

NodePieSpy

<https://github.com/tuhoojabotti/NodePieSpy>

1. Sisälllys

[1. Sisälllys](#)

[2. Johdanto](#)

[3. Ominaisuudet](#)

[3.1. Päättelyheuristiikat](#)

[3.1.1. Vierekkäisyys](#)

[3.1.2. Suora puhuttelu](#)

[3.1.3. Epäsuora maininta](#)

[3.5. Analysointialgoritmit](#)

[3.5.1. BFS-puu](#)

[3.5.2. Ryvästäminen](#)

2. Johdanto

NodePieSpy on työkalu, jonka avulla voi analysoida sosiaalisia verkkoja. Ohjelmisto toimii tällä hetkellä vain irssin (<http://irssi.org/>) lokien perusteella, mutta se on tulevaisuudessa laajennettavissa esimerkiksi Facebook- tai Github-suhteiden kanssa toimivaksi.

Tällä hetkellä ohjelmisto koostuu kahdesta komponentista. NodePieSpy on palvelinohjelma, joka hoitaa tiedon lataamisen ja käsittelyn, ja tarjoaa HTTP:n yli rajapinnan, jolla tietoja voi pyytää JSON-formaatissa. NodePieSpy-Client on HTML-sivu, joka käyttää kyseistä rajapintaa ja visualisoi tulokset. Tämä käyttöliittymä ei ole harjoitustyön osana, vaikka sitä tullaankin käyttämään projektin testauksessa ja esittelyssä.

Projektin inspiraationa toimii PieSpy: <http://www.jibble.org/piespy/>

3. Ominaisuudet

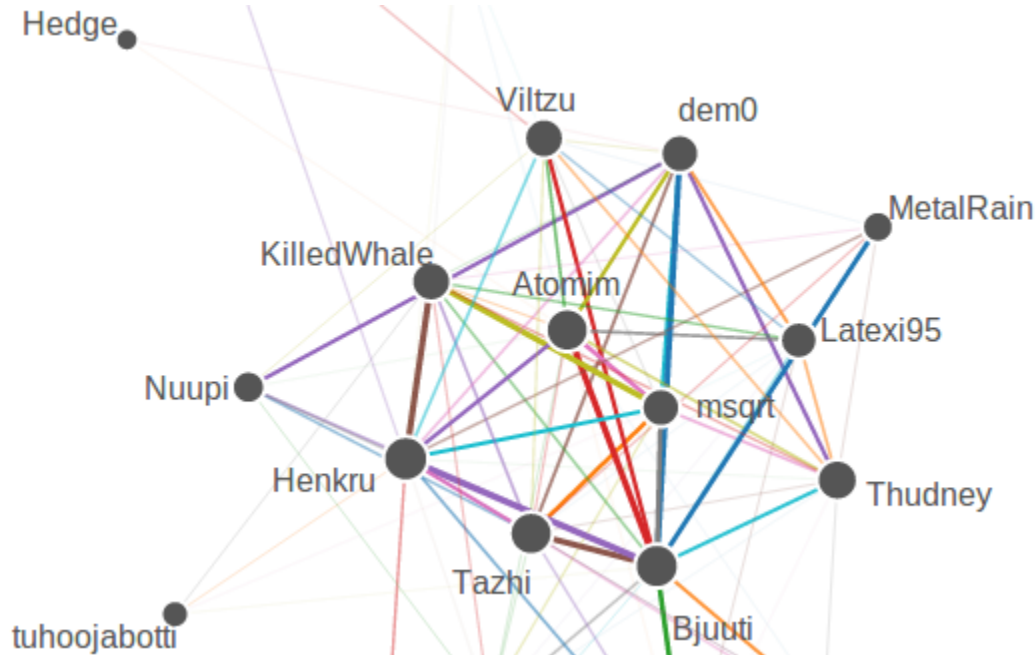
Tässä osiossa selitän miten mitäkin ominaisuutta on testattu.

3.1. Päättelyheuristiikat

Tärkein osa koko ohjelmistossa on itse tietojen jäsentäminen sosiaalseksi verkoksi. Tiedot tulevat Irssi IRC-asiakasohjelman lokitiedostoista.

3.1.1. Vierekkäisyys

Testattu lähinnä silmämääräisesti asiakasohjelmalla ja konsolitulostuksilla.



Yksikkötestaus on työn alla.

3.1.2. Suora puhuttelu

Tekemättä

3.1.3. Epäsuora maininta

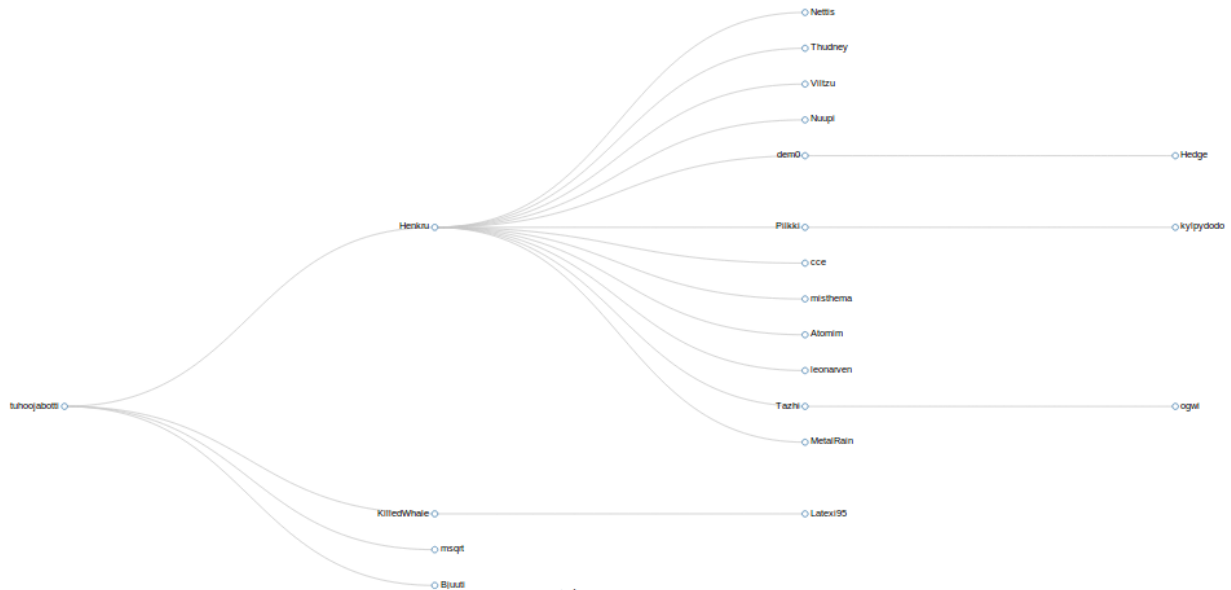
Tekemättä

3.5. Analysointialgoritmit

Minkälaista tietoa tästä verkosta saa sitten irti.

3.5.1. BFS-puu

Puun testaus on myös toteutettu silmämääräisesti ensin HTTP-vastauksia tulkitsemalla ja sitten graafisen implementaation tarkastamisella.



Implementaatiosta löytyikin bugi, joka aiheutti väärän puun generoimisen.

<https://github.com/tuhoojabotti/NodePieSpy/blob/ff705a192fb6cee5ed465de9cd9ea4661083dbb5/lib/algorithms/bfs.js#L25>

visited-taulukon arvon asetus oli alunperin virheellisesti toistorakenteen jälkeen.

Tästä aiheutui se, että jo läpikäytyjä solmuja käytiin läpi useammin ja puun järjestys meni rikki.

3.5.2. Ryvästäminen

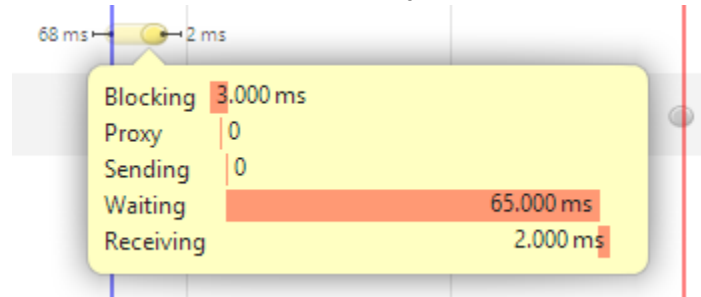
Tekemättä

4. Suoritusteho

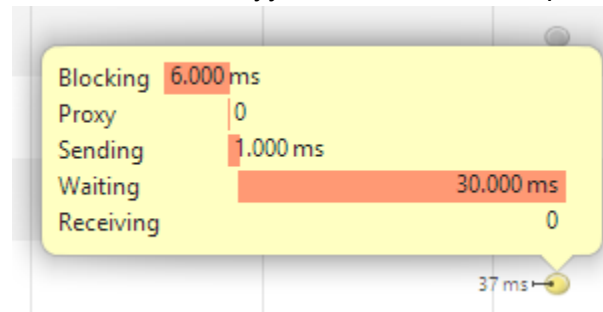
Ohjelman oltua päällä 2 viikkoa se oli kerännyt huomattavat määrät dataa, joka aiheutti käyttöliittymän hidastumista suuremmilla keskustelukanavilla. Kuitenkin itse backend pysyy suorituskyykyisenä. Mikäli ohjelmasta tulee suositumpi, on myös mahdollista ottaa käyttöön aliprosessien käyttö, eli jokaista suoritinydintä kohden luotaisiin oma prosessi.

4.1. rivejä 10 000

Ladataan lokeista 10 000 riviä ja tutkitaan kuinka nopeasti palvelin vastaa:



Kuten kuvasta näkyy, verkon lataaminen palvelimelta ei kestä kovin kauaa.



Samoin BFS:n generoiminen tapahtuu silmänräpäyksessä.