



**Universiteit  
Leiden**  
The Netherlands

## **Algorithms for analyzing evolving networks on the Dark Web & in science**

Boekhout, H.D.

### **Citation**

Boekhout, H. D. (2026, March 17). *Algorithms for analyzing evolving networks on the Dark Web & in science*. Retrieved from <https://hdl.handle.net/1887/4297227>

Version: Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/4297227>

**Note:** To cite this publication please use the final published version (if applicable).

# **Algorithms for Analyzing Evolving Networks on the Dark Web & in Science**

Hanjo D. Boekhout

Copyright ©2026 Hanjo Daniël Boekhout

Without written permission of the author, no part of this thesis may be reproduced, stored, or published in any form.

Typeset using L<sup>A</sup>T<sub>E</sub>X, figures generated using TIKZ, draw.io, Matplotlib and Geopandas.

Cover design: Lynn Putman and Hanjo D. Boekhout. Made with Natural Earth.

Open-access: <https://scholarlypublications.universiteitleiden.nl/>

Printed: Ridderprint, [www.ridderprint.nl](http://www.ridderprint.nl)

ISBN: 978-94-6537-300-3

# Algorithms for Analyzing Evolving Networks on the Dark Web & in Science

## Proefschrift

ter verkrijging van  
de graad van doctor aan de Universiteit Leiden,  
op gezag van rector magnificus prof.dr. S. de Rijcke,  
volgens besluit van het college voor promoties  
te verdedigen op dinsdag 17 maart 2026  
klokke 14:30

door

Hanjo Daniël Boekhout  
geboren te Noordwijkerhout  
in 1991

**Promotores:**

Dr. F.W. Takes

Dr. W.A. Kusters

**Promotiecommissie:**

Prof.dr. T.H.W. Bäck

Prof.dr.mr. A.A.J. Blokland

Prof.dr. M.M. Bonsangue

Prof.dr. L.R. Waltman

Prof.dr. E.M. Heemskerk — Universiteit van Amsterdam

Dr. C.E.S. Mattsson — CENTAI, Turin, Italy



This research was conducted while the author was employed at the Leiden Institute of Advanced Computer Science (LIACS) and held a guest employment at the Centre for Science and Technology Studies (CWTS) at Leiden University.



Part of this research was funded by Politie & Wetenschap.

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Real-world networks . . . . .	1
1.2	Network science theory and metrics . . . . .	4
1.2.1	Network types . . . . .	6
1.2.2	Network metrics . . . . .	7
1.3	Introduction to two real-world networked systems . . . . .	11
1.3.1	Dark web cryptomarkets . . . . .	11
1.3.2	Scientific research and publication system . . . . .	14
1.4	Data quality . . . . .	15
1.5	Thesis outline . . . . .	17
<b>I</b>	<b>Dark web cryptomarket networks</b>	<b>21</b>
<b>2</b>	<b>A large-scale longitudinal structured dataset of the dark web cryptomarket Evolution (2014–2015)</b>	<b>23</b>
2.1	Background and summary . . . . .	24
2.2	Methods . . . . .	26
2.2.1	Step 1: Extraction of forum and market data from source files . . . . .	26
2.2.2	Step 2: Resolving data duplication and inconsistencies in extracted forum and market data . . . . .	28
2.2.3	Step 3: Forum user and market vendor matching . . . . .	33
2.3	Data records . . . . .	33

---

2.3.1	Forum . . . . .	33
2.3.2	Market . . . . .	37
2.3.3	User matching . . . . .	39
2.4	Technical validation . . . . .	39
2.4.1	Forum and market data completeness . . . . .	40
2.4.2	Data reliability . . . . .	47
2.5	Communication network extraction and analysis . . . . .	48
2.5.1	Network extraction . . . . .	48
2.5.2	Network records . . . . .	50
2.5.3	Network analysis . . . . .	51
2.6	Usage notes . . . . .	53
<b>3</b>	<b>Early warning signals for predicting cryptomarket vendor success using dark net forum networks</b>	<b>55</b>
3.1	Introduction . . . . .	56
3.2	Results . . . . .	58
3.2.1	Data . . . . .	58
3.2.2	Network measures and activity indicators . . . . .	60
3.2.3	Distinguishing vendors and their level of success . . . . .	61
3.2.4	Detecting vendors in the user base . . . . .	66
3.2.5	Key player identification . . . . .	72
3.3	Discussion and conclusion . . . . .	75
3.4	Methods . . . . .	77
3.4.1	Dataset . . . . .	78
3.4.2	Network extraction . . . . .	78
3.4.3	Activity indicators . . . . .	80
3.4.4	Network centrality measures . . . . .	81
3.4.5	Evaluation metrics . . . . .	84
3.5	Supplementary results . . . . .	86
3.5.1	Robustness of results . . . . .	86
3.5.2	Post activity and sales recall . . . . .	92
3.5.3	Sales and post activity of recalled and non-recalled users . . . . .	93
3.5.4	Performance at different thresholds . . . . .	98
<b>II</b>	<b>Temporal weighted maximal clique enumeration</b>	<b>101</b>
<b>4</b>	<b>Fast maximal clique enumeration in weighted temporal networks</b>	<b>103</b>
4.1	Introduction . . . . .	104
4.2	Related work . . . . .	106

---

4.3	Preliminaries . . . . .	107
4.3.1	Basic notation and definitions . . . . .	107
4.3.2	Maximal cliques . . . . .	107
4.4	Fast $(\delta, \gamma)$ -maximal clique enumeration . . . . .	110
4.4.1	Overall approach . . . . .	110
4.4.2	Stretching phase . . . . .	111
4.4.3	Bulking phase . . . . .	116
4.5	Data . . . . .	124
4.6	Experiments . . . . .	126
4.6.1	Experimental setup . . . . .	126
4.6.2	Results . . . . .	128
4.7	Conclusions . . . . .	138
4.8	Appendix . . . . .	141
4.8.1	Algorithm pseudo code . . . . .	141
4.8.2	Supplementary result figures . . . . .	149

### III Scientific co-authorship networks 153

<b>5</b>	<b>Evolution of the world stage of global science from a scientific city network perspective</b>	<b>155</b>
5.1	Introduction . . . . .	156
5.2	Definitions, measures, and background . . . . .	157
5.2.1	Network notation and definitions . . . . .	157
5.2.2	Ranking measures . . . . .	158
5.2.3	Rank comparison measures . . . . .	159
5.3	Materials and methods . . . . .	160
5.3.1	Bibliographic database . . . . .	160
5.3.2	Delineation of scientific city agglomerations - nodes . . . . .	160
5.3.3	Publication sets and counting methods - edge weights . . . . .	162
5.3.4	Network formation . . . . .	163
5.3.5	Evolving degree respecting rewired networks . . . . .	164
5.4	Results . . . . .	164
5.4.1	Experimental setup . . . . .	164
5.4.2	Centrality changes over time at various levels of impact . . . . .	164
5.4.3	Real-world rank correlation vs. null model . . . . .	167
5.4.4	Limitations . . . . .	168
5.5	Conclusions . . . . .	168

<b>6</b>	<b>Freshness, persistence and success of scientific teams</b>	<b>171</b>
6.1	Introduction . . . . .	172
6.2	Results . . . . .	175
6.2.1	Success of persistent teams . . . . .	175
6.2.2	Team freshness . . . . .	175
6.2.3	Team composition . . . . .	178
6.2.4	Team openness and success . . . . .	180
6.3	Conclusion . . . . .	184
6.4	Materials and Methods . . . . .	185
6.4.1	Dataset . . . . .	185
6.4.2	Identifying persistent teams . . . . .	187
6.4.3	Overlapping persistent teams . . . . .	189
6.5	Supplementary results . . . . .	191
<b>7</b>	<b>Conclusion</b>	<b>197</b>
7.1	Methodological and algorithmic advances . . . . .	197
7.2	Dark Web Market findings . . . . .	200
7.3	Scientific research system findings . . . . .	201
	<b>Bibliography</b>	<b>203</b>
	<b>Summary</b>	<b>217</b>
	<b>Samenvatting</b>	<b>221</b>
	<b>Curriculum Vitae</b>	<b>225</b>
	<b>Acknowledgments</b>	<b>227</b>
	<b>Publication List</b>	<b>229</b>