IS THERE A MOUS IN YOUR HOUSE?

Robyn Pascoe

Eastern Institute of Technology
Hawke's Bay, NZ
rpascoe@eit.ac.nz

ABSTRACT

No, not a small rodent, or a pointing device, a Microsoft Office User Specialist. MOUS certification is globally recognised as the skill standard for Microsoft's suite of business productivity applications.

In the United States, there has been an ever increasing push towards the use of Microsoft certification as a standard qualification. The MOUS certification program is being offered as an elective in dozens of U.S. colleges and universities, and many accounting firms and recruitment agencies in the US are using the core Word and Excel MOUS certification as a baseline when employing new staff. In Britain, Microsoft and City & Guilds announced in June 2000 that they were collaborating in a venture designed to improve training standards in the information technology industry.

New Zealand has lagged somewhat behind this trend but this will surely change if, as seems likely, NZQA announces that it has registered MOUS certification on the national framework. This will allow MOUS certification to be credited to Unit Standards and NCEA.

This paper explores some of the ramifications in the N.Z. tertiary sector and documents the results of a survey of student knowledge of, and attitude toward MOUS certification.

KEYWORDS

computer education, qualifications, student satisfaction, future employment.

1. INTRODUCTION

This paper explores by means of a questionnaire the level of knowledge regarding the Microsoft Office Specialist certification programme, within a group of current EIT Business and Computing faculty students. Additionally the questionnaire attempts to gauge the level of interest in attaining certification.

While we all indulge in a spot of Microsoft bashing from time to time, the reality is that the Office suite of applications is used by a majority of companies worldwide. In most cases the employees who use the applications do not even get close to the competence and productivity possible with these tools.

Gaining Office Specialist certification acknowledges the expertise to work with Microsoft Office programmes and provides a valid and reliable measure of technical proficiency and expertise by evaluating overall comprehension of Office applications, ability to use advanced features and integrate with other software applications.

Office Specialist exams are globally developed and validated by industry experts and are available at over 9,000 Authorised Testing Centres worldwide. Certified individuals report increased competence and productivity with Microsoft Office programs as well as increased credibility with their employers, co-workers, and clients.

To earn Microsoft Office Specialist certification the applicant must pass one or more certification exams. The exams come in three flavours, Core, Comprehensive and Expert.

The Master Office Specialist certification is achieved by passing core or comprehensive exams in Access, Outlook and PowerPoint, plus expert exams in Word and Excel.

The recent registration by NZQA of the Microsoft Office Specialist certification means that students are able to credit an Office Specialist certification toward the NCEA This move has implications for polytechnics in that New Zealand secondary schools are being encouraged to become testing centres and large numbers of high school students are sitting Microsoft Office Specialist exams. It is likely therefore that there will be students with higher levels of software skills entering tertiary programmes.

Note: Since writing the abstract for this paper I have discovered that Microsoft has dropped the word User, and the certification is to be henceforth known as Microsoft Office Specialist. Microsoft have also ceased to use the acronym MOUS to describe the certification programme, but as it's so difficult to come up with a snappy title I have decided for titular purposes to continue the use of MOUS in this paper's title, thus both the name Microsoft Office User Specialist and the acronym MOUS appear in the abstract and in quotes and excerpts from publications throughout this paper.

2. ACADEMIC ALLIANCES

Worldwide there have been many academic alliances with Microsoft's certification partner, Certiport, in the last five years, notably the British City & Guilds training establishments and several Colleges and Universities in the United States.

Microsoft and City & Guilds collaboration, designed to improve training standards in the information technology industry, has been developed to provide those taking part in the City & Guilds e-Quals series with the skills to sit Office User Specialist exams without the need for further training. (Hook 2002).

A growing number of universities in the US now require students to attain Microsoft certification to graduate. In April 2002 Certiport Inc., the exclusive worldwide administrator of the programme announced the adoption of Microsoft Office User Specialist certification throughout institutions of higher learning in the U.S. Pepperdine University and the University of Memphis are among dozens of colleges and universities leveraging office certification to prepare students for the technology-centred workplace. In recognition of the universal use of technology in the workplace and the fact that businesses want assurance that new graduates they hire are skilled in using these tools, both of these universities and many others now require Microsoft Office Specialist certification as part of their graduation requirements.

These universities collectively state that a unified standard of computer software proficiency:

- ♦ simplifies the teaching process
- ♦ enables professors to demand more of their students in assignments
- ♦ boosts student confidence and lowers student anxieties about using basic software applications
- ♦ signals to employers that graduates will come equipped with the computer skills needed to be immediately effective at work versus requiring expensive on-the-job training

The impetus for requiring certification first came from business lecturers who noticed that many students lacked the computer software skills needed to complete general business assignments, such as preparing PowerPoint slides to serve as visual aids to presentations, tracking a company's cash flow using an Excel spreadsheet or producing professional looking documents using Microsoft Word.

Dr. James Goodrich, associate dean of full-time graduate programmes at Pepperdine University who was instrumental in the decision to require certification, says that thanks to training courses, the university boasts a 95 percent exam pass rate and has certified more than 500 full-time MBA students.

The University of Memphis, requires undergraduate business students to enrol in a Microsoft Office course as part of the core curriculum. The students are required to take an exam in either Word or Excel, and the exam score is then averaged as a part of their grade for the class.

Kyle Root, training & certification coordinator for EDS Global Support Services, said he is glad to see these universities require certification. "Microsoft Office Specialist certification not only benefits students, but is also great for businesses like ours because it ensures graduates will be proficient in some of the most commonly used business computer applications," said Root. "By hiring students who are certified in computer software applications such as Microsoft Word, Excel or PowerPoint, our company reduces the total costs involved in training new employees by $300 per employee per exam." (PR Newswire 2002)

3. INDUSTRY VIEW OF OFFICE CERTIFICATION

In 1998 Elliott Masie, an internationally recognised speaker, author and consultant on the topics of technology, business, learning and workplace productivity authored a white paper based on research conducted at the Masie Centre, New York.

The Masie Centre's study of Microsoft's Office User Specialist programme highlighted an enormous demand for skill certification, and a workforce marketplace eager for global desktop performance standards.

"We believe that organisations need to become more specific about computer skills. Workers, targeting a rapidly changing workplace and marketplace, deserve a standard for documenting their skills and competencies. Computer training can use a much-needed refocus on performance and work." Masie (1998)

Its research and interview processes persuaded the Masie Centre that there was widespread demand, interest and support of the Microsoft Office User Specialist program.

These views are supported by accounting firms in the United States, as reported by L Gary Boomer, a featured writer with the New York publication Accounting Today, Boomer wrote in May 2000;

"A Microsoft Office User Specialist offers some great potential benefits to CPA firms. Although Word and Excel are essential work tools in most offices, training and evaluation criteria vary widely from firm to firm. People who go through the MOUS certification process are ready to get the most out of these tools and enhance their efficiency and their companies' profitability along the way. Perhaps the least known and least understood of all the Microsoft certifications, this one offers some great potential benefits to CPA firms."

Results of an independent research study of participants in the Microsoft Office User Specialist (MOUS) programme indicate MOUS certification improves employee competency, productivity and credibility.

♦ 87 percent of employers observe increased competency in their MOUS-certified employees

♦ 83 percent of employers feel their MOUS-certified employees are more productive

♦ 77 percent of employers feel MOUS certification has a positive effect on employee professional credibility among co-workers

♦ 74 percent of employers feel MOUS certification has a positive effect on employee professional credibility with customers and clients

♦ 82 percent of employers believe MOUS directly benefits an organisation

♦ 67 percent of employers feel MOUS simplifies hiring and advancement decisions

(Computer Press 2002)

4. WHAT ABOUT NEW ZEALAND?

Up until recently, the Office certification programme has languished at the bottom of the heap in New Zealand, being virtually ignored by most polytechnics and universities. It has been relegated to the realms of the private training provider, and even among these, only a select few have had exclusive rights as Microsoft Testing Centres. But this is scene is changing. Computer Press, the New Zealand partner in the Certiport/Microsoft certification alliance has been quietly introducing the concept to New Zealand secondary schools.

Some schools, such as Lytton High School in Gisborne have been very enthusiastic about the Microsoft programme and have achieved amazing results. Over 200 individual Microsoft Office specialist exams have been passed by Lytton students in the last few months, and a number have achieved Microsoft Master Office Specialist certification.

While not all high schools have embraced the concept so wholeheartedly, there are growing numbers of high school students leaving school with certification. This has implications for the tertiary sector in that there may well be numbers of students enrolling in certificate, degree and diploma programmes who already have higher skill levels in Microsoft Office applications than many existing modules/papers demand. This may in turn be followed by demand for more advanced training, particularly in the areas of Excel and Access.

5. WHERE DOES NZQA FIT?

Official information from NZQA has not been forthcoming to date, however the Microsoft, Certiport, Computer Press alliance has lobbied NZQA for some time regarding the addition of Microsoft Office Specialist certification to the New Zealand Register of Quality Assured Qualifications. The Register contains all qualifications in New Zealand which have been quality assured by various organisations including NZQA and Polytechnic, University and College of Education bodies.

Microsoft qualifications cannot be placed on the National Qualifications Framework to lead to Unit Standards or Achievement Standards, they can however be made to count towards the National Certificate of Educational Achievement (NCEA) by credit inclusion if they are registered. Besides Unit Standards and Achievement Standards, NCEA can include quality assured qualifications which are accepted for credit inclusion.

A Microsoft Office Specialist qualification will need to be mapped against Unit Standards or Achievement

Standards at the same level and with the same outcomes. For example, achievement of the Microsoft Office Specialist proprietary qualification for Excel could be mapped as being equivalent to a Unit Standard in using a spreadsheet.

6. STUDENT SURVEY

In a student survey (n = 122) conducted at the Eastern Institute of Technology Hawke's Bay in April 2003, students were asked firstly what they knew of Microsoft Office Specialist certification. Those who knew little or nothing about it were then supplied with a short explanation of the certification. Almost all the students who took part in the survey were full time business students, the majority being computing students so it was a little surprising that 68% had never heard of the certification. Of the remaining respondents 29% were familiar with the name, and only 3% professed to know more than a little about it.

Somewhat more surprising was the level of interest generated after the students read the explanatory leaflet accompanying the questionnaire. An overwhelming 84% expressed interest in gaining Office Specialist qualifications. They were then presented with a list of the exams that make up the Microsoft Master Office Specialist certification and asked to check those they would be most interested in sitting.

The table below illustrates the percentage of students expressing interest in each exam.

Access Core, 91%

Excel Core, 73%

Word Core, 70%

Word Expert, 58%

Excel Expert, 55%

PowerPoint Comprehensive, 52%

Outlook Core, 48%

6. CONCLUSION

While there has not been any official announcement from NZQA, it is apparent that many more secondary school pupils will enter tertiary institutions with Microsoft certifications, or at least a higher level of office application skills than previously. The implications are that many of these students will skip introductory level computing courses such as National level 2 and 3 computing and office systems certificates, leading to less demand for the lower level qualifications and more enrolments in higher level diploma and degree courses.

These students will have the expectation that their office certification could, and should be cross credited towards software packages type modules such as SP500 and SP510 and first year degree papers with similar learning outcomes. Consequently tertiary institutions may need to develop strategies for dealing with cross crediting issues and further upstream, a demand for higher level learning in these areas.

REFERENCES

PR Newswire (April 29, 2002) It Takes Technology Proficiency to Make the Grade; Growing Number of Universities Require Microsoft Certification to Graduate.

Hook, S. (2002) "Joint IT venture points way to higher skill levels" The Times Educational Supplement; London; Issue: 4484 p32.

Masie, E. (1998) "Creating The Next Generation Of Users: Skilled & Certified" White Paper, The Masie Centre, NY 12866

Boomer, G.L. (2000) "The kind of MOUS you'll actually want in your house". Accounting Today; New York; Volume14 Issue 8, pp24-28

Computer Press (2003) Certification Benefits. Accessed April 20th 2003. <http://www.microsoftofficespecialist.co.nz/benefits/commercial.htm>

Approximate Computation of Multidimensional Aggregates in a Data Warehouse

Russel Pears
School of Information Technology

Megan Mansfield
Dept. of Electrotechnology,

Auckland University of Technology
Auckland , NZ

rpears@aut.ac.nz

ABSTRACT

This paper presents a new technique for compression and recovering data in a Data Warehousing environment. Although the compression technique presented can be used for compressing data in any large-scale data storage medium, it is optimised for use in Data Warehouses where queries range against multiple dimensions.

Keywords:

Dimension, Data Warehouse, Aggregate Queries, Prime Numbers, Prime Index Tree, Prime Tree.

1. INTRODUCTION

Data Warehouses are increasingly being used by decision makers to discern trends in data. For example, a Marketing Analyst may be interested in identifying the top selling products over the last financial year. While data warehouses provide valuable information they are costly to build and to query. This paper investigates a technique for compressing warehouse data as a solution to overcoming poor query performance in large-scale data warehousing systems.

The standard technique for improving query performance is to build aggregate tables that are targeted at known queries [Rama 200], [Elm 2000], [Ora9i]. In the above case of the "top selling products" query a summary table can be created that contains the total sales value (in dollar terms) for each of the products and sort the result in decreasing value of sales. It would then be a simple matter of querying the summary table and retrieving the first ten rows. The main problem with this approach is the lack of flexibility. If the analyst now chooses to identify the bottom ten products an expensive sort would have to be performed to answer this new query. Worst still, if the information is to be tracked by sales area, then the summary table would be of no value at all.

Our approach avoids these issues by compressing the original data in the warehouse in such a way so as to preserve the structure of the data. Thus for example a 3-dimensional warehouse that tracks sales by product, time and sales area could be compressed along the product dimension that reduces the volume of data to be stored and queried to a fraction of the original size. This is achieved by a novel data compression algorithm based on prime numbers.

The paper presents an overview of the algorithm, reports on results of our experimentation on different types of data sets and indicates some avenues for further research.

2. THE PRIME FACTOR SCHEME

2.1 Pre-Processing

The input stream is first pre-processed by approximating each value by its nearest prime number. The rationale for using prime numbers is two-fold. Firstly,

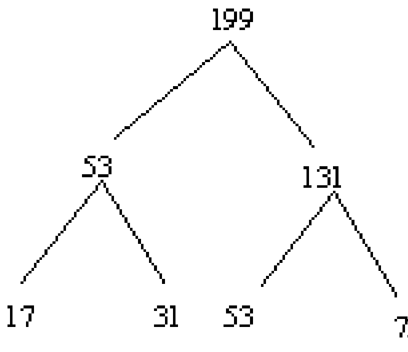

Figure 2.1 the Prime Index Tree

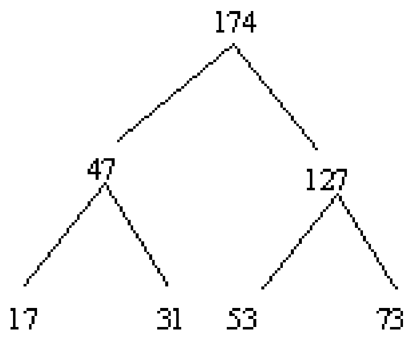

Figure 2.2 the Prime Tree

a prime number can be referenced in a space-effective manner by its index. For example the prime 71 happens to be the 20th prime number and can be referenced as such. Secondly, prime numbers become less dense towards the higher ends of the number scale – there are 24 primes for the first 100 positive integers, in contrast the next 900 integers yields only 142 primes. Taken together these two facts give us a storage scheme that scales very well up the number range.

Once the input stream has been approximated as above, we divide it into a number of equal-sized blocks. These blocks are then input to the algorithm described in sections 2.1 and 2.2. The algorithm works by using the primes in a block as leaf nodes to construct two binary trees. Once the trees are constructed, we retain only the two root nodes of the trees and a small array of pointers to aid in the reconstruction phase. The two trees are named the Prime Index Tree (PIT) and Prime Tree (PT).

2.2 Construction

The PIT is constructed by taking each pair of node values and adding their numerical values to a term that represents the difference in their index positions. This is done recursively until the root node is reached (see figure 2.1). The PT is constructed in a similar fashion except that at each stage we simply add the numerical values of the nodes and do not include the term that accounts for the index positions (see figure 2.2).

In order to reconstruct the leaf node values (i.e. the original pre-processed numbers) we store for each PIT node a pointer that references the children of that node. This is done for all levels except for the level just above the leaf level. This involves a total of N/2 pointers for a block size of N. We were able to avoid the storage overhead of a further N/2 pointers (that would otherwise be needed to navigate to the leaf level nodes) by using the PT to construct the leaf level nodes. The storage overhead for the PT is minimal since it only involves storing the root node value. All other (internal) nodes of the PT is constructed by using a scoring function that obtains a best fit with the corresponding parent node of the PIT; the details are in [Pears 2003].

Thus in general, a block of N numbers will be replaced by 2 numbers + N/2 pointers. For large values of N, this approximates to N/2 (a reduction factor of 2 over the original data set). It turns out that a pointer only requires a few bits of storage (a maximum of 5 over the data sets that we investigated) [Pears 2003]. This means that we were able to achieve a relatively high compression ratio of 8:1which compares well with standard compression techniques such as Huffman coding [Huff 52].

We shall illustrate the working of the algorithm with the help of an example. The algorithm consists of two phases: firstly the construction of the compressed version of the input, and secondly, the reconstruction of the original set of numbers from the compressed version.

Suppose that we have the primes 17,31, 53 and 73 in the input stream after the pre-processing step. We now add 17+31+(11-7) to give 52, which is approximated by its nearest prime which in turn gives 53 for the parent node on the left. Similarly we obtain 131 as the right parent. In the next step we repeat the same process on 53 and 131 to yield the root node value of 199.

The Prime Tree is constructed on the same block and yields a root value of 174. We have now replaced the values in our block by two values, the PIT and the PT root node values, and a single pointer associated with the root level node of the PIT.

2.3 Reconstruction

The reconstruction phase begins at the root node of the two trees. The two trees are reconstructed in parallel with each other. At each step we descend one level down the PIT to give its children. At the same time we apply the scoring function to the PT root node to yields its children. This is repeated for all levels until we descend to the level just above the leaf level. At this point we can calculate the pointer for each PIT node value at this level. This is done by subtracting the PT node value from its corresponding PIT node value. This enables us to navigate to the leaf level, at which point the reconstruction is complete.

The root value of the PIT with value 200 and the associated (stored) pointer are used to identify the children of the node. This yields the left and right children with values 53 and 131 respectively. We now apply the scoring function on the PT root value of 174 to give us the left and right children with values 47 and 131 respectively. We now calculate the pointer for the PIT node with value 53 as 53-47=6; similarly the pointer for the other PIT node gives a value of 131-127=4. These two pointers are then used to recover the original set of values.

2.4 Query Processing

In a Data Warehousing environment the above compression scheme can be used to answer range-sum queries very effectively. For the sake of simplicity let us assume that we have such a query Q (23:100) that requires us to calculate the sum across a range with starting position 23 and ending position 100. If we use a block size of 16, then all that is needed is to reconstruct only the 2^{nd} and 7^{th} blocks. After recovery of these two blocks the answer is simply the addition of the PT root node values for blocks 1, 3, 4, 5, and 6 together with the sum of the relevant leaf node values in blocks 2 and 7. This result generalises to any query Q(N:M), which means that any 1-dimensional query will need to recover only (a maximum) of two blocks corresponding to the blocks at the two ends of the query.

This has two important implications. Firstly, response time improves dramatically. Preliminary comparisons with other compression techniques such as the Wavelet algorithm [Vitt 98] indicate a significant improvement (see section 3). Secondly, since only two blocks at most are recovered the margin of error is small. Again, our experimentation so far has revealed that the error rate compares very favourably with that of the Wavelet technique.

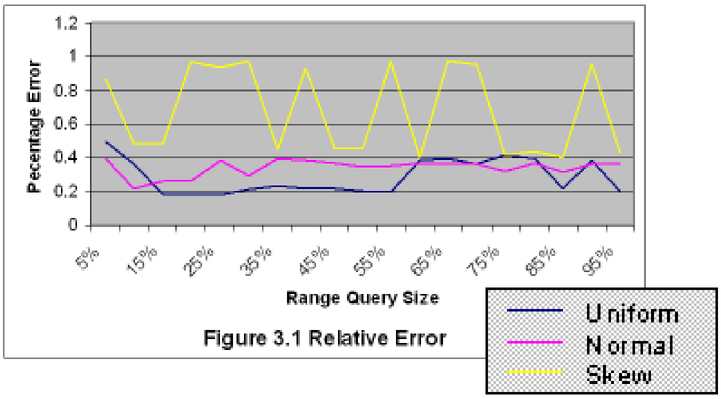

Figure 3.1 Relative Error

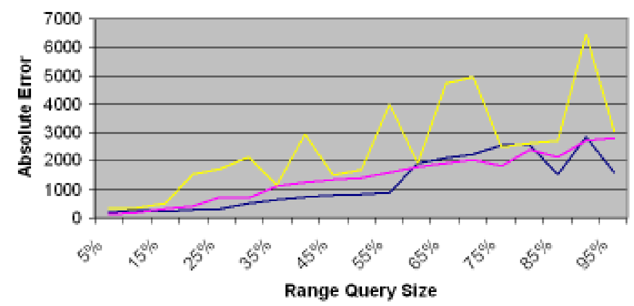

Figure 3.2 Absolute Error

3 EMPIRICAL RESULTS

In this section we report on some of the experiments that were performed on the Prime Factor algorithm. Our simulation data sets were sets of randomly generated values using three different types of distributions: Uniform, Normal and Skew distribution. Each experiment ran a range-sum query on the one of the data sets and the resulting average error rate (over 15 runs) was graphed against the size of the range sum query. All experiments were run with a compression ratio of 8:1. The error metrics we used were the standard measures of error, namely Relative Error and Absolute Error [Vitt 98].

Our experimentation revealed that the Prime Factor method performs well across a wide spectrum of range sum query sizes. This is apparent in Figure 3.1 where the Relative Error only deviates between 0.18% to 0.98% accuracy. Another characteristic of the Prime Factor method is the speed with which it processes the queries. When compared to the Wavelet method, the Prime Factor algorithm processed each query on average 8 times faster than the Wavelet method.

3.1 Normal Distribution

After analysis, the results revealed that the algorithm was most stable on the Normalised data. The Prime Factor method was found to mimic the normal distribution bell curve. The relative and absolute errors shown in Figures 3.1 and 3.2 are stable in the centre of the figures, which can be associated with the centre of a bell curve, while at the outer ends of the bell curve the method is less stable.

3.2 Uniform Distribution

Queries on the Uniformly distributed data set, as expected, exhibited lower error rates when compared with data from the Normal distribution The stability however was not as good as both Figure 3.1 and Figure 3.2 illustrate where deviations occur around the 55% and 80 to 95% query values.

3.3 Skew Distribution

The performance of algorithm on skewed data was less satisfactory when compared to the other two data distributions. Preliminary investigation discovered that most of the variation in error rates shown corresponded to queries which required reconstruction of almost all values on the two blocks at the two ends of the query.

4. CONCLUSIONS AND FUTURE RESEARCH

Our results so far indicate that the Prime Factor algorithm performs well across both Uniform and Normalised data distributions. The error rate on skewed data while being small at less than 1%) oscillates throughout the query range that we investigated. This is one aspect that we will be investigating further.

We are currently in the process of extending the algorithm to work for an arbitrary number of dimensions. On completion of this we will test its performance against the Wavelet algorithm on real-life data sets such as the US Census data [Bur].

REFERENCES

[Rama 2000], Ramkrishnan R, and Gehrke J, “Database Management Systems” (2nd edition), McGraw Hill, 2000

[Elm 2000], Elmasri R and Navathe S, “Fundamentals of Database Systems”, Addison Wesley, 2000.

[Ora9i], “Oracle Data Warehousing Guide”, Oracle Corporation, 2003.

[Pears 2003], Pears R and Mansfield M, “The Prime Factor Algorithm for Processing Range-Sum Queries in an OLAP Environment” in preparation.

[Huff 52] Huffman, D, *Proceedings of the Institute of Radio Engineers*, vol. 40, 1952.

[Vitt 98] Vitter, J and Wang M, “Data Cube Approximation and Histograms via Wavelets”, Proceedings of the Seventh International Conference on Information and Knowledge Management, Washington D.C, 1998.

[Bur] U.S. Census Bureau. Census bureau databases, online at http://www.census.gov/