



evropský
sociální
fond v ČR



EVROPSKÁ UNIE



MINISTERSTVO ŠKOLSTVÍ,
MLÁDEŽE A TĚLOVÝCHOVY



OP Vzdělávání
pro konkurenceschopnost

INVESTICE
DO ROZVOJE
VZDĚLÁVÁNÍ

Twitter data potential

– case studies (Ukraine Crisis and others)

Vít PÁSZTO

vit.paszto@gmail.com

StatGis Team

Reg. no.: CZ.1.07/2.3.00/20.0170



Outline

- Microblogs (MB)
- Social Networks (SN)
- Data Feeding (Perl script)
- Case study 1 (Results)
- Case study 2 (Results)

Microblogs (MB)

- Services allowing sending short messages (microposts)
- Usually from one author
- Microposts are approx. 160 characters long
- Microposts can contain text, links, images etc.
- Simple message about „what I am doing now“

- Twitter, Facebook (status), Tumblr, Instagram, Google+

Microblogs (MB)

Wszystkie wpisy

Aktywne

🔥 Gorące dyskusje



Popularne tagi

- #ukraina 5k
- #heheszki 3k
- #muzyka 3k
- #pytanie 2k
- #nswf 2k
- #ladnapani 2k
- #mecz 1k
- #humor 1k
- #kiciochpyta 1k
- #oswiadczenie 1k
- #humorobrazkowy 1k
- #mikroreklama 1k
- #gif 1k
- #wojna 1k
- #rosja 992
- #gry 849
- #kryptowaluty 811
- #suchar 718
- #roknasilowni 681
- #gimbynieznajo 670
- #sztafeta 660
- #cycki 639
- #rowerowyrownik 616
- #bukmacherka 567
- #plusman 561

Gorące dyskusje

ostatnie 12h

miud 7 godz. temu

+71

#miud

Karolka przepisali do mojej klasy jakoś w połowie pierwszego roku gimbazy. W poprzedniej pełnej patusów nie mógł już dłużej wytrzymać ze względu na gnębienie [pokaż całość](#)

[odpowiedzi \(7\)](#)



PedroLeon 55 sek. temu via Android

0

Idę spać #dobranoc #oswiadczenie



aturt 57 sek. temu

0

Mirki czy ktoś się z was orientuje czy dostane coś takiego tylko z mniejszą ilością gniazd? <http://korporal.pl/index.php?p2621,listwa-zasilajaca-ip-power-socket-rack-19-5x230v-1xr45-zaradzalna>

Szukam od dłuższego czasu takiej listwy tylko z mniejszą ilością [pokaż całość](#)



Mehgods 1 min. temu via Android

0

33 godziny nie spałem, naczytałem się o OOB i chyba nie zasne :<



jankotron 1 min. temu

0

#samochodowykonkurs

4, przedostatnia zagadka.

Ten motocykl zbudowano w celu ładnego prezentowania się przy pewnym aucie. Czy po kolorze, sposobie malowania, wykończeniu detali domyślicie się, jakie auto symbolizuje?

Po drugie, na podstawie jakiego innego auta zbudowano to auto:)

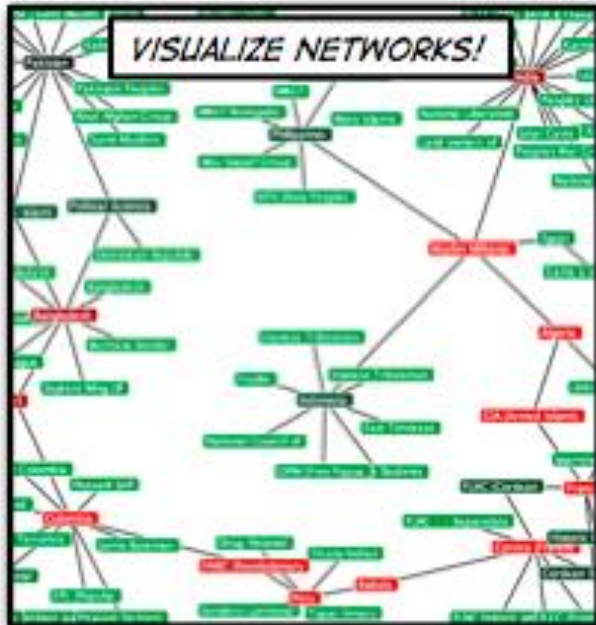


Social Networks (SN)

- Finite set of actors (social entities) and relations among them
- Individuals are not studied separately and their behaviour is influenced by their relations with others
- Information about the relations is thus the most important part of analyses
- In fact, relations define properties of a social network
- Facebook, LinkedIn, QQ空间, ВКонтакте, (Google+)

SocialAction

MAKING SENSE OF SOCIAL NETWORKS...

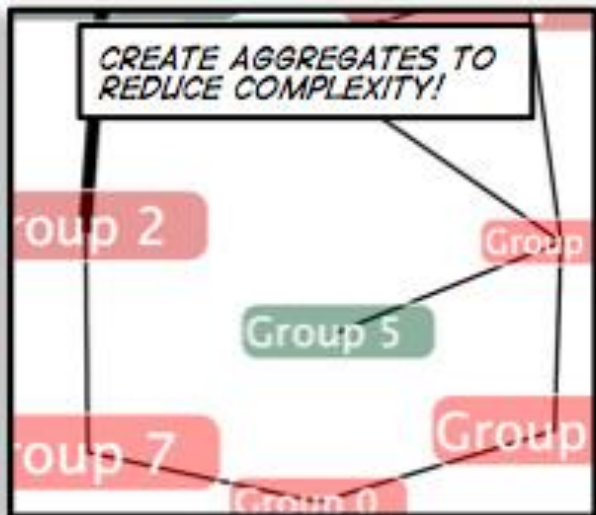
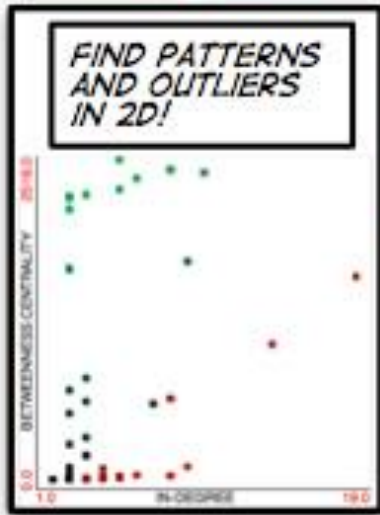
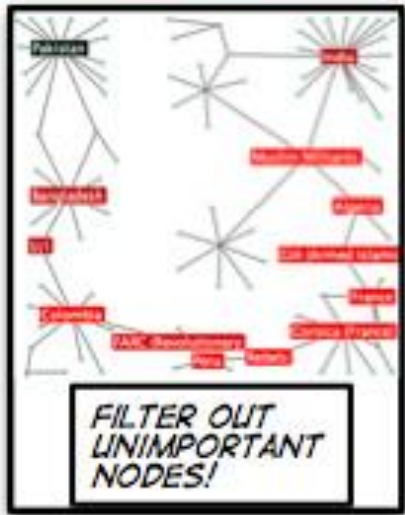


Betweenness Centrality

The number of shortest paths between pairs of nodes that pass through a given node.

Rank	Node	Type
2,516.00	Muslim Militants	Terrorist Group
2,436.50	Corsica (France)	Country
2,413.00	Colombia	Country
2,368.00	Peru	Country
2,280.50	France	Country
2,239.00	Algeria	Country
2,226.00	Rebels	Terrorist Group
2,214.00	CIA (Armed Islamic Group)	Terrorist Group
2,124.00	FARC (Revolutionary Armed For...	Terrorist Group
1,718.00	Bangladesh	Country
1,658.00	Uji	Terrorist Group
1,598.00	India	Country
1,063.00	Pakistan	Country
798.00	Corsican Separatists	Terrorist Group
704.00	RNC (Corsican National Libera...	Terrorist Group
704.00	Historic Wing RNC	Terrorist Group
637.00	Indonesia	Country
614.00	Political Activists	Terrorist Group
518.00	Indonesia	Country

RANK ACCORDING TO IMPORTANCE METRICS!



Twitter

- Microblogging service
- Users share messages with length less than 160 char.
- Via both webpages and smart phones (SMS, apps)
- 284 M active users (30.9.2014)
- 58 M tweets per day
- 9,000 tweets per second
- In 2013 entered the stock market

Twitter – searching tool



- Script in interpret language Perl
- We can set searching parameters
 - Any text string/specific phrase – e.g. „StatGIS 2014“
 - Using logical connectors (AND, OR etc.)
 - Key words with hashtag (#) – e.g. #Crimea
 - Search for user (@) - @vit_paszto
 - Newest tweets, most popular tweets, tweets time scale, location, author info etc.
- Output in .txt form

Twitter – searching tool

- We can download tweet content as well as additional information about the user
- Tweets are geocoded -> we can set a place on Earth
- Limitations:
 - Data with 7 days history
 - Last 100 tweets
- Additional tool for automatic downloading

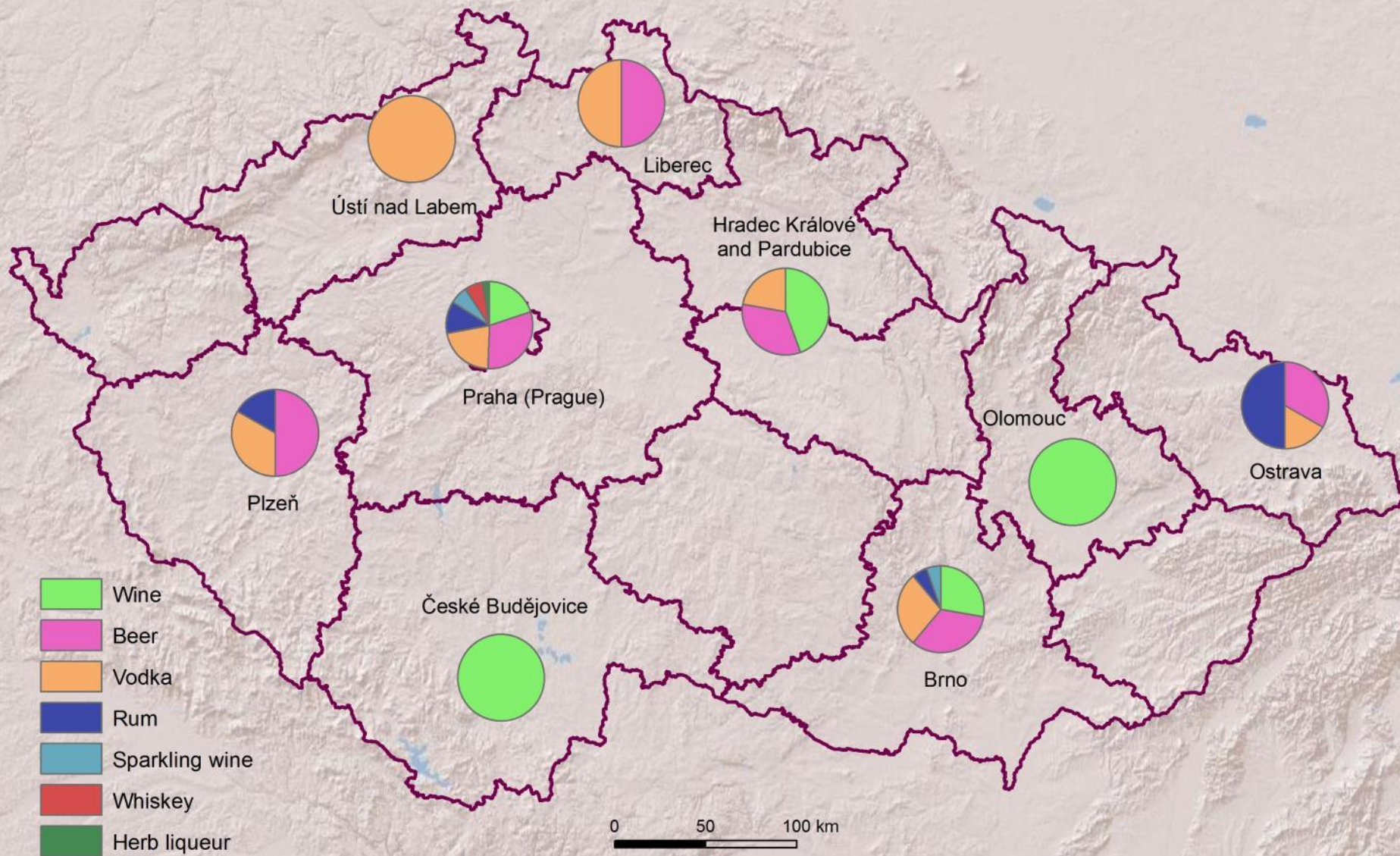
Twitter – searching tool

Live demo

Micro_Case_study_1

- Carried out at the turn of the year (31.12.2013 – 3.1.2014)
- Every regional major city (i.e. Praha, Brno, Ostrava, Plzeň, Liberec, Olomouc, Zlín, Jihlava, Karlovy Vary, Ústí nad Labem, České Budějovice, Pardubice + Hradec Králové)
- We tracked occurrence of alcoholic beverages in tweets:
 - Wine
 - Beer
 - Vodka
 - Rum
 - Sparkling wine
 - Whiskey
 - Fernet

STRUCTURE OF TWEETS ACCORDING TO ALCOHOLIC BEVERAGES (31.12.2013 - 3.1.2014)

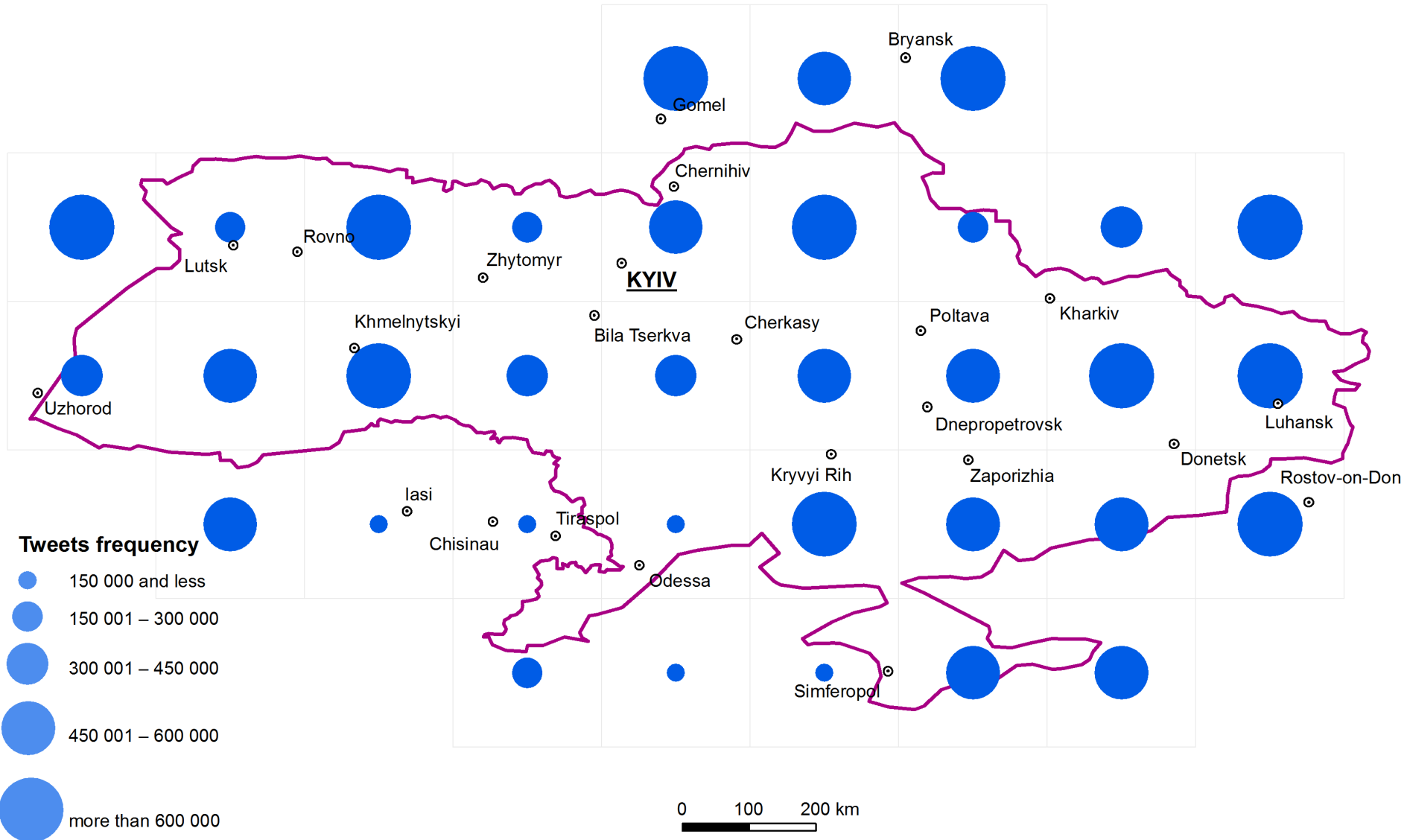


Case_Study_2

- Ukraine crisis
- Data collected since 26th August 2014
- In sum – over **16 M** Tweets
- Ukraine divided into 34 cells

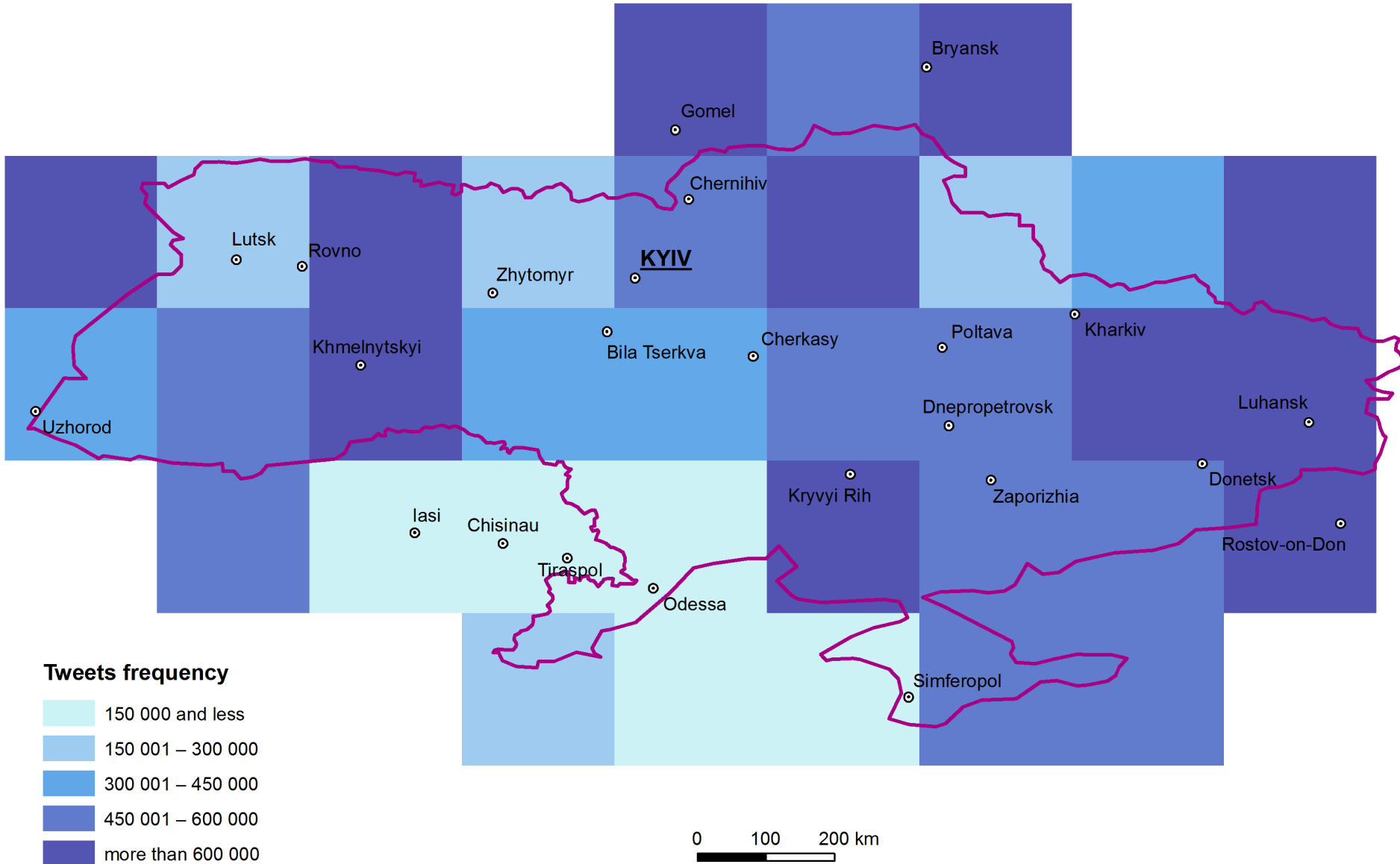
SPATIAL DISTRIBUTION OF SUM OF TWEETS

(Ukraine, collected between 26.8. - 30.10. 2014)

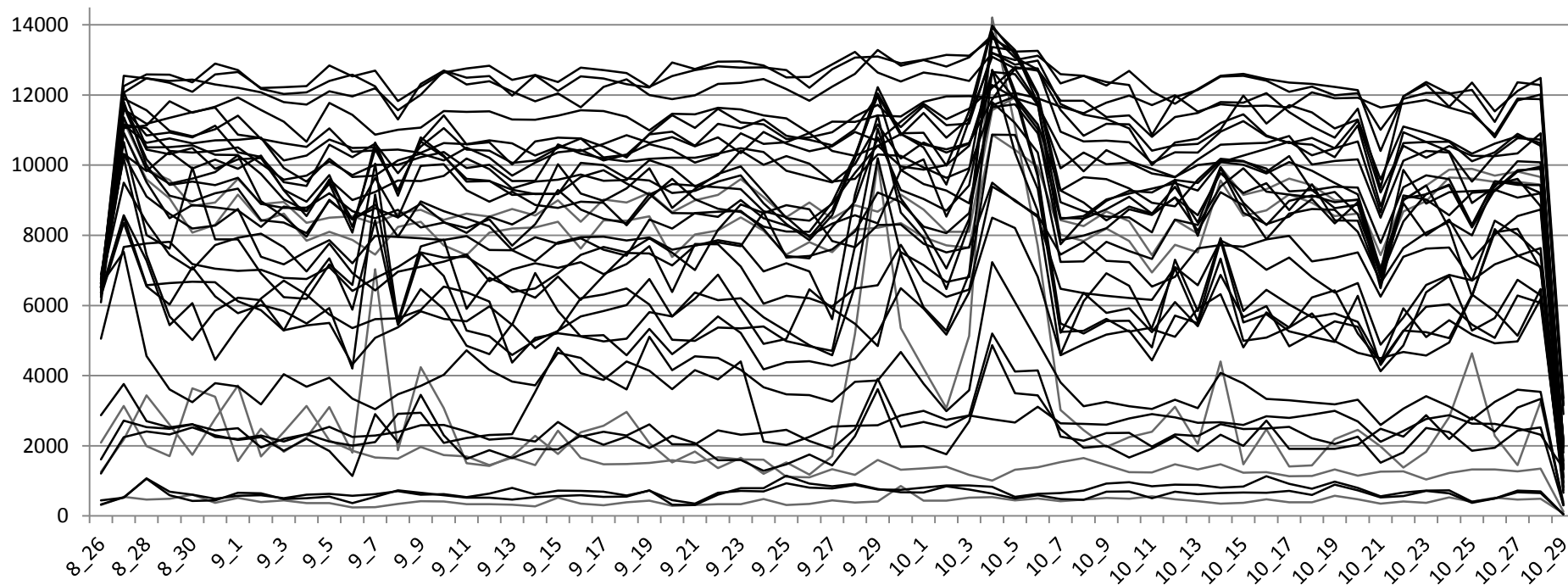


SPATIAL DISTRIBUTION OF SUM OF TWEETS

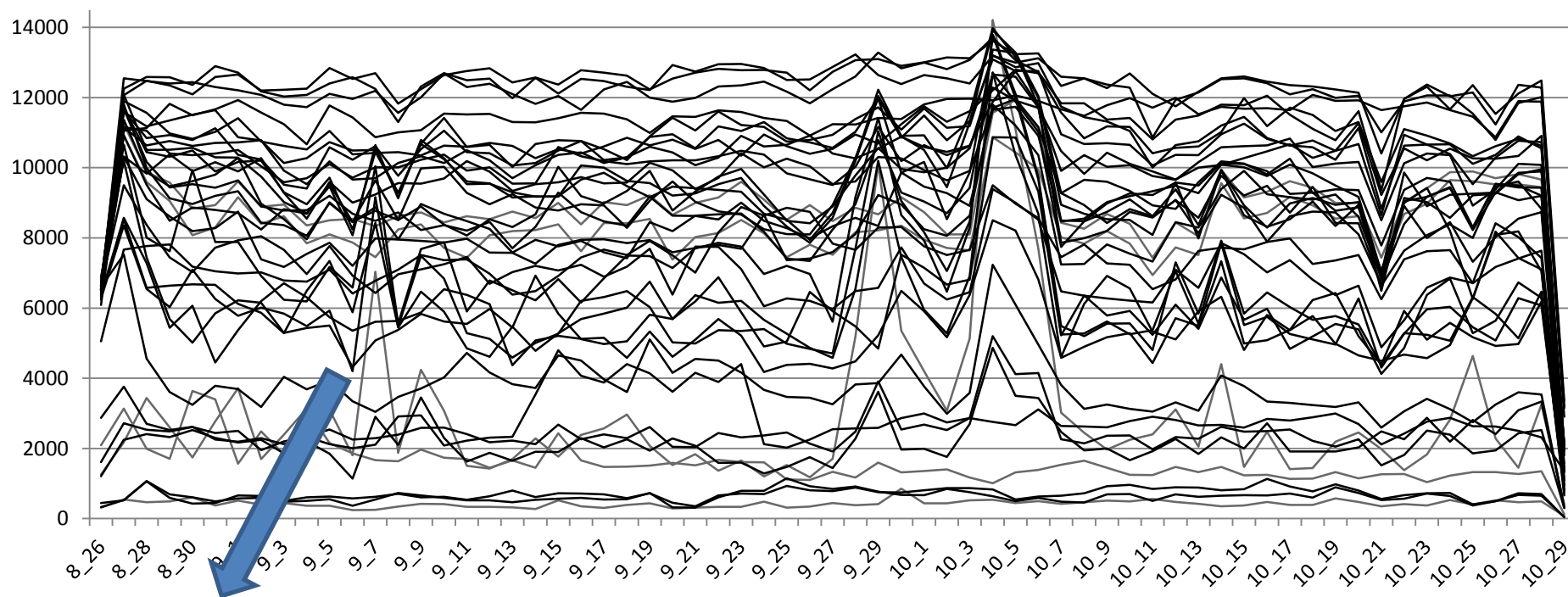
(Ukraine, collected between 26.8. - 30.10. 2014)



Case_Study_2

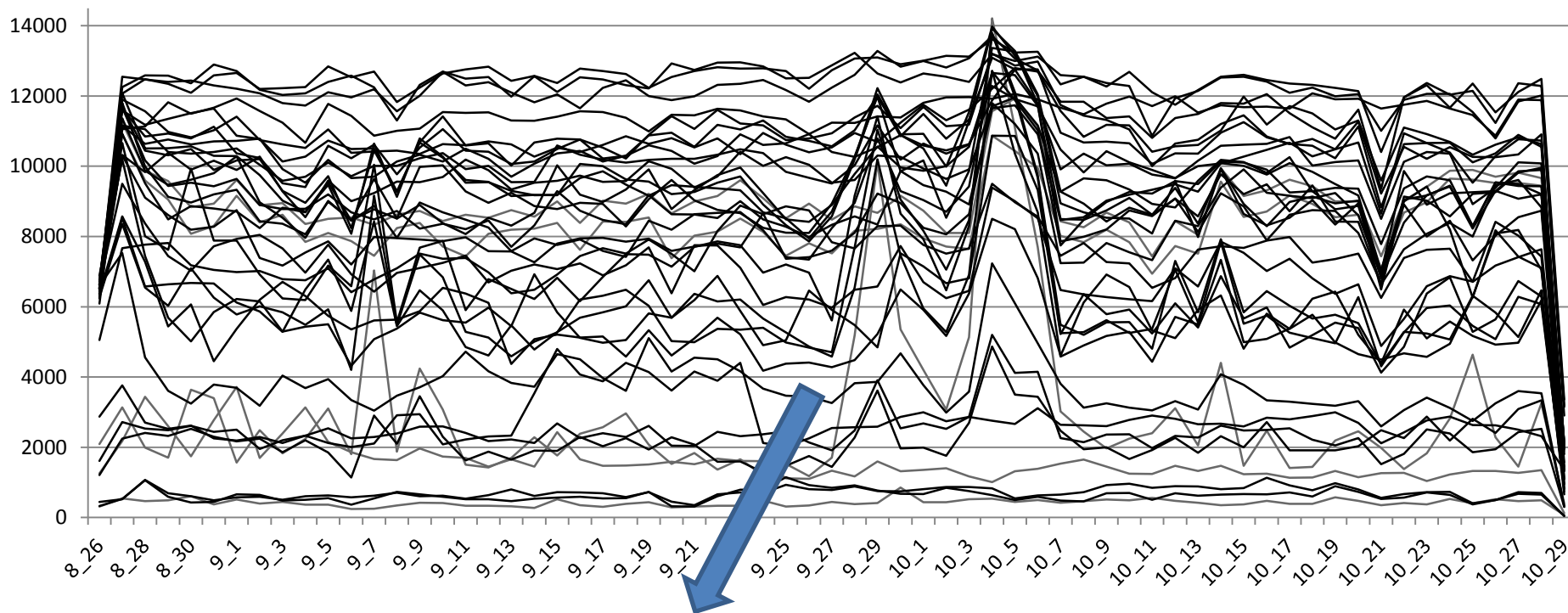


Case_Study_2



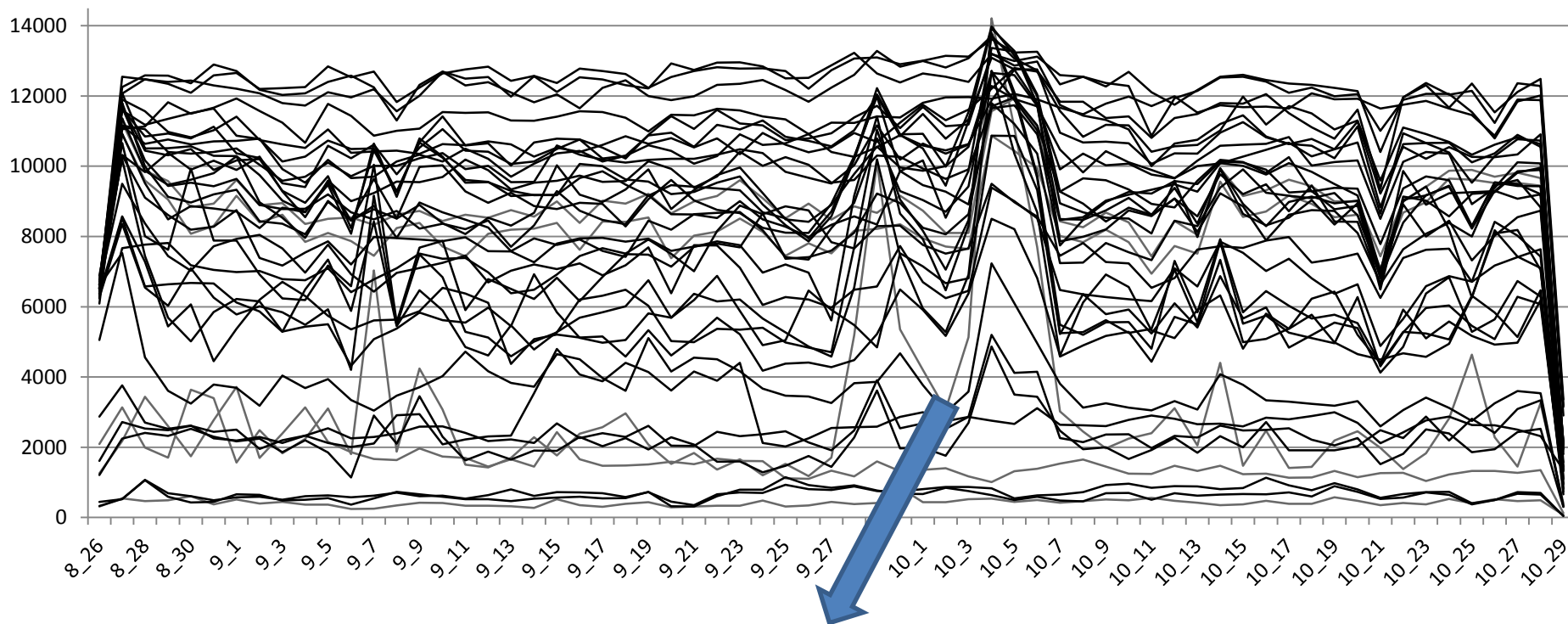
5th Sept.: Cease-fire Announcement; **7th Sept.:** shelling near the Donetsk airport

Case_Study_2



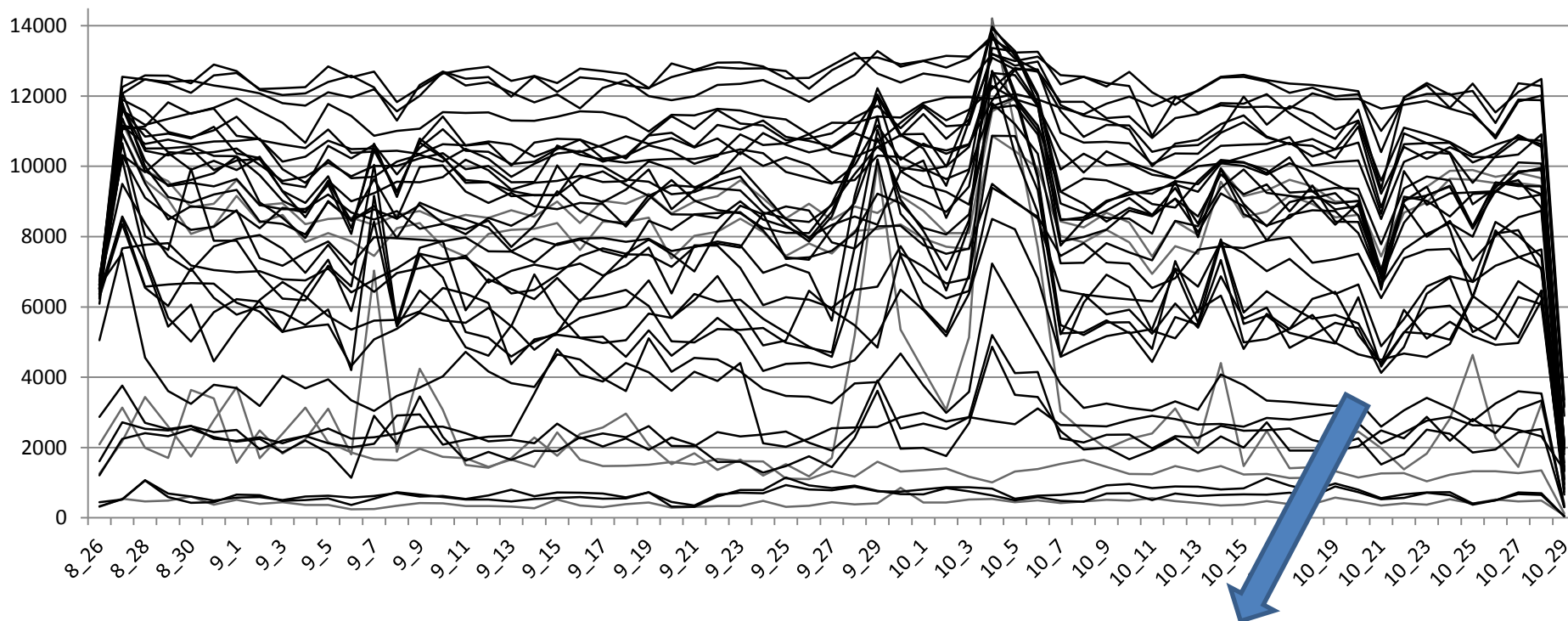
29th Sept.: Clashes near Donetsk – 9 soldiers killed, 27 injured

Case_Study_2



Since 3rd Oct.: Fighting for control of Donetsk airport intensifies – 12 separatists killed

Case_Study_2



21st Oct.: UEFA Champions League – Bate Borisov vs. Shakhtar Donetsk... **0:7** 😊

What's next

- Detailed time-visualization & animation
- Words statistics and analyses
- Mood analysis
- Overall context & interpretation

Twitter data potential

– case studies (Ukraine Crisis and others)



MINISTERSTVO ŠKOLSTVÍ,
MLÁDEŽE A TĚLOVÝCHOVY



OP Vzdělávání
pro konkurenceschopnost

INVESTICE
DO ROZVOJE
VZDĚLÁVÁNÍ

Vít PÁSZTO

vit.paszto@gmail.com

&

František DAŘENA, Lukáš MAREK (volunteer)

StartGis Team