A Survey on Research in Software Reliability Engineering in the Last Decade

† Joicymara Xavier, †Autran Macêdo, †Rivalino Matias, ‡Lucio Borges
† School of Computer Science, ‡ School of Mathematics
Federal University of Uberlandia, Brazil
joicymara@mestrado.ufu.br, autran@fc.ufu.br, rivalino@fc.ufu.br, lucio@famat.ufu.br

ABSTRACT
In this paper, we present the findings of a systematic review of the SRE literature, covering 300 papers in 95 publication venues (conferences and journals) in the last twelve years. We employ statistic techniques to help mining patterns and trends from the analyzed data set. We identified a prevalence of purely theoretical studies in the surveyed period, where no more than 31% are experimental research. Also, the number of research works in "Reliability Assessment" has clearly increased since 2002, and especially during the last three years. On the other hand, works in "Reliability Modeling" and "Fault Analysis" have been gradually reduced during the last decade.

Categories and Subject Descriptors
D.2.4 [**Software Engineering**]: Software/Program Verification – *reliability*. G.3 **[Mathematics of Computing]**: Probability and Statistics – *reliability and life testing.*

General Terms
Reliability.

Keywords
Software reliability, systematic review, trends.

1. INTRODUCTION
The software complexity has grown dramatically over the past few decades, and this tendency will certainly continue [1]. To deal with this challenging scenario, the discipline of software reliability engineering (SRE) [1] has continually evolved its body of theoretical and practical knowledge. Consequently, many research works have been developed, considering different methodological approaches, techniques, and applications. Due to the importance of this research area, as well as its fast growth in the recent years, it is natural that prospective-related questions be raised, such as: Which SRE topics have been more recently investigated? What research approaches have been adopted in SRE in the last decade? Where the most cited SRE works come from and which problems they solve?

In order to answer these questions we searched for SRE papers using the sentence "*software reliability*" in the following repositories: ACM Digital Library [2], IEEE Xplore [3], Science Direct [4], and Google Scholar [5], in the period from 2002 to 2013. After analyzing every paper, we selected 300 papers more closely related to the SRE area. Later, we applied the Ward method [6] to evaluate the similarities among the selected papers creating 7 distinct clusters for analysis, as shown in Figure 1.

SAC'14, March 24-28, 2014, Gyeongju, Korea.

http://dx.doi.org/10.1145/2554850.2555161
978-1-4503-2469-4/14/03…$15.00.

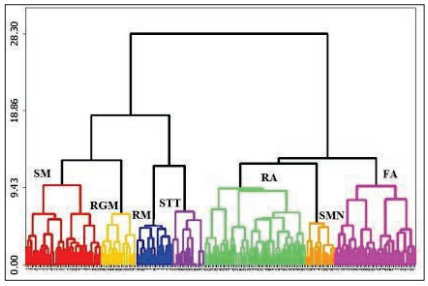

Figure 1. Papers clustered according to keywords.

We named each cluster according to its most frequent keyword. The resulting names (and corresponding colors) are the following: red = Stochastic Modeling (SM); yellow = Software Reliability Growth Model (RGM); blue = Reliability Modeling (RM); purple = Software Testing (STT); green = Reliability Assessment (RA); orange = Software Maintenance (SMN); violet = Fault Analysis (FA).

2. RESULT ANALYSIS
The 300 selected papers covered 304 authors from 92 different affiliations spread into 15 countries. Figure 2 shows the distribution of papers across the clusters.

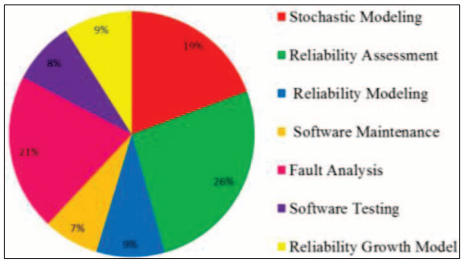

Figure 2. Distribution of papers per clusters (KG).

We observe a pattern where every two or three years there is an increase on the number of papers (i.e., 2006, 2008, and 2011). We hypothesize that this occurs due to the necessary time authors need to maturate their research works, which seems to be two years in average.

Table 1 presents the affiliation of the top ten most productive authors, in terms of number of papers in the samples analyzed. Each line of this table corresponds to one author's affiliation; the

third column, NPP, indicates the number of the author's published papers in the surveyed period (2002-2013); the fourth column contains the average number of citations to the author's papers.

Table 2 presents the affiliation of the top ten most cited authors in our data set, where the fourth column, NC, corresponds to the sum up of the citations to the author's papers (column NPP) according to Google Scholar [5].

Table 1. Affiliations of the top ten most productive authors

Affiliation	Country	NPP	Average
Hiroshima University	JPN	56	6.80
Duke University	USA	49	34.04
National Tsing Hua University	TWN	48	22.02
Hiroshima University	JPN	34	6.79
Chinese Univ. of Hong Kong	HKG	27	33.90
University of Connecticut	USA	25	25.70
Tottori University	JPN	25	9.08
Indian Inst. of Tech. of Kanpur	IND	18	15.44
North Carolina State University	USA	18	46.83
Beihang University	CHN	16	2.00

Table 2. The affiliation of the top ten most cited authors

Affiliation	Country	NPP	NC
Duke University	USA	49	1668
National Tsing Hua University	TWN	48	1057
Chinese Univ. of Hong Kong	HKG	27	916
North Carolina State University	USA	18	843
University of Connecticut	USA	25	643
Microsoft Research	USA	4	436
Duke University	USA	5	436
National Tsing Hua University	TWN	10	423
Hiroshima University	JPN	56	381
Duke University	USA	3	297

We can conclude that researchers from Duke University (USA) and National Tsing Hua University (TWN) had important influence on this research field, considering both number of published papers and citations. Indeed, the most cited paper of each cluster comes from these both countries, where USA relates to clusters SM, RA, RM, SMN, FA, and STT, and Taiwan to RGM. Additionally, we notice that one author from Duke has three papers among the most cited.

In terms of research approach, we considered experimental and theoretical works, as well as both approaches applied together. We observe that 31% (93 out 300 papers) relates to experimental works, where only 13.6% are purely experimental. We hypothesize that this low number is due to two reasons: (1) public experimental data sets in SRE are very limited; (2) producing software reliability data through experiments typically require long time, which sometimes is unattractive for the nowadays-scientific publication process. The clusters "Software Testing" and "Reliability Modeling" bring together the majority of experimental works. On the opposite, the cluster "Software Maintenance" contains the lesser number of experimental works.

In Figure 3, we can observe the annual publication per clusters. It is possible to see that researches in "Reliability Assessment" have made significant progress since 2002, and especially during the last three years. Improvements can also be observed in "Software Testing" research works.

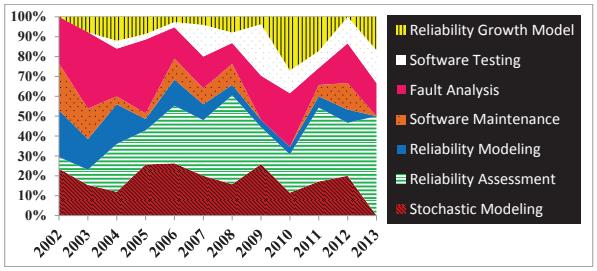

Figure 3: Evolution of annual publications per cluster topic.

The top three most cited experimental related papers are the following: [7] - a theoretical and experimental work with 141 citations; [8] - a purely experimental work with 115 citations; and [9] - another purely experimental work with 104 citations. Similarly, the most frequently cited theoretical related papers are: [10] – with 284 citations; [11] with 282 citations; and [12] with 140 citations.

3. REFERENCES

[1] Lyu, M. R. Software Reliability Engineering: A Roadmap. In *Future of Software Engineering*. IEEE Computer Society, Washington, DC, USA, 2007, 153-170.

[2] ACM. *ACM Digital Library*. ACM, http://dl.acm.org.

[3] IEEE. *IEEE Xplore - Home.* IEEE, http://ieeexplore.ieee.org/ Xplore/home.jsp.

[4] Elsevier. *Science Direct.* Elsevier B. V. http://www. sciencedirect.com.

[5] Google. *Google Scholar*. Google, http://scholar.google.com.

[6] Johnson, R. A. and Wichern, D. W. *Applied Multivariate Statistical Analysis*, Upper Sandler River, New Jersey, 2007.

[7] Nagappan, N., Murphy, B. and Basili, V. R. The influence of organizational structure on software quality: an empirical case study. In *Proc. of the Int'l Conference on Software Engineering*. ACM, NY, USA, 2008, 521-530.

[8] Li, Z., Tan, L., Lu, S., Zhou, Y. and Zhai, C. Have things changed now?: an empirical study of bug characteristics in modern open source software. *Workshop on Architectural and System Support for Improving Software Dependability*. ACM, NY, USA, 2006, 25-33.

[9] Saff, D. and Ernst, M. D. An experimental evaluation of continuous testing during development. In *Proc. of the 2004 ACM Int'l Symp. on Software Testing and Analysis*. ACM, NY, USA, 2004, 76-85.

[10] Nicol, D. M., Sanders, W. H. and Trivedi, K. S. Model-based evaluation: from dependability to security. *IEEE Transactions on Dependable and Secure Computing,* 1,1 (Oct. 2004), 48-65.

[11] Huang, C.-Y., Lyu, M. R.-T. and Kuo, S.-Y. A unified scheme of some Nonhomogenous Poisson process models for software reliability estimation. *IEEE Transactions on Software Engineering,* 29, 3 (March 2003), 261-269.

[12] Nagappan, N. and Ball, T. Use of relative code churn measures to predict system defect density. In *Proc. of the Int'l Conf. on Software Eng*. USA, 2005, 284 - 292.