

Komunitná katalogizácia roztrúsených informácií



Martin Kováčik (mato.kovacik@gmail.com)

<http://student.fiit.stuba.sk/~kovacic03>

Ciele a motivácia projektu

- ◆ Predstavenie otvoreného webového kolaboratívneho prostredia - Wiki
- ◆ Analýzy a zaradenie problémov spojených s otvorenosťou, identifikácia spoločných znakov
- ◆ Návrh ochranných mechanizmov aplikovateľných v prostredí wiki
- ◆ Návrh postupov na hodnotenie relevantnosti stránok vo wiki

Plán

- ◆ **Otvorená webová kolaborácia – Wiki**
- ◆ *Problémy s otvorenosťou Wiki*
- ◆ *Požiadavky a charakteristika ochranných mechanizmov*
- ◆ *Vybrané návrhy riešenia problémov*
- ◆ *Zhodnotenie*
- ◆ *Ďalšie smerovanie*

Otvorená webová kolaborácia

Wiki

WikiWiki = rýchlo

Wiki:

- ◆ Voľne rozšíriteľná kolekcia prepojených webových stránok.
- ◆ Hypertextový systém na ukladanie a modifikáciu informácií – databáza informácií a ich prepojení.

(Leuf, Bo: The WikiWay)

Wiki - Vlastnosti

- ◆ Otvorené
 - *Hocikto môže editovať a pridávať stránky do wiki*
- ◆ Organické
 - *Štruktúra a forma sa môže meniť. Prepojenia sú súčasťou stránky*
- ◆ Kovergujúce
 - *Vytváranie kompromisu medzi jednotlivými príspevkami – vytváranie názorovo neutrálnej stránky*
- ◆ „Dôveryhodné“
 - *Veľké množstvo participantov*

Plán

- ◆ *Otvorená webová kolaborácia – Wiki*
- ◆ **Problémy s otvorenosťou Wiki**
- ◆ *Požiadavky a charakteristika ochranných mechanizmov*
- ◆ *Vybrané návrhy riešenia problémov*
- ◆ *Zhodnotenie*
- ◆ *Ďalšie smerovanie*

Problémy

- ◆ Narastá množstvo informácií a je potrebná uľahčiť rozpoznanie a zabránenie vkladania nevhodných informácií.
- ◆ Súčasné riešenia sú silne integrované s komunitou.
- ◆ Menšie wiki majú menej informácií na stráženie, ale aj menej ľudí, ktorí ich sledujú
- ◆ Veľké wiki majú veľké množstvo informácií a je problém ich ustriehnuť.

Typy útokov

- ◆ Programové útoky
Rýchly sled zmien od jedného používateľa, globálny útok na celé wiki, kombinácie viacerých typov útokov
- ◆ Mazanie obsahu
- ◆ Spam
Explicitný spam, Vkladanie liniek za cieľom zlepšenia hodnotenia prehliadačom.
- ◆ Vkladanie nerelevantných informácií
Pripisovanie zásluh, šírenie klebiet...
- ◆ Rušenie zmien „Edit wars“
Vzájomné rušenie zmien ak sa autori nezhodnú na danej informácii, častý návrat k pôvodnej verzii informácie

Plán

- ◆ *Otvorená webová kolaborácia – Wiki*
- ◆ *Problémy s otvorenosťou Wiki*
- ◆ **Požiadavky a charakteristika ochranných mechanizmov**
- ◆ *Vybrané návrhy riešenia problémov*
- ◆ *Zhodnotenie*
- ◆ *Ďalšie smerovanie*

Existujúce ochranné mechanizmy

- ◆ Autentifikácia a autorizácia používateľa
- ◆ Iné mechanizmy len podporujú rozpoznanie a riešenie problémov používateľom
 - ◆ *Možnosť vrátiť sa ku ktorejkoľvek staršej verzii stránky*
 - ◆ *Rýchle náhľady zmien na stránke, zobrazenie diferencie*
 - ◆ *Sumarizácia zmien asociovaných s používateľom, kategóriou, stránkou, ...*

Problémy pri používaní ochranných mechanizmov

- ◆ Označenie nesprávnych zmien
- ◆ Obmedzovanie používateľov

Dôsledky:

- ◆ Používatelia strácajú motiváciu prispievať
- ◆ Zníženie aktivity smeruje k pozastaveniu vývoja informácií a tým sa obsah stáva nerelevantný a neaktuálny

Činnosti ochranných mechanizmov

- ◆ Rozpoznanie útoku
- ◆ Reakcia na útok
- ◆ Získanie informácií potrebných na určenie útoku
- ◆ Pasívna ochrana pred útokom

Rozpoznanie útoku

- ◆ Identifikácia útoku na základe obsahu informácie
 - *Vyhľadávanie znakov priamo v texte – nevhodné výrazy, linky, mechanizmy používané pri emailovom spame*
- ◆ Identifikácia útoku na základe správania sa používateľov
 - *Rýchly sled zmien, opakované pridávanie rovnakých výrazov, sledy špecifických operácií, časté vracanie autorových zmien*
- ◆ Hybridný prístup
 - *Využívanie znakov získaných počas sledovania správania pri analýze na základe obsahu a naopak*

Reakcia na útok

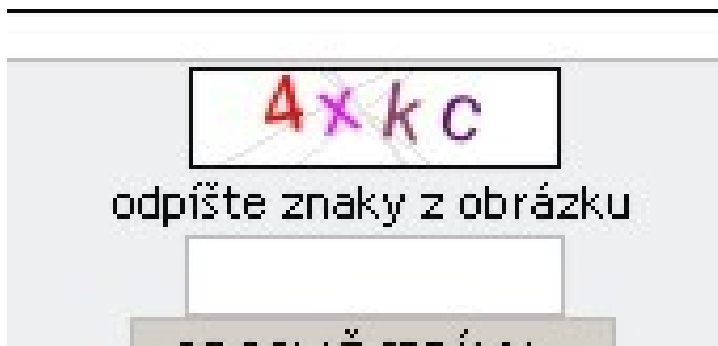
- ◆ Upozornenie administrátora
- ◆ Obmedzenie prístupu
 - ◆ *Určenie pôvodcu – obmedzenie konkrétneho používateľa, IP adresy, siete*
 - ◆ *Obmedzenie operácií – obmedzenie špecifických operácií, obmedzenie editačných operácií, úplné zamedzenie prístupu*
 - ◆ *Trvanie obmedzení – dočasné obmedzenie, trvalé obmedzenie*
- ◆ Obnovenie pôvodnej verzie stránky

Získanie informácií potrebných na určenie útoku

- ◆ Explicitné získanie ohodnotenia stránky od používateľa
 - ◆ *Vhodnosť*
 - ◆ *Ohodnotenie relevantnosti*
- ◆ Získanie informácií z externých zdrojov
 - ◆ *Databázy znakov identifikujúcich útok – výrazy, linky, súbory*
- ◆ Získanie informácií z interných zdrojov
 - ◆ *História stránky*
 - ◆ *História stránok v systéme*

Pasívne ochranné prvky

- ◆ Označovanie externých odkazov pre vyhľadávače – atribút `nofollow`
- ◆ Určenie minimálneho času medzi editáciami stránky a stránok vo wiki
- ◆ CAPTCHA: *Completely automated public Turing test to tell computers and humans apart*
 - ◆ Zavedenie ďalších prvkov
 - ◆ Prepojenie s existujúcimi prvkami



To Save page: Click tree



Plán

- ◆ *Otvorená webová kolaborácia – Wiki*
- ◆ *Problémy s otvorenosťou Wiki*
- ◆ *Požiadavky a charakteristika ochranných mechanizmov*
- ◆ **Vybrané návrhy riešenia problémov**
- ◆ *Zhodnotenie*
- ◆ *Ďalšie smerovanie*

Linkový spam – hybridný prístup

- ◆ Overovanie výskytu nových liniek v posledných aktualizáciách
- ◆ Ak výskyt zhodných liniek v nových príspevkoch prekročil určenú hranicu je linka považovaná za spam

Edit Wars – analýza správania

- ◆ Analýza histórie stránky
- ◆ V prípade, že je stránka často editovaná (nová verzia / revert) je editácia zakázaná
- ◆ Editácia je po určitom čase opäť sprístupnená
- ◆ Je ťažké určiť vinníka, pri opatreniach proti vinníkovi musí zasiahnuť administrátor

Hodnotenie relevantnosti

- ◆ Súčasnú wiki nie sú vhodné na zisťovanie relevantnosti obsahu
- ◆ Cieľová skupina systémov sú *semantické wiki*
- ◆ Prepoklad je semantická podobnosť pôvodnej informácie s novou verziou.
- ◆ Stránky bez určenia semantického významu:
Využitie verejných databáz na určenie semantickej podobnosti

Využitie verejných databáz na určenie semantickej podobnosti

- ◆ Využívané databázy: *Fulltextové vyhľadávače*
- ◆ Určenie znakov určujúcich semantický kontext:
Kľúčové slová a výrazy
 - *extrakcia pomocou algoritmov používaných vo fulltextových vyhľadávačoch*
 - *štatistická extrakcia*
 - *iné spôsoby extrakcie*
- ◆ Vyhľadanie prepojení s kľúčovými slovami a porovnanie množín pre pôvodnú informáciu a novú

Plán

- ◆ *Otvorená webová kolaborácia – Wiki*
- ◆ *Problémy s otvorenosťou Wiki*
- ◆ *Požiadavky a charakteristika ochranných mechanizmov*
- ◆ *Vybrané návrhy riešenia problémov*
- ◆ **Zhodnotenie**
- ◆ *Ďalšie smerovanie*

Zhodnotenie

Výstup projektu

- ◆ Identifikácia charakteristických črt útokov na wiki pre potrebu ich rozpoznania
- ◆ Návrh postupov na rozpoznanie útoku
- ◆ Návrh postupov na ohodnotenie relevantnosti dokumentu v prostredí bez semantických

Implementovaný systém

- ◆ Jednoduché wiki
- ◆ Implementácia mechanizmov na rozpoznanie explicitného spamu a zabránenie automatizovaným útokom.
- ◆ Podpora pre implementáciu ďalších mechanizmov

Plán

- ◆ *Otvorená webová kolaborácia – Wiki*
- ◆ *Problémy s otvorenosťou Wiki*
- ◆ *Požiadavky a charakteristika ochranných mechanizmov*
- ◆ *Vybrané návrhy riešenia problémov*
- ◆ *Zhodnotenie*
- ◆ **Ďalšie smerovanie**

Ďalšie smerovanie projektu

- ◆ Semantické určovanie podobnosti dokumentov
- ◆ Behaviorálna analýza používateľov v otvorených prostrediach – sledovanie neštandardného správania
- ◆ Semanticky orientované wiki