ISSN: 2277-3754
ISO 9001:2008 Certified
International Journal of Engineering and Innovative Technology (IJEIT)
Volume 4, Issue 4, October 2014

On How Matrix Multiplication with Two Dense Matrices of Order n Achieves Empirical O (n^2) Complexity

Brajnandan Kumar[1], Suman Kumar Sourabh[2], Basant Kumar Das[3] & Dipo Mahto[4]
[1,3]University Department of Statistics & Computer Applications, T.M.B.U. Bhagalpur
[2]University Department of Computer Applications, T.M.B.U. Bhagalpur
[4]Department of Physics, Marwari College Bhagalpur, T.M.B.U. Bhagalpur

Abstract*—Average complexity analysis forms an interesting and intriguing part of algorithm theory since it explains why some algorithm such as matrix multiplication with bad worst-case complexity can better themselves in performance on the average. Statistical bound (asymptotic) and their empirical estimate over a finite range, the so called empirical O, makes average complexity more meaningful. The present paper shows that on how matrix multiplication with two dense matrices of order n achieves empirical O(n^2) complexity with a case study on Amir Schoor's Algorithm.

Keywords: **Amir Schoor's algorithm, sparse matrices, dense matrices, average case complexity**

I. INTRODUCTION

This is an aggressive research paper based on the philosophy: "As far as the laws of mathematics are certain, they do not refer to reality. And as far as they refer to reality, they are not certain" – Albert Einstein.

Average complexity analysis forms an interesting and intriguing part of algorithm theory since it explains why some algorithm such as matrix multiplication with bad worst-case complexity can better themselves in performance on the average. Statistical bound (asymptotic) and their empirical estimate over a finite range, the so called empirical *O*, makes average complexity more meaningful.

The present paper shows that Empirical $O(n^2)$ Complexity is convincingly gettable with two dense matrices in *nxn* matrix multiplication with a case study on Amir Schoor's Algorithm. We already know that for square matrices of order *n*, the average case complexity of Amir Schoor's algorithm is $O(d1d2n^3)$ where *d*1 and *d*2 are the densities(fraction of non-zero elements) of the pre factor and the post factor matrices respectively. For a formal proof, see ref [1]. If the product of these densities is kept at 1/*n*, we automatically have an $O(n^2)$ complexity trivially. For example, we may keep the pre factor matrix a sparse matrix with density 1/*n* and the post factor matrix fully dense with density unity. Our aim is to get similar complexity under more robust input conditions. Therefore, in the present paper, we keep the pre factor matrix approximately as dense as triangular [see ref(2)] except that the zeroes are randomly allocated and the post factor matrix is fully dense. Using "smart statistics" we observe that fits to quadratic and cubic are equally good.

II. DISCUSSION

The present paper can be discussed into two sections:

1. Section I describe the Amir Schoor's algorithm.
2. Section II gives the observations followed by the statistical analysis.

Section - I

Amir Schoor's Algorithm

Let A, B, and C be pre-factor, post-factor and product matrices respectively. Amir Schoor's algorithm states that for every non-zero *a*(*i* , *k*), multiply the k^{th} row of B by *a*(*i* , *k*) and add it to the i^{th} row of C. See also ref.[1]The pseudo code for the computational version only is as follows:-

```
for i = 1 to n
        for k = 1 to n
                while ( a(i , k) <> 0)
                        r = a(i , k)
                        for j = 1 to n
                                b(k , j) = b(k , j) * r
                        end for
                        for j = 1 to n
                                c(i , j) = c(i , j) + b(k , j)
                        end for
                end while
        end for
end for
```

We have built the product matrix with all zero entries before starting the algorithm. Also, we would be using the code for dense matrices only rather than sparse and hence Schoor's original data structure (the "row-column-value" structure) normally used for sparse matrices need not be adhered to. Recall that a triangular matrix is dense (see ref.[2]) with density (*n*+1)/(2*n*). Since the borderline between sparse and dense matrices is not well defined, we would agree to call the pre-factor matrix dense in which the fraction of zeroes is approximately equal to that in a triangular matrix.

Section-II

The observations on time computation (where time consumed on computation can be considered as its response) of Amir Schoor's Algorithm for multiplication of dense

ISSN: 2277-3754
ISO 9001:2008 Certified
International Journal of Engineering and Innovative Technology (IJEIT)
Volume 4, Issue 4, October 2014

matrices are given in Table-1(run times in seconds) for $d1= (n+1)/(2n)$ and $d2= 1$.

III. Table 1

n Trial	1000	1250	1500	1750	2000	2250	2500	2750	3000
1	0.030	0.060	0.070	0.110	0.140	0.170	0.210	0.250	0.270
2	0.030	0.050	0.080	0.111	0.130	0.180	0.220	0.260	0.270
3	0.040	0.060	0.070	0.100	0.140	0.190	0.220	0.291	0.300
4	0.030	0.060	0.080	0.100	0.160	0.180	0.280	0.321	0.291
5	0.040	0.050	0.080	0.120	0.181	0.231	0.231	0.310	0.280
6	0.040	0.050	0.081	0.120	0.160	0.210	0.260	0.320	0.300
7	0.030	0.070	0.070	0.110	0.151	0.211	0.260	0.291	0.291
8	0.030	0.050	0.080	0.140	0.171	0.220	0.260	0.260	0.271
9	0.040	0.060	0.091	0.121	0.150	0.220	0.270	0.260	0.291
10	0.040	0.071	0.110	0.110	0.160	0.240	0.220	0.260	0.300

The following plots show the regression analysis on linear, quadratic and cubic fit.

IVA.Graph1

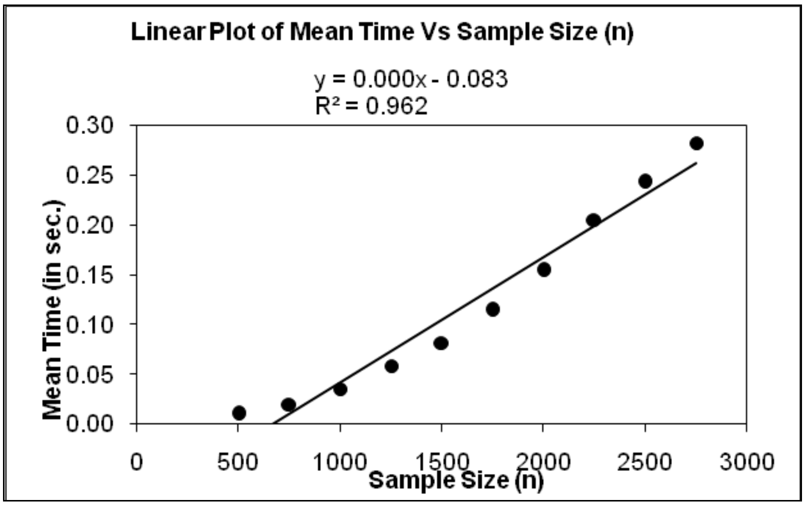

Fig. 1

IV B. Graph 2

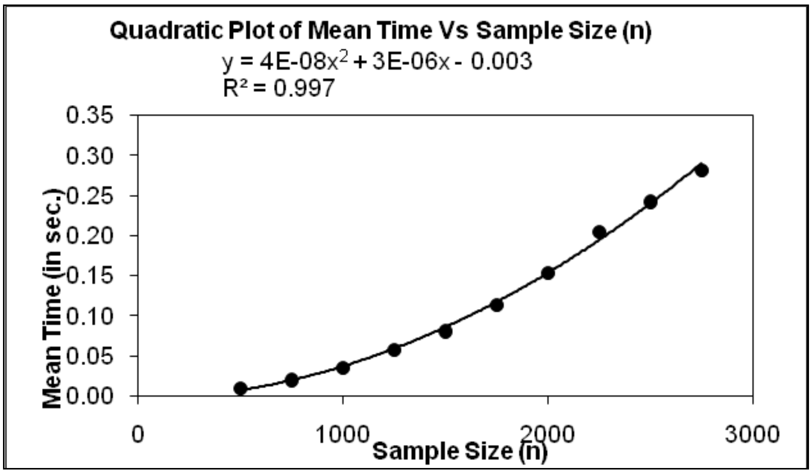

Fig. 2

ISSN: 2277-3754
ISO 9001:2008 Certified
International Journal of Engineering and Innovative Technology (IJEIT)
Volume 4, Issue 4, October 2014

IVC. Graph 3

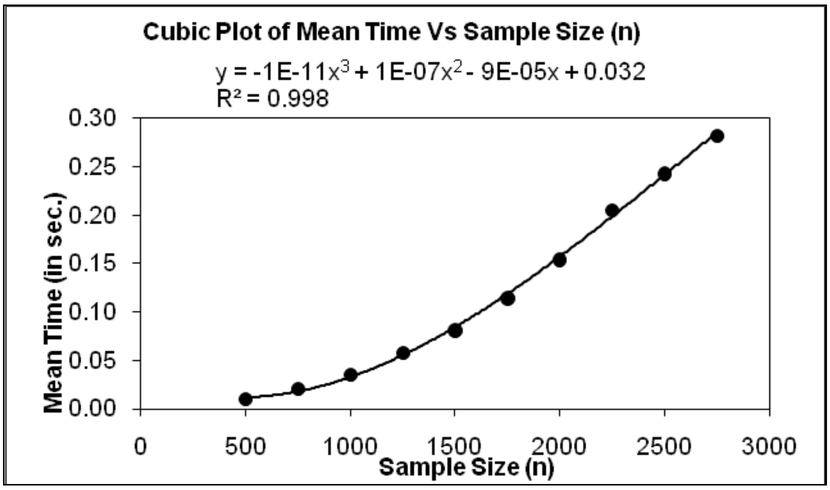

Fig. 3

V. CONCLUSIONS AND SUGGESTIONS FOR FUTURE WORK

It is easy to see that an empirical $O(n^2)$ complexity is certainly gettable with both matrices dense, the pre-factor matrix approximately as dense as triangular and the post factor matrix fully dense. Future work includes studying whether the same can be had if the density of the pre-factor matrix is increased even further (this increases robustness of input conditions), the post factor matrix remaining fully dense. Perhaps the quadratic fit will have to be "Kussmaul-protected" (use of D-optimality design is proposed). This means that *we would be fitting a quadratic fit with a data set where observations are taken at those points which are optimal points for a cubic fit*. The quadratic fit is then said to be *"Kussmaul-protected*" against the error for not fitting a cubic fit (see ref. [8]). The protection is suggested because we already know that when both matrices are fully dense, the complexity becomes $O(n^3)$ theoretically. It follows therefore that, with one density full and the other approaching fullness; *there will be a certain theoretically proven drift towards a cubic pattern*. As such, any adventure with a quadratic fit demands some kind of protection. Note further that to get an empirical $O(n^2)$ complexity for any two matrices convincingly, Coppersmith-Winograd's algorithm should be used. But as this algorithm is not very practicable, we used Winograd's algorithm instead for two arbitrary nxns matrices in ref [5] which shows how close we got!

To the question why we are able to work with a quadratic fit without sacrificing predictive power in the present case, two arguments come to our minds. First, Schoor has clearly claimed that his algorithm is not only fast for sparse matrices but also fast for dense matrices because it is faster to work with rows only than both rows and columns (see p. 89 of ref [1]). But this interesting claim demanded an empirical verification and hence our study. Second, when we say $T_{avg}(n) = O(f(n))$ we merely mean there exist two positive constants c and N such that $T_{avg}(n) < cf(n)$ whenever $n > N$. It is to be remembered that the constant c, which depends on implementation, can get very close to zero. If that happens, it is quite possible to work with an empirical complexity lower than the theoretical counterpart without sacrificing predictive power. The concept can be useful in some situations like chain matrix multiplication with a large number of matrices in the chain (see ref [2]), preventing ill-conditioning brought about by unnecessary complicated modeling in a *computer experiment* in general (ref [3]) and in tracking down empirically the complexity of an arbitrary algorithm for which determining a precise upper bound is a "deep intellectual challenge" (see ref [6] and [4]).

REFERENCES

[1] A. Schoor, Information Processing Letters, vol. 15, pp. 87-89, 1982.

[2] Sartaj Sahni, "Data Structure and Algorithms in C++", Tata McGrawHill, 2000.

[3] Jerome Sacks et. al., "Design and Analysis of Computer Experiments", Statistical Science, vol. 4, no. 4, 1989.

[4] Soubhik Chakraborty et. al., "On How Statistics Can Provide a Reliable and Valid Measure for an Algorithm's Complexity", InterStat Statistics Journal, Dec 2004#2.

[5] Soubhik Chakraborty and Kiran Kumar Sundararajan, "A Statistical Adventure towards Getting an Empirical O (n^2) Complexity in nxn Matrix Multiplication", InterStat Statistics Journal, March 2006#5.

[6] Aho, Hopcroft, Ullman, "Data Structures and Algorithms", Pearson Education Reprint, 2000.

[7] Soubhik Chakraborty and Suman Kumar Sourabh, "On why an algorithmic time complexity measure can be system invariant rather than system independent", Applied Mathematics and Computation (AMC), Volume 190, Issue 1, Pages 195-204, 2007.

[8] K. Kussmaul, Protection against assuming the wrong degree in polynomial regression, Technometrics, vol.11, 1966.

ISSN: 2277-3754
ISO 9001:2008 Certified
International Journal of Engineering and Innovative Technology (IJEIT)
Volume 4, Issue 4, October 2014

AUTHOR BIOGRAPHY

Brajnandan Kumar (1987): He is a research scholar in the University Department of Statistics & Computer Application, T.M.B.U. Bhagalpur under the supervision of third author. He has passed M. Sc. in 2010 from T.M.B.U. Bhagalpur, Bihar, India. He is Gold Medalist in M.Sc. His research work is associated with the retrieval of structured text from XML database, analysis of query and crawl based plan in terms of both execution time and output completeness and Query Optimization for Text Databases.

Dr. Suman Kumar Sourabh (1982): He is a guest lecturer in the the University Department of Computer Application, T.M.B.U. Bhagalpur India. He has passed M. Sc. in 2004 from T.M.B.U. Bhagalpur, **MBA** in 2008 from Sikkim Manipal University, Manipal and **MCA** in 2009 from Punjab Technical University, Jalandhar Sikkim Manipal University, Manipal and awarded by PhD degree in 2009 from T.M.B.U. Bhagalpur University.His research work is associated with Complexity Theory, especially average case complexity measure and Stochastic modeling of deterministic computer algorithm with Algorithm Design and Analysis parameterized complexity analysis. There are thirteen papers published in repute international journals.

Dr. Basant Kumar Das (1955): He is associate Professor in the University Department of Statistics and computer application, T.M.B.U. Bhagalpur. He has passed M.Sc. in 1979 from T. M. B. U. Bhagalpur, India and awarded by PhD degree in 1996 from the same University. There are eight research papers published in repute national and international Journals. He was former H.O.D. of Statistics and computer application, T.M.B.U. Bhagalpur from 2007 to 2013.

Dr. Dipo Mahto (1965): He is assistant Professor in the Department of Physics, Marwari College, Bhagalpur under T.M.B.U. Bhagalpur from 2003. He has passed M.Sc. in 1990 and awarded by PhD degree in 2005 on the topic-Contribution of Mathematical Physics in the development of theory of relativity with special reference to rotating media under the supervision of Dr. Gopi Kant Jha, Former Head & Professor, University Department of Physics of L.N. Mithila University, Darbhanga, India. He is also reviewers of many International Journals. He is invited as resource person for special classes in Physics. His research work is associated with relativity and Astrophysics, especially black holes. There are forty nine research papers credit to his name published in repute national and international journals. He has been invited as speaker in Space meeting to be held in USA organized by OMICS Conference on 12-13 Aug 2014.