



Idea: Visual Analytics for Web Security

Victor Le Pochat, Tom Van Goethem, Wouter Joosen
victor.lepochat@cs.kuleuven.be
imec-DistriNet, KU Leuven, Celestijnenlaan 200A, 3001 Leuven, Belgium
ESSoS 2018, June 27, 2018



DistriNet



Web security issues are increasing in
ubiquity • scale • complexity

Web security issues are increasing in
ubiquity • scale • complexity

leading researchers to conduct
large-scale experiments and observations

Large-scale experiments consist of two phases

Data collection	Analysis
<ul style="list-style-type: none">• large scale• distributed• fast to set up• reusable	<ul style="list-style-type: none">• inefficient• ad hoc• labor-intensive• duplicated

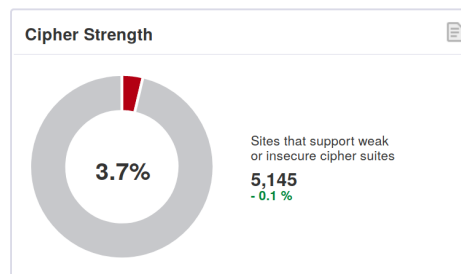
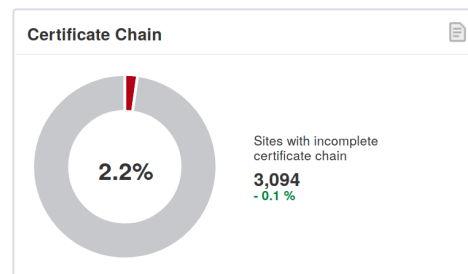
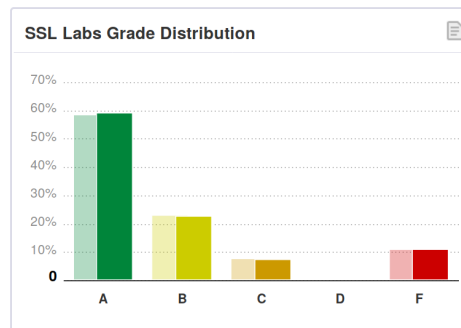
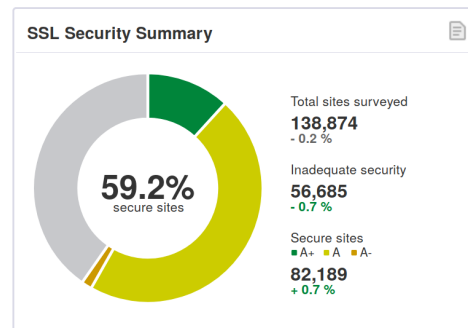
Visual analytics is a promising approach [Kei01, TC05]

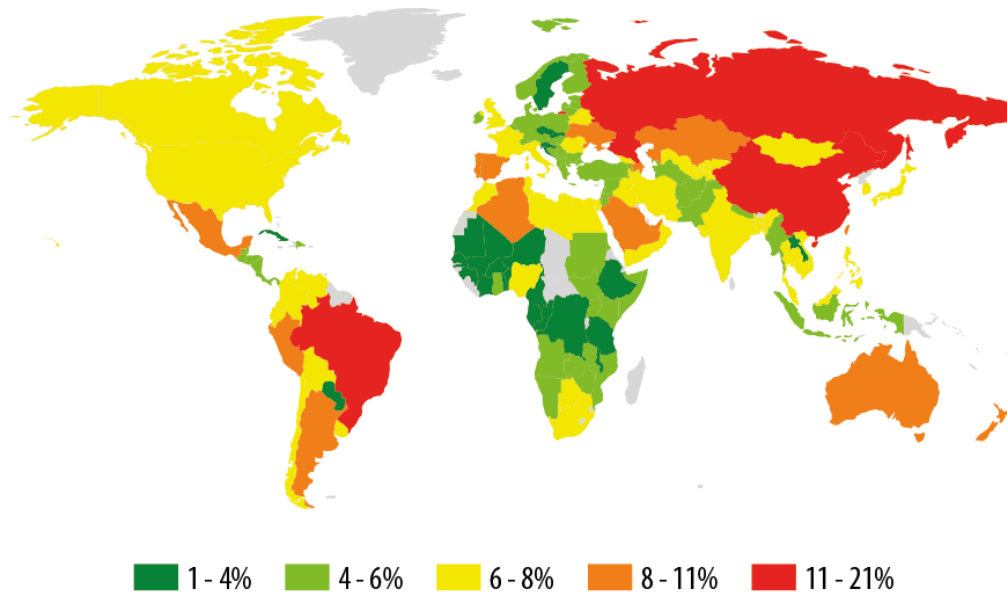
Visualization

leverage increased processing
power of the human perception

Interaction

encourage exploration of data
by highlighting/zooming/linking





© 2017 Kaspersky Lab. All Rights Reserved.

<https://securelist.com/spam-and-phishing-in-q1-2017/78221/>

Visual analytics can benefit our research
but how should it be implemented?

Applying visual analytics in web security research

- ① The benefits and challenges
- ② Our design
- ③ A preliminary evaluation

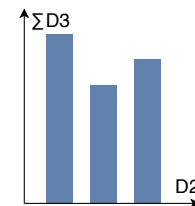
Visual analytics allows to gain insights
into large and diverse web security data sets

- Measurements of large sections of the Internet

Visual analytics allows to gain insights into large and diverse web security data sets

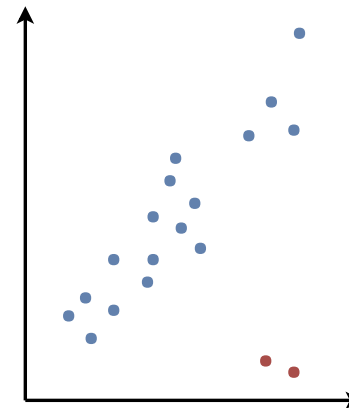
- Measurements of large sections of the Internet
 - Get a visual **overview** of the data
Challenge: scale data processing + representation

D1	D2	D3
1.5	AXS	3
2.7	AXS	9
1.4	BGH	4
3.1	PLU	4



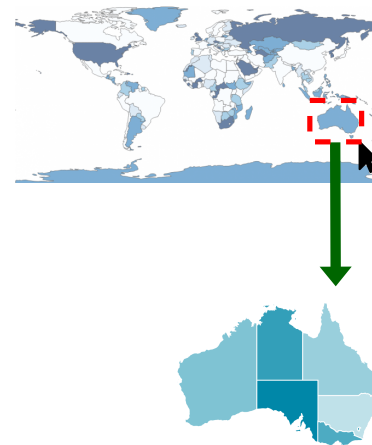
Visual analytics allows to gain insights into large and diverse web security data sets

- Measurements of large sections of the Internet
 - Get a visual **overview** of the data
 - Detect **patterns and outliers** visually
Challenge: avoid misrepresenting data



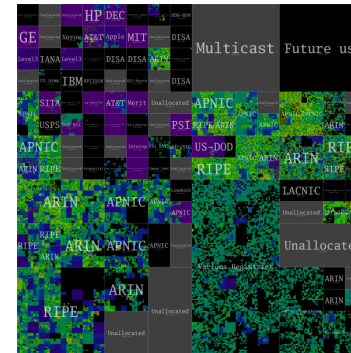
Visual analytics allows to gain insights into large and diverse web security data sets

- Measurements of large sections of the Internet
 - Get a visual **overview** of the data
 - Detect **patterns and outliers** visually
 - **Interact** to study parts in more detail
Challenge: load detailed data on demand



Visual analytics allows to gain insights into large and diverse web security data sets

- Measurements of large sections of the Internet
 - Get a visual **overview** of the data
 - Detect **patterns and outliers** visually
 - **Interact** to study parts in more detail
 - Use **domain-specific visualizations**
Challenge: integrate data types/structures

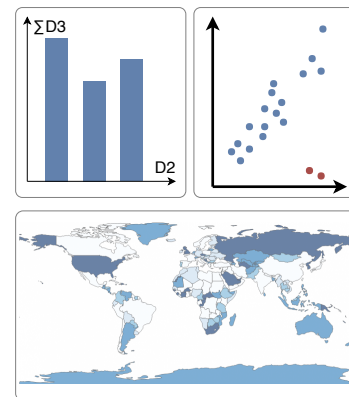


Visual analytics allows to gain insights
into large and diverse web security data sets

- Relations between collected data sets

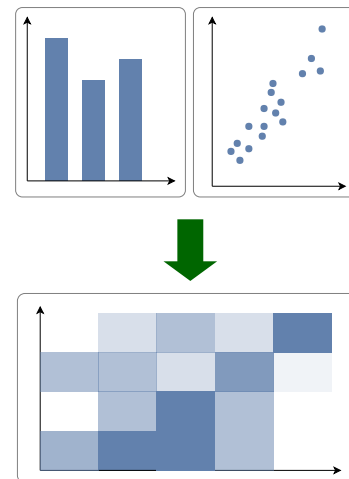
Visual analytics allows to gain insights into large and diverse web security data sets

- Relations between collected data sets
 - Explore simultaneously through a **dashboard**
Challenge: difficulty of creating visualizations [FNER09]



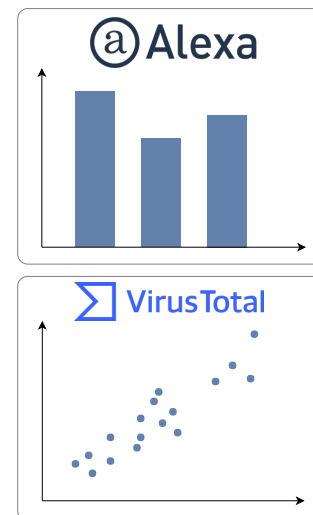
Visual analytics allows to gain insights into large and diverse web security data sets

- Relations between collected data sets
 - Explore simultaneously through a **dashboard**
 - **Combine** data sets interactively
- Challenge: heterogeneous data storage*



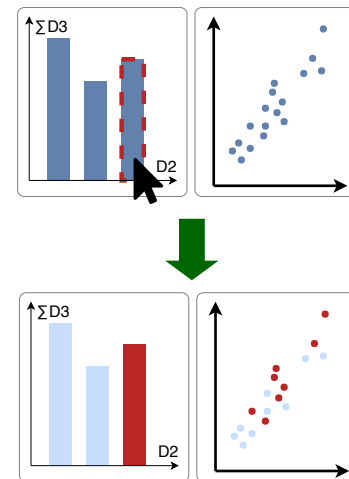
Visual analytics allows to gain insights into large and diverse web security data sets

- Relations between collected data sets
 - Explore simultaneously through a **dashboard**
 - **Combine** data sets interactively
 - Augment with **public data sets**
Challenge: select/integrate interesting sources



Visual analytics allows to gain insights into large and diverse web security data sets

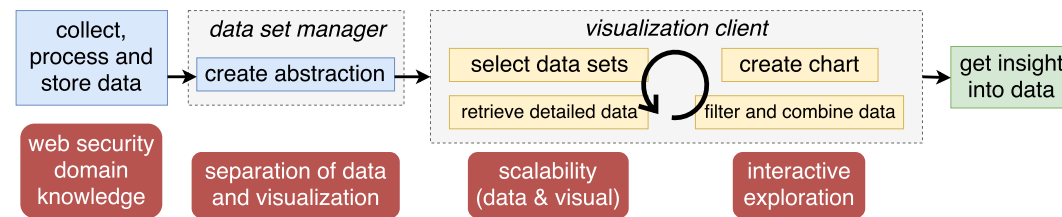
- Relations between collected data sets
 - Explore simultaneously through a **dashboard**
 - **Combine** data sets interactively
 - Augment with **public data sets**
 - **Synchronize** selections across charts
- Challenge: enable interactions in interface*



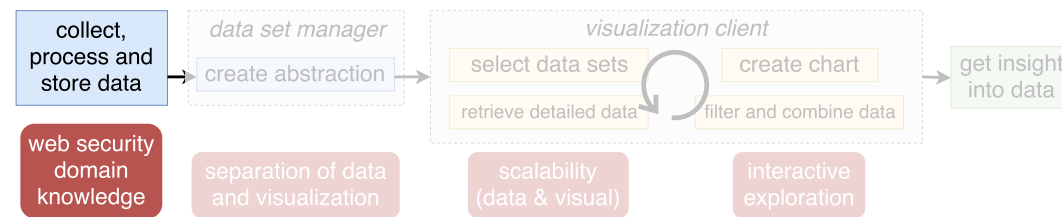
Applying visual analytics in web security research

- ① The benefits and challenges
- ② Our design
- ③ A preliminary evaluation

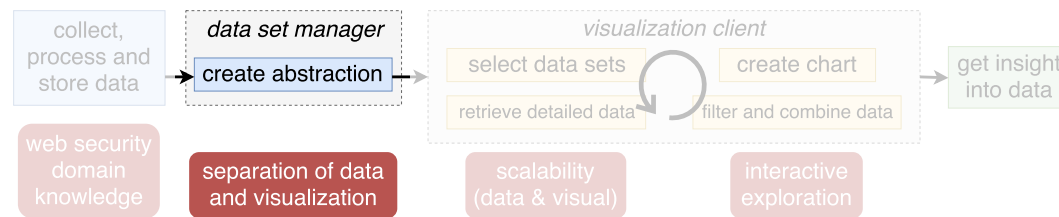
Our solutions to the challenges



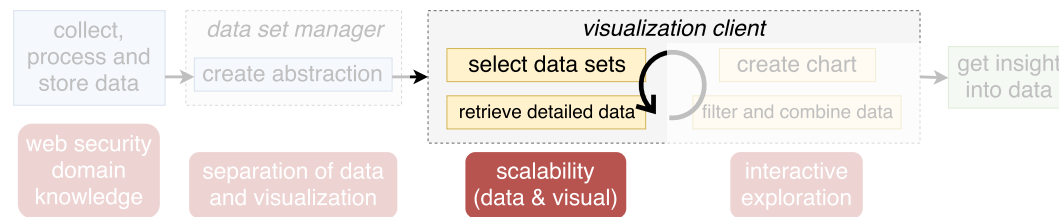
Our solutions to the challenges



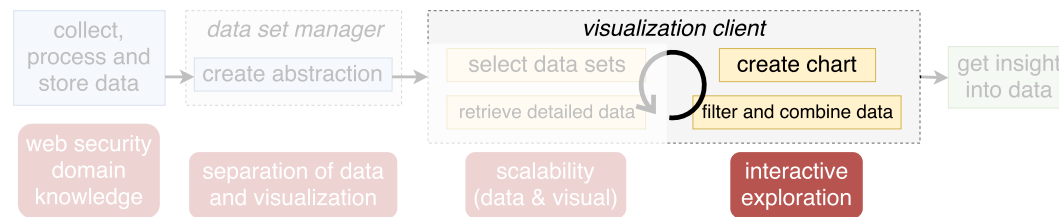
Our solutions to the challenges



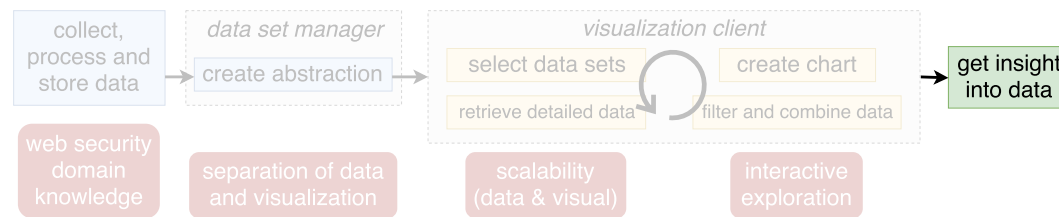
Our solutions to the challenges



Our solutions to the challenges



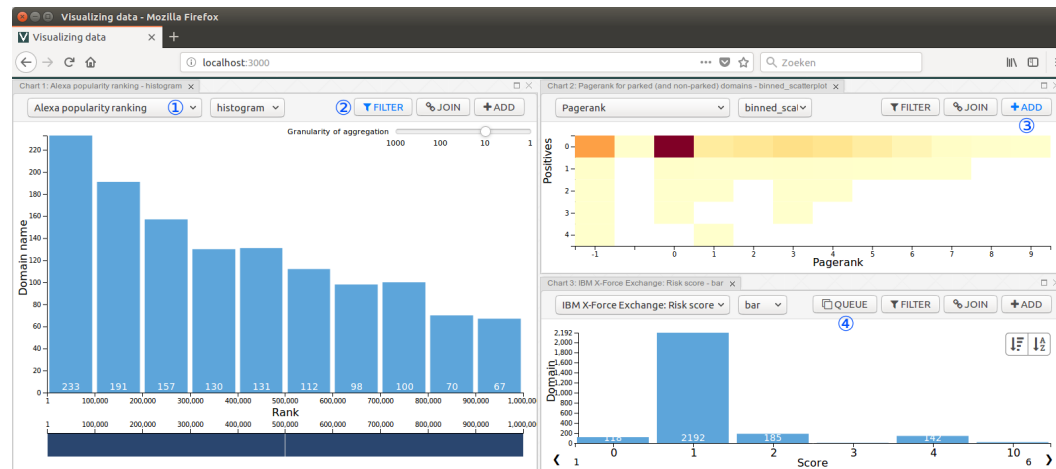
Our solutions to the challenges



Applying visual analytics in web security research

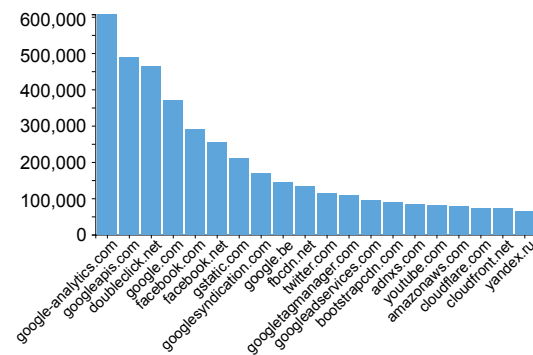
- ① The benefits and challenges
- ② Our design
- ③ A preliminary evaluation

Evaluating our solution to the challenges



A case study of studying web security data using visual analytics

- Online tracking across one million websites [EN16]
 - Set up inclusions data set
 - Display popular sites from Alexa
 - Queue sites for inclusions crawling
 - Visualize inclusions data
 - Explore data set through interaction



Evaluating our solution to the challenges

- Functionality
 - Combine transformed data sources
 - Modular approach to chart types
 - Interactively queue items for data collection
 - Add interactive data analysis

Evaluating our solution to the challenges

- Functionality
- Performance
 - Aggregated view by default
 - Case study: 10x faster, 10^6 x smaller document

Evaluating our solution to the challenges

- Functionality
- Performance
- Productivity
 - Transforming a data set: higher effort (code or parameters)
 - Visualizing data: low effort (handful operations)
 - Further simplify to reduce effort

Evaluating our solution to the challenges

- Functionality
- Performance
- Productivity
- Utility and usability
 - User study with researchers and experts

Applying visual analytics in web security research

- ① The benefits and challenges
- ② Our design
- ③ A preliminary evaluation

Visual analytics can benefit our research...

- Analyzing data more **efficiently**
- Extracting **more insights**
- Interpreting the web more **accurately**
- Improving the **reproducibility** of studies

...but we need to address some challenges before implementing it

- Coping with the **scale and diversity of data**
- Reducing the effort in **creating interactive visualizations**
- Supporting the **domain-specific properties** of web security

Visual analytics will benefit our research

- Goal: further develop our visual analytics approach
- Create models (developing security metrics & entity mappings)
- Involve researchers in eliciting requirements and evaluating tools
- Release tool to fellow researchers



Idea: Visual Analytics for Web Security

Victor Le Pochat, Tom Van Goethem, Wouter Joosen

victor.lepochat@cs.kuleuven.be

imec-DistriNet, KU Leuven, Celestijnenlaan 200A, 3001 Leuven, Belgium

ESSoS 2018, June 27, 2018



DistriNet



References I

- [EN16] [S. Englehardt and A. Narayanan](#).
Online tracking: A 1-million-site measurement and analysis.
In Proceedings of the 2017 ACM Conference on Computer and Communications Security (CCS 2017), pp. 1388–1401. ACM, 2016.
- [FNER09] [G. A. Fink, C. L. North, A. Endert, and S. Rose](#).
Visualizing cyber security: Usable workspaces.
In Proc. VizSec, pp. 45–56, 2009.
- [Kei01] [D. A. Keim](#).
Visual exploration of large data sets.
Communications of the ACM, 44(8):38–44, August 2001.
- [TC05] [J. J. Thomas and K. A. Cook](#), editors.
Illuminating the Path: The Research and Development Agenda for Visual Analytics.
IEEE Computer Society Press, 2005.