

Analýza návštěvnosti z vyhledávačů

Ondrej Podolinský

Bakalářská práce
2006



Univerzita Tomáše Bati ve Zlíně
Fakulta aplikované informatiky

Univerzita Tomáše Bati ve Zlíně
Fakulta aplikované informatiky
Ústav aplikované informatiky
akademický rok: 2005/2006

ZADÁNÍ BAKALÁŘSKÉ PRÁCE

(PROJEKTU, UMĚLECKÉHO DÍLA, UMĚLECKÉHO VÝKONU)

Jméno a příjmení: **Ondrej PODOLINSKÝ**
Studijní program: **B 3902 Inženýrská informatika**
Studijní obor: **Informační technologie**

Téma práce: **Analýza návštěvnosti z internetových vyhledávačů**

Zásady pro vypracování:

Provedte analýzu internetových vyhledávačů.

Vytvořte třídu v PHP 5, která umožní shromažďování potřebných dat a jejich uložení do databáze MySQL.

Získané výsledky prezentujte graficky pomocí PHP 5, XHTML a CSS, použijte šablony.

Vytvořená aplikace bude testována na firemní prezentaci firmy ZETR SK

<http://www.zetr.sk>

Rozsah práce:

Rozsah příloh:

Forma zpracování bakalářské práce: **tištěná/elektronická**

Seznam odborné literatury:

GILMORE, W. Jason. Velká kniha PHP5 a MySQL : kompendium znalostí pro začátečníky i profesionály. RNDr. Jan Pokorný. 1. vyd. Brno : Zoner Press, 2005. 711 s. ISBN 80-86815-20-X.

SCHLOSSNAGLE, G. Pokročilé programování v PHP 5. ZonerPress, 2005. ISBN 80-86815-14-5

Prokop, M.: CSS – kaskádové styly pro webdesignery. Mobil Media, 2003.

Maslakowski, M.: Naučte se MySQL za 21 dní. ComputerPress, 2001.

Vedoucí bakalářské práce:

Ing. Martin Sysel, Ph.D.

Ústav aplikované informatiky

Datum zadání bakalářské práce:

14. února 2006

Termín odevzdání bakalářské práce:

16. června 2006

Ve Zlíně dne 14. února 2006

prof. Ing. Vladimír Vašek, CSc.
pověřený děkan



doc. Ing. Ivan Zelinka, Ph.D.
ředitel ústavu

ABSTRAKT

Obsahem této práce je souhrn základních principů optimalizace internetových prezentací pro vyhledávače a potažmo i uživatele, kteří se na stránky z vyhledávačů dostanou.

Praktická část je tvořena popisem aplikace sloužící pro analýzu návštěvnosti z těchto vyhledávačů, zahrnující souhrny za určitý počet dnů, statistiky návštěvnosti z jednotlivých vyhledávačů v grafické a číselné podobě, informace o návštěvnosti na jednotlivé dotazy a návštěvnost z externích serverů, které nejsou identifikovány jako známé vyhledávače.

Klíčová slova: seo, optimalizace pro vyhledávače, analýza návštěvnosti, smarty

ABSTRACT

The content of this work is a summary of basic principles of optimization of internet presentations for internet search engines and also for users, which get on the pages due to these search engines.

The practical part is made up by the description of the application, which serves for analysis of the visit rate from these search engines, including summaries for a specific count of days, statistics of the visit rate from individual search engines both in graphical and numerical version, information of the visit rate for individual requests and the visit rate from external servers, which are not identified as known search engines.

Keywords: seo, search engine optimization, analysis of the visit rate, smarty

Poděkování:

Děkuji tímto svému vedoucímu bakalářské práce, panu Ing. Martinovi Syslovi, Ph.D za odborné vedení a věcné připomínky a rodině a všem přátelům za morální podporu a připomínky k aplikaci.

Motto:

Člověk je nedokonalý, ale dokonale si s tím poradí. (Leszek Kumor)

Prohlašuji, že jsem na celé bakalářské práci pracoval samostatně a použitou literaturu jsem citoval.

V Uh. Hradišti, 15.6.2006

.....

Ondrej Podolinský

OBSAH

OBSAH	7
ÚVOD	9
I TEORETICKÁ ČÁST	10
1 OPTIMALIZACE PRO VYHLEDÁVAČE	11
1.1 FAKTORY OVLIVŇUJÍCÍ POZICE VE VYHLEDÁVAČÍCH.....	11
1.1.1 Obsah klíčových slov v titulu stránky <title>	12
1.1.2 Obsah klíčových slov v url.....	12
1.1.3 Obsah klíčových slov na stránce	12
1.1.4 Struktura stránek	12
1.1.5 Zpětné odkazy	12
1.2 FAKTORY OVLIVŇUJÍCÍ CHOVÁNÍ ENGINE VYHLEDÁVAČŮ.....	13
1.2.1 Frekvence aktualizací.....	13
1.2.2 Jazyk použitý na stránkách.....	13
1.2.3 Podvodné techniky	13
1.2.4 Použití souboru robots.txt	14
II PRAKTICKÁ ČÁST	15
2 NÁVRH APLIKACE	16
3 ŠABLONOVACÍ SYSTÉMY	17
3.1 ŠABLONY SMARTY	17
3.1.1 Instalace SMARTY	17
3.1.2 Spuštění jednoduché šablony	18
3.1.3 Funkce ve SMARTY.....	18
4 NÁVRH DATABÁZE	20
4.1 TABULKA APPROACHES	20
4.2 TABULKA SEARCH_ENGINES	21
4.3 TABULKA OTHER_REFERER_SERVERS	22
5 TŘÍDY POUŽITÉ V PRÁCI	23
5.1 TŘÍDA DETEKCE_VYHLEDAVACE (DETEKCE_VYHLEDAVACE.CLASS.PHP).....	23
5.1.1 kontrola_vstupu()	23
5.1.2 overeni_zapisu_pred_hodinou().....	24
5.1.3 ulozeni_vysledku().....	24
5.2 TŘÍDA ANALYZA (ANALYZA.CLASS.PHP).....	24
5.2.1 vrat_spojnicovy_graf_vyhledavace().....	25
5.3 TŘÍDA INSTALACE (INSTALACE.CLASS.PHP)	25
5.4 TŘÍDA POWER GRAPHIC (GRAPHIC.CLASS.PHP)	25
5.4.1 Ukázkový graf.....	29
5.4.2 Vlastnosti třídy	30
5.5 TŘÍDA EZSQL_MYSQL (EZ_SQL_MYSQL.CLASS.PHP)	30
5.5.1 Ukázková práce s databází	30
6 POPIS APLIKACE A JEJÍ OVLÁDÁNÍ	33
6.1 INSTALACE APLIKACE.....	33
6.1.1 Automatická instalace	33

6.1.2	Ruční instalace	33
6.2	STRÁNKA SOUHRNNÉ INFORMACE (=ÚVODNÍ STRÁNKA APLIKACE)	34
6.3	STRÁNKA NÁVŠTĚVY Z VYHLEDÁVAČŮ	34
6.4	STRÁNKA VYHLEDÁVANÉ ŘETĚZCE	34
6.5	STRÁNKA NÁVŠTĚVY Z JINÝCH ADRES	35
	ZÁVĚR	36
	SEZNAM POUŽITÉ LITERATURY.....	38
	SEZNAM POUŽITÝCH SYMBOLŮ A ZKRATEK	40
	SEZNAM OBRÁZKŮ	41
	SEZNAM TABULEK.....	42
	SEZNAM PŘÍLOH.....	43

ÚVOD

V dnešní době už není při tvorbě internetových stránek důležité myslet pouze na jejich vzhled, správně strukturovaný XHTML kód a vhodné použití jednotlivých XHTML značek, které vytváří sémanticky správně strukturovanou stránku, jenž je poté mnohem přístupnější pro vyhledávače i čtečky pro zrakově postižené či pro zařízení nepodporující CSS formátování.

Sémanticky správná stránka nezaručí ještě dobrou pozici ve vyhledávačích. Kromě obecných pravidel optimalizace je také velmi důležitá průběžná analýza a zpracování informací o návštěvách z vyhledávačů. Velmi dobrý nástroj na sledování a vyhodnocování návštěvnosti Google Analytics, který byl v roce 2005 uvolněn k bezplatnému použití širokou veřejností, nabízí kromě mnoha obecných funkcí a statistik s množstvím nastavení, rovněž možnosti pro sledování návštěvnosti z vyhledávačů. I přes svoje obrovské možnosti mi v této aplikaci chybělo jednoduché zobrazení naměřených dat a snažil jsem se proto navrhnout koncept aplikaci, která by srozumitelně interpretovala získané informace.

Jelikož se zabývám programováním a čekala mne bakalářská práce, rozhodl jsem se požádat o téma Analýza návštěvnosti z vyhledávačů, které rozšířilo mé znalosti o objektově orientované PHP5 a práci se šablonovacími systémy. Cílem bylo vytvoření aplikace, která by splňovala mé požadavky na vyhodnocování dat o návštěvnících přicházejících z vyhledávačů.

I. TEORETICKÁ ČÁST

1 OPTIMALIZACE PRO VYHLEDÁVAČE

V dnešní době, kdy jsou na internetu k dispozici informace téměř o všem, je pro fulltextové vyhledávače velmi důležitá schopnost nalézt stránky s nejrelevantnějším obsahem, který uživatel vzhledem ke klíčovému slovu požaduje.

Všichni vývojáři fulltextů se snaží optimalizovat engine vyhledávače tak, aby se jeho chování co nejvíce podobalo chování uživatele, tzn. laicky: co je dobré pro uživatele, je dobré i pro vyhledávače.

Vyhledávací robot neboli crawler je automatizovaná aplikace, která neustále prochází stránky zaindexované v databázi fulltextového vyhledávače a hledá odkazy na nové stránky. Rovněž zanáší do indexu jakékoliv změny stránek. Každý fulltext má vlastního robota a lze ho identifikovat v přístupech na stránky podle proměnné se jménem prohlížeče, kde je jeho jméno uvedeno. Uvádím zde, v našich oblastech, nejznámější:

Tab. 1 Seznam nejznámějších vyhledávacích robotů

SEZNAM.CZ	SeznamBot/1.0 (+http://fulltext.seznam.cz)
CENTRUM.CZ	holmes/3.8 (http://morfeo.centrum.cz/bot.php)
JYXO.CZ	Jyxobot/1
GOOGLE.COM	Googlebot/2.1
MSN.COM	msnbot/1.0 (+http://search.msn.com/msnbot.htm)
YAHOO.COM	Yahoo! Slurp

1.1 Faktory ovlivňující pozice ve vyhledávačích

Všechny fulltextové vyhledávače s vlastním engine, řeší od svého vzniku problém dodání co nejkvalitnějších výsledků uživateli [6]. Je mnoho faktorů, kterými se engine vyhledávače řídí při určování relevance jednotlivých stránek. Při popisu faktorů pro optimalizaci stránek lze vycházet jen z doporučení lidí pracujících v tomto oboru, jelikož způsoby, jak jednotlivé engine vyhledávačů přiřazují důležitost jsou věci tajnou.

1.1.1 Obsah klíčových slov v titulku stránky <title>

Délka titulku stránky může být až kolem 200 znaků, poté některé enginy vyhledávačů následující slova ořezávají (jyxo, morfeo). Důležitost slova v title je závislá na pozici v titulku. Opakování klíčového slova nemá na lepší výsledek žádný větší vliv.

1.1.2 Obsah klíčových slov v url

Je dobré, pokud se klíčové slovo vyskytuje i v url stránky, např. hledaný na výraz **levná Felicie** bude vhodnější www.levna-auta.cz/skoda/felicie/do-10-tisic nežli www.autohlohovec.cz/index.php?auto=1245.

1.1.3 Obsah klíčových slov na stránce

Klíčová slova se podle důležitosti pro vyhledávače řadí následovně.

- nadpisy H1-H3 (H4-H6 mají mizivou ne-li žádnou váhu)
- text zvýrazněný tagy nebo . Tagy , <i> slouží pouze k odlišení textu, nikoliv ke zvýšení důležitosti označených slov
- ostatní text na stránce: tagy p, div a span

1.1.4 Struktura stránek

Pro co nejlepší indexaci stránek je vhodná dobře strukturovaná navigace. Tzn. www.levna-auta.cz/skoda/felicie/do-10-tisic je mnohem vhodnější nežli www.autohlohovec.cz/index.php?kategorie=12&auto=1245. V prvním případě url přesně kopíruje způsob, jakým se návštěvník může k nabídce auta „proklikat“ pomocí menu.

1.1.5 Zpětné odkazy

Jenom kvalitní obsah u konkurenčnějších frází zdaleka nestačí k dobré pozici ve vyhledávačích. Velmi důležité jsou odkazy, které mohou přivést návštěvníky z jiných, nejlépe tématicky podobných, stránek. Důležitý je nejenom počet, ale i text, který je v odkazech uveden. Jako jeden ze známějších příkladů zneužití takového odkazu mohu uvést příklad kauzy „Namyšlený buran“. Na jaře 2006 proběhla internetem zpráva o Gogole bombě, jejíž obětí se stal Ing. Jiří Paroubek. Google bomba není nic jiného než odkaz, ve kterém je negativní text odkazující na stránky, které takto poškozují. Z toho příkladu je patrné, jakou sílu zpětné odkazy mají.

1.2 Faktory ovlivňující chování engine vyhledávačů

Kromě výše uvedených činitelů, které mají více či méně důležitý význam pro dobrou pozici ve fulltextových vyhledávačích, by se mělo zdůraznit ještě několik výrazů, které s tímto tématem úzce souvisí. [6]

1.2.1 Frekvence aktualizací

Pokud na stránkách probíhají často aktualizace obsahu, „naučí se“ na ně robot chodit a při jakémkoliv jejich změně bude do několika dnů jejich aktuální obsah v databázi vyhledávače. Toto se týká zejména vyhledávače google či jyx. Fulltext seznamu je v tuto chvíli ještě mladý a zřejmě nemá ani dostatek strojového výkonu, aby mohl často aktualizované stránky procházet častěji než 1x za týden. Toto je ovšem velmi závislé na čase a za půl roku může být frekvence návštěvnosti stránek robotem seznamu každodenní.

1.2.2 Jazyk použitý na stránkách

Při vyhledávání je většinou relevantní jazyk, ve kterém mají být hledány stránky, proto mají vyhledávače schopnost rozpoznání jazyku použitého na stránkách, který jim pomáhá třídit a filtrovat nevyžádané jazyky. Každý vyhledávač s vlastním hledacím engine si tyto algoritmy, stejně jako algoritmus vyhledávací vyvíjí sám a proto je kvalita rozpoznání jazyka různá. U vyhledávače jako google se může stát, že při hledání stránek „Pouze česky“ najde i stránky ve slovenštině, což je způsobeno jeho universálností pro vyhledávání a třídění v mnoha jazycích celého světa, tzn. pracuje s množinami výrazů charakteristickými pro každý jazyk a pokud dojde k překrývání některých těchto výrazů, tzn. jsou stejné pro více jazyků, může dojít ke špatnému rozpoznání jazyka použitého na stránkách.

Oproti tomu má např. seznam.cz mnohem kvalitnější algoritmus na rozpoznávání jazyka a dokáže určit, zda se jedná či nejedná o češtinu velmi přesně – má výhodu v tom, že vyhledává pouze v jednom jazyce(češtině) a může si proto dovolit dokonalejší heuristiku pro rozlišení použitého jazyka.

1.2.3 Podvodné techniky

Při použití podvodných technik, jejichž cílem je téměř vždy zvýšit návštěvnost z vyhledávačů na klíčové výrazy pro danou doménu, může dojít k zablokování stránek ve

výsledcích hledání, a to na základě „udání“ takovéto stránky a nebo odhalením těchto technik samotným vyhledávačem.

Každý fulltextový vyhledávač by měl obsahovat formulář na nahlášení zneužití jeho výsledků, kde je možné nahlásit stránky používající podvodné techniky. Jelikož by se mohlo stát, že bude nahlášena i stránka, která žádné podvodné techniky nepoužívá, jsou tyto formuláře s informacemi o zneužití ručně kontrolovány.

Jelikož se vyhledávače vyvíjí a probíhá jejich neustálé zdokonalování, které sleduje současně trendy internetu, zvětšují se rovněž i jejich možnosti detekce podvodných technik.

Vyřazení stránek z vyhledávačů může být na několik měsíců či let. Po tuto dobu nejsou stránky ve vyhledávači k dispozici a ten za určitou periodu provádí kontrolu odstranění či neodstranění technik, které vedly k vyřazení z výsledků.

1.2.4 Použití souboru robots.txt

Tento soubor určuje, kam všude na doméně má crawler přístup. Tzn. lze mu zakázat indexování adresářů, které obsahují citlivá data, která by mohla být zneužita. Samozřejmostí by mělo být zabezpečení těchto adresářů i programově, ale to už není otázkou této kapitoly. Pokud např. chcete zakázat přístupu do adresáře private, použijete tento kód. [15]

```
User-agent: *  
Disallow: /private/
```

Každý crawler má vlastní identifikaci User-agent, takže lze označit i adresáře, ke kterým nebude mít přístup určitý robot. Např. crawler googlu má označení Googlebot.

II. PRAKTICKÁ ČÁST

2 NÁVRH APLIKACE

PHP byla zvolena díky dlouhodobým zkušenostem s tímto jazykem. Při výběru verze PHP, pro kterou bude aplikace navržena, padl výběr na PHP5, které poskytuje větší podporu práce s objekty. Verze PHP5 má oproti verzi 4 resp. 3 tyto následující výhody. [4]

- rereference objektů - volání metody z objektu, která vrací další objekty
- modifikátory pro proměnné a metody (private, protected, static a abstract)
- lepší práce s konstruktory a destruktory objektu
- vícenásobná dědičnost
- používání výjimek (exceptions)

Z důvodu, že s OOP v PHP teprve začínám, bylo využito jen velmi málo z výše uvedených výhod PHP5, avšak je lepší pracovat v OOP v PHP5 nežli v nižších verzích PHP, jelikož je zde používání objektů rozhodně blíže tvorbě objektů v jiných objektově orientovaných jazycích. [1]

Protože aplikace používá k analýzám průběžně zaznamenávaná data, je na jejich uložení použito databáze. Stejně jako u výběru programovacího jazyka byla zvolena díky dlouhodobým zkušenostem a dobrým vlastnostem (snadná komunikace s PHP, dostupnost na drtivě většině hostingů s PHP) databáze MySQL. [3]

Jelikož jsem doposud nepracoval s žádným šablonovacím systémem, byla pro mne práce se systémem SMARTY výzvou, které nešlo odolat. [7]

Aplikace byla vytvořena za použití XHTML a formátována CSS, což je dnes zcela standardní způsob tvorby stránek. [5]

3 ŠABLONOVACÍ SYSTÉMY

Šablony se používají všude tam, kde je potřeba u webových stránek oddělit kód pro zpracování dat od HTML kódu, který slouží ke zobrazení výstupních dat. Je vhodné ho používat ke zvýšení přehlednosti kódu.

Na rozdíl od klasické dynamické stránky, kdy je programová část i její zobrazení v jednom souboru, je potřeba u šablony hned několika souborů.

- třída se šablonou
- třída pro práci s daty (může dědit z třídy pro práci se šablonou)
- samotná šablona

3.1 Šablony SMARTY

Tento šablonovací systém je jeden z nejrozšířenějších a to díky několika faktorům. [7]

- na jeho vývoji se neustále pracuje
- je to systém vyvíjený širokým spektrem programátorů
- programátor má možnost vytvoření vlastních pluginů
- systém je podporován přímo PHP Group

3.1.1 Instalace SMARTY

Instalaci SMARTY zvládne i začátečník s PHP. Stačí jej nahrát na server přes ftp a povolit práva na zapisování pro aplikace do adresáře, který bude obsahovat dočasné zkompilevané soubory. V případě, že je zkompilevaný soubor starší než šablona nebo PHP zdrojový soubor, dochází ke kompilaci nového souboru. Výchozí adresář je `templates_c` ve složce s nainstalovaným SMARTY. Samotné jádro šablonovacího systému je v `./libs/smarty.class.php` [11]

Pro běh SMARTY lze použít několik proměnných. Namátkově jedny z nejdůležitějších.

- `$template_dir` – adresář se šablonami (výchozí `./templates/`)
- `$compile_dir` – adresář se zkompilevanými soubory (výchozí `./templates_c/`)
- `$plugins_dir` – adresář pro nové pluginy (výchozí je v `./libs/plugins/`)
- `$debugging=true` – zobrazení všech proměnných poslaných šabloně

- `$caching=false` - zákaz ukládání zkompileovaných stránek do `$compile_dir`

Všechny uvedené proměnné jsou členské proměnné třídy smarty, tzn. musí být volány přes objekt. Tzn. pokud se vytvoří instance třídy `$smarty=new Smarty`, musí se jednotlivé proměnné volat `$smarty->template_dir`, `$smarty->compile_dir` apod.

3.1.2 Spuštění jednoduché šablony

Nejdůležitějšími věcmi při používání SMARTY je vložení třídy souboru se SMARTY, vytvoření instance, načtení šablony a její zobrazení. Jednoduchý příklad:

```
require_once('Smarty.class.php');
$smarty = new Smarty();
$smarty->display('index.tpl');
```

Aby byla šablona vůbec k něčemu a plnila svůj úkol, tzn. zobrazení dat generovaných PHP, je potřeba před zavolání metody `display` ještě přiřadit proměnným SMARTY nějakou hodnotu, a to pomocí funkce `assign`.

PHP

```
require_once('Smarty.class.php');
$smarty = new Smarty();
$smarty->assign('promenna', 'SMARTY');
$smarty->display('index.tpl');
```

ŠABLONA (index.tpl):

Toto je šablona použitá pro systém `{$promenna}`.

HTML VÝSTUP:

Toto je šablona použitá pro systém SMARTY.

3.1.3 Funkce ve SMARTY

Jak je z ukázky šablony patrné, proměnné se uzavírají do složených závorek. Pro doplnění: nejsou to pouze proměnné, ale všechny příkazy šablonovacího systému SMARTY, tzn. i podmínky, cykly, podmíněné komentáře či includování jiných souborů, pokud tyto chceme provést v šabloně a ne ve zdrojovém kódu PHP skriptu. Opět uvedu jednoduchý příklad použití podmínky `if`:

```
{if $login==1}
Povedlo se přihlásit
{else}
Přihlášení se nezdařilo
```

```
{/if}
```

V šablonách lze používat proměnných poslaných metodou GET, POST či uložených SESSION či COOKIES přes pole `$smarty.get` resp. `$smarty.post` atd. Tečka odděluje v šabloně jednotlivé části pole, tzn. co by se v PHP napsalo jako `$pole['udaj1']['udaj2']` se ve SMARTY napíše jako `$pole.udaj1.udaj2`.

Pokud pole nemá indexy, lze jej procházet cykly `foreach` nebo `section`, který má více možností v nastavení parametrů.

PHP

```
$types = array('555','666','777');  
$smarty->assign('types',$types);
```

ŠABLONA

```
{section name=num loop=$types}  
    {$types[num]} <br />  
{/section}
```

HTML VÝSTUP

```
555<br />  
666<br />  
777<br />
```

4 NÁVRH DATABÁZE

Při návrhu aplikace se vycházelo z předpokladů, že bude potřeba tabulka s informacemi o návštěvníkovi a stránce, ze které přišel, tabulku s informacemi o vyhledávači a tabulku, kam se budou zaznamenávat servery, které nejsou označeny jako vyhledávače. Jelikož bylo použito v PHP skriptech kódování UTF-8, bylo toto kódování i použito i v databázích, které v tomto kódování ukládají data (typ varchar). Přesný typ použitého kódování je utf8_general_ci. [9]

4.1 Tabulka approaches

Tab. 2 Struktura tabulky approaches

Sloupec	Typ
id	int(10), unsigned
referer	varchar(255)
time	datetime
ip	varchar(39)
prohlizec	varchar(255)
retezec	varchar(255)
url	varchar(255)
id_search_engine	smallint(4), unsigned
id_other_referer_server	smallint(4), unsigned

Pro sloupec `time` byl použit formát `datetime`. Mnoho lidí používá zápis v unixovém formátu, tzn. v sekundách od půlnoci 1.ledna 1970, avšak `datetime` má oproti unixovému formátu času výhodu v práci s časem přímo na úrovni dotazu na databázi, což ve výsledku znamená, že se nemusí provádět žádné výpočty či převody sekund v PHP, ale stačí pouze zadat podmínku času do sql dotazu.

U ip adresy je uvedena délka 39 znaků, což je max. délka IPv6, který je nástupce stávajícího IPv4, jenž už přestává kapacitám internetu v dnešní době dostáčet. [14]

Po dokončení aplikace bylo zjištěno, že pokud je zadán dotaz typu `SELECT * FROM approaches WHERE retezec='odlucovac'`, najdou se všechny řádky, které obsahují ve sloupci `retezec` jak slovo `odlucovac`, tak i slovo `odlučovač`, `odlucovač` či `odlučovac`. Tento sloupec byl tedy převeden na `utf8_bin`, který porovnává bitovou hodnotu znaků.

4.2 Tabulka `search_engines`

Tab. 3 Struktura tabulky `search_engines`

Sloupec	Typ
<code>id</code>	<code>smallint(4),unsigned</code>
<code>search_engine</code>	<code>varchar(50)</code>
<code>var_url_string</code>	<code>varchar(10)</code>
<code>name</code>	<code>varchar(20)</code>
<code>encoding</code>	<code>set('UTF-8', 'ISO-8859-2', 'CP1250')</code>

Sloupec `search_engine` obsahuje url vyhledávače bez `www`, např. `google.com`, či `jyxo.cz`. Název vyhledávače uvedený ve sloupci `name` je zobrazen v aplikaci a slouží k přehlednější identifikaci vyhledávače, např. `fulltext seznamu`, `google – Bulharsko` apod.

`var_url_string` udává název proměnné, který obsahuje hledaný řetězec. Obvykle mívá 1-6 znaků.

Jelikož používají různé vyhledávače různá vnitřní kódování (jazykové sady), musí tabulka rovněž obsahovat sloupec s informací o tomto kódování, aby bylo možné správně interpretovat vstupní data. Pro konverzi slouží modul `iconv`, který je ve verzi PHP5 standardně obsažen a u verze PHP4 nebývá jen zřídka mezi standardní moduly PHP.

4.3 Tabulka other_referer_servers

Tab. 4 Struktura tabulky other_referer_servers

Sloupec	Typ
id	smallint(4), unsigned
server	varchar(255)
name	varchar(255)

Tato tabulka slouží k uložení serverů, které nepatří mezi vyhledávače a nejsou tedy v databázi. Lze tak tedy zjistit, ze kterých konkrétních stránek kromě vyhledávačů na analyzovanou doménu chodí.

5 TŘÍDY POUŽITÉ V PRÁCI

Tato kapitola obsahuje popis tříd, které byly vytvořeny (`Detekce_vyhledavace`, `Analyza`, `Instalace`) a také tříd, které jsou použity jako externí, tzn. určeny k volnému použití (`ezSQL_mysql`, `PowerGraphic`). Všechny jsou uloženy v aplikaci v adresáři `classes`. Kromě tříd samotných se budu věnovat i funkci jednotlivých klíčových metod, které jsem používal či vytvářel.

Jednotlivé třídy byly psány s co největším využitím nových funkcí OOP, které přináší PHP5 a nejsou zpětně kompatibilní s PHP4. Externí třídy jsou vytvořeny v PHP4.

Kromě výše uvedených tříd obsahuje aplikace ještě další 2 třídy (`Prace_s_textem`, `Prace_s_url`), které ale díky jejich jednoduchosti nejsou zahrnuty do manuálu. V závorce za názvem třídy je jméno souboru v adresáři `classes`, ve kterém jsou všechny uloženy. Některé třídy používají konstanty pro práci s databází či zobrazením počtu výsledků. Tyto konstanty jsou uloženy v souboru `pristup.php`.

5.1 Třída `Detekce_vyhledavace` (`detekce_vyhledavace.class.php`)

Třída je inicializována při úspěšném načtení stránky sledované domény a slouží pro identifikaci adresy, ze které se návštěvník na stránky dostal. Identifikace probíhá rozborem http hlavičky `HTTP_REFERER`.

Rozbor je proveden v konstruktoru třídy, kdy se ověří, zda nepřichází návštěvník ze stránek analyzované domény nebo nepoužívá tlačítko nazpět (viz metoda `kontrola_vstupu`). Pokud přišel návštěvník z externí stránky, je provedena kontrola, zda se jedná o známý vyhledávač či o jiný server. V případě, že není zanesen v databázi, uloží se do ní společně s analyzovanými daty z http hlavičky.

5.1.1 `kontrola_vstupu()`

Hlavním úkolem této metody je ověření a případný zápis cookie s informací o stránce, ze které se návštěvník na analyzovanou doménu dostal. Odfiltruje všechny návštěvníky, kteří se dostaly na stránky, klikli na odkaz hlouběji do struktury webu a poté se vrátili přes tlačítko Zpět na stránku vyhledaných odkazů. Při opětovném požadavku na analyzovanou stránku je již COOKIE zapsána a tudíž nedochází ke spuštění další části ověřovacího mechanismu a instance třídy je ukončena.

5.1.2 overeni_zapisu_pred_hodinou()

Metoda ověřuje, zda není v databázi již uložena za poslední hodinu stejná IP adresa s internetovou adresou, ze které se na doménu přišlo. Jelikož je pro kontrolu potřebné vytvořit dotaz na databázi, je metoda spuštěna až po ověření členskou metodou `kontrola_vstupu()`.

5.1.3 ulozeni_vysledku()

Pokud je server, ze kterého návštěvník přišel v seznamu vyhledávačů, zjistí se, jaké kódování server používá, převede se z něj na UTF-8, které používá aplikace (použitím modulu `iconv`), a výsledek se uloží do databáze. Tento systém je použit, aby se zamezilo zkomolení slov, jak je uvedeno v Tab. 5.

Tab. 5 Zobrazení různých kódování při nastavení jazyka na windows-1250

WINDOWS-1250	Příliš žluťoučký kůň úpěl ďábelské ódy.
ISO-8859-2	Přílią łlu»oučký kůń úpěl ďábelské ódy.
UTF-8	PĀ,Ā-liĹ ĹłluĀťouĀ“kĀ” kĀąĀ ĀşpĀ-l ĀŽĀ~belskĀ© Āłdy.

5.2 Třída Analyza (analyza.class.php)

Třída pro vyhodnocování dat uložených v databázi a předávání proměnných a jejich hodnot šablonovacímu systému SMARTY.

Jak je ze zápisu třídy patrné, bylo hojně jsem použito lokálních proměnných, které používaly pouze jednotlivé metody a nebyly přístupné třídě samotné.

U dotazů na databázi MySQL byly pokládány co nejkompexnější dotazy, aby se tím omezilo jejich množství.

Jelikož je u jednotlivých metod uvedeno relativně srozumitelně jejich použití a všechny pracují pouze s výběrem dat z databáze a s jejich uložením do pole, je uveden pouze popis jedné z posledních metod ve zdrojovém kódu, která se od těchto odlišuje práci s grafickou třídou Power Graphic. [13]

5.2.1 vrat_spojnicovy_graf_vyhledavace()

Název metody je mírně zavádějící, jelikož nevrací přímo graf, ale pouze url, kterou se volá třída Power Graphic (v url obrázky) a při volání této třídy se vygeneruje graf.

Po výběru dat z tabulky jsou provedeny drobné modifikace, pokud nejsou za měřené období (1 rok) dostupné všechny měsíce, jsou tyto nahrazeny nulovými hodnotami a jsou nastaveny jednotlivé hodnoty třídy, bez kterých by nebylo možné vygenerovat správné url. Více informací o jednotlivých proměnných a nastavení třídy v kapitole 5.4.1.

5.3 Třída Instalace (instalace.class.php)

Pokud se uživatel rozhodne pro automatizovanou instalaci aplikace, tzn. spustí aplikaci bez vytvořeného souboru `nastaveni.php`, inicializuje se třída pro instalaci aplikace, ve které se krom připojení k databázi zadávají ještě dodatečné informace.

- Prefix tabulek – slouží pro odlišení od ostatních tabulek, pokud je k dispozici pouze jedna databáze a nemůžeme mít data uložena v separované databázi
- Doména – orientační hodnota, aby bylo zřejmé, na jaké doméně byla analýza provedena
- Vyhledávačů – počet vyhledávačů, který se zobrazí při jejich výpisu (viz kapitola 6.3)
- Jiných serverů – počet serverů, které se zobrazí při jejich výpisu (viz kapitola 6.5)
- Frázi – počet frází, který se zobrazí při jejich výpisu (viz kapitola 6.4)
- Adres z jiných serverů – počet odkazů, které se zobrazí při výpisu serverů (viz kapitola 6.5)

Po získání všech hodnot nutných pro instalaci aplikace se vytvoří soubor `nastaveni.php` a všechny potřebné tabulky v databázi.

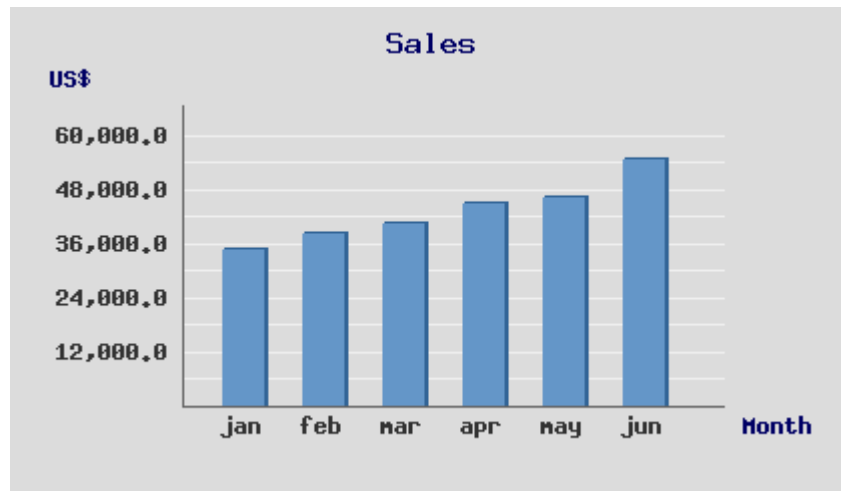
5.4 Třída Power Graphic (graphic.class.php)

Třída vykresluje grafy pomocí grafické knihovny GD, která se přikompiluje k PHP běžícímu na serveru. Tato knihovna je standardní součástí všech hostingů s PHP.

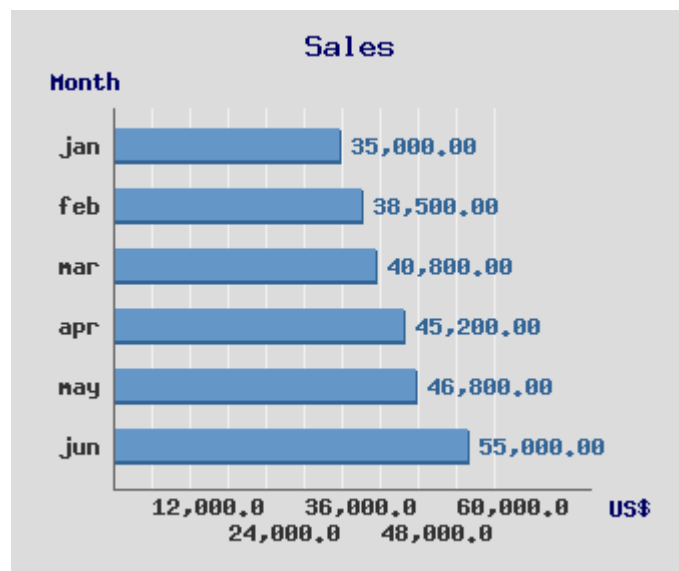
Třída pracuje na principu vygenerování url po zadání hodnot potřebných proměnných a následně volání třídy s takto vygenerovanou internetovou adresou, např:

.

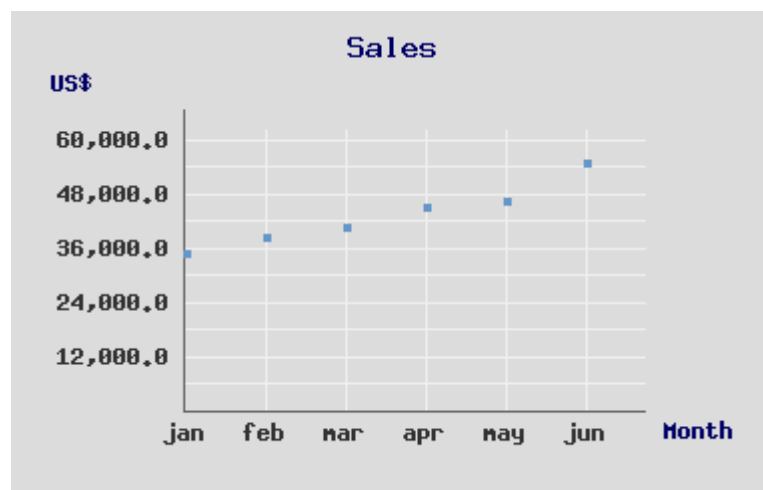
Třída má možnost vykreslení 6 typů grafů. [13]



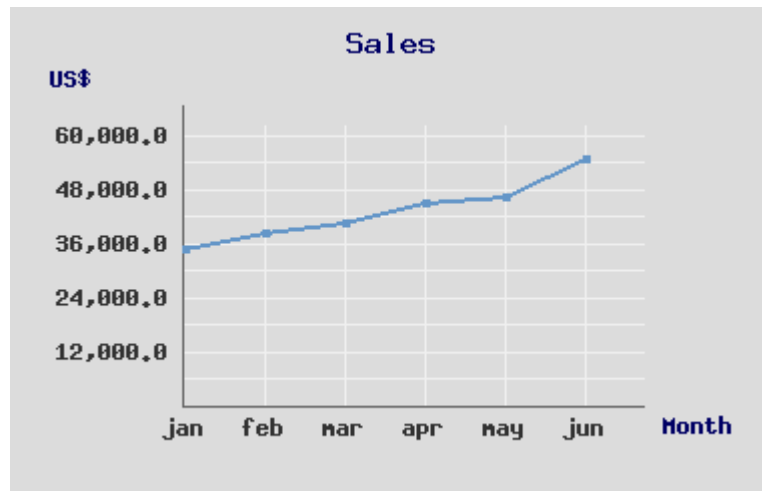
Obr. 1 Graf se svislými pruhy



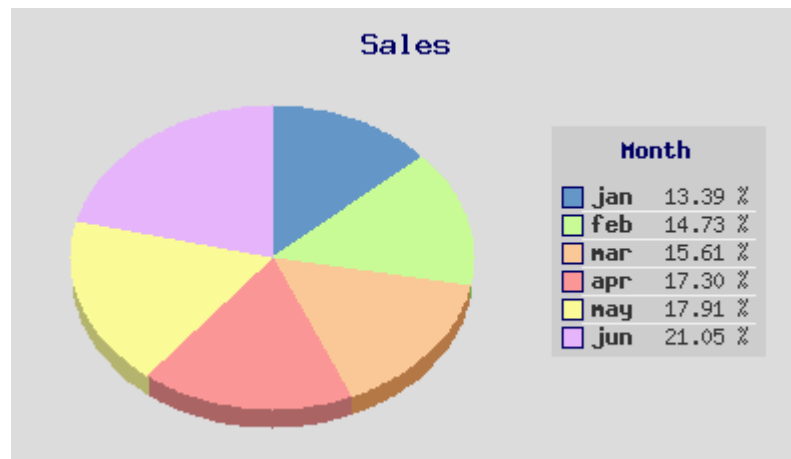
Obr. 2 Graf s vodorovnými pruhy



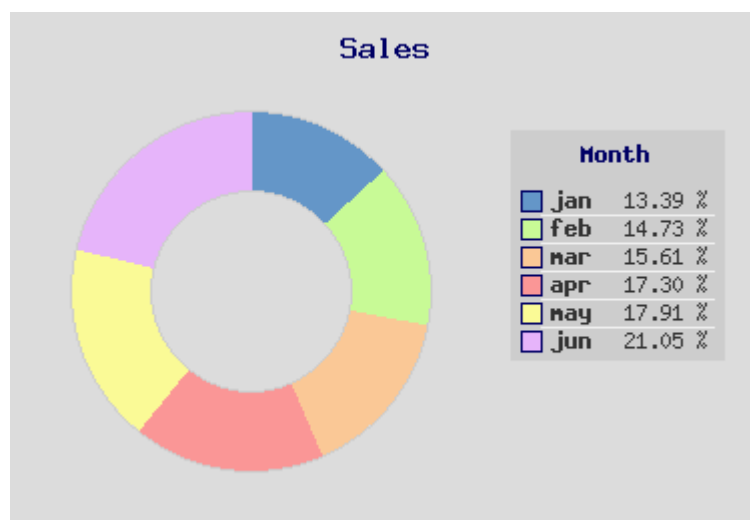
Obr. 3 Bodový graf



Obr. 4 Spojnicový graf

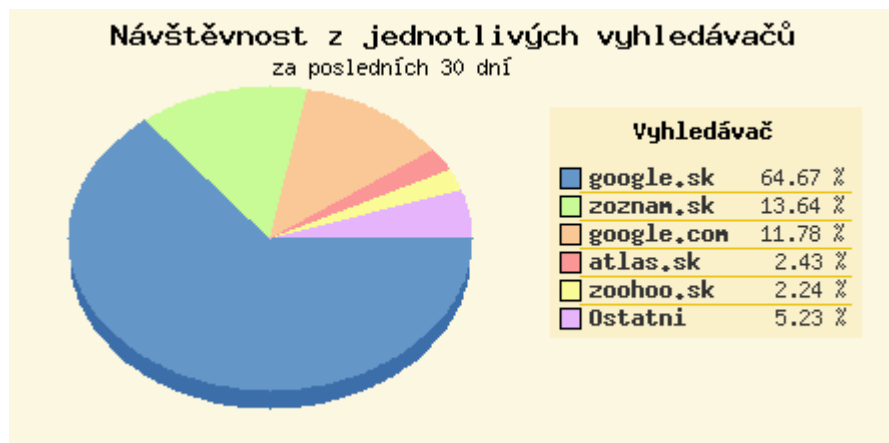


Obr. 5 Koláčový graf

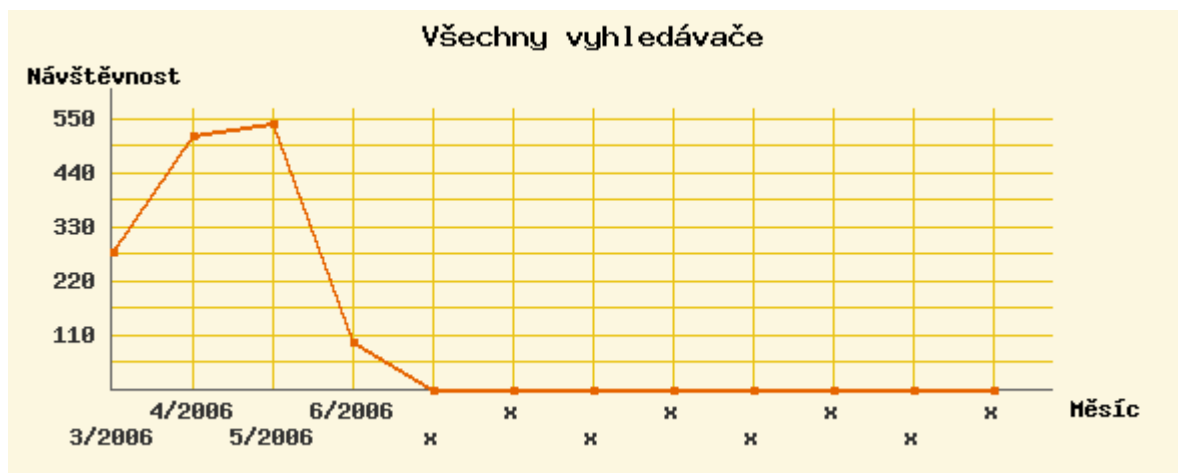


Obr. 6 Prstencový graf

Všechny grafy je možné vykreslit ve 3 základních barevných variacích a jdou jednoduše modifikovat změnou hexadecimálního tvaru čísla, udávajícího množství RGB. Jak je patrné z obrázku 7 a 8, výše uvedená barevná variace byla modifikována, aby více ladila s barevným pojetím celé aplikace.



Obr. 7 Koláčový graf s modifikovanými barvami



Obr. 8 Spojnicový graf s modifikovanými barvami

Kromě této nepatrné modifikace bylo nutné provést úpravy rovněž ve zpracování češtiny, a to z důvodu zpracování znaků s diakritikou pouze v kódování ISO-8859-2. Problém spočíval v předávání znaků graf. knihovně GD, která nezpracovává používané kódování UTF-8, ale pouze středoevropské ISO-8859-2. Úprava byla provedena použitím modulu iconv v úvodu třídy, kdy jsou načteny hodnoty pro Titulek grafu a popis os.

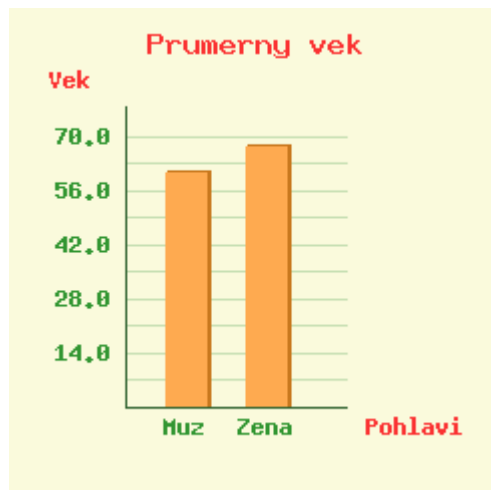
Poslední úprava se týkala vykreslování osy Y u spojnicového grafu, která při malých číslech zobrazovala i desetiny, pokud toto vyšlo na relativní velikost grafu. Např. při max. hodnotě 8 návštěv v měsíci byly zobrazeny popisky dat o velikostech 1,6; 3,2 ; 4,8; 6,4 a 8.

5.4.1 Ukázkový graf

Před generováním samotné url, uvedené v adrese obrázku, je potřeba načíst nadpis, popisky os a vlastní hodnoty, které se budou zobrazovat. To se provádí přiřazováním hodnot jednotlivým proměnným:

```
$graf = new PowerGraphic; //inicializace třídy
$graf->type      = 1; //výběr typu grafu, 1=graf se svisl. pruhu
$graf->skin      = 3; //výběr barevné variace grafu
$graf->title     = 'Prumerny vek'; //titulek grafu
$graf->axis_x    = 'Pohlavi'; //popis osy x
$graf->axis_y    = 'Vek'; //popis osy y
$graf->x[0]      = 'Muž'; //název prvního pruhu
$graf->y[0]      = '61'; //hodnota prvního pruhu
$graf->x[1]      = 'žena'; // název druhého pruhu
$graf->y[1]      = '68'; //hodnota druhého pruhu
```

Po zavolání metody `$graf->create_query_string()` se vytvoří řetězec: `title=Prumerny+vek&axis_x=Pohlavi&axis_y=Vek&type=1&skin=3&graphic_1=&graphic_2=&credits=&x0=Muz&y0=61&x1=Zena&y1=68`, který se použije jako seznam proměnných vložený za název třídy např. `graphic.class.php?...`, a vygeneruje následující graf:



Obr. 9 Ukázkový graf třídy Power Graphic

5.4.2 Vlastnosti třídy

Hlavní devizou této třídy je její jednoduchá inicializace a přiřazení hodnot popiskům a jednotlivým hodnotám v grafu. Další výhodou je také výše uvedený počet grafů, které třída generuje.

Hlavním nedostatkem třídy je nemožnost vykreslit v bodových či spojnicových grafech více než dva průběhy. Druhou chybou je zatížení serveru, které knihovna GD při zpracování obrázků způsobuje. Řešením by bylo pouze automatizované generování obrázků za určitou periodu a jejich uložení na server. Toto by bylo možné zajistit pod linuxem aplikací cron, která by spustila příslušný PHP skript. Jelikož je však třída použita pouze na vykreslování několika obrázků za den, byl by čas modifikace třídy neúměrný výsledku. Výborně má tento problém řešený systém pro analýzu návštěvnosti Google Analytics, který všechny grafy generuje pomocí FLASH objektů na stránce, kterým pouze posílá předzpracovaná data a odpadá tím serveru zátěž v podobě generování jakýchkoliv obrázků. Druhým plusem metody použité Google Analytics je, že lze u FLASHe použít vyhlazování, stíny a jiné grafické efekty, které s modulem GD docílíme jen obtížně, či vůbec.

5.5 Třída ezSQL_mysql (ez_sql_mysql.class.php)

Tato třída slouží k práci s mysql databází a slouží jako nástavba nad klasické mysql příkazy. Tedy, ve výsledku, s mírnou modifikací této třídy lze při zachování ostatních částí aplikace přistupovat i k jiným typům databází, které umožňují přístup k uloženým datům přes internet. [10]

5.5.1 Ukázková práce s databází

Po nezbytném includování souboru s třídou je nutné se připojit k databázi, což lze provést při vytvoření instance třídy:

```
$db=new db(DBUSER, DBPASSWORD, DBNAME, DBSERVER);
```

Kde DBUSER, DBPASSWORD, DBNAME, DBSERVER jsou konstanty, které si můžeme definovat v externím souboru, který budeme vkládat funkcí include. Po připojení k databázi je možné pracovat s daty. [12]

Pokud je potřeba pracovat s daty z databáze, je možné je načíst do asociativní či číselného pole nebo do objektu.

Tab. 6 Tabulka s ukázkovými hodnotami pro demonstraci
výběru z databáze

id	hodnota
1	První
2	Druhý

- Výběr s konstantou ARRAY_N

```
$result = $db->get_results('SELECT * FROM tabulka',ARRAY_N)
$result = Array (0=>Array(0=>1,1=>První),1=>Array(0=>2,1=>Druhý))
```

- Výběr s konstantou ARRAY_A

```
$result = $db->get_results('SELECT * FROM tabulka',ARRAY_A)
$result = Array
(0=>Array(id=>1,hodnota=>První),1=>Array(id=>2,hodnota=>Druhý))
```

- Výběr s konstantou OBJECT (defaultně nastaven)

```
$result = $db->get_results('SELECT * FROM tabulka',OBJECT)
K proměnným přistupujeme jako k členským metodám:
$result->0->id, apod.
```

Stejný způsob se používá i pro výběr jednoho řádku z databáze, kdy se místo metody `get_results` použije `get_row`.

Pokud je z nějakého důvodu potřeba výpis dat, které byly získány posledním požadavkem, slouží pro tento výpis metoda `debug`:

```
$result = $db->get_results('SELECT * FROM tabulka',ARRAY_A)
$db->debug();
```

Metoda v místě volání vygeneruje následující kód:

```
ezSQL (v2.0) Debug..  
Query [1] -- [SELECT * FROM tabulka]  
Query Result..
```

(row)	int 1 id	string 5 hodnota
1	1	První
2	2	Druhý

Jak je z tabulky patrné, zobrazuje funkce `debug` výpis posledního příkazu volaného pomocí třídy a to včetně výsledků. Pomocný první sloupec `(row)` slouží pro zpřehlednění výsledků, jelikož ne vždy máme ve vybraných datech i čísla řádků.

6 POPIS APLIKACE A JEJÍ OVLÁDÁNÍ

Celá aplikace je dostupná na adrese http://www.zetr.sk/analyza_navstevnosti_z_vyhledavacu/, kde je možné si prohlédnout všechny její funkce (kap. 6.2 - 6.5). Jelikož nebylo možné veřejně zpřístupnit celou MySQL databázi firmy zetr.sk, byly používané 3 tabulky včetně získaných dat exportovány na moji mysql databázi mysql.podolinsky.net, kde jsou do konce června 2006 k dispozici k nahlédnutí.

6.1 Instalace aplikace

Instalace aplikace je možná buďto ručně, či automaticky. Pokud se z jakéhokoliv důvodu neprovede instalace automatická, která by s měla spustit při otevření souboru `index.php`, doporučuje se použít instalaci ruční.

6.1.1 Automatická instalace

Viz přílohy P I a P II

Po vyplnění všech potřebných údajů (viz kapitola 5.3) dojde k automatickému vytvoření konfiguračního souboru `nastaveni-default.php` a vytvoření tabulek v mysql databázi. Před potvrzením je ještě nutné povolit zápis do adresáře s aplikací, aby mohl skript vytvořit soubor `nastaveni.php`.

6.1.2 Ruční instalace

Instalace sestává ze dvou kroků - vyplnění všech konstant používaných programem a vytvoření tabulek v databázi.

Soubor pro vyplnění konfiguračních konstant programu je v adresáři s aplikací a má název `nastaveni-default.php`. Stačí v něm změnit první 4 konstanty (`EZSQL_DB_HOST`, `EZSQL_DB_NAME`, `EZSQL_DB_USER`, `EZSQL_DB_PASSWORD`, `DOMAIN`) a zbytek může zůstat vyplněn výchozími hodnotami.

Tabulky se vytvoří zkopírováním obsahu souboru `analyza_navstevnosti_z_vyhledavacu.sql` do MySQL.

Konfiguraci aplikace lze kdykoliv změnit úpravou souboru `nastaveni-default.php`.

6.2 Stránka Souhrnné informace (=úvodní stránka aplikace)

Viz příloha P III.

Tato stránka obsahuje nejdůležitější informace analýzy na jednom místě: procentuální a absolutní návštěvnost z vyhledávačů, celkový počet přístupů z vyhledávačů, návštěvnost z jiných serverů a nejvyhledávanější fráze, které vedly na analyzovanou doménu.

Jednotlivé údaje jsou provedeny pomocí odkazů a vedou na detailnější informace o dané položce. Dle mého se nejvíce informací dozví uživatel při kliknutí na vyhledávač v jejich seznamu, kdy se mu zobrazí jeho detailní popis. Více informací u Návštěv z vyhledávačů.

6.3 Stránka Návštěvy z vyhledávačů

Viz příloha P IV.

Vypíše se seznam vyhledávačů o určitém počtu (konstanta `ZOBRAZENYCH_VYHLEDAVACU`), spojnicový graf s návštěvností za jednotlivé měsíce s historií počtu přístupů, rozbalovací seznam detailních informací pro jednotlivý měsíc a seznamem vyhledávaných slov (konstanta `ZOBRAZENYCH_VYHLEDAVANÝCH_FRAZI`) za určité období (konstanta `DNU_NAZPET`).

Po výběru měsíce ze seznamu a potvrzení výběru dojde k načtení detailnějších informací z vyhledávače, na kterou bychom se dostali i ze stránky Souhrnných informací, kliknutím na vyhledávač. Krom stejného grafu jako u seznamu všech vyhledávačů (data se nemění), je na stránce zobrazen i seznam vyhledávaných slov za dané období, které je vybráno v rozbalovacím menu měsíců.

6.4 Stránka Vyhledávané řetězce

Viz příloha P V.

Na této stránce se dozvíte návštěvnost z jednotlivých vyhledávačů na hledanou frázi. Jelikož je oproti návštěvnosti z vyhledávačů relativně nestálá návštěvnost na jednotlivé fráze (do seznamu se mohou zařadit i jiný měsíc nezobrazené), je zbytečné zde mít pro ně grafy. Pro orientační hodnoty v jednotlivých měsících je opět vhodné použít rozbalovací menu s výběrem jednotlivých měsíců.

6.5 Stránka Návštěvy z jiných adres

Viz příloha P VI.

Stránka obsahuje určitý počet odkazů (konstanta ZOBRAZENYCH_ADRES_Z_JINYCH_SERVERU) pro určitý počet serverů (konstanta ZOBRAZENYCH_JINYCH_SERVERU), které na analyzované stránky odkazují a zvyšují tak jejich Page Rank hodnotu.

ZÁVĚR

Jelikož je použitá aplikace zaměřená na lidi, kteří mají alespoň minimální zkušenosti s tvorbou stránek (naprostému laikovi zřejmě aplikace mnoho neřekne), pojmul jsem tak i teoretickou část, kde jsem se snažil uvést stručně opravdu jenom ty nejdůležitější body při optimalizaci pro fulltextové vyhledávače.

Při tvorbě jsem narazil na několik drobných problémů, které však neměly ve výsledku žádný vliv na funkce aplikace.

Testování probíhalo na 3 hostinzích s PHP5 s MySQL databází.

- hosting pipni.cz, program FREE, systém Linux www3 2.6.16.5, PHP5.1.2
- hosting pipni.cz, program PROFI, systém Linux www6 2.6.16.5, PHP5.1.2
- hosting czechia.com, program LinuxPHP5, systém Linux eunomia 2.6.8-2-686-smp, PHP5.1.1.-1.dotdeb.2

Pokud bych měl shrnout vlastnosti aplikace, splnila všechny požadavky, které jsem na ni při vymýšlení konceptu kladl.

- grafické zobrazení analyzovaných dat
- aktuální informace za určitou dobu, nejenom za jednotlivé měsíce
- zobrazení vyhledávaných frází bez zkomolení češtiny a tím i omezení zkreslení výsledků

Při vytváření databáze jsem opomenul podstatnou věc, a tou je dostupné místo pro databázi na hostingu. Většinou bývá u hostingů místo omezeno na 50 – 100 MB, což je pro záznamy až rok zpětně v tomto formátu značně málo. Hrubý odhad odvozený z dosud naměřených údajů značí limit obsahu dat v tabulce approaches (viz kap. 4.1) na 200 000 řádků na 50 MB databáze, což ve výsledku znamená cca. 550 návštěv z vyhledávačů denně. Pro zvětšení kapacity záznamů by bylo možné provést odstranění několika nadbytečných sloupců, které jsou v tabulce pro administrátory, kteří mají zájem si zjistit detailnější informace o návštěvníkovi, jako je verze internetového prohlížeče nebo ip adresa návštěvníka. Též by bylo možné odstranit informaci refereru u řádků, které obsahují informace o příchodu z vyhledávače – adresa ve vyhledávači je téměř zcela zbytečný údaj. Těmito kroky lze docílit až dvojnásobení kapacity tabulky, tzn. do 50MB databáze by bylo

možné zaznamenat až 400 000 řádků. Všechny výše uvedené odhady nepočítají se zaplněním ostatních tabulek v databázi.

Použití této aplikace je vhodné u prezentací, kde administrátora či majitele stránek zajímá návštěvnost z vyhledávačů na určitá klíčová slova a nejenom pouze návštěvnost celková.

SEZNAM POUŽITÉ LITERATURY

- [1] Gilmore, W. Jason. Velká kniha PHP5 a MySQL : kompendium znalostí pro začátečníky i profesionály. RNDr. Jan Pokorný. 1. vyd. Brno: Zoner Press, 2005. 711 s. ISBN 80-86815-20-X.
- [2] Prokop, Marek. CSS - kaskádové styly pro webdesignéry. Mobil Media, 2003.
- [3] Maslakowski, Mark. Naučte se MySQL za 21 dní. Computer Press, 2001.
- [4] Schlossnagle, Georgie. Pokročilé programování v PHP5. Zoner Press, 2004. 640 s. ISBN 80-86815-14-5
- [5] Stáníček, Petr. CSS kaskádové styly: kompletní průvodce. Computer Press, 2003. 178 s. ISBN 80-7226-872-4
- [6] Nawebu.cz: Archiv konference o SEO (Search Engine Optimization) [online]. [cit. 2006-6-10]. Dostupný z www: <<http://seo.nawebu.cz/>>
- [7] Smarty: Template Engine [online]. [cit. 2006-6-10]. Dostupný z WWW: <<http://smarty.php.net/>>
- [8] PHP: PHP Documentation [online]. [cit. 2006-6-10]. Dostupný z WWW: <<http://www.php.net/manual/en/>>
- [9] MySQL: MySQL Documentation [online]. [cit. 2006-6-10]. Dostupný z WWW: <<http://dev.mysql.com/doc/>>
- [10] Olšavský, Marek: PHP a SQL - EZ_SQL [online]. Linuxsoft.cz: PHP a SQL [cit. 2006-6-10]. Dostupný z WWW: <http://www.linuxsoft.cz/article.php?id_article=664>
- [11] Olšavský, Marek: PHP knihovny - Smarty [online]. Linuxsoft.cz: PHP knihovny [cit. 2006-6-10]. Dostupný z WWW: <http://www.linuxsoft.cz/article.php?id_article=714>
- [12] JV Multimedia: ezSQL Database Class (2.00) [online]. [cit. 2006-6-10]. Dostupný z WWW: <<http://www.jvmultimedia.com/portal/node/6>>
- [13] Reche, Carlos: Power Graphic [online]. PHP Classes [cit. 2006-6-10]. Dostupný z WWW: <<http://www.phpclasses.org/browse/package/1993.html>>
- [14] Vrána, Jakub: Ukládání IP adresy [online]. PHP triky [cit. 2006-6-10]. Dostupný z WWW: <<http://php.vrana.cz/ukladani-ip-adresy.php>>

- [15] Janovský, Dušan: Zákaz přístupu vyhledávačům [online]. Jak psát web [cit. 2006-6-10]. Dostupný z WWW: <<http://www.jakpsatweb.cz/robots-txt.html>>

SEZNAM POUŽITÝCH SYMBOLŮ A ZKRATEK

CSS	(Cascading Style Sheets). Kaskádovací styly. Slouží pro naformátování stránky, která obsahuje XHTML tagy.
FTP	(File Transfer Protocol). Protokol sloužící pro nahrávání souborů na jiný počítač (nejčastěji hosting).
GD	Celým názvem GD Graphic Library. Grafická knihovna pro generování a úpravu obrázků na serveru ve formátu PNG, JPG či GIF.
XHTML	(The Extensible HyperText Markup Language). Značkovací jazyk pro tvorbu www stránek.
IPv4	Internetový protokol pro adresaci dat. Každý počítač obsahuje svoji IP adresu. Je zastaralý a probíhá jeho nahrazování protokolem IPv6.
ISO-8859-2	Středoevropské kódování češtiny. Starší a více podporované (i na jiných platformách než OS Windows) kódování středoevropského jazyka než windows-1250.
MySQL	Jedna z nejrozšířenějších open source databází.
OOP	Zkratka pro sousloví Objektově orientované programování. Používá se pro všechny jazyky, využívající modelu objektového programování.
Page Rank	Hodnocení stránek vyhledávačem. Je přímo závislé na množství odkazů směřujících na danou stránku.
PHP	(PHP: Hypertext Processor). Programovací jazyk zpracovávaný na straně serveru.
SEO	(Search Engine Optimization). Optimalizace stránek pro internetové vyhledávače.
UTF-8	(UCS Transformation Format). Tzv. unicode kódování znaků. Nejuniversálnější kódování znaků obsahujících češtinu a další znaky, které nemusí být obsaženy v místní znakové sadě (např. ruština).
WINDOWS-1250	Středoevropské kódování češtiny od společnosti Microsoft. Používané na OS Windows.

SEZNAM OBRÁZKŮ

Obr. 1 Graf se svislými pruhy.....	26
Obr. 2 Graf s vodorovnými pruhy	26
Obr. 3 Bodový graf	27
Obr. 4 Spojnicový graf.....	27
Obr. 5 Koláčový graf	27
Obr. 6 Prstencový graf.....	27
Obr. 7 Koláčový graf s modifikovanými barvami	28
Obr. 8 Spojnicový graf s modifikovanými barvami	28
Obr. 9 Ukázkový graf třídy Power Graphic.....	29

SEZNAM TABULEK

Tab. 1 Seznam nejznámějších vyhledávacích robotů	11
Tab. 2 Struktura tabulky approaches	20
Tab. 3 Struktura tabulky search_engines	21
Tab. 4 Struktura tabulky other_referer_servers	22
Tab. 5 Zobrazení různých kódování při nastavení jazyka na windows-1250.....	24
Tab. 6 Tabulka s ukázkovými hodnotami pro demonstraci výběru z databáze	31

SEZNAM PŘÍLOH

- P I: 1. Instalační stránka aplikace - zadání dat
- P II: 2. Instalační stránka aplikace - ověření údajů
- P III: Úvodní stránka aplikace
- P IV: Stránka aplikace - Návštěvy z vyhledávačů
- P V: Stránka aplikace - Vyhledávané řetězce
- P VI: Stránka aplikace - Návštěvy z jiných adres

PŘÍLOHA P I: 1. INSTALAČNÍ STRÁNKA APLIKACE – ZADÁNÍ DAT

Instalace aplikace

Připojení k databázi

Název MySQL serveru	<input type="text"/>
Název databáze	<input type="text"/>
Přihlašovací jméno	<input type="text"/>
Heslo	<input type="text"/>
Prefix tabulek	<input type="text"/>

Konstanty na stránkách

Doména	<input type="text"/>
Vyhledávačů	<input type="text"/>
Jiných serverů	<input type="text"/>
Frází	<input type="text"/>
Adres z jichých serverů	<input type="text"/>
Dnů nazpět	<input type="text"/>

PŘÍLOHA P II: 2. INSTALAČNÍ STRÁNKA APLIKACE – OVĚŘENÍ ÚDAJŮ

Instalace aplikace

Ještě jednou si prosím zkontrolujte všechny níže uvedené údaje. Po potvrzení je bude možné změnit pouze ručním zásahem do konfiguračního souboru. Pro případnou úpravu stiskněte tlačítko zpět ve vašem prohlížeči.

Spuštění analýzy na stránkách:

Provedte vložení následujícího zdrojového kódu do vaší hlavní stránky s php (index.php):

```
include_once "../analiza_navstevnosti_z_vyhledavacu/nastaveni.php";  
include_once "../analiza_navstevnosti_z_vyhledavacu/classes/ez_sql_mysql.class.php";  
include_once "../analiza_navstevnosti_z_vyhledavacu/classes/detekce_vyhledavace.class.php";  
include_once "../analiza_navstevnosti_z_vyhledavacu/classes/prace_s_url.class.php";  
$detekce = new Detekce_vyhledavace ();
```

Připojení k databázi

Název MySQL serveru server

Název databáze database

Přihlašovací jméno login

Heslo heslo

Prefix tabulek prefix_

Konstanty na stránkách

Doména www.domena.cz

Vyhledávačů 5

Jiných serverů 10

Frází 15

Adres z jejich serverů 20

Dnů nazpět 30

Potvrdit

PŘÍLOHA P III: ÚVODNÍ STRÁNKA APLIKACE

Souhrnné informace

Návštěvy z vyhledávačů

Vyhledávané řetězce

Návštěvy z jiných adres

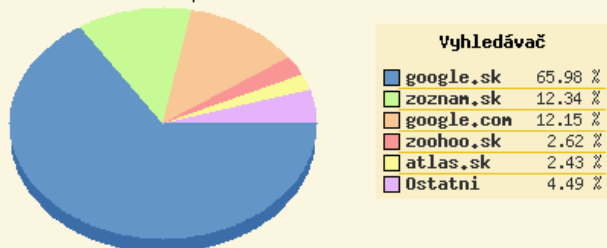
Souhrnné informace

Tato stránka slouží k zobrazení souhrnných informací o návštěvnosti z vyhledávačů za poslední měsíc.

Analyzovaná doména: **www.zetr.sk**

Za posledních 30 dní bylo zaznamenáno 535 návštěv z vyhledávačů.

Návštěvnost z jednotlivých vyhledávačů
za posledních 30 dní



fráze

zetr	[30]
odlučovač ropných látek	[19]
odvodňovacie žľaby	[19]
vodomerná šachta	[18]
pororošť	[11]
svetlik	[11]
poklopy	[8]
odlučovače ropných látek	[8]
podlahy	[7]
stropy	[7]
odvodnovacie zľaby	[6]
svetlik	[6]
poklop	[5]
odlučovače	[5]
kariera	[5]

vyhledávač

google.sk	[353]
zoznam.sk	[66]
google.com	[65]
zoohoo.sk	[14]
atlas.sk	[13]

jiný server

stavebnykatalog.sk - databáze archinfo.sk	[57]
forehand.sk - databáze archinfo.sk	[29]
Žilinská univerzita - stavební fakulta	[4]
techneau.com	[3]
dir.pozri.sk	[2]
firmy.centrum.sk	[2]
hermelock.be	[2]
search.arabia.msn.com	[2]
termiter.cz	[2]
google cache	[2]

PŘÍLOHA P IV: STRÁNKA APLIKACE - NÁVŠTĚVY Z VYHLEDÁVAČŮ

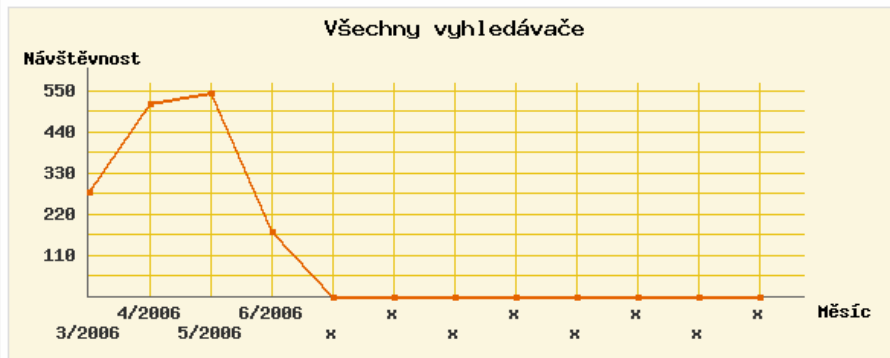
Souhrnné informace

Návštěvy z vyhledávačů

Vyhledávané řetězce

Návštěvy z jiných adres

Návštěvy z vyhledávačů



server: google.sk

adresa: google.sk

nejvyhledávanější fráze za měsíc 5/2006

zetr	[30]
odvodňovacie žľaby	[28]
vodomerná šachta	[17]
poklopy	[11]
pororošt	[10]
odlučovač ropných látok	[9]
odlučovače ropných látok	[8]
podlahy	[8]
poklop	[8]
svetlík	[8]
odvodnovacie zľaby	[6]
bukocel	[5]
hermelock	[5]
svetlík	[5]
POKLOPY	[5]

detailnější statistiky pro jednotlivé měsíce

měsíc 5/2006

PŘÍLOHA P V: STRÁNKA APLIKACE - VYHLEDÁVANÉ ŘETĚZCE

Souhrnné informace | **Návštěvy z vyhledávačů** | Vyhledávané řetězce | Návštěvy z jiných adres

Vyhledávané řetězce

detailnější statistiky pro jednotlivé měsíce

měsíc -- posledních 30 dní -- [zobrazit detail](#)

zetr [30]

nejvyhledávanější fráze za posledních 30 dní

google.sk	14
google.com	7
zoznam.sk	5
google - francie	2
msn.com	1

odlučovač ropných látek [19]

nejvyhledávanější fráze za posledních 30 dní

google.sk	13
google.com	4
zoznam.sk	2

odvodňovacie žľaby [19]

nejvyhledávanější fráze za posledních 30 dní

google.sk	15
zoznam.sk	2
altavista.com	1
google.com	1

vodomerná šachta [18]

nejvyhledávanější fráze za posledních 30 dní

google.sk	11
zoznam.sk	5
google.com	1
msn.com	1

detailnější statistiky pro jednotlivé měsíce

měsíc -- posledních 30 dní -- [zobrazit detail](#)

PŘÍLOHA P VI: STRÁNKA APLIKACE - NÁVŠTĚVY Z JINÝCH ADRES

Souhrnné informace	Návštěvy z vyhledávačů	Vyhledávané řetězce	Návštěvy z jiných adres
Návštěvy z jiných adres			
dir.pozri.sk [2]			
nejaktivnější odkazy posledních 30 dní			
/index.php?od=0&id=f017&znak=z /aktuality/ [2]			
firmy.centrum.sk [2]			
nejaktivnější odkazy posledních 30 dní			
/search/index.php?op=domain&did=500406 /aktuality/ [2]			
hermelock.be [2]			
nejaktivnější odkazy posledních 30 dní			
/www/official_dealers%20En.htm /aktuality/ [2]			
termiter.cz [2]			
nejaktivnější odkazy posledních 30 dní			
/ /aktuality/ [2]			
google cache [2]			
nejaktivnější odkazy posledních 30 dní			
/search?q=cache:Pb9ckgOgD5UJ:www.vsevedko.sk/zoznam-stavebne-m... /aktuality/ [1]			
/search?q=cache:oUEgeK2_P6EJ:www.techneau.sk/2%2555CORL_11.htm+... /aktuality/ [1]			