



Content Types Subjects

English

Athens Authentication Point

**Recognized as:**

The University of Western Australia (565-28-928)

GO8 Universities Australia (647-04-468)

**Welcome!**

To use the personalized features of this site, please **log in** or **register**.

If you have forgotten your username or password, we can **help**.

**My SpringerLink**

Marked Items

Alerts

Order History

**Saved Items**

All

Favorites

**Book Chapter**



Term Frequency Normalization via Pareto Distributions

Book Series

Publisher

ISSN

Subject

Volume

Book

Copyright

Page

SpringerLink Date

Lecture Notes in Computer Science

Springer Berlin / Heidelberg

0302-9743 (Print) 1611-3349 (Online)

Computer Science

Volume 2291/2002

Advances in Information Retrieval : 24th BCS-IRSG European Colloquium on IR Research Glasgow, UK, March 25-27, 2002. Proceedings

2002

183

Thursday, February 19, 2004

**Add to marked items**

Add to shopping cart

Add to saved items

Recommend this chapter



**Authors**

Gianni Amati, Cornelis Joost van Rijsbergen

**Abstract**

We exploit the Feller-Pareto characterization of the classical Pareto distribution to derive a law relating the probability of a given term frequency in a document and its the length. A similar law was derived by Mandelbrot. We exploit the paretian distribution to obtain a term frequency normalization to substitute for the actual term frequency in the probabilistic models of Information Retrieval recently introduced in TREC-10. Preliminary results show that the unique parameter of the framework can be eliminated in favour of the the term frequency normalization derived by the Paretian law.

**Find** **more options**
  

- Within all content
- Within this book series
- Within this book

**Export this chapter**

Export this chapter as RIS | Text

**Text****PDF**

The size of this document is 145 kilobytes. Although it may be a lengthier download, this is the most authoritative online format.

Open: Entire document

Ads by Google**Transportation Models**

Download Free Trial & Examples Optimize Product Distribution  
[www.solver.com](http://www.solver.com)

**Prestige Fulfillment**

Fulfillment Services/Outsourcing Located near San Diego & Long Beach  
[www.pfclogistics.com](http://www.pfclogistics.com)

**Monte Carlo Simulation**

Download a Monte Carlo simulation engine for Excel spreadsheet models  
[www.riskamp.com](http://www.riskamp.com)

**WeibullPro from Isograph**

Weibull Software for the Analysis of Failure and Repair Data  
[www.isograph-software.com](http://www.isograph-software.com)

**Probability Distribution**

Online Help With Your Dissertation, Thesis or any Question. From \$0.17!  
[Statistics.kasamba.com](http://Statistics.kasamba.com)

[Frequently asked questions](#) | [General information on journals and books](#) | [Send us your feedback](#) | [Impressum](#)

© Springer. Part of Springer Science+Business Media

[Privacy](#), [Disclaimer](#), [Terms and Conditions](#), [© Copyright Information](#)

Remote Address: 130.95.1.118 • Server: mpweb20

HTTP User Agent: Mozilla/5.0 (Windows; U; Windows NT 5.1; en-US; rv:1.8.1.3) Gecko/20070309 Firefox/2.0.0.3