

Západočeská univerzita v Plzni
Fakulta aplikovaných věd
Katedra informatiky a výpočetní techniky

Bakalářská práce

Personalizace elektronického obchodu

Prohlášení

Prohlašuji, že jsem bakalářskou práci vypracoval samostatně a výhradně s použitím citovaných pramenů.

V Plzni dne 5. května 2015

Ondřej Hovjcký

Poděkování

Rád bych poděkoval firmě RTsoft za ochotu a cenné rady při vypracování této práce a vstřícnost při konzultacích.

Abstract

This work deals with personalization of the electronic shop. It points out the possibilities in this area, how a customer needs to be tracked and based on that, how to do the personalization itself. The Prestashop System, which is used for Internet shops, is described. In the practical part, some of these principles are implemented and tried out. Apart from the algorithm for choosing products which would be offered to a customer, e-newsletter sending is dealt with too. Their deliverability is examined and a customer's view of the e-mail is tracked.

Abstrakt

Tato práce pojednává o personalizaci elektronického obchodu. Ukazuje, jaké jsou v této oblasti možnosti, co je třeba o zákazníkovi sledovat a na základě toho pak provádět samotnou personalizaci. Následně je zde v krátkosti popsán systém Prestashop, který se používá pro internetové obchody. V praktické části jsou některé principy v tomto systému implementovány a vyzkoušeny. Kromě algoritmu výběru produktů, které se zákazníkovi nabídnou, se také řeší rozesílání e-newsletterů. Zkoumá se jejich doručitelnost a sleduje se zobrazení e-mailu zákazníkem.

Obsah

1	Úvod	3
2	Možnosti personalizace	4
2.1	Aspekty personalizace	4
2.1.1	Personalizace internetových stránek	5
2.1.2	Personalizace e-newsletteru	5
2.2	Sledování zákazníka	6
3	Prestashop	8
3.1	Architektura Prestashopu	8
3.2	Modul Prestashopu	13
4	Návrh systému	15
4.1	Modul sledování zákazníka	15
4.1.1	Sbírání dat	15
4.1.2	Statistiky	16
4.2	Modul personalizované nabídky produktů na stránkách	17
4.2.1	Nabízení produktů zákazníkovi	17
4.2.2	Statistiky proklikovosti nabízených produktů	17
4.3	Modul obsluhující personalizovaný newsletter	18
4.3.1	Posílání newsletterů	18
4.3.2	Sledování zobrazení vzdáleného obsahu	19
4.3.3	Statistiky proklikovosti nabízených produktů	19
5	Popis implementace	20
5.1	Modul sledování zákazníka	20
5.1.1	Sbírání dat	20
5.1.2	Statistiky	21
5.2	Modul personalizované nabídky produktů na stránkách	22
5.2.1	Nabízení produktů zákazníkovi	23
5.2.2	Sledování kliknutí na nabízený produkt	24
5.3	Modul obsluhující personalizovaný newsletter	24
5.3.1	Posílání newsletterů	25

5.3.2	Sledování kliknutí na nabízený produkt	25
5.3.3	Sledování zobrazení vzdáleného obsahu	26
6	Testování a zhodnocení výsledků	27
7	Závěr	30

1 Úvod

Cílem této práce je navrhnout a implementovat systém pro personalizaci internetového obchodu – konkrétně zobrazování personalizované nabídky produktů na stránkách a rozesílání personalizovaných newsletterů. Jelikož je v dnešním světě nabídka různých produktů obrovská, každý obchodník se snaží maximalizovat své zisky. Aby toho docílil, obvykle si také založí internetový obchod, do kterého má přístup mnohem více zákazníků, než do obchodu kamenného. Tam se jim pak snaží nabízet produkty, které je mohou zaujmout, a pokud se zákazník v nějakém obchodě zaregistruje, např. při objednávce produktu, obchodník se ho poté snaží přimět k opětným návštěvám obchodu a ke koupi dalších produktů.

K tomu používá elektronické propagační materiály zasílané e-mailem – tzv. e-newslettery. Těch se v dnešní době rozesílá značné množství a mnoho jich končí rovnou v koši. Proto je pro obchodníky aktuální téma zavedení personalizace nabídky, která má za cíl upoutat uživatele něčím, co jej pravděpodobně zajímá. Nedostává tedy v newsletteru obecný obsah, ale takový o kterém si obchodník myslí, že ho nejvíce zaujme.

Aby se toto dalo realizovat, používají se různé techniky pro zjištění, jaká nabídka se bude danému zákazníkovi líbit. V této práci budeme sledovat zákazníkův pohyb na stránkách internetového obchodu a z nasbíraných dat se snažit určit, jaké produkty se mu líbí, jaké kategorie navštěvuje nejvíce, na jakých stránkách tráví více času apod.

2 Možnosti personalizace

V této kapitole se seznámíme s možnostmi personalizace internetového obchodu. Jedná se o velmi rozsáhlé téma, které se neustále vyvíjí a rozšiřuje. Obchodník se snaží vytvořit zákazníkovi (nebo uživateli, dále jsou tato slova brána jako synonyma) pohodlné a uživatelsky přívětivé prostředí internetového obchodu a zároveň se snaží přimět zákazníka ke koupi a tím maximalizovat svůj zisk [Quah et al.(2008)].

Vyjmenujme si některé z nejpoužívanějších aspektů. Tyto aspekty se dají rozdělit do dvou větších kategorií:

1. personalizace internetových stránek – jedná se o úpravu zobrazení stránky. Každý zákazník vidí stránku zobrazenou jinak.
 - výběr produktů na domovskou stránku
 - řazení produktů v jednotlivých kategoriích
 - značka, velikost, cena
 - oblíbené produkty
 - geologická lokace uživatele
 - řazení menu
2. personalizace e-newsletteru – jedná se o posílání e-mailů upravených podle dat nasbíraných sledováním chování registrovaného zákazníka s cílem motivovat jej k návštěvě obchodu a ke koupi dalších produktů či služeb.
 - obsah
 - čas doručení
 - předmět e-mailu

2.1 Aspekty personalizace

Nyní si podrobněji rozebereme aspekty vypsané výše.

2.1.1 Personalizace internetových stránek

Aby se zákazníkovi nakupovalo pohodlně a aby obchodník maximalizoval své zisky, stránky se dají upravit různými způsoby.

- Výběr produktů na domovskou stránku – když zákazník vstoupí do obchodu, na domovské stránce se většinou zobrazí několik produktů. Jedná se o různé akce a slevy či nové produkty. Pokud však na stránky přijde zákazník, který už na nich jednou byl a víme, že se mu líbila určitá kategorie, můžeme mu zobrazit nějaké produkty právě z této kategorie.
- Řazení produktů v jednotlivých kategoriích – při vstupu do určité kategorie je třeba produkty nějak seřadit. Opět mohou hrát roli slevy nebo nové produkty, ale také následující faktory:
 - značka, velikost, cena – víme-li, že zákazník v minulosti hledal produkty podle těchto parametrů, můžeme je při řazení upřednostnit.
 - oblíbené produkty – můžeme mu zobrazit produkty, které koupil více než jedenkrát a je pravděpodobné, že je bude chtít koupit znovu.
 - geologická lokace uživatele – pokud nabízíme např. služby, které jsou poskytovány jen v určitých městech, zobrazíme spíše ty, které jsou blízko uživatelova bydliště (v případě že známe jeho polohu).
- Řazení menu – je možné také seřadit menu pro každého zákazníka trochu jinak, podle preferencí, které jsme zjistili při jeho minulých návštěvách.

2.1.2 Personalizace e-newsletteru

Obchodník má dále možnost zasílat registrovaným zákazníkům e-maily s akcemi. Pokud máme o zákazníkovi dostatek informací, můžeme mu e-maily přizpůsobovat následovně.

- Obsah – zákazníkovi pošleme jen produkty, o kterých si myslíme, že o ně bude mít zájem – tj. např. prohlížel si dané kategorie.

- Čas doručení – víme-li, že zákazník navštívuje stránky např. ve večerních hodinách, e-newsletter mu také pošleme někdy večer. Zvýšíme tak pravděpodobnost, že si jej přečte.
- Předmět e-mailu – aby zákazník e-newsletter vůbec otevřel, je také důležitý předmět e-mailu, který může obsahovat krátké heslo či delší nabídku. Podle měření, jestli se zákazníkovi e-mail otevřel, můžeme zjistit, jaké předměty ho upoutají více.

2.2 Sledování zákazníka

Abychom mohli personalizaci provádět, musíme se o zákazníkovi něco dozvědět – nasbírat si o něm určitá data. Nyní se podíváme na možnosti sledování zákazníka.

Zákazník, který vstoupí na stránky obchodu má z hlediska sledování dva základní módy – může být přihlášen nebo nepřihlášen. V každém případě (vyjma případu, kdy nemá zákazník povoleny cookies) jej však můžeme sledovat a to za pomoci cookie, do které zapíšeme zákazníkům jedinečný identifikátor. Cookie je informace uložená na uživatelském počítači, kterou si tam uložila webová stránka, kterou navštívil [coo(2015)]. Při prvním vstupu na stránky se cookie vytvoří a zapíše se do ní vygenerovaný řetězec. Při dalším přístupu už se jen ověří, že cookie existuje, a k zápisu dat do databáze se použije identifikační řetězec, který je v ní obsažen.

Když se zákazník zaregistruje, vytvoří se mu záznam s více informacemi, jako je jeho jméno, příjmení, e-mailová adresa a opět identifikační číslo. Před registrací už ale měl také vytvořenou cookie, můžeme ho tedy nyní identifikovat dvojnásobem. To nám pomůže např. když se zaregistrovaný zákazník při vstupu na stránky nepřihlásí – stále víme o koho se jedná a můžeme mu posílat personalizovaný newsletter, což můžeme dělat jen u registrovaných zákazníků, u kterých známe e-mailovou adresu.

Dále se může stát, že se zákazník přihlásí na jiném zařízení či prohlížeči. Pak se mu vytvoří nová cookie s novým identifikátorem. Díky jeho přihlášení ovšem víme, že se opět jedná o našeho zákazníka.

Do databáze pak ukládáme čas vstupu a opuštění jednotlivých stránek, díky čemuž jsme pak schopni určit, jaké kategorie navštívuje nejvíce a kde tráví nejvíce času.

Výše uvedené možnosti teď můžeme pro jasnost zjednodušeně vypsát

takto:

- přihlášený – známe u něj jeden či více cookie identifikátorů,
- nepřihlášený
 - neregistrovaný – vytvoříme či máme cookie identifikátor,
 - registrovaný – máme cookie identifikátor spojený se zákazníkem.

3 Prestashop

Jelikož cílem této práce není implementace celého internetového obchodu ale jen jeho specifických částí, použijeme jako základ Prestashop



(viz [Borderie(2014)]), což je zdarma stáhnutelný software, využívající architekturu podobnou Model-View-Controller. Je používán pro množství internetových obchodů, jelikož obsahuje mnoho funkcí a zároveň umožňuje obrovské možnosti přizpůsobení. Mezi tyto patří:

- šablony,
- moduly,
- přepisování.

Šablony umožňují vytvoření vzhledu, který je pak možno aplikovat na jakýkoliv obchod (stejně jako další dvě možnosti).

Moduly jsou prostředkem použitým v implementaci této bakalářské práce. Každý modul může být v administraci jednoduše nainstalován, přičemž oficiální stránky Prestashopu nabízejí velké množství modulů, které jsou zdarma ke stažení nebo ke koupi.

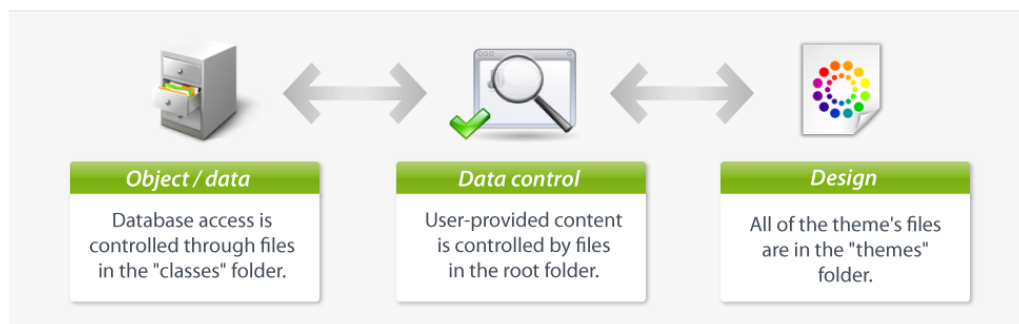
Přepisování je poslední možnost, kdy se již existující třídy či metody přepíše jinými a tím se změní jejich původní chování.

3.1 Architektura Prestashopu

Jak již bylo řečeno, Prestashop je založen na třívrstvé architektuře.

- Objekt/Data – Přístup do databáze skrz soubory ve složce *classes*.
- Ovládání dat – Obsah poskytovaný uživateli je ovládán soubory v kořenovém adresáři.
- Design – Všechny soubory šablony jsou ve složce *themes* [Borderie(2014)].

PrestaShop's 3-tier architecture



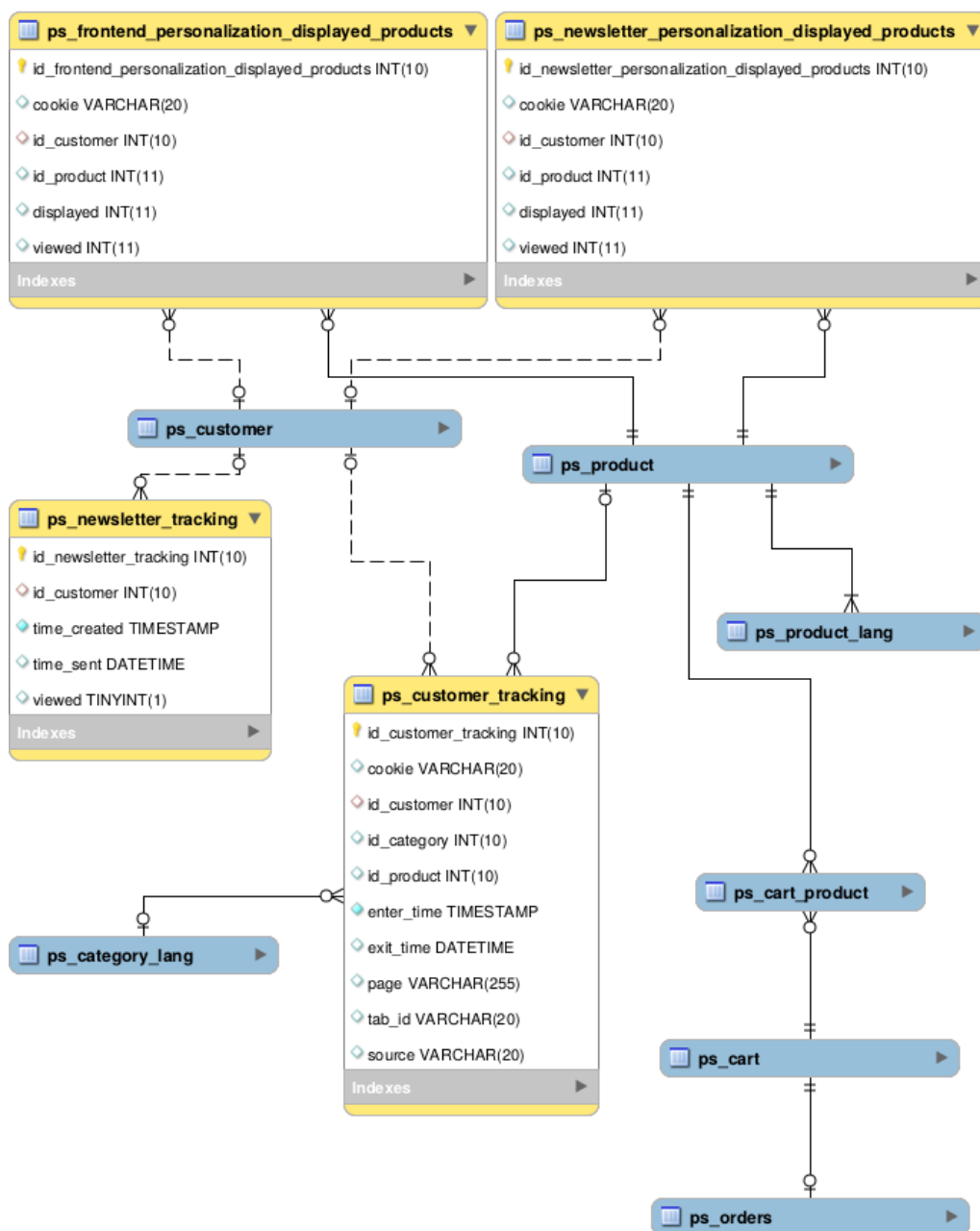
Obrázek 3.1: Architektura Prestashopu

Tvůrci Prestashopu se rozhodli nepoužít PHP framework jako Zend Framework, Symfony nebo CakePHP, aby umožnili lepší přehlednost a tak rychlejší editaci. Pomáhá to také lepšímu výkonu, jelikož je software tvořen pouze řádky kódu, které potřebuje a neobsahuje spousty obecných knihoven navíc [Borderie(2014)].

Moduly, které programátoři vytvoří, se vloží do adresáře *modules*. V administraci je pak možné je nainstalovat – samotné nakopírování nestačí. Moduly pak mohou mít svou konfiguraci a také zobrazovat určitý obsah ve frontendu (zákazníky viditelná část stránek) či backendu (administraci).

Níže je velmi zjednodušený model databáze (Obr. 3.2) obsahující tabulky, které byly použity v tomto projektu. Žluté tabulky byly vytvořeny v rámci tohoto projektu, modré tabulky jsou obsaženy v Prestashopu samotném. Model se všemi sloupci těchto tabulek a model základní instalace Prestashopu naleznete na příloženém CD.

V zobrazeném modelu je vidět, že všechny čtyři vytvořené tabulky jsou navázány na tabulku *ps_customer* přes sloupec *id_customer*. Tabulky, které zaznamenávají počet zobrazení jednotlivých produktů a případné prokliknutí, jsou navázány také na tabulku *ps_product*. Tabulka *ps_customer_tracking* je pak svázána s tabulkami *ps_category_lang* a *ps_product*, jelikož se do ní zaznamenává id kategorie a produktu, které uživatel prohlíží.



Obrázek 3.2: Zjednodušený ERA model, zobrazující použité tabulky

Základní adresářová struktura Prestashopu je následující:

- /admin (jméno je přizpůsobeno při instalaci): obsahuje všechny soubory Prestashopu vztahující se k administraci. Když se k tomuto adresáři přistupuje z prohlížeče, je třeba se z bezpečnostních důvodů přihlásit.
- /cache: obsahuje dočasné složky které jsou vygenerovány a znovu používány, aby se odlehčilo zatížení serveru.
- /classes: obsahuje všechny soubory patřící k objektovému modelu (některé jsou používány pro frontend (stránky viditelné zákazníkem) a některé pro administraci). Každý soubor představuje (a obsahuje) PHP třídu a její metody a proměnné.
- /config: obsahuje všechny konfigurační soubory Prestashopu. Pokud to není nezbytně nutné, neměly by se měnit, jelikož je obstarává instalátor a administrace.
- /controllers: obsahuje všechny soubory patřící ke controllerům Prestashopu – podle softwarové architektury *Model-View-Controller* (nebo *MVC*) používané Prestashopem (popsáno v kapitole *Architektura Prestashopu*). Každý soubor ovládá určitou část Prestashopu.
- /css: obsahuje všechny CSS soubory, které nejsou připojeny k šablonám – jsou tedy většinou používány v administraci.
- /docs: obsahuje dokumentaci, licence a příklady souborů k importu. Pozn.: adresář by měl být vymazán v produkčním prostředí.
- /download: obsahuje virtuální produkty, které mohou být stažené uživatelem, který za ně zaplatil. Soubory jsou uloženy s md5 jménem.
- /img: obsahuje všechny základní obrázky a ikony Prestashopu – tzn. ty které nepatří k šabloně. Zde se nacházejí obrázky kategorií produktů (podadresář /*c*), obrázky produktů (podadresář /*p*) a obrázky pro administraci samotnou (podadresář /*admin*).
- /install: obsahuje všechny instalační soubory související s instalátorem Prestashopu. Je požadováno je smazat po instalaci pro zvýšení bezpečnosti.
- /js: obsahuje všechny JavaScriptové soubory, které nejsou svázány se šablonami. Většina z nich patří k administraci. Zde se také nachází jQuery framework.

- /localization: obsahuje všechny lokalizační soubory – tzn. soubory, které obsahují informace jako je měna, jazyk, daňová pravidla a daňové skupiny, státy a různé jednotky, které se používají v dané zemi (tj. objem v litrech, váha v kilogramech, atd.).
- /log: obsahuje soubory logu generované Prestashopem v různých fázích, např. během procesu instalace.
- /mails: obsahuje všechny HTML a textové soubory, související s e-maily posílanými Prestashopem. Každý jazyk má svoji specifickou složku, kde se může upravovat jejich obsah. Prestashop obsahuje nástroj k upravování e-mailů v administraci pod Lokalizace > Překlady.
- /modules: obsahuje všechny moduly Prestashopu, každý ve svém vlastním adresáři. Pokud má být modul úplně odstraněn, nejprve se musí odinstalovat v administraci a teprve pak se můžou odstranit tyto soubory.
- /override: toto je speciální adresář, který byl zaveden v Prestashopu 1.4. Používáním konvence pojmenovávání souborů je možné vytvořit soubory, které přepisují základní třídy nebo controllery. To umožňuje změnit základní chování Prestashopu, bez úpravy původních souborů, aby se pak systém mohl jednoduše aktualizovat.
- /pdf: obsahuje všechny soubory šablony (.tpl) sloužící ke generování PDF souborů (faktury, doručovací lístky, atd.). Soubory lze upravit pro změnu vzhledu generovaných PDF souborů.
- /themes: obsahuje všechny nainstalované šablony, každou ve svém adresáři.
- /tools: obsahuje externí nástroje, které byly integrovány do Prestashopu. Zde se nachází např. Smarty (*template/theme*), TCPDF (generátor PDF souborů), Swift (posílání e-mailů), PEAR XML Parser (PHP nástroj), atd.
- /translations: obsahuje podadresář pro každý dostupný jazyk. Nicméně pokud je třeba upravit překlad, musí se tak učinit použitím interního nástroje Prestashopu a neupravovat jej přímo zde.
- /upload: obsahuje soubory, které jsou nahrány zákazníky pro přizpůsobitelné produkty (např. obrázek, který chce zákazník vytisknout na hrneček).
- /webservice: obsahuje soubory, které umožňují aplikacím třetí strany přistupovat k Prestashopu skrze jeho API [Borderie(2014)].

3.2 Modul Prestashopu

Jelikož v této práci budou z výše uvedeného používány jen moduly a v rámci nich jen jednoduché šablony, popíšeme si teď modul důkladněji.

Modul Prestashopu se skládá z hlavního PHP souboru, dalších PHP souborů, které jsou potřeba a obrázků a šablon, které jsou třeba k zobrazení dat. Každý modul Prestashopu, který je do obchodu nainstalován může používat jeden nebo více „hooků“. Hook umožňuje připojit kód k aktuálnímu zobrazení stránky.

Moduly jsou ideální způsob jak nechat vývojáře vyjádřit svůj talent a představitivost, jelikož kreativních možností je mnoho. Mohou zobrazovat různý obsah (bloky, text, atd.), provádět různé úlohy (dávkové aktualizace, import, export, atd.), poskytovat rozhraní pro další nástroje a mnoho mnoho dalšího.

Moduly mohou být konfigurovatelné podle potřeby; čím konfigurovatelnější modul je, tím bude lehčí jej používat a uspokojí tak potřeby více uživatelů.

Jedna z hlavních výhod modulu je přidávat funkcionality bez nutnosti úpravy vlastních souborů Prestashopu, a tak umožňuje jednodušší aktualizaci bez nutnosti převedení úprav do souborů jádra. Nikdy by se neměly upravovat soubory jádra při vytváření modulu, i když to někdy může vypadat jako nezbytně nutné [Borderie(2014)].

Nyní se podíváme na strukturu adresáře modulu, přičemž vyjmenujeme jen položky, které se používají v této práci. Celý seznam lze nalézt ve vývojářské dokumentaci Prestashopu ([Borderie(2014)]). První tři soubory jsou povinné:

- *nazevmodulu.php* – hlavní soubor modulu,
- *config.xml* – konfigurační soubor, Prestashop jej vygeneruje při instalaci modulu,
- *logo.png* – logo modulu, zobrazující se v seznamu modulů v administraci,
- *classes* – adresář se třídami objektového modelu,
- *controllers* – controllery, obsahuje podadresáře *admin* a/nebo *front*,

- *views* – obsahuje soubory šablony,
- *css* – CSS soubory,
- *js* – JavaScriptové soubory.

4 Návrh systému

Cílem této práce je navrhnout a implementovat jednoduchý systém, který bude sbírat data o zákazníkovi a použije je k zobrazování personalizované nabídky produktů na stránkách e-shopu a při rozesílání personalizovaného newsletteru. Zákazníkovi budou nabízeny pouze produkty z kategorií, o které se nejvíce zajímal.

Tento systém tedy můžeme rozdělit do tří modulů, které pak budou v Prestashopu implementovány. Správce obchodu si pak může instalovat pouze ty moduly, které bude chtít:

- modul sledování zákazníka,
- modul personalizované nabídky produktů na stránkách,
- modul obsluhující personalizovaný newsletter.

4.1 Modul sledování zákazníka

Tento modul bude umožňovat sledování chování zákazníků a zobrazování statistik. Dalším dvěma modulům bude také poskytovat metody společně pro oba moduly (např. získávání produktů, které se zákazníkovi mohou líbit se používá na stránkách i v newsletteru). Proto pro instalaci dalších dvou modulů je třeba mít nainstalovaný tento a pro odinstalaci tohoto modulu nejprve odinstalovat zbylé dva.

4.1.1 Sbíráání dat

Je třeba si uvědomit, že zákazník může otevřít stránky obchodu ve více panelech, přepínat mezi nimi či prohlížení přerušit jinou činností. Aby bylo sledování zákazníka co nejpřesnější, je nejlepší použít JavaScript, který reaguje na různé změny a neprovádět měření pouze při kliknutí na nějaký odkaz. Proto můžeme vykonat kód když je stránka přenesena do popředí (funkce *focus*) a opět když zákazník klikne na jiné okno či záložku (funkce *blur*) [jq(2015)].

Při těchto akcích si zapíšeme do databáze údaje:

- vstup na stránku (*focus*) – vytvoříme nový záznam v databázi obsahující identifikaci uživatele (identifikátor z cookie, popř. id registrovaného zákazníka), čas vstupu, stránku, na kterou uživatel vstoupil a ze které přišel, id kategorie a produktu a identifikátor záložky,
- opuštění stránky (*blur*) – k existujícímu záznamu připíšeme čas opuštění stránky.

Jak je uvedeno výše, abychom mohli identifikovat jednotlivé záložky, je třeba jim taky přiřadit jednoznačný identifikátor, který se pak zapíše do databáze – ten se vytvoří při načtení stránky.

4.1.2 Statistiky

Správce obchodu bude moci v administraci sledovat návštěvnost jednotlivých kategorií jak obecně tak i u jednotlivých uživatelů.

Pro zobrazení individuálních statistik bude třeba z nasbíraných dat identifikovat zákazníka, kterého můžeme poznat podle identifikátoru cookie a podle id zákazníka. Je třeba mít na paměti, že jeden zákazník může používat více zařízení a jedno zařízení může být používáno více zákazníky. Navíc může uživatel používat zařízení i jako nepřihlášený. Proto pokud budeme chtít statistiky o uživateli se známým id, můžeme do statistik započítat i záznamy se stejným cookie identifikátorem, které u sebe nemají uvedeno id zákazníka, když nemají nikde jinde id jiného zákazníka. To znamená, že se jedná o zařízení, které pravděpodobně používá jeden zákazník, který je někdy přihlášený a jindy ne. Pokud se na daném zařízení přihlašuje více uživatelů (existují záznamy se stejným cookie identifikátorem a různými id uživatelů), pak už u záznamů bez id uživatele nemůžeme určit, kterému patří.

Může se stát, že zařízení je používáno různými uživateli, přičemž pouze jeden se někdy přihlásil a při získání statistik pak budeme uvažovat všechny záznamy jako jeho, ale sledování zákazníka samozřejmě není stoprocentní.

4.2 Modul personalizované nabídky produktů na stránkách

Tento modul umožňuje zobrazení produktů v levém nebo pravém sloupci na stránkách obchodu. Kde a jestli vůbec se tyto sloupce zobrazují záleží na samotném nastavení Prestashopu. Většinou se ovšem samozřejmě zobrazuje jen jeden sloupec nebo minimálně daný modul jen v jednom z nich.

4.2.1 Nabízení produktů zákazníkovi

Tento bod platí pro oba moduly nabízející produkty.

Výběr produktů pro zobrazení zákazníkovi může být prováděn různými způsoby, jelikož roli může hrát např. čas strávený zákazníkem v kategorii, počet přístupů do kategorie a další. Pokud použijeme pouze čas strávený v kategorii, může se stát, že oblíbenou kategorii určíme chybně, neboť zákazník může stránku zobrazit a pak na určitou dobu od počítače odejít. V tom případě nepoznáme, jestli ji doopravdy četl.

Počet přístupů do kategorie zase signalizuje, že zákazník je u počítače doopravdy přítomen, i když se může stát, že stránky jen rychle prohlíží.

Další možností je nabízet zákazníkovi produkty, které vložil do nákupního košíku, ale nikdy je nekoupil.

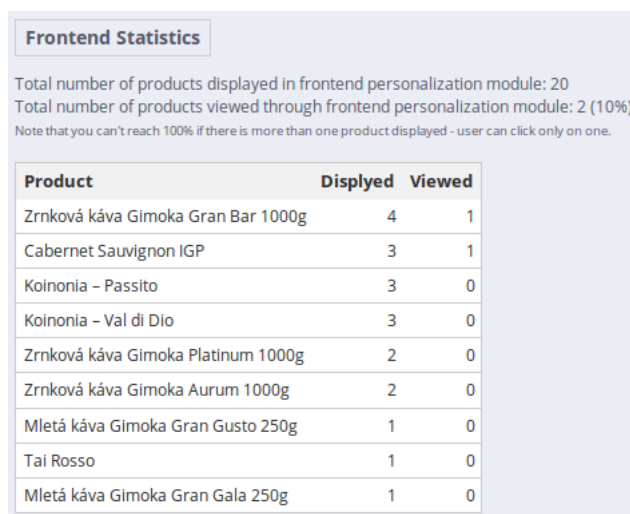
Jelikož je možností více, nabízí se přenechat výběr částečně na administrátorovi stránek. Výběr se udělá náhodně, ovšem každá z uvedených možností bude mít jinou pravděpodobnost výběru. Administrátor bude moci určit, která z možností bude mít pravděpodobnost největší a která nejmenší seřazením možností pod sebe.

Dále také bude moci určit, kolik produktů se má zobrazovat. Tato nastavení se provádějí vždy v konfiguraci v daném modulu.

4.2.2 Statistiky proklikovosti nabízených produktů

V administraci se také budou zobrazovat statistiky, které ukazují kolik produktů bylo zákazníkovi nabídnuto a na kolik z nich ve skutečnosti kliknul. U jednotlivých zákazníků pak bude i tabulka s výčtem těchto produktů

a počty zobrazení a kliknutí.



Frontend Statistics

Total number of products displayed in frontend personalization module: 20
Total number of products viewed through frontend personalization module: 2 (10%)
Note that you can't reach 100% if there is more than one product displayed - user can click only on one.

Product	Displyed	Viewed
Zrnková káva Gimoka Gran Bar 1000g	4	1
Cabernet Sauvignon IGP	3	1
Koinonia - Passito	3	0
Koinonia - Val di Dio	3	0
Zrnková káva Gimoka Platinum 1000g	2	0
Zrnková káva Gimoka Aurum 1000g	2	0
Mletá káva Gimoka Gran Gusto 250g	1	0
Tai Rosso	1	0
Mletá káva Gimoka Gran Gala 250g	1	0

Obrázek 4.1: Statistika zobrazení a kliknutí u jednotlivých uživatelů

4.3 Modul obsluhující personalizovaný newsletter

Administrátor může v tomto modulu rozposílat newslettery, sledovat jejich doručitelnost (jestli byl e-mail doručen či nikoliv) a sledovat, jestli uživatel e-mail zobrazil (resp. jestli zobrazil i vzdálený obsah (obrázky) nebo měl místo obrázků jen prázdná místa).

Na rozdíl od předchozího modulu je tento použitelný pouze pro registrované zákazníky, u kterých známe e-mailovou adresu.

4.3.1 Posílání newsletterů

Pro předvedení této práce bude rozesílání uděláno velmi jednoduše – po stisku tlačítka se e-mail najednou pošle všem zákazníkům, kteří mají povoleno zasílání newsletteru. Pro další vylepšení by se například mohlo udělat plánované rozposlání podle času, kdy zákazník chodí na stránky, ale to už přesahuje rámec této práce.

Pro každého zákazníka se tedy vygeneruje obsah e-mailu a odešle se. Po odeslání všech e-mailů se pak podle odpovědi serveru v maillogu pro

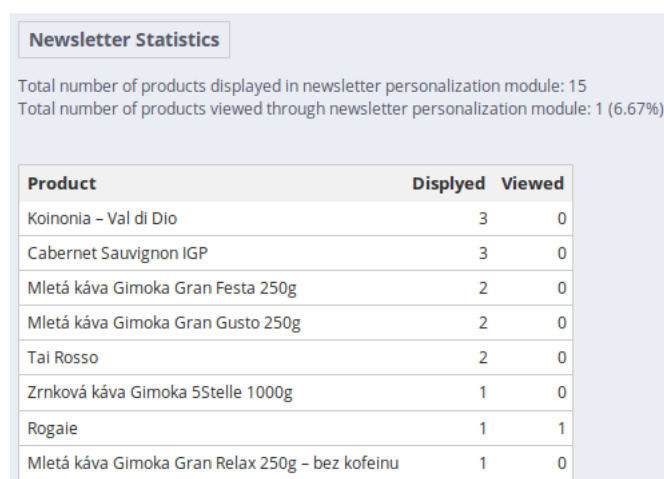
každý e-mail určí, jestli byl úspěšně doručen. Maillog zaznamenává čas, kdy došlo k odeslání e-mailu a následně odpověď poštovního serveru. Na stránce s přehledem odeslaných e-mailů se u úspěšně doručených objeví datum a čas odeslání. Opět by rozšířením práce mohla být identifikace chyby při posílání a následné vyhodnocení – jestli schránka neexistuje vyřadit ji ze seznamu, pokud má např. uživatel pouze plnou schránku zkusit poslat e-mail později znovu apod.

4.3.2 Sledování zobrazení vzdáleného obsahu

Pro zjištění jestli byl e-mail otevřen do něj zakomponujeme tzv. *měřící pixel*. Jedná se o fiktivní obrázek, na který odkaz v e-mailu ukazuje. Server ovšem požadavek přeměruje na PHP skript, který do databáze uloží shlédnutí e-mailu a následně vygeneruje onen požadovaný obrázek, který vrátí klientovi k zobrazení [pix(2015)].

4.3.3 Statistiky proklikovosti nabízených produktů

Stejně jako v předchozím modulu i zde budou v administraci statistiky proklikovosti produktů.



The screenshot shows a 'Newsletter Statistics' section. It includes two summary lines: 'Total number of products displayed in newsletter personalization module: 15' and 'Total number of products viewed through newsletter personalization module: 1 (6.67%)'. Below this is a table with three columns: 'Product', 'Displyed', and 'Viewed'. The table lists eight products with their respective counts.

Product	Displyed	Viewed
Koinonia - Val di Dio	3	0
Cabernet Sauvignon IGP	3	0
Mletá káva Gimoka Gran Festa 250g	2	0
Mletá káva Gimoka Gran Gusto 250g	2	0
Tai Rosso	2	0
Zrnková káva Gimoka 5Stelle 1000g	1	0
Rogaie	1	1
Mletá káva Gimoka Gran Relax 250g - bez kofeinu	1	0

Obrázek 4.2: Statistika zobrazení a kliknutí u jednotlivých uživatelů

5 Popis implementace

5.1 Modul sledování zákazníka

Tento modul je nazván *customertracking* a kromě základního souboru a několika méně podstatných souborů obsahuje tyto důležité prvky:

- *Chart* – externí knihovna pro vytváření grafů.
- *Classes* – adresář se třídami obsahuje dvě třídy. Jedna slouží pro tento modul (*CustomerTracking.php*) a druhá poskytuje služby, které jsou společné i s ostatními moduly (*CustomerTrackingModules.php*).
- *Controllers* – modul obsahuje jeden controller *AdminCustomerTrackingController.php*, který obsluhuje zobrazování statistik v administraci.
- *Js* – adresář obsahuje soubor *customertracking.js*, který slouží k sledování zákazníka.
- *Views* – zde jsou umístěny jednotlivé soubory šablony, které slouží k zobrazení statistik (adresář *templates/admin*) a soubor *templates/hook/customertracking.tpl*, který obsahuje některé důležité proměnné používané v JavaScriptu pro sledování zákazníka.

5.1.1 Sbíráání dat

Jak již bylo naznačeno, data se budou zaznamenávat pomocí JavaScriptu (knihovny jQuery) použitím funkcí `$(window).focus()` a `$(window).blur()`. V těchto funkcích se vyvolá pomocí AJAXu PHP skript *ajax.php* z adresáře modulu, který do databáze uloží příslušná data získaná z proměnných nastavených při načtení stránky šablonou *templates/hook/customertracking.tpl*.

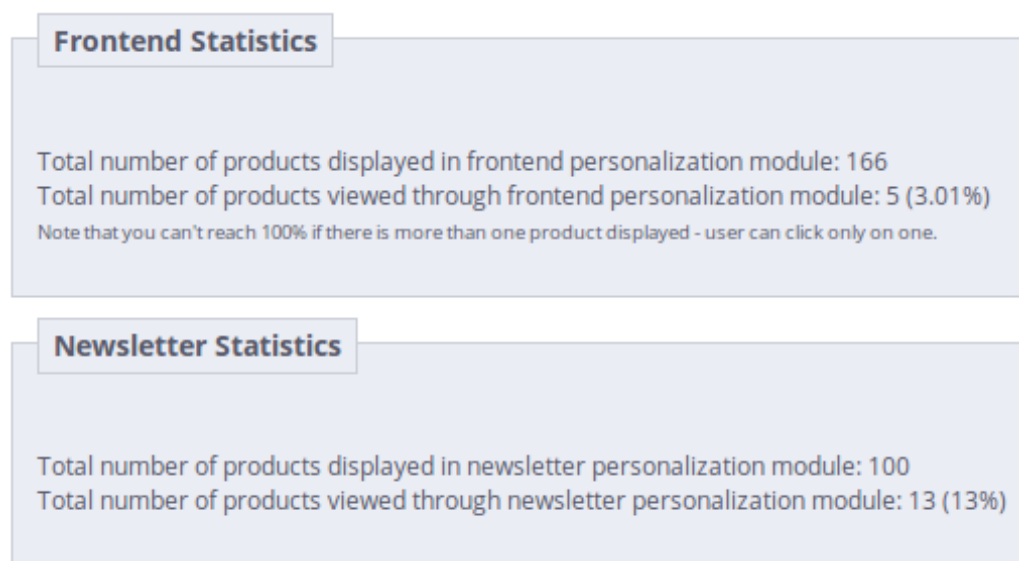
Jelikož může zákazník otevřít stránky ve více záložkách, potřebujeme identifikovat každou zvlášť. Některé prohlížeče umožňují dostat jakési id záložky, ale jelikož se toto liší mezi různými prohlížeči, není to dobré řešení. Vygenerujeme si proto při načtení stránky vlastní id záložky, které uložíme spolu s ostatními proměnnými, jako je např. identifikátor z cookie, id zákazníka, kategorie a produktu apod.

5.1.2 Statistiky

Pro zobrazování statistik v administraci potřebujeme vytvořit záložku s naším modulem. Ta se vytvoří při instalaci modulu a povede na akce zprostředkovávané controllerem.

Základní akce je *renderList()*, která zobrazí seznam uživatelů s časem, který na stránkách strávili a počtem přístupů, který udělali.

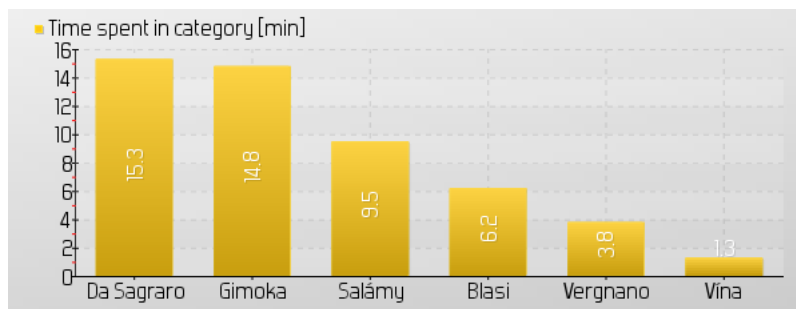
Pokud jsou nainstalovány i další dva moduly, zobrazuje také statistiky kliknutí na nabízený produkt. Je zde uveden počet nabídnutých produktů a počet, na které z nich zákazník klikl. Toto je rozděleno zvláště na modul personalizované nabídky produktů na stránkách a modul newsletteru, přičemž v prvním případě v podstatě nelze docílit stoprocentní úspěšnosti kliknutí na nabízený produkt, jelikož se většinou nabízí více produktů, ale zákazník může kliknout pouze na jeden.



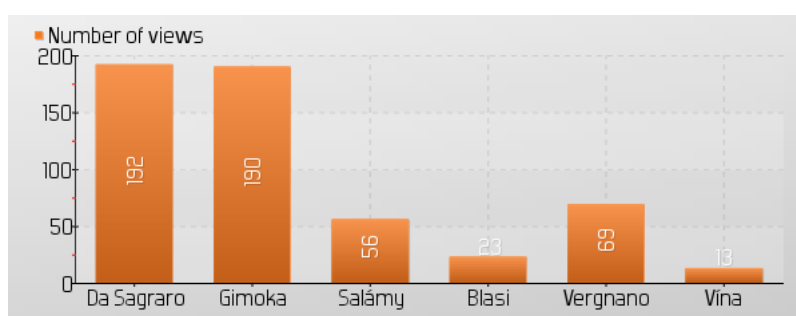
Obrázek 5.1: Statistiky zobrazení a kliknutí

Kromě tabulky zobrazí také grafy, kde jsou vidět nejnavštěvovanější kategorie.

U každého zákazníka je pak možné zobrazit zvláště jeho statistiky. Toto zajišťuje akce *renderView()*. Z těchto statistik také vychází výběr produktů v ostatních modulech. Zobrazují se podobné informace jako u souhrnných statistik, časy a počty přístupů jsou však rozděleny na jednotlivé kategorie a u statistik proklikovosti nabízených produktů se navíc zobrazuje počet zobrazení a kliknutí u jednotlivých produktů.



Obrázek 5.2: Graf nejnavštěvovanějších kategorií podle času



Obrázek 5.3: Graf nejnavštěvovanějších kategorií podle počtu přístupů

5.2 Modul personalizované nabídky produktů na stránkách

Soubory tohoto modulu se nachází v adresáři *frontendpersonalization*. Tento modul nemá svoji záložku v administraci, obsahuje tedy méně souborů v porovnání s ostatními.

- *Classes* – v adresáři je třída *FrontendPersonalizationTrackingModel* komunikující s databází.
- *Views* – adresář *templates/admin* obsahuje šablony použité v administraci pro zobrazení statistik v předchozím modulu (*overall_statistics.tpl* a *statistics.tpl*) a šablonu použitou v konfiguraci modulu pro řazení pravděpodobností jednotlivých možností, podle jakých se vyberou produkty (*priorities.tpl*). Ve druhém adresáři *templates/hooks* se nachází šablona *frontendpersonalization.tpl* vykreslující samotné produkty vybrané pro zobrazení v levém nebo pravém sloupci na stránkách.

5.2.1 Nabízení produktů zákazníkovi

Pro výběr produktů, které se mají zobrazit zákazníkovi, se používá společná třída *CustomerTrackingModulesModel* z modulu pro sledování zákazníka, jelikož stejný postup se aplikuje i pro modul newsletteru. Pro výběr produktů se zavolá metoda *getPersonalizedProductIds(\$cookie, \$id_customer, \$id_lang, \$module, \$current_id_category = NULL)*, jejíž parametry jsou:

- identifikátor z cookie,
- id zákazníka,
- id jazyka (obchod může fungovat multilinguálně),
- modul je jedna z konstant definovaných v této třídě a udává, jestli se jedná o modul obsluhující frontend nebo o modul newsletteru,
- id kategorie, ve které se zákazník aktuálně nachází, aby se mu nenabízely produkty z této kategorie dvakrát.

Tato metoda vrátí pouhá id vygenerovaných produktů, které pak tvoří parametr metody *getPersonalizedProductsInfo(\$productsViewed, \$id_lang)*, která vrátí všechny informace o vybraných produktech, které jsou potřeba pro jejich vykreslení.

Vygenerování produktů probíhá následujícím způsobem:

1. Z konfigurace se načte počet produktů, které se mají vygenerovat a priority metod, jakými se mají generovat.
2. Náhodně se vybere metoda která se má použít, podle intervalu kam padne náhodně vygenerované číslo. Intervaly jsou různě velké podle priorit metod (vážený průměr). V této práci byly použity tři různé metody, ovšem jednoduše by se daly implementovat další. Pro tři metody tedy intervaly vypadají následovně:

$$\left\langle 0 - \frac{1}{6} \right\rangle, \left\langle \frac{1}{6} - \frac{3}{6} \right\rangle, \left\langle \frac{3}{6} - 1 \right\rangle$$

Největší pravděpodobnost výběru má samozřejmě poslední interval a postupně pravděpodobnosti klesají.

3. Pro metodu která byla vybrána se získají produkty. Pokud počet produktů po tomto kroku neodpovídá požadovanému, pokračuje se zbylými metodami podle jejich priorit.
4. Do databáze se uloží shlédnutí jednotlivých produktů.

Vygenerování produktů probíhá podle těchto metod:

- nejdéle strávený čas v kategorii – z databáze se načtou kategorie seřazené podle času v nich stráveném, postupně se prochází a náhodně se z nich vybírají produkty,
- největší počet přístupů do kategorie – obdobně jako předchozí,
- produkty, které byly v nákupním koši, ale nebyly zakoupeny – z tabulky *cart_product*.

5.2.2 Sledování kliknutí na nabízený produkt

Při personalizovaném zobrazení produktů zákazníkovi si do tabulky *ps_frontend_personalization_displayed_products* uložíme cookie identifikátor, id zákazníka, pokud je zákazník přihlášen, a id produktu, který mu nabízíme. Pokud už takový záznam v databázi je, přičteme pouze jedničku ve sloupci počet nabídnutí.

Abychom mohli sledovat, jestli zákazník na nějaký nabízený produkt kliknul, přidáme si za odkaz na produkt speciální parametr, který nám umožní zapsat tuto informaci do databáze. Jelikož používáme dva moduly, u kterých lze toto sledovat, bude tento parametr určovat, o který modul se jedná (*source*). Při načtení stránky s produktem pak, pokud je tento parametr nastaven na *front*, záznamu v databázi s tímto produktem a identifikací uživatele navýšíme počet kliknutí.

5.3 Modul obsluhující personalizovaný newsletter

Modul je umístěn v adresáři *newsletterpersonalization*, kde mezi nejdůležitější patří:

- *Classes* – v adresáři jsou dvě třídy komunikující s databází. Třída *NewsletterPersonalizationTrackingModel* se podobá třídě z předchozího modulu – obstarává ukládání zobrazení produktů a získání statistik zobrazených a prokliknutých produktů. Druhá třída *NewsletterPersonalizationModel* obstarává samotné posílání e-mailů a ukládání informací spojených s jejich odesláním.
- *Controllers* – modul obsahuje controller *AdminNewsletterPersonalizationController*, který zajišťuje možnost odesílat newslettery a zobrazovat informace o odeslaných e-mailech.
- *Views* – v adresáři *templates/email* je šablona *content.tpl*, která se používá k zobrazení obsahu newsletteru.
- *emailviewed.php* – skript zajišťující uložení shlédnutí newsletteru do databáze.

5.3.1 Posílání newsletterů

Z administrace se zavolá metoda *sendEmails()*, která pošle e-mail každému zákazníkovi, který má povoleno zasílání newsletteru. Každému newsletteru je pak zvlášť vygenerován personalizovaný obsah stejnou metodou jako v předchozím modulu a odeslán. Následně se pro každý z odeslaných e-mailů nalezne odpověď poštovního serveru v maillogu. V této práci je pouze zaznamenáváno úspěšné odeslání, hledáme tedy kód 250 [smt(2013)]. Pro rozšíření práce by se mohly také rozlišit kódy jednotlivých chyb.

5.3.2 Sledování kliknutí na nabízený produkt

V tomto modulu je v odkazu na produkt kromě parametru *source* s hodnotou *newsletter* ještě uveden parametr id zákazníka *url_id* a parametr *hash*, ve kterém je toto id spolu s tajným řetězcem zahashováno, abychom mohli zkontrolovat, že zákazník nezměnil id v odkazu. Ostatní postup zůstává stejný jako v předchozím modulu, jen se tyto statistiky ukládají do tabulky *ps_newsletter_personalization_displayed_products*.

5.3.3 Sledování zobrazení vzdáleného obsahu

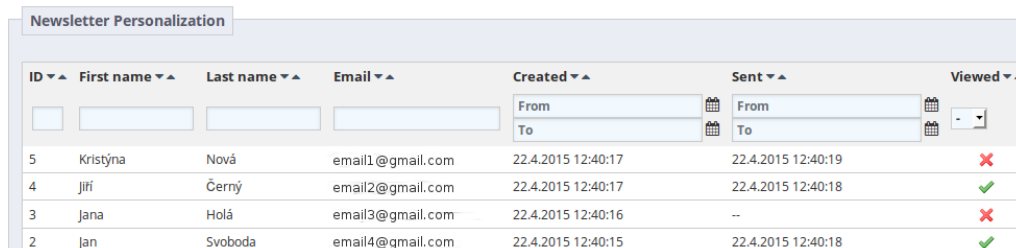
Abychom mohli sledovat, jestli byl uživatelem zobrazen vzdálený obsah, vytvoříme tzv. měřicí pixel, jak již bylo řečeno výše. Jedná se o obrázek velmi malých rozměrů, obvykle 1x1 pixel, ale jelikož některé softwary již toto blokují, je lepší použít obrázek o velikosti např. 3x3 pixely.

Odkaz na tento obrázek obsahuje také parametr *id*, s pomocí kterého identifikujeme záznam odeslaného e-mailu v tabulce *ps_newsletter_tracking*. V kořenovém adresáři na serveru ovšem v souboru *.htaccess* vytvoříme řádek, který bude vypadat nějak takto:

```
Redirect url_obrazku url_souboru
```

a který zajistí přsměrování na PHP skript, který uloží informaci o načtení obsahu u daného záznamu do databáze a vygeneruje průhledný obrázek, který vrátí e-mailovému klientovi k zobrazení.

Na následujícím obrázku můžete vidět seznam odeslaných newsletterů, kde třetí e-mail pro Janu Holou nemá vyplněn čas odeslání *Sent* – tzn. že nebyl doručen. U druhého a čtvrtého e-mailu pak došlo k zobrazení vzdáleného obsahu (sloupeček *Viewed*).



ID	First name	Last name	Email	Created	Sent	Viewed
5	Kristýna	Nová	email1@gmail.com	22.4.2015 12:40:17	22.4.2015 12:40:19	✘
4	Jiří	Černý	email2@gmail.com	22.4.2015 12:40:17	22.4.2015 12:40:18	✔
3	Jana	Holá	email3@gmail.com	22.4.2015 12:40:16	--	✘
2	Jan	Svoboda	email4@gmail.com	22.4.2015 12:40:15	22.4.2015 12:40:18	✔

Obrázek 5.4: Přehled odeslaných newsletterů

6 Testování a zhodnocení výsledků

Testování bylo provedeno na několika uživateli, kteří potvrdili nabídku personalizovaných produktů jako relevantní prohlíženému obsahu stránek. Nyní si ukážeme příklad výběru produktů u vybraného uživatele.

Pro nabídku produktů na stránkách i do newsletteru byla konfigurace nastavena tak, jak můžete vidět na obrázku 6.1, s výjimkou počtu vybraných produktů do newsletteru, kde jich bylo pět.

MODULES / FRONTENDPERSONALIZATION / CONFIGURE

Configure "Front-end Personalization" module

Back Translate Check u Manage

FRONT-END PERSONALIZATION

* Number of products to be proposed to customer

4

* Number of products per category to be proposed to customer

1

Save

SET PRIORITIES FOR CHOOSING PERSONALIZED PRODUCTS

Sorted by probability of being selected from highest (drag to set priority)

Longest time spent in a category

Most access to a category

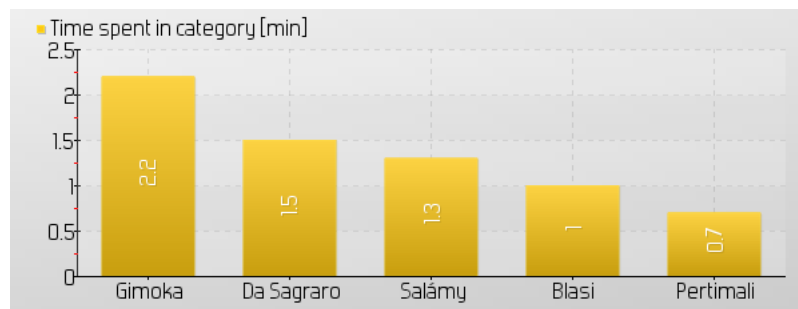
Products that were in the abandoned cart

Obrázek 6.1: Konfigurace

Jelikož má největší prioritu metoda řazení kategorií podle času v nich stráveném, podíváme se na graf (Obr. 6.2) zobrazující statistiky pro tohoto uživatele.

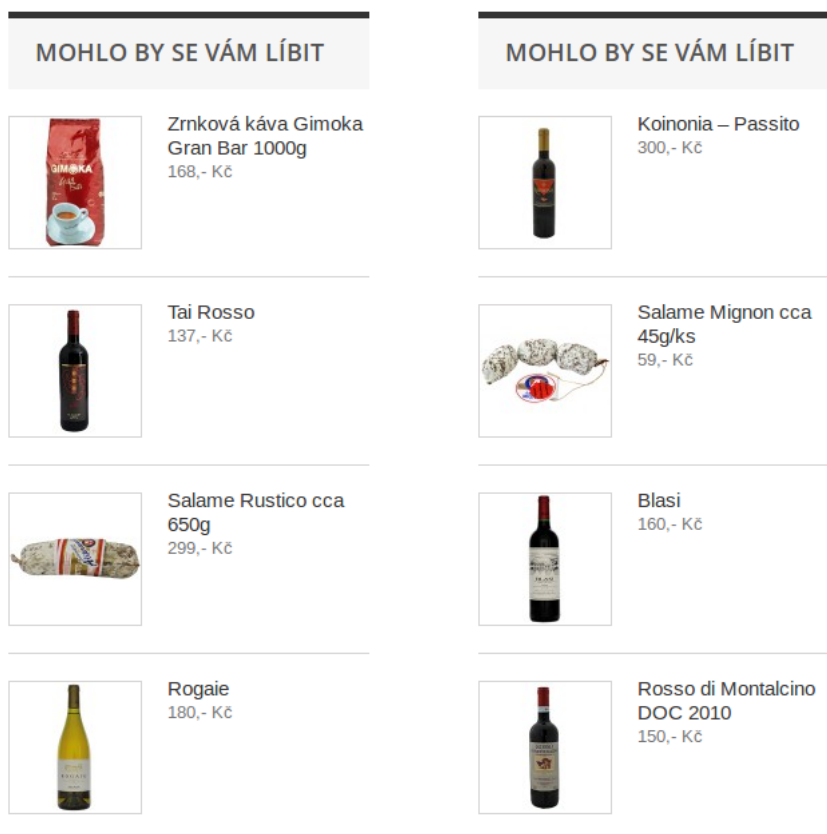
Pro lepší přehlednost se zmíníme, že *Gimoka* je káva a kategorie *Da Sagraro*, *Blasi* a *Pertimali* jsou vína.

Nyní můžeme porovnat nabídku zobrazenou v kategorii *Vergnano* (Obr.



Obrázek 6.2: Graf nejnavštěvovanějších kategorií podle času

6.3a), která ve výpisu není a tudíž se produkty zobrazují v daném pořadí, z každé kategorie vždy jeden (podle konfigurace) a nabídku zobrazenou v kategorii *Gimoka* (Obr. 6.3b), kde můžeme vidět, že první produkt nabídky je z kategorie *Da Segraro*, která je ve statistikách až druhá. Kategorie *Gimoka* byla vynechána, aby se produkty nezobrazovaly dvakrát.




(a) V kategorii Vergnano

(b) V kategorii Gimoka






Obrázek 6.3: Porovnání zobrazení nabídky produktů ve dvou kategoriích

Na závěr si ukážeme newsletter (Obr. 6.4), který byl odeslán stejnému zákazníkovi, kde opět můžeme vidět, že výběr produktů proběhl podle kategorií vypsanych výše.



DOBRÝ DEN JIŘÍ ČERNÝ,

posíláme Vám nějaké tipy na zboží, které by se Vám mohlo líbit:

	Mletá káva Gimoka Gran Gala 250g 70,- Kč
	Cabernet Sauvignon IGP 137,- Kč
	Salame Mignon cca 45g/ks 59,- Kč
	Blasi 160,- Kč
	Brunello di Montacino DOCG 2007 160,- Kč

Obchod [Aggeo](#) funguje v systému [PrestaShop™](#)

Obrázek 6.4: Ukázka newsletteru

Otestovat úspěšnost personalizace sledováním zvýšení prodeje bohužel není možné, jelikož se jedná pouze o obchod vytvořený pro potřeby této práce, a který tedy není (alespoň v době vydání práce) v provozu.

7 Závěr

Tato práce je spíše jakýmsi úvodem do problematiky personalizace internetových aplikací. Implementovaný systém umožňuje vyzkoušet některé základní principy, nicméně poskytuje široký rozsah možností pro další rozšiřování.

Po otestování jednotlivých modulů můžeme vidět, že splňují požadavky kladené na tuto práci. Chování zákazníka je úspěšně sledováno a na jeho základě je pak prováděn výběr produktů, které se mu nabízí. Princip tohoto výběru se ukázal taktéž úspěšný a u newsletterů, které jsou odesílány, se správně zobrazuje jejich doručení a následné zobrazení.

Testování rozesílání newsletterů nebylo provedeno na větším objemu zákazníků (stovky až tisíce), kde by pravděpodobně začalo docházet k problémům s přijetím zpráv u jednotlivých poštovních serverů, které by je mohly pokládat za spam. To je ale nicméně další z možností rozšíření této práce.

Literatura

- [coo(2015)] *Práce s cookies* [online]. 2015. [cit. 10. 4. 2015]. Dostupné z: <<https://support.mozilla.org/cs/kb/Pr%C3%A1ce%20s%20cookies>>.
- [jqu(2015)] *jQuery API* [online]. 2015. [cit. 19. 4. 2015]. Dostupné z: <<https://api.jquery.com>>.
- [pix(2015)] *Pixel tracking in third-party and custom creatives* [online]. 2015. [cit. 19. 4. 2015]. Dostupné z: <https://support.google.com/dfp_premium/answer/1347585>.
- [smt(2013)] *SMTP errors and reply codes* [online]. 2013. [cit. 22. 4. 2015]. Dostupné z: <<http://www.serversmtp.com/en/smtp-error>>.
- [Borderie(2014)] BORDERIE, X. *PrestaShop 1.6 Developer Guide* [online]. 2014. [cit. 15. 4. 2015]. Dostupné z: <<http://doc.prestashop.com/display/PS16/Developer+Guide>>.
- [Quah et al.(2008)] QUAH, J. T. S. – LEOW, W. C. H. – YONG, K. L. Auto-Personalization Web Pages. In ANG, C. S. – ZAPHIRIS, P. (Ed.) *Human Computer Interaction: Concepts, Methodologies, Tools, and Applications*, s. 807–815. Hershey: IGI Global, 2008.