

Analiza danych z Twittera

Liga Mistrzów

1. Cel projektu

Celem projektu było zbadanie zależności oraz społeczności pojawiających się w danych uzyskanych z portalu Twitter. Aby ułatwić analizę zawężono tematykę do Ligi Mistrzów UEFA.

2. Wyjaśnienie pojęć

- Twitter – portal społecznościowy pozwalający na publikowanie krótkich wiadomości – www.twitter.com
- Tweet – pojedyncza wiadomość o długości nie przekraczającej 140 znaków
- Hashtag – konstrukcja zaczynająca się od znaku „#”, pozwalająca na powiązanie tweeta z jakimś tematem – np. #cristiano
- Retweet – tweet, który został ponownie udostępniony (najczęściej przez innego użytkownika)
- User mention - konstrukcja zaczynająca się od znaku „@”, pozwalająca na nawiązanie do jakiegoś użytkownika – np. @MUFC
- LM – Liga Mistrzów - najbardziej prestiżowe, europejskie rozgrywki klubowe w piłkę nożną

3. Schemat postępowania

Pierwszym etapem było zebranie danych – tweetów. Zostały one umieszczone w relacyjnej bazie danych.

Kolejnym krokiem była analiza ogólna. Pozwoliła ona zorientować się w kwestii rozmiaru zebranych danych i zaplanować poruszanie się po nich. Ponadto, umożliwiła określenie które powiązania między danymi są obiecujące do zbadania.

Kluczowym etapem była analiza konkretnych zjawisk społecznościowych w sieciach tworzonych przez użytkowników i ich tweety. W tym celu wybierano reprezentatywne podzbiory danych – za pomocą podziału czasowego lub zawężenia tematyki, np. tweety dotyczące pojedynczego meczu.

4. Źródło danych i metoda ich pozyskiwania

Twitter oferuje tzw. „Streaming API”, które pozwala na nawiązanie połączenia z portalem i odbieranie w czasie rzeczywistym nowych tweetów publikowanych przez użytkowników. Jest to bardzo wydajne i wygodne rozwiązanie, ponieważ nie wymaga wykonywania wielu zapytań HTTP. Wprawdzie w ten sposób można uzyskać tylko te tweety, które są publikowane w trakcie trwania połączenia, jednak na potrzeby tego projektu było to zupełnie wystarczające. Połączenie było utrzymywane od dnia rozpoczęcia 3. kolejki grupowej LM do następnego dnia po zakończeniu kolejki.

Do zebrania danych została użyta biblioteka „tweepy”, która implementuje między innymi autoryzację i nawiązanie połączenia. Dzięki zastosowaniu filtrowania, możliwe było odsianie tweetów dotyczących Ligi Mistrzów na podstawie słów kluczowych i hashtag'ów zawartych w ich treści.

Tweety przesyłane są w formacie JSON, a każdy z nich zawiera około 5 tysięcy znaków. Wynika to z dużej różnorodności atrybutów powiązanych z każdym tweetem. Surowe dane zapisywane były do pliku, następnie dokonano redukcji nadmiarowych informacji i migracji do bazy danych.

Okres zbierania tweetów:

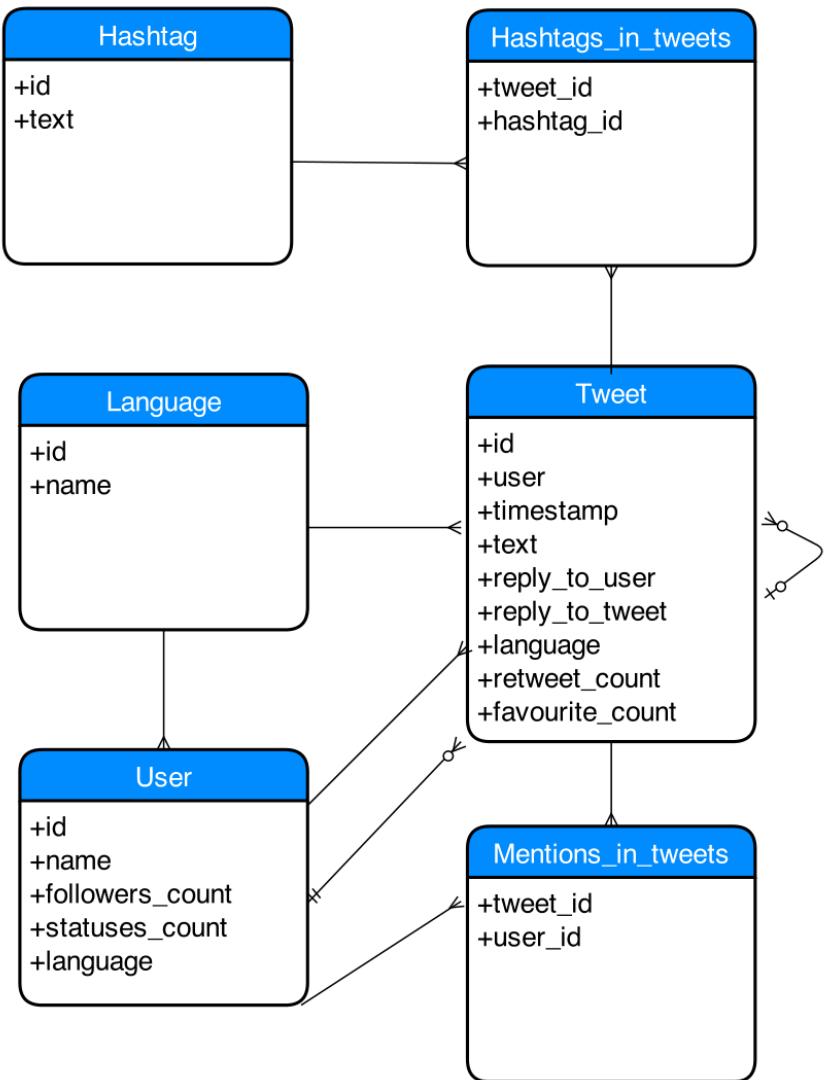
21.10.2014, g. 16:15 – 23.10.2014, g. 10:47

Kolejka LM rozgrywana była 21.10 oraz 22.10 o godzinie 20:45

Lista hashtagów i słów kluczowych użytych do filtrowania:

| | |
|------------------|------------------|
| #championsleague | #lfc |
| #ucl | #ludogorecsbg |
| champions league | #malmo_ff |
| uefa | #mcfc |
| #afcajax | #nkmaribor |
| #apoelfc | #olympiacos |
| #as_monaco | #olympiacossfp |
| #arsenal | #psg |
| #athletic_eng | #roma |
| #athleticdbilbao | #realmadrid |
| #atleti | #romabayern |
| #bayer04 | #rscanderlecht |
| #bvb | #s04_en |
| #chelsea | #s04 |
| #chelseafc | #sl_benfica |
| #cska | #szachtar |
| #fc_basel | #sportingptcom |
| #fcbarcelona | #sportingliisboa |
| #fcbate | #vfbb04 |
| #fcbayern | #cr7 |
| #fcporto | #messi |
| #galatasaraysk | #ibrahimovic |
| #alamadrid | #daretozlatan |
| #juventusfc | #zlatan |
| #juventusfcen | #suarez |

5. Struktura bazy danych



Hashtag – przechowuje nazwy hashtagów - np. „cristiano”.

Language – przechowuje identyfikatory języków - np. „pl”.

User – przechowuje dane o użytkownikach.

Tweet – przechowuje tweety i informacje o nich, takie jak autor, dokładny czas publikacji, treść, język czy też id użytkownika oraz jego tweeta jeśli jest to odpowiedź na jakiegoś tweeta.

Hashtags in tweets – tabela łącznikowa, przechowuje pojedyncze wystąpienia konkretnego hastaga w konkretnym tweecie.

Mentions in tweets – tabela łącznikowa, przechowuje pojedyncze wystąpienia konkretnego user mention w konkretnym tweecie.

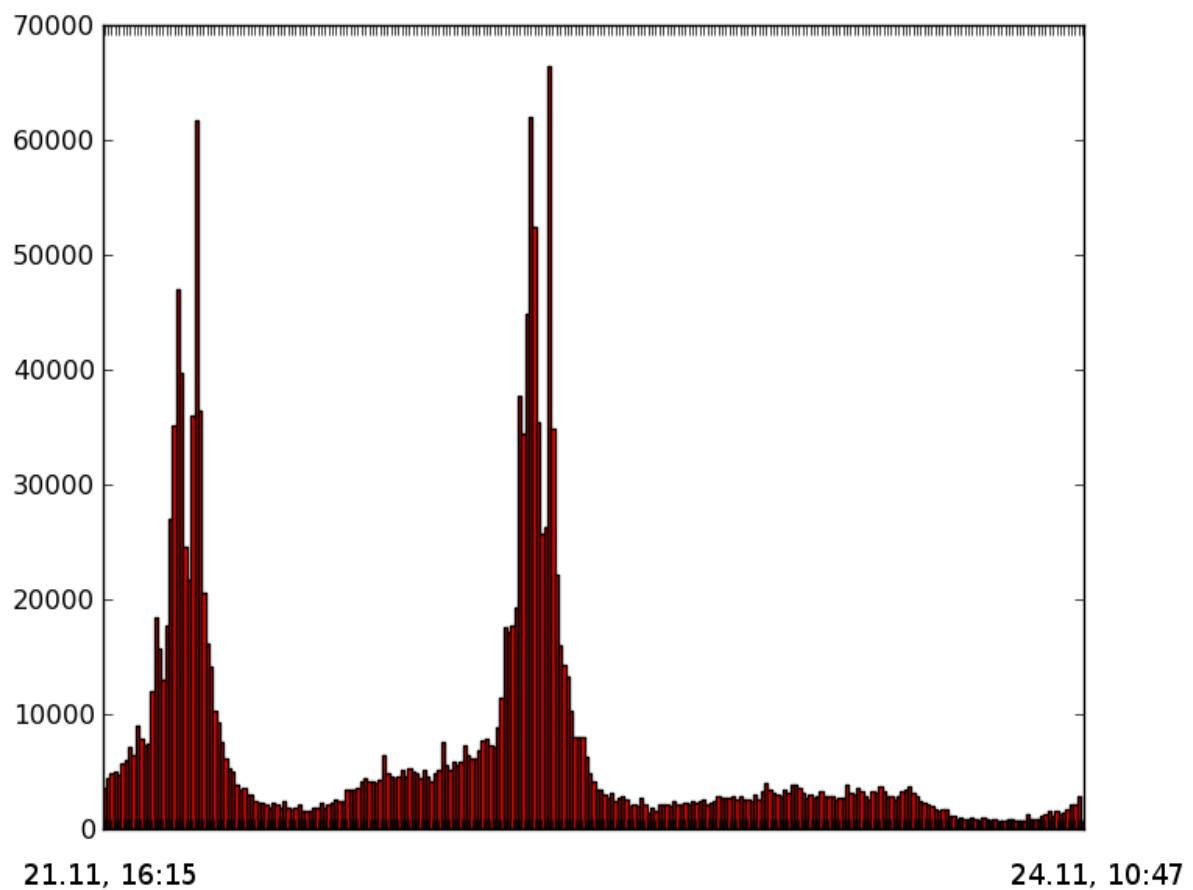
6. Analiza danych

6.1. Analiza wstępna

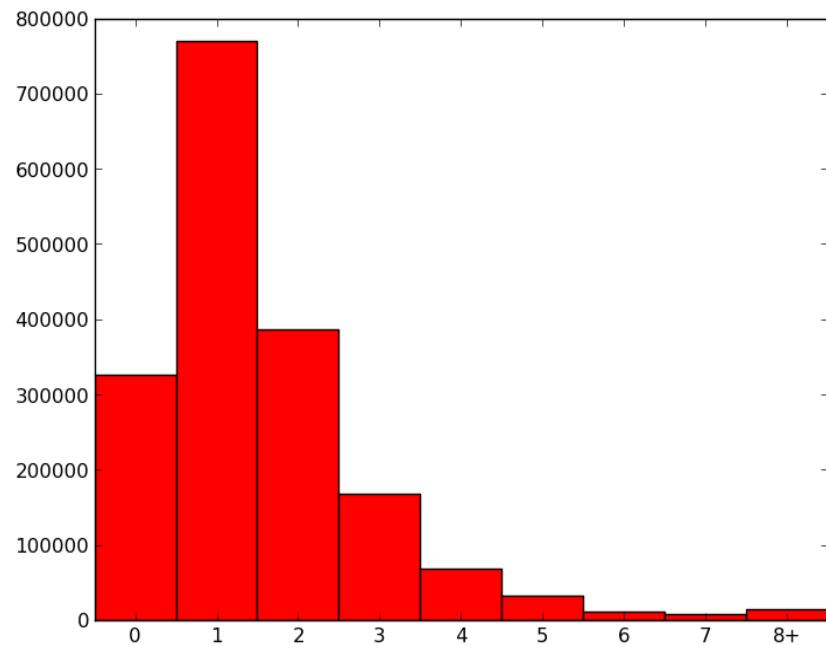
Liczliwość rekordów w bazie danych:

- hashtag_in_tweet: **2 739 238**
- hashtag: **68 979**
- tweet: **1 784 265**
- language: **53**
- user: **673 460**
- mention_in_tweet: **1 577 217**

Rozkład tweetów w czasie:



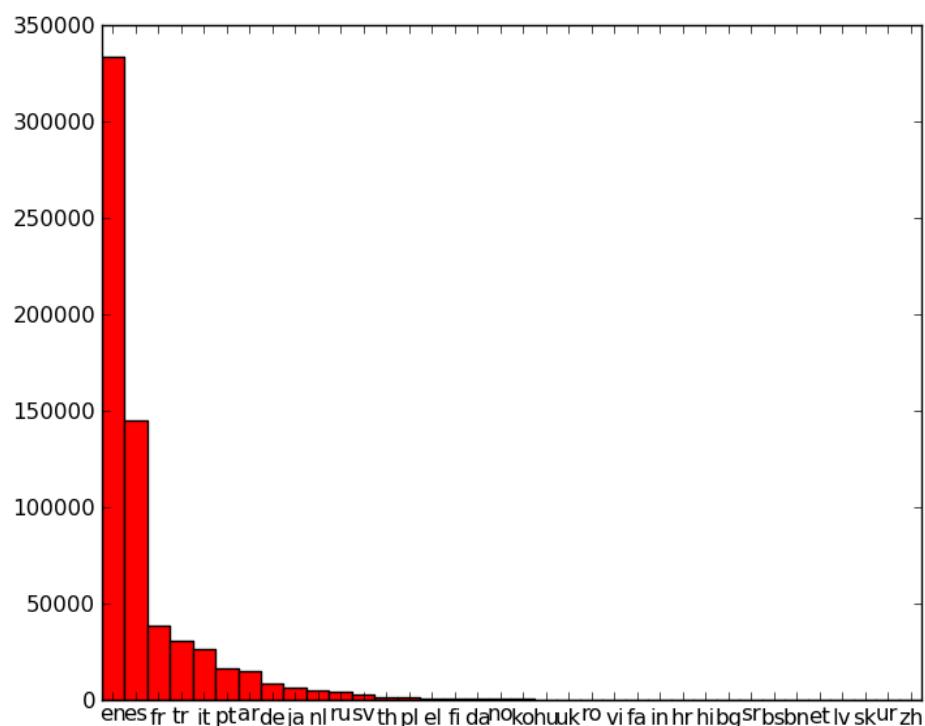
Histogram dla ilości hashtagów w tweecie:



Najpopularniejsze hashtagi:

- #UCL: 399 384
- #HalaMadrid? : 267 256
- #LFC : 159 313
- #ChampionsLeague? : 104 690
- #RomaBayern? : 78 728
- #RealMadrid? : 48 884

Język użytkowników:

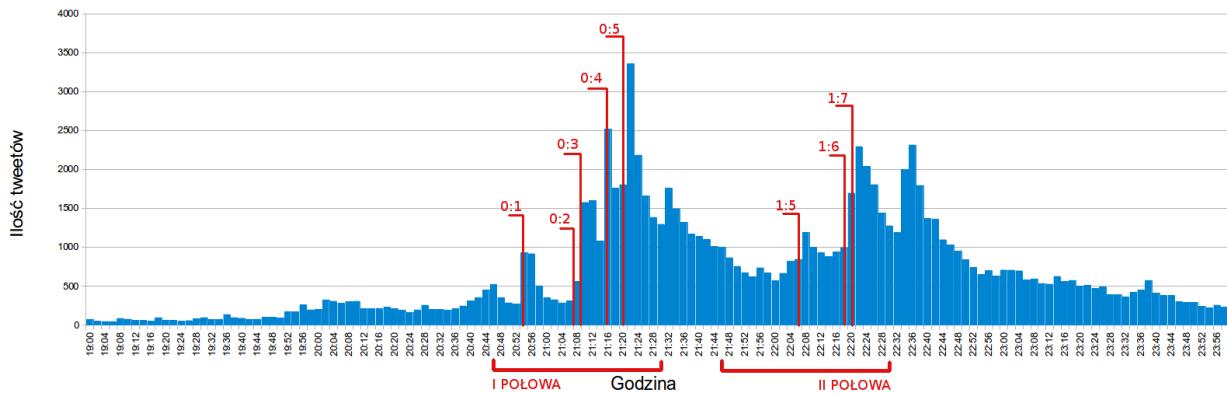


6.2. Tweety w szlagierach kolejki

Przeprowadzono analizę częstości występowania tweetów z wybranymi hashtagami, dotyczącymi najciekawszych meczów kolejki Ligi Mistrzów. Poniżej znajdują się histogramy obrazujące zebrane statystyki. Na wykresach zaznaczono minuty, w których padały gole. W ramach poszczególnych meczów wykonano wykresy zależne od hashtagów branych pod uwagę.

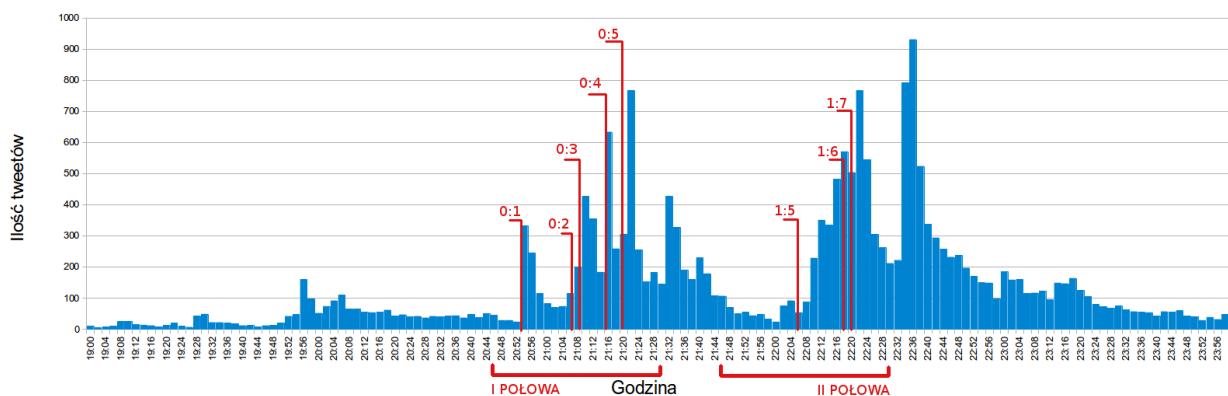
Roma vs. Bayern - wszystkie tweety

Wtorek, 21.10.2014



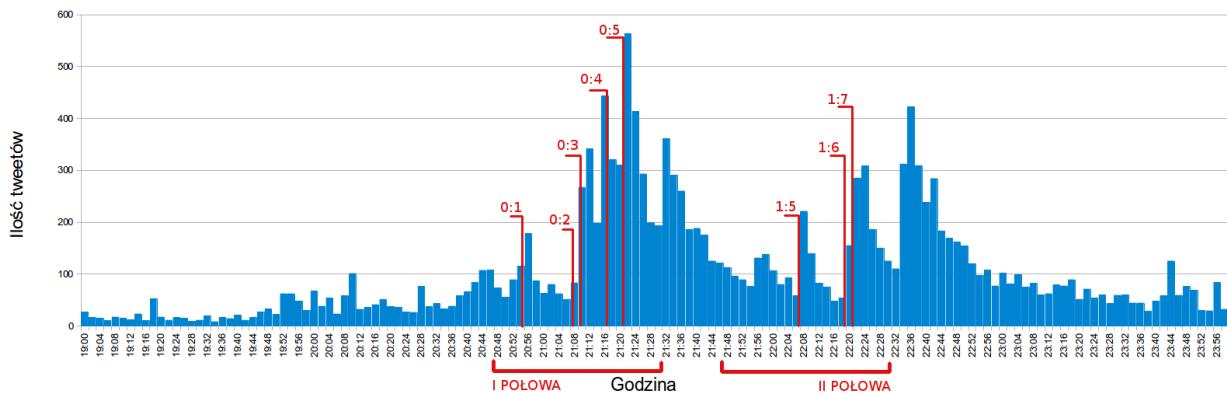
Roma vs. Bayern - #fcbayern

Wtorek, 21.10.2014



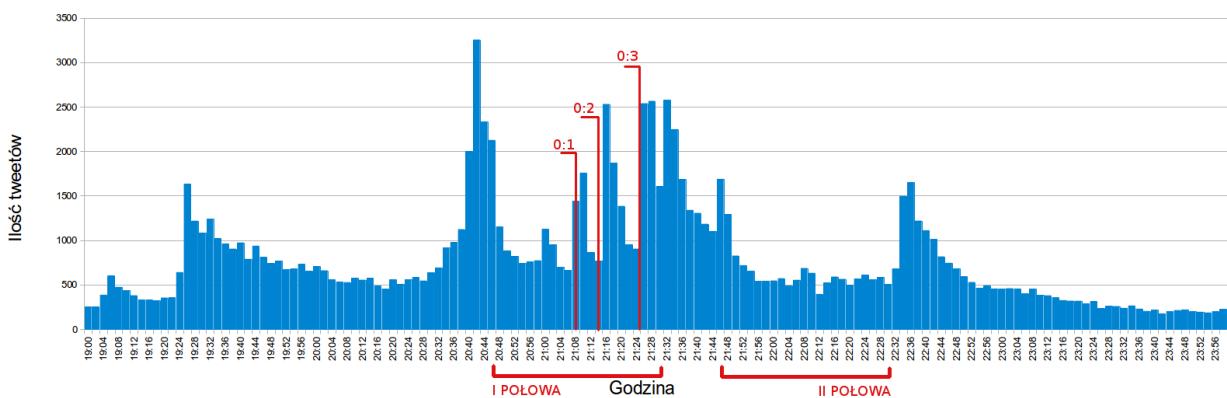
Roma vs. Bayern - #roma

Wtorek, 21.10.2014



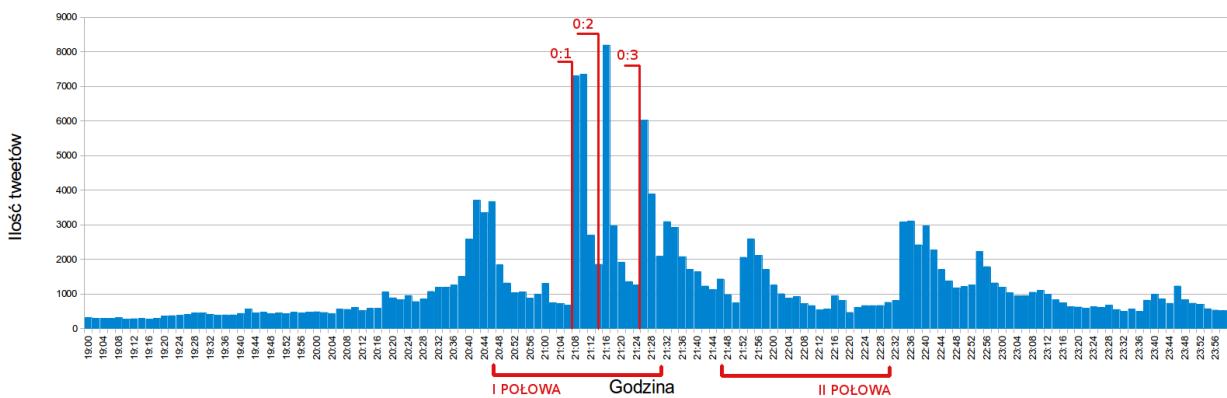
Liverpool vs. Real Madrid - #lfc & #realmadrid

Środa, 22.10.2014



Liverpool vs. Real Madrid - #cr7 & #realmadrid & #halamadrid

Środa, 22.10.2014



6.3. Analiza user mentions

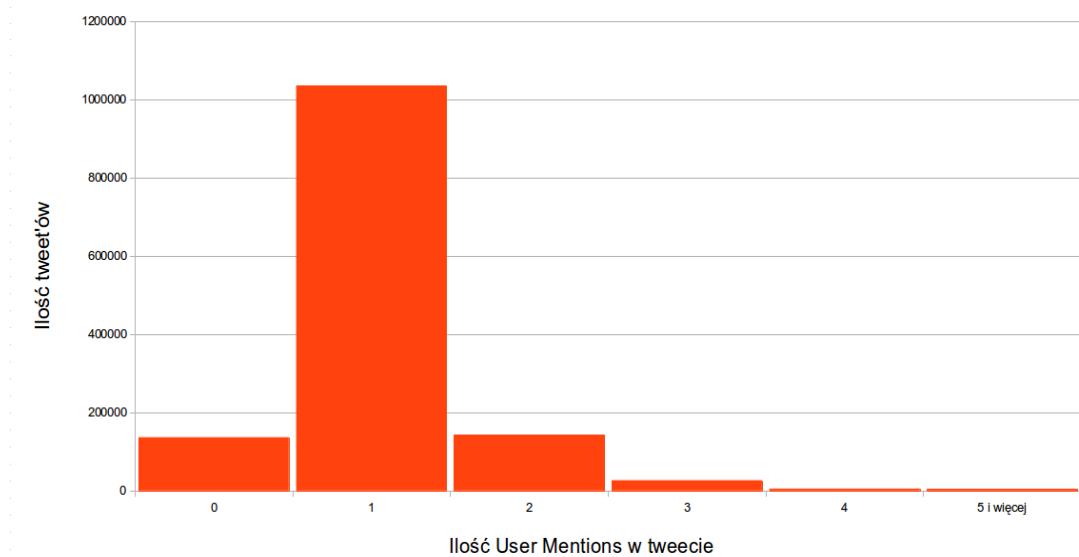
Całkowita liczba User Mentions w bazie danych: **1 577 217**

Ilość tweetów, zawierających co najmniej jedno user mention: **1 212 551**

Maksymalna ilość User Mentions w tweecie: **10**

Najczęściej wspominani użytkownicy:

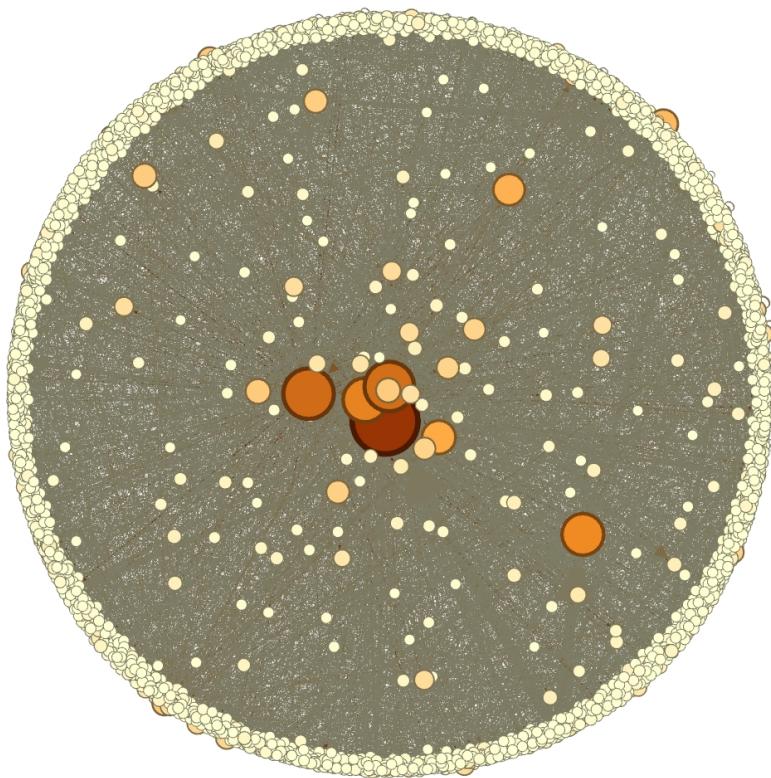
- Champions League **99393**
- Real Madrid C. F **60870**
- Liverpool FC **48971**
- Cristiano Ronald **26170**
- Mesut Özil **24591**
- MisterChip (Alex **23835**
- Real Madrid C.F. **18406**
- FC Barcelona **18395**
- Chelsea FC **17464**
- Gareth Bale **16994**



6.4. Grafowa analiza user mentions

Poniżej przedstawiono analizę User Mentions dla tweetów dotyczących meczu **AS Roma vs. Bayern FC**. Do analizy wybrano tweety z przedziału czasowego 20:30 - 23:00 w dniu 21.10.2014, w których pojawiał się dowolny z hashtagów: **#romabayern, #roma, #fcbayern**.

Wierzchołki grafu przedstawiają użytkowników Twittera. Wybrano tylko takich użytkowników, którzy chociaż raz wspominali o kimś w swoich tweetach i choć raz byli wspominani.



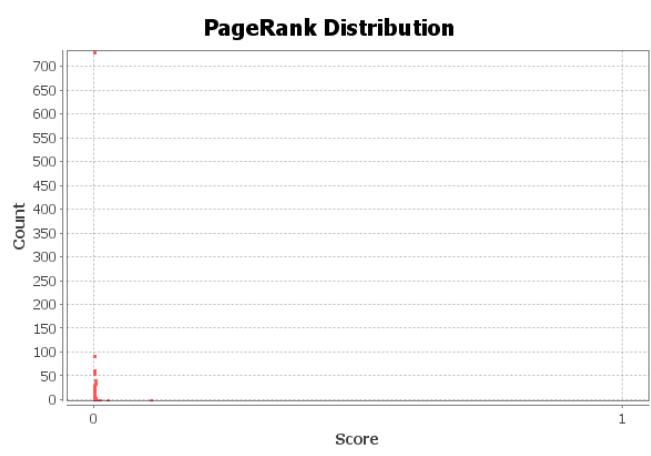
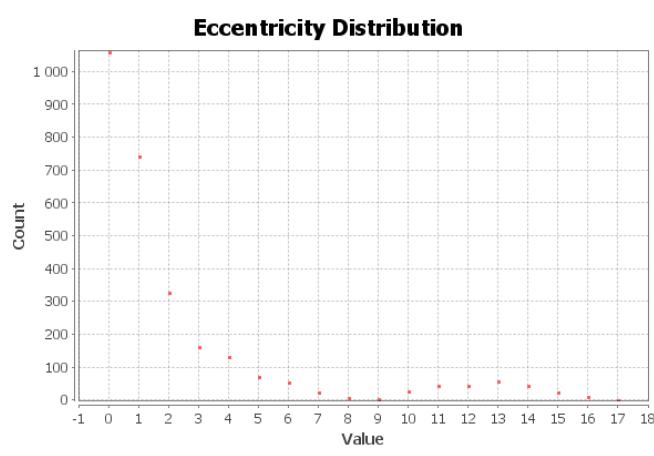
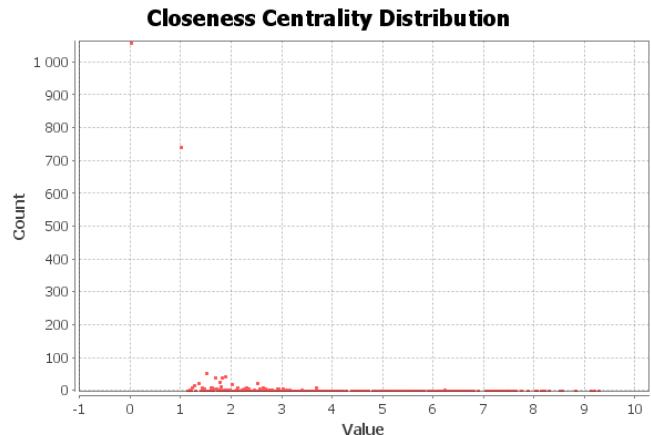
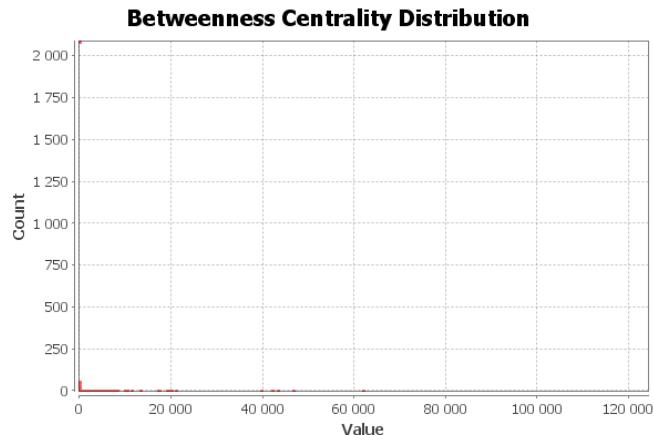
Kompletny graf:

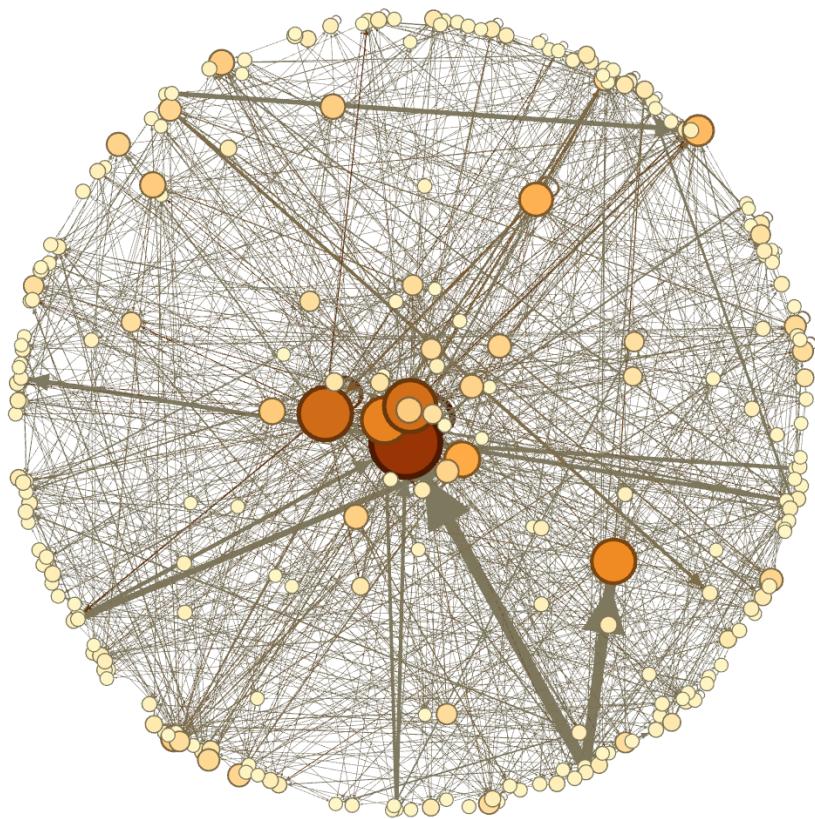
Liczba wierzchołków: **2832**

Liczba krawędzi: **6173**

Średni stopień wierzchołka: **2.18**

Średnia długość ścieżki: **5.644**





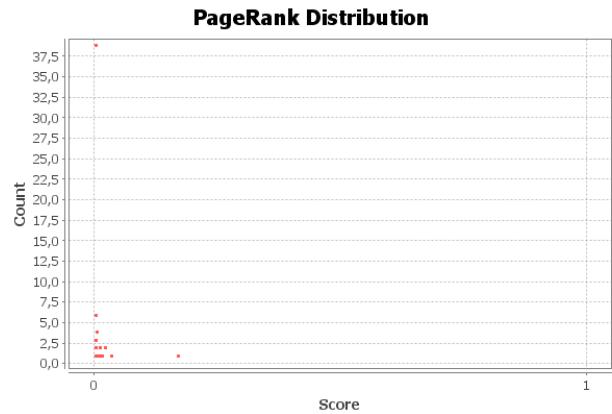
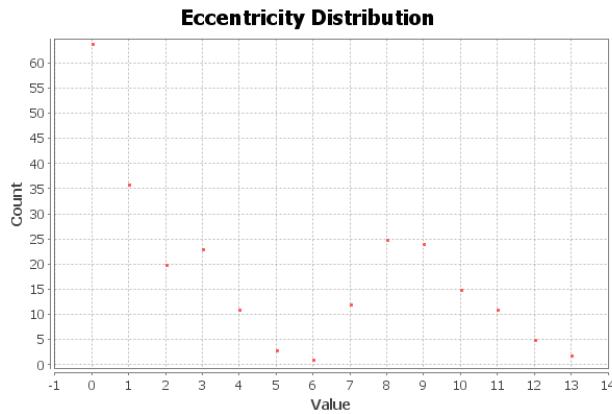
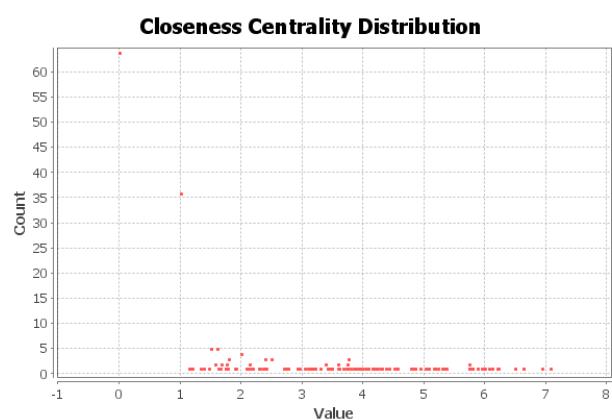
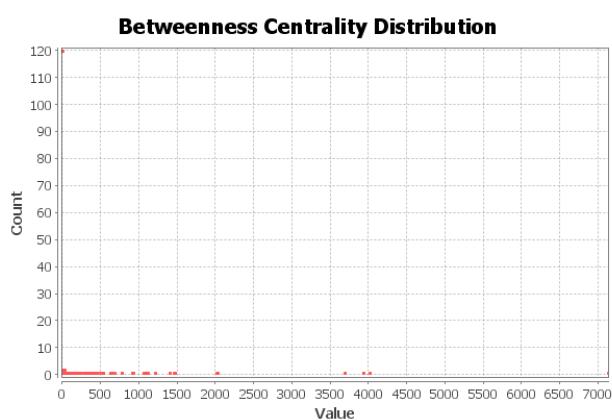
Wierzchołki o stopniu ≥ 10 :

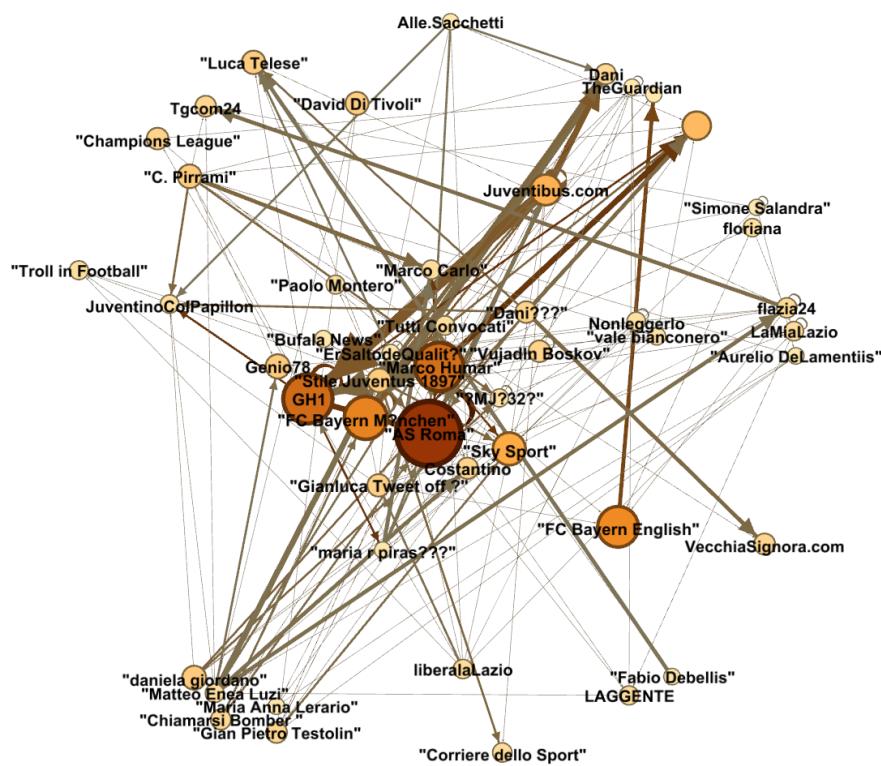
Liczba wierzchołków: **252**

Liczba krawędzi: **1187**

Średni stopień wierzchołka: **4.71**

Średnia długość ścieżki: **4.31**





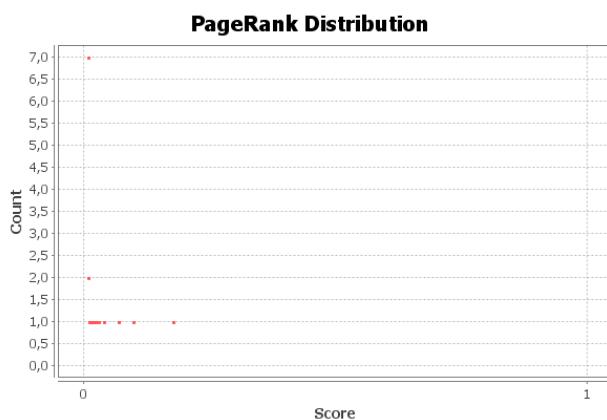
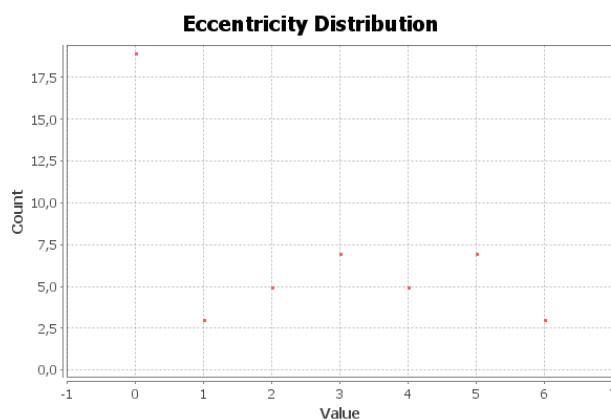
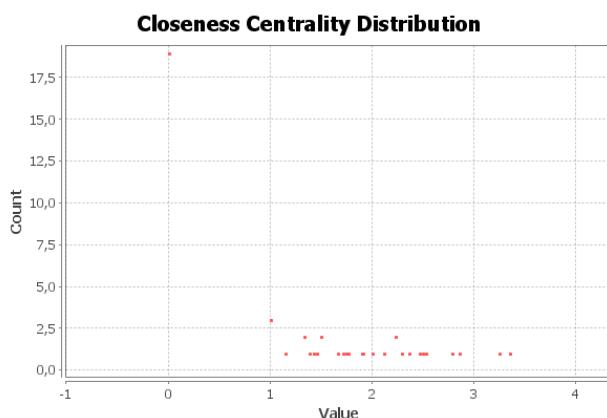
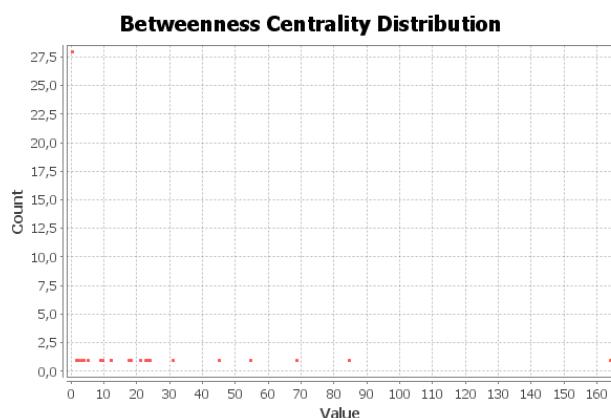
Wierzchołki o stopniu >= 25:

Liczba wierzchołków: **49**

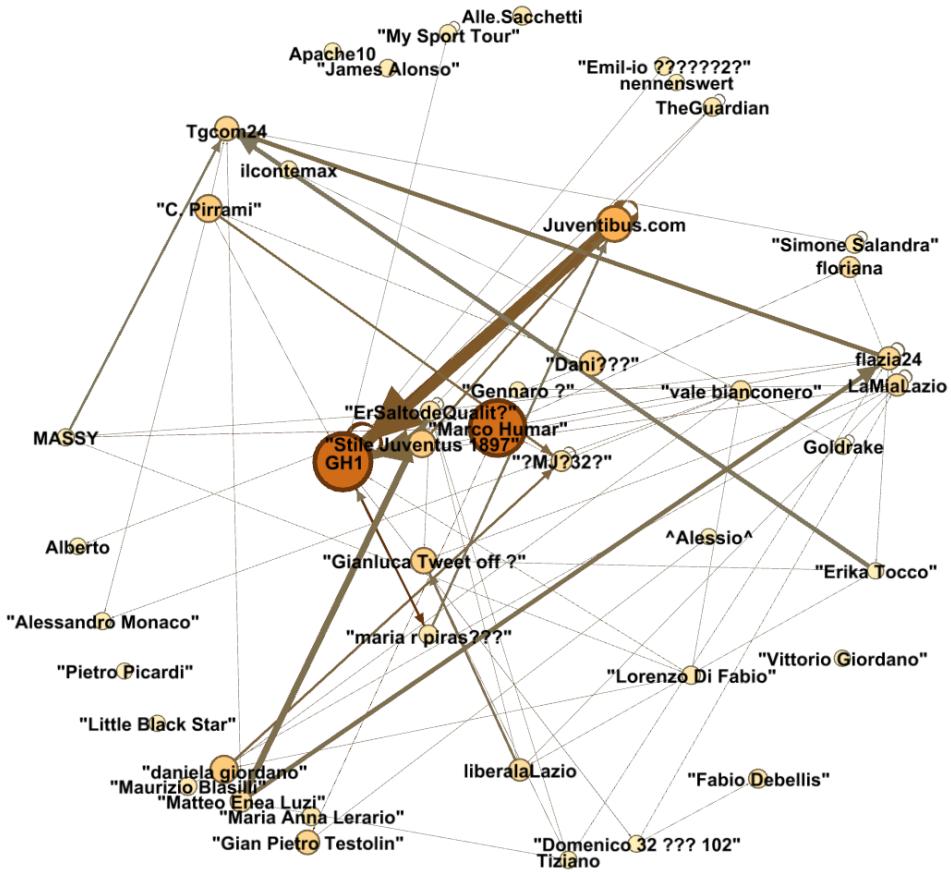
Liczba krawędzi: **171**

Średni stopień wierzchołka: **3,49**

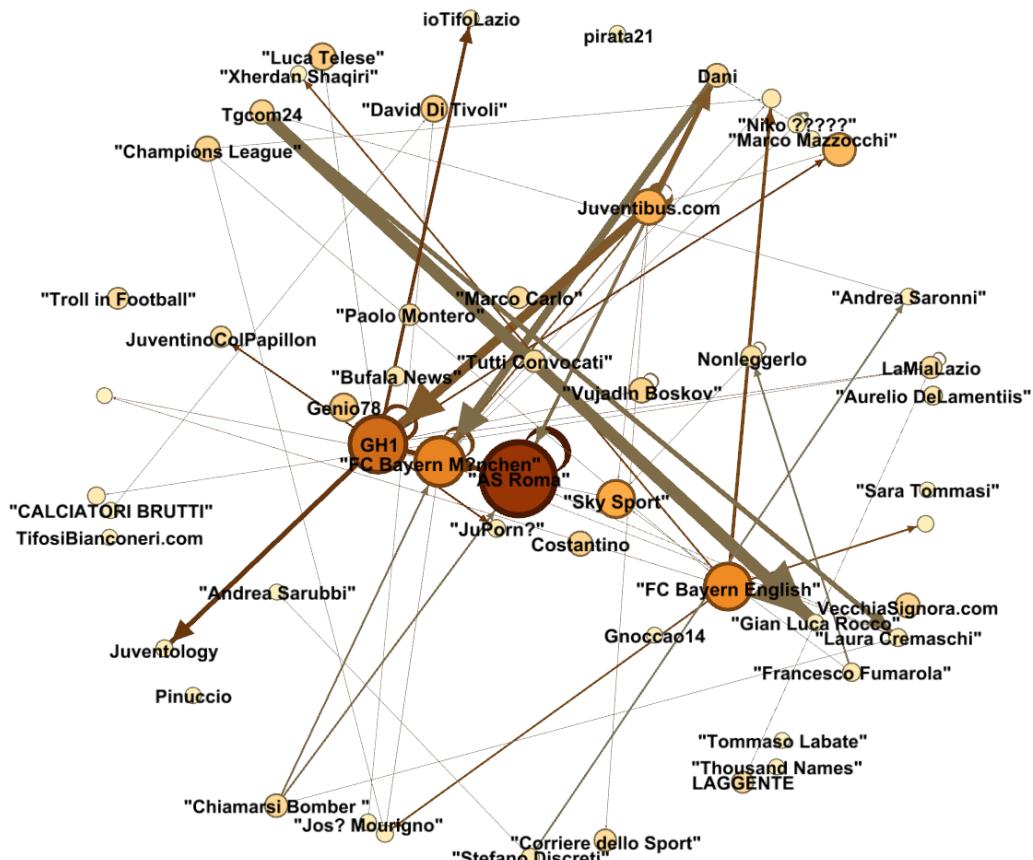
Średnia długość ścieżki: **2,294**



Wierzchołki o stopniu wyjściowym ≥ 15 (ile raz ktoś "wspominał"):



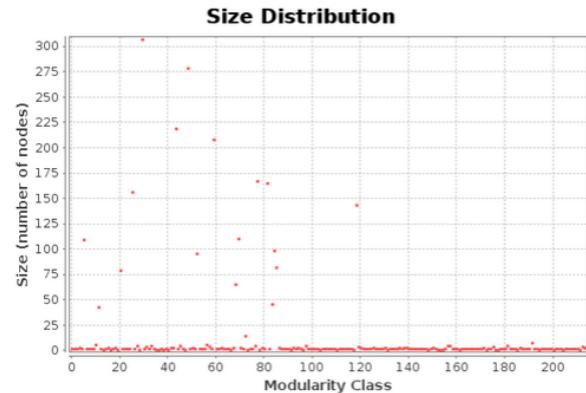
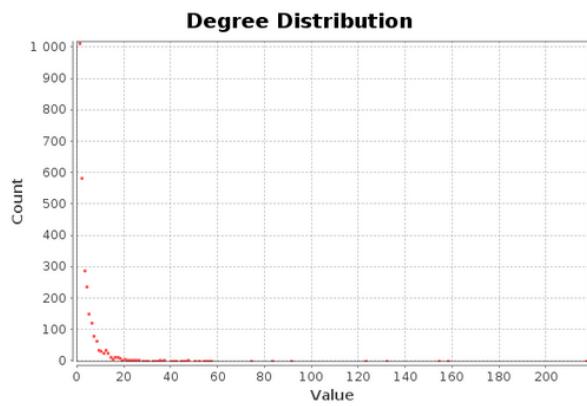
Wierzchołki o stopniu wejściowym ≥ 15 (ile raz ktoś był "wspominany"):



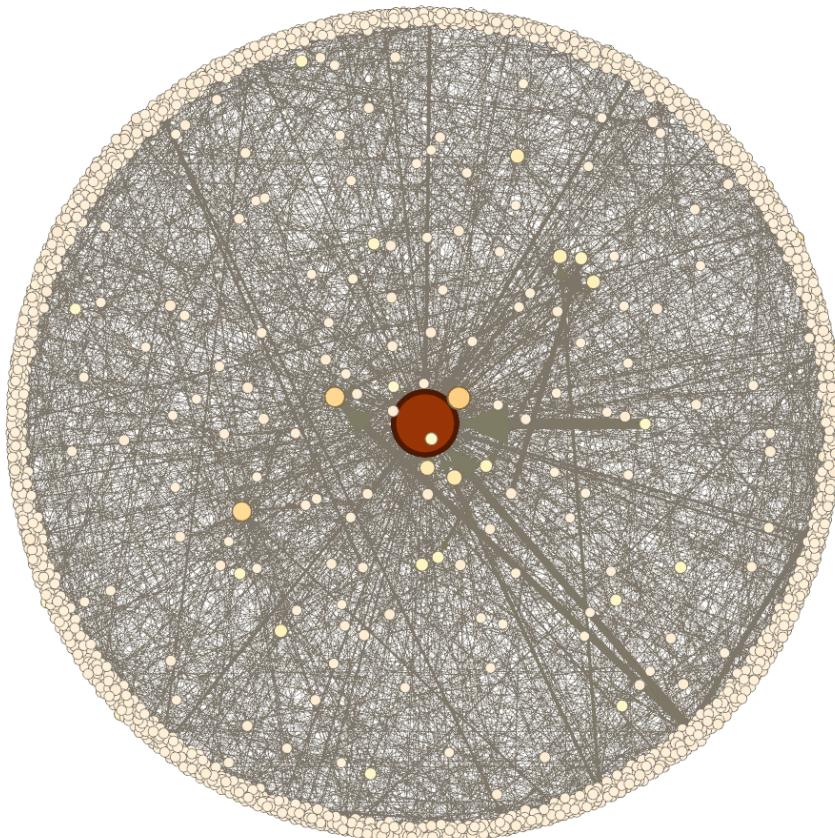
Analiza rozkładu stopni i modularności wierzchołków:

Average Degree: 2,180

Modularity: 0,567
Modularity with resolution: 0,567
Number of Communities: 214



Poniżej przedstawiono analizę User Mentions dla tweetów dotyczących meczu **Real Madryt vs. Liverpool FC**. Do analizy wybrano tweety z przedziału czasowego 20:30 - 23:00 w dniu 22.10.2014, w których pojawiał się dowolny z hashtagów: **#realmadrid, #lfc.**



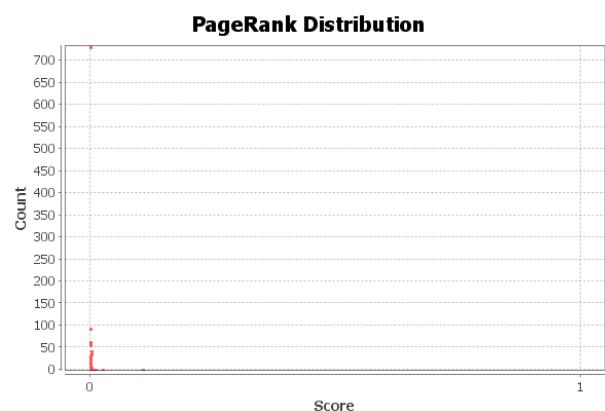
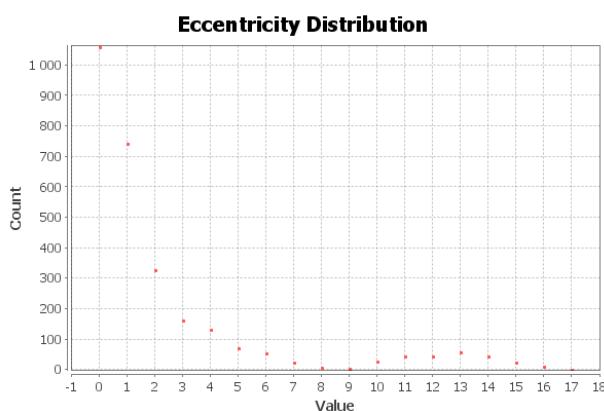
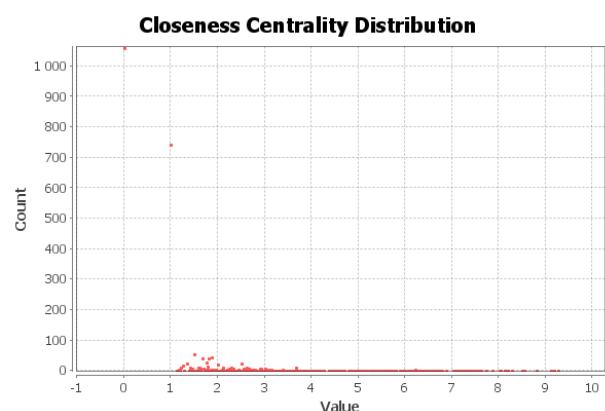
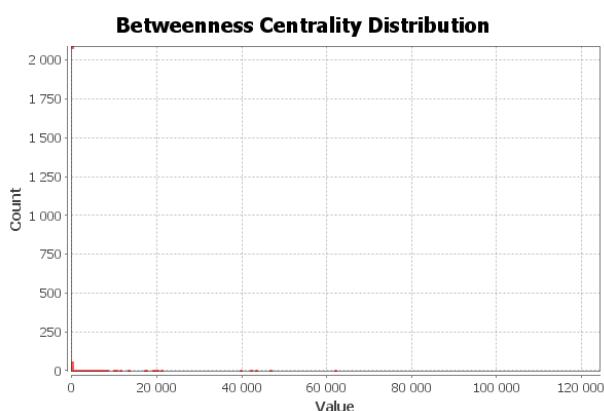
Kompletny graf:

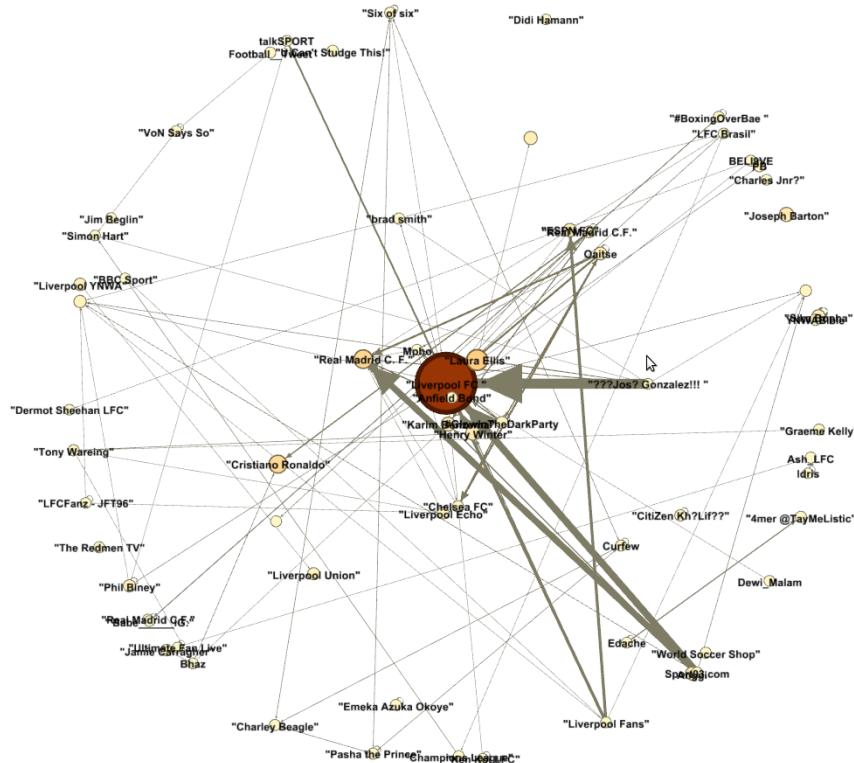
Liczba wierzchołków: **2405**

Liczba krawędzi: **3456**

Średni stopień wierzchołka: **1.437**

Średnia długość ścieżki: **2.011**





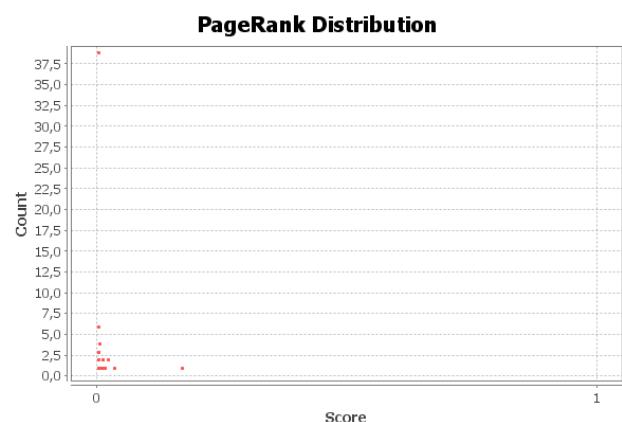
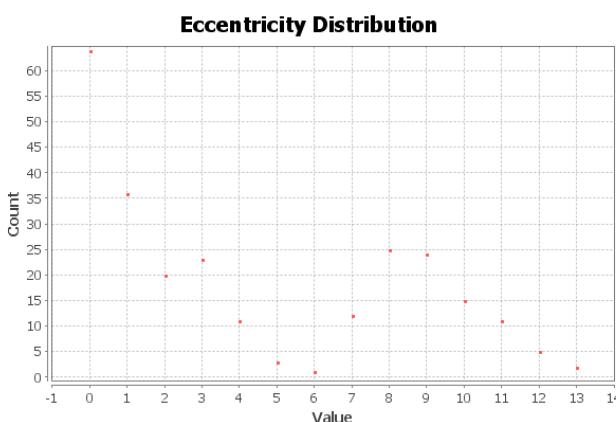
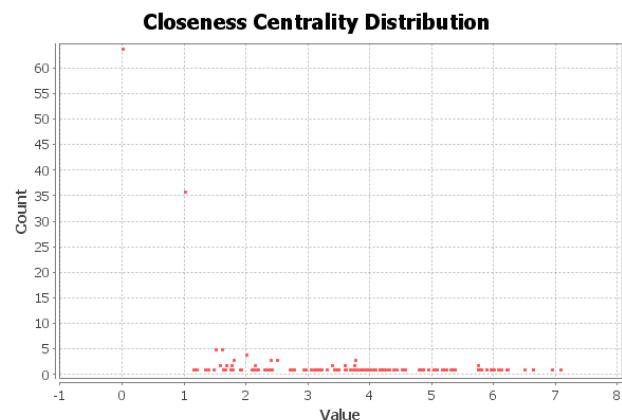
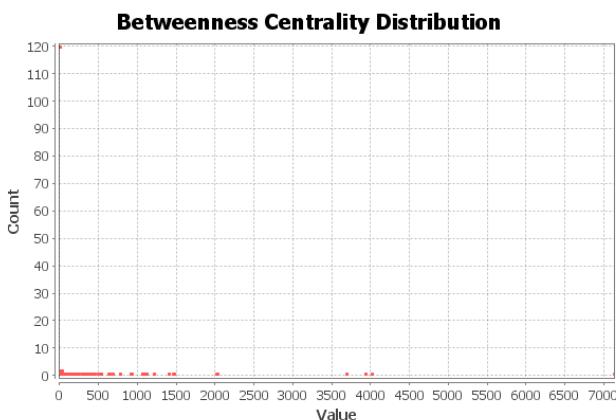
Wierzchołki o stopniu ≥ 10 :

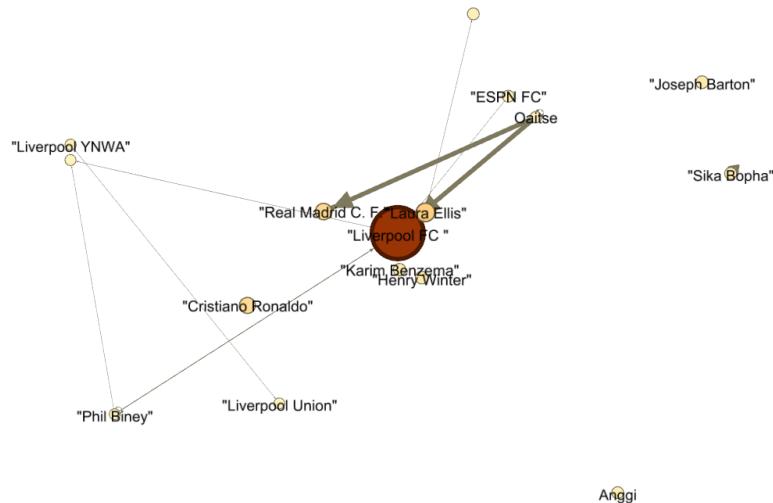
Liczba wierzchołków: **66**

Liczba krawędzi: **110**

Średni stopień wierzchołka: **1.667**

Średnia długość ścieżki: **1.569**





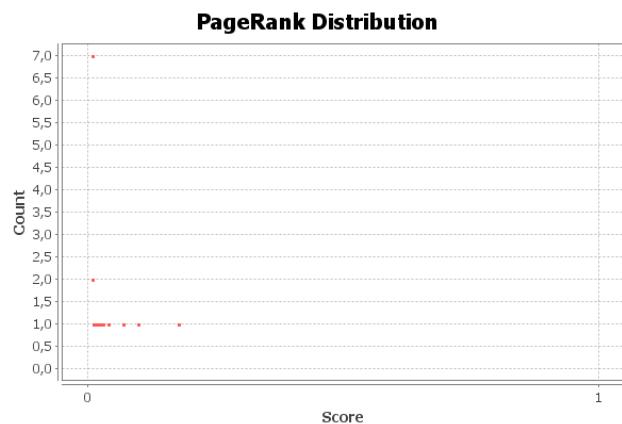
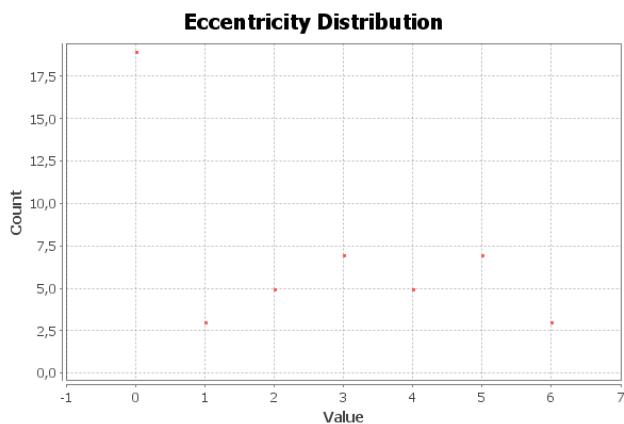
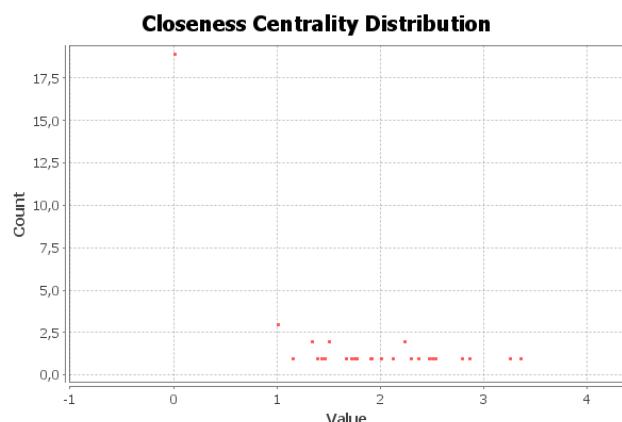
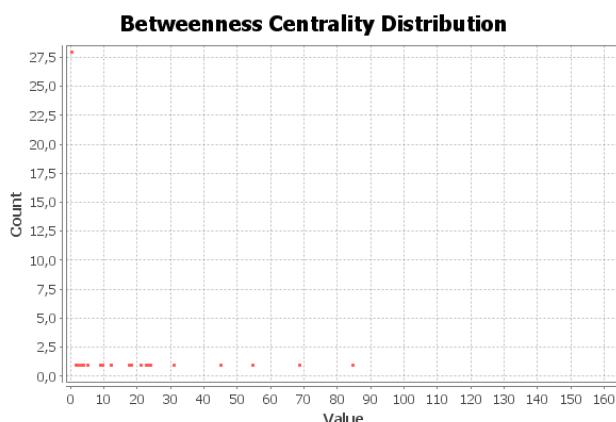
Wierzchołki o stopniu ≥ 20 :

Liczba wierzchołków: **16**

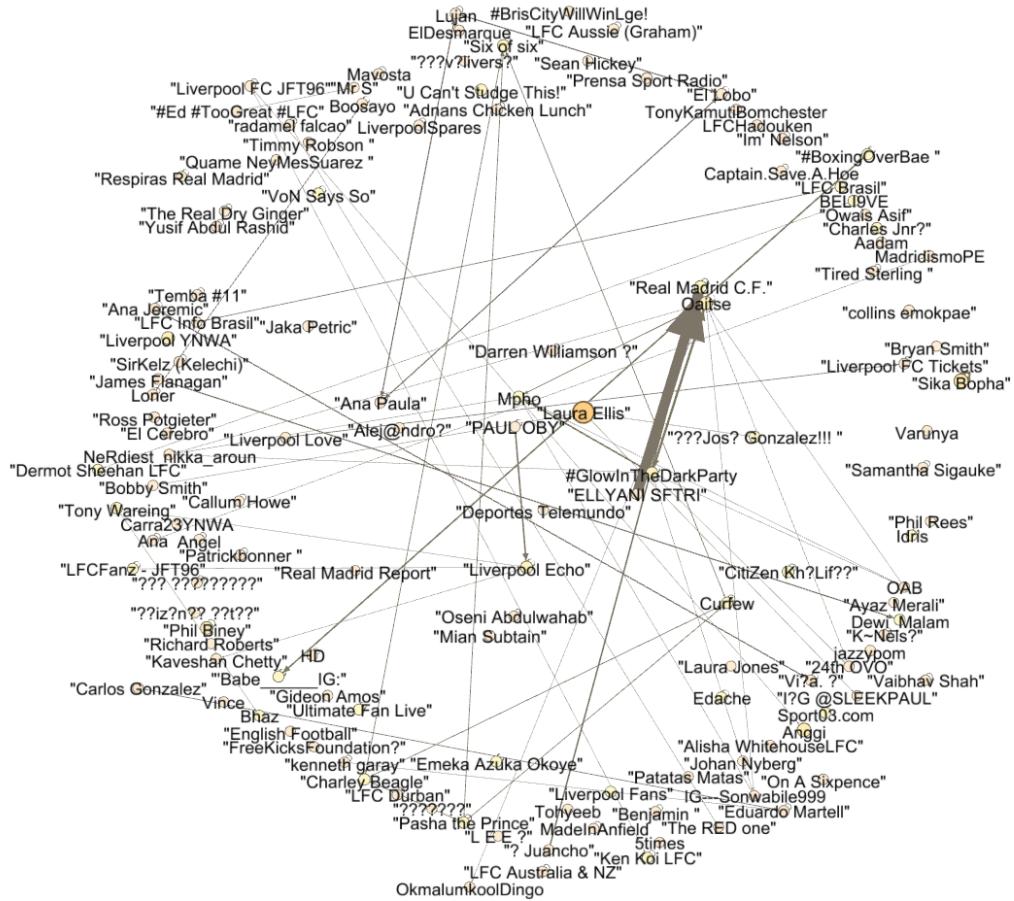
Liczba krawędzi: **12**

Średni stopień wierzchołka: **0.75**

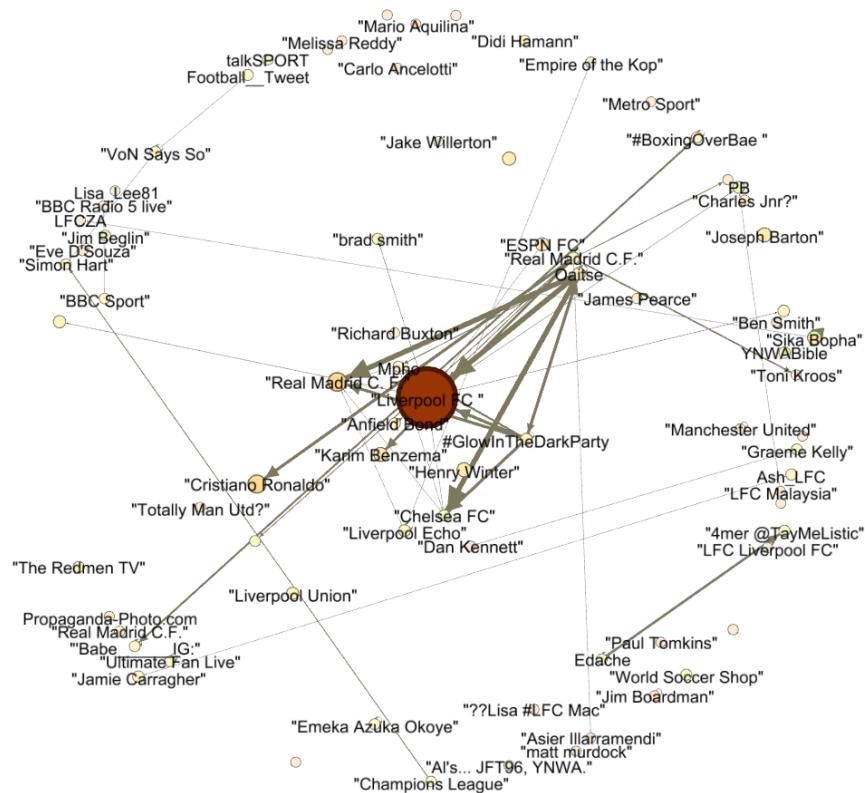
Średnia długość ścieżki: **1.25**



Wierzchołki o stopniu wyjściowym ≥ 5 (ile raz ktoś "wspominał"):



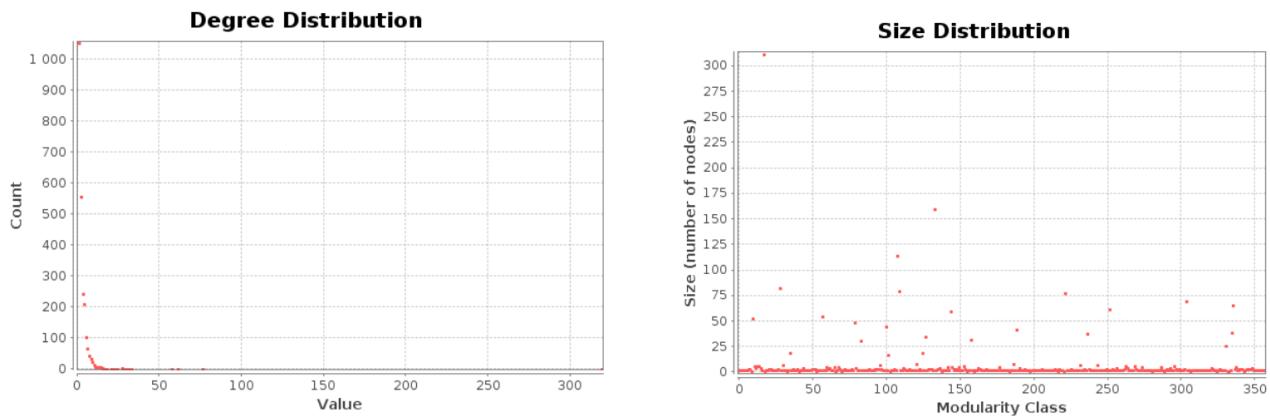
Wierzchołki o stopniu wejściowym ≥ 5 (ile raz ktoś był "wspominany"):



Analiza rozkładu stopni i modularności wierzchołków:

Average Degree: 1,437

Modularity: 0,817
Modularity with resolution: 0,817
Number of Communities: 357



6.5. Analiza retweet'ów

Poniżej przedstawiono dane statystyczne dotyczące tweetów będących "retweet'ami" - czyli postami, które zostały ponownie udostępnione przez innego użytkownika.

Liczba tweetów w bazie: **1784265**

Liczba tweetów będących retweetami: **1123033**

Procentowy udział retweetów: **62,94%**

Liczba retweetów, dla których oryg. tweet znajduje się w bazie*: **1071311**

*Z racji, że tweety zbierane były z określonego przedziału czasowego, część oryginalnych postów nie znajduje się w bazie danych.

Najchętniej retweetowane tweety:

12005 „gooooooooooooal!!! yeeeeeee! 2:1 Arsenaaaal!!! what a finish! very important!!! #ANDvAFC #UCL”

10057 „Gutted not to be there, it's been a great first 45!! Got to keep it up for the second half #Hal”

7891 „'This is Anfield' It's an honour to play in a legendary stadium like this one in charge of Real”

6543 „happy birthday, #ArseneWenger :-) wish you a blissful day and three #ucl points in belgium toni”

6242 „¡El Real Madrid ha tomado Anfield! ;) ¡Noche inolvidable! #HalaMadrid <http://t.co/GMBBtILIXC>”

Najczęściej retweetowani użytkownicy:

Champions League **81409**

Liverpool FC **35106**

FC Barcelona **30882**

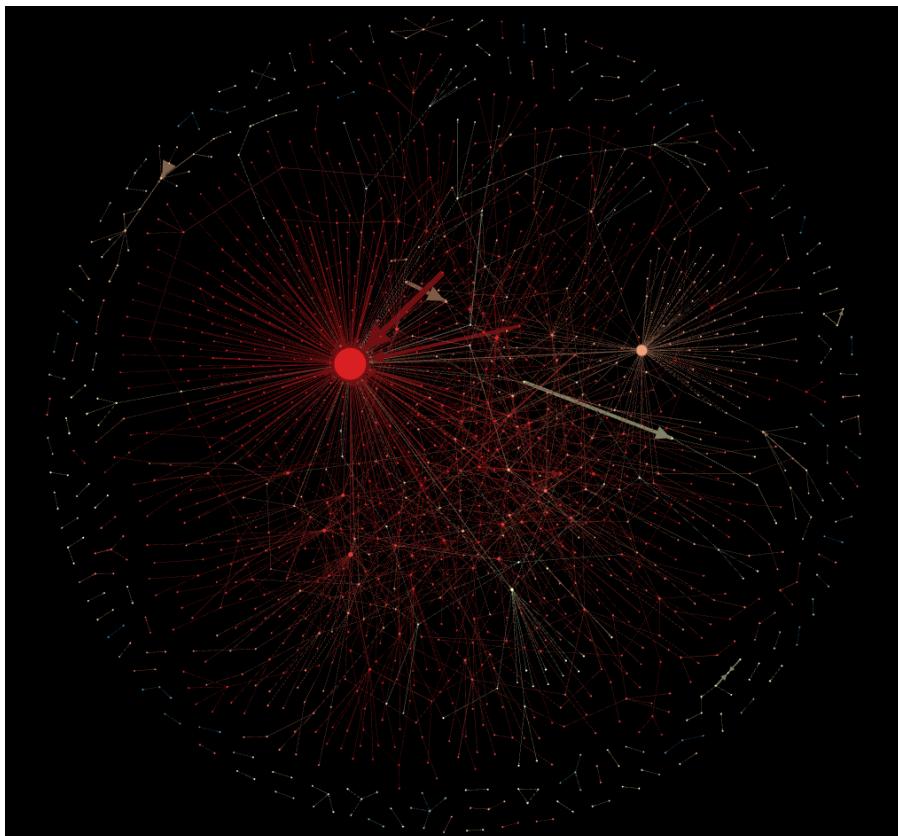
Real Madrid C. F **30660**

Mesut Özil **23533**

6.6. Grafowa analiza retweet'ów

Poniżej przedstawiono analizę retweet'ów dotyczących meczu **AS Roma vs. Bayern FC**. Do analizy wybrano tweety z przedziału czasowego 20:30 - 23:00 w dniu 21.10.2014, w których pojawiał się dowolny z hashtagów: **#romabayern, #roma, #fcbayern**.

Wierzchołki grafu przedstawiają użytkowników Twittera. Wybrano tylko takich użytkowników, którzy chociaż raz retweetowali i choć raz byli retweetowani.



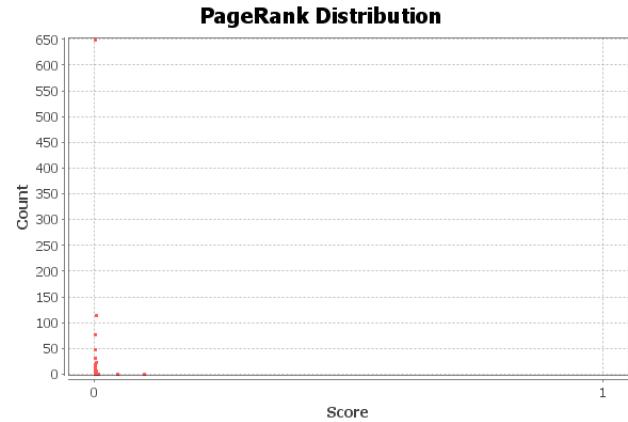
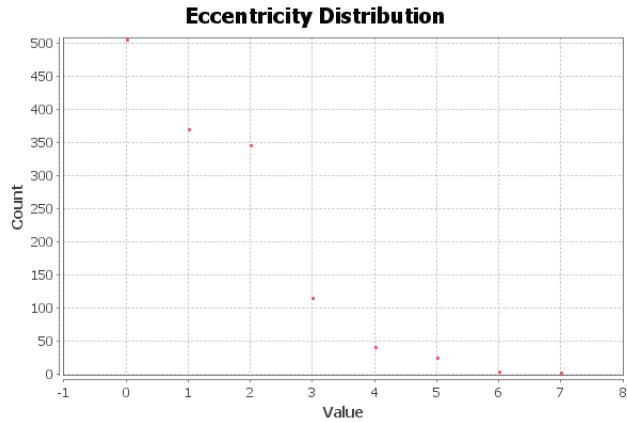
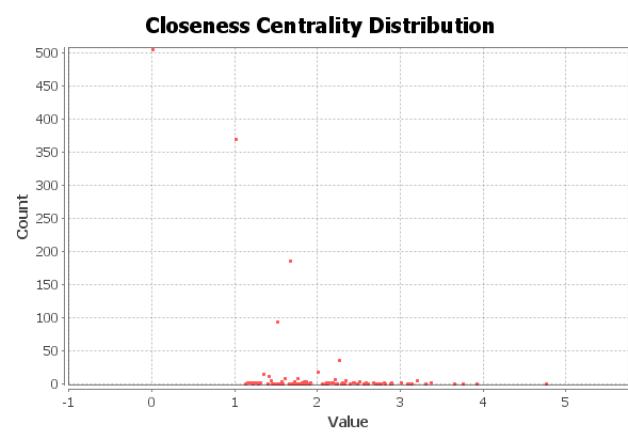
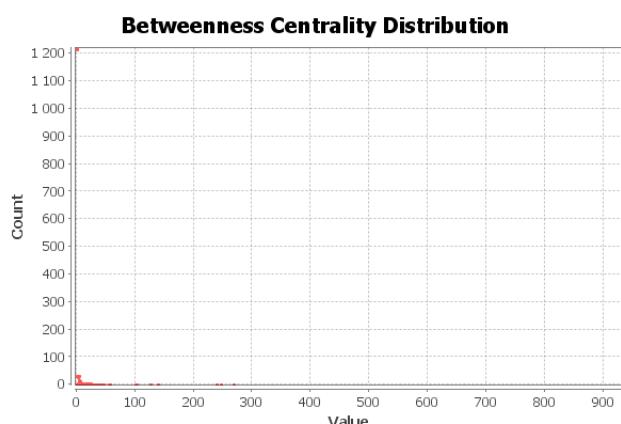
Kompletny graf:

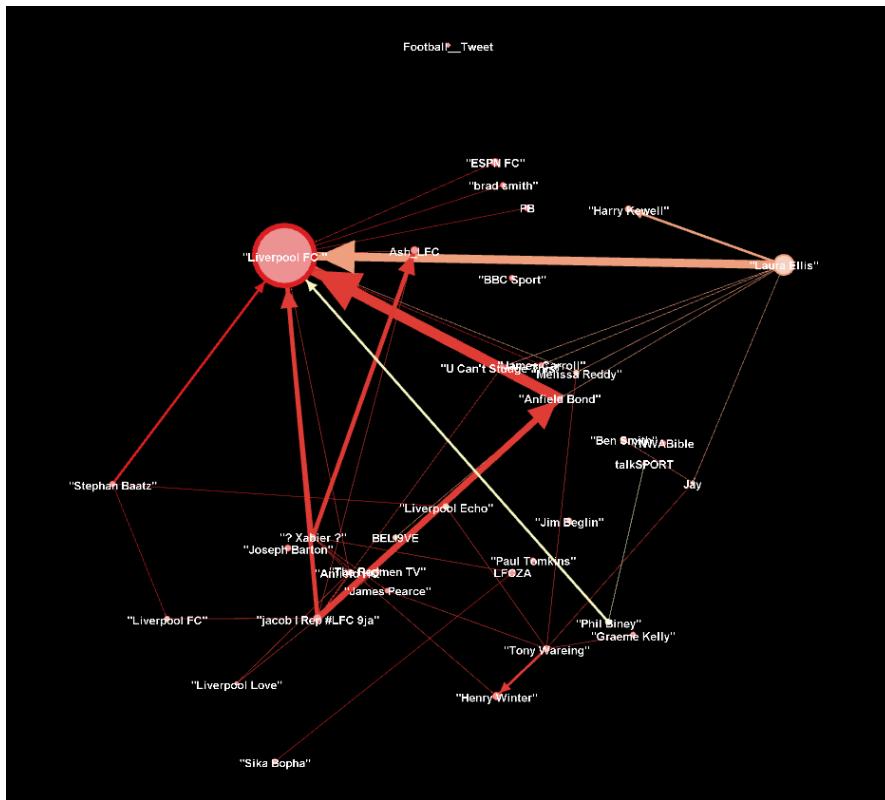
Ilość wierzchołków: **1412**

Ilość krawędzi: **1767**

Średni stopień wierzchołka: **1.251**

Średnia długość ścieżki: **1.878**





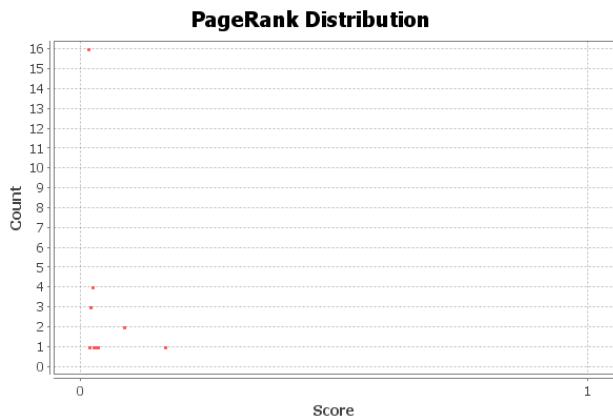
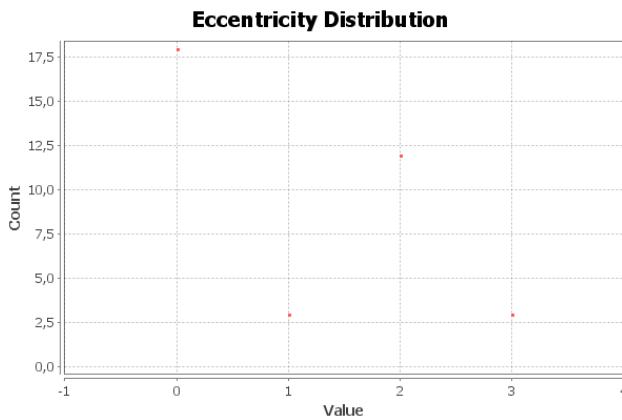
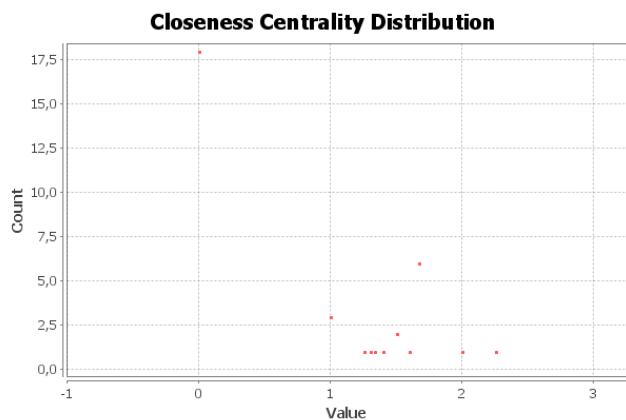
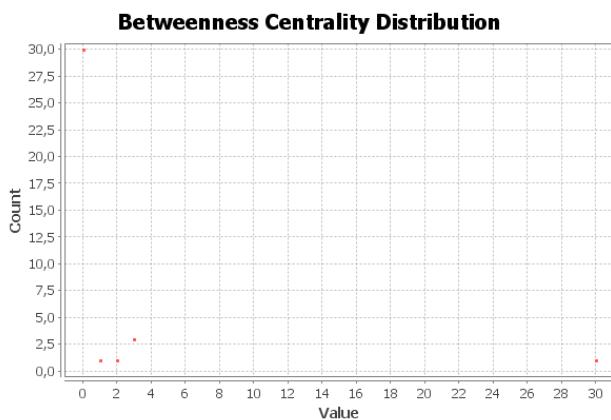
Wierzchołki o stopniu 10 lub wyższym:

Ilość wierzchołków: **36**

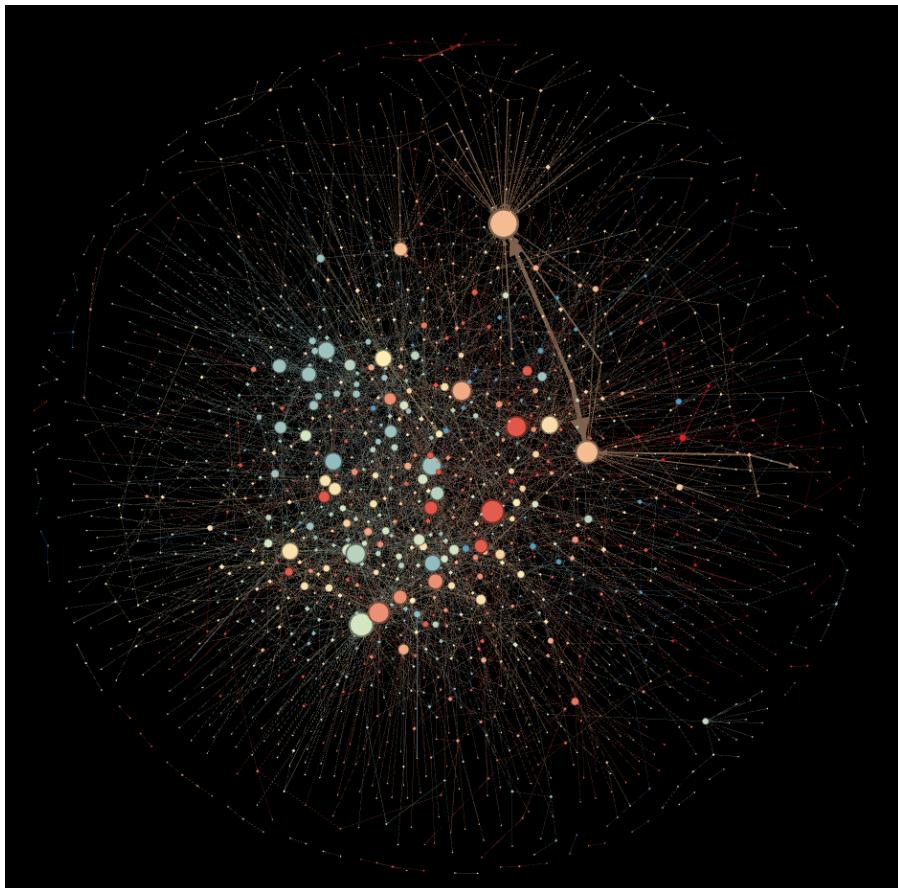
Ilość krawędzi: **45**

Średni stopień wierzchołka: **1.25**

Średnia długość ścieżki: **1.519**



Poniżej przedstawiono analizę retweet'ów dotyczących meczu **Real Madryt vs. Liverpool FC**. Do analizy wybrano tweety z przedziału czasowego 20:30 - 23:00 w dniu 22.10.2014, w których pojawiał się dowolny z hashtagów: **#realmadrid, #lfc**.



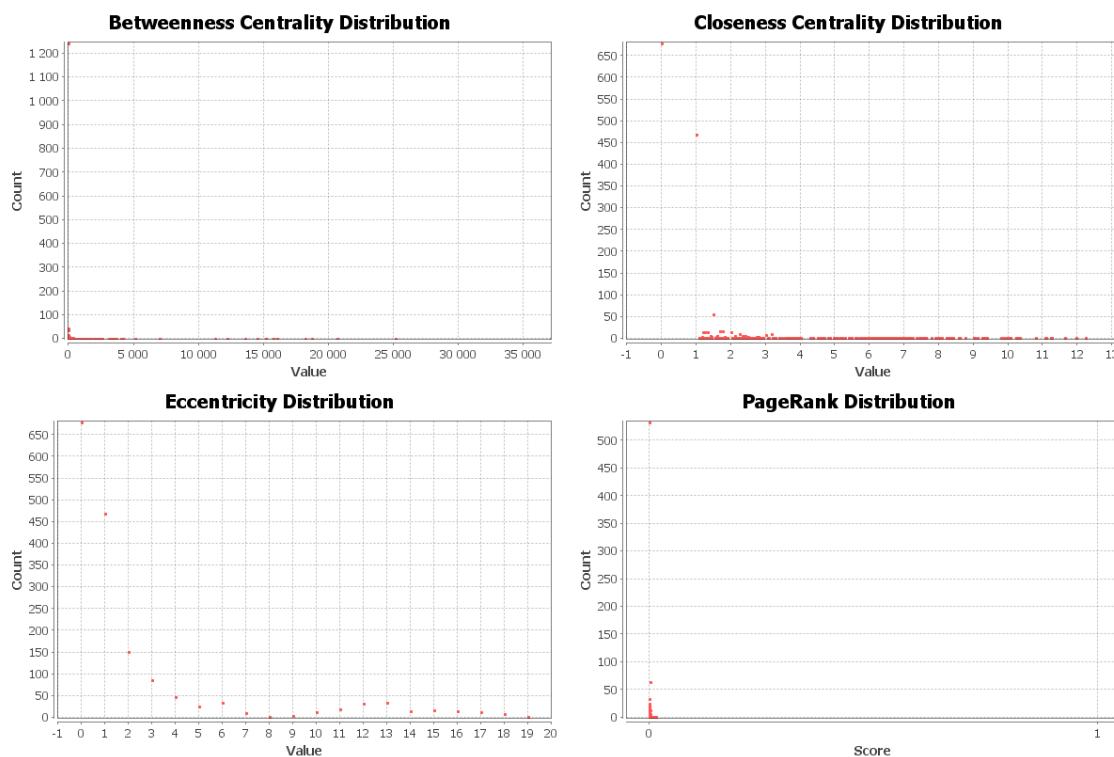
Kompletny graf:

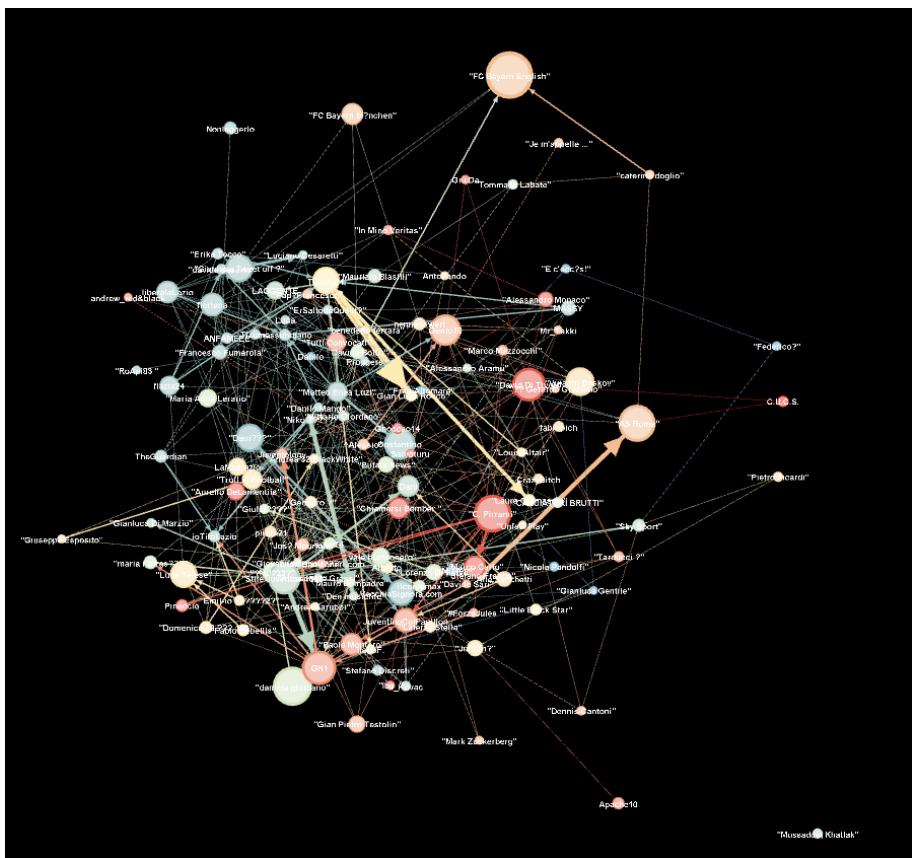
Ilość wierzchołków: **1683**

Ilość krawędzi: **3139**

Średni stopień wierzchołka: **1.863**

Średnia długość ścieżki: **6.726**





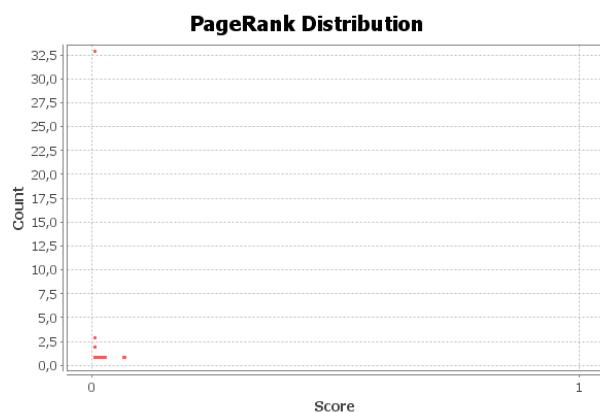
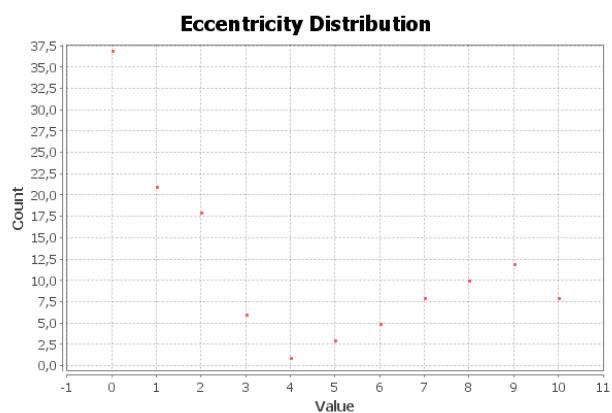
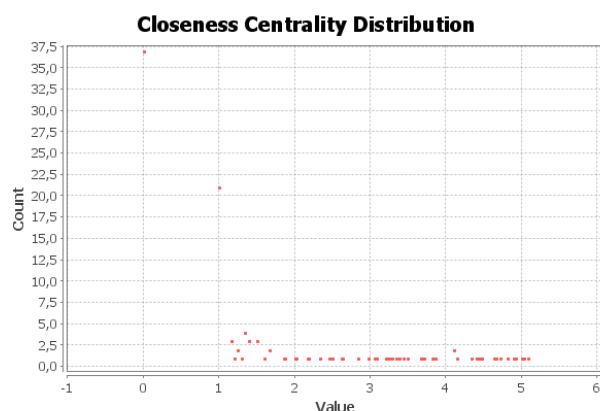
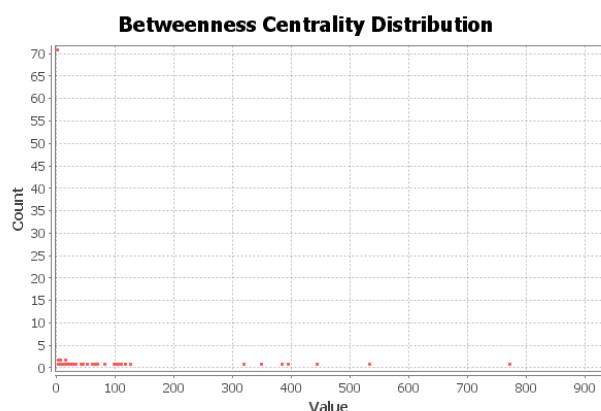
Wierzchołki o poziomie 10 i wyższym:

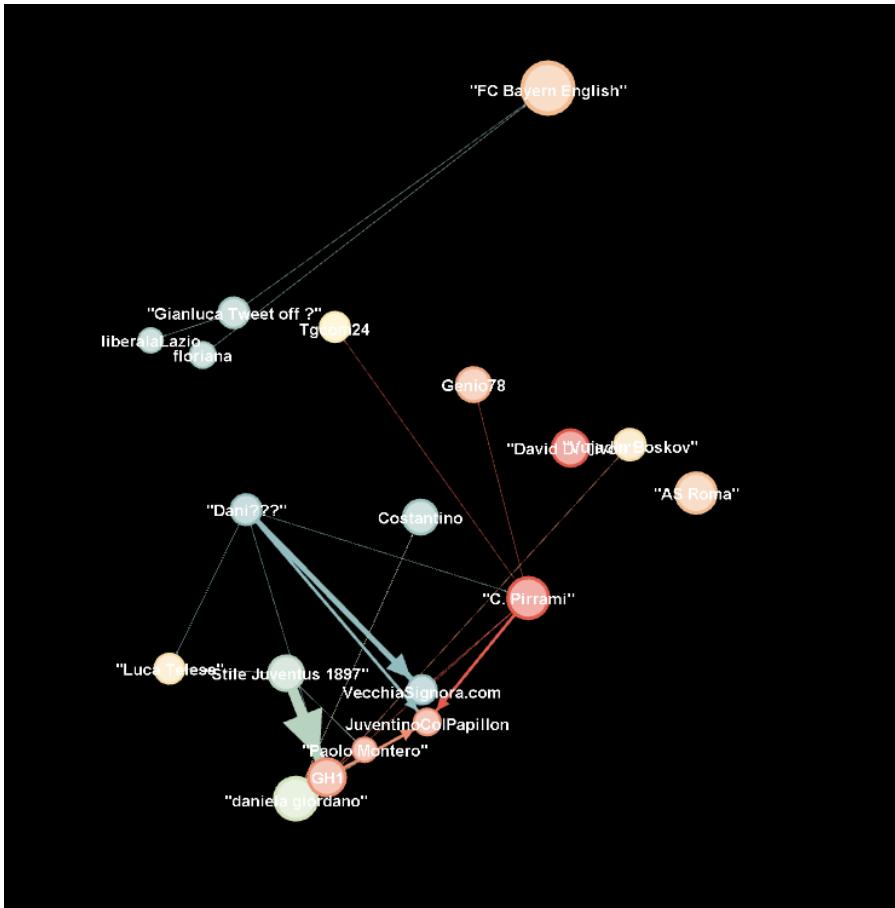
Ilość wierzchołków: **129**

Ilość krawędzi: **405**

Średni stopień wierzchołka: **3,14**

Średnia długość ścieżki: **3,483**





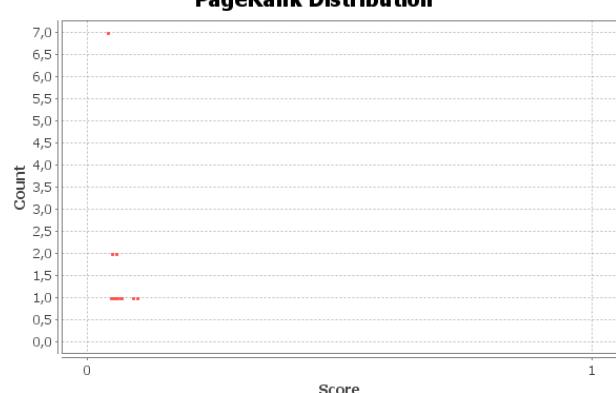
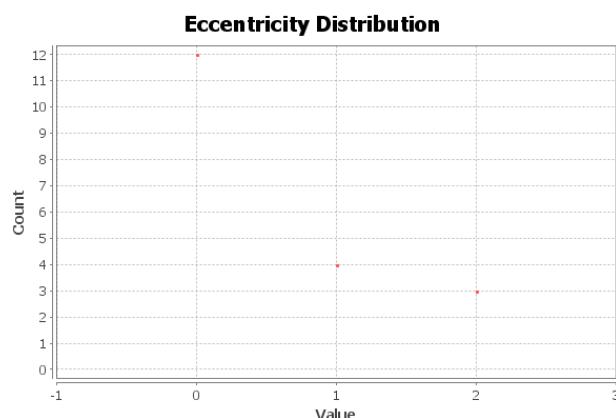
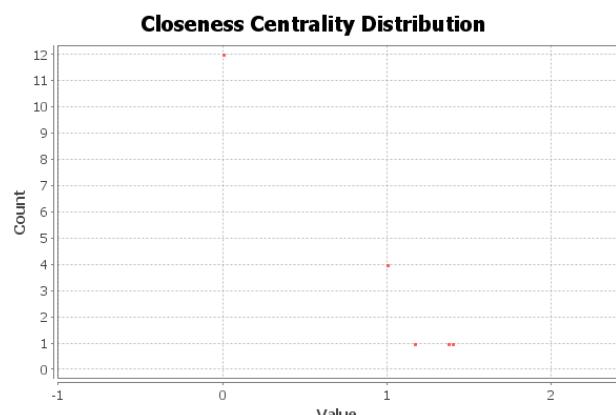
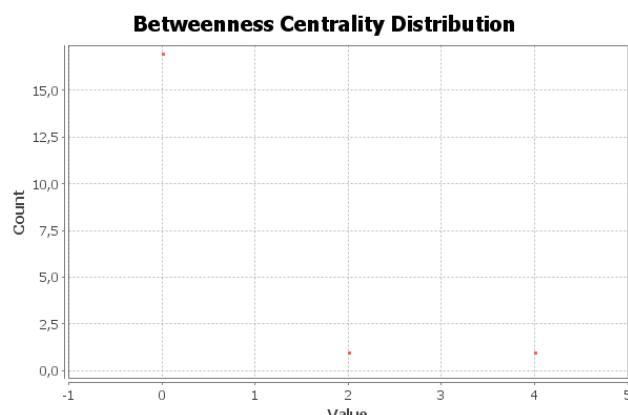
Wierzchołki o poziomie 20 i wyższym:

Ilość wierzchołków: **19**

Ilość krawędzi: **20**

Średni stopień wierzchołka: **1.053**

Średnia długość ścieżki: **1.231**



6.7. Wyszukiwanie społeczności w programie „Gephi”

Przeanalizowano wcześniej stworzone grafy retweet'ów pod kątem społeczności.

AS Roma - FC Bayern Monachium

W całym grafie znaleziono **81** społeczności.

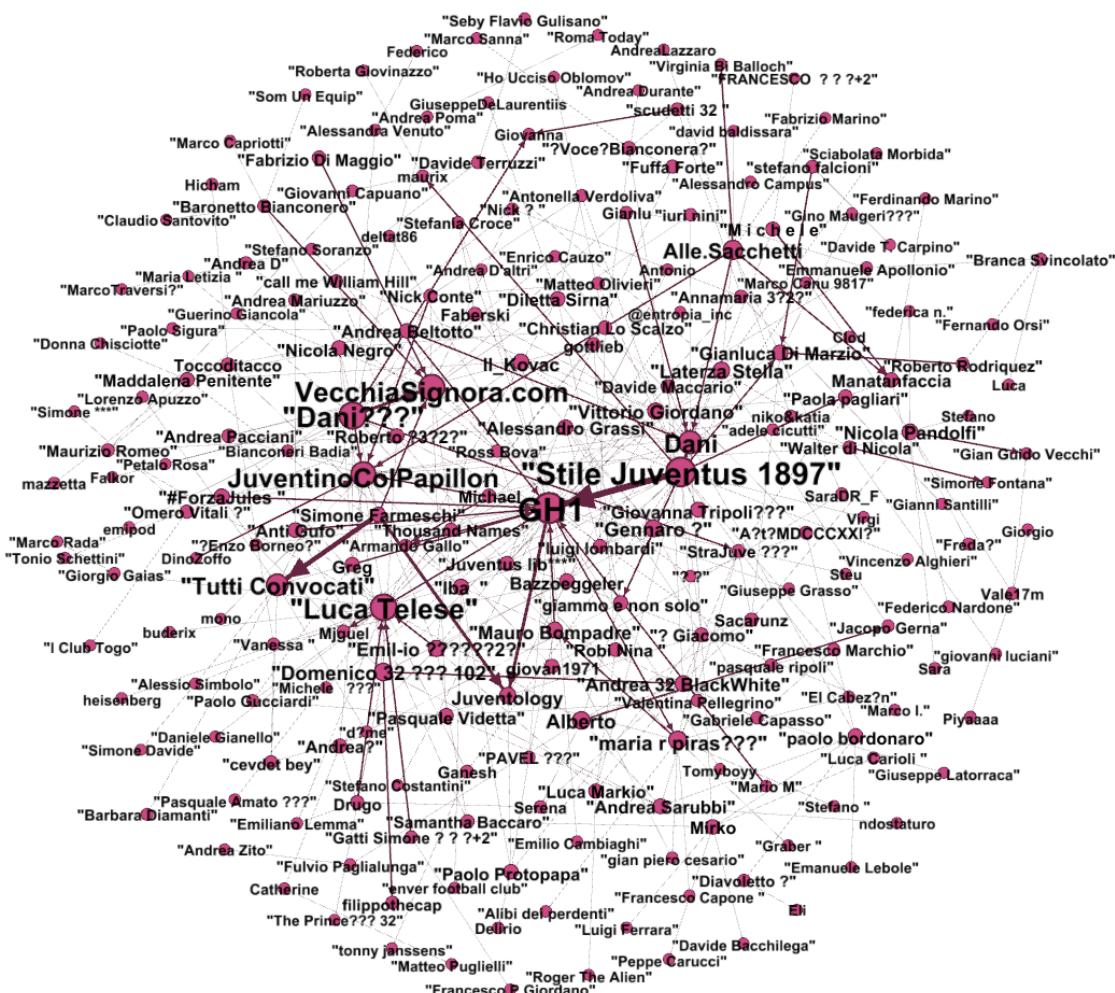
Największe społeczności liczą kolejno **220**, **187**, **135**, oraz **124** węzły.

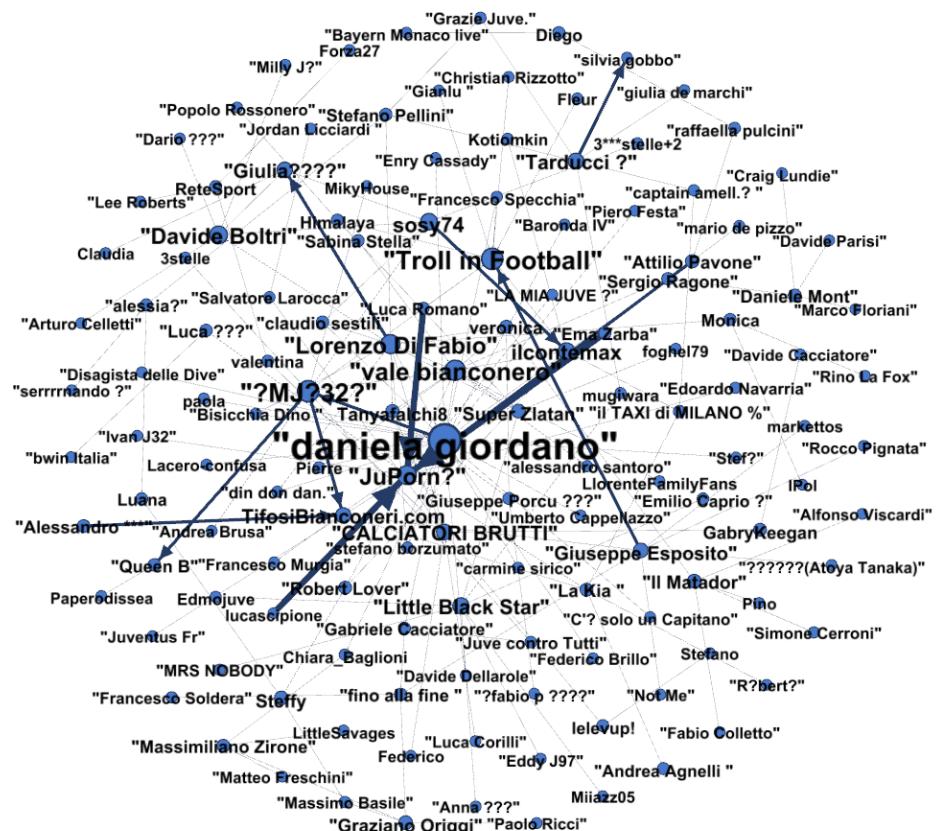
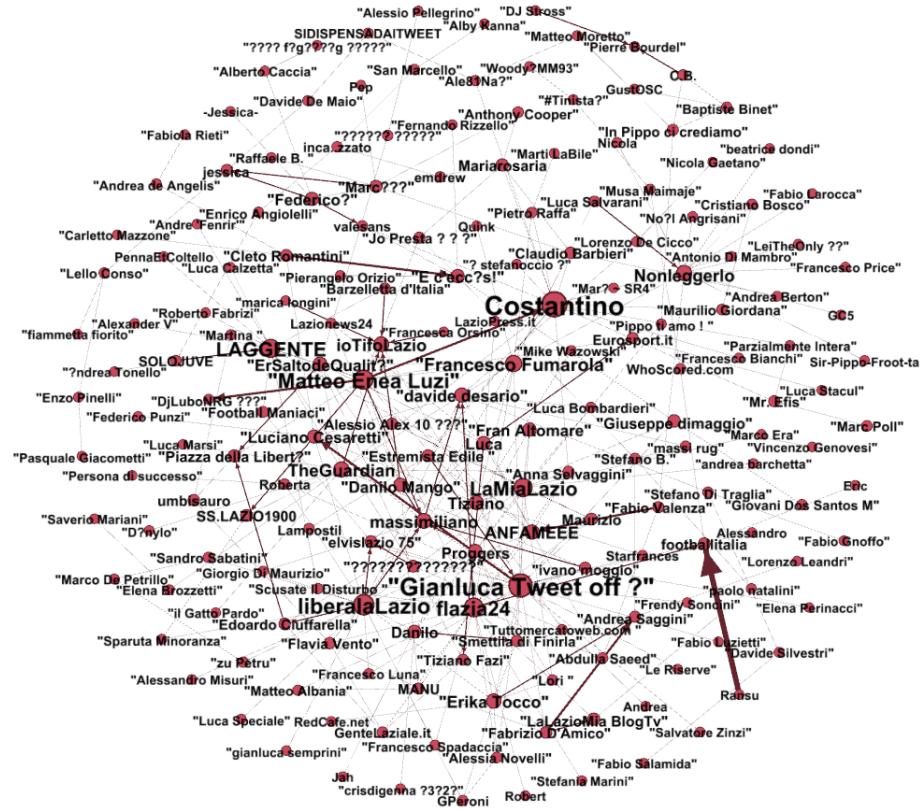
Modularność grafu wyniosła **0,61**.

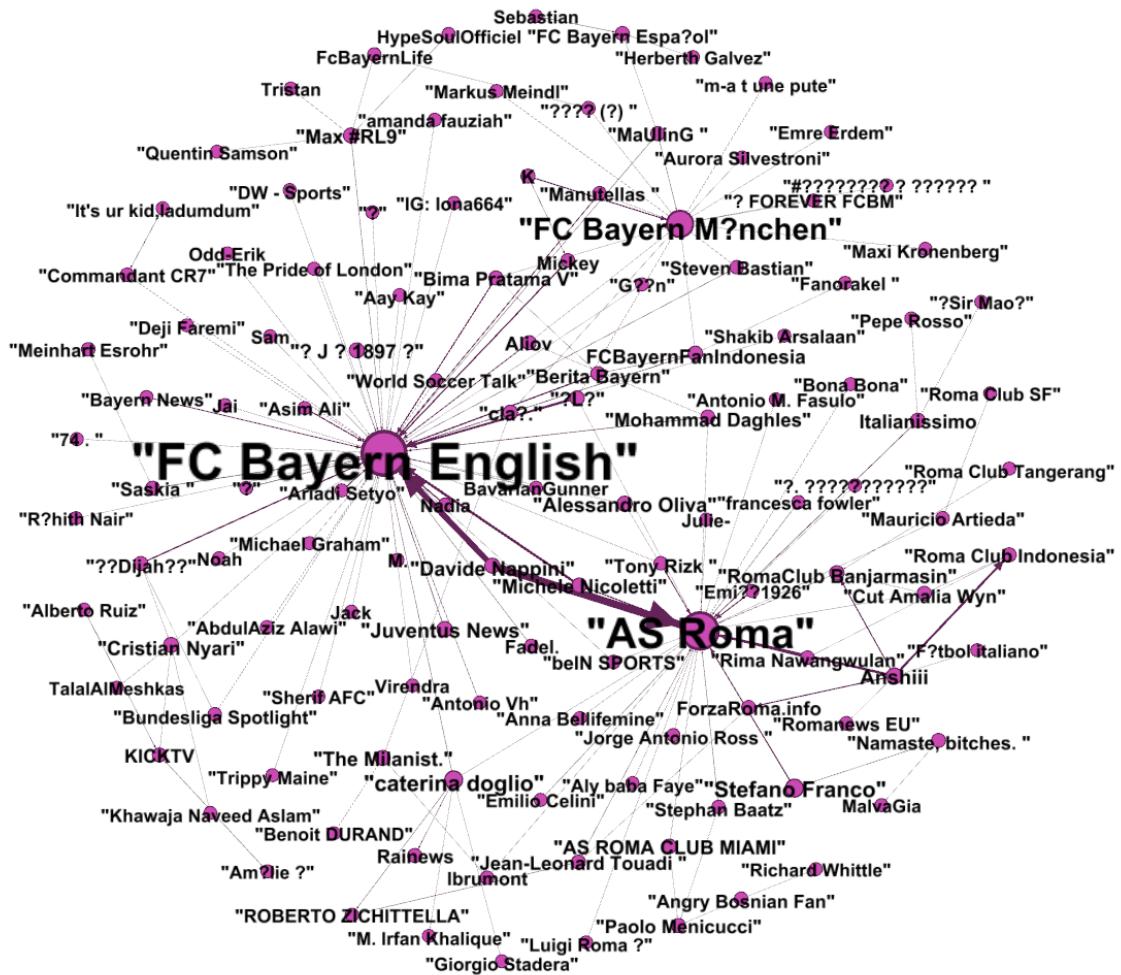
W społecznościach wyraźnie dominują użytkownicy o narodowości włoskiej.

- Wszyscy użytkownicy: **1683**
 - IT: **1319**
 - EN: **227**
 - inne: **137**

Poniżej umieszczono grafy dla 4 największych społeczności. Wielkość węzła jest zależna od jego stopnia.







Liverpool FC - Real Madryt

W całym grafie znalezioneo **161** społeczności.

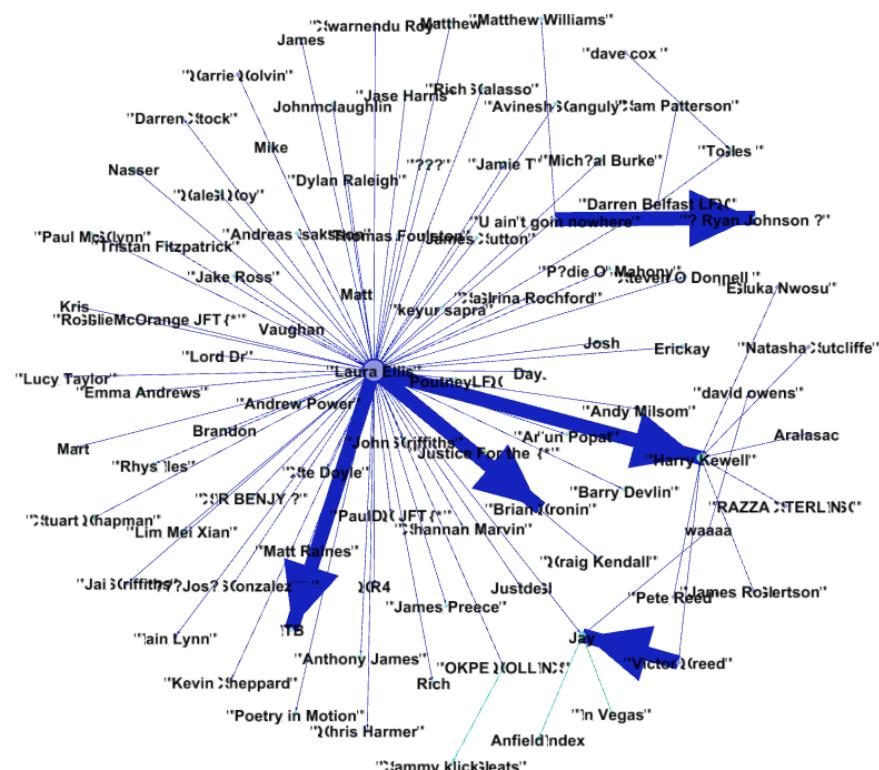
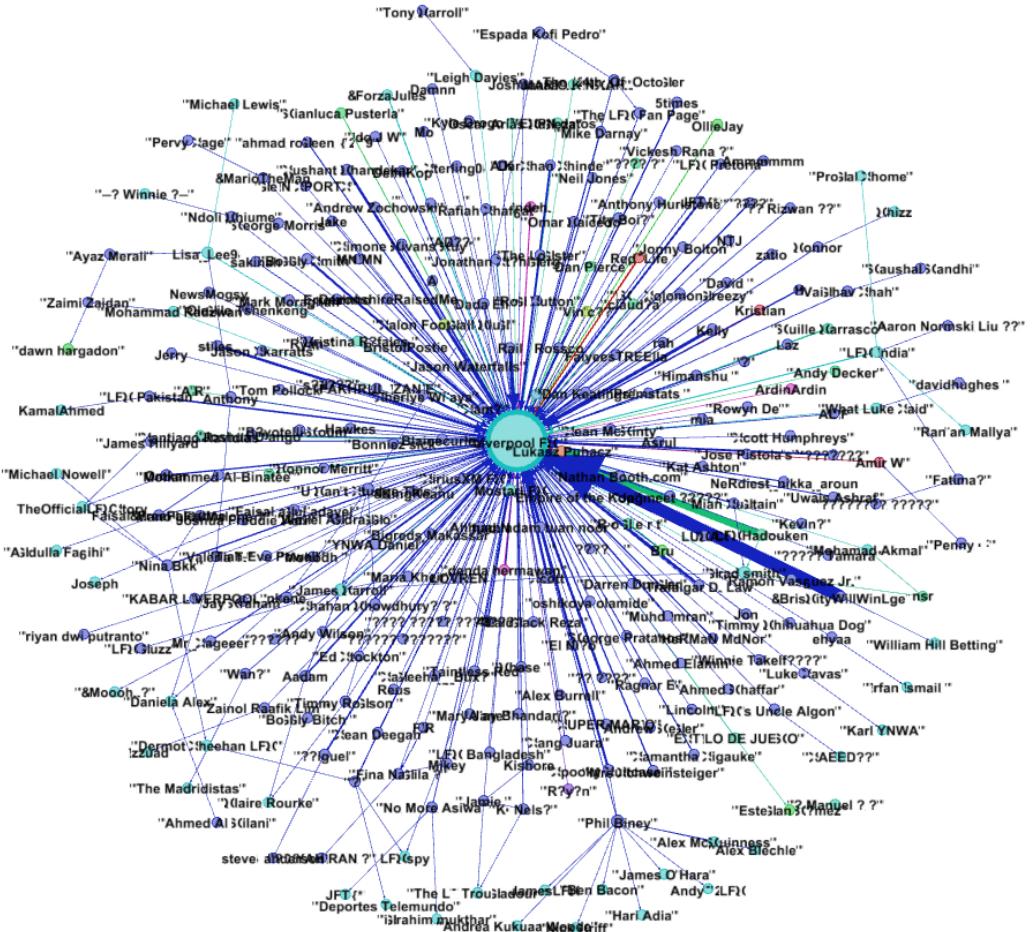
Największe społeczności liczą kolejno **282**, **90**, oraz **65** węzłów.

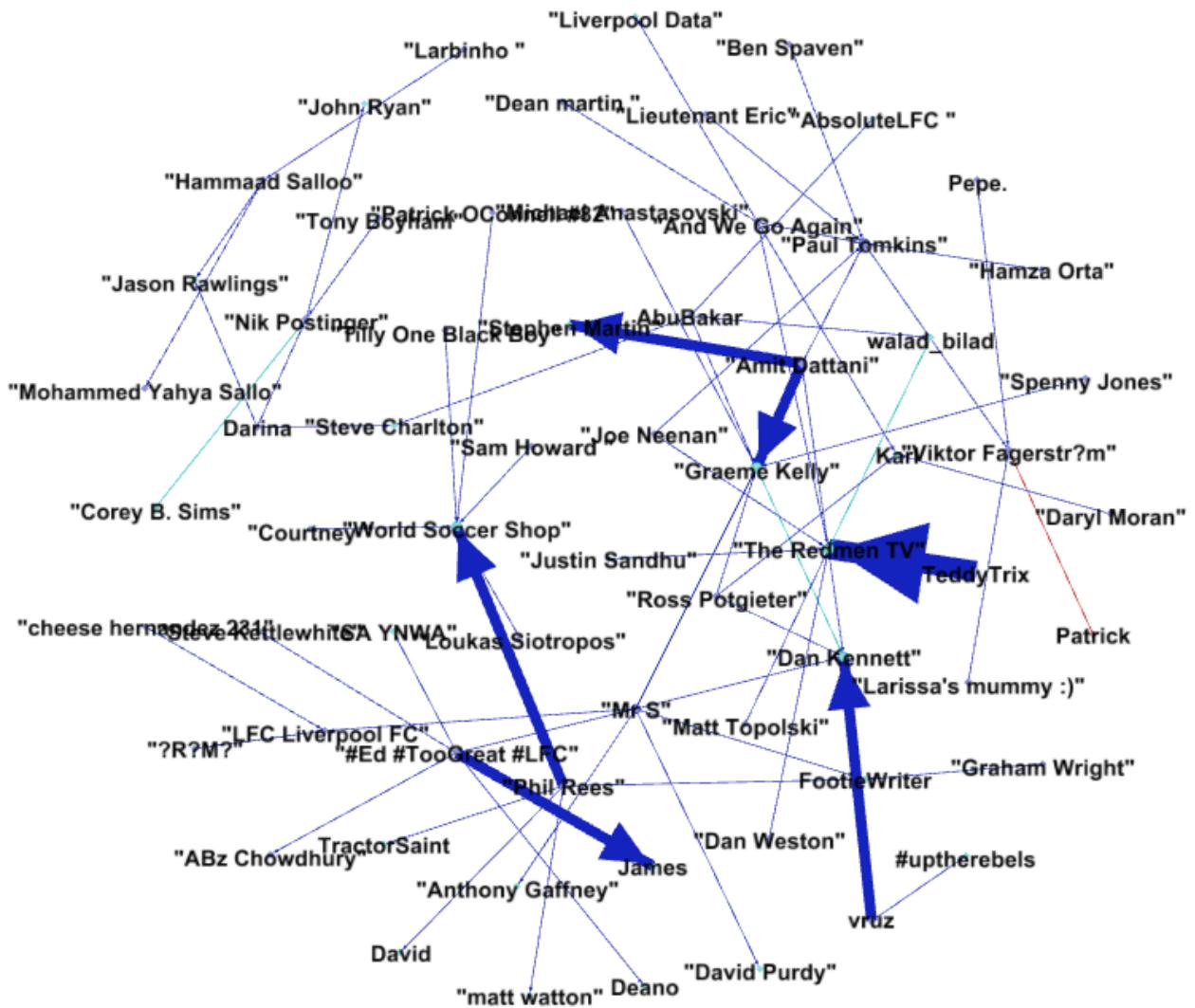
Modularity dla grafu wyniosło **0,689**.

W społecznościach wyraźnie dominują użytkownicy o narodowości brytyjskiej (język angielski).

- Wszyscy użytkownicy: **1412**
- EN: **1229**
- ES: **79**
- inne: **104**

Poniżej umieszczone grafy dla 3 największych społeczności. Wielkość węzła jest zależna od jego stopnia.



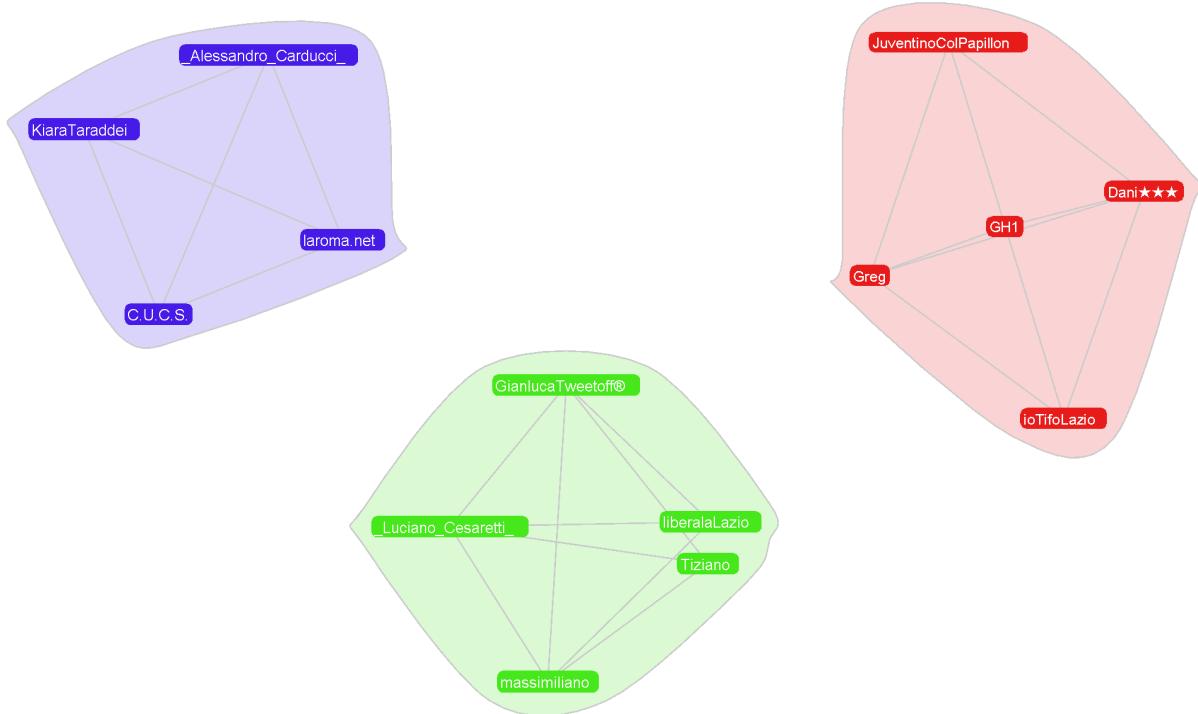


6.8. Wyszukiwanie grup zachodzących w programie „Cfinder”

Przeanalizowano wcześniej stworzone grafy retweet'ów pod kątem społeczności.

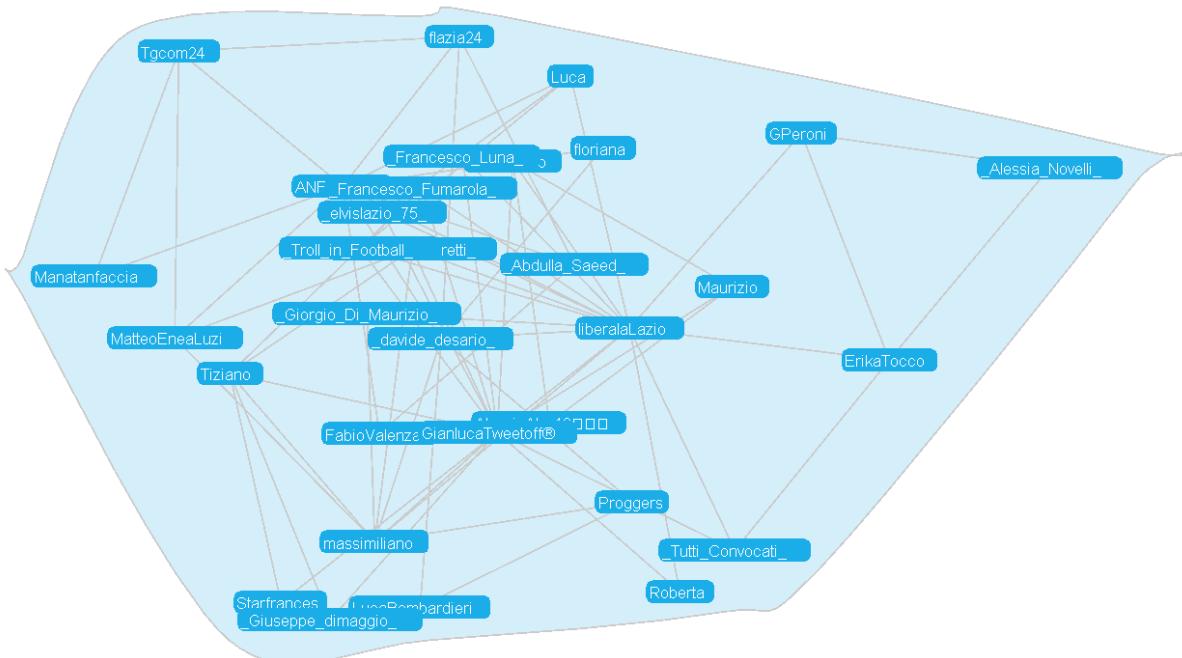
AS Roma - FC Bayern Monachium

Znaleziono trzy grupy o rozmiarze kliki = 4:

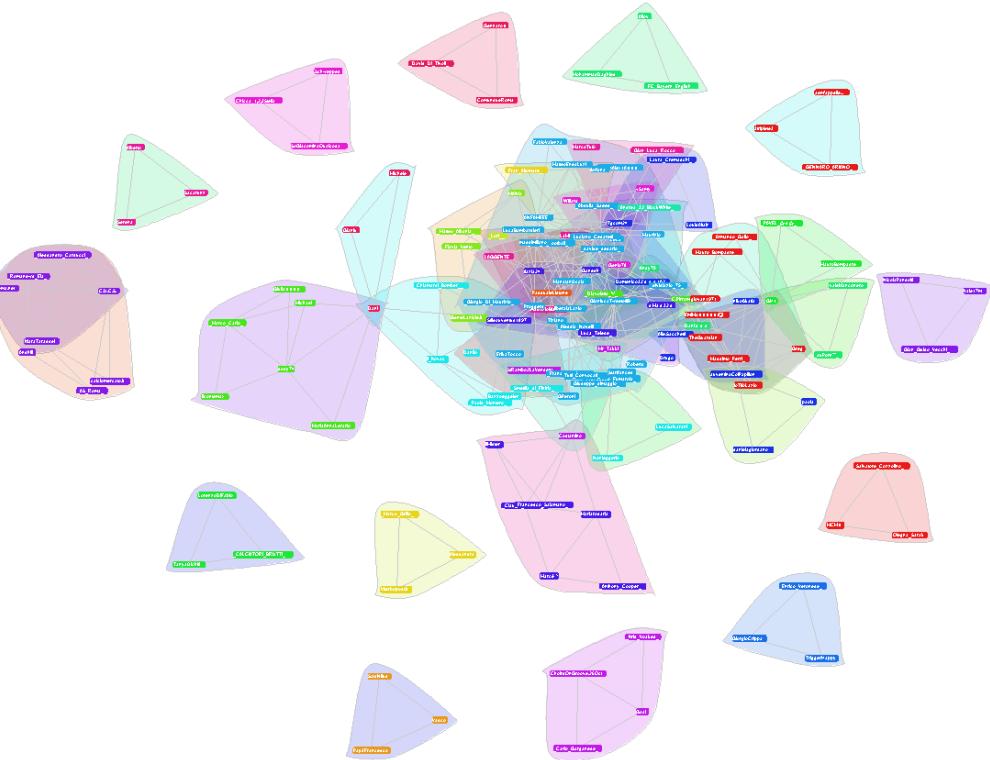


Jeżeli chodzi o liczebność grup z $k = 3$, to znaleziono 39 grup. Największa miała 32 elementy.

Najliczniejsza grupa:



Wszystkie grupy na jednym grafie:



TOP5 Betweenness:

vale bianconero - **37 106,224**

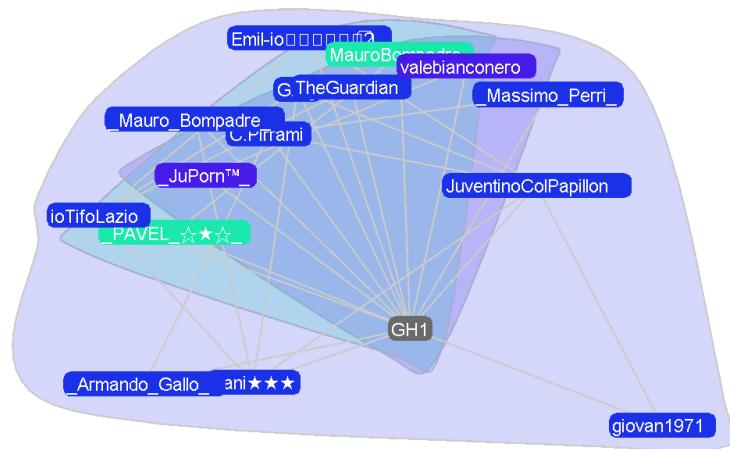
GH1 - **25 139,436**

ilcontemax - **20 655,017**

Alessandro Monaco - **18 684,452**

sosy74 - **18 183,917**

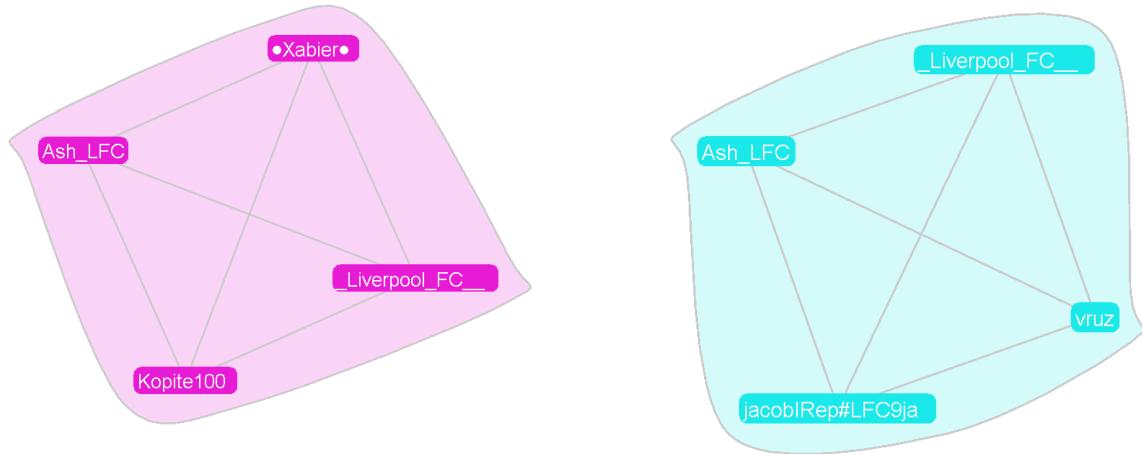
GH1 (4 społeczności):



| Name | In-Degree | Out-Degree | Degree | Closeness | Betweenness | Eccentricity |
|-------------|-----------|------------|--------|-----------|-------------|--------------|
| GH1 | 16 | 31 | 47 | 5,404 | 25139,436 | 14 |
| LaMiaLazio? | 17 | 7 | 24 | 8,392 | 12195,267 | 15 |
| Tgcom24 | 22 | 15 | 37 | 1,808 | 3336,7 | 4 |

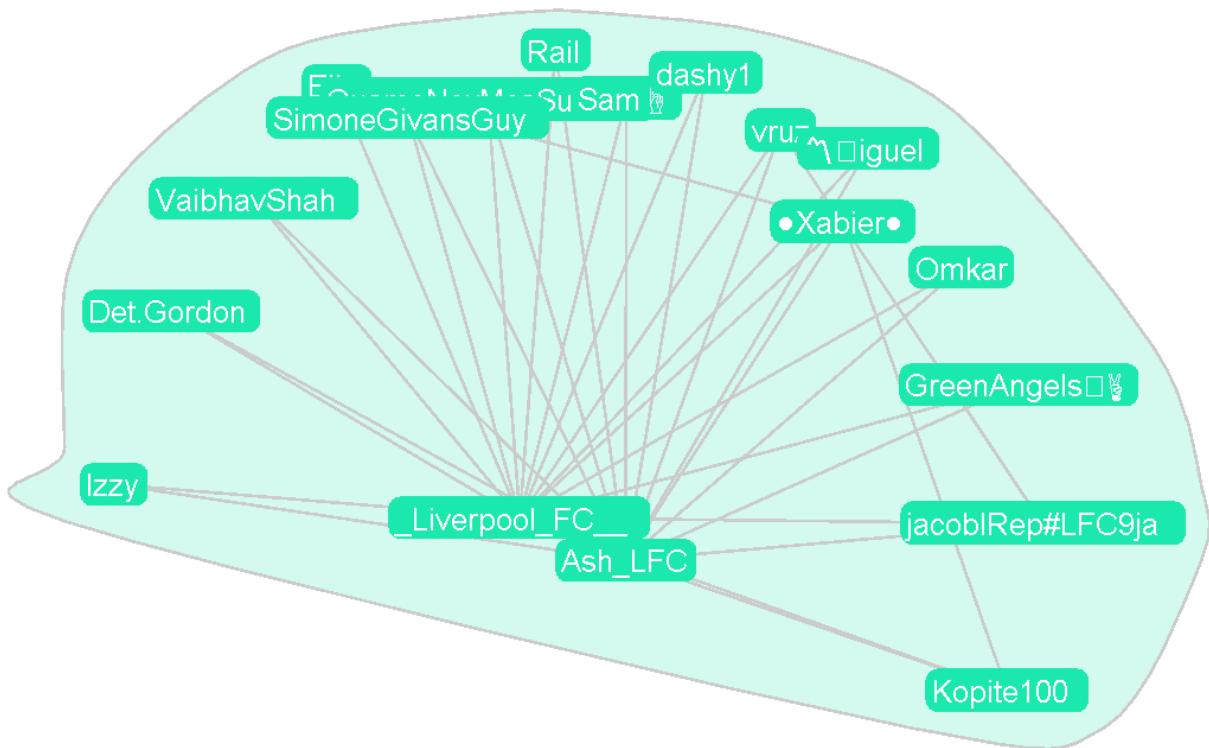
Liverpool FC - Real Madryt

Znaleziono dwie grupy o rozmiarze kliki = 4:

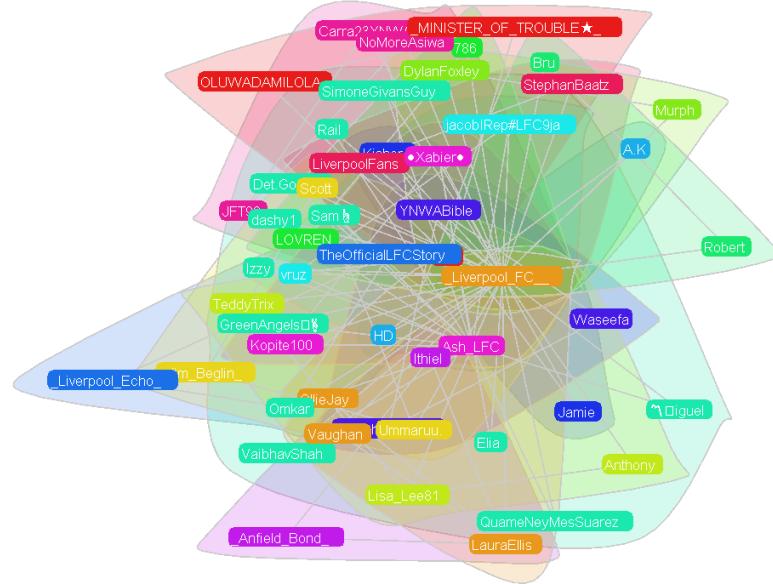


Jeżeli chodzi o liczebność grup z $k = 3$, to było 10 grup 3-elementowych, 4 grupy 4-elementowe i jedna grupa 18-elementowa.

Najliczniejsza grupa:



Wszystkie grupy na jednym grafie:



TOP 5 Betweenness:

Liverpool FC – 933,5

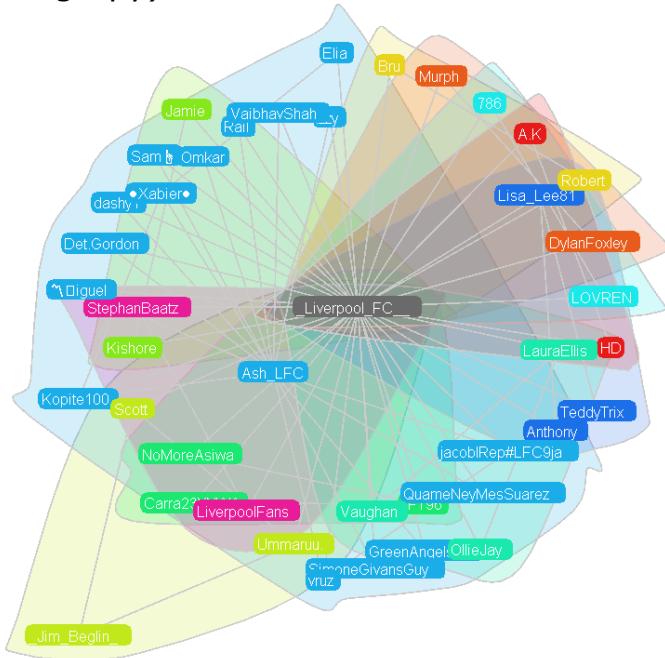
Liverpool Echo - 268,6

Stephan Baatz – 245,5

Kaveshan Chetty - 239,5

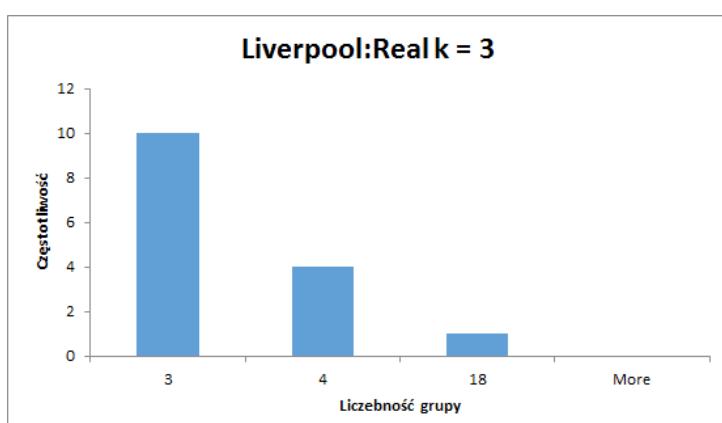
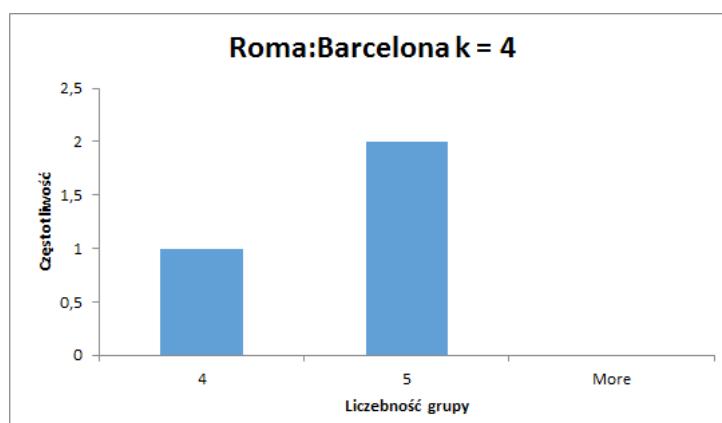
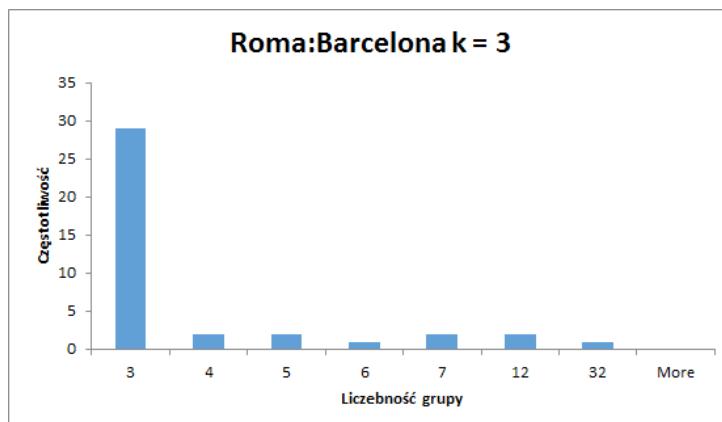
Laura Ellis - 138

Z powyższych Liverpool FC zawiera się w 13 grupach, reszta użytkowników w 1 (oprócz Laury Ellis - 2 grupy).



| Name | In-Degree | Out-Degree | Degree | Closeness | Betweeness | Eccentricity |
|--------------|-----------|------------|--------|-----------|------------|--------------|
| Liverpool FC | 326 | 2 | 328 | 1 | 933,5 | 1 |
| Laura Ellis | 1 | 95 | 96 | 1504 | 138 | 6 |

6.9. Histogramy liczebności grup

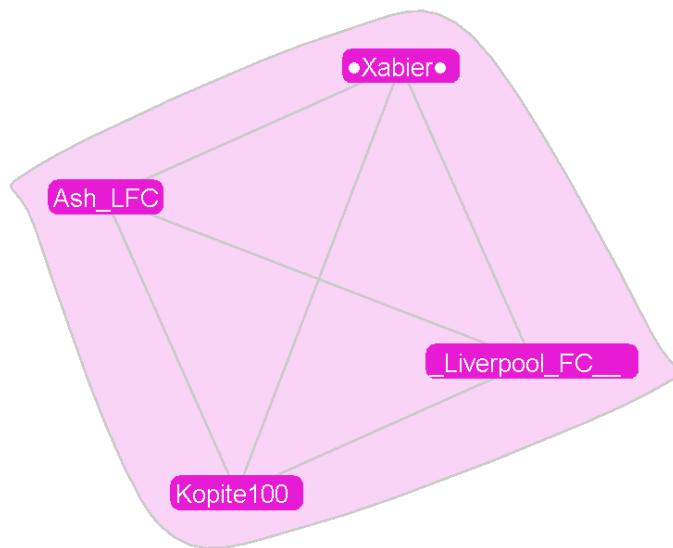


6.10. Analiza wybranych społeczności

Przeprowadzono analizę kilku wybranych społeczności (wyznaczonych przy użyciu programu CFinder), które tworzyły kliki pod względem retweetów - to znaczy użytkownicy pochodzący z takich grup udostępniali ponownie nawzajem swoje posty. W ramach analizy zidentyfikowano użytkowników i zapoznano się z ich kontami w serwisie Twitter.

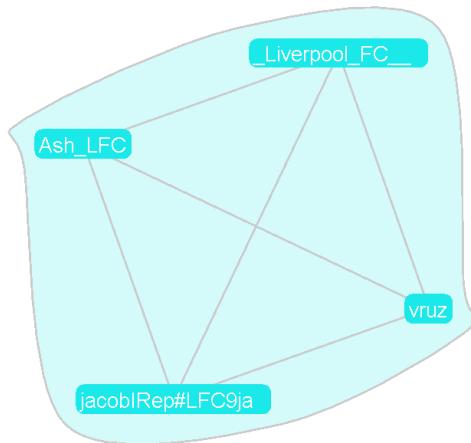
Real Madryt vs. Liverpool FC

Grupa I



| ID | Username | Displayed Name | Krótki opis |
|------------|------------|----------------|---|
| 51427676 | zodman100 | Ash_LFC | Zagorzały fan LFC, postuje newsy i ciekawostki |
| 19583545 | LFC | Liverpool FC | Oficjalne konto twitterowe LFC |
| 86251414 | nudananya | Xabier | Tweetuje absolutnie o wszystkim, okazjnie o LFC |
| 2488768159 | Kopite2014 | Kopite100 | Zagorzały fan LFC, postuje newsy i ciekawostki |

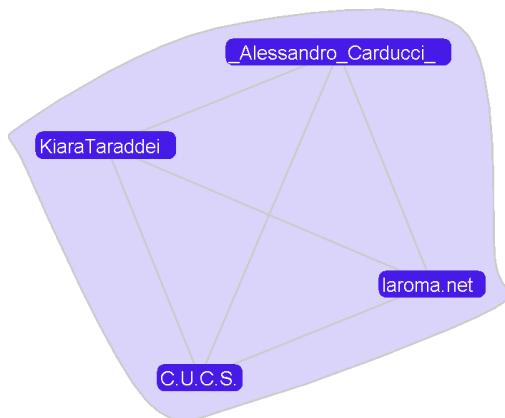
Grupa II



| ID | Username | Displayed Name | Krótki opis |
|-----------|----------------|----------------------|--|
| 51427676 | zodman100 | Ash_LFC | Zagorzały fan LFC, postuje newsy i ciekawostki |
| 19583545 | LFC | Liverpool FC | Oficjalne konto twitterowe LFC |
| 28009632 | vruz | vruz | Posty o wszystkim, podczas LM retweetował LFC |
| 259868944 | speechlessjay9 | jacob I Rep #LFC 9ja | Fan LFC, sporo tweetów na wiele innych tematów |

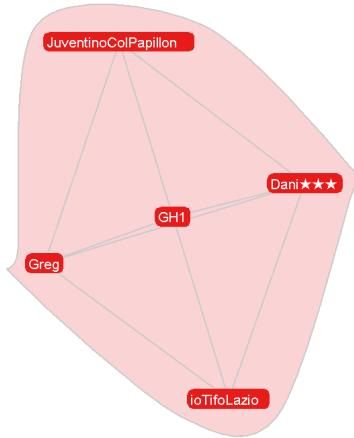
AS Roma vs. Bayern Monachium

Grupa I



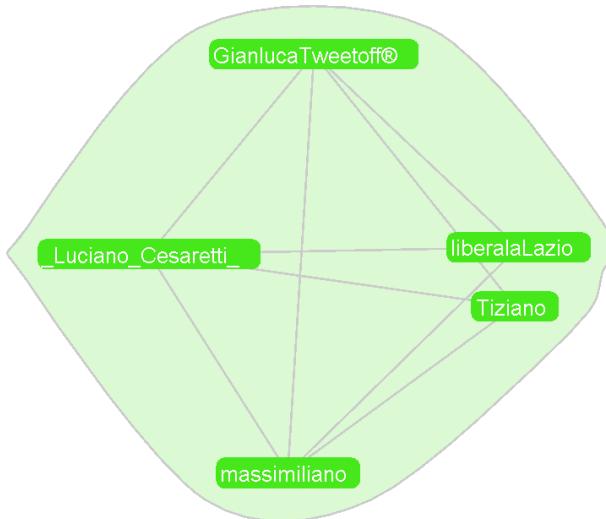
| ID | Username | Displayed Name | Krótki opis |
|------------|-----------------|---------------------|---|
| 358140712 | alecarducci | Alessandro Carducci | dziennikarz vocegiallorossa.it - strona poświęcona AS Roma |
| 82952446 | LAROMA24 | laroma.net | Fanpage Romy, aktualności |
| 2575158406 | ASROMA1927kiara | Kiara Taraddei | Fanka romy, w szczególności Tottiego |
| 707305230 | CUCS_- | C.U.C.S. | Stowarzyszenie najzagorzalszych kibiców Romy (sektor Curva Sud) |

Grupa II



| ID | Username | Displayed Name | Krótki opis |
|------------|--------------|------------------------|--|
| 1563077322 | juvePapillon | Juventino Col Papillon | Nieoficjalny fanpage Juventusu |
| 571133532 | wertjb | Dani | Fan piłki nożnej, tweety nt. Romy, Juventusu, psów i kotów |
| 2358432521 | ju20no | Greg | Fan piłki nożnej, tweety nt. Romy, Juventusu |
| 1077998070 | ioTifoLazio | ioTifoLazio | Fanpage Lazio |
| 615714862 | gabryhouse1 | GH1 | Fan piłki nożnej, tweety głównie nt. Romy |

Grupa III



| ID | Username | Displayed Name | Krótki opis |
|-----------|-----------------|----------------------|--|
| 335240263 | GLaurenzi | Gianluca Tweet off ® | Rozmaite tematy tweetów, fan Lazio. |
| 722641382 | lucianoCesarett | Luciano Cesaretti | Profesjonalny dziennikarz sportowy |
| 313652711 | giulioguazzugli | liberalaLazio | Rozmaite tematy tweetów, fan Lazio. |
| 14314686 | maxcontri | massimiliano | Fan Lazio, sporo tweetów na wiele innych tematów. Dużo zdjęć kawy i rogalików. |
| 152787627 | cholo81 | Tiziano | Fan Lazio, zwierząt i kobiet. |

6.11 Eccentricity, Closeness, Betweenness, PageRank dla TOP10 wierzchołków

Wyliczono statystyki dla wierzchołków w grafach dotyczących retweet'ów.

AS Roma vs. Bayern Monachium

| id | username | degree |
|-----|-----------------------|--------|
| 1 | "FC Bayern English" | 66 |
| 258 | "daniela giordano" | 54 |
| 384 | "C. Pirrami" | 52 |
| 3 | "AS Roma" | 51 |
| 494 | GH1 | 47 |
| 382 | "David Di Tivoli" | 46 |
| 843 | "Stile Juventus 1897" | 44 |
| 123 | Costantino | 43 |
| 786 | Genio78 | 43 |
| 429 | "Vujadln Boskov" | 39 |

| id | username | eccentricity |
|------|-------------------------|--------------|
| 858 | "Pietro Picardi" | 19.0 |
| 1000 | "StRambo2 La Vendetta" | 19.0 |
| 1571 | "KillerdellaComicita?!" | 19.0 |
| 79 | Menad | 18.0 |
| 124 | C.B. | 18.0 |
| 135 | "Ignazio Simone Pizzo" | 18.0 |
| 234 | emdrew | 18.0 |
| 323 | "Namaste, bitches. " | 18.0 |
| 493 | "DomenicoBasciano?★★★2" | 18.0 |
| 558 | "Alex Capolupi" | 18.0 |

| id | username | closeness |
|------|-------------------------|--------------------|
| 1571 | "KillerdellaComicita?!" | 12.228070175438596 |
| 1000 | "StRambo2 La Vendetta" | 11.956140350877194 |
| 858 | "Pietro Picardi" | 11.633053221288515 |
| 323 | "Namaste, bitches. " | 11.263157894736842 |
| 79 | Menad | 11.261764705882353 |
| 558 | "Alex Capolupi" | 11.258064516129032 |
| 234 | emdrew | 11.24561403508772 |
| 493 | "DomenicoBasciano?★★★2" | 11.230994152046783 |
| 701 | "Ema Zarba" | 11.105571847507331 |
| 135 | "Ignazio Simone Pizzo" | 11.07514450867052 |

| id | username | betweeness |
|------|-------------------------|--------------------|
| 597 | "vale bianconero" | 37106.22380952381 |
| 494 | GH1 | 25139.435714285715 |
| 1029 | ilcontemax | 20655.016666666663 |
| 1327 | "Alessandro Monaco" | 18684.452380952378 |
| 223 | sosy74 | 18183.916666666664 |
| 78 | "Gianluca Tweet off ®" | 16024.26904761906 |
| 937 | "Raskol'nikov" | 15718.185714285726 |
| 506 | "Andrea 32 BlackWhite?" | 15193.076190476195 |
| 326 | "davide desario" | 15134.790476190476 |
| 826 | "Davide Boltri" | 14477.192857142863 |

| id | username | Page Rank |
|-----|---------------------|-----------------------|
| 1 | "FC Bayern English" | 0.014806380545823939 |
| 382 | "David Di Tivoli" | 0.009116979520616167 |
| 200 | "Laura Cremaschi" | 0.008992318788317159 |
| 3 | "AS Roma" | 0.008850922183800255 |
| 419 | "Chiamarsi Bomber " | 0.008436613679141738 |
| 123 | Costantino | 0.0067268514388136745 |
| 786 | Genio78 | 0.006469812328228789 |
| 47 | "Luca Telese" | 0.006190348831993433 |
| 429 | "Vujadln Boskov" | 0.006085240398814864 |
| 138 | "FC Bayern München" | 0.005319565620091691 |

Real Madryt vs. Liverpool FC

| id | username | degree |
|-----|------------------------|--------|
| 24 | "Liverpool FC " | 328 |
| 181 | "Laura Ellis" | 96 |
| 455 | "jacob I Rep #LFC 9ja" | 30 |
| 66 | "ESPN FC" | 28 |
| 11 | "Henry Winter" | 26 |
| 167 | Ash_LFC | 24 |
| 214 | "• Xabier •" | 19 |
| 318 | LFCZA | 19 |
| 396 | "Ben Smith" | 19 |
| 707 | "Tony Wareing" | 19 |

| id | username | eccentricity |
|------|------------------------|--------------|
| 970 | "U ain't goin nowhere" | 7.0 |
| 1272 | "Larbinho " | 7.0 |
| 181 | "Laura Ellis" | 6.0 |

| | | |
|------|---------------------|-----|
| 532 | "Hammaad Salloo" | 6.0 |
| 1137 | . | 6.0 |
| 1264 | Scorpion | 6.0 |
| 40 | Anthony | 5.0 |
| 62 | "Liverpool Echo" | 5.0 |
| 93 | TheOfficialLFCStory | 5.0 |
| 100 | "James Hillyard" | 5.0 |

| id | username | closeness |
|------|------------------------|--------------------|
| 1272 | "Larbinho " | 4.75 |
| 532 | "Hammaad Salloo" | 3.9130434782608696 |
| 1264 | Scorpion | 3.75 |
| 1137 | . | 3.642857142857143 |
| 62 | "Liverpool Echo" | 3.3636363636363638 |
| 516 | "Al's... JFT96, YNWA." | 3.3636363636363638 |
| 1401 | "Rich R" | 3.3636363636363638 |
| 93 | TheOfficialLFCStory | 3.3 |
| 429 | Kevin | 3.1904761904761907 |
| 432 | "Total LFC" | 3.1904761904761907 |

| id | username | betweenness |
|-----|-------------------|-------------|
| 24 | "Liverpool FC " | 933.5 |
| 63 | "Liverpool Echo" | 268.5 |
| 286 | "Stephan Baatz" | 245.5 |
| 552 | "Kaveshan Chetty" | 239.5 |
| 181 | "Laura Ellis" | 138.0 |
| 255 | "Liverpool FC" | 125.0 |
| 451 | "Anfield Talk" | 101.5 |
| 111 | Waseefa | 100.5 |
| 194 | "Paul Tomkins" | 57.5 |
| 66 | "ESPN FC" | 54.0 |

| id | username | Page Rank |
|------|-----------------|-----------------------|
| 24 | "Liverpool FC " | 0.09682311103354911 |
| 627 | PB | 0.043960695216193835 |
| 402 | "brad smith" | 0.0427071667732594 |
| 66 | "ESPN FC" | 0.005666294247311436 |
| 11 | "Henry Winter" | 0.005328042649846534 |
| 396 | "Ben Smith" | 0.004965200257052002 |
| 250 | YNWABible | 0.004161075826959178 |
| 167 | Ash_LFC | 0.004043209662295818 |
| 1113 | "Jim Beglin" | 0.003786792397593293 |
| 851 | "Joseph Barton" | 0.0037753042646854287 |

7. Podsumowanie

W ramach projektu pobrano ok. 1,7 mln tweetów zawierających hasztagi związane z rozgrywkami w Lidze Mistrzów. Tweety zostały zebrane w ciągu trzech dni: 21.10 - 23.10 (kolejka fazy grupowej rozgrywana była w dniach 21.10 oraz 22.10).

Z racji ogromnej różnorodności zdecydowano się analizować zestawy tweetów, które powiązane były z konkretnymi meczami w lidze mistrzów. Do tego celu wybrano mecze:

- FC Liverpool vs. Real Madryt
- AS Roma vs. FC Bayern Monachium

Mecze te cechowały się największym zainteresowaniem wśród użytkowników Twittera. Selekcja dokonywana była na podstawie hasztagów zawartych w treści tweetów.

Analiza pozwoliła wyciągnąć następujące wnioski:

- Hasztagi na twitterze są powszechnie uznawanym i stosowanym narzędziem komunikacji. Jedynie 1/6 wszystkich badanych tweetów nie posiadała hasztagów w swojej treści. Podobnie, chętnie stosuje się tzw. "user mention", czyli nawiązanie do użytkownika za pomocą znaku '@'. 70% tweetów zawierało co najmniej jedno "user mention".
- Zgodnie z oczekiwaniemi, natężenie tweetów było największe w godzinach trwania meczów. W chwilach, gdy padały gole, rosło ono nawet kilkukrotnie.
- Funkcjonalność "reply to" jest rzadko używana w serwisie Twitter. Pozwala ona na odpowiedź na konkretnego tweeta innego użytkownika. Użytkownicy znacznie częściej wybierają prostszą opcję którą jest używanie "user mentions" w celu nawiązania do innego użytkownika.
- Sieci zależności pomiędzy włoskimi użytkownikami Twittera są znacznie bardziej rozbudowane niż te u Niemców, Anglików czy Hiszpanów. Powód może być kilka. Po pierwsze, Twitter jest wyjątkowo popularny pośród włoskich fanów piłki. Po drugie, wykształciły się u nich pewne trendy w użytkowaniu portalu Twitter, które powodują chętnie używanie możliwości "retweetowania" (ponownego udostępniania tweeta) oraz "user mentions". Co więcej, w przypadku meczu AS Romy można zauważać wyraźne zainteresowanie fanów współzawodniczących włoskich zespołów (Lazio, Juventus), którzy mieli duży wpływ na otrzymane statystyki.
- Interesujące zależności zostały znalezione w społecznościach użytkowników, którzy "retweetedowali" nawzajem swoje posty. W ten sposób powstała np. społeczność złożona z kilku fanów Liverpool FC oraz oficjalnej strony tegoż klubu. Podobna sytuacja wystąpiła wśród fanów AS Romy - społeczność składała się z dziennikarza sportowego klubu z Rzymu, strony dla najzagorzalszych kibiców, nieoficjalnego fanpage'a oraz oddanej fanki Romy i Francesco Tottiego. Widać zatem, że osoby i organizacje o podobnych preferencjach klubowych "trzymają się" razem. Co ciekawe, odnaleziono również grupę fanów Lazio, którzy publikowali niepochlebne komentarze nt. Romy podczas gdy ta przegrywała.