SAND2017-7576C

The Center for Cyber Defenders

Expanding computer security knowledge

Current Cryptographic Interests

Yongwoon Lee Escobar, University of Illinois, Urbana-Champaign

Project Mentor: N. Andrew Fisher, Org. 5845

Problem Statement:

Given the research activity in cryptography, how does one determine current research interests and activities in this field?

Objectives:

- What is the next "big thing"?
- Who and from which institutions?
- What are the current implementations?

Approach:

1. Identify data sources and extract
2. Determine which data fields to use:
   - Frequency of terms over publications
   - Citations vs. Time
3. Process the data fields through Citrus + plugin tools:
   - "Sticky / Sparky": How a term goes viral through publications and citations.
   - Citation Analysis
4. Generate and visualize the results

Scientific meme: how popular is a term?

"Sticky": How far the term spreads. Example: "Advanced Encryption Standard" should have high value; "perpetual energy" should have a low value.

$$\text{Sticky} = \frac{\text{Articles with the term \& cites others with term}}{\text{Articles that cite others with term}}$$

"Sparky": How general or common the term is. Example: "Security" should have a high value; "Supersingular Elliptic Curves" should have a low value

$$\text{Sparky} = \frac{\text{Articles with the term \& does not cite others with term}}{\text{Articles that does not cite others with term}}$$

Propagation score: How interesting the term is

$$\text{Propogation score} = \frac{\text{Sticky}}{\text{Sparky}}$$

Meme Score: How popular and significant the term is

$$\text{Meme score} = \text{Propagation Score} \cdot (\%\ \text{articles with the term})$$

Example: Top 10 scientific terms by Meme score in physics using scientific meme tool on 500,000+ journal articles. All terms considered significant.

1. Loop quantum cosmology
2. Unparticle
3. Sonoluminescence
4. MgB_2
5. Stochastic Resonance
6. Carbon Nanotubes
7. NbSe_3
8. Black hole
9. Nanotubes
10. Lattice Boltzmann

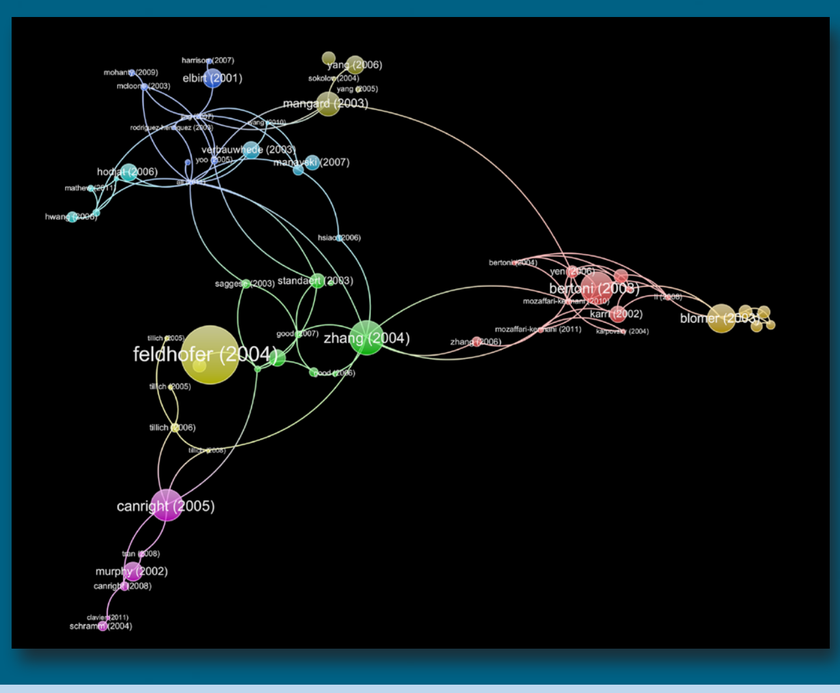
Tobias Kuhn, Matjaz Perc, Dirk Helbing. "Inheritance Patterns in Citation Networks Reveal Scientific Memes". Table 1. 2014. Physical Review X 4, 041036 (2014).

Objectives:

- Extract Data: Downloaded (10000+) articles from Intern. Assoc. for Cryptologic Research
- Prepare Data: generate metadata + citations for each article using Grobid tool
- Identified a tool for Sticky/Sparky process: Scientific Memes

Current Work (30 June 2017)

- Adapt Grobid and Sticky/Sparky method as a Citrus Plugin
- Implement method to visualize the results
- Automate data extraction of data sources



The dream: Visualizing a network of terms, individuals, and institutions in cryptography. Above: Visualization of citation network on 500+ articles from IACR related to Advanced Encryption Standard from 2010 – 2017 using VosViewer tool.

Impact and Benefits:

- Overall picture of current research and latest implementations
- Assess current literature coverage and interests:
  - What do we know
  - What do we need to learn
- Highlight individuals, institutions, and articles of recent high impact

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC., a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525. SAND2017-XXXXC.