

A4B99RPH: Řešení problémů a hry
2. úloha: Spam filtr

Petr Pošík

Spam filtr	2
SPAM?	3
Ne, vážně, co to tedy ten SPAM je?	4
SPAM nebo HAM? (1)	5
SPAM nebo HAM? (2)	6
Tak do třetice... ..	7
Spam filtr	8
Filtrování spamu.	9
Cíle, organizace a náplň	10
Spam filtr jako 2. úloha?	11
Co budeme dělat?	12
Hodnocení filtru	13
Dotazy?	14

SPAM?

Kdy?

- 1937

Kdo?

- Hormel Foods Corporation

Co?

- "Lančmít"!!!!



Význam acronymu SPAM:

- *Shoulder of Pork And Ham* nebo *Spiced Ham*
- "Specially Processed American Meats" (USA nabídly SPAM Anglii během 2. sv. války)
- "Something Posing As Meat"
- "Specially Processed Artificial Meat"
- "Stuff, Pork and Ham"
- "Special Product of Austin Minnesota"

P. Pošík © 2014

A4B99RPH: Řešení problémů a hry – 3 / 14

Ne, vážně, co to tedy ten SPAM je?

70. léta:

- v USENET skupinách se začala objevovat sdělení, která nikoho nezajímala
- uživatelům situace připomněla scénku **Monty Python's Flying Circus**:
 - k vidění např. na youtube.com
 - manželský pár si chce dát v bistru snídani
 - obsluha jim čte jídelní lístek:
 - egg and bacon
 - egg, sausage and bacon
 - egg and spam
 - egg, bacon and spam
 - egg, bacon, sausage and spam
 - spam, bacon, sausage and spam
 - spam, egg, spam, spam, bacon and spam
 - manželka nechce jídlo obsahující SPAM
 - obsluha i manžel jí přesto stále SPAM vnucují
 - "I DON'T LIKE SPAM!!!!"
- spam, spam, spam, egg and spam
- spam, spam, spam, spam, spam, spam, baked beans, spam, spam, spam and spam
- lobster thermidor aux crevettes, with a mornay sauce garnished with truffle paté, brandy and a fried egg on top and spam
- SPAM - nezajímavé, nechtěné, vnucované, opakující se zprávy
- HAM - korektní zprávy týkající se tématu

P. Pošík © 2014

A4B99RPH: Řešení problémů a hry – 4 / 14

SPAM nebo HAM? (1)

As seen on NBC, CBS, CNN, and even Oprah! The health discovery that actually reverses aging while burning fat, without dieting or exercise! This proven discovery has even been reported on by the New England Journal of Medicine. Forget aging and dieting forever! And it's Guaranteed!

Click here: <http://web.kuhleersparnis.ch/hgh/index.html>

Would you like to lose weight while you sleep!?

No dieting!

No hunger pains!

No Cravings!

No strenuous exercise!

Change your life forever!

100% GUARANTEED!

1.Body Fat Loss	82% improvement.
2.Wrinkle Reduction	61% improvement.
3.Energy Level	84% improvement.
4.Muscle Strength	88% improvement.
5.Sexual Potency	75% improvement.
6.Emotional Stability	67% improvement.
7.Memory	62% improvement.

P. Pošík © 2014

A4B99RPH: Řešení problémů a hry – 5 / 14

SPAM nebo HAM? (2)

Filled with useful examples and the depth, clarity, and attention to detail that made the first edition so popular with web developers, the just-released "JavaServer Pages, 2nd Edition" (Bergsten, \$44.95) is completely revised and updated to cover the substantial changes in the 1.2 version of the JSP specifications, and includes coverage of the new JSTL Tag libraries--an eagerly anticipated standard set of JSP elements for the tasks needed in most JSP applications, as well as thorough coverage of Custom Tag Libraries.

What people said about the first edition:

"an excellent printed resource on JSPs...I have been extremely impressed by its depth, clarity, and attention to detail. "

--Reuven M. Lerner, Linux Journal

"This is a great book: it was written by a key contributor not only to the JSP specification, but also to the JSP and Servlet reference implementations. Filled with useful examples, it stands as an important text in the adoption of JSP in the market."

--Eduardo Pelegri-Llopert, lead JSP Specification Engineer.

To order your copy or for more information, see:

<http://www.oreilly.com/catalog/jservletpages2/?CMP=7337>

or call 1-800-998-9938

or email orders@oreilly.com

P. Pošík © 2014

A4B99RPH: Řešení problémů a hry – 6 / 14

Tak do třetice...

Co je podle vás spam? Jak byste jej definovali?

- Nevyžádaná pošta?
- Nevyžádaná komerční pošta?
- Hromadná nevyžádaná komerční pošta?
- Spam je *nevyžádaná zpráva hromadně rozesílaná mnoha lidem bez rozdílu, velmi často za komerčním účelem.*
 - definice je vágní, nemusí vyhovovat všem

Další otázky o spamu:

1. Označili by všichni lidé za spam ty samé emaily?
2. Kdyby stejnému člověku přišla stejná sada emailů v současnosti a za pět let, označil by za spamy stejné emaily?

Spam filtr

Co je podle vás spam filtr?

- Automat, který každému emailu umí přiřadit štítek SPAM nebo HAM.

Jaké má vstupy a jaké má výstupy?

- Vstup: všechny informace o emailu, které můžeme získat (tělo emailu, přílohy, hlavičky). Velmi často nestrukturovaný text.
- Výstup: štítek SPAM nebo HAM

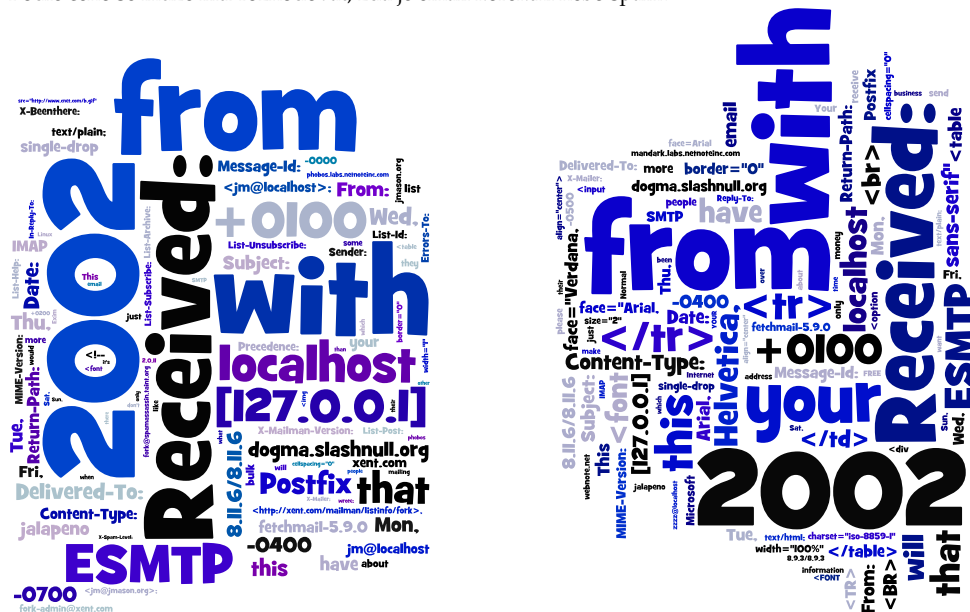
Další otázky o spam filtru:

1. Bude jeden a tentýž spam filtr vyhovovat všem lidem?
2. Co musí spam filtr umět, aby si udržel svou efektivitu v průběhu času?
3. Co musí umět, aby byl efektivní pro různé lidi?

Filtrování spamu

Další souvislosti, které bude třeba vyjasnit v průběhu řešení úlohy:

- Jaké jsou rysy spamu? Pomůže nám definice rozpoznat spam?
- Podle čeho se může filtr rozhodovat, zda je email korektní nebo spam?



- Může se spam filtr dopustit nějakých chyb? Jakých?
- Budou pro vás všechny typy chyb stejně špatné, nebo se jednoho druhu chyb budete obávat více než jiných?
- Kdy je jeden filtr lepší než druhý?

Cíle, organizace a náplň

10 / 14

Spam filtr jako 2. úloha?

Proč byl spam filtr vybrán jako 2. úloha?

- Praktická netriviální úloha, přesto řešitelná s našimi znalostmi
- Ukázka základních principů a problémů strojového učení:
 - datová sada a její vlastnosti
 - učení a testování klasifikátoru (spam filtru)
 - přeučení

Co mi tato úloha přinese?

- Zajímá mě, jak se může program učit z dat — strojové učení.
- Zajímá mě, jak spam filtry pracují.
- Chci si prohloubit znalosti Pythonu.
- Zajímám se o strojové zpracování textů.

Co budeme dělat?

Plán činností:

- Na práci na spam filtru máme cca 4 týdny
- 2 týdny: práce na "zázemí", které nám umožní filtry
 - jednoduše psát,
 - snadno testovat a
 - porovnávat a
 - vyjasnit si, jak že vlastně ten spam filtr bude vypadat.
- 2 týdny: práce na vlastním filtru
 - nepožadujeme super-anti-kontra-multi-extra-unikátní filtry; metody pro jejich tvorbu si osvojíte v dalším průběhu studia
 - cíl: vyzkoušíte si nějakou vaši jednoduchou metodu, která podle vás bude na filtrování spamu alespoň trochu fungovat

Co od nás dostanete:

- Unit testy k jednotlivým dílčím podúlohám

Co budete odevzdávat?

- Implementaci funkce pro hodnocení kvality filtru
- Implementaci vašeho spam filtru

Hodnocení filtru

Složky:

- nápad
- čistota kódu
- účinnost filtru na skutečných datech
 - V minulých letech:
 - SOUTĚŽ!
 - Filtry seřazeny podle dosaženého skóre
 - Prvních 20 % dostalo 4 body, druhých 20 % dostalo 3 body, ..., posledních 20 % dostalo 0 bodů.
 - Studentům vadilo, že jejich hodnocení závisí na úspěšnosti jejich soupeřů.
 - Letos:
 - Pevně stanovené hranice účinnosti:
 - více než 0.9: 4 body
 - více než 0.7: 3 body
 - více než 0.5: 2 body
 - více než 0.3: 1 bod
 - méně než 0.3: 0 bodů

