

Cubit-hakuprotokolla pyrkii syrjäyttämään torrent-sivustot

18.6.2008 13:45 — Hannu Nokso-Koivisto

Cubitiksi nimetty vertaisverkko (p2p) -hakuprotokolla nousee haastamaan torrent-sivustot. Cubitin tavoitteena on tarjota käyttäjille toimiva haku tiedostojenkäyttöön p2p-verkoissa.

[Cornellin yliopiston projektissa](#) pyritään kehittämään pätevä hakutekniikka, joka pystyisi käsittelemään epätarkkoja merkkijonohakuja tiedostojenkäyttö-ohjelmissa.

Cubitin yksi kehittäjistä, **Bernard Wong**, kertoo [Ars Technica](#)-sivustolla projektin motivaatioksi kirjoitusvirheet sekä hauissa, että tiedostojen nimissä. Wong esitteli sivustolle esimerkiksi Googlen tilastoja Britney Spears-hauista. Tilastot osoittivat, että laulajan nimi oli 20 prosentissa hauista kirjoitettu väärin.

Hakutulokset koostuvatmerkkijonojen muutoksista

Teknisesti Cubit luo alkuun tunnistetiedot p2p-verkoston kaikille jaossa oleville tiedostoille.

Tietokone luo tämän jälkeen sisäisen kartan tunnistetietojen perusteella. Kartta verkottaa esimerkiksi sanat talo ja valo lähemmäksi, koska niiden merkkijonojen välinen muutos on yksi. Sisäinen kartta sijoittaa sanan valas kohtalaisen lähelle sanaa talo, sillä merkkijonojen välinen muutos on kolme.

Tällöin haun tapahtuessa väärinkirjoitetulla sanalla ralo, tulokset sanoille talo ja valo sijoittuisivat korkealle merkkijonojen välisen muutoksen ollessa yksi.

Sisäisen kartan koontitavan vuoksi hakujen ei tarvitse tarkistaa jokaista verkon tiedostoa.

P2P-ohjelmat sisältävät jo hajautustauluun pohjautuvia hakutoimintoja, mutta ne yleisesti toimivat vain tarkkojen hakusanojen kanssa.

Arvioiva haku ei vielähaasta torrent-sivustoja

Cubitin päivä parrasvaloissa on kuitenkin vielä kaukana, torrent-sivustojen pystyessä vielä tällä hetkellä tarjoamaan paremmin metatietoa ja parempaa organisointia. Cubit pohjaa hakunsa tiedostojen nimiin ja kommenttiosioihin.

Laittoman materiaalin jakamisen tutkimista Cubit ei varsinaisesti helpota tai vaikeuta. Hakuprotokolla pystyy kuitenkin kohdentamaan laittoman sisällön jäljittämisen käyttäjiin sivustojen sijaan.

<http://www.digitoday.fi/data/2008/06/18/cubit-hakuprotokolla-pyrkii-syrjayttamaan-torrent-sivustot/200816518/66>