upf. Universitat Pompeu Fabra Barcelona

# VENT

***Express your feelings and connect with people who care***

March 2020

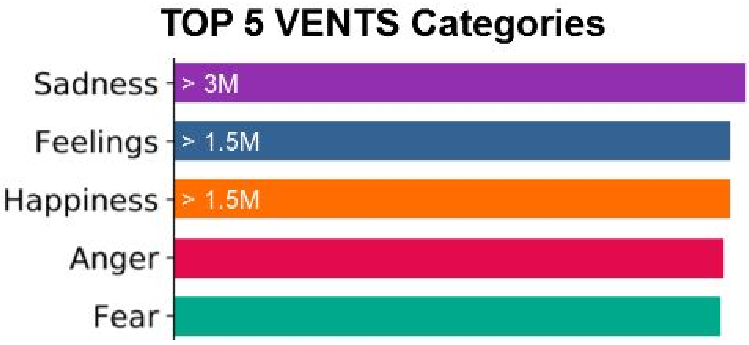
Data-Driven Social Analytics

Jonatan Koren

## Sharing emotions at scale: The Vent dataset

Nikolaos Lykousas, Costantinos Patsakis, Andreas Kaltenbrunner, Vicenç Gómez (2019) [pdf]
Sharing emotions at scale: The Vent dataset. International AAAI Conference on Web and Social Media

- **The paper highlighs several key points:**
  - The largest annotated dataset of text, emotions, and social connections to date (24.03.19)
  - The benefits of emotion analysis beside sentiment analysis
  - Ethical considerations after data collection and analysis
  - Methods to infer emotion networks based on the user activity



# The data

the data contains 4 files:

- **emotions.csv** [ 705 x 4 ]
  - id (string)
  - emotion_category_id (string)
  - name (string)
  - enabled (boolean)

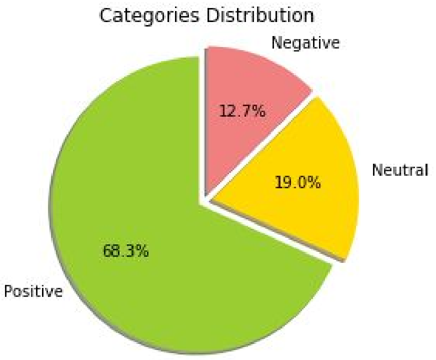
- **emotions_categories.csv** [ 63 x 2 ]
  - id (string)
  - emotion_category_id (string)
  - name (string)
  - enabled (boolean)

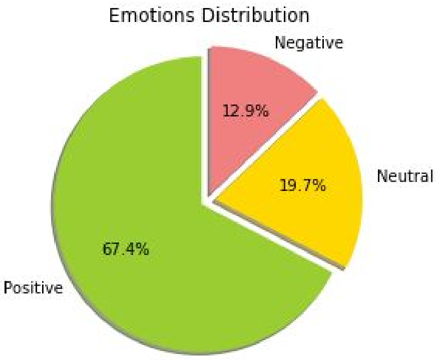
- **vents_metadata.csv** [ 33.6M x 4 ]
    - emotion_id (string)
    - user_id (string)
    - created_at (string YYYY:MM:DD hh:mm:ss.sss)
    - enabled (boolean)

- **edges** [ 13.6M x 2 ]
    - source (string)
    - target (string)

## The following pie charts explain the properties of the data:

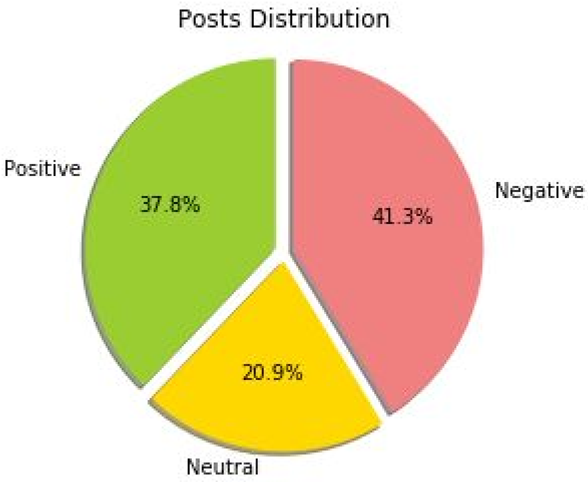
The given emotions distribution by the Vent App:





The given categories and emotions, the sentiment types are distributing pretty the same way. A little bit more than $2/3$ of the emotions and categories are positive, almost $1/5$ are neutral sentiments and the rest are negative sentiments.

Emotions distribution of users posts:



In contrary to the distribution offered by the Vent App, user posts distribution are completely different.

Overall, most posts that are posted are either positive or negative, but slightly more posts with negative emotion (about $4/5$), but only about $1/5$ of the posts are of a neutral emotion.

# Thought

- **User Profiling:**

  → Categorize manually emotions to positive (+1), neutral (0) and negative (-1)

  → Aggregating data and giving scores to nodes according to the mean profile score:

  - positive: x > 0.2
  - neutral: -0.2 < x < 0.2
  - negative: x < -0.2

- **User Impact:**

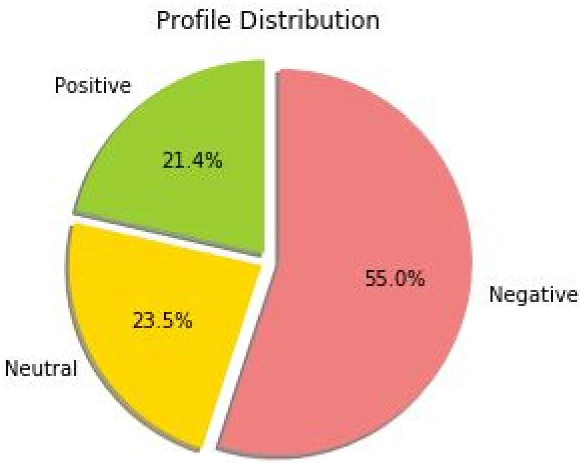
  → Categorize nodes impact by sum aggregation of reactions.

    → According to mean( x ) = 220, median( x ) = 15

    - Top user: x > 330
    - High: 110 < x < 330
    - Soft: 15 < x < 110
    - Low: x < 15

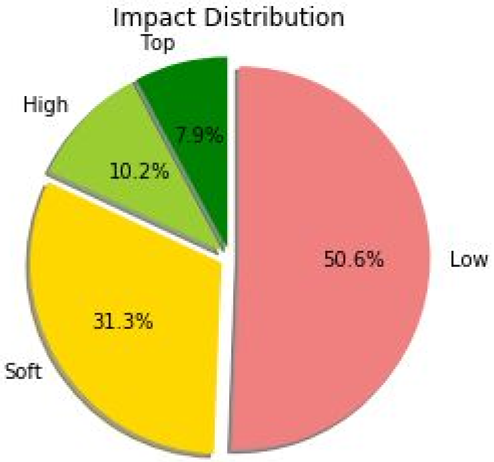
After understanding the data structure i thought to label each emotion category into different category sentiments types: positive, negative and neutral. This allows me to map the different emotions also to these different sentiments types.

General user profiling according to their posts:



More than half of the users are determined as negative users. In one hand surprising because of the emotions distribution that are given by the application and the actual situation, but in the other hand we have already seen in the last pie chart that people are posting more negative posts rather than positive ones despite the numerical closeness between them ( ~40% each ).

General user's post popularity / impact:



About half of the users have low impact / posts popularity. Almost 1/5 of the users have big impact by their posts.

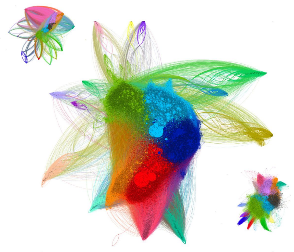
In order to analyze the graph with SNAP, i had to convert the strings of the user_id in the nodes dataset into 'int' types by giving them the index number as according to their id.

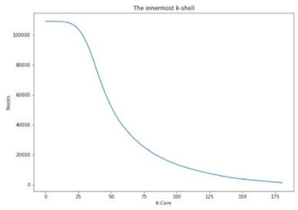
This allowed me to map the users id in the edges dataset into the corresponding numbers in the new nodes id's.

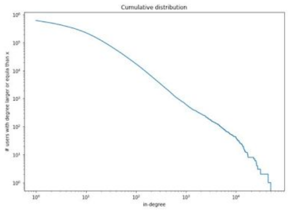
## Giant Component:

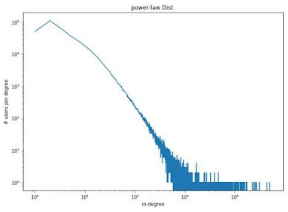
I decided to go a little deeper than the giant component and to stay with nodes that are higher than 25 (in and out degrees) after trying different nodes removal by different degrees.

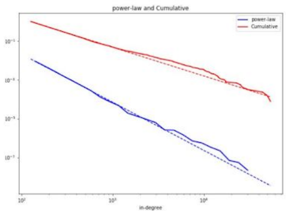
- Staying with nodes with degree > 25
- Nodes: 109,293 | Edges: 6,423,647
- Modularity: 0.444 | 22 Communities
- Triads: 25.5M | Clustering Coefficient: 0.125
- innermost k-shell: 180 (with 954 nodes)



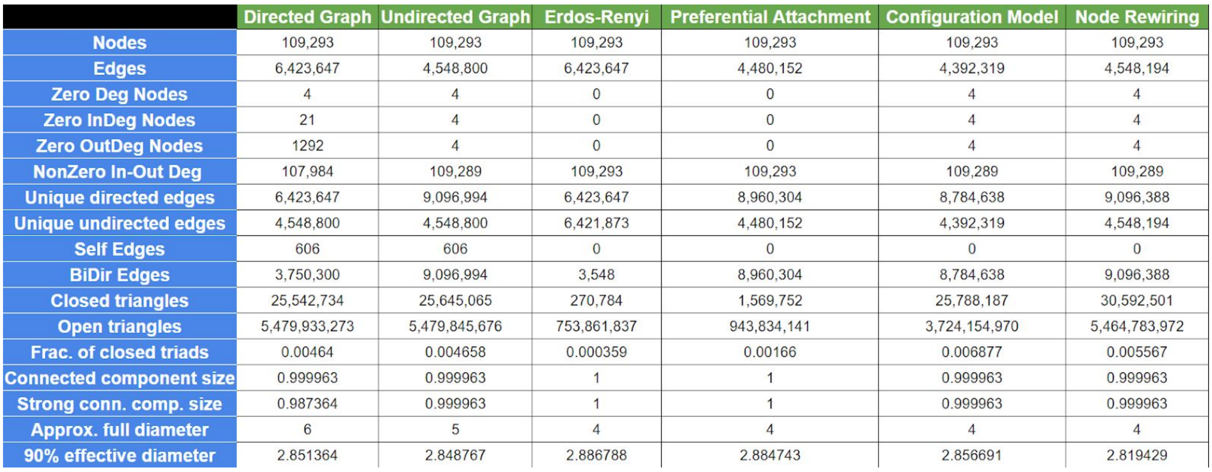








## Comparing to other random graphs:

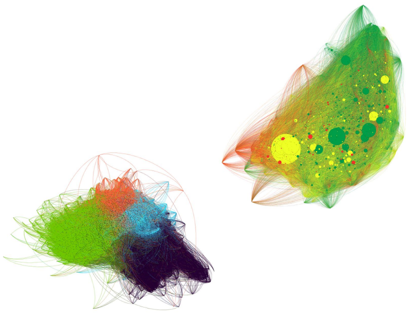


| | Directed Graph | Undirected Graph | Erdos-Renyi | Preferential Attachment | Configuration Model | Node Rewiring |
|---|---|---|---|---|---|---|
| Nodes | 109,293 | 109,293 | 109,293 | 109,293 | 109,293 | 109,293 |
| Edges | 6,423,647 | 4,548,800 | 6,423,647 | 4,480,152 | 4,392,319 | 4,548,194 |
| Zero Deg Nodes | 4 | 4 | 0 | 0 | 4 | 4 |
| Zero InDeg Nodes | 21 | 4 | 0 | 0 | 4 | 4 |
| Zero OutDeg Nodes | 1292 | 4 | 0 | 0 | 4 | 4 |
| NonZero In-Out Deg | 107,984 | 109,289 | 109,293 | 109,293 | 109,289 | 109,289 |
| Unique directed edges | 6,423,647 | 9,096,994 | 6,423,647 | 8,960,304 | 8,784,638 | 9,096,388 |
| Unique undirected edges | 4,548,800 | 4,548,800 | 6,421,873 | 4,480,152 | 4,392,319 | 4,548,194 |
| Self Edges | 606 | 606 | 0 | 0 | 0 | 0 |
| BiDir Edges | 3,750,300 | 9,096,994 | 3,548 | 8,960,304 | 8,784,638 | 9,096,388 |
| Closed triangles | 25,542,734 | 25,645,065 | 270,784 | 1,569,752 | 25,788,187 | 30,592,501 |
| Open triangles | 5,479,933,273 | 5,479,845,676 | 753,861,837 | 943,834,141 | 3,724,154,970 | 5,464,783,972 |
| Frac. of closed triads | 0.00464 | 0.004658 | 0.000359 | 0.00166 | 0.006877 | 0.005567 |
| Connected component size | 0.999963 | 0.999963 | 1 | 1 | 0.999963 | 0.999963 |
| Strong conn. comp. size | 0.987364 | 0.999963 | 1 | 1 | 0.999963 | 0.999963 |
| Approx. full diameter | 6 | 5 | 4 | 4 | 4 | 4 |
| 90% effective diameter | 2.851364 | 2.848767 | 2.886788 | 2.884743 | 2.856691 | 2.819429 |

Graph’s Core:

I decided even more deeper than before and to stay with nodes that are higher than 200 (in and out degrees) that are the core of the VENT application.

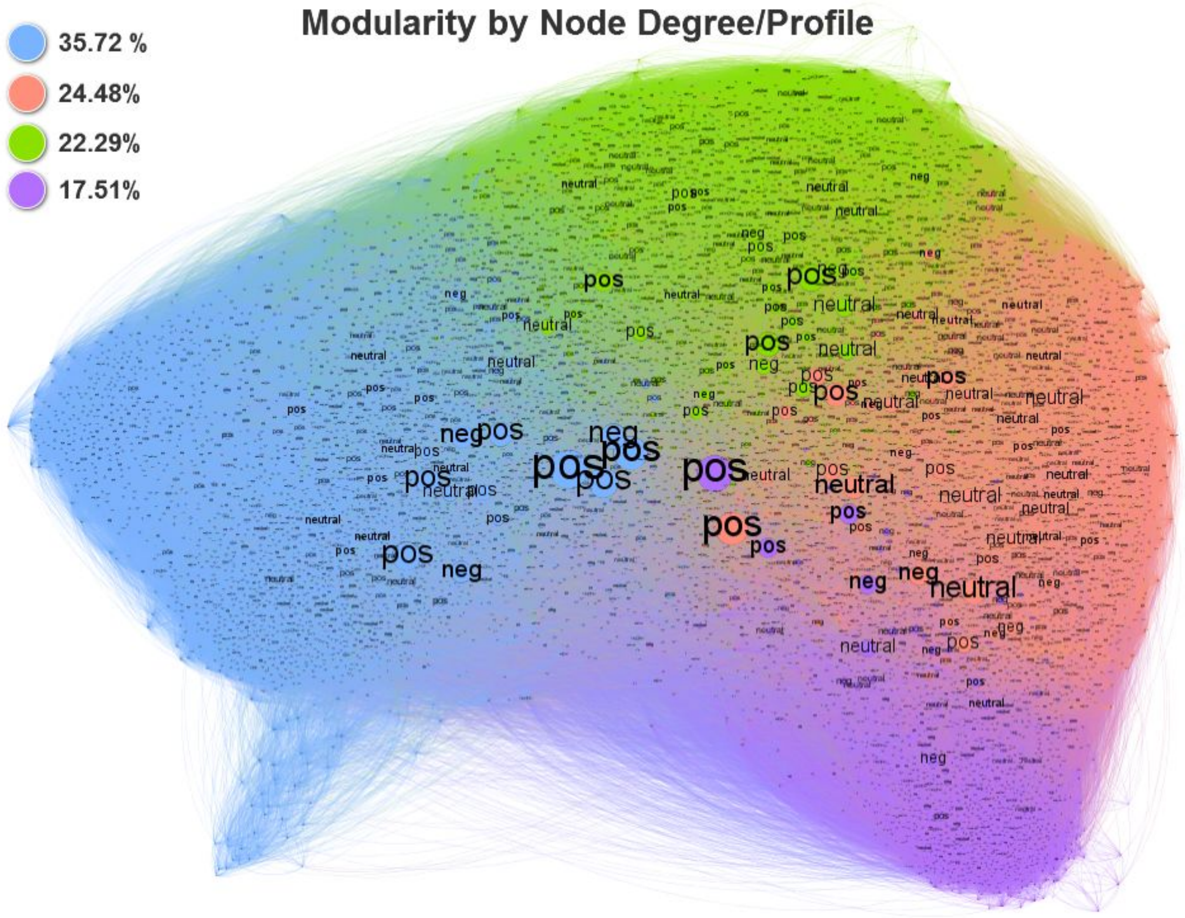
- Staying with nodes with degree > 200
- Nodes: 7,852 | Edges: 1,018,120
- Modularity: 0.326 | 4 Communities
- Triads: 10.01M | Clustering Coefficient: 0.170
- innermost k-shell: 180 (with 898 nodes)

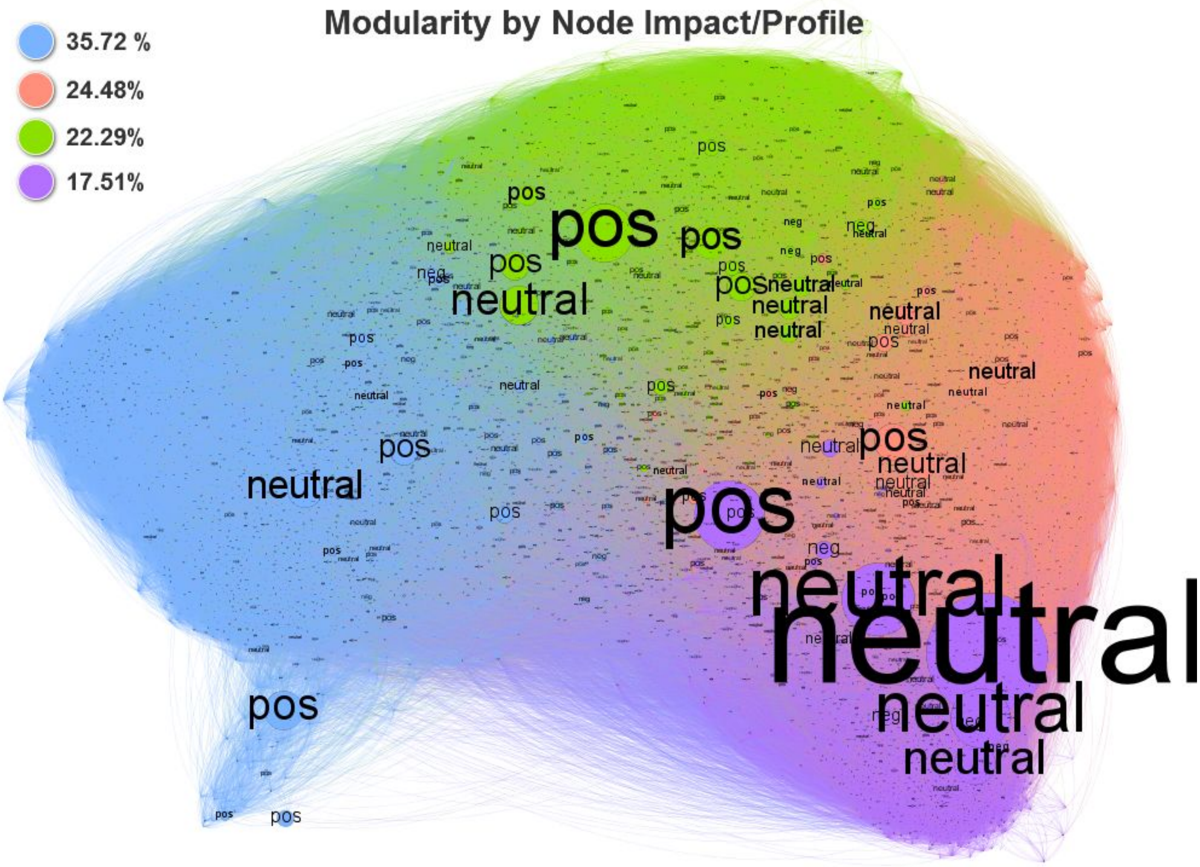


# Modularity

Modularity = 0.326

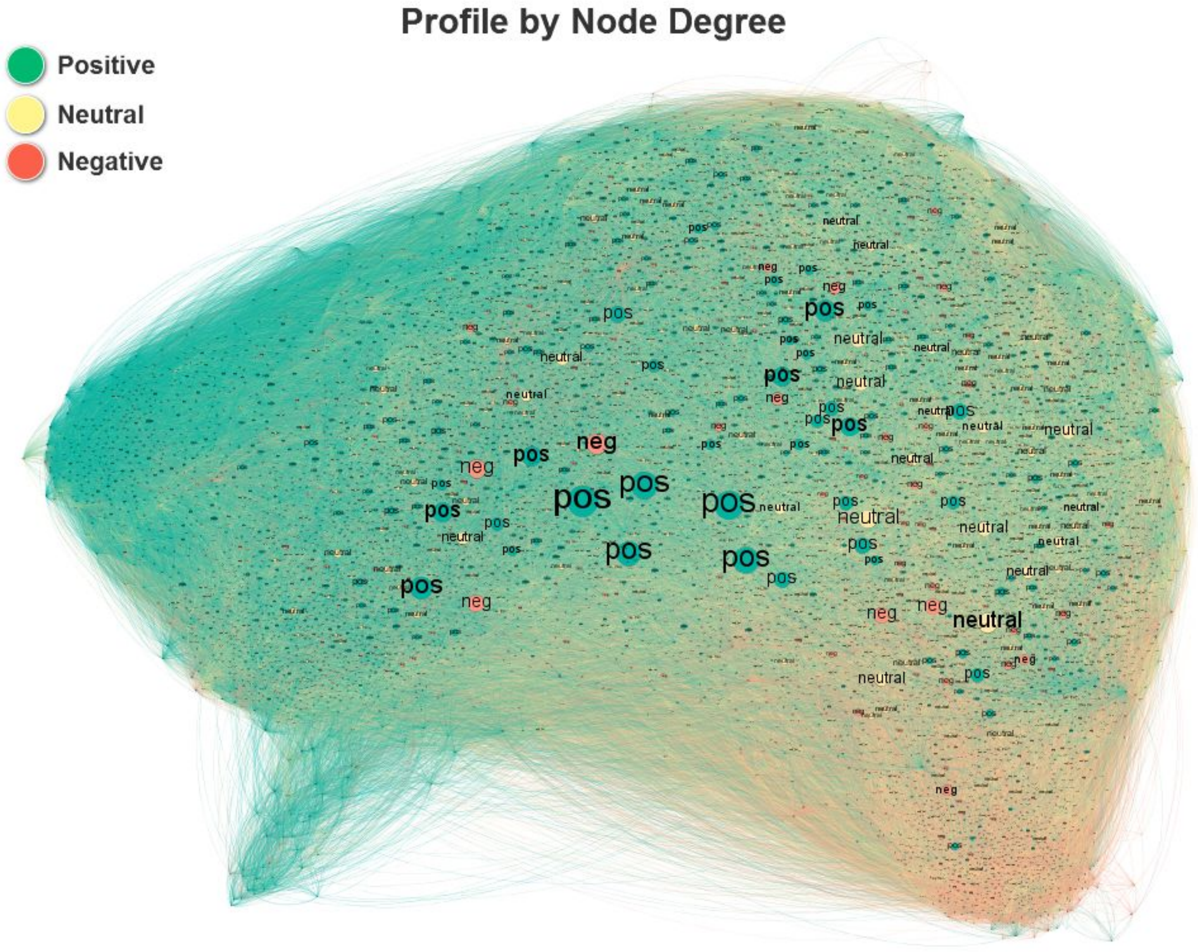
Highly degree nodes, separated into 4 communities using the louvain method for community detection



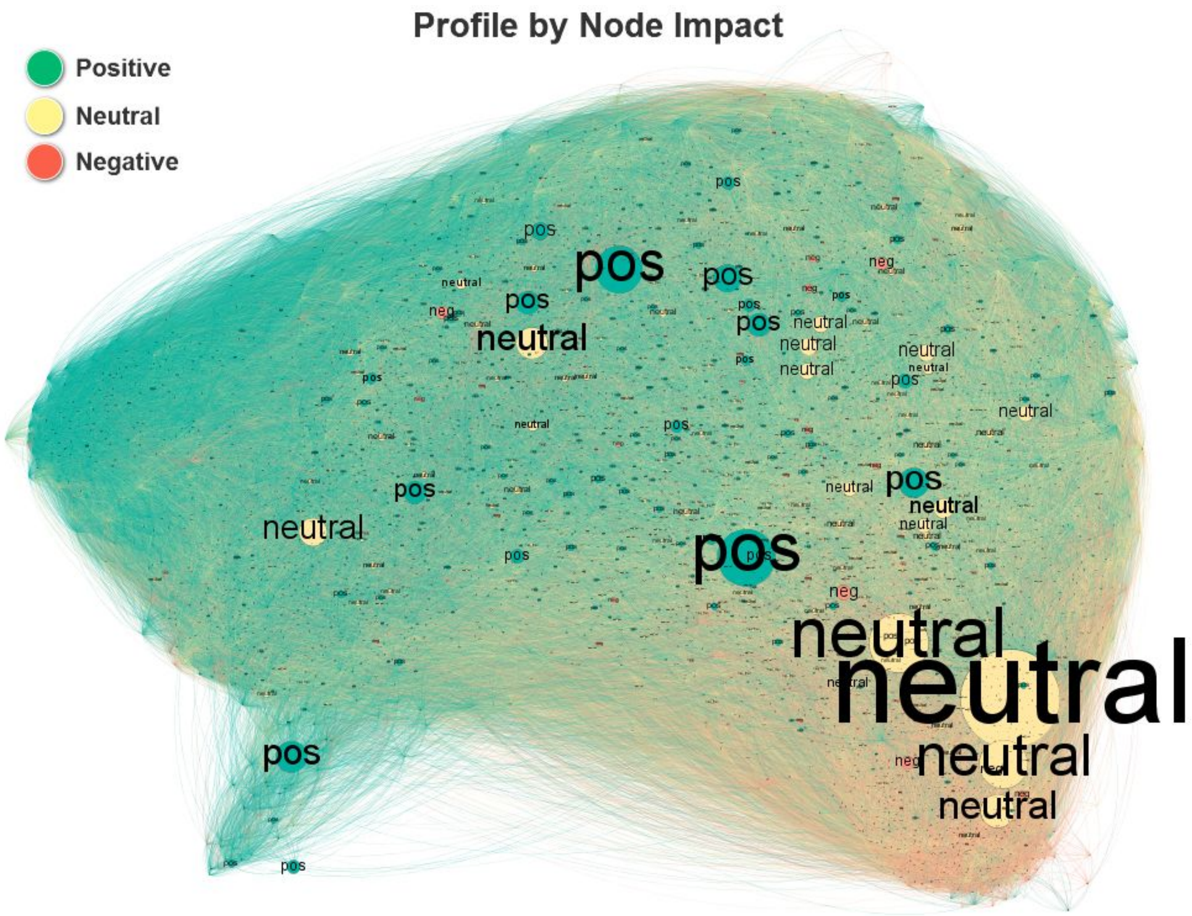


Highly attractive users that get a lot of reactions to their vents are reflected in the graph.
the neutral users in the purple community seems very attractive and uniformly neutral community

# Profile



Most of the highly degree nodes in the core graph are positive. Some negative users with high degree are located aside bunches of very high positive users.

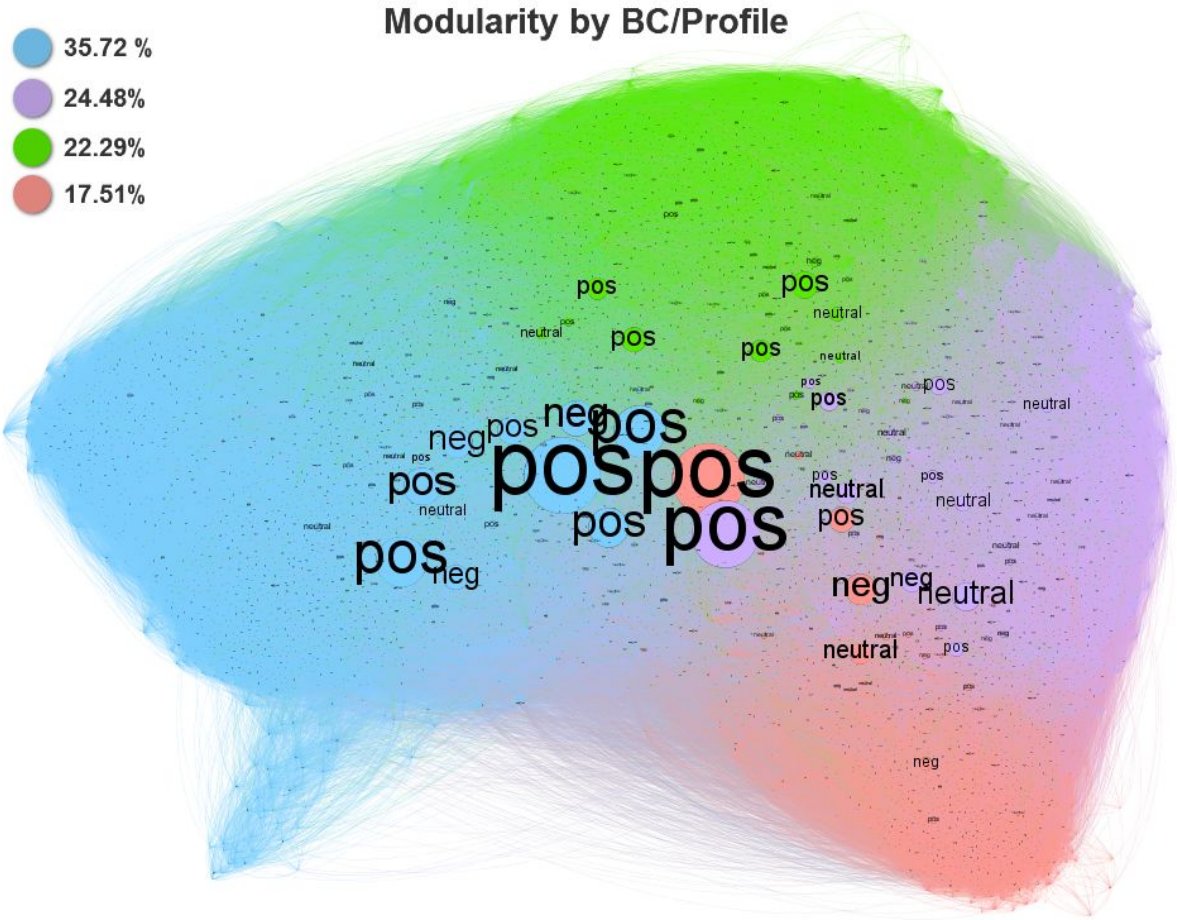


It can be seen clearly that negative users are not likely to be attractive in terms of reactions to their posts. Moreover, they are the least users in the core of the application.

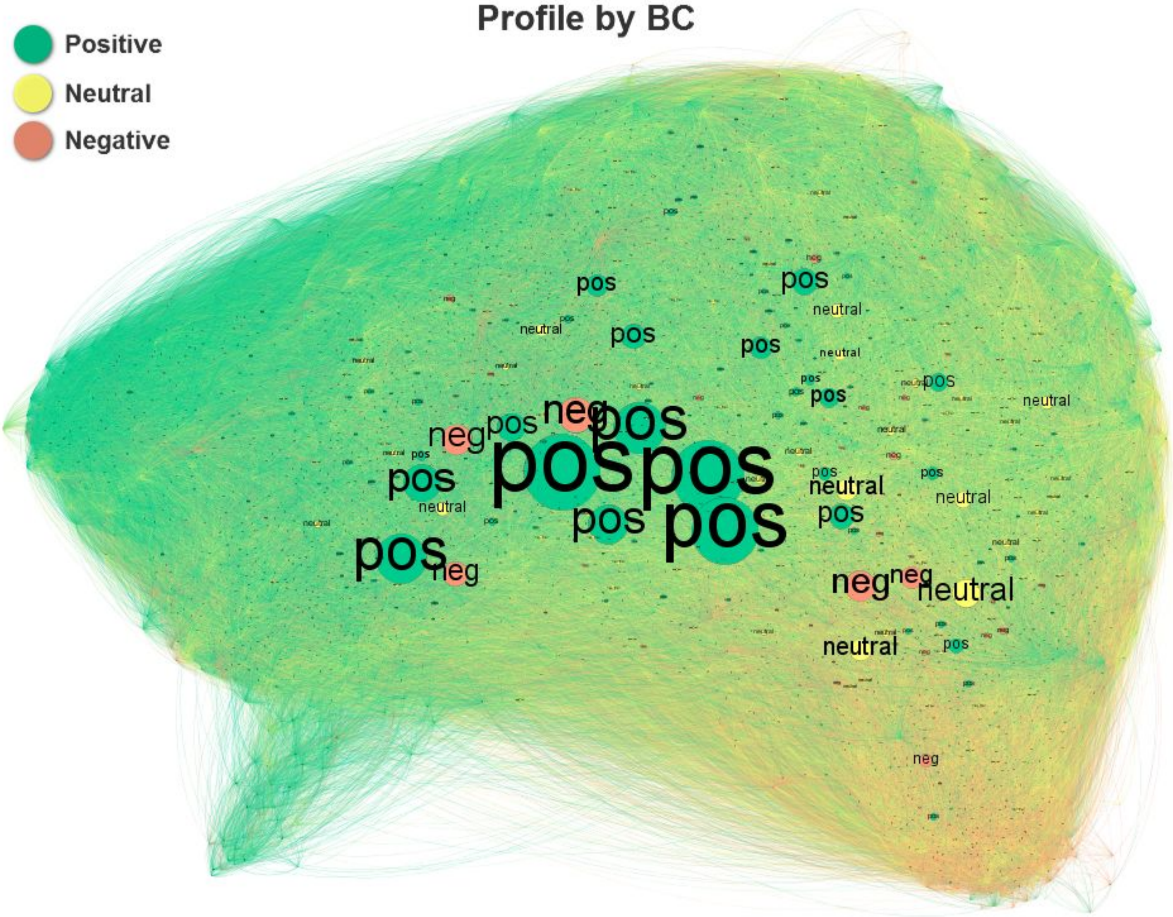
# Centrality Metrics

## Betweenness Centrality:

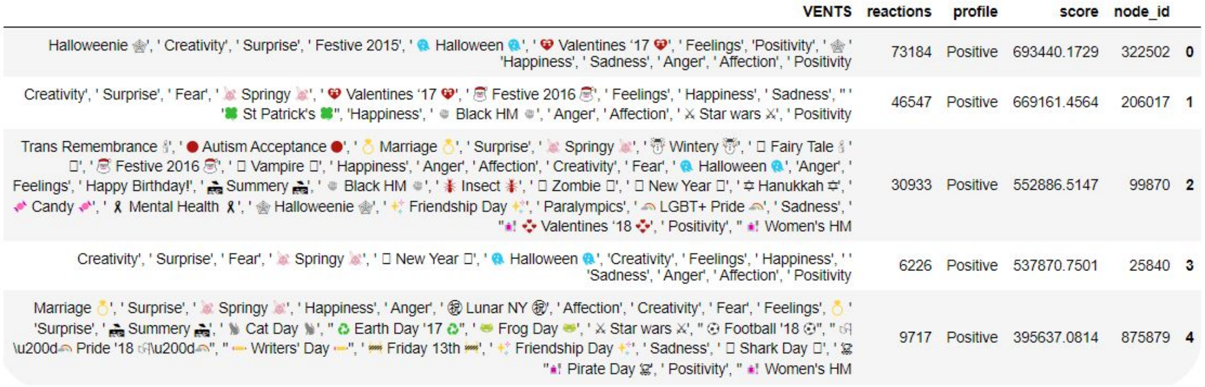
Displays nodes rank for each pair of vertices in a connected graph according to the shortest path. As expected, as moving away from the core's center , the lower the rank (in the center it can be found more high-ranked nodes).



As moving toward the center it can be found more users with high degree. when centralizing the high degree nodes we can see more clearly which users have more impact in terms of degree



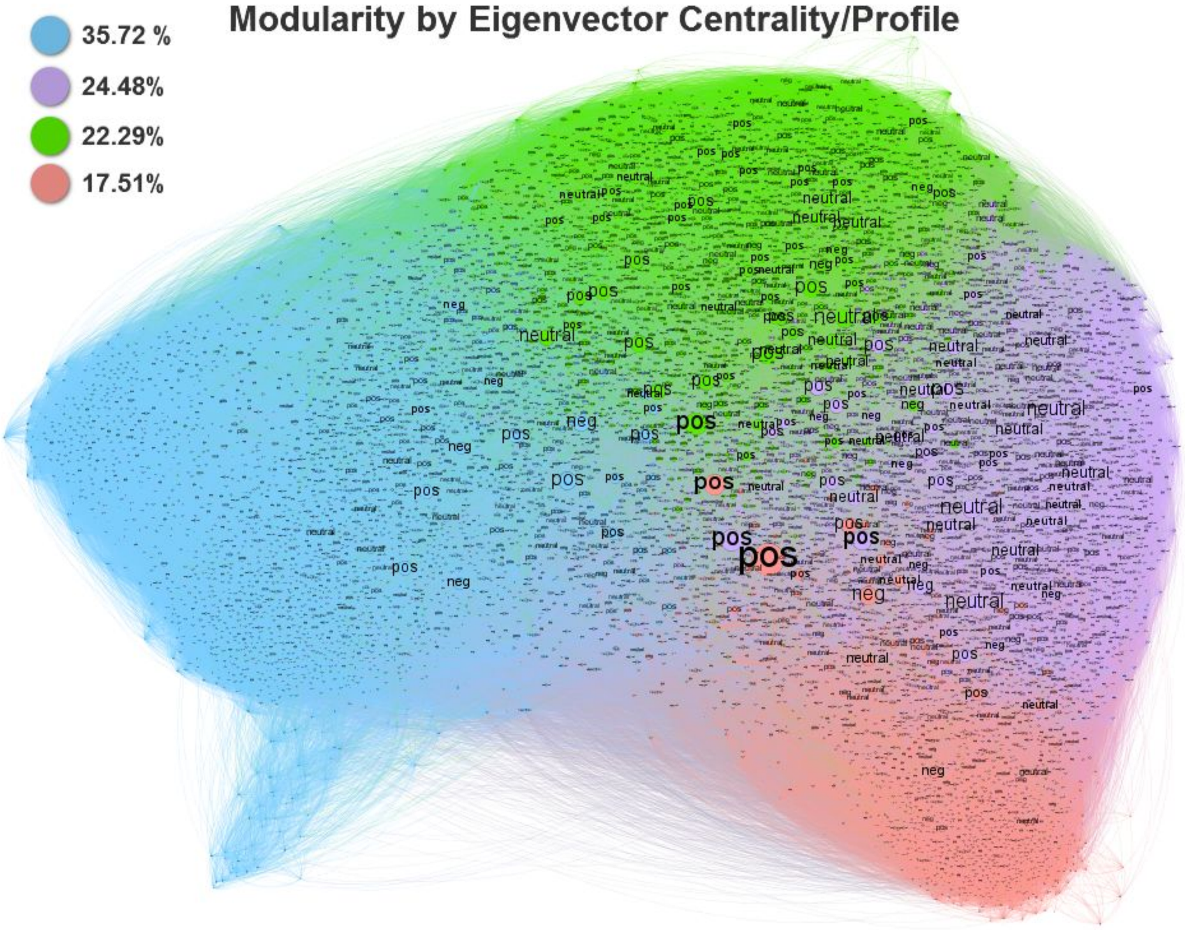
## Top Nodes - Betweenness Centrality



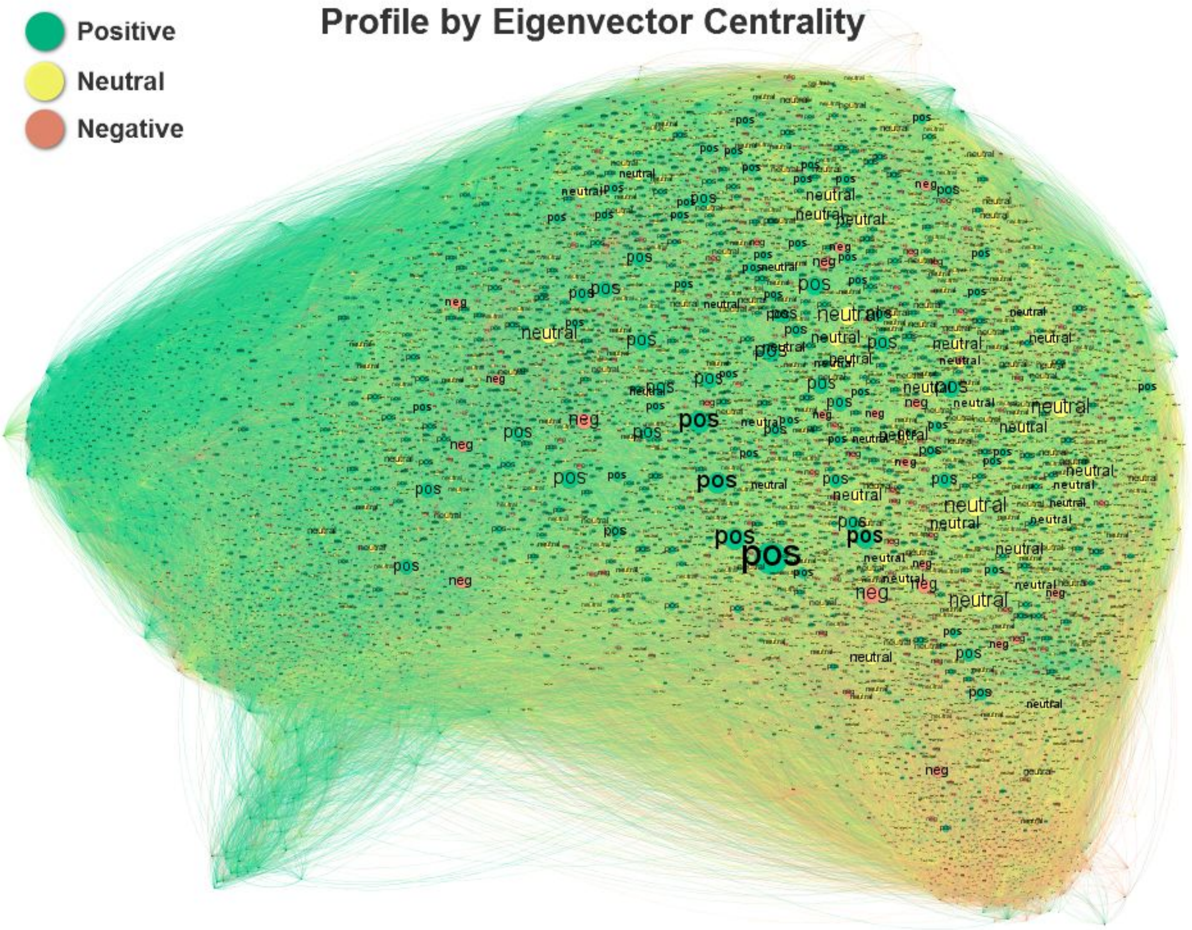
| | VENTS | reactions | profile | score | node_id |
|---|---|---|---|---|---|
| | Halloweenie 🕸', ' Creativity', ' Surprise', ' Festive 2015', ' 👻 Halloween 👻', ' 💔 Valentines '17 💔', ' Feelings', 'Positivity', ' 🕸 ' 'Happiness', ' Sadness', ' Anger', ' Affection', ' Positivity | 73184 | Positive | 693440.1729 | 322502 | **0** |
| | Creativity', ' Surprise', ' Fear', ' 🐰 Springy 🐰', ' 💔 Valentines '17 💔', ' 🎅 Festive 2016 🎅', ' Feelings', ' Happiness', ' Sadness', " ' '🍀 St Patrick's 🍀", 'Happiness', ' ✊ Black HM ✊', ' Anger', ' Affection', ' ⚔ Star wars ⚔', ' Positivity | 46547 | Positive | 669161.4564 | 206017 | **1** |
| | Trans Remembrance 🕯', ' ● Autism Acceptance ●', ' 💍 Marriage 💍', ' Surprise', ' 🐰 Springy 🐰', ' ⛄ Wintery ⛄', ' □ Fairy Tale 🕯 ' □', ' 🎅 Festive 2016 🎅', ' □ Vampire □', ' Happiness', ' Anger', ' Affection', ' Creativity', ' Fear', ' 👻 Halloween 👻', 'Anger', ' Feelings', ' Happy Birthday!', ' 🚣 Summery 🚣', ' ✊ Black HM ✊', ' 🐜 Insect 🐜', ' □ Zombie □', ' □ New Year □', ' ✡ Hanukkah ✡', ' 🍬 Candy 🍬', ' 🎗 Mental Health 🎗', ' 🕸 Halloweenie 🕸', ' ✨ Friendship Day ✨', ' Paralympics', ' 🌈 LGBT+ Pride 🌈', ' Sadness', ' "🚺! 💞 Valentines '18 💞', ' Positivity', " 🚺! Women's HM | 30933 | Positive | 552886.5147 | 99870 | **2** |
| | Creativity', ' Surprise', ' Fear', ' 🐰 Springy 🐰', ' □ New Year □', ' 👻 Halloween 👻', 'Creativity', ' Feelings', ' Happiness', ' ' 'Sadness', ' Anger', ' Affection', ' Positivity | 6226 | Positive | 537870.7501 | 25840 | **3** |
| | Marriage 💍', ' Surprise', ' 🐰 Springy 🐰', ' Happiness', ' Anger', ' 🐲 Lunar NY 🐲', ' Affection', ' Creativity', ' Fear', ' Feelings', ' 💍 ' 'Surprise', ' 🚣 Summery 🚣', ' 🐈 Cat Day 🐈', " ♻ Earth Day '17 ♻", ' 🐸 Frog Day 🐸', ' ⚔ Star wars ⚔', " ⚽ Football '18 ⚽", " 🏳\u200d🌈 Pride '18 🏳\u200d🌈", " ✏ Writers' Day ✏", ' 🚧 Friday 13th 🚧', ' ✨ Friendship Day ✨', ' Sadness', ' □ Shark Day □', ' ☠ "🚺! Pirate Day ☠', ' Positivity', " 🚺! Women's HM | 9717 | Positive | 395637.0814 | 875879 | **4** |

# Eigenvector Centrality:

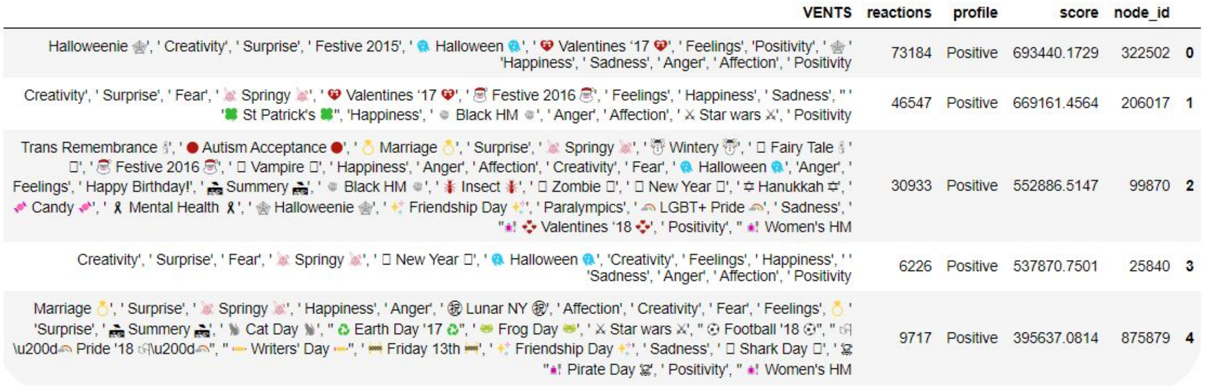
Eigenvector centrality is used to measure the level of influence of a node within a network. The Eigenvector Centrality score is relative to the number of connections a node have with other nodes.



In terms of connections, as you're more positive user you are more likely to have higher connections with other users, then neutral users and then negative ones.



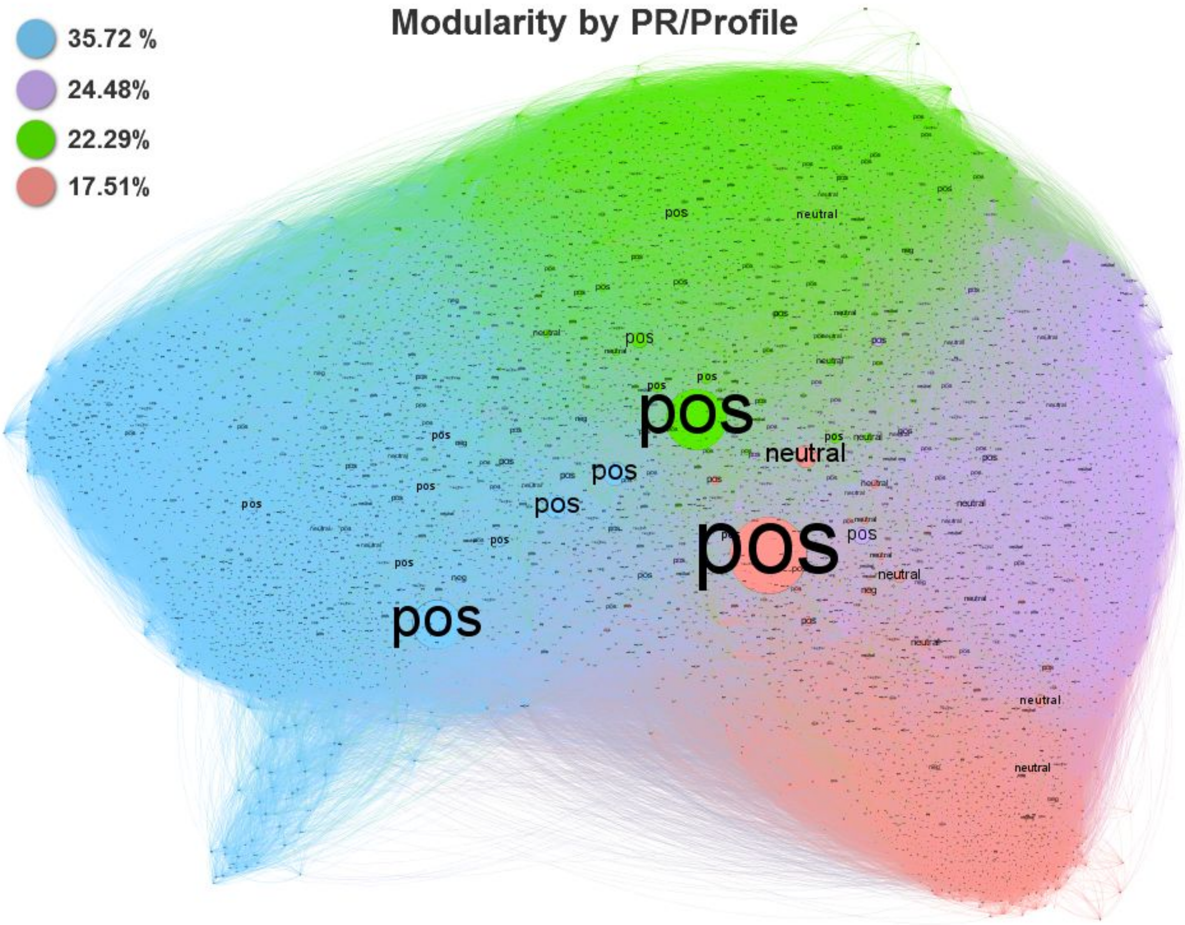
## Top Nodes - Eigenvector Centrality



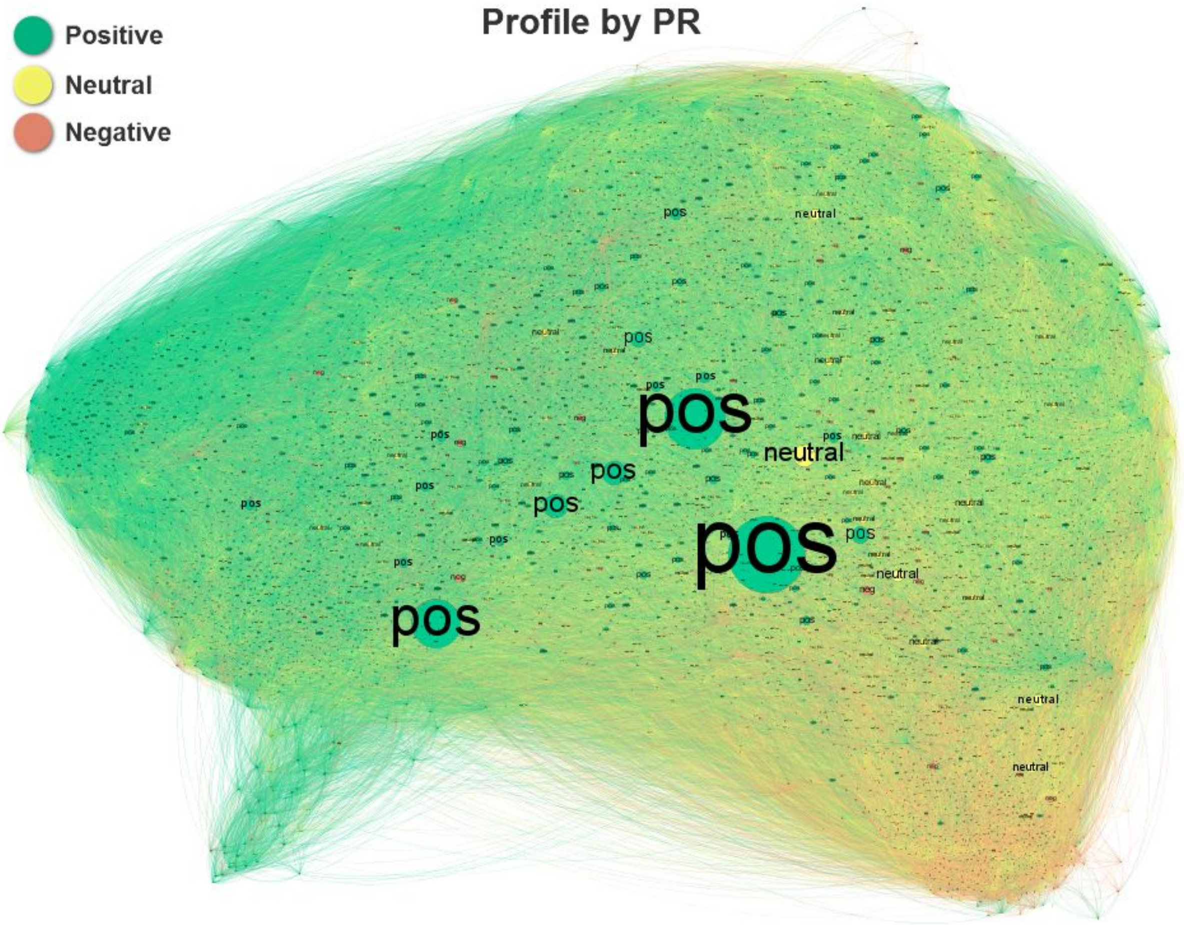
| VENTS | reactions | profile | score | node_id | |
|---|---|---|---|---|---|
| Halloweenie 🕸', ' Creativity', ' Surprise', ' Festive 2015', ' 👻 Halloween 👻', ' 💔 Valentines '17 💔', ' Feelings', 'Positivity', ' 🕸 ' 'Happiness', ' Sadness', ' Anger', ' Affection', ' Positivity | 73184 | Positive | 693440.1729 | 322502 | **0** |
| Creativity', ' Surprise', ' Fear', ' 🐰 Springy 🐰', ' 💔 Valentines '17 💔', ' 🎅 Festive 2016 🎅', ' Feelings', ' Happiness', ' Sadness', " ' '🍀 St Patrick's 🍀", 'Happiness', ' ✊ Black HM ✊', ' Anger', ' Affection', ' ⚔ Star wars ⚔', ' Positivity | 46547 | Positive | 669161.4564 | 206017 | **1** |
| Trans Remembrance ⚧', ' 🔴 Autism Acceptance 🔴', ' 💍 Marriage 💍', ' Surprise', ' 🐰 Springy 🐰', ' ☃ Wintery ☃', ' □ Fairy Tale 🧚 ' □', ' 🎅 Festive 2016 🎅', ' □ Vampire □', ' Happiness', ' Anger', ' Affection', ' Creativity', ' Fear', ' 👻 Halloween 👻', 'Anger', ' Feelings', ' Happy Birthday!', ' 🏖 Summery 🏖', ' ✊ Black HM ✊', ' 🐜 Insect 🐜', ' □ Zombie □', ' □ New Year □', ' ✡ Hanukkah ✡', ' 🍬 Candy 🍬', ' 🎗 Mental Health 🎗', ' 🕸 Halloweenie 🕸', ' ✨ Friendship Day ✨', ' Paralympics', ' 🌈 LGBT+ Pride 🌈', ' Sadness', ' "💅! 💞 Valentines '18 💞', ' Positivity', " 💅! Women's HM | 30933 | Positive | 552886.5147 | 99870 | **2** |
| Creativity', ' Surprise', ' Fear', ' 🐰 Springy 🐰', ' □ New Year □', ' 👻 Halloween 👻', 'Creativity', ' Feelings', ' Happiness', ' ' 'Sadness', ' Anger', ' Affection', ' Positivity | 6226 | Positive | 537870.7501 | 25840 | **3** |
| Marriage 💍', ' Surprise', ' 🐰 Springy 🐰', ' Happiness', ' Anger', ' 🧧 Lunar NY 🧧', ' Affection', ' Creativity', ' Fear', ' Feelings', 💍 ' 'Surprise', ' 🏖 Summery 🏖', ' 🐈 Cat Day 🐈', " ♻ Earth Day '17 ♻", ' 🐸 Frog Day 🐸', ' ⚔ Star wars ⚔', " ⚽ Football '18 ⚽", " 🏳\u200d🌈 Pride '18 🏳\u200d🌈", " ✏ Writers' Day ✏", ' 🚧 Friday 13th 🚧', ' ✨ Friendship Day ✨', ' Sadness', ' □ Shark Day □', ' ☠ "💅! Pirate Day ☠', ' Positivity', " 💅! Women's HM | 9717 | Positive | 395637.0814 | 875879 | **4** |

## Page Rank:

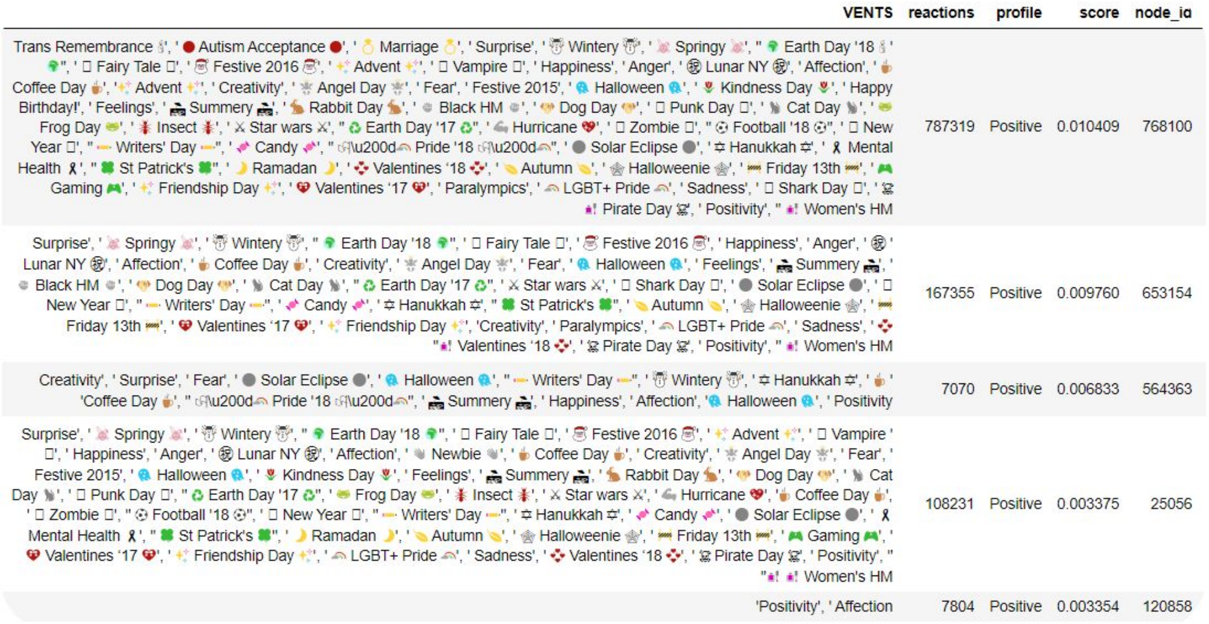
Another form of Eigenvector Centrality designed for ranking user’s content, using edges between ‘nodes’ as a measure of importance. It can be used for any kind of network, though.



The top ranked users by their vents content are mostly positive and they have much more impact than neutral users and negative ones.



## Top Nodes - Page Rank

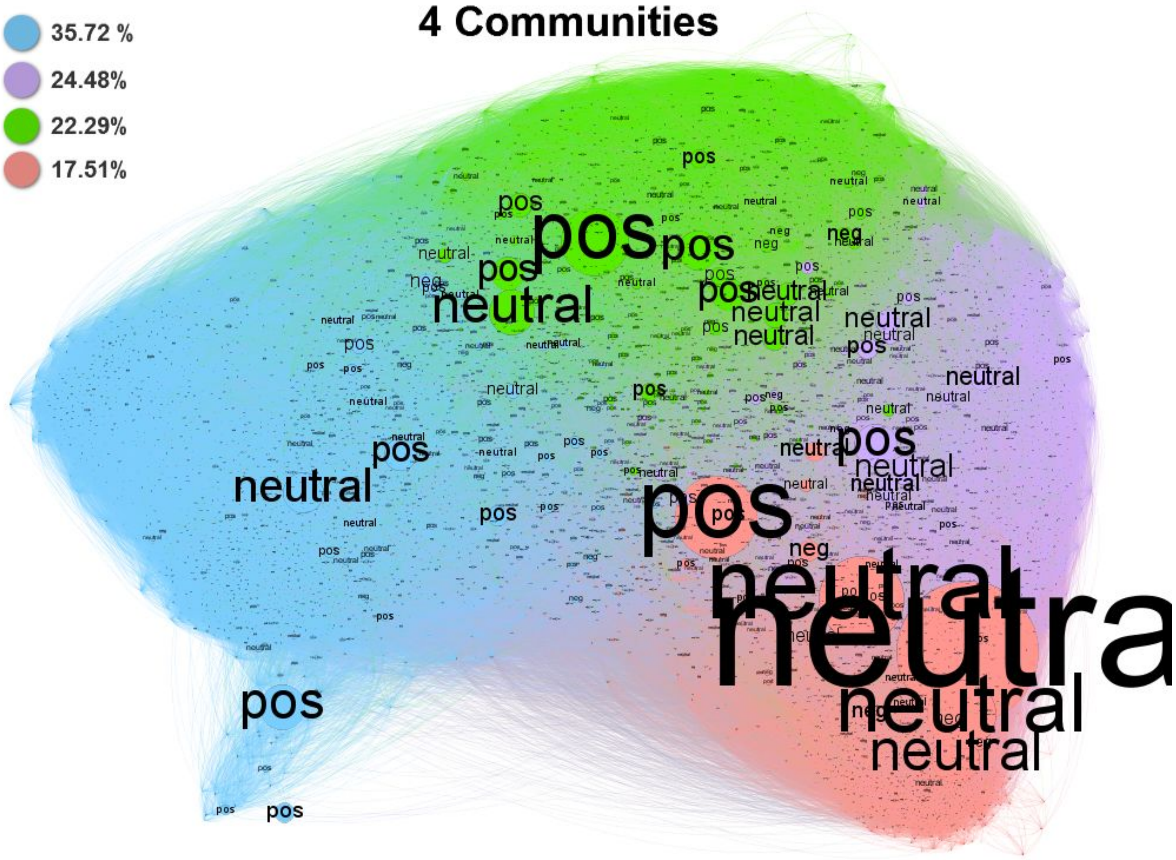


| VENTS | reactions | profile | score | node_id |
|---|---|---|---|---|
| Trans Remembrance 🏳️‍⚧️', ' 🔴 Autism Acceptance 🔴', ' 💍 Marriage 💍', ' Surprise', ' ☃️ Wintery ☃️', ' 🐰 Springy 🐰', " 🌍 Earth Day '18 🏳️‍⚧️ ' 🌍", ' □ Fairy Tale □', ' 🎅 Festive 2016 🎅', ' ✨ Advent ✨', ' □ Vampire □', ' Happiness', ' Anger', ' ㊗ Lunar NY ㊗', ' Affection', ' ☕ Coffee Day ☕', ' ✨ Advent ✨', ' Creativity', ' 👼 Angel Day 👼', ' Fear', ' Festive 2015', ' 👻 Halloween 👻', ' 🌷 Kindness Day 🌷', ' Happy Birthday!', ' Feelings', ' 🏖 Summery 🏖', ' 🐇 Rabbit Day 🐇', ' ✊ Black HM ✊', ' 🐶 Dog Day 🐶', ' □ Punk Day □', ' 🐈 Cat Day 🐈', ' 🐸 Frog Day 🐸', ' 🐜 Insect 🐜', ' ⚔ Star wars ⚔', " ♻ Earth Day '17 ♻", ' 💪 Hurricane 💔', ' □ Zombie □', " ⚽ Football '18 ⚽", ' □ New Year □', " ✏ Writers' Day ✏", ' 🍬 Candy 🍬', " 🏳\u200d🌈 Pride '18 🏳\u200d🌈", ' ● Solar Eclipse ●', ' ✡ Hanukkah ✡', ' 🎗 Mental Health 🎗', " 🍀 St Patrick's 🍀", ' 🌙 Ramadan 🌙', ' 💞 Valentines '18 💞', ' 🍂 Autumn 🍂', ' 🕸 Halloweenie 🕸', ' 🚧 Friday 13th 🚧', ' 🎮 Gaming 🎮', ' ✨ Friendship Day ✨', ' 💔 Valentines '17 💔', ' Paralympics', ' 🌈 LGBT+ Pride 🌈', ' Sadness', ' □ Shark Day □', ' ☠ 💅! Pirate Day ☠', ' Positivity', " 💅! Women's HM | 787319 | Positive | 0.010409 | 768100 |
| Surprise', ' 🐰 Springy 🐰', ' ☃️ Wintery ☃️', " 🌍 Earth Day '18 🌍", ' □ Fairy Tale □', ' 🎅 Festive 2016 🎅', ' Happiness', ' Anger', ' ㊗ Lunar NY ㊗', ' Affection', ' ☕ Coffee Day ☕', ' Creativity', ' 👼 Angel Day 👼', ' Fear', ' 👻 Halloween 👻', ' Feelings', ' 🏖 Summery 🏖', ' ✊ Black HM ✊', ' 🐶 Dog Day 🐶', ' 🐈 Cat Day 🐈', " ♻ Earth Day '17 ♻", ' ⚔ Star wars ⚔', ' □ Shark Day □', ' ● Solar Eclipse ●', ' □ New Year □', " ✏ Writers' Day ✏", ' 🍬 Candy 🍬', ' ✡ Hanukkah ✡', " 🍀 St Patrick's 🍀", ' 🍂 Autumn 🍂', ' 🕸 Halloweenie 🕸', ' 🚧 Friday 13th 🚧', ' 💔 Valentines '17 💔', ' ✨ Friendship Day ✨', 'Creativity', ' Paralympics', ' 🌈 LGBT+ Pride 🌈', ' Sadness', ' 💞 "💅! Valentines '18 💞', ' ☠ Pirate Day ☠', ' Positivity', " 💅! Women's HM | 167355 | Positive | 0.009760 | 653154 |
| Creativity', ' Surprise', ' Fear', ' ● Solar Eclipse ●', ' 👻 Halloween 👻', " ✏ Writers' Day ✏", ' ☃️ Wintery ☃️', ' ✡ Hanukkah ✡', ' ☕ 'Coffee Day ☕', " 🏳\u200d🌈 Pride '18 🏳\u200d🌈", ' 🏖 Summery 🏖', ' Happiness', ' Affection', '👻 Halloween 👻', ' Positivity | 7070 | Positive | 0.006833 | 564363 |
| Surprise', ' 🐰 Springy 🐰', ' ☃️ Wintery ☃️', " 🌍 Earth Day '18 🌍", ' □ Fairy Tale □', ' 🎅 Festive 2016 🎅', ' ✨ Advent ✨', ' □ Vampire □', ' Happiness', ' Anger', ' ㊗ Lunar NY ㊗', ' Affection', ' 👋 Newbie 👋', ' ☕ Coffee Day ☕', ' Creativity', ' 👼 Angel Day 👼', ' Fear', ' Festive 2015', ' 👻 Halloween 👻', ' 🌷 Kindness Day 🌷', ' Feelings', ' 🏖 Summery 🏖', ' 🐇 Rabbit Day 🐇', ' 🐶 Dog Day 🐶', ' 🐈 Cat Day 🐈', ' □ Punk Day □', " ♻ Earth Day '17 ♻", ' 🐸 Frog Day 🐸', ' 🐜 Insect 🐜', ' ⚔ Star wars ⚔', ' 💪 Hurricane 💔', '☕ Coffee Day ☕', ' □ Zombie □', " ⚽ Football '18 ⚽", ' □ New Year □', " ✏ Writers' Day ✏", ' ✡ Hanukkah ✡', ' 🍬 Candy 🍬', ' ● Solar Eclipse ●', ' 🎗 Mental Health 🎗', " 🍀 St Patrick's 🍀", ' 🌙 Ramadan 🌙', ' 🍂 Autumn 🍂', ' 🕸 Halloweenie 🕸', ' 🚧 Friday 13th 🚧', ' 🎮 Gaming 🎮', ' 💔 Valentines '17 💔', ' ✨ Friendship Day ✨', ' 🌈 LGBT+ Pride 🌈', ' Sadness', ' 💞 Valentines '18 💞', ' ☠ Pirate Day ☠', ' Positivity', " "💅! 💅! Women's HM | 108231 | Positive | 0.003375 | 25056 |
| 'Positivity', ' Affection | 7804 | Positive | 0.003354 | 120858 |

The centrality metrics graphs demonstrates how positive users are highly impact on the VENT application from different aspects like closeness, popularity, connections and more.
The neutral have little effect and the negative nodes are low impact for the parameters mentioned above.
Furthermore, for the 3 closeness metrics that i calculated, the top 5 users were positive.

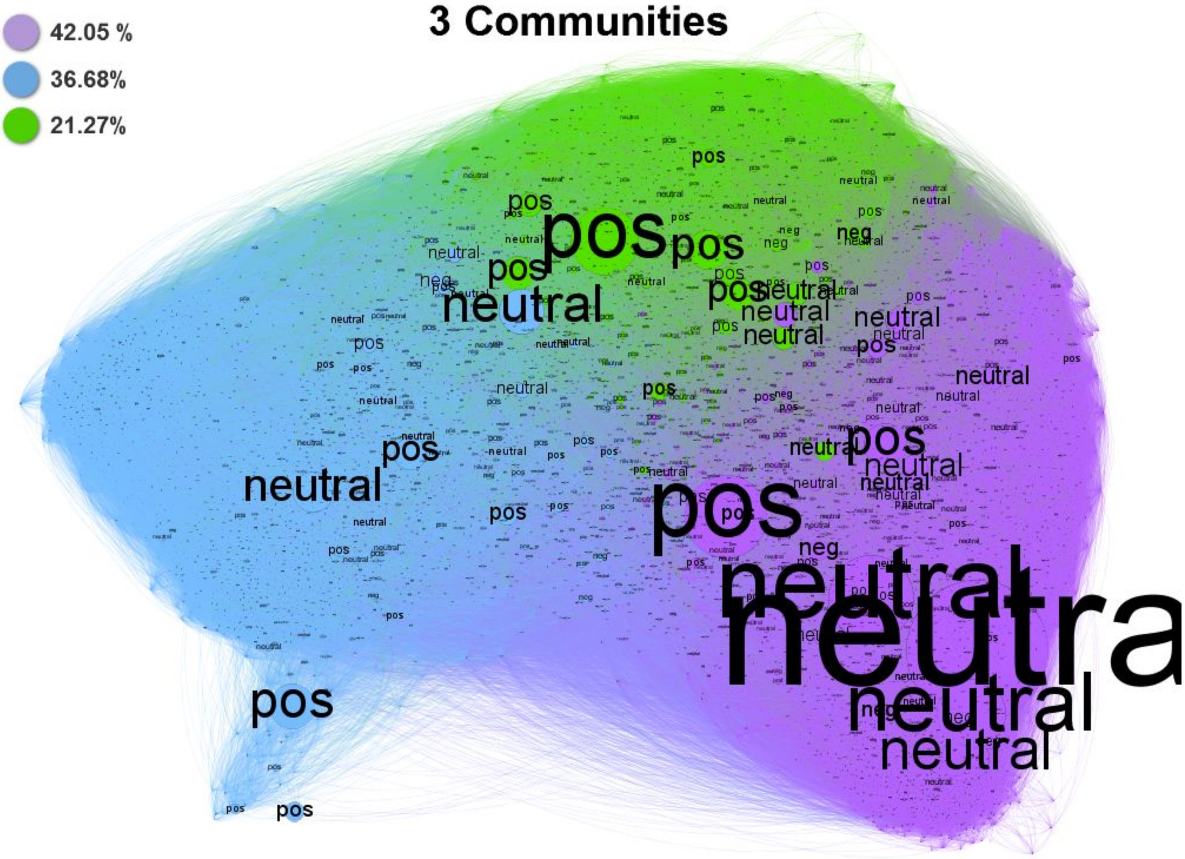
## Communities Hierarchy:

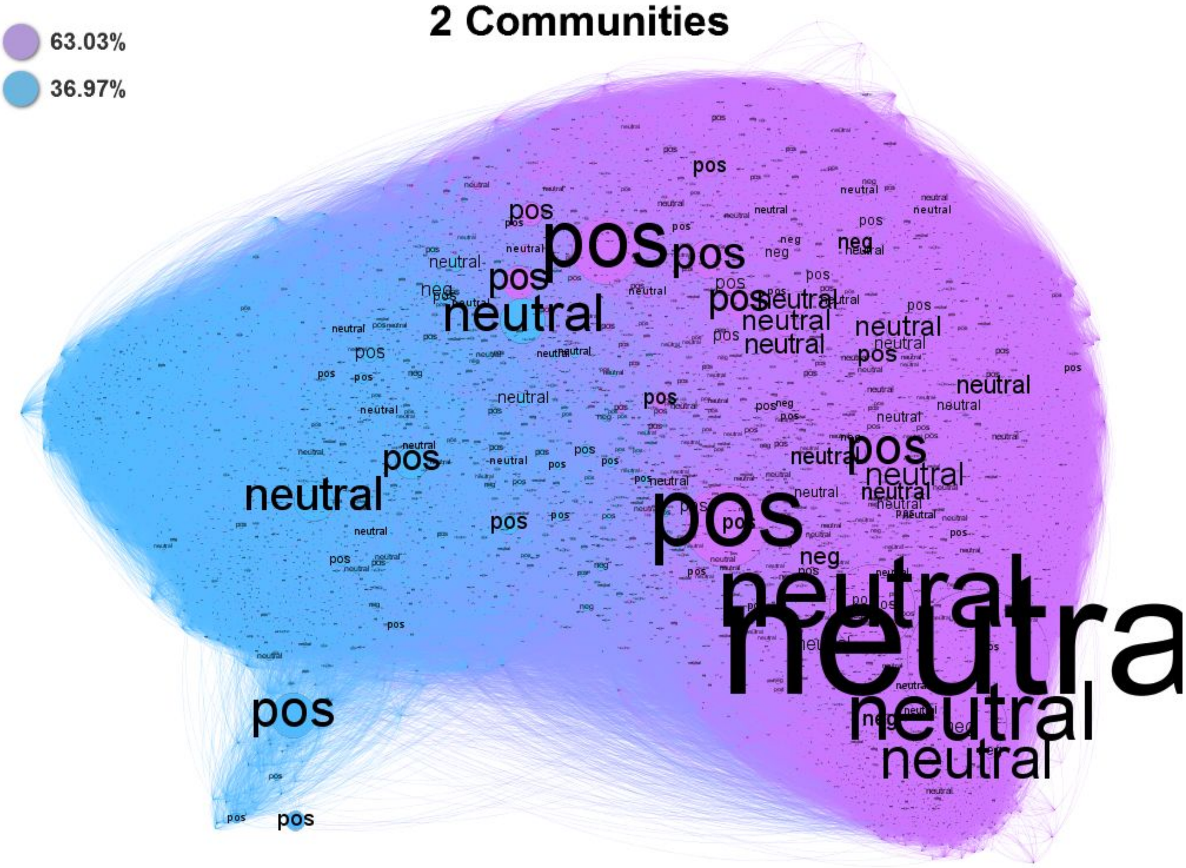
| | | |
|---|---|---|
| **2 communities**: | Resolution = 1.0 | Modularity = 0.326 |
| **3 communities**: | Resolution = 1.1 | Modularity = 0.323 |
| **4 communities**: | Resolution = 2.0 | Modularity = 0.252 |



4 Communities

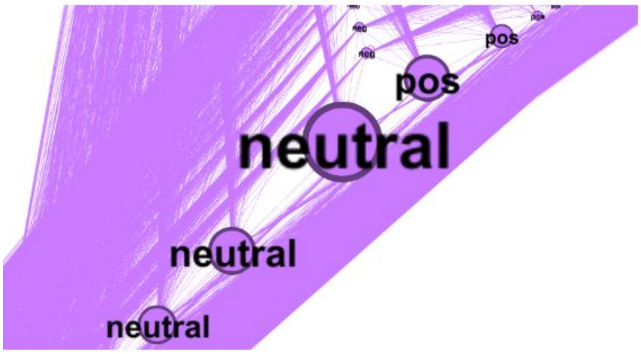
Most of the communities detected by Louvain method are positive-neutral except 1 community that is significant  mostly neutral.

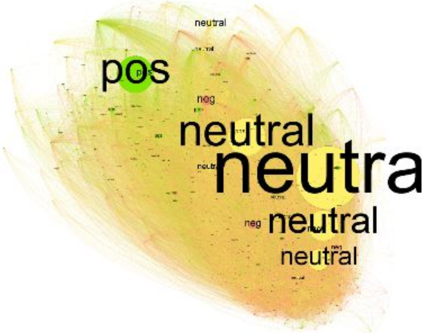

3 Communities
42.05 %
36.68%
21.27%
pos
neutral
neg


2 Communities
63.03%
36.97%
pos
neutral
neg

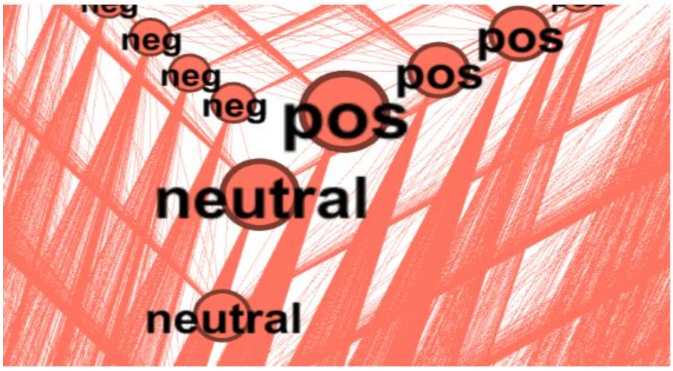
# Communities Properties:

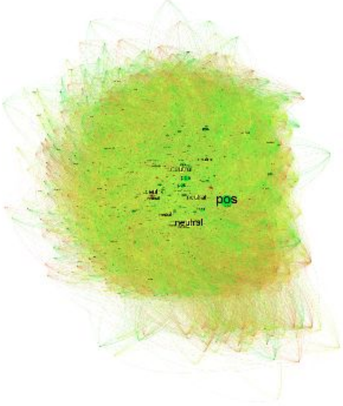
1





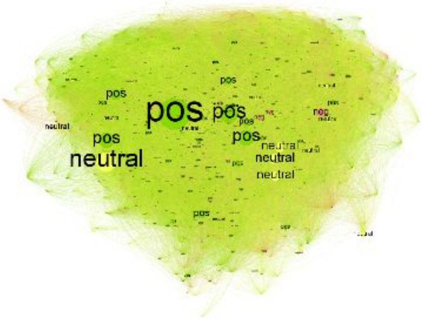
2



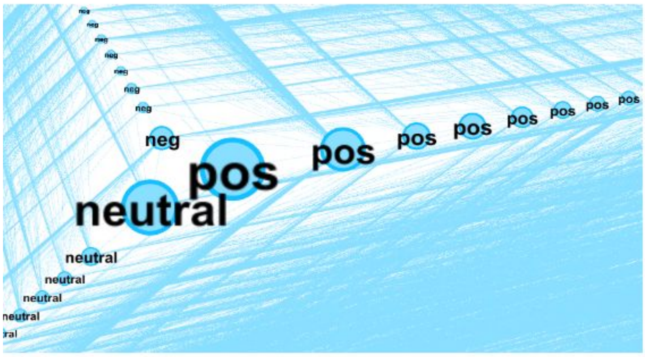


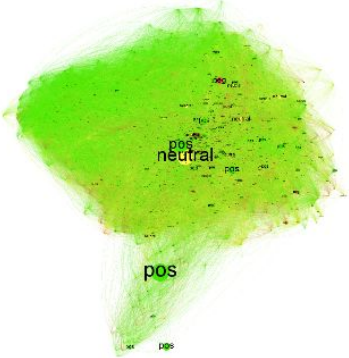
3

neg
neg
neg
pos
pos
pos
pos
neutral
neutral
neutral
neutral
neutral
neutral
neutral

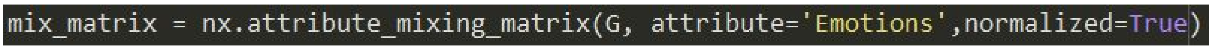


4

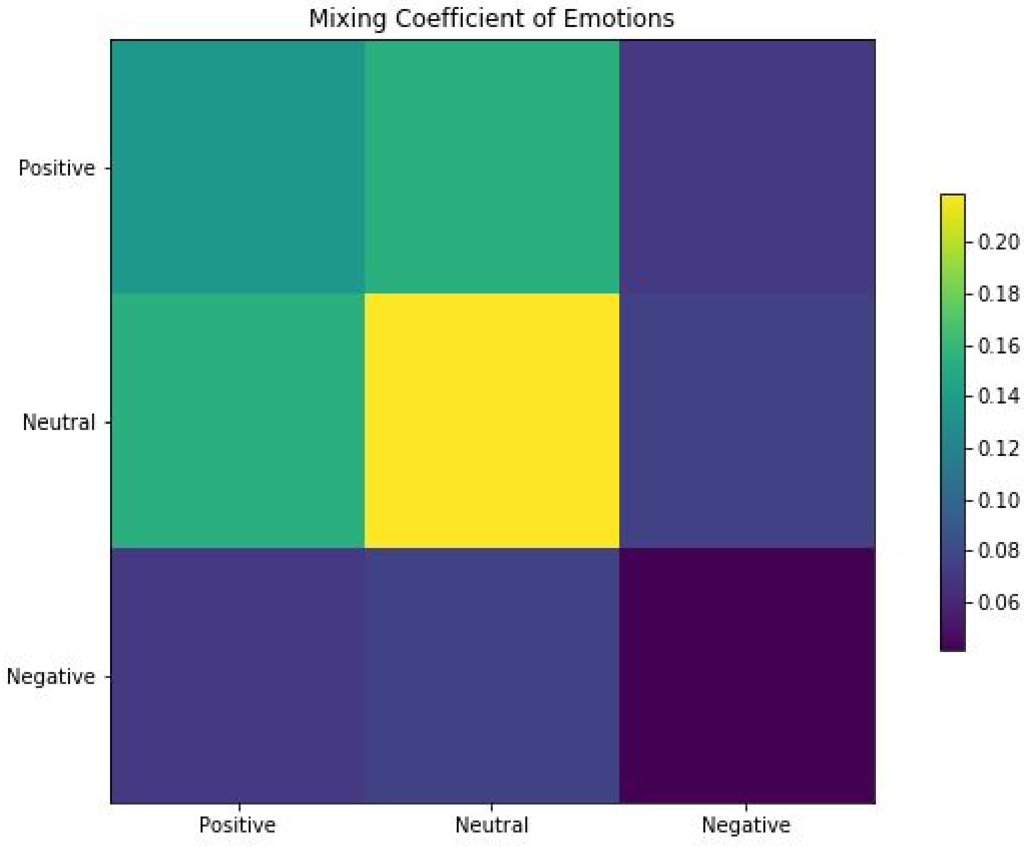




Using ‘networkx’ package, python I calculated the mixing coefficient of the emotions: positive, neutral and negative



```
mix_matrix = nx.attribute_mixing_matrix(G, attribute='Emotions',normalized=True)
```



## Thought

The Negative nodes have low coefficient probabilities with positive and neutral and especially with themselves, but they have a little bit more co-occurrence with Neutral nodes than Positive nodes.
Positive nodes are have more co-occurrence with Neutral nodes than with themselves, and the Neutral nodes have the biggest joint probability co-occurrence with themselves.

The purple community that is characterized as a largely neutral community connects with a community that is neutral-positive, then connects with a community that is essentially positive-neutral and finally with a community that is essentially positive.

# Conclusion

- As analyzing only the core of the data I conclude that the leaders of this application (nodes) that are characterized as neutral users are more likely to get a lot of reactions to their posts, and especially from neutral users like them. These neutral users have more effect on positive users than themselves. Positive and neutral users don't have much joint probability co-occurrence with negative users because “nobody likes them” and especially they don’t like themselves.
  It also reflected in the centrality metrics graphs how the positive users are the ‘leaders’ and a little bit after the neutral users, and the negative users are ‘trying to be pushed’.

- It’s all a question of perspective - the given emotions by the VENT application are mostly positive (more than ⅔ of the total vents), but the metadata shows different situation: about ⅖ of the total vents posted were negative while only ⅓ were positive (and the rest are neutral).
  Furthermore, when getting into the core of the application’s users, we realize that most of the influential users are positive and it proved by different centrality metrics: Page Rank, Eigenvector Centrality and Betweenness Centrality. (analysis for Closeness Centrality is not necessary in this case since we are dealing with the core of the graph).
  Positive users are more likely to have more connections, much more ranked by their content they post and are more close to the other users.
  Analyzing only the core cannot tell the whole story, further analysis and conclusion for lower degree of users (nodes) must be done.

- Only at the end of the project i realized that i could benefit much more from the data by analyzing the linking between different users, communities, the features that i have created to the date that these vents were posted. In the future, it can provide me a different perspective about different factors from outside the application that affects users behavior or on the other hand some vents that have connection about the real world.

- I would recommend the VENT application to add more vents that are defined as positive-neutral: we have seen how positive and neutral users are connecting between themselves and how they are responsible and highly impact when focusing on the communities hierarchy.
  Adding vents that are positive-neutral will reduce the gap between positive and negative communities, allowing us to strengthen the user community to more dense communities that have a stronger connection and less negativity- It would allow less clustering and adjust the business marketing strategy accordingly: advertisements, offers and more.

- Respecting data privacy as they do in the paper, inspired me alot for future work - the fact that in very clever way the data derived and legally, does not mean that we should publish also the usernames and their exact posts (vents).