš pozitivní ohlas ze strany školních koordinátorů se domníváme, že námi popisovaný přístup je možnou cestou, jak získat vhled do problematiky podvádění v online soutěžích a jak nalézt strategie pro eliminaci tohoto jevu.

Význam zjišťování podvádění při online testech má jistě uplatnění i v maturitních testech nebo při zvažování zapojení online soutěží do ministerského programu Excellence.

10 Literatura

- [1] PRŮCHA, J., E. WALTEROVÁ a J. MAREŠ. *Pedagogický slovník*. 6. rozšířené a aktualizované vyd. Praha: Portál, 2009. 395 s. ISBN 978-807-3676-476.
- [2] MAREŠ, J.. Tradiční a netradiční podvádění ve škole. *Pedagogika*. 2005, roč. 55, č. 4, s. 310–335. ISSN 0031-3815.
- [3] ROWE, N. C. Cheating in Online Student Assessment: Beyond Plagiarism. In: *Distance Learning Administration*. 2004, roč. 7, č. 2. Dostupné z: <http://www.westga.edu/~distance/ojdla/summer72/rowe72.html>
- [4] MAREŠ, J. a J. KŘIVOHLAVÝ. *Komunikace ve škole*. Brno: Masarykova univerzita, 1995. 210 s. ISBN 80-210-1070-3.
- [5] VANÍČEK, J. Potenciální a skutečný dopad informatické soutěže do změn kurikula ICT v České republice. In: *DidInfo 2012* [CD-ROM]. Banská Bystrica: Univerzita Mateja Bela, Fakulta přírodních věd v Banskej Bystrici, 2012. s. 15–24. ISBN 978-80-557-0342-8.
- [6] GUJBEROVÁ, M. Názory učitelův na kategorii Bobříci soutěže iBobor. In: *DidInfo 2013*. Banská Bystrica: Univerzita Mateja Bela, Fakulta přírodních věd v Banskej Bystrici, 2013, s. 82–87. ISBN 978-80-557-0527-9.
- [7] VANÍČEK, J. Ústní sdělení dne 4. 11. 2013.
- [8] HAUSNER, M. Počítače. Strašák? Modla? Pomůcka? In: *Rodina a škola*. 2002, č. 6, s. 12–13.

- [9] CIZEK, G. J. *Detecting and Preventing Classroom Cheating*. Thousand Oaks: Corwin Press, 2003. 167 s. ISBN 0-7619-4655-1.
- [10] BLISS, T. J. *Statistical Methods to Detect Cheating on Tests: A Review of the Literature* [online]. 2012 [cit. 2013-12-06]. Dostupné z: http://www.ncbex.org/assets/media_files/Research/2012Bliss.pdf
- [11] HANSON, B. A., D. J. HARRIS a R. L. BRENNAN. *A comparison of several statistical methods for examining allegations of copying* [online]. 1987 [cit. 2013-11-25]. Dostupné z: http://www.act.org/research/researchers/reports/pdf/ACT_RR87-15.pdf
- [12] VANÍČEK, J. Rozvoj informatických kompetencí žáků pomocí soutěže: kauza Bobřík informatiky [online]. In: *Sborník konference Počítač ve škole*. Nové Město na Moravě: Gymnázium Vincence Makovského, 7.–9. 4. 2009 [cit. 2013-11-13]. Dostupné z: http://www.pocitacveskole.cz/system/files/uzivatel/9/clanky/vanicek_pdf_43005.pdf
- [13] TOMCSÁNYI, P. a J. VANÍČEK. Implementation of informatics contest Bebras in Czechia and Slovakia. In: *Information and communication technology in education. Proceedings*. Ostrava: Ostravská univerzita, 2009. s. 214–218. ISBN 80-7368-459-4.

Mgr. Václav Šimandl
Katedra informatiky
Pedagogická fakulta, Jihočeská univerzita
Jeronýmova 10
371 15 České Budějovice, ČR
Tel: +420 387 773 078
E-mail: simandl@pf.jcu.cz
Www pracoviště: wvc.pf.jcu.cz/ki/