

Juhaňák, Libor

Analýza 3: Chování studentů v průběhu plnění online testů

In: Juhaňák, Libor. *Analytika učení a data mining ve vzdělávání v kontextu systémů pro řízení výuky*. Vydání první Brno: Masarykova univerzita, 2023, pp. 167-190

ISBN 978-80-280-0184-1; ISBN 978-80-280-0185-8 (online ; pdf)

Stable URL (handle): <https://hdl.handle.net/11222.digilib/digilib.77697>

Access Date: 29. 11. 2024

Version: 20230228

Terms of use: Digital Library of the Faculty of Arts, Masaryk University provides access to digitized documents strictly for personal use, unless otherwise specified.

9 ANALÝZA 3: CHOVÁNÍ STUDENTŮ V PRŮBĚHU PLNĚNÍ ONLINE TESTŮ⁹⁰

Ačkoli je využívání systémů pro řízení výuky (LMS) zvláště na univerzitách a vysokých školách již poměrně běžné (srov. Ferguson, 2012a; Macfadyen & Dawson, 2012; Poullová, 2010), učitelé využívající LMS ve své výuce mají relativně často jen omezené možnosti sledovat, co přesně se v jejich online kurzech děje. Ať jde o to, jak se v kurzech studenti chovají, jakým způsobem přistupují ke studiu online materiálů či jak postupují při plnění výukových aktivit. Důvodem je to, že běžné systémy typu LMS prozatím stále neobsahují nativní analytické či data miningové nástroje umožňující učitelům snadno a rychle získat vhled do dění v kurzu. Existují sice různé specializované externí nástroje, které jsou schopny dát vyučujícím potenciálně přínosné výstupy, obvykle jsou však pro učitele příliš komplexní, jelikož jsou určeny primárně pro výzkumníky a nabízí příliš širokou nabídku funkcí významně přesahující potřeby učitelů (srov. Romero, Ventura, & García, 2008; Romero & Ventura, 2010; Romero, Cerezo, Bogarín, & Sánchez-Santillán, 2016). Otázky související s rozvojem nových analytických a data miningových nástrojů a přístupů, které by učitelům umožnily jednoduše, a přitom podrobně sledovat chování a interakci studentů při vykonávání online výukových aktivit, jsou stále na pořadu dne a tvoří významnou část výzkumu v oblasti analytiky učení a data miningu ve vzdělávání (viz kapitolu 6.1).

Nemožnost dostatečně podrobně sledovat, jak se studenti při učení v LMS chovají, může přitom mít poměrně zásadní důsledky především v kontextu hodnocení online aktivit studentů. A ačkoli je (především) v oblasti data miningu ve vzdělávání věnována problematice chování studentů v online prostředí poměrně

90 Tato kapitola vychází z autorem publikované studie s názvem *Using process mining to analyze students' quiz-taking behavior patterns in a learning management system* (Juhaňák, Zounek, & Rohlíková, 2019).

velká pozornost (srov. Papamitsiou & Economides, 2014 či Peña-Ayala, 2014b, ale také kapitole 6.3.2 Modelování chování studentů), většina výzkumů se zaměřuje na analýzu dat pocházejících z inteligentních tuteurských systémů (ITS), spíše než ze systémů typu LMS (viz Baker & Gowda, 2010; Baker, Goldstein, & Heffernan, 2011; Muldner, Burleson, Van de Sande, & VanLehn, 2011). Pokud už se výzkumy zaměřují přímo na systémy typu LMS, věnují pozornost zejména predikci úspěchu studentů či identifikaci rizikových studentů (např. Smith, Lange, & Huston, 2012; Jayaprakash, Moody, Lauría, Regan, & Baron, 2014; Baker, Lindrum, Lindrum, & Perkowski, 2015), případně se pohybují spíše na obecnější úrovni a sledují chování studentů v průběhu celého kurzu (např. Agudo-Peregrina, Iglesias-Pradas, Conde-González, & Hernández-García, 2014; Hershkovitz & Nachmias, 2011).

Výrazně méně často se však výzkumy zaměřují na detailní analýzu interakcí studentů v kontextu konkrétních výukových aktivit v LMS (Phillips et al., 2011), či dokonce specificky na analýzu chování studentů v rámci aktivit sloužících k hodnocení (Papamitsiou & Economides, 2016). Přitom se stále více ukazuje, že právě specifický kontext, ve kterém online učení a výuka probíhá, je pro porozumění chování studentů zcela zásadní. A bez dostatečného zohlednění kontextu (tj. nejen konkrétní technologie či nástroje, ale také způsobu jejich zapojení do výuky), zřejmě nelze budovat fungující a přitom dostatečně obecné prediktivní modely použitelné v analytických či data miningových nástrojích pro učitele (Gašević et al., 2016).

V této analýze se proto zaměřují právě na problematiku chování studentů v systémech typu LMS. Konkrétně zaměřují pozornost na jeden určitý typ výukových aktivit sloužících pro účely hodnocení, jelikož se věnují specificky chování studentů v průběhu plnění online testů. Pro analýzu chování studentů je přitom využita metoda dolování procesů (*process mining*). Ta prozatím v kontextu analytiky učení a data miningu ve vzdělávání patří spíše mezi méně vyžívané metody. Dosavadní studie využívající tuto metodu v kontextu systémů pro řízení výuky (Romero, Cerezo, Bogarín, & Sánchez-Santillán, 2016), a dokonce i přímo v souvislosti s analýzou online testů (Papamitsiou & Economides, 2016) však naznačují potenciál využití této metody právě pro účely mapování a modelování procesu vyplňování online testů studenty. Cílem analýzy je proto především zhodnocení, nakolik je metoda dolování procesů využitelná pro účely analýzy chování studentů v průběhu plnění online testů a případně pro účely detekce specifických forem chování studentů při plnění testů v LMS.

9.1 Přehled dosavadních výzkumů

V úvodu této analýzy a stejně tak v kapitole 6.3, resp. 6.3.2, byla problematika analýzy, resp. modelování chování studentů v různých online vzdělávacích systémech (LMS, ITS apod.) spojována především s oblastí data miningu ve vzdělávání

a analytiku učení, kde je této problematice v současnosti věnována značná pozornost. Zaměříme-li se však na chování studentů přímo v průběhu vyplňování online testů, zjistíme, že tomuto tématu je věnována pozornost i v jiných kontextech a mimo výzkumné oblasti data miningu ve vzdělávání a analytiku učení. Chování studentů při plnění testů (tzv. *quiz-taking behavior*) může mít totiž zásadní důsledky pro validitu testů, a tudíž i hodnocení studentů. Proto je možností odhalování různých typů neobvyklého či nestandardního chování studentů v průběhu plnění testů věnována pozornost i v širší oblasti hodnocení studentů (*student assessment*) a v oblasti didaktických testů. V souvislosti s hodnocením či testováním prostřednictvím systémů typu LMS se pak obvykle hovoří o oblastech označovaných jako online hodnocení (*online assessment*), počítačové hodnocení (*computer-based assessment*) či elektronické hodnocení (*electronic assessment*) (srov. Buchanan, 2002; Gaytan & McEwen, 2007; Kim, 2015; Terzis & Economides, 2011; Whitelock, 2006).

Při mapování dosavadního výzkumu chování studentů v průběhu plnění online testů, resp. problematiky detekce specifických či nestandardních forem chování při vyplňování online testů, považují proto za nezbytné poskytnout širší přehled a věnovat pozornost i dalším oblastem mimo data mining ve vzdělávání a analytiku učení. Konkrétně tak v rámci přehledu dosavadních výzkumů nastíním tři základní skupiny přístupů k analýze chování studentů při plnění online testů:

1. přístupy založené na tzv. *person-fit* indexech,
2. přístupy založené na rychlosti odpovídání a
3. přístupy vycházející z data miningu ve vzdělávání a z analytiku učení.

Než se však přesunu k představení jednotlivých přístupů, je nutné provést vymezení několika základních pojmů souvisejících s různými typy testů. Důvodem je jednak to, že tyto pojmy budou dále v textu používány, jednak také to, že v různých typech testů se mohou vyskytovat odlišné formy chování studentů, a tudíž se v souvislosti s různými typy testů objevují různé způsoby detekce a modelování chování studentů při jejich plnění.

Předně je třeba rozlišit formativní a sumativní hodnocení. Cílem formativního či průběžného hodnocení je obvykle získat informaci o aktuálním stavu určitého aspektu učení již v průběhu učebního procesu. Zjišťovaným aspektem učení pak může být např. aktuální výkon studentů, zjišťování porozumění určitému problému, identifikace typických chyb apod. Formativní hodnocení slouží především jako podpora učení a pomůcka pro zlepšování výkonu studentů. Naopak cílem sumativního hodnocení není již ani tak podpora učení, jako spíše zhodnocení výsledku učebního procesu, tj. zjištění, zda student dosáhl požadované úrovně znalostí (Kim, 2015; Průcha, Walterová, & Mareš, 2003, 2009).

V kontextu online hodnocení a systémů typu LMS je pak jedním z nejčastěji využívaných způsobů realizace formativního i sumativního hodnocení v LMS nástroj pro tvorbu testů, kvízů či cvičení. V oficiálním českém překladu systému Moodle je pro tento nástroj použito označení Test, v originálním anglickém znění

se pak využívá pojem *Quiz*. V tomto ohledu je nutno upozornit na relativně široké a nejednoznačné pojetí pojmu „test“ v českém kontextu. Je-li řeč o „testech“ v odborné pedagogické literatuře, mají se tím obvykle na mysli didaktické testy⁹¹. Didaktický test je pak zpravidla chápán jako zkouška orientující se na objektivní zjišťování úrovně zvládnutí učiva (Chráška, 1999; Maňák, Švec, & Švec, 2005), případně jako nástroj pro systematické měření výsledků výuky (Byčkovský, 1982). Tato pojetí směřují primárně k sumativnímu hodnocení a ověřování úrovně znalostí. Při zaměření pozornosti na obecné porozumění pojmu „test“ v pedagogické praxi však zjistíme, že jsou pod tento pojem zahrnovány i další typy aktivit, které mají blíže k formativnímu hodnocení a jejichž primárním účelem je spíše podpora učení než měření znalostí. S oporou o anglickou terminologii tak může být vhodné rozlišovat mezi testy (*tests*) a kvízy (*quizzes*). Zatímco o testech se hovoří primárně v kontextu sumativního hodnocení, označení „kvízy“ se používá spíše v rámci formativního hodnocení (Peat & Franklin, 2002; Kibble, 2007). V českém odborném kontextu však toto rozlišení není běžné, proto dále používám pouze označení „test“ a případně rozlišuji, zda mám na mysli test pro účely formativního či sumativního hodnocení.

Posledním důležitým rozlišením, které je třeba nyní provést, je rozlišení na tzv. *low-stakes* a *high-stakes* testy. Toto rozlišení je založeno na tom, jaké mají výsledky testu důsledky pro testovaného studenta. Výsledky *low-stakes* testů obvykle nemají žádné nebo jen minimální důsledky pro testovaného studenta, zatímco výsledky *high-stakes* testů (v češtině také tzv. rozhodných testů) jsou z hlediska důsledků pro studenta velmi důležité či zcela zásadní. Příkladem *high-stakes* testů může být např. postupová zkouška, bez jejíhož splnění nemůže student pokračovat dále ve studiu. Příkladem *low-stakes* testu může být např. průběžný test, který se nepodílí na výsledné známce za kurz. *Low-stakes* testy tak bývají obvykle spojovány spíše s formativním hodnocením, zatímco *high-stakes* testy s hodnocením sumativním (Davis, 2013; Zhang & Henderson, 2015).

9.1.1 Přístupy založené na person-fit indexech

Jedním z přístupů, který je používán především v oblasti sumativního hodnocení a *high-stakes* testů, je přístup založený na tzv. person-fit indexech. V českém prostředí této problematice zatím není věnována přílišná pozornost. Krátce se jí věnují snad jen Jelínek, Květon a Vobořil (2011), kteří hovoří o PF (person-fit) indexech umožňujících odhalovat nekonzistentní testové chování. Zároveň zmiňují, že „mimo oblast osobnostních dotazníků lze PF indexy využít i v oblasti výkon-

91 Druhou významnou oblastí, kde se v pedagogice, a zvláště pedagogickém výzkumu pracuje s pojmem „test“, je samozřejmě statistika s různými typy statistických testů. Toto využití pojmu „test“ však nyní není předmětem zájmu.

vých testů, například k identifikaci v testu podvádějících osob nebo ke zjištění osob se specifickými schopnostmi a diagnóze kognitivních chyb konkrétních probandů“ (Jelínek, Květon, & Vobořil, 2011, s. 83).

Analýzy založené na person-fit indexech mají již poměrně dlouhou tradici a jsou využívány nejen pro účely hodnocení studentů (tj. v rámci pedagogické evaluace), ale také (a možná i primárně) v oblasti psychologického hodnocení, tj. např. u osobnostních testů, měření postojů apod. (Meijer & Sijtsma, 2001; Rupp, 2013). Person-fit metody jsou rozvíjeny především v rámci tzv. *item response theory* (IRT) neboli teorie odpovědi na položku. Obecně lze přitom říci, že se používají pro zhodnocení toho, nakolik je individuální výkon studenta na jednotlivých položkách testu v souladu či naopak nesouladu s určitým IRT modelem, resp. s výkony ostatních studentů, kteří vyplňovali daný test. Zjednodušeně řečeno jsou person-fit indexy využívány ke zhodnocení toho, nakolik je způsob vyplňování testu studentem očekávatelný či „normální“ a nakolik je neobvyklý, nestandardní či neočekávaný (Meijer, Niessen, & Tendeiro, 2016).

Vezmeme-li si pro ilustraci didaktický test, tak výsledné skóre určitého žáka nemusí být vždy adekvátní vzhledem k jeho skutečným znalostem. To znamená, že v určitých případech nemusí didaktický test správně změřit skutečné znalosti žáka. Nepřesnost měření samozřejmě může být dána mnoha faktory. Nechme nyní stranou faktory týkající se samotného testu i jeho konstrukce a zaměříme se na faktory týkající se osoby vyplňující test. Student může v průběhu testu např. hádat (tzv. *guessing behavior*) a uměle si tak navyšovat výsledné skóre. Nebo naopak může student trpět tzv. *test anxiety* čili trémou či úzkostí spojenou s vyplňováním testu. V takovém případě student kvůli přílišné nervozitě či strachu z neúspěchu odpovídá špatně i na otázky testující ty znalosti, které ve skutečnosti má. Student také nemusí být dostatečně seznámen s formátem testu (tzv. *test unfamiliarity*), což může ve výsledku způsobit nižší skóre, než by odpovídalo skutečným znalostem. Jiným příkladem může být tzv. *“sleeping” behavior*, kdy má student problémy s odpovídáním na první otázky v testu. Samozřejmě může docházet i k různým formám podvádění (*cheating*), kdy je skóre studenta uměle navyšováno např. tím, že došlo k prozrazení správných odpovědí na otázky v testu, anebo student získal správnou odpověď od jiného studenta vyplňujícího stejný test apod. (Emons, Sijtsma & Meijer, 2005; Meijer & Sijtsma, 2001).

Takové specifické případy chování při vyplňování testu pak mohou vyústit v neobvyklý či nestandardní způsob vyplňování, resp. v nestandardní vzorec odpovědí⁹², což způsobuje neadekvátní výsledky studentů v testu. Person-fit analýzu je pak možné využít právě pro odhalování těchto neobvyklých případů, u kterých může být výsledné skóre nevalidní a u kterých tudíž mohlo docházet k některému z výše uvedených typů chování v průběhu vyplňování testu.

92 V anglicky psané literatuře se v této souvislosti používá spojení *aberrant item response pattern*.

Podrobný popis jednotlivých person-fit indexů a způsobu jejich výpočtu jde již nad rámec této analýzy, pro základní přehled lze ale odkázat na metodologickou přehledovou studii Meijera a Sijtsmy (2001), kteří podávají přehled širokého spektra používaných person-fit indexů. Je však nutné zmínit, že při výpočtech person-fit indexů se vychází pouze ze skóre studentů na jednotlivých položkách testu. Pracuje se tedy pouze s informací o správnosti či nesprávnosti studentovy odpovědi na jednotlivé otázky testu a jiné okolnosti nejsou brány v úvahu. Obecný princip výpočtu většiny person-fit indexů lze pak zjednodušeně vysvětlit jako kvantifikaci toho, nakolik student zodpovídá obtížné testové položky správně a jednoduché testové položky nesprávně (viz Meijer & Sijtsma, 2001; Meijer, Niessen & Tendeiro, 2016).

Z hlediska praktického využití pro detekci různých typů chování studentů v průběhu vyplňování testů má však přístup založený na person-fit indexech několik nevýhod. Především, person-fit indexy jsou citlivé na řadu různých typů neobvyklého chování, přičemž výše byly zmíněny příklady jen některých z nich (další uvádí např. Rupp, 2013). Person-fit indexy lze tak sice využít pro detekci neobvyklého chování při plnění testu, ovšem u většiny z nich nelze na základě samotného person-fit indexu určit, o který typ netradičního či neobvyklého chování se jedná. Pro řešení tohoto problému tak např. Emons, Sijtsma & Meijer (2005) navrhnou analytický postup, kde je zapojena grafická analýza (*graphical analysis*) a lokální analýza (*local analysis*), jež umožňují hlubší vhled do možných příčin daného neobvyklého chování. Podrobnější popis tohoto analytického postupu jde však již za rámec této analýzy.

9.1.2 Přístupy založené na rychlosti odpovídání

Zatímco přístupy založené na person-fit indexech mají již relativně dlouhou historii a jsou využívány i mimo oblast počítačového testování, přístupy založené na rychlosti odpovídání začaly získávat na větší popularitě jednak s rozvojem počítačového testování (CBT), jednak s rozvojem inteligentních tuteurských systémů (ITS). V souvislosti s tím můžeme zároveň sledovat dvě do určité míry paralelní oblasti výzkumu. Na jedné straně rozvíjeli tento typ přístupů výzkumníci věnující se primárně problematice testování, kteří pocházejí zejména z oblasti psychologie a pedagogiky (např. Wise & Kong, 2005; Wise, Kong & Pastor, 2009). Na druhé straně jsou pak výzkumníci pracující primárně s daty z inteligentních tuteurských systémů, tedy pocházející spíše z oblasti informatiky a data miningu (např. Beck, 2005; Mostow, Beck, Cuneo, Gouvea, & Heiner, 2005). Druhý ze zmiňovaných výzkumných směrů lze přitom do značné míry považovat za jeden z hlavních inspiračních zdrojů současných výzkumů v oblasti data miningu ve vzdělávání.

Ačkoli se terminologie ve zmiňovaných oblastech poněkud liší, základní myšlenka přístupů založených na rychlosti odpovídání je obdobná a vychází z mož-

ností nástrojů počítačového testování (resp. ITS), které dovolují přesně zaznamenávat čas od zobrazení konkrétní úlohy v testu po zadání odpovědi studentem⁹³. Z hlediska odpovídajícího studenta se pak obvykle očekává, že v zásadě volí jednu z následujících dvou strategií:

1. Buď se aktivně snaží určit správnou odpověď na zadanou úlohu (tzv. *solution behavior*, resp. *solution-seeking behaviour*),
2. anebo pouze (rychle) vybírá některou z odpovědí, aniž by se aktivně snažil zjistit, jaká odpověď je správná (tzv. *guessing behavior*, resp. *rapid-guessing behavior*).

Délka odpovědi (tj. *item response time*) pak může být využita právě pro rozlišení mezi standardním chováním spojeným se snahou správně zodpovědět zadanou úlohu (*solution behavior*) a nestandardním chováním spojeným s tipováním správné odpovědi (*rapid-guessing behavior*). Za tímto účelem tak Wise a Kong (2005) zavedli metriku označovanou jako *response time effort* (RTE), jež byla posléze využívána různými autory i v mnoha dalších studiích (např. Wise, Kong & Pastor, 2009; Slim, Must, & Täht, 2013; Lee & Jia, 2014).

Určitou nevýhodou RTE a obdobných metrik založených na rychlosti odpovídání je ovšem to, že je nutné pro každou z testových úloh stanovit konkrétní práh, který určuje, jaká rychlost bude považována za standardní chování a kdy už naopak půjde o nestandardní chování (Wise & Kong, 2005). Přičemž přístupy k určení odpovídajících prahů mohou být různé, a ne vždy dostatečně dobře automatizovatelné. Navíc, jak je z výše uvedeného patrné, z hlediska detekce různých typů chování studentů v průběhu plnění testů je využití čistě časového přístupu problematické, jelikož lze pomocí těchto přístupů detekovat vesměs jen dva obecné typy chování (tj. *solution behavior* vs. *rapid-guessing behavior*). Zvláště v kontextu výzkumů věnujících se inteligentním tutorským systémům se tak postupně indikátor délky odpovědi (tj. *response time*) začal využívat pouze jako jeden z více různých faktorů vstupujících do komplexnějších modelů (např. Romero & Ventura, 2007).

9.1.3 Přístupy vycházející z data miningu ve vzdělávání a analytiku učení

Jak bylo uvedeno v úvodu představované analýzy, detekce a modelování různých typů chování studentů v online výukových prostředích je zvláště v data miningu ve vzdělávání jednou ze stěžejních oblastí zájmu výzkumníků. Dle přehledové studie Peña-Ayala (2014b) se přímo problematice modelování chování studentů (*student behavior modeling*) věnuje více jak 21 % výzkumných publikací, přičemž

⁹³ V anglicky psané literatuře se v této souvislosti používá spojení *item response time* či jen *response time*.

modelování studentů obecně se v určitém ohledu věnuje dokonce 82 % výzkumů. Předmětem modelování chování studentů jsou přitom v dosavadních výzkumech různé typy chování, jako například:

- hádání či tipování (tzv. *guessing behavior*),
- chování označované jako “*sleeping*” *behavior* (viz výše),
- chování označované jako *gaming the system*, které by šlo přeložit jako obcházení systému,
- chování spojené s hledáním pomoci (*help-seeking behavior*) či naopak se zneužíváním pomoci, resp. nápovědy (*misusing help*) a další (Baker & Inventado, 2014; Peña-Ayala, 2014a).

Z hlediska doposud používaných metod jsou přitom pro identifikaci různých forem chování studentů využívány především metody shlukování a klasifikace (Bousbia & Belamri, 2014).

Co se týká přímo metody dolování procesů, je tato metoda zmíněna již v první stěžejní publikaci týkající se data miningu ve vzdělávání (*Handbook of Educational Data Mining*) jakožto jedna ze základních technik používaných v této oblasti (viz Romero, Ventura, Pechenizkiy, & Baker, 2010). Zároveň je ale zmíněna jakožto metoda, která zatím není příliš využívána, jelikož dosavadní přístupy v kontextu data miningu ve vzdělávání se jen zřídka zaměřují na proces jako celek (viz Trčka, Pechenizkiy, & van der Aalst, 2010). Přesto metoda dolování procesů získává postupně i v oblasti vzdělávání určitou pozornost výzkumníků. Dolování procesů a jeho možné využití ve vzdělávání představují např. Reimann a Yacefová (2013). Ve svém kritickém příspěvku se pak Reimann, Markauskaiteová a Bannertová (2014) věnují metodě dolování procesů z pohledu metodologických výzev spojených s datově intenzivními výzkumnými metodami.

Zaměříme-li pozornost přímo na empirické studie, lze zmínit především využití dolování procesů v kontextu počítačem podporovaného kolaborativního učení (*computer-supported collaborative learning – CSCL*), kde jej Schoorová a Bannertová (2012) využily pro mapování sociálních regulačních procesů. Autorky zároveň dochází k závěru, že metody dolování procesů se ukázaly jako užitečné pro získání vhledu do samotného procesu učení a doporučují využití těchto metod pro další analýzy. Zajímavé využití dolování procesů nabízí i Bannertová, Reimann a Sonnenberg (2014), kteří se ve své studii věnují problematice seberegulovaného učení a využívají metody dolování procesů na původně kvalitativních datech získaných pomocí think-aloud protokolu. Aplikaci dolování procesů na datech pocházející z prostředí LMS Moodle se pak věnují např. Romero et al. (2016). Přímou souvislostí s problematikou chování studentů v průběhu plnění online testů využívají metodu dolování procesů Papamitsiouová a Economides (2016). Konkrétně se autoři zaměřují na využití dolování procesů pro detekci chování spojeného s hádáním či tipováním (*guessing behavior*). A ačkoli pracují s daty pocházejícími ze systému vlastní výroby, což může mít jisté důsledky z hlediska přenositelnosti na jiné systémy,

výsledky jejich studie naznačují, že využití metody dolování procesů může přinést nové možnosti v oblasti detekce a modelování různých typů chování studentů v průběhu vyplňování online testů v rámci systémů typu LMS.

9.2 Řešený problém a otázky

Výzkumný problém této analýzy se zaměřuje na možnosti využití metody dolování procesů pro analýzu chování studentů v rámci systému pro řízení výuky (LMS), specificky pak na analýzu chování studentů v průběhu vyplňování online testů. Primárním cílem analýzy je především obecné zhodnocení využitelnosti metody analýzy procesů pro tento typ výzkumného problému. Sekundárním cílem je pak zmapování a popis případných specifických způsobů chování studentů při plnění online testů, které se podaří za využití dolování procesů odhalit. Hlavní výzkumnou otázkou (HVO) této analýzy lze tak formulovat následovně:

- *HVO: Jak je využitelná metoda dolování procesů pro účely detekování specifických způsobů chování studentů v průběhu plnění online testů v LMS?*

Výše uvedenou hlavní výzkumnou otázku jsem přitom pro účely řešení této analýzy rozdělil na následující tři specifické výzkumné otázky (SVO):

- *SVO1: Zda a případně nakolik je využitelná metoda dolování procesů pro analýzu a detekci různých forem chování studentů v průběhu plnění online testů v LMS Moodle?*
- *SVO2: Jaké způsoby chování studentů v průběhu plnění online testů v LMS Moodle lze detekovat za využití dolování procesů?*
- *SVO3: Jaké je četnostní zastoupení jednotlivých forem či způsobů chování studentů v průběhu plnění online testů v LMS Moodle?*

Jak ukázal přehled dosavadního výzkumu, prozatím je k dispozici jen velmi málo studií využívajících dolování procesů jakožto metodologický přístup. Zvláště pak, pokud se zaměříme přímo na systémy typu LMS a následně pouze na aktivitu studentů v průběhu plnění online testů. Analýze chování studentů při plnění testů za použití dolování procesů se věnují jen Papamitsiouová a Economides (2016). Ti se však zaměřují pouze na využití dolování procesů pro detekci chování spojeného s hádáním či tipováním (*guessing behavior*). Přímou v kontextu LMS Moodle využívají dolování procesů Romero et al. (2016), nevěnují však pozornost přímo procesu vyplňování online testů, nýbrž sledují spíše chování studentů v kurzu jako celku. Otázka po možnostech využití metody dolování procesů pro analýzu a detekci různých forem chování studentů v průběhu plnění online testů v LMS Moodle tak zůstává stále otevřená, a proto se na ni zaměřuje první specifická výzkumná otázka.

Druhá otázka se zaměřuje na potenciálně odhalitelné způsoby chování studentů v průběhu plnění testů. Jak bylo zmíněno výše, Papamitsiouová a Economides (2016) se zaměřují pouze na tzv. *guessing behavior*. V přehledu dosavadních výzkumů však

byla uvedena řada dalších typů chování, kterým výzkumníci věnují pozornost v souvislosti s plněním online testů (viz Baker & Inventado, 2014; Emons, Sijtsma & Meijer, 2005; Meijer & Sijtsma, 2001; Peña-Ayala, 2014a; Wise & Kong, 2005). Doposud však zřejmě neexistují studie, jež by zjišťovaly, které z těchto dalších typů chování studentů v průběhu plnění testů jsou detekovatelné za využití metody proces miningu. Na tento problém se proto zaměřuje druhá specifická výzkumná otázka.

Poslední specifickou výzkumnou otázkou (SVO3) lze chápat především jako rozšíření SVO2, jelikož se zaměřuje na četnost výskytu jednotlivých forem či způsobů chování v průběhu plnění testů. Vzhledem k doposud chybějícímu výzkumu totiž k tomuto zatím nemáme žádné informace.

9.3 Metodologie

Představovaná analýza má vzhledem ke stanoveným cílům především exploratorní povahu, kdy se za využití relativně inovativních metod dolování procesů snaží odhalit a zmapovat způsoby chování studentů v průběhu plnění online testů v systémech typu LMS, a to na datech z vybraných online testů různých typů. V následujících metodologických podkapitolách nejprve představím dolování procesů jakožto základní metodologický přístup této studie, následně popíši analyzovaná data, resp. jednotlivé testy, kterým byla v rámci analýzy věnována pozornost. V posledních dvou podkapitolách pak nastíním proces sběru a předzpracování dat a následně způsob jejich analýzy.

9.3.1 Metodologický přístup – dolování procesů

Metoda dolování procesů (*process mining*) pochází původně z podnikatelské sféry. Jak ale bylo naznačeno v předchozí kapitole, začíná se tento metodologický přístup stále více využívat i v kontextu vzdělávání. Dolování procesů přitom označuje sadu metod, technik a nástrojů zaměřujících se na extrakci informací a znalostí o zkoumaném procesu z tzv. logů, které jsou běžně ukládány v různých typech online informačních systémů včetně vzdělávacích systémů typu LMS (Pechenizkiy, Trčka, Vasilyeva, van der Aalst, & De Bra, 2009; Trčka, Pechenizkiy & van der Aalst, 2010; van der Aalst, 2016). Obecně může být metoda dolování procesů použita pro analýzu jakýchkoli dat, která mají charakter sekvence různých typů aktivit, jež pramení z opakovaného zapojování se různých aktérů do určitého procesu. Dolování procesů je tedy možné použít v takových případech, v nichž je možné určitou sekvenci aktivit či událostí chápat jakožto výsledek nějakého procesu s vnitřní strukturou (Reimann & Yacef, 2013). V případě této analýzy je pak oním procesem vyplňování online testu studentem.

Využití metody dolování procesů obvykle slouží jednomu ze tří základních účelů: objevení (*discovery*), potvrzení (*conformance*) nebo rozšíření (*extension*) procesního modelu na základě analyzovaných procesních dat. V prvním případě neexistuje žádný předem známý či apriori model, a tudíž je cílem jeho konstrukce na základě dostupných dat. V druhém případě naopak existuje apriori model a cílem je ověřit, nakolik reálné procesy zaznamenané pomocí logovacího mechanismu určitého online systému skutečně odpovídají tomuto modelu. V případě rozšíření jde pak do určité míry o kombinaci obou předchozích přístupů. Sice existuje apriori model, ale cílem je ho rozšířit na základě dat o reálném průběhu zkoumaného procesu (srov. Pechenizkiy et al., 2009; Trčka, Pechenizkiy & van der Aalst, 2010). Využití metody dolování procesů v rámci této analýzy spadá do první kategorie, jelikož jde o exploratorně pojatý výzkum a cílem je konstruovat model různých typů či způsobů chování studentů v průběhu vyplňování testů.

Využití metody dolování procesů je obvykle realizováno ve třech základních krocích či fázích (srov. Trčka, Pechenizkiy & van der Aalst, 2010; van der Aalst, 2011, 2016). První fází je předzpracování dat (tzv. *pre-processing*), kdy je třeba upravit data do takové podoby, aby bylo možné aplikovat vybrané algoritmy pro dolování procesů. Nutnost předzpracování dat je dána především tím, že u většiny systémů nemají sbíraná data takovou podobu, která by bez dalších úprav odpovídala požadavkům konkrétního výzkumu.

Druhou fází je identifikace a dolování procesního modelu z analyzovaných dat. V tomto kroku se jednak využívají různé postupy pro exploraci a deskriptivní popis dat, jednak se aplikují vybrané algoritmy pro dolování procesů, jejichž výsledkem je pak určitá forma modelu. Jako příklad explorativních a deskriptivních technik lze zmínit vizualizaci variant procesů (tzv. *trace variants visualization*) či analýzu procesu za využití tečkového grafu (*dotted chart analysis*). Příkladem algoritmů pro dolování procesů pak může být např. *alpha algorithm* či *Heuristic Miner*. Výsledný model pak může mít podobu např. Petriho sítě (*Petri net*), kauzální sítě (*causal net*) či závislostního grafu (*dependency graph*).

Posledním krokem pak bývá testování vytvořeného modelu. Zde se používají různé postupy v závislosti na použitém algoritmu a formě modelu. V případě modelů v podobě Petriho sítě se obvykle provádí kontrola souladu modelu s analyzovanými daty (tzv. *conformance checking*), kdy se vyhodnocuje počet sekvencí, které jsou v souladu s modelem a které naopak modelu neodpovídají (Reimann & Yacef, 2013).

9.3.2 Analyzované případy

V této analýze byla využita data pocházející z celkem pěti online testů realizovaných v pěti různých kurzech v prostředí LMS Moodle provozovaném na FF MU. Všechny testy jsou součástí předmětů vyučovaných na FF MU, přičemž tematické

9 Analýza 3: Chování studentů v průběhu plnění online testů

zaměření těchto předmětů spadá do oblasti filozofie, kultury a historie. Výběr testů pro analýzu byl dán dvěma primárními kritérii:

1. Zahrnutí testů různých typů, a to jak z hlediska jejich nastavení v rámci LMS, tak i z hlediska toho, zda se jedná o testy formativní či sumativní.
2. Zahrnutí testů s co možná největším množstvím pokusů (resp. odpovědí) od studentů, jelikož velké množství pokusů dává větší šanci výskytu různých způsobů chování studentů při jejich plnění.

Základní charakteristiky analyzovaných testů jsou uvedeny v tabulce níže (viz tabulka 17). Testy A a B mají charakter spíše průběžných testů (tj. slouží pro účely formativního hodnocení), zatímco testy C, D a E jsou naopak koncipovány jako závěrečné (zápočtové) testy, a tudíž je možné je vnímat jako součást sumativního hodnocení. Ve všech těchto testech (tj. C, D a E) totiž studenti musí splnit test na určitou minimální úroveň, aby úspěšně ukončili studovaný předmět. Test C a test D pak mají podobu závěrečného testu v kurzech, které probíhají kompletně online, ostatní testy jsou součástí blended learningových kurzů, které vedle online části obsahují také prezenční výuku. Test E jako jediný studenti vyplňují přímo v počítačové učebně na fakultě (tj. *on-site*) a pod dohledem vyučujícího, ostatní z analyzovaných testů jsou vyplňovány online.

Tabulka 17: Přehled základních charakteristik analyzovaných testů

	Test A	Test B	Test C	Test D	Test E
Typ testu	průběžný	průběžný	závěrečný (zápočtový)	závěrečný (zápočtový)	závěrečný (zápočtový)
Počet studentů	14	66	311	176	43
Počet pokusů v testu	356	241	472	186	43
Počet otázek v testu	7	12	40	20	25
Časový limit	ano	ne	ano	ano	ano
Náhodný výběr otázek	mix	ano	ne	ano	ne
Počet možných pokusů	neomezeno	5	3	2	1
Požadováno splnění	ne	ne	ano	ano	ano
Stanovena minimální požadovaná známka	ne	ne*	ano	ano	ano
Zpětná vazba**	vše mimo správnou odpověď; vždy	vše mimo správnou odpověď; jen po dokončení pokusu	vše mimo správnou odpověď; dokud je test dostupný	jen body; vždy	vše; jen po dokončení pokusu
Online/On-site	online	online	online	online	on-site

* Přímo v testu není stanovena požadovaná známka čili student nemusí v tomto testu získat určitou

známku. Ovšem jde o jeden z průběžných testů, ze kterých musí student v průběhu kurzu jako celku získat určitý počet bodů.

** V rámci LMS Moodle se při nastavování zpětné vazby u testů určuje jednak to, jaké informace se mají studentům zobrazit (např. body, správnost položky, správná odpověď apod.), jednak také v jaké fázi se jim mají dané informace zobrazit (např. jen ihned po dokončení pokusu či dokud je test studentům přístupný apod.).

9.3.3 Sběr a předzpracování dat

Pro sběr dat o chování studentů v průběhu vyplňování testů byl využit standardní logovací mechanismus (tzv. *logging system*) implementovaný v LMS Moodle (verze 3.1)⁹⁴. Pro účely analýzy tak musela být data nejprve extrahována z databáze systému, konkrétně pak z tabulky *mdl_logstore_standard_log*. Extrahována přitom byla data o veškeré aktivitě studentů v rámci systému Moodle v průběhu plnění vybraných testů. Data z tabulky *mdl_logstore_standard_log* byla následně předzpracována pro účely analýzy. Po předzpracování dat obsahoval analyzovaný datový soubor následující proměnné:

- identifikátor testu (*qui.id*),
- identifikátor studenta (*student.id*),
- číslo pokusu v testu (*attempt.number*),
- název aktivity (*activity.name*),
- název skupiny aktivit (*activity.group*),
- zkratka skupiny aktivit (*activity.group.abbreviation*) a
- čas vytvoření záznamu (*timestamp*).

V průběhu předzpracování došlo především k vytvoření proměnné obsahující pojmenování zaznamenaných aktivit studentů a následně k vytvoření skupin aktivit stejného typu. Pojmenování aktivit vzniklo kombinací a úpravou proměnných používaných v rámci standardního logovacího mechanismu LMS Moodle a obsažených v datech před zpracováním. Základní výčet aktivit studenta v průběhu plnění online testů v LMS Moodle a způsob jejich pojmenování lze vidět v následující tabulce (viz tabulka 18). Jak je patrné, označení aktivity je kombinací jednak informace o provedené akci (*viewed*, *started*, *submitted*, *reviewed*), jednak informace o tom, čeho se provedená akce týká (*quiz*, *quiz_attempt*, *attempt_summary*).

⁹⁴ Podrobné informace o fungování logovacího mechanismu lze nalézt v oficiální dokumentaci k systému na adrese docs.moodle.org/31/en/Logging.

Tabulka 18: Přehled základních názvů aktivit v průběhu plnění testu po předzpracování dat

Názvy aktivit v testu	Zkratka	Význam
quiz_viewed	qv	Student si zobrazil úvodní stranu testu se základními informacemi a pokyny k jeho vyplňování.
quiz_attempt_started	qas	Student začal nový pokus v testu.
quiz_attempt_viewed	qav	Student si zobrazil stránku s otázkami testu. (Otázky v testu mohou být na jedné či více stránkách.)
attempt_summary_viewed	asv	Student se dostal na závěrečnou přehledovou stranu testu.
quiz_attempt_submitted	qasub	Student odeslal pokus v testu k vyhodnocení.
quiz_attempt_reviewed	qar	Student si zobrazil zpětnou vazbu k předchozímu pokusu v testu. (Tato možnost je dostupná jen v případě, že to vyučující u daného testu povolí.)

Obdobným způsobem jako záznamy v logovacím systému týkající se přímo aktivity v testech byly předzpracovány a pojmenovány záznamy související s jinými částmi kurzu a LMS jako celku. Zde bylo pouze nutné navíc rozlišit jiné testy dostupné v kurzu a test, který podléhal analýze. Důvodem bylo to, aby bylo možné odlišit aktivitu přímo v daném analyzovaném testu od případné aktivity v jiných testech v kurzu. Výše zmíněná aktivita *quiz_viewed* tudíž znamená, že si student zobrazil úvodní stranu analyzovaného testu, kdežto aktivita *o.quiz_viewed* značí zobrazení úvodní strany jiného testu v daném kurzu. Obdobně je pak zobrazení testu v úplně jiném kurzu označeno jako *oc.quiz_viewed* (viz tabulka 19).

Související aktivity byly zároveň seskupovány do obecnějších kategorií, jak je naznačeno rovněž v tabulce 19. Seskupení se týkalo jednak plnění různých druhů aktivit studenta v kurzu (např. odevzdávání úkolu, přispívání do fóra, plnění jiných testů apod.), jednak zobrazování různých druhů studijních materiálů v kurzu (soubor, stránka, odkaz apod.).

Tabulka 19: Příklady názvů dalších aktivit (tj. aktivit mimo samotný test) po předzpracování dat

Příklady názvů dalších aktivit	Název skupiny aktivit	Zkratka
course_viewed	course_viewed	cv
other_course_viewed	other_course_viewed	ocv
o.quiz_viewed	activity_viewed	av
oc.quiz_viewed	activity_viewed	av
page_viewed	material_viewed	mv
book_chapters_viewed	material_viewed	mv

Příklady názvů dalších aktivit	Název skupiny aktivit	Zkratka
resource_viewed	material_viewed	mv
user_viewed	user_info_viewed	uiv
user_grades_viewed	user_info_viewed	uiv
lesson_started	activity_performed	ap
questionnaire_submitted	activity_performed	ap
discussion_post_uploaded	activity_performed	ap
oc.quiz_attempt_viewed	activity_performed	ap

9.3.4 Analýza dat

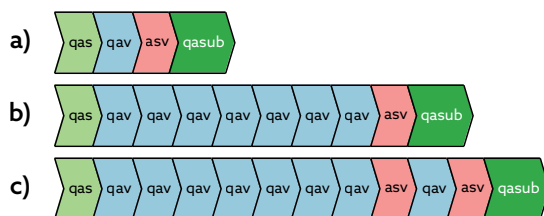
V rámci analýzy dat jsem v prvním kroku provedl exploraci variant procesů (*trace variants*), tj. případů, které obsahují stejné sekvence aktivit. Zde bylo základním cílem jednak zjistit, jaké varianty procesů se v analyzovaném datovém souboru vyskytují, jednak následně hledat podobnosti mezi různými variantami procesů a seskupovat takové varianty procesů, které vykazují obdobné vzorce chování studentů při vyplňování testů. V druhém kroku pak bylo přistoupeno k tvorbě a analýze tzv. procesních map, k jejichž průzkumu byl využit nástroj *Disco*, používající vlastní upravenou a rozšířenou implementaci algoritmu *Fuzzy Miner* (Günther & van der Aalst, 2007) zvanou *Disco miner*.

Důvodů pro využití procesních map spíše než jiných forem modelů bylo několik. Vzhledem k exploratornímu charakteru analýzy šlo primárně o průzkum a pokud možno komplexní poznání reálného procesu než o budování obecného modelu a jeho ověřování. Procesy zachycující reálné chování lidí přitom obvykle bývají méně strukturované, s čímž mívají „tradiční“ algoritmy (např. zmiňované *alpha algorithm* či *Heuristic Miner*) určité problémy, a nedokážou tak poskytnout dostatečný vhled do charakteru analyzovaného procesu. Buď poskytnou příliš obecný model, který tudíž není schopen zachytit méně často se vyskytující formy chování, nebo vyprodukují tzv. *spaghetti-like model*, který je příliš složitý, a tudíž jen obtížně srozumitelný a prakticky nepoužitelný (srov. Günther & van der Aalst, 2007; van der Aalst, 2016). Využití algoritmu *Fuzzy Miner* a modelů v podobě procesních map naopak dovoluje interaktivní úpravu úrovně abstrakce a následnou volbu odpovídajícího, a přitom smysluplného popisu studovaného procesu.

9.4 Výsledky

Na obrázku 25 lze vidět příklady průchodů testem, které vykazují standardní chování studentů při plnění testů (*standard quiz-taking behavior*). Ve všech třech

9 Analýza 3: Chování studentů v průběhu plnění online testů

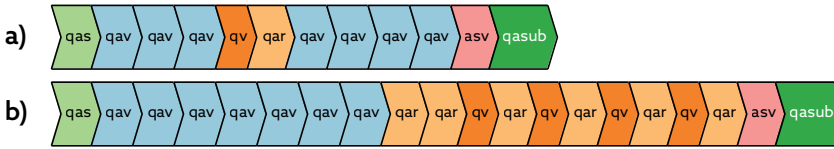


Obrázek 25: Příklad variant procesů vykazujících standardní typ chování studentů při vyplňování testů

případech student postupuje tak, že nejprve spustí test (*qas*), poté vyplňuje úlohy v testu (*qav*) a následně si zobrazuje závěrečnou přehledovou stránku testu (*asv*), načež svůj pokus v testu odesílá k vyhodnocení (*qasub*). Varianta a) přitom ukazuje případ, kdy jsou všechny úlohy v testu na jedné stránce, kdežto varianty b) a c) pochází z testů obsahujících úlohy na více stránkách. Varianta c) pak zároveň znázorňuje případ, kdy se student ze závěrečné přehledové strany (*asv*) vrací zpět k vyplňování úloh (*qav*). To může být způsobeno např. tím, že si na závěrečné přehledové straně všiml nějaké opomenuté nezodpovězené úlohy, a tudíž se vrátil zpět k jejímu vyplnění. Je zároveň důležité zmínit, že v testu E se vyskytoval pouze tento typ chování studentů (tj. standardní vyplňování testu). Naopak ve všech ostatních analyzovaných testech se vyskytovaly i typy chování studentů, které lze označit za nestandardní vyplňování testu (*non-standard quiz-taking behavior*). Tento výsledek naznačuje, že v kontrolovaných podmínkách (tj. vyplňování testu přímo v počítačové učebně a pod dohledem vyučujícího) se studenti spíše neuchylují k nestandardnímu vyplňování testů, a naopak postupují tak, jak je očekáváno.

Obrázek 26 ilustruje způsob chování studentů při plnění online testů, který by bylo možné označit jako zneužití zpětné vazby (*feedback misuse*). Tento typ chování studentů lze charakterizovat tak, že si student v průběhu vyplňování testu zobrazí stránku s přehledem podrobné zpětné vazby k předchozímu pokusu v daném testu. Na této stránce se přitom zobrazují jednak odpovědi studenta z předchozího pokusu, jednak také informace o tom, zda byla uvedená odpověď správně či nikoli. Zároveň se může zobrazovat také doplňující zpětná vazba k jednotlivým úlohám, či dokonce správná odpověď (v závislosti na nastavení testu a zpětné vazby učitelem). Tento typ chování studenta se tedy může vyskytovat jen v takových testech, které jsou využívány pro opakované vyplňování (tj. dovolují více pokusů v testu) a zároveň jsou nastaveny takovým způsobem, že studentům povolují zobrazování zpětné vazby k předchozím pokusům. V případě analyzovaných testů umožňují tento typ chování testy A a C, přičemž v obou testech se tento typ chování studentů vyskytl. Častěji se přitom objevoval v testu A.

Další identifikovaný typ chování studentů v průběhu vyplňování testů lze označit jako zneužití studijních materiálů (*study materials misuse*). V případě zneužití



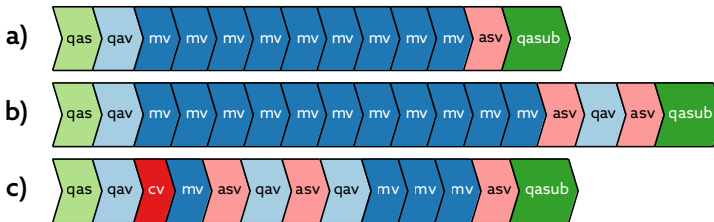
Obrázek 26: Příklad variant procesů vykazujících chování typu zneužití zpětné vazby

studijních materiálů (viz obrázek 27) student v průběhu plnění testu nahlíží do studijních materiálů obsažených v kurzu. Obvykle jde přitom o opakovanou činnost, kdy student postupně otevírá různé studijní materiály dostupné v kurzu. Lze tedy předpokládat, že se student v průběhu pokusu snaží dohledávat správné odpovědi na úlohy v testu. Zkratka *mv* zde přitom zahrnuje různé typy studijních materiálů, které se mohou v různých kurzech lišit.

V souvislosti s výše identifikovaným typem chování studentů (tj. zneužití zpětné vazby) lze přitom zmínit zajímavé zjištění, že ačkoli se oba typy chování mohou teoreticky vyskytovat v průběhu jediného pokusu, v analyzovaných testech se tyto typy chování v rámci jediného pokusu nikdy nevyskytovaly. To by mohlo naznačovat, že studenti spíše volí buď jednu, nebo druhou strategii nestandardního chování. Pro podporu této teze by však bylo třeba provést obdobnou analýzu na mnohem širším vzorku testů.

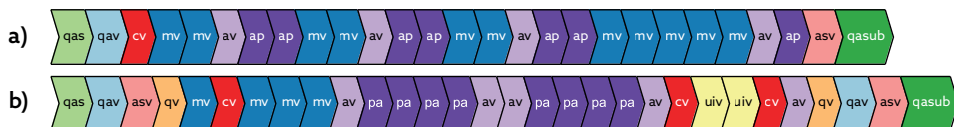
Poslední detekovaný typ chování studentů při plnění testů by se dal označit jako *multitasking*. Toto označení je zvoleno proto, že se studenti v průběhu plnění hlavní aktivity (tj. plnění pokusu v daném testu) zároveň zapojují do jiných (vedlejších) aktivit, které s plněním testu přímo nesouvisí. Na obrázku 28 jsou znázorněny dva příklady takového chování, přičemž v obou případech bylo onou vedlejší aktivitou vyplňování jiného testu v kurzu.

Při podrobnějším průzkumu těchto variant přitom bylo možné odhalit, že se jedná o dvě různé situace. V první situaci (varianta a) student začne vyplňovat analyzovaný test (*qas*) a následně prohlíží studijní materiály (*mv*). Poté ovšem otevírá jiný test (*ap*) a začíná pokus v něm. Přičemž toto chování se opakuje ještě třikrát a až poté student ukončuje pokus v původním test (*qasub*). V průběhu plnění



Obrázek 27: Příklad variant procesů vykazujících chování typu zneužití studijních materiálů

9 Analýza 3: Chování studentů v průběhu plnění online testů



Obrázek 28: Příklad variant procesů vykazujících chování typu multitasking

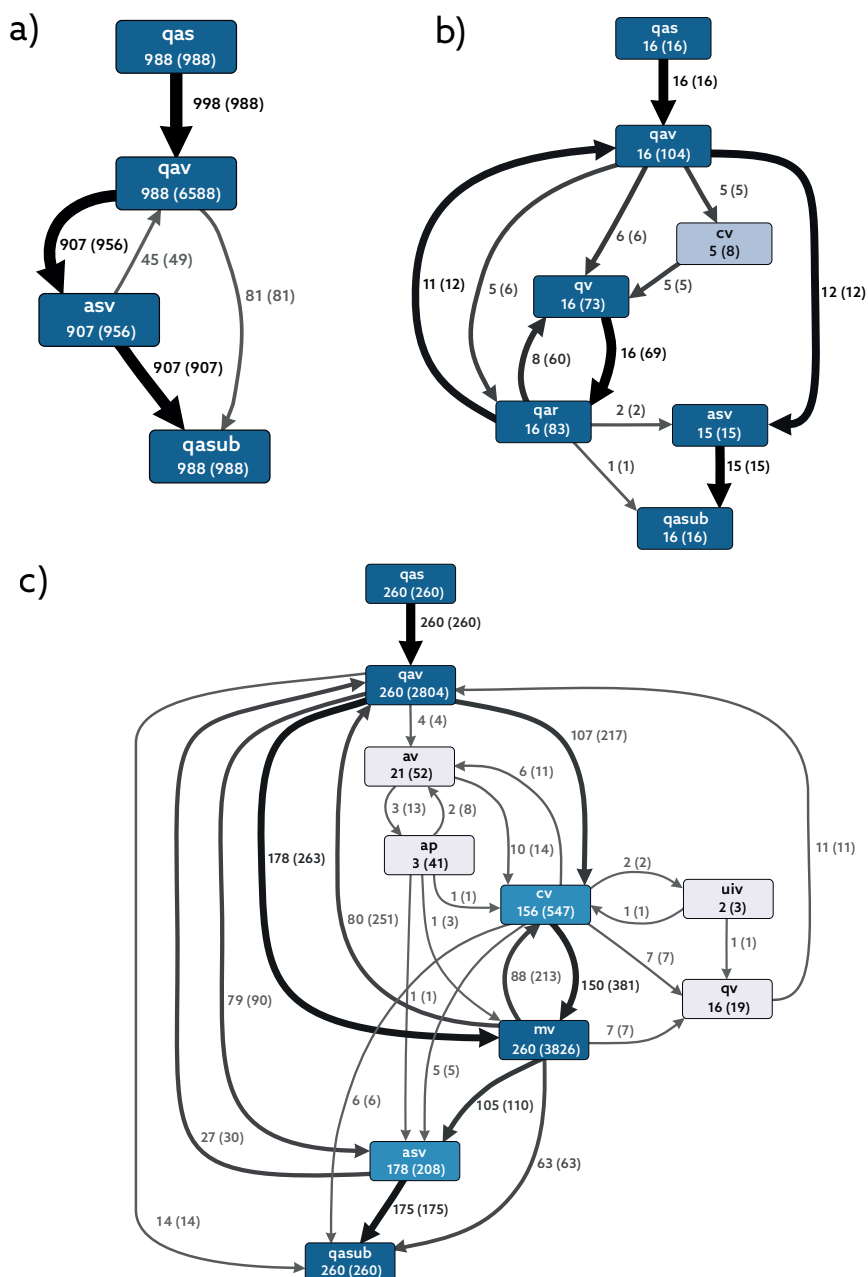
původního testu však student ve vedlejších testech pokusy pouze odstartoval, ale (prozatím) nedokončil. V druhém příkladu (varianta b) naopak student odstartuje i dokončí pokusy ve dvou jiných testech, a teprve poté se vrací k hlavnímu testu a ukončuje svůj pokus i zde.

Zajímavým zjištěním u tohoto typu chování bylo především to, že se nevyskytoval pouze u testu B. Test B je totiž oproti ostatním analyzovaným testům specifický v tom, že nemá nastaven pevný časový limit. Studenti jej tak mohou vyplňovat v průběhu delší doby. Nabízelo by se tedy vysvětlení, že někteří studenti v průběhu plnění tohoto testu mohou začít zároveň plnit jiné aktivity v kurzu a k dokončení testu se vrátit později. Za poměrně překvapivé lze však považovat to, že se tento typ chování objevil i v testu C, u něhož jde o závěrečný test s pevným časovým limitem a poměrně velkým množstvím otázek k zodpovězení.

Doplňme také, že mimo výše uvedené typy chování studentů v průběhu plnění testů bylo možné v datech nalézt i takové varianty procesů, které nezapadají do žádného z uvedených typů. Zároveň jde o případy, které lze jen obtížně interpretovat čistě na základě dat zaznamenávaných systémem (tj. na základě logů). Příkladem mohou být situace, kdy si student v průběhu vyplňování testu zobrazí hlavní stranu kurzu či hlavní stranu testu, ale již nepokračuje nikam dál (např. na nějaký studijní materiál či na stránku se zpětnou vazbou). Pro porozumění těmto zvláštním případům by tak zřejmě bylo třeba využít některou z kvalitativních výzkumných metod.

Zatímco výše byly představeny jednotlivé typy chování studentů v průběhu vyplňování testů (tj. *quiz-taking behavior*) na příkladech konkrétních variant procesů, obrázek 29 shrnuje všechny varianty procesů jednotlivých typů chování studentů v podobě tří procesních map⁹⁵. Procesní mapy A a B zobrazují proces kompletně, u procesní mapy C jde o zjednodušený model, který vynechává některé méně významné cesty. U všech tří procesních map zároveň nejsou pro větší přehlednost zobrazeny tzv. *self-loops* (tj. případy, kdy vede cesta z aktivity určitého typu k aktivitě stejného typu – typicky např. postupné zodpovídání testových otázek na více stranách). U každé aktivity a cesty jsou pak zobrazeny frekvence případů (*case frequency*) a následně v závorce absolutní frekvence (*absolute frequency*) výskytu dané aktivity či cesty.

⁹⁵ Procesní mapa C zahrnuje jak chování typu zneužití studijních materiálů, tak i chování typu multitasking. Důvodem je to, že v analyzovaných datech se ve všech případech vyskytovalo chování typu multitasking společně s chováním typu zneužití studijních materiálů.



Obrázek 29: Procesní mapy různých typů chování studentů při plnění testů: a) standardní chování, b) zneužití zpětné vazby, c) zneužití studijních materiálů spolu s typem multitasking

9 Analýza 3: Chování studentů v průběhu plnění online testů

Z uvedených procesních map lze zároveň vyčíst četnosti jednotlivých typů chování studentů při vyplňování testů. Z celkového počtu 1 298 pokusů ve všech pěti analyzovaných testech se v 988 pokusech (tj. 76,12%) vyskytovalo standardní chování studentů (*standard quiz-taking behavior*). Chování typu zneužití studijních materiálů (*study materials misuse*) se pak vyskytlo v celkem 260 pokusech (20,03%), a jde tak o hlavní typ nestandardního chování studentů při plnění testů (*non-standard quiz-taking behavior*), který byl v analyzovaných datech z LMS Moodle odhalen. Zbývající dva typy nestandardního chování studentů v testech se pak vyskytovaly v mnohem menším počtu případů. Chování typu zneužití zpětné vazby (*feedback misuse*) se objevilo pouze v 16 pokusech (1,23%) a chování typu *multitasking* se vyskytlo dokonce pouze ve třech pokusech (tj. 0,23%). Celkově přitom bylo možné pomocí výše uvedených typů popsat 97,38% všech analyzovaných případů.

Zároveň je vhodné zmínit, že jednotlivé procesní mapy dovolují odhalovat i další zajímavé skutečnosti o chování studentů v průběhu vyplňování testů. Studenti např. poměrně často neodesílají test k vyhodnocení sami, nýbrž je jejich pokus odeslán k vyhodnocení automaticky, tedy poté, co vyprší stanovený časový limit pro vyplnění testu (v procesních mapách jde o všechny případy, kdy cesta na stránku *qasub* nevede přes stránku *asv*). To si lze vysvětlit buď tím, že studenti spoléhají na funkcionalitu systému, která pokus po vypršení časového limitu odešle automaticky (což ovšem nemusí platit vždy, jelikož je to dáno nastavením každého individuálního testu), nebo se do poslední chvíle snaží hledat odpověď např. ve studijních materiálech (viz 63 takových případů v procesní mapě C). Nezanedbatelná část studentů se pak ze závěrečné přehledové strany testu (*asv*) vrací zpět k vyplňování otázek (*qav*). To může být dáno nejen tím, že si student všiml opomenuté úlohy, nýbrž to může naznačovat i zapojení specifických strategií plnění testů (tzv. *quiz taking strategies*). Student může například nejprve rychle vyplnit úlohy, na které zná okamžitě odpověď, a teprve poté se postupně vracet k řešení obtížnějších úloh.

V neposlední řadě pak procesní mapy ukazují, že studenti poměrně běžně při vyplňování testů pracují s více okny/panely internetového prohlížeče. Je to patrné např. z procesní mapy C, kde v 178 případech pokračuje student ze strany s otázkami v testu (*qav*) přímo na stránku konkrétního studijního materiálu (*mv*), což je možné jen v případě, že má student v jednom okně/panelu spuštěný test, zatímco v druhém okně/panelu má zobrazenou např. hlavní stranu kurzu či stránku s přehledem studijních materiálů, ze které si pak již otevírá konkrétní studijní materiál.

9.5 Diskuze a závěr

Tato analýza využívala procesně orientovaný přístup k analýze dat o chování studentů v LMS Moodle při plnění online testů. V rámci tohoto přístupu přitom byla využita metoda dolování procesů, která umožňuje analýzu dat procesního charakteru. Přestože již existuje několik studií, které tuto metodu v kontextu vzdělávání využívají (např. Schoor & Bannert, 2012; Bannert, Reimann & Sonnenberg, 2014; Papamitsiou & Economides, 2016), či ji dokonce využívají přímo k analýze dat z LMS Moodle (Romero et al., 2016), obecně jsou procesně orientované přístupy a specificky metoda dolování procesů ve vzdělávání využívány spíše jen zřídka (srov. van der Aalst, Guo & Gorissen, 2013; Peña-Ayala, 2014b).

Doposud však neexistuje žádná studie, která by se věnovala aplikaci metod dolování procesů přímo v souvislosti s analýzou a detekcí různých forem chování studentů při plnění testů v rámci LMS Moodle. Primárním cílem této exploratorně pojaté analýzy tak bylo prozkoumat možnosti dolování procesů jakožto metody analýzy k tomuto účelu. Zároveň byla analyzována data pocházející z testů různých typů, čímž tato analýza usilovala o odhalení různých způsobů chování studentů při vyplňování širší škály různých testů. V následujících dvou podkapitolách nejprve shrnuji odpovědi na stanovené výzkumné otázky, a poté zmiňuji hlavní limity realizované analýzy, jakož i možnosti dalšího výzkumu a důsledky zjištěných výsledků pro praxi.

9.5.1 Shrnutí odpovědí na stanovené otázky

První výzkumná otázka se týkala obecně využitelnosti metody dolování procesů pro detekci specifických forem či způsobů chování studentů v průběhu plnění online testů v LMS Moodle. Přesněji pak bylo hlavním zájmem to, zda využití výhradně procesního přístupu k analýze dat o chování studentů umožní rozlišit různé typy či vzorce chování studentů při plnění testů. Čili zda je možné detekovat různé chování studentů při vyplňování testů jen za využití metod dolování procesů a dat procesního charakteru (logů) bez kombinace s dalšími analytickými postupy či typy dat (např. výše zmiňované přístupy založené na person-fit indexech a přístupy založené na rychlosti odpovídání). Výsledky analýzy přitom naznačují kladnou odpověď. S využitím metody dolování procesů bylo možné identifikovat čtyři různé typy chování studentů při plnění online testů. Zároveň bylo možné odlišit standardní chování studentů (*standard quiz-taking behavior*) od nestandardního chování (*non-standard quiz-taking behavior*) a v rámci nestandardního chování dále rozlišit tři různé typy. Lze tak souhlasit se závěry dosavadních studií (Papamitsiou & Economides, 2016; Romero et al., 2016), že využití metody dolování procesů k analýze dat o chování studentů v LMS může být přínosné zvláště pro objevování, mapování

a monitorování různých učebních procesů. Tento metodologický přístup se tak jeví být vhodným doplněním dosavadních analytických postupů a nástrojů.

Ve druhé výzkumné otázce byla zaměřena pozornost na to, jaké typy chování studentů v průběhu plnění online testů (tj. *quiz-taking behavior*) je možné s využitím metody dolování procesů odhalit. Jak bylo naznačeno v předchozím odstavci, celkem bylo možné v analyzovaných datech identifikovat a odlišit čtyři různé způsoby chování studentů při plnění testů. V prvním případě šlo o standardní chování, kdy studenti postupují dle očekávání, tzn.: spustí test, splní jednotlivé úlohy a odešlou své odpovědi k vyhodnocení. Ve zbývajících třech případech pak šlo o různé typy nestandardního testového chování. Konkrétně pak byly rozlišeny následující tři formy nestandardního chování studentů: a) chování typu zneužití zpětné vazby (*feedback misuse behavior*), při kterém studenti v průběhu plnění otázek v testu využívají zpětnou vazbu k předchozím pokusům v testu, b) chování typu zneužití studijních materiálů (*study materials misuse behavior*), u kterého studenti při vyplňování testů nahlíží do studijních materiálů dostupných v kurzu a c) chování typu multitasking (*multitasking behavior*), kdy se studenti při vyplňování testu zároveň zapojují do jiných výukových aktivit v rámci kurzu.

Zde lze přitom vidět určité paralely s existujícími výzkumy, zvláště pak s výzkumy realizovanými v kontextu inteligentních tuteurských systémů (např. Baker, 2007; Baker et al., 2009; Baker, Goldstein, & Heffernan, 2011; Baker & Gowda, 2010; Muldner, Burleson, Van de Sande, & VanLehn, 2011). Příkladem může být odhalený typ chování označený jako zneužití zpětné vazby (*feedback misuse*), který lze do značné míry chápat jako odpovídající tomu, co je v kontextu systémů typu ITS označováno jako zneužívání nápovědy (*misusing help*) či obcházení systému (*gaming the system*).

Přesto je třeba zároveň upozornit na to, že výše identifikované typy chování nelze zcela ztotožnit s typy chování, které byly identifikovány v jiných studiích. Mimo jiné proto, že systémy typu ITS a systémy typu LMS jsou poměrně zásadně odlišné výukové technologie, tudíž umožňují různé formy nestandardního chování studentů. Např. výše odhalený typ chování označený jako zneužití studijních materiálů lze zřejmě považovat za specifický právě pro systémy typu LMS. Navíc využití procesně orientovaného přístupu k detekci nestandardního chování studentů při plnění testů ve své podstatě sleduje velmi odlišnou rovinu či odlišný aspekt chování studenta než například častěji používané přístupy založené na délce odpovědi (*response time*). Lze tak do značné míry předpokládat, že využití takto výrazně odlišných přístupů bude zároveň odhalovat odlišné typy či způsoby chování studentů, a to i přesto, že se na první pohled mohou zdát podobné.

Poslední výzkumná otázka se pak týkala četnosti výskytu jednotlivých typů chování studentů při vyplňování online testů. V analyzovaných datech pocházejících z pěti testů různých typů se standardní testové chování vyskytovalo v cca 76 % všech případech (tj. pokusech studentů v testech), zatímco ve více jak pětina případů

(cca 21,5 %) docházelo k výskytu nestandardní chování při plnění testů. V naprosté většině šlo přitom o nestandardní chování typu zneužití studijních materiálů (cca 20 %), kdežto ostatní typy nestandardního chování studentů se vyskytovaly jen zřídka. Zbývajících cca 2,5 % případů pak odpovídá takovému chování studentů, které zůstalo v rámci realizované analýzy nevysvětleno.

Zároveň je však třeba zdůraznit, že z hlediska výskytu jednotlivých typů nestandardního chování studentů v testech se ukázal být zcela zásadním kontext. A to jednak co se týče typu testu a jeho „role“ v kurzu, jednak co se týče konkrétního způsobu nastavení testu v LMS, jakož i konkrétních podmínek, za jakých je test studenty vyplňován. Například výskyt chování typu zneužití zpětné vazby je možný jen v případě určitého nastavení testů. Bez odpovídajícího nastavení tak vůbec nemůže nastat. Zde se tudíž projevuje role učitele (či tvůrce kurzu), který vlastně zvoleným nastavením testu může určit, zda vůbec určitý typ nestandardního testového chování studentů bude možný. Jiným příkladem je analyzovaný test E, který byl studenty vyplňován tzv. *on-site*, tj. v počítačové učebně a pod dohledem vyučujícího. V tomto testu se totiž vyskytovalo pouze standardní chování studentů, což může naznačovat, že v takovém kontextu se studenti spíše neuchylují k nestandardnímu chování, přestože by třeba technicky (tj. z hlediska vlastností či nastavení systému) bylo možné. Zbývá dodat, že k obdobným závěrům o důležitosti vzdělávacího kontextu a zohlednění konkrétních podmínek kurzu či nastavení daného výukového nástroje dospívají také Agudo-Peregrina et al. (2014) či Gašević et al. (2016).

9.5.2 Limity, možnosti dalšího výzkumu a důsledky pro praxi

V souvislosti s limity analýzy lze rozlišit na jedné straně limity dané pojetím a zaměřením analýzy, na druhé straně pak limity dané podstatou a rozsahem analyzovaných dat. Především je třeba uvést, že v této analýze byla věnována pozornost výhradně mapování „pohybu“ studentů v LMS v průběhu plnění testu. Z toho automaticky plyne nutné omezení související s tím, jaké typy chování studentů je vůbec možné tímto přístupem detekovat. Zároveň je velmi pravděpodobné, že kombinací dolování procesů s dalšími technikami (např. shlukování) a zahrnutím dalších typů dat by bylo možné dosáhnout mnohem podrobnějšího a přesnějšího vhledu do chování studentů v průběhu plnění testů.⁹⁶ Další výzkum je tak možné směřovat především ke kombinaci procesně orientovaného přístupu s dalšími používanými přístupy k detekci nestandardního chování studentů při plnění testů.

96 Např. pokud bychom navíc sledovali časový aspekt (tj. délku pokusu či přímo délku odpovídání na jednotlivé úlohy), tak lze očekávat, že u určitého množství případů, které jsou s využitím výhradně dolování procesu zařazeny do kategorie standardní chování, by bylo odhaleno chování typu hádání, resp. tipování (tzv. *rapid-guessing behavior*).

Jinou možností rozšíření výsledků této analýzy by bylo sledování rozdílů v chování u úspěšných a neúspěšných studentů (srov. Romero et al., 2016). Zajímavý směr dalšího výzkumu pak nabízí i van der Aalst, Guo a Gorissen (2013), kteří se věnují možnostem tzv. srovnávacího dolování procesů (*comparative process mining*).

Co se týče podstaty analyzovaných dat, je třeba zmínit, že logy ve skutečnosti nezaznamenávají reálné chování studentů v LMS ve své úplnosti, nýbrž vždy nutně zachycují pouze určitý fragment či výsek možného chování (srov. van der Aalst, 2016). A to jednak v závislosti na technickém řešení logovacího mechanismu v rámci daného systému (tj. ne všechno chování musí být zaznamenáváno), jednak také proto, že vždy pracujeme jen s určitým časovým úsekem. Nelze tak předpokládat, že by analýza zachytila všechny možné varianty procesů. Z hlediska rozsahu dat byl také analyzován jen omezený počet testů a jejich typů, takže je možné, že určité typy chování vůbec nebyly v datech zastoupeny, jelikož se vyskytují jen ve specifických podmínkách konkrétního způsobu využití či nastavení testu (podobně jako výše zmiňované zneužití zpětné vazby, které se může vyskytovat jen při určitém nastavení testu). Potenciál budoucích studií tak spočívá i v analýze výrazně většího množství dat z širšího spektra různých typů testů.

I přes uvedené limity lze výsledky analýzy považovat za přínosné z hlediska praxe. Realizaci analýzy lze totiž chápat jako určitou demonstraci toho, že dolování procesů lze využít k identifikaci (ne)standardního chování studentů při plnění testů, a to na základě dat, které v současnosti již LMS Moodle implicitně sbírá. Bylo by tak možné vytvořit analytický nástroj pro učitele, který by jim například v podobě vizualizace procesních map či variant procesů umožnil zjistit, zda studenti při plnění zadaného testu postupovali standardně či nikoli. To by pro učitele mohlo mít zásadní přidanou hodnotu zvláště u testů, které se podílejí na finálním hodnocení. Vizualizace v podobě procesních map má navíc oproti jiným způsobům vizualizace procesů tu výhodu, že je poměrně intuitivní a může být dobře srozumitelná i pro učitele. Využití procesně orientovaného přístupu představeného v této studii, tak může v kontextu praktického využití sloužit jako první krok ke vzniku analytického nástroje poskytujícího učitelům lepší vhled do procesu vyplňování testů studenty.