REPORT DOCUMENTATION PAGE

Form Approved OMB NO. 0704-0188

1. REPORT DATE (DD-MM-YYYY)	2. REPORT TYPE	3. DATES COVERED (From - To)
30-07-2010	Final Report	1-Jun-2007 - 31-May-2010

4. TITLE AND SUBTITLE
Final report: Prob/Stat Modeling of Complex Systems Exhibiting LRD and Heavy Tails

5a. CONTRACT NUMBER
W911NF-07-1-0078

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER
611102

6. AUTHORS
S. Resnick, G. Samorodnitsky

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAMES AND ADDRESSES
Cornell University
Office of Sponsored Programs
Cornell University
Ithaca, NY 14853 -2801

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Research Office
P.O. Box 12211
Research Triangle Park, NC 27709-2211

10. SPONSOR/MONITOR'S ACRONYM(S)
ARO

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
52286-MA.23

12. DISTRIBUTION AVAILIBILITY STATEMENT
Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES
The views, opinions and/or findings contained in this report are those of the author(s) and should not contrued as an official Department of the Army position, policy or decision, unless so designated by other documentation.

14. ABSTRACT
We and our students conducted a research program in probability and statistics about applications of non-standard stochastic models, emphasizing their large deviations, extreme values and dependence properties. We focussed on structural and distributional properties that explain important relationships and consequences. We emphasized realistic settings by often testing models with data. The applications areas covered complex networks of var-

15. SUBJECT TERMS
risk, extremes, conditional model, multivariate analysis, detection, returns, heavy tails, long range dependence, random geometry, inference

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	15. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Sidney Resnick
UU	UU	UU	UU		19b. TELEPHONE NUMBER: 160-725-5121

Report Title

Final report: Prob/Stat Modeling of Complex Systems Exhibiting LRD and Heavy Tails

ABSTRACT

We and our students conducted a research program in probability and
statistics about applications of non-standard stochastic models, emphasizing their large de-
viations, extreme values and dependence properties. We focussed on structural and distributional properties that explain important relationships and consequences. We emphasized realistic settings by often testing models with data. The applications areas covered complex networks of var-
ious types including data networks and explored reliability issue involving estimation, risk analysis and financial control. Our models were typically non-Gaussian, often driven by Poisson or Levy noises, and often possessed heavy tails and/or long range dependence; often models exhibited unusual fractal and scaling behavior.

List of papers submitted or published that acknowledge ARO support during this reporting period. List the papers, including journal references, in the following categories:

(a) Papers published in peer-reviewed journals (N/A for none)

Mitra, A. and S.I. Resnick, (2009)."Aggregation of risks and asymptotic independence". Advances in Applied Probability, 41(3), 797-828.

Fasen, V. and Samorodnitsky, G. (2009)``The workload process with a Poisson cluster input can look like a Fractional Brownian motion even in the slow growth regime". Advances in Applied Probability, 41(2), 393-427.

Yeghiazarian, L., G. Samorodnitsky, and C. D. Montemagno. (2009). A Poisson random field model of pathogen transport in surface water. Water Resour. Res., 45, W11415, doi:10.1029/2009WR007896, 10 pages, 4 figures, 2 tables.

Grabchak, M. and Samorodnitsky, G. (2010) ``Do financial returns have finite or infinite variance? A paradox and an explanation". Quantitative Finance, 1, 11 pages.

Hult, H. and Samorodnitsky, G. (2010) ``Large deviations for point processes based on stationary sequences with heavy tails". Advances in Applied Probability, 47(1), 1-40.

Adler, R., Samorodnitsky, G. and Taylor, J. (2010) ``Excursion sets of three classes of stable random fields". Advances in Applied Probability 42(2), 293-318.

Ghosh, S. and Resnick, S. (2010) "A discussion on mean excess plots." Stochastic Processes and their Applications, 120(8), 1492-1517.

Lopez-Oliveros, L. and Resnick, (2010) S. "Extremal dependence analysis of network sessions." Extremes, DOI: 10.1007/s10687-009-0096-4.

Das, B. and Resnick, S.(2010) "Detecting a conditional extreme value model." Extremes. DOI: 10.1007/s10687-009-0097-3.

Das. B. and Resnick, S. (2010). Conditioning on an extreme component: Model consistency with regular variation on cones. Bernoulli. http://isi-web.org/bj/future

Number of Papers published in peer-reviewed journals: 10.00

(b) Papers published in non-peer-reviewed journals or in conference proceedings (N/A for none)

Demichel, Y., Estrade, A., Kratz, M. and Samorodnitsky, G. (2010)``How Fast Can the Chord-Length Distribution Decay? Submitted.

Ghosh, S. and Samorodnitsky, G. (2010) ``Long strange segments, ruin probabilities and the effect of memory on moving average processes". Submitted.

Roueff, F., Samorodnitsky, G. and Soulier, P. (2010). ``Function-indexed empirical processes based on an infinite source Poisson transmission stream". submitted.

A. Mitra and S.I. Resnick, (2010). "Hidden regular variation: Detection and estimation". Submitted.

Number of Papers published in non peer-reviewed journals: 4.00

(c) Presentations

Resnick, S. "Modeling Data Network Sessions."
New Topics at the Interface Between Probability and Communications
11 January to 15 January 2010. Isaac Newton Institute for Mathematical Sciences, Cambridge, UK

Resnick, S. "Bringing some coherence to multivariate extremes, hidden regular variation and conditional extremes"; 3 hours. September 14-18, 2009, `Risk, Rare Events and Extremes', the Bernoulli Centre, EPFL, Lausanne, Switzerland. Workshop on High-dimensional extremes.

Resnick, S. "Modeling Data Network Sessions: Multivariate Heavy Tails and Conditional Extremes." December 7, 2009. Workshop on Extreme Values in honor of Prof Ishay Weissman. Technion, Haifa, Israel.

Resnick, S. "Heavy tailed statistical analysis of network session." March 1-5, 2010. 9th German Open Conference on Probability and Statistics, Leipzig, Germany.

Samorodnitsky, G. Stochastic Networks and Related Topics II, a conference organized by the Institute of Mathematics of the Polish Academy of Sciences, Bedlewo, Poland, May 17-22, 2009. Invited talk "Large deviations for point processes based on stationary sequences with heavy tails".

Samorodnitsky, G. Organized session: "Large Deviations" and gave invited talk "Geometric characteristics of the excursion sets over high levels of non-Gaussian infinitely divisible random fields" at the 6th International Conference on Extreme Value Analysis, Fort Collins, CO June 22-26, 2009.

Samorodnitsky, G. International Conference on self-similar processes and their applications, Angers, France, July 20-24, 2009. Keynote talk ``What self-similar processes best describe the input to communication network models?"

Samorodnitsky, G. 33d Conference on Stochastic Processes and their Applications, Berlin, July 27-31, 2009. Invited talk ``Long range dependence and large deviations".

Samorodnitsky, G. Conference on Recent Advances in Heavy Tailed Modeling in Finance, Brussels, April 28-30, 2010. Talk: Do financial return have finite or infinite variance? A paradox and an explanation.

Samorodnitsky, G. Workshop on Time Series, Extremes and Dependence, Luminy, France, May 25-30, 2010. Talk: The effect of memory on large deviations and ruin probabilities.

Number of Presentations: 10.00

Non Peer-Reviewed Conference Proceeding publications (other than abstracts):

Number of Non Peer-Reviewed Conference Proceeding publications (other than abstracts): 0

Peer-Reviewed Conference Proceeding publications (other than abstracts):

Number of Peer-Reviewed Conference Proceeding publications (other than abstracts): 0

(d) Manuscripts

Number of Manuscripts: 0.00

Patents Submitted

Patents Awarded

Graduate Students

NAME	PERCENT_SUPPORTED
U. Can $3000, summer 09	0.10
A. Mitra summer 10	0.10
A. Chakrabarty $3000, summer 09	0.10
FTE Equivalent:	**0.30**
Total Number:	**3**

Names of Post Doctorates

NAME	PERCENT_SUPPORTED
FTE Equivalent:	
Total Number:	

Names of Faculty Supported

NAME	PERCENT_SUPPORTED	National Academy Member
Gennady Samorodnitsky (one month,	0.11	No
Sidney Resnick (one month, summer	0.11	No
FTE Equivalent:	**0.22**	
Total Number:	**2**	

Names of Under Graduate students supported

NAME	PERCENT_SUPPORTED
FTE Equivalent:	
Total Number:	

Student Metrics

This section only applies to graduating undergraduates supported by this agreement in this reporting period

The number of undergraduates funded by this agreement who graduated during this period:	0.00
The number of undergraduates funded by this agreement who graduated during this period with a degree in science, mathematics, engineering, or technology fields:.....	0.00
The number of undergraduates funded by your agreement who graduated during this period and will continue to pursue a graduate or Ph.D. degree in science, mathematics, engineering, or technology fields:.....	0.00
Number of graduating undergraduates who achieved a 3.5 GPA to 4.0 (4.0 max scale):	0.00
Number of graduating undergraduates funded by a DoD funded Center of Excellence grant for Education, Research and Engineering:	0.00
The number of undergraduates funded by your agreement who graduated during this period and intend to work for the Department of Defense	0.00
The number of undergraduates funded by your agreement who graduated during this period and will receive scholarships or fellowships for further studies in science, mathematics, engineering or technology fields:	0.00

Names of Personnel receiving masters degrees

NAME
D. Zeber
Total Number: 1

Names of personnel receiving PHDs

NAME
A. Chakraborty
B. Das
Total Number: 2

Names of other research staff

NAME	PERCENT_SUPPORTED
FTE Equivalent:	
Total Number:	

Sub Contractors (DD882)

Inventions (DD882)

W911NF0710078
Probabilistic and Statistical Modeling of Complex Systems Exhibiting Long Range Dependence and Heavy Tails

PI's: S.I. Resnick and G. Samorodnitsky, School of Operations Research and Industrial Engineering, Cornell University, Rhodes Hall, Ithaca NY 14853; sir1@cornell.edu, (607) 255 1210; gennady@orie.cornell.edu, (607) 255 9141.

Technical POC: Harry Chang, Mathematical Sciences Division, Probability and Statistics Program, US Army Research Office.

Problems studied.

General statement. We and our students conducted a research program in probability and statistics concentrating on applications of non-standard stochastic models subject to influence of large deviations, extreme values and dependence. We focussed on structural and distributional properties that explain important relationships and consequences. We emphasized realistic settings by often testing models with data. The applications areas covered complex networks of various types including data networks and explored reliability issues involving estimation, risk analysis and financial control. Our models were typically non-Gaussian, often driven by Poisson or Levy noises, and often possessed heavy tails and/or long range dependence; often models exhibited unusual fractal and scaling behavior.

Current budget year. During the final budget year, G. Samorodnitsky worked on the analysis of stochastic systems, concentrating on the effects of long range dependence and heavy tails. A surprising onclusion from the work "How Fast Can the Chord-Length Distribution Decay" with Demichel, Estrade and Kratz [4] is that in the threshold model for the bi-phasic media the effect of memory on the size of the phases is minimal. On the other hand, long memory has a major effect on the length of long strange segments and on the ruin probabilities for moving average processes, as was shown in a joint work with Ghosh, "Long strange segments, ruin probabilities and the effect of memory on moving average processes" [5]. A similar phenomenon was shown to exist in telecommunication transmission processes, in the work "Function-indexed empirical processes based on an infinite source Poisson transmission stream" with Roueff and Soulier [13].

Mitra and Resnick have concentrated on accurate estimation of risk regions using a technique called *hidden regular variation*. The paper "Hidden Regular Variation: Detection and Estimation" [11] shows how more accurate estimation could be conducted in higher dimensional problems. This work is ongoing and currently being extended in two directions. One direction extends to *hidden domain of attraction* where lower order behavior may be more general than heavy tailed type but still influential [10]. The other direction is to develop techniques for exploring more general regions which can influence risk calculations [1].

Lopez-Oliveros and Resnick are in the process of studying the effect of merging different traffic streams in order to explain the following phenomena: theory [9] predicts cumulative Internet traffic ought to look Gaussian or heavy tailed but in practice one does not observe the heavy tailed possibility. We present empirical evidence and theoretical explanation [8].

This project will morph into a project examining the statistical properties of Internet data segmented by application.

Zeber and Resnick have been considering foundations of the conditional extreme value model [12] with a view to explaining relationships between different proposed definitions [7, 6, 2, 3]. The study illuminates the connection of Markov dependence and the conditional extreme value model and this is the subject of current intensive study.

Important results

Our goal has been to understand and quantify the effects of heavy tails and long range dependence on complex systems subject to risk in order to better estimate failure and exceedance probabilities. This has led to improved risk estimates using a unified theory of heavy tails on cones that relates hidden regular variation, conditional limit laws and traditional theory. We have developed sensitive and sensible empirical tools to validate new and classical heavy tailed models from data. We have clarified how long range dependence and heavy tails affect frequency of rare but potentially disastrous events. We have been able to infer the structure of a spatial heavy tailed random field using the severity and clustering of peaks and valleys. By example we have shown that in the presence of long memory, what was thought to be a 100 year event assuming lack of memory becomes a 10year event when properly accounting for long memory. We have also demonstrated by example that classical estimation methods based on the Gaussian copula and multivariate extreme value theory result in an inaccurate estimate of the probability of simultaneous exceedance of critical thresholds to be zero when in actual fact, our improved methods yield significant non-zero probabilities.

References

[1] B. Das, A. Mitra, and S.I. Resnick. Hidden regular variation on cones. Technical report, Cornell University, 2010. In preparation.

[2] B. Das and S.I. Resnick. Conditioning on an extreme component: Model consistency and regular variation on cones. Technical report, Cornell University, School of ORIE, 2008. `http://arxiv.org/abs/0805.4373`; to appear: *Bernoulli*.

[3] B. Das and S.I. Resnick. Detecting a conditional extreme value model. Technical report, Cornell University, School of ORIE, 2008. `http://arxiv.org/abs/0902.2996`; to appear: *Extremes*.

[4] Y. Demichel, A. Estrade, M. Kratz, and G. Samorodnitsky. How fast can the chord-length distribution decay? `http://hdl.handle.net/1813/13700`, 2009.

[5] S. Ghosh and G. Samorodnitsky. Long strange segments, ruin probabilities and the effect of memory on moving average processes. *Arxiv preprint arXiv:1002.2751*, 2010.

[6] J.E. Heffernan and S.I. Resnick. Limit laws for random vectors with an extreme component. *Ann. Appl. Probab.*, 17(2):537–571, 2007.

[7] J.E. Heffernan and J.A. Tawn. A conditional approach for multivariate extreme values (with discussion). *JRSS B*, 66(3):497–546, 2004.

[8] L. López-Oliveros and S.I. Resnick. On the superposition of traffic components. Technical report, Cornell University, 2010. In preparation.

[9] T. Mikosch, S.I. Resnick, H. Rootzén, and A.W. Stegeman. Is network traffic approximated by stable Lévy motion or fractional Brownian motion? *Ann. Appl. Probab.*, 12(1):23–68, 2002.

[10] A. Mitra and S.I. Resnick. Hidden domain of attraction: extension of hidden regular variation. Technical report, Cornell University, 2010. In preparation.

[11] A. Mitra and S.I. Resnick. Hidden Regular Variation: Detection and Estimation. *Arxiv preprint arXiv:1001.5058*, 2010.

[12] S.I. Resnick and D. Zeber. Foundations of conditional extreme value modeling. Technical report, Cornell University, 2010. In preparation.

[13] F. Roueff, G. Samorodnitsky, and P. Soulier. Function-indexed empirical processes based on an infinite source Poisson transmission stream. *Arxiv preprint arXiv:1004.2182*, 2010.