

Design and implementation of a social network for making acquaintances

Obhajoba bakalářské práce

Marek Bryša

Masarykova Univerzita
Fakulta informatiky

22. června 2012

- Prozkoumat existující společenské sítě pro vytváření známostí a vyhodnotit jejich klady a zápory
- Shrnout relevantní aspekty ochrany osobních údajů v těchto sítích
- Narhnout novou společenskou síť
- Zvolit nejvhodnější technologie pro implementaci
- Implementovat prototyp
- Popsat implementaci s uživatelského a programátorského pohledu

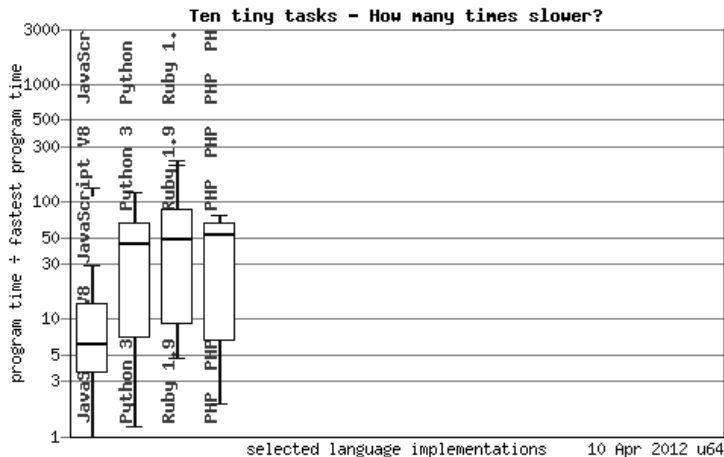
PlentyofFish, Match.com, OkCupid, eHarmony,...

Společné vlastnosti:

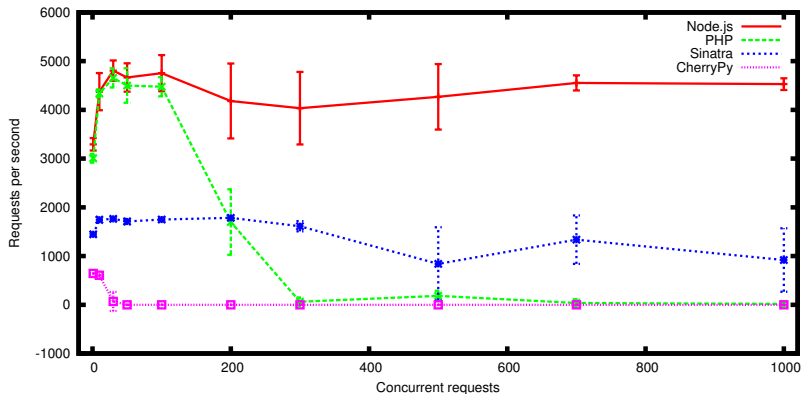
- Nutno vyplňovat zdlouhavé dotazníky citlivými (časti i intimními) osobními údaji
⇒ Možnost zneužití nebo úniku těchto údajů
- Matematický nebo nějak "záračný" systém na automatické vyhledání vhodného partnera
Podle Finkel et al. (2012) tyto systémy nejsou úspěšné
- Nutno sebrat odvahu a iniciativně kontaktovat nalezeného partnera

Návrh nové společenské sítě

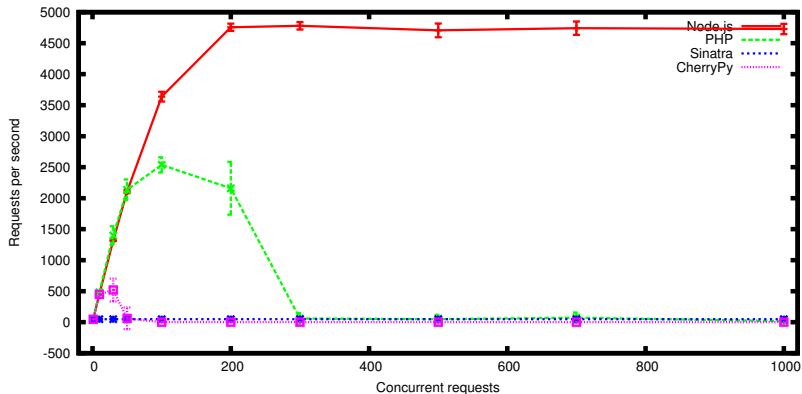
- 1 Uživatel poskytne pouze tyto údaje:
rok narození, pohlaví, přibližná poloha na úrovni okresu, jedna fotografie
- 2 Ostatní uživatele vyhledá podle kritérií:
pohlaví, věkové rozpětí, relativní poloha (stejný okres, sousední okresy, stejný kraj, sousední kraje, stejný stát)
- 3 Prohlíží si fotografie nalezených uživatelů jednu po druhé a označuje ty, kteří se mu líbí
- 4 Pouze až se dva uživatelé označí, že se sobě vzájemně líbí, oba se o tom dozví a je jim nabídnuta možnost real-time chatu, kde se mohou dále poznat



Obrázek: Benchmark jazyků V8 Engine, Python 3, Ruby 1.9 and PHP 5.4.0. Menší znamená rychlejší. Zdroj: The Computer Language Benchmarks Game



Obrázek: Benchmark metody add.




Obrázek: Benchmark metody sleep.

Podle výkonu a dalších kritérií (celková vhodnost, stabilita, škálovatelnost, faktor inovace,...) byl jako základ zvolen framework Node.js a jeho HTTP nádstavbový modul Express.js.


Jako úložiště dat byl zejména z důvodu rychlosti a inovativnosti zvolen Redis (NoSQL in-memory pokročilá key-value storage)

- Vzhledem ke zvolené technologii primární zaměření na výkon
Kompletní zpracování požadavku vyhledání v souboru 200 000 uživatelů trvá na starším serveru 2 ms
Drobné zápisy do databáze (např. označení líbí/nelíbí) do 1 ms
- HTML5, CSS, kompletně AJAX – domovská stránka uživatele se načítá pouze jednou po přihlášení, další požadavky jsou asynchronní
- Možnost ořezání nahrané fotky přímo na serveru
- Real-time chat pomocí socket.io (XHR long-poll i WebSocket), e-mailové upozornění na zmeškané zprávy
- Geolokační databáze Yahoo! GeoPlanet
- Plně funkční implementace dostupná na <http://pickdo.com>

 **Pickdo**

Help mode enabled testbp1@pickdo.com Logout

Pick Selection Chat Me


Edit

testbp1@pickdo.com

Male born 1975 [Change password...](#)

Old password

New password

Retype password

Change

Brno-Město, Jihomoravský, CZ [Change...](#)

Country

State

County

Apply

Looking for women [Change...](#)

Your account is **enabled**. Other people may pick you.

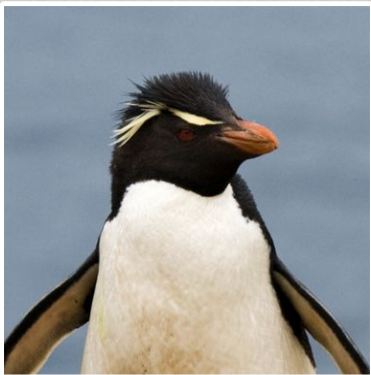
Obrázek: Stránka profilu a nastavení. (Author of the penguin photo: Hannes Grobe/AWI, CC-BY-SA 3.0)

Screenshoty

Age from to

Area:

- ☒ **My county**
- ☐ **Neighbouring counties**
- ☐ **My state**
- ☐ **Neighbouring states**
- ☐ **My country**



Obrázek: Stránka hledání a výběru. (Author of the penguin photo: Samuel Blanc, CC-BY-SA 3.0)



Obrázek: Stránka real-time chatu

- Ruby a EventMachine

Smyslem benchmarku v metodě `sleep` bylo simulovat požadavek na databázi. V Ruby jsou moduly pro asynchronní přístup k databázi buď nestabilní, nebo původním vývojářem opuštěné.

- PHP vs. Ruby jako nejstabilnější jazyk

Uznávám, že by bylo přesnější zůstat v případě PHP u označení "léty prověřený". Přesto se nedomnívám, že by v PHP v budoucnu došlo k radikálním změnám, které se přeci jen v Ruby ještě plánují.

- Napojení na Facebook atd.

Není implementováno záměrně, protože všechny tyto sítě poskytují při propojení jako absolutní minimum údajů skutečné jméno, které v této společenské síti vůbec není vyžadováno a z pohledu uživatele je každé takové propojení rizikem pro únik osobních údajů.

- Cloud hosting Hosting na Heroku by byl pohodlný, ale poněkud drahý a především zatím používá výhradně Amazon Web Services data centra v USA, kde latence z ČR dosahuje přes 200 ms, což už je znát.

Děkuji za pozornost!