

# NodePieSpy

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

## 1. Sisälllys

- [1. Sisälllys](#)
- [2. Johdanto](#)
- [3. Ominaisuudet](#)
  - [3.1. Päättelyheuristiikat](#)
    - [3.1.1. Vierekkäisyys](#)
  - [3.5. Analysointialgoritmit](#)
    - [3.5.1. BFS-puu](#)
    - [3.5.2. Ryvästäminen](#)
- [4. Puutteet](#)
- [5. Yhteenveto](#)

## 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 määritellään kaikki erilaiset ominaisuudet, mitä sain toteutettua aikataulullisista ongelmista huolimatta.

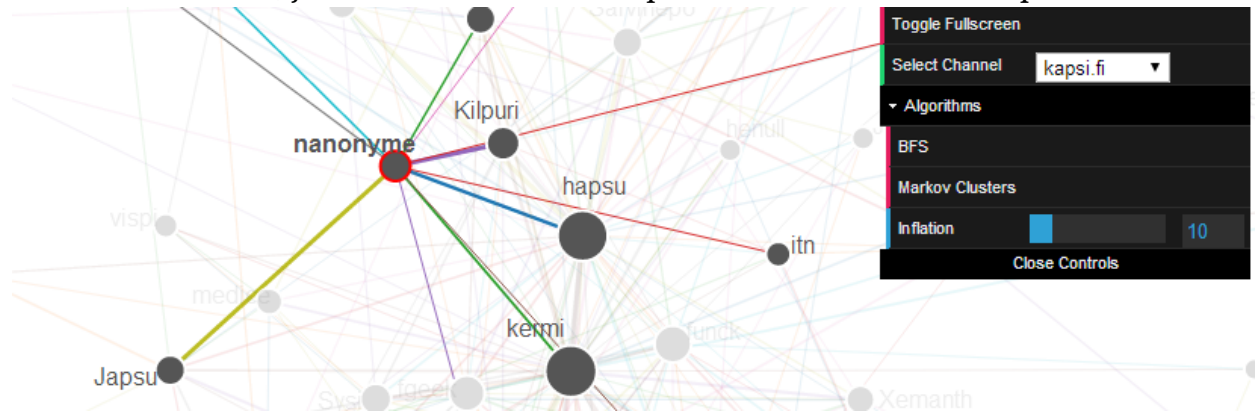
12.1.2014  
Ville Lahdenvuo  
Toteutusdokumentti

### 3.1. Päättelyheuristiikat

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

### 3.1.1. Vierekkäisyys

Peräkkäiset viestit irkissä kuuluvat yleensä yhteen, joskin on myös todennäköistä, että keskustelu vaihtuu, joten tämän heuristiikan painoarvon tulee olla aika pieni.

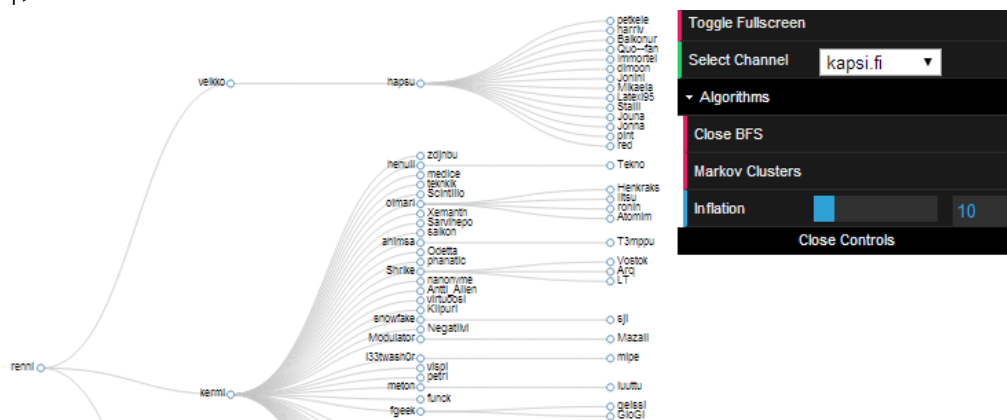


### 3.5. Analysointialgoritmit

Minkälaista tietoa tästä verkosta saa sitten irti.

### 3.5.1. BFS-puu

Generoimalla BFS-puu jostain käyttäjästä, selviää monen henkilön kautta jokaisella irkkaajalla on yhteys muihin kanavalla irkkaaviin. BFS:n generoiminen onnistuu ajassa  $O(|V| + |E|)$ .



### 3.5.2. Ryvästäminen

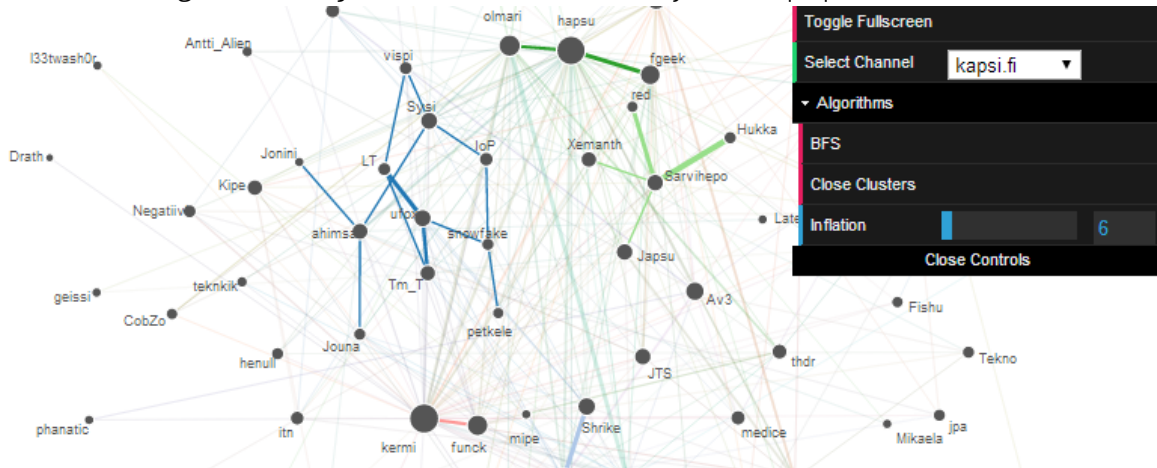
12.1.2014

Ville Lahdenvuo

Toteutusdokumentti

Olisi mukava nähdä millaisia ryhmiä irkkaajat muodostava keskenään. Nämä ongelmat ovat NP-täydellisiä, mutta jokin yksinkertaisempi algoritmi lienee riittävän hyvä, sillä verkon rakennekin perustuu olettamuksiin.

Markov Cluster -algoritmillä ryvästäminen onnistuu ajassa  $O(|V|^3)$



## 4. Puutteet

Aikalailla ne ominaisuudet, mitkä olen kirjoittanut määrittelydokumenttiin.

Lisäksi olisi kiva ehkä helpottaa asennustyötä ja asetusten säätämistä vielä enemmän.

Yksikkötestit jäi toteuttamatta toistaiseksi, mutta algoritmien oikeellisuuden merkitys on pienekkö, kunhan visuaalisesti tulos näyttää hyvältä, sillä verkko on luotu tulkitsemalla ihmisten interaktioita, eikä se ole kovin eksaktia.

## 5. Yhteenveto

Minusta projekti oli menestys, vaikka aikaa ei niin paljon ollutkaan. Tulen varmasti jatkamaan projektin kehitystä tulevaisuudessa.

Minulla on jo muutamia kiinnostuneita osapuolia, jotka tulevat varmaan auttamaan jatkokehityksessä tiralabran jälkeen.