On the Development of Automated Forensic Analysis Methods for Mobile Devices

Panagiotis Andriotis, Theo Tryfonas, George Oikonomou, Shancang Li,
Zacharias Tzermias, Konstantinos Xynos, Huw Read and Vassilis Prevelakis

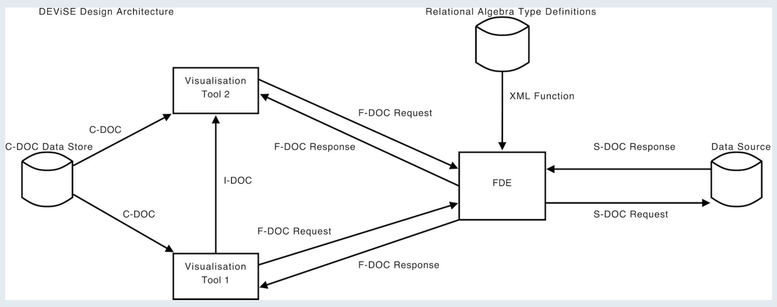
Motivation and Objectives

We live in a connected world where devices communicate with each other and humans use them as valuable tools. The use of mobile devices leaves traces that can be treasured assets for a forensic analyst. As mobile devices become more powerful, they can hold numerous personal data.

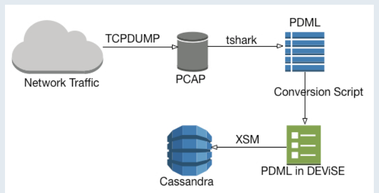
Our aim is to investigate methods and exercise techniques that will merge all these valuable information in a way that will be efficient for a forensic analyst, via graphical representation of the underlying data structures.

The project brings together different research areas and aggregates algorithms in order to automate procedures in forensic examinations while diminishes the work payload by focusing on specific areas of interest.

The DEViSE Framework works as a data aggregator. It is a middleware between data sources and visualization tools.



Network traffic aggregator plugin for DEViSE.



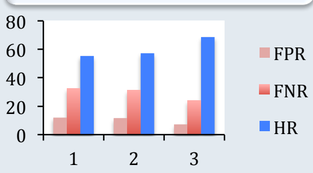
Data Mining from Mobile Devices And Sentiment Analysis Timeline View

We developed a data migration mechanism able to work with sms data and perform sentiment analysis. We propose the use of a popular lexicon and we added features like the word valence and the presence of emoticons to strengthen the success rates of the scheme.

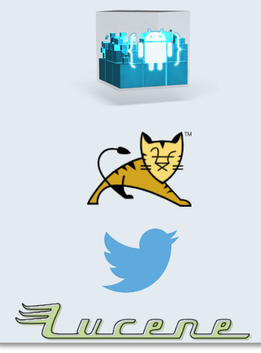
	Datasets	Total	False Positive Rate (FPR)	False Negative Rate (FNR)	Correctly Identified
AFINN	Positive	1867	12.1%	32.6%	55.3%
	Negative	919	22.7%	32.4%	44.8%
Emoticons	Positive	1867	11.5%	31.1%	57.4%
	Negative	919	22%	31.7%	46.3%
Emoticons & Valence	Positive	1867	7.3%	23.9%	68.8%
	Negative	919	29.2%	25%	45.8%

The Sentiment Score calculation algorithm and the effectiveness of our approach.

$$s(t_k) = \Sigma_i l_{ni}(t_k) - \Sigma_j l_{pj}(t_k)$$

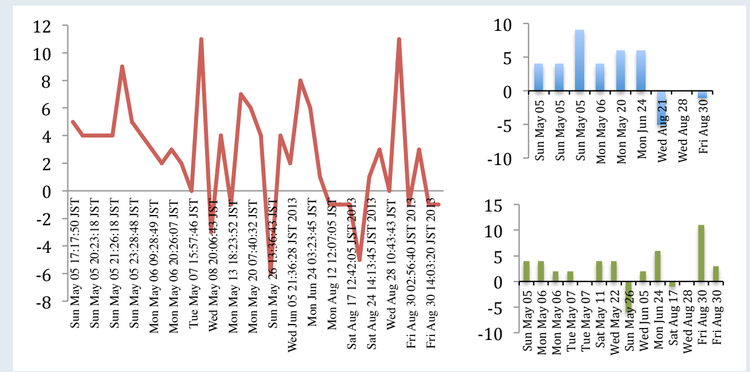


We use well tested open source tools providing costless solutions.



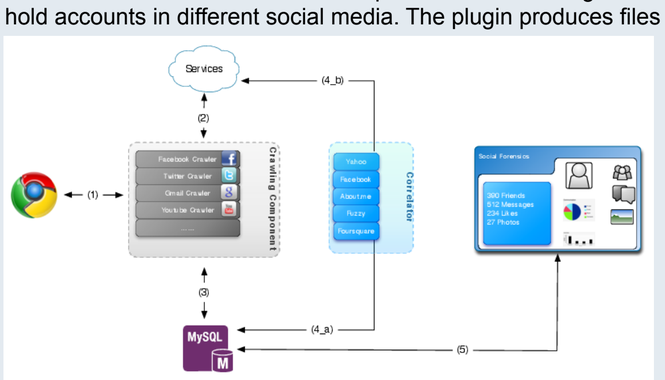
id	date	address	t...	body	keywords	sentiment
1	Sun May 05 1...	+3[redacted]	8 1	The End is Nigh, well not nigh, it is actually here. The Worlds End is out...	end, nigh, wel...	5
2	Sun May 05 1...	+3[redacted]	8 2	Andy Murray winning Wimbledon is my 1966 World Cup moment. What...	andy, murray,...	4
3	Sun May 05 2...	+3[redacted]	8 1	I should clarify Bilderberg is a secret meeting of the worlds powerful eli...	i, should, clari...	4
4	Sun May 05 2...	+3[redacted]	0 1	Seriously, an ad for The Mormon Church in the programme of The Book...	seriously, ad,...	4
5	Sun May 05 2...	+4[redacted]	4 1	Watching the Mamet HBO flick PHIL SPECTOR. Its oddly the least Mamet...	watching, ma...	4
6	Sun May 05 2...	+3[redacted]	8 1	Thank you for all of the love and support! And thank you @CraigZadan,...	thank, you, all...	9
7	Sun May 05 2...	+3[redacted]	8 1	Nothing says Thursday like tiny, tiny adorable animals. http://t.co/qb5...	nothing, says,...	5
8	Mon May 06 0...	+3[redacted]	8 2	Watching my audience dance is one of my favorite parts of the day. Esp...	watching, my,...	4
9	Mon May 06 0...	+4[redacted]	4 1	I love this performance by @EmeliSande. http://t.co/rmT96hL4N0	i, love, perfor...	3
10	Mon May 06 1...	+3[redacted]	0 2	Which state has the smallest drinks? Mini-soda. #ClassicJokeTuesday	which, state,...	2
11	Mon May 06 2...	+3[redacted]	8 1	It is a boy! So happy for my cousin Kate and the future King of England!	boy, so, happ...	3
12	Tue May 07 1...	+4[redacted]	4 2	My favorite thing about Heads Up! is that it films you. This is me brillia...	my, favorite, t...	2
13	Tue May 07 1...	+3[redacted]	8 2	Thinking of my friend. http://t.co/ibKVOrDVm7	thinking, my,...	0
14	Wed May 08 1...	+3[redacted]	0 1	Congratulations to my dear friend @JimmyKimmel and his beautiful ne...	congratulatio...	11

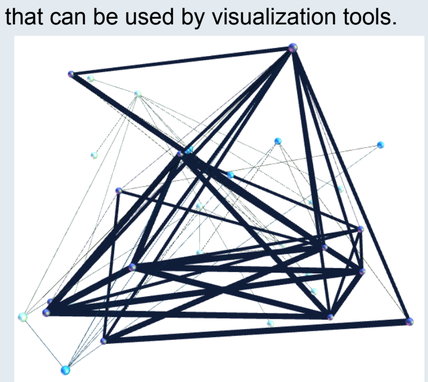
Found 4 hits.
1. It is a boy! So happy for my cousin Kate and the future King of England!
---- Database id: 11 on Mon May 06 20:26:07 JST 2013 received from: +3[redacted]18
2. Happy Halloween from me & the dummy http://t.co/jQxPvHqi
---- Database id: 32 on Mon Aug 26 03:10:23 JST 2013 received from: +44[redacted]94
3. Happy 4th of July! In 1776, Betsy Ross signed the Gettysburg Address and celebrated at the Boston Tea Party. It was a magical day.
---- Database id: 19 on Mon May 20 07:40:32 JST 2013 received from: +3[redacted]8
4. Congratulations to my dear friend @JimmyKimmel and his beautiful new wife! I am so happy for him. I had no idea he was straight.
---- Database id: 14 on Wed May 08 14:43:32 JST 2013 received from: +3[redacted]40



Multilevel Visualization of Social Interactions Merging Data from Social Media and Mobile Devices

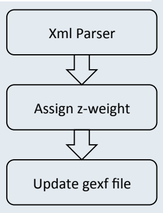
The use of APIs for developers provided by social media companies can produce weighted graphs like the following. Our crawlers need the credentials of the person under investigation. We correlate the collected data and link entities that hold accounts in different social media. The plugin produces files that can be used by visualization tools.

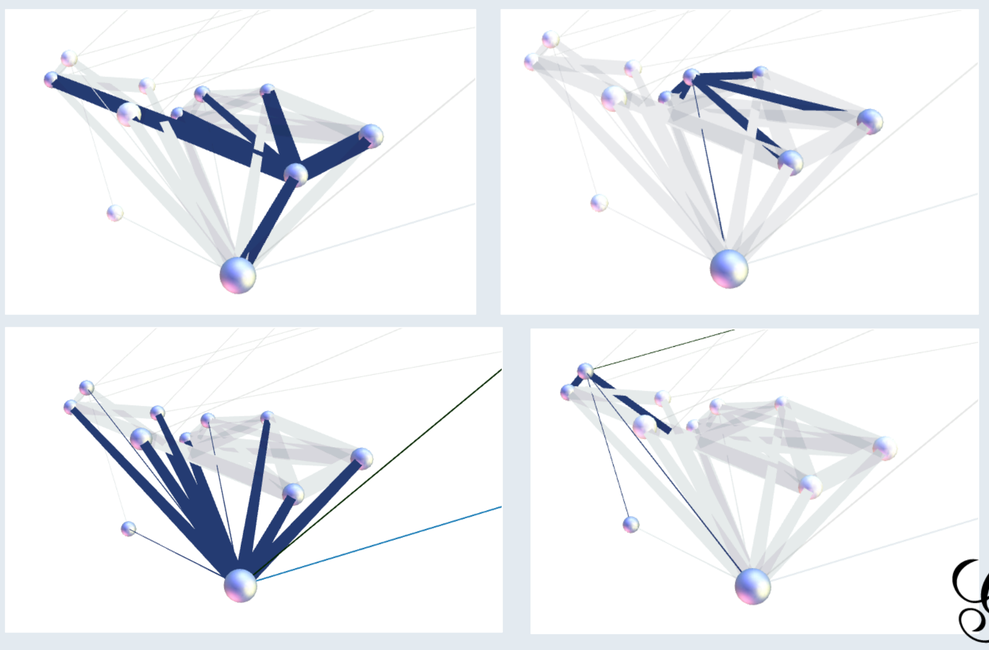




These graphs might be overpopulated and the analysis task might seem daunting. Taking into account that we can use data provided by a smartphone, such as the contact list or other interactions (calls, instant messenger logs, chats and sms texts) we provide various levels of proximity among the different entities.

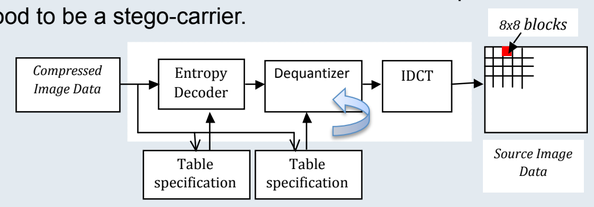
Here, entities that exchange messages and make calls with the person under investigation are placed closest to him/her. The 3D visualization scheme shows relations between the closest and distant 'friends'.



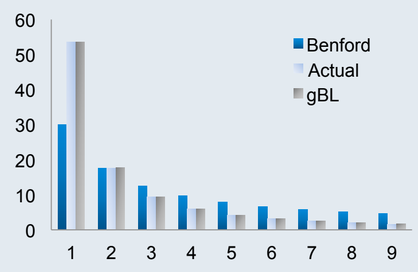


Steganalysis with Benford's Law and Classification of Malicious JPEG Images

This module aims to classify JPEG images stored in mobile devices as suspicious stego-carriers or as pure JPEG images. We examine the internal structure of the images during the dequantization phase and use the empirical Benford's Law to build a model that describes the distributions of the first digits of the quantized coefficients (f.d.). The model is used to evaluate variations of the expected deviations of f.d. on the tested image and estimate the likelihood to be a stego-carrier.



$$p(x) = N \cdot \log_{10}\left(1 + \frac{1}{s + x^q}\right)$$



QF	N	q	s	SSE
q100	1.608	1.605	0.0702	5,129e-06
q90	1.25	1.585	-0.405	7.235e-07
q80	1.344	1.685	-0.376	3.007e-06
q75	1.396	1.731	-0.3549	3.986e-06
q70	1.434	1.766	-0.339	4.455e-06
q60	1.514	1.843	-0.3114	5.464e-06
q50	1.584	1.909	-0.2875	5.119e-06

The classification module evaluates the variations from the expected digit deviations.

Our results on the training set and on a different photoset produced by a smartphone.

Acknowledgements

This work has been supported by the European Union's Prevention of and Fight against Crime Programme "Illegal Use of Internet" – ISEC 2010 Action Grants, grant ref. HOME/2010/ISEC/AG/INT-002 and the Systems Centre of the University of Bristol.

Digital Forensics Tools