

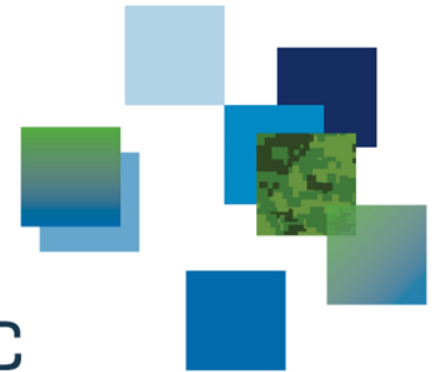


# Understanding Social Media Networks Through Trending Analysis

Marie-Eve Jobidon, PhD  
Bruce Forrester, PhD

*21<sup>st</sup> ICCRTS – 6 Sept 2016*

DRDC | RDDC



# Outline

- Social media 101
- Trends
  - General background
  - Trends in social media
- Trend analysis on SM platforms
- Case study
  - Freddie Gray Jr. and #blacklivesmatter
- Discussion and conclusion

# Social Media 101



# What to monitor?



# What makes SM different?

## Traditional OSINT Sources

Academic research, books, encyclopaedias, business and government documents, grey literature, images, journals, periodicals, broadcast media, maps, newspapers, radio

Edited

Written by professional authors

Use of proper grammar, spelling and punctuation

Minimal use of sarcasm, street language, profanity

## SM Sources

Blogs and micro-blogs, Internet forums, user-generated FAQs, Chat, podcasts, online games, tags, ratings, comments, social networking sites, online video, wikis, search engines, social bookmarking

Not edited

Written by anyone and everyone

Anything goes

Anything goes

# Social media exploitation for intelligence

- Monitor many channels of SM for SA and cueing
- Examine trends and provide early warnings
- Find and discover activity and potential areas of danger (force protection)
- Identify influential users (e.g., radicalization)
- Identify and collect information about particular persons-groups-organizations of interest (targeting)



# Analytical challenges are many

- Big messy data
- Natural language
- Social in nature
- Veracity
- Deception
- Determination of ground-truth
- Demographics
- Real-time
- Cultural impacts/biases

# Are we ready to deal with modern warfare?





# Trends – Background

- “A general development or change in a situation or in the way that people are behaving”
- Various natures: Mathematical, psychological, behavioural...
- How they manifest vs. how – or whether – they are perceived
  - We may not perceive trends that exist or perceive trends that don't
- Time-based
  - Time scale is critical in defining and detecting trends

# Trend analysis for intelligence

- Trend analysis:
  - Trying to determine whether an underlying pattern exists in data
  - To forecast future events or better understand past or current events
  - Signal in noise
- For INT analysts, SM trend analysis is challenging:
  - May not be aware that a change is coming (monitoring vs. discovering)
  - Difficulty in determining what a SM trend means in the “real” world

# Trends in social media

- Several SM platforms track what is *trending*
  - Trending topic: A general idea that is gaining popularity at a given time
- Main characteristics from which trends are determined:
  - Topic
  - User engagement (individual and/or population)
  - Proprietary algorithm
  - Timeliness
  - User determined location



\* Exactly what counts as a trend in SM is neither clear nor consistent across platforms

Characteristic used in trend determination	Twitter	Google (+)	Facebook	Pinterest	YouTube
Topic	X	X	X	X	
Individual user engagement		X	X	X	
Population user engagement		X		X	X
Proprietary algorithm	X	X	X		
Timeliness	X	X	X		X
User determined location	X	X	X		X
Location of post					X
Overall trending topics in platform			X		
Things user liked			X		
List of followers	X				
Search volume		X			X
Language					
Age range					X

## Case study – Framework

*Can we detect an emerging trend and track its evolution?*

- Explores how trend analysis can be used in an INT context
- Post-hoc analysis of a situation that has trended in the past
- NexaMaster: A social data analysis platform

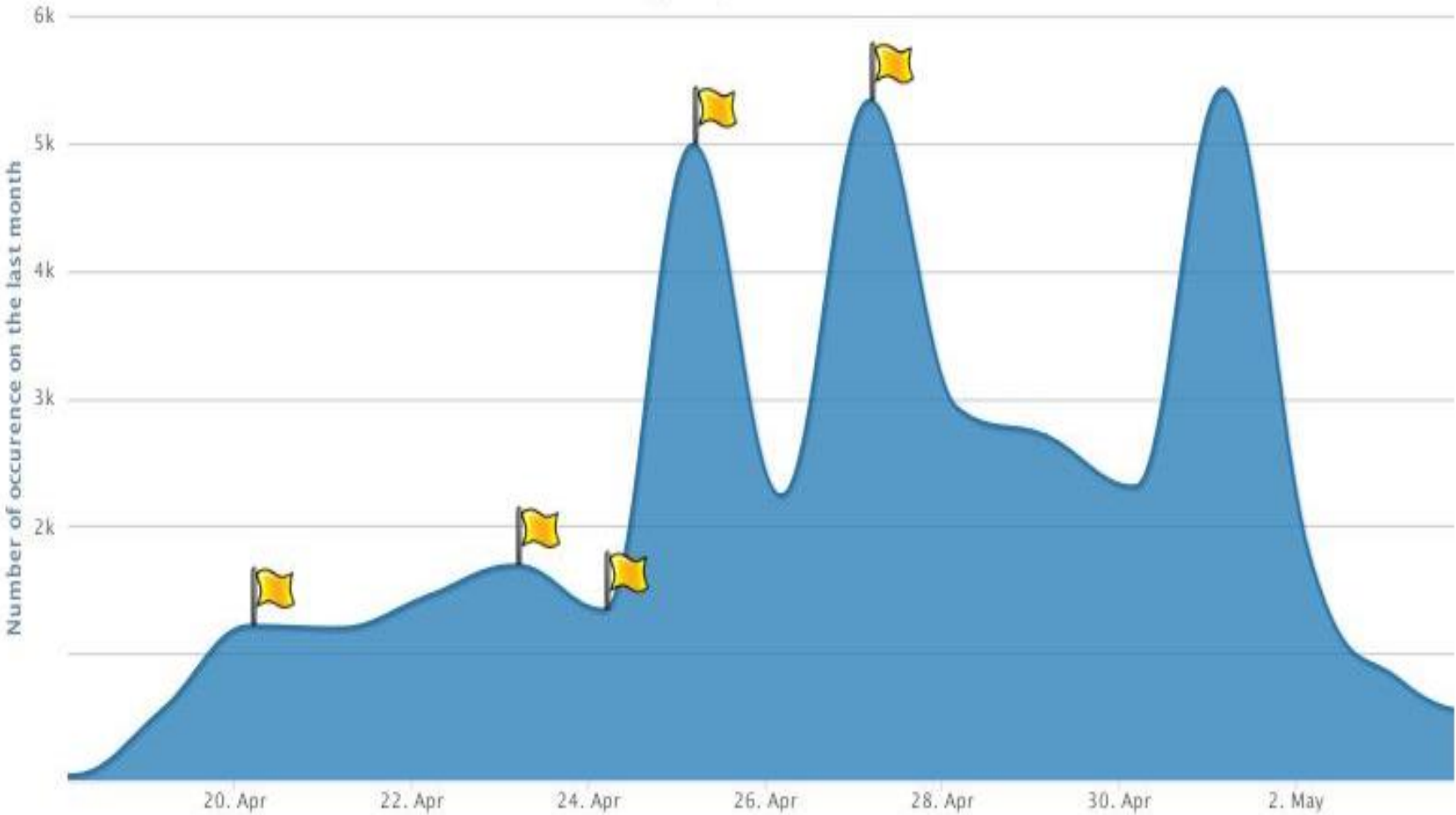
# Case study – Freddie Gray Jr. and #blacklivesmatter

- Events surrounding the arrest and death of Freddie Gray Jr. in Baltimore, Maryland, in April 2015
  - Fuelled the Black Lives Matter movement on SM
- Collected a sample of relevant tweets (~40K tweets)
- Analyzed the sample with focus on protests
- SM data were compared to offline events

# Case study – Timeline of events

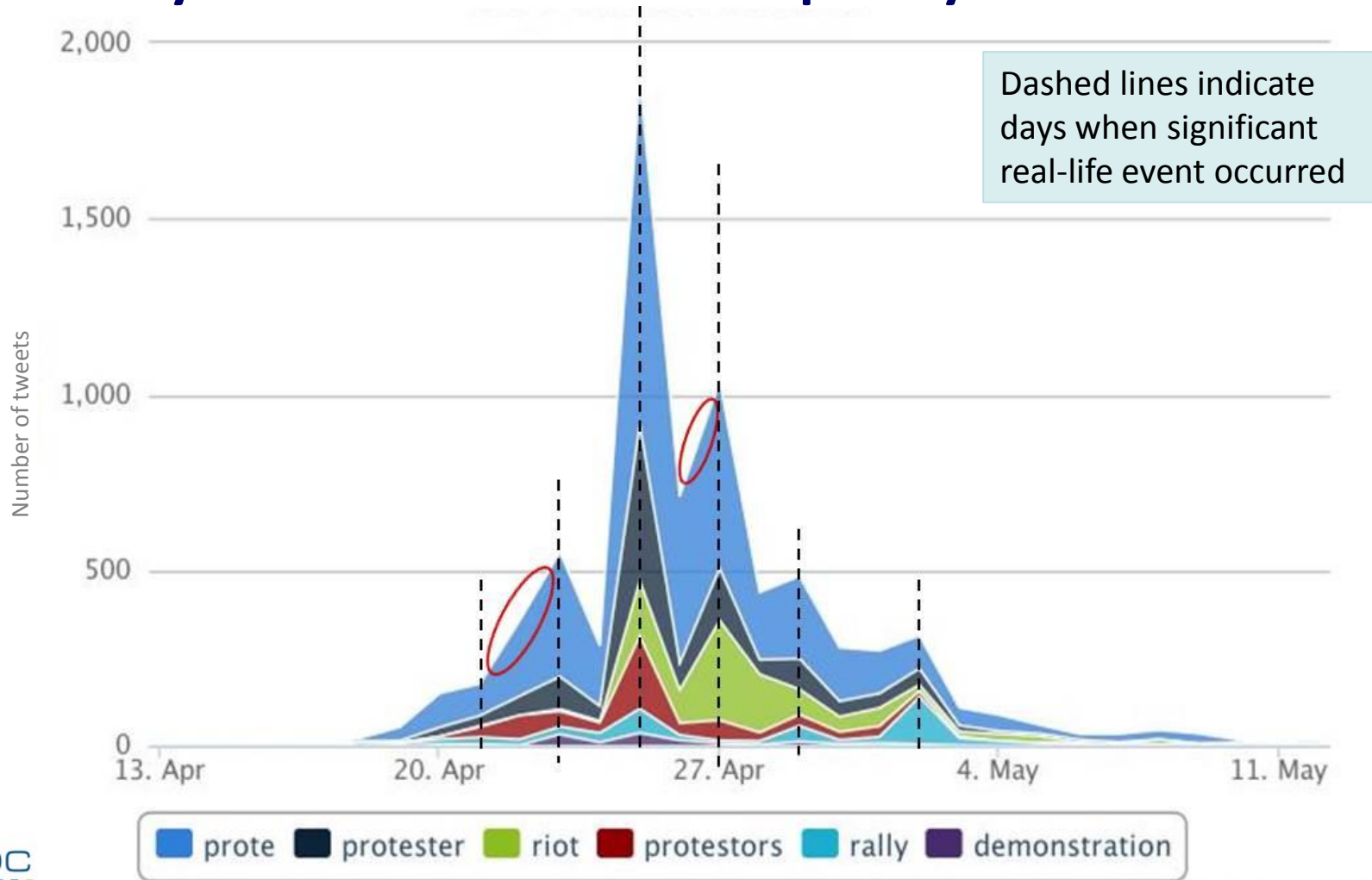
Date	Event
12 Apr	Arrest and transportation unsecured in police van
19 Apr	Death of Freddie Gray Jr.
23 Apr	Protests in downtown Baltimore
25 Apr	Major protests in downtown Baltimore that turned violent
27 Apr	Funeral and burial – Civil disorder intensifies in Baltimore (e.g., riots) State of emergency is declared and the National Guard is called in
29 Apr	Protests in other U.S. cities in response to Freddie Gray Jr.'s death
30 Apr	Medical Examiner reports that Gray sustained injuries as a result of slamming inside the van
1 May	Charges filed against the 6 officers involved

# Case study – Category peaks





# Case study – Timeline of word frequency and trends



# Discussion

- It seems possible to identify SM trends that correlate with offline events
  - Post-hoc analysis - Need for further research
- Face validity of trending factors used by platforms seems good
  - But clarity of how trends are actually determined needs more definition
- Representativeness of trends is in question
  - Analysts need to understand how each individual platform calculates trends
- Many platforms use individual preferences: Biases trends

# Conclusion

- Trend analysis in SM is challenging for INT analysts
  - Huge set of complex social data – Makes detecting/tracking trends difficult
  - Number of factors can influence – or manipulate – SM trends
  - Challenging to determine the “so what”
- Difficult for INT analysts to have good level of confidence in trends from SM platforms
  - With regards to population meaning, timescales used for baseline, magnitude and significance of change, etc.
- Analysts require independent trend analysis tools rather than relying on trending topics identified by platforms

# DRDC | RDDC

**SCIENCE, TECHNOLOGY AND KNOWLEDGE**  
FOR CANADA'S DEFENCE AND SECURITY

**SCIENCE, TECHNOLOGIE ET SAVOIR**  
POUR LA DÉFENSE ET LA SÉCURITÉ DU CANADA

