Software Complexity Analysis Using Halstead Metrics

4 authors, including:

Chandra Segar Thirumalai
VIT University
63 PUBLICATIONS   **1,395** CITATIONS

Hariprasad Thirunavukkarasu
VIT University
1 PUBLICATION   **19** CITATIONS

Some of the authors of this publication are also working on these related projects:

Project   Software metrics assessment and predictions View project

Project   Fuzzy based Decision making system View project

Software Complexity Analysis Using Halstead Metrics

Hariprasad T
MS Software Engineering,
School of Information Technology and Engineering,
VIT University, Vellore, India.
hariprasad.t2013@vit.ac.in

Seenu K
MS Software Engineering,
School of Information Technology and Engineering,
VIT University, Vellore, India.
seenu.k2013@vit.ac.in

Vidhyagaran G
MS Software Engineering,
School of Information Technology and Engineering,
VIT University, Vellore, India.
vidhyagaran.g2012@vit.ac.in

Chandrasegar Thirumalai, IEEE Member
Assistant Professor Senior,
School of Information Technology and Engineering,
VIT University, Vellore, India.
chandru01@gmail.com

Abstract— **Software Complexity influences inward connections. Higher the multifaceted nature, bigger the deformities. Programming complexity for any product or a program is hard to discover without utilizing any measurements. The unpredictability, time and exertion fluctuate starting with one program then onto the next. For this reason, Halstead measurements are presented which recognizes the product complexity of a program by utilizing source line with the assistance of operands and operators. This metric was produced by Maurice Halstead to decide a quantitative measure of complexity specifically from the operands and operators in the module. This article gives a correlation between a program that was composed in two unique dialects to distinguish unpredictability, time and exertion of both and furthermore to gauge which programming dialect was better in wording less time to execute, least exertion.**

Keywords-component; Software Complexity, Halstead Metrics.

I. INTRODUCTION

The primary objective of any item is to fulfill client with high quality. Innovation is expanding each day. Many numbers of programming are being produced by numerous product organizations. Each of those organizations is some way or another attempting to stay aware of whatever is left of the organizations as their rivals and to rest of the world in giving top notch programming. The nature of programming is accomplished when the clients or the partner's needs are fulfilled. Maurice Halstead acquainted Halstead measurements with measure programming complexity which would be difficult to quantify without utilizing any measurements. Halstead measurements check the program in recognizing testing time, mistakes, multifaceted nature and how much exertion it takes to execute a similar program in various dialects et cetera with the assistance of operators and operands.

II. HALSTEAD METRICS

Complexity must be characterized for the whole program. One can't state that a program with minimum lines of code has less multifaceted nature than a program with more lines of code. Indeed, even a little program can be more mind boggling and expands the hazard while a major program can be basic. Multifaceted nature can't be resolved by simply considering the span of the program in light of the fact that a program can be made short by composing numerous source code in a solitary line. Complexity can be measured by examining the operators, operands and so on of each and every line. Consider these two illustrations "k=5" and "k= (a-b*ab)/control (a, b)" these two can be composed of a solitary line yet the main case is straightforward and the second case is more intricate than the first. This is sufficiently basic to demonstrate that any program can be intricate and unpredictability must be characterized by operators and operands in each and every line of code.

Halstead measurements for measuring complexity were presented by Maurice Howard Halstead in 1977. He watched that measurements mirror the usage or articulation of calculations in various dialects in light of the fact that there is a distinction in many-sided quality, time and exertion when it is composed in many programming dialects. Halstead measurements are utilized as a part of a few business apparatuses that tally programming lines of code. Many creators have received Halstead measurements either as a premise or as a center with minor changes in growing new philosophy. One of those fruitful accomplishments is Albrecht in 1984 that created work focuses technique which depends on Halstead's volume measurements. The source code and the calculation of that source code to which the Halstead measurements are connected are the elements. Operators and operands are the credits that are to be numbered. This record is used as the feeling duty in regards to consider the straight slip alongside wrongdoing [21], [22] and Pearson [5], [9], [11], [17], [19], [44]. In the present days, there are goliath measures of data recorded from the patients and a while later taking a gander at them requires complex estimations. We played out the thing metric examination on the given educational gathering [42], [48], [50].

From the data examination [4], [7], [12], [13], [15], [54] we can pick which quality can be considered and which

property can be disregarded. For example, in the Pearson structure if the estimation of r is more than 0.5 then the assigns are thought to be astoundingly related and if it is underneath 0.3 the qualities are barrenly related.

A touch of the past procedures to figure the decisions in setting of their relationship of basic worth is Spearman [6], [45] Analytical Hierarchical Process (AHP) [8], [10] and Traveling Salesman Problem (TSP) [33]. The test information's among various substances [14], [16], [23], [25], [27], [29], [31], [37] among the patient's and genius joint effort and decision demonstrate [40] are directed by late secured structures [18], [20], [24], [26], [28], [30], [32], [34], [36], [39], [41]. To have trust and shock guaranteeing late fundamental structures [43], [46], [47], [49], [51], [53] are existed. A liberal piece of the lightweight frameworks are passes on less fundamental of rigging with its optimal programming. To settle on a sharp choice on this light structure, many machine learning based strategy are associated in a couple of vocations.

Halstead's measurements rely on program execution and its measures, which are analyzed specifically from the operators and operands from source code. Halstead measurements permit assessing testing time, vocabulary, estimate, trouble, mistakes, and endeavors for the source code of any programming dialect. Halstead measurements were intended to think each program as a grouping of operators with its related operands. The targets of Halstead measurements are to gauge certain qualities, for example, vocabulary, volume, level, trouble, programming exertion and required programming time. As per Halstead, "A PC program is an execution of a calculation thought to be an accumulation of tokens which can be named either operators or operands". Dissecting the program line by line with a specific end goal to distinguish what are the operators and operands and what number of operators and operands are accessible. The measures of Halstead are based on

n_1 = number of unique or distinct operators.

n_2 = number of unique or distinct operands.

N_1 = total number of occurrences of operators.

N_2 = total number of occurrences of operands.

Halstead measurements think a program as an arrangement of operators and their related operands. Aside from these amounts there are others as well

n_1* = number of potential operators

n_2* = number of potential operands

Without operators and operands, a program can't be a program. Fundamental operators and operands must be characterized for each program keeping in mind the end goal to continue. These operators and operands characterize the whole procedure of the program. Break down the whole program; distinguish every one of the operators and operands accessible and the events of these operators and operands. These qualities are taken and are connected in the accompanying equation for length, vocabulary, volume, trouble, mistake, exertion, time and so on.

Halsted Metrics are as follows:

Measure N (Length of the program):

Is the sum of the total number of the operators and operands in the program?

$$N = N_1 + N_2$$

Measure n (Vocabulary of the program):

Is the sum of the number of unique operators and operands?

$$n = n_1 + n_2$$

Measure V (Volume of the program):

Is for the size of any implementation of any algorithm?

$$V = (N_1 + N_2) \log_2 (n_1 + n_2)$$

$$V = N \log_2 (n)$$

Measure D (Difficulty of the program):

Is proportional to a number of unique operators and total usage of operands.

$$D = (2/n_1) * (n_2/N_2)$$

Measure V* (Potential or minimal Volume V*):

$$V* = (2 + n_2) \log_2 (2 + n_2)$$

Measure L (Implementation Level):

$$L = V*/V$$

Measure L' (Program Level Estimator L'):

$$L' = (2/n_1) * (n_2/N_2)$$

Measure I (Intelligent Content):

$$I = L'V$$

Measure E - Effort required implementing or understanding the program is directly proportional to difficulty and volume.

$$E = D*V$$

Measure B - Number of bugs expected in the program:

Is proportional to the effort?

$$B = E0.667/3000$$

Measure T –Estimated time is taken to write the program:

Is proportional to the effort? $T = E/S$ [S=18]

III.   SAMPLE CODE

In this segment, we have taken two example programs in various languages to assess its complexity through Halstead Metrics and a similar specimen codes are demonstrated as follows.

(a) Sample Code 1 in C++:

```cpp
int main()
{
 int n, num, sum = 0,digit;
 cout << "Enter a positive  integer: ";
 cin >> n;
 num = n;
 while(num != 0)
 {
    digit = num % 10;
    sum += digit * digit * digit;
    num /= 10;
 }
 if(sum == n)
  cout << n << " is an Armstrong number.";
 else
  cout << n << " is not an Armstrong number.";
 return 0;
}
```

(b) Sample Code 2 in Python:

```python
num = int(input("Enter a number: "))
sum = 0
temp = num
while temp > 0:
  digit = temp % 10
  sum += digit ** 3
  temp /= 10
if num == sum:
  print(num,"is an Armstrong number")
else
  print(num,"is not an Armstrong number")
```

IV.   NUMERICAL ANALYSIS

TABLE I.   HALSTEAD METRICS ANALYSIS ON CODE 1 VS CODE 2.

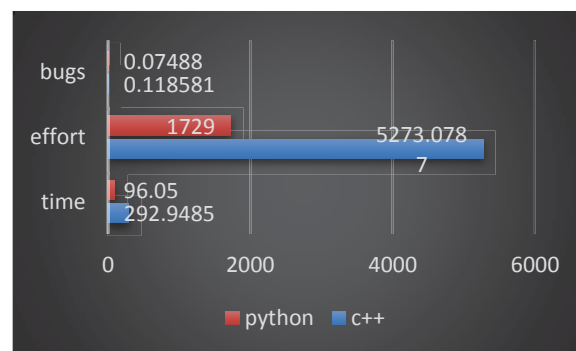| Halsted metrics	Code 1	Code 2
N_1	50	30
N_2	24	19
n_1	22	17
n_2	6	7
N	74	49
n	28	24
V	355.744	224.663
V*	24	28.529
D	0.022727	7.8721
L	0.06746	0.127
L'	0.02272	0.04334
I	8.085	9.7377
Effort E	5273.075	1729
Bugs Expected B	0.118581	0.07488
Time T	292.9485	96.05

Figure 1.   Comparison Chart of Time, Effort and Bugs.

V.   CONCLUSION

With numerous framework interfaces and complex necessities, the software complexity of programming frameworks some of the time develops out of hand which prompts increment in support cost and hazard in the upgrade. Halstead measurements require operators and operands which can be found in each program. It is basic, simple to actualize and doesn't require inside and out the investigation of the structure of the program. General quality can be enhanced by recognizing the software complexity of the product, presentation to hazard, time, exertion, estimate and so on. Indeed, even the program was given in various dialects at various circumstances Halstead measurements will have the capacity to ascertain the specified amounts and gives a proof of those lines of code and the extent of the program has very little to do with unpredictability. Halstead measurements have an inseparable tie to the code, can't be utilized at the plan level. It is exceptionally valuable in planning and announcing ventures and can be utilized for any programming dialect. As of late an extensive variety of ventures backings the utilization of Halstead in anticipating programming exertion and programming bugs.

REFERENCES

[1] Hauke J., Kossowski T., Comparison of values of Pearson's and Spearman's correlation coefficient on the same sets of data. Quaestiones Geographicae 30(2), Bogucki Wydawnictwo Naukowe, Poznań 2011, pp. 87–93, 3 figs, 1 table. DOI 10.2478/v10117-011-0021-1, ISBN 978-83-62662-62-3, ISSN 0137-477X.

[2] Piovani J.I., 2008. The historical construction of correlation as a conceptual and operative instrument for empirical research. Quality & Quantity 42: 757–777.

[3] Rashad Manzoor, Chandrasegar Thirumalai, "Investigating the Breast Cancer Tissue utilizing Semi-Supervised Learning and Similarity Measure," IEEE ICECA, 2017.

[4] Halstead Metric for Intelligence, Effort, Time predictions, DOI:10.13140/RG.2.2.17988.42881

[5] Fisher R.A., 1921. On the "probable error" of a coefficient of correlation deduced from a small sample. Metron 1: 3–32.

[6] Spearman C.E, 1904b. General intelligence objectively determined and measured. American Journal of Psychology 15: 201–293.

[7] Software metric Numerical Data analysis using Box plot and control chart methods, VIT University, DOI:10.13140/RG.2.2.27422.95041

[8] Vaishnavi B, Karthikeyan J, Kiran Yarrakula, Chandrasegar Thirumalai, "An Assessment Framework for Precipitation Decision Making Using AHP", International Conference on Electronics and

Communication Systems (ICECS), IEEE & 978-1-4673-7832-1, Feb. 2016

[9] Griffith D.A., 2003. Spatial autocorrelation and spatial filtering. Springer, Berlin.

[10] Chandrasegar Thirumalai, Senthilkumar M, "An Assessment Framework of Intuitionistic Fuzzy Network for C2B Decision Making", International Conference on Electronics and Communication Systems (ICECS), IEEE & 978-1-4673-7832-1, Feb. 2016

[11] Rodgers J.L. & Nicewander W.A., 1988. Thirteen ways to look at the correlation coefficient. The American Statistician 42 (1): 59–66.

[12] F. Fioravanti, P. Nesi, "A method and tool for assessing object-oriented projects and metrics management," Journal of Systems and Software, Volume 53, Issue 2, 31 August 2000, Pages 111-136

[13] Galton F., 1875. Statistics by intercomparison. Philosophical Magazine 49: 33–46

[14] Chandrasegar Thirumalai, Viswanathan P, "Diophantine based Asymmetric Cryptomata for Cloud Confidentiality and Blind Signature applications," JISA, Elsevier, 2017.

[15] Galton F., 1877. Typical laws of heredity. Proceedings of the Royal Institution 8: 282–301.

[16] Chandrasegar Thirumalai, Sathish Shanmugam, "Multi-key distribution scheme using Diophantine form for secure IoT communications," IEEE IPACT 2017.

[17] Galton F., 1888. Co-relations and their measurement, chiefly from anthropometric data. Proceedings of the Royal Society of London 45: 135–145.

[18] Chandrasegar Thirumalai, Senthilkumar M, "Spanning Tree approach for Error Detection and Correction," IJPT, Vol. 8, Issue No. 4, Dec-2016, pp. 5009-5020.

[19] Galton F., 1890. Kinship and correlation. North American Review 150: 419–431.

[20] Chandrasegar Thirumalai, Senthilkumar M, "Secured E-Mail System using Base 128 Encoding Scheme," International journal of pharmacy and technology, Vol. 8 Issue 4, Dec. 2016, pp. 21797-21806.

[21] Yule G.U., 1897a. On the significance of Bravais' formulae for regression, in the case of skew correlation. Proceedings of the Royal Society of London Ser. A 60: 477–489

[22] Chandramowliswaran N, Srinivasan.S, and Chandra Segar.T, "A Note on Linear based Set Associative Cache address System" International J. of Computer Science and Engg. (IJCSE) & India, Engineering Journals & 0975-3397, Vol. 4 No. 08 / pp. 1383-1386 / Aug. 2012.

[23] T Chandra Segar, R Vijayaragavan, "Pell's RSA key generation and its security analysis," in Computing, Communications and Networking Technologies (ICCCNT) 2013, pp. 1-5

[24] Chandrasegar Thirumalai, Senthilkumar M, Vaishnavi B, "Physicians Medicament using Linear Public Key Crypto System," in International conference on Electrical, Electronics, and Optimization Techniques, ICEEOT, IEEE & 978-1-4673-9939-5, March 2016.

[25] Chandrasegar Thirumalai, "Physicians Drug encoding system using an Efficient and Secured Linear Public Key Cryptosystem (ESLPKC)," International journal of pharmacy and technology, Vol. 8 Issue 3, Sep. 2016, pp. 16296-16303

[26] E Malathy, Chandra Segar Thirumalai, "Review on non-linear set associative cache design," IJPT, Dec-2016, Vol. 8, Issue No.4, pp. 5320-5330

[27] Hoadley, Katherine A., et al. "Multiplatform analysis of 12 cancer types reveals molecular classification within and across tissues of origin." Cell 158.4 (2014): 929-944.

[28] Chandrasegar Thirumalai, Senthilkumar M, Silambarasan R, Carlos Becker Westphall, "Analyzing the strength of Pell's RSA," IJPT, Vol. 8 Issue 4, Dec. 2016, pp. 21869-21874.

[29] Chandramowliswaran N, Srinivasan.S, and Chandra Segar T, "A Novel scheme for Secured Associative Mapping" The International J. of Computer Science and Applications (TIJCSA) & India, TIJCSA Publishers & 2278-1080, Vol. 1, No 5 / pp. 1-7 / July 2012

[30] Chandrasegar Thirumalai, "Review on the memory efficient RSA variants," International Journal of Pharmacy and Technology, Vol. 8 Issue 4, Dec. 2016, pp. 4907-4916.

[31] Vinothini S, Chandra Segar Thirumalai, Vijayaragavan R, Senthil Kumar M, "A Cubic based Set Associative Cache encoded mapping," International Research Journal of Engineering and Technology (IRJET), Volume: 02 Issue: 02 May -2015

[32] Chandrasegar Thirumalai, Himanshu Kar, "Memory Efficient Multi Key (MEMK) generation scheme for secure transportation of sensitive data over Cloud and IoT devices," IEEE IPACT 2017.

[33] M.Senthilkumar, T.Chandrasegar, M.K. Nallakaruppan, S.Prasanna, "A Modified and Efficient Genetic Algorithm to Address a Travelling Salesman Problem," in International Journal of Applied Engineering Research, Vol. 9 No. 10, 2014, pp. 1279-1288

[34] Nallakaruppan, M.K., Senthil Kumar, M., Chandrasegar, T., Suraj, K.A., Magesh, G., "Accident avoidance in railway tracks using Adhoc wireless networks," 2014, IJAER, 9 (21), pp. 9551-9556.

[35] Kahlert, Christoph, et al. "Prognostic impact of a compartment-specific angiogenic marker profile in patients with pancreatic cancer." Oncotarget 5.24 (2014): 12978.

[36] Chandrasegar Thirumalai, Anudeep Duba, Rajasekhar Reddy, "Decision Making System using Machine Learning and Pearson for Heart Attack," IEEE ICECA 2017.

[37] Layton, Robert, and Paul A. Watters. "A methodology for estimating the tangible cost of data breaches." Journal of Information Security and Applications 19.6 (2014): 321-330.

[38] P Viswanathan, P Venkata Krishna, "A Joint FED Watermarking System Using Spatial Fusion for Verifying the Security Issues of Teleradiology," IEEE Journal of Biomedical and Health Informatics, Vol.8, Issue 3, pp. 753-764

[39] Chandrasegar Thirumalai, Vignesh M, Balaji R, "Data analysis using Box and Whisker plot for Lung Cancer," IEEE IPACT, 2017.

[40] Kalaiarassan G, Krishan, Somanadh M, Chandrasegar Thirumalai, Senthilkumar M, "One-Dimension Force Balance System for Hypersonic Vehicle an experimental and Fuzzy Prediction Approach," Elsevier, ICMMM - 2017.

[41] Sonchack, John, Adam J. Aviv, and Jonathan M. Smith. "Cross-domain collaboration for improved IDS rule set selection." Journal of Information Security and Applications 24 (2015): 25-40.

[42] Chandrasegar Thirumalai, Abhiram Reddy P, Jaswanth Kishore Y, "Evaluating Software Metrics of Gaming Applications using Code Counter Tool for C and C++ (CCCC)," IEEE International Conference on Electronics, Communication and Aerospace Technology, 2017.

[43] Amintoosi, Haleh, Salil S. Kanhere, and Mohammad Allahbakhsh. "Trust-based privacy-aware participant selection in social participatory sensing." Journal of Information Security and Applications 20 (2015): 11-25.

[44] Chandrasegar Thirumalai, Aarthi B, Abhinaya V, "Analyzing the Forest Fire using Correlation Methods." IEEE International Conference on Electronics, Communication and Aerospace Technology, 2017.

[45] Chandrasegar Thirumalai, Vaishnavi M, Swapna Anupriya Chandhini, "Analysing the Concrete Compressive Strength using Pearson and Spearman," IEEE International Conference on Electronics, Communication and Aerospace Technology, 2017.

[46] Nguyen, Thi Hanh, Yi Wang, and Renfa Li. "An improved ridge features extraction algorithm for distorted fingerprints matching." Journal of Information Security and Applications 18.4 (2013): 206-214.

[47] P Viswanathan, "Fusion of cryptographic watermarking medical image system with reversible property," in Computer Networks and Intelligent Computing 2011 pp. 533-540.

[48] Chandrasegar Thirumalai, Shridharshan R R, Ranjith Reynold L, "An Assessment of Halstead and COCOMO Model for Effort Estimation," International Conference on Innovations in Power and Advanced Computing Technologies, 2017.

[49] P Viswanathan, P Venkata Krishna, "Fingerprint enhancement and compression method using Morlet wavelet," in Inderscience Publishers 2010, Vol. 3, Issue 4, pp. 261-268