IShare

Monthly Magazine

Enlight IT with Global Vision

K.S.Rangasamy College of Arts & Science (Autonomous), Tiruchengode

Technical Forum created by students for students

Inside this issue

- **Blue Gene**

- **Myself in Book Shelf**

- **Motion Capture**

- **Academic Forum**

Explore inside......

PATRON

Lion.Dr.K.S.Rangasamy, MJF

Founder & President

ADVISORS

Ms. Kavithaa Srinivashaan, *M.A.,M.B.A.,*
Executive Director

Dr. N. Kannan, *M.Sc., Ph.D.*
Principal

Prof.K.Natarajan, *M.Sc.,M.Phil.,*
Director,
Center for Information Technology

Mr. R. Pugazendi, *M.Sc.,M.Phil.,M.E.,*
HOD,
Department of Computer Science

EDITORIAL TEAM

Ms. S.Sasikala, *M.Sc., M.Phil., M.C.A.,PGDPM &IR.,*

Ms. S.Nithya., *M.Sc(CS & IT).,*

Ms. M.M.Kavitha., *M.Sc., M.Phil*

Ms. R.Priyanka.,*M.Sc. (CST),*

Ms. S.Ramya., *M.C.A.,*

Ms. M.Gomathi., *M.Sc. (IT),*

Ms. S.Shalini., *M.C.A.,*

& Office bearers of TRACE and ACAI

EDITORIAL …

This issue presents

- Top Seven most Dangerous Computer Viruses

- "Blue Gene"

- Events that occurred in January – This month that Day

- Future trends - Data Mining techniques and methods

- Many more useful and interesting information inside……

Explore it !!!

CONTENTS

GUIDE TO SEARCH ENGINE

Author: M.GOMATHI

Lecturer, Dept. of Computer Science

This article deals with tools and mechanisms about Search engines.

Search engines crawl the World Wide Web to gather information about websites and their content. This is usually done through robots or crawlers or bots, complex mechanisms that can roam the internet with incredible speed doing what ordinary browsers can do but with much greater efficiency, speed and capacity. The information so gathered is then passed on to indexers who index the content according to a set of business rules, algorithms and other criteria and store them as indexed data in huge databases. Each search engine company develops its own set of rules and criteria to organize the data they collect based upon the business model they have chosen. Once the data is organized, then client mechanisms such as search forms can be used to access it using keywords and various other criteria.

A typical search engine usually has four components. Together they constitute what we understand as search engine mechanism.

1. Information gathering mechanism.
2. Indexing mechanism.
3. Ranking mechanism and
4. Retrieval mechanism.

Limitations of Search Engines

Search engine business is very cost intensive because of the amount of work involved in gathering and indexing information and keeping it

up-to-date. To accomplish this task search engine companies have to invest heavily in the state of the art technology and technically qualified staff to maintain, manage and manipulate the information and make it useful, convenient and meaningful for the end users. The fast expanding World Wide Web, with its complexity and incongruity poses a multitude of problems and challenges to the search engine companies in managing information and keeping their technologies scalable and effective. Government interference, internet threats, cyber crime, linguistic and regional variations, cross-cultural issues, absence of uniform global internet policies and usability issues and people's unwillingness to pay for search are some of the serious issues which threaten the viability of search engine business and make it one of the most difficult to manage on a long term basis without recourse to search based ads and paid listings.

While these alternatives save the companies from financial problems, there is no guarantee that they do not undermine the quality of the information they provide. Despite the advances made in search engine technology, most search engines do not have necessary means to keep pace with the vast amount of data that is being added constantly to the World Wide Web and the new websites that are hosted every day. This results in some inefficiencies in the manner in which the search engines work which are discussed below.

- Search engines have built in limitations in responding to users' queries due to the limitations in their indexing mechanism or the algorithms they use. They may also respond differently to each keyword or combination of keywords or letters and symbols depending upon how they are programmed.
- Because of the limitations in processing and indexing information and the time and costs involved in removing irrelevant and

useless information to keep the indexes clean and up-to-date, a substantial portion of the content available on the world wide web is either outdated or outside the reach of the search engines and the public who use them. The so called invisible web is considered to be two to three times larger than the visible web.

- Search engines distribute information on several servers to manage load problems and not all of them are updated or available at the same time. So the results of a search query may vary depending upon which server received your query.

- Most search engines limit the number of pages they crawl on a website. Even in respect of pages they crawl they index only a certain portion of content and links on a page. Google for example indexes the first 101KB of a Web page, and 120KB of PDF's.

- Since most of the websites do not keep reliable records of date stamps or the dates on which they add or modify their content, date searching capability of search engine content is unreliable.

- The indexing is usually a long drawn process and may involve days and weeks before the information is processed and made available to the public. So the information is not always the latest.

- Spamming, keyword manipulation, search engine optimization techniques dilute and slow down the efforts of search engines in maintaining quality.

- Paid submission policies used by Yahoo and other companies and paid listings compromise the quality and the actual ranking of websites based on merit.

- The rules and regulations evolved by search engines to deal with duplicate content on the web often go against the original

providers of the information. Search engines do not have a reliable mechanism to distinguish original content from the duplicate because of limitations in date stamping. As a result, providers of original content often suffer due to illegal copying and reproduction.

- The ranking criteria used by search engines do not necessarily bring up the best websites in each category. Hindu website is one good example.

Directory Services

A Directory is a database of information about websites and their pages are organized alphabetically into categories, usually done by humans, instead of machines and automated software, using a set of predefined criteria. Users can navigate through the directory through a series of menus organized in a predictable manner to find the information they want. Unlike the search engines which require state of the art technology to gather and index information, the creation and maintenance of directory requires the involvement of huge manpower to organize, evaluate and categorize information. Hence they are slow to develop and usually smaller in size compared to the indexes created by commercial search engine companies.

One of the best examples of a web directory is the one maintained by dmoz.org, which being a public domain non-commercial directory is used by several search engines and websites like Hinduwebsite.com. Among the commercial directory Yahoo's directory is perhaps the best known and the largest. Besides general directories, there are also specialized directories dealing with a specific subject or category, also known as metasites.

The Directory vs. Search Engine

Directories are very useful when you are researching on a general topic, a popular category or a particular subject. For example if you are looking for information on religion, you can go to the society and culture part of a directory to begin your search. If you are looking for information on a particular religion such as Hinduism or Buddhism you can scroll down the category on religion in the directory and locate links to them easily.

Besides categories of information, the directory services usually provide an internal search engine with which you can easily look for information with in the directory using a keyword or combination of keywords. Search engines are more useful when you are looking for in-depth information or more recent information or more specialized information on a subject, or information that is beyond the scope of the categories in a directory. The standard practice is to begin your search with directories and then move on to search engines.

Meta Search Engines

Meta search engines do not use their own crawlers or databases to gather and index information. Instead they use a complex set of routines to access the databases publicly made available by various search engines to gather information and provide them to the public in an organized way. The advantage with Meta search tools is that you can simultaneously access various search engine databases and subject directories without doing individual searches and see the results displayed in one place. The main disadvantage is that the results are not necessarily comprehensive.

A Meta search tool can only fetch results from as many search engines as time, technology and resources permit. Secondly, Due to the

limitations placed by each search engine in retrieving information, you may not always get the best results or all the results. Besides, Meta search tools retrieve information basically through simple search routines. So these tools are not ideal for advanced search. Despite these limitations, if you are aiming to have an overview or comparative view of how each search engine is reacting to a particular keyword or a set of keywords, Meta search engines are the best place to start with.

Best Search Engine Tools			
Google	Scubtheweb	37.com	Kanoodle
All the Web	Jayde	OneSeek	NBCi/
Ask Jeeves	AOL Search	MetaSpider	Snap
Alta Vista	HotBot	Vivisimo	Go
Gigablast	Search.com	PlanetSearch	InfoSeek
Lycos	Metacrawler	surfwax	7Search
Teoma	Dogpile	qbSearch	Acclaim Search
Yahoo	Mamma	ProFusion	AllCrawl
AOL Search	C4	Proteus	Amnesi
MSN Search	Canada.com	Go2 Net	Ampleo
Netscape	ixquick	MegaGo.com	Deja.Com
Dipsie	Infogrid	WebFile	Deoji
Fybersearch	WebInfoSearch	myGO	DevSearch
Mozdex	Query Server	Megacrawler	Frequent Finders
Whatuseek	800go	Search Climbers	iBound
Wisenut	Debriefing	IX Quick	Info Hiway
ExactSeek	Highway 61	Northern Light	Infomak
Lost Link/ Web	Link Master	Subjex	GoshDarn!
Links	Splat Search	Zen Search	
Link Centre			
Meta Search Engines			
clusty.com	IxQuick	Ithaki	Turbo10
HighBeam	Search.com	KillerInfo	Weblens
Research	Fazzle	Mamma	Widow
Dogpile	Infogrid	Profusion	philb.com
Surfwax	Vivismo	Kartoo	Zapmeta Searchy
Copernic	Infonetware	QueryServer	
Metacrawler			

This Month That Day

ACM/GAMM committee
January 11, 1960

ACM/GAMM Committee Convenes to Develop Algol ACM/GAMM committee, a team of computer industry luminaries, convenes to develop Algol 60, the first block-structured language and one that eventually led to the more widely used Pascal. Algol (Algorithmic Language) and Algol 60 were designed to solve scientific computations and were meant to be more portable than most languages in existence at the time. Alan Perlis described Algol as the lingua franca of computer science.

Geoffrey Marcy (right) and Paul Butler
January 17, 1996

Computer is used in the Discovery of New Planets. Paul Butler and Geoffrey Marcy announced to the American Astronomical Society that they had discovered two new planets using an unconventional computer technique to analyze the movement of stars. Butler and Marcy let computers analyze spectrographic images of stars for eight years, looking for shifts in the light that would imply it is being pulled by the gravity of a planet. The first discovery, a planet orbiting the star 47 Ursae Majoris, was announced in December 1995 and, since then, this team found 12 planets outside of our solar system.

Console of GE's ERMA
(Electronic Recording Machine - Accounting)
January 28, 1952

Bank of America and SRI Sign a Contract to Develop ERMA's Pilot Model. Bank of America and SRI signed a contract for phase 3 of the proposal covering the development, construction, and testing of a pilot

model ERMA to provide service to 12 branches. The contract specified that Bank of America would pay SRI no more than $850,000 over four years, with an additional $25,000 for subcontracts. Although the final expenses were never released, most engineers estimate that the grand total was actually around $10 million.

January 30, 1952

Two New Primes Found with SWAC.

Using the Standards Western Automatic Computer (SWAC), researchers found two new prime numbers the first time they attempted a prime-searching program on the computer. Within the year, three other primes had been found. The National Bureau of Standards funded construction of the SWAC in Los Angeles in 1950 and it ran, in one form or another, until 1967.

January 31, 1995

AT&T and VLSI Protect Against Eavesdropping

AT&T Bell Laboratories and VLSI Technology announce plans to develop strategies for protecting communications devices from eavesdroppers. The goal would be to prevent problems such as insecure cellular phone lines and Internet transmissions by including security chips in devices.

GET ON NET

- http://homepage.ntlworld.com/kayseycarvey/ - Javascript Tutorials for Beginner's
- http://homepage.ntlworld.com/kayseycarvey/JavascriptIntro.html - Introduction to Javascript.
- http://www.homeandlearn.co.uk/JS/javascript.html - Javascript for beginners
- http://www.homeandlearn.co.uk/WD/WebDesign.html - Web Design Tutorial
- http://www.homeandlearn.co.uk/php/php.html - Beginners PHP

BLUE GENE

AUTHOR

Ms. S. Ramya,
Lecturer, Dept. of Computer Science

Blue gene is powerful super computer which describes the features and its architecture

Blue Gene is a computer architecture project to produce several supercomputers, designed to reach operating speeds in the PFLOPS (petaFLOPS) range, and currently reaching sustained speeds of nearly 500 TFLOPS (teraFLOPS). It is a cooperative project among IBM (particularly IBM Rochester and the Thomas J. Watson Research Center), the Lawrence Livermore National Laboratory, the United States Department of Energy (which is partially funding the project), and academia. There are four Blue Gene projects in development: Blue Gene/L, Blue Gene/C, Blue Gene/P, and Blue Gene/Q. The project was awarded the National Medal of Technology and Innovation by U.S. President Barack Obama on September 18, 2009. The president bestowed the award on October 7, 2009.

A BLUE GENE/L CABINET

The first computer in the Blue Gene series, **Blue Gene/L**, developed through a partnership with Lawrence Livermore National Laboratory (LLNL), originally had a theoretical peak performance of 360 TFLOPS, and scored over 280 TFLOPS

sustained on the Linpack benchmark. After an upgrade in 2007 the performance increased to 478 TFLOPS sustained and 596 TFLOPS peak. The term *Blue Gene/L* sometimes refers to the computer installed at LLNL; and sometimes refers to the architecture of that computer. As of November 2006, there are 27 computers on the Top500 list using the Blue Gene/L architecture. All these computers are listed as having an architecture of *eServer Blue Gene Solution*.
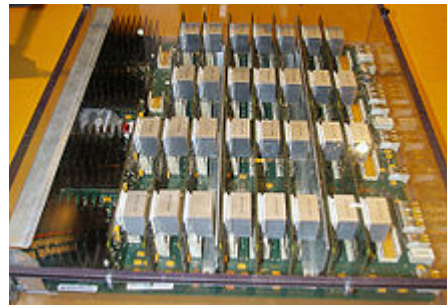
MAJOR FEATURES

The Blue Gene/L supercomputer is unique in the following aspects:

- Trading the speed of processors for lower power consumption.
- Dual processors per node with two working modes: co-processor (1 user process/node: computation and communication work is shared by two processors) and virtual node (2 user processes/node)
- System-on-a-chip design
- A large number of nodes (scalable in increments of 1024 up to at least 65,536)
- Three-dimensional torus interconnect with auxiliary networks for global communications, I/O, and management
- Lightweight OS per node for minimum system overhead (computational noise)

ARCHITECTURE

A schematic overview of a Blue Gene/L supercomputer, each Compute or I/O node is a single ASIC with associated DRAM memory chips. The ASIC integrates two 700 MHz PowerPC 440 embedded processors, each with a double-pipeline-double-precision Floating Point Unit (FPU), a cache sub-

system with built-in DRAM controller and the logic to support multiple communication sub-systems. The dual FPUs give each Blue Gene/L node a theoretical peak performance of 5.6 GFLOPS (gigaFLOPS). Node CPUs are not cache coherent with one another.

Compute nodes are packaged two per compute card, with 16 compute cards plus up to 2 I/O nodes per node board. There are 32 node boards per cabinet/rack.[10] By integration of all essential sub-systems on a single chip, each Compute or I/O node dissipates low power (about 17 watts, including DRAMs). This allows very aggressive packaging of up to 1024 compute nodes plus additional I/O nodes in the standard 19" cabinet, within reasonable limits of electrical power supply and air cooling. The performance metrics in terms of FLOPS per watt, FLOPS per m^2 of floorspace and FLOPS per unit cost allow scaling up to very high performance.

Each Blue Gene/L node is attached to three parallel communications networks: a 3D toroidal network for peer-to-peer communication between compute nodes, a collective network for collective communication, and a global interrupt network for fast barriers. The I/O nodes, which run the Linux operating system, provide communication with the world via an Ethernet network. The I/O nodes also handle the filesystem operations on behalf of the compute nodes. Finally, a separate and private Ethernet network provides access to any node for configuration, booting and diagnostics.

Blue Gene/L compute nodes use a minimal operating system supporting a single user program. Only a subset of POSIX calls are supported, and only one process may be run at a time. Programmers need to implement green threads in order to simulate local concurrency.Application development is usually performed in C, C++,

or Fortran using MPI for communication. However, some scripting languages such as Ruby have been ported to the compute nodes.

To allow multiple programs to run concurrently, a Blue Gene/L system can be partitioned into electronically isolated sets of nodes. The number of nodes in a partition must be a positive integer power of 2, and must contain at least $2^5 = 32$ nodes. The maximum partition is all nodes in the computer. To run a program on Blue Gene/L, a partition of the computer must first be reserved. The program is then run on all the nodes within the partition, and no other program may access nodes within the partition while it is in use. Upon completion, the partition nodes are released for future programs to use. With so many nodes, component failures are inevitable. The system is able to electrically isolate faulty hardware to allow the machine to continue to run.

GPRS – An Overview

Author

Ms. M. M. Kavitha,
Lecturer, Dept. of Computer Science

This articles describes about the architecture of general packet radio service

The General Packet Radio Service (GPRS) is an enhancement to the existing GSM network infrastructure and provides a connectionless packet data service. The same cellular base-stations that support voice calls are used to support GPRS and as a consequence GPRS can be used wherever it is possible to make a voice call. GPRS roaming

agreements exist with a large number of countries and this means users can use GPRS devices whilst abroad.

GPRS is based on internet Protocols (IP) and enables users to utilize a wide range of applications – email and internet and/or intranet resources for instance. With throughput rates of up to 40 Kbit/s, users have a similar access speed to a dial-up modem, but with the convenience of being able to connect from anywhere.

GPRS is classed as being a packet switched network whereby radio resources are used only when users are actually sending or receiving data. Rather than dedicating a radio channel to a mobile data user for a fixed period of time, the available radio resource can be concurrently shared between several users. This efficient use of scarce radio resources means that large numbers of GPRS users can potentially share the same bandwidth and be served from a single cell. The actual number of users supported depends on the applications being used and how much data is being transferred.

The term "always on always connected" is often used when people are describing GPRS and means once users have logged on they can remain connected to the data network for the working day. It should be noted that unlike GSM circuit switched data working, where the cost of the data call is related to the time spent connected to the network, this is not an issue when using GPRS as the cost of a GPRS data session is dependant on the amount of data sent and received not the time spent connected to the network.

GPRS introduces a number of new functional elements that support the end to end transport of IP based packet data. GPRS was developed by the GSM standards bodies, resulting in a system with defined functionality, interfaces and inter-network operation for roaming support. The GPRS network architecture is shown in Figure.
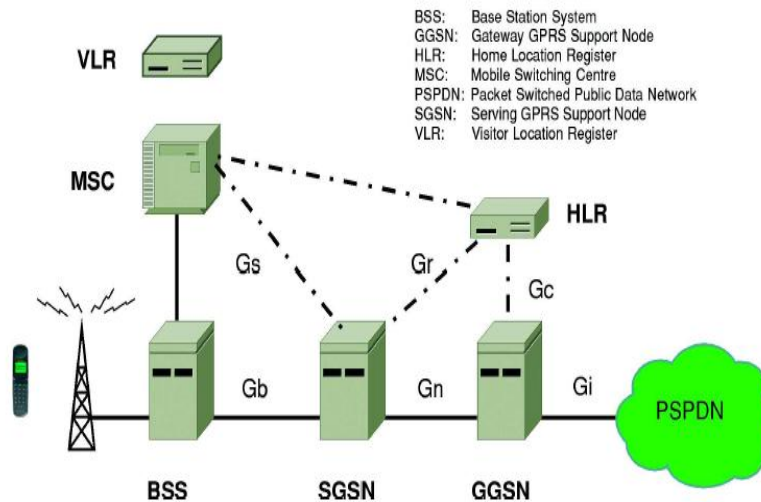
Fig: GSM Network Architecture

**Two major new core network elements are introduced:**

The Serving GPRS Support Node (SGSN) and the Gateway GPRS Support node (GGSN). The SGSN monitors the state of the mobile station and tracks its movements within a given geographical area. It is also responsible for establishing and managing the data connections between the mobile user and the destination network.

The GGSN provides the point of attachment between the GPRS domain and external data networks such as the internet and Corporate Intranets. Each external network is given a unique Access Point Name (APN) which is used by the mobile user to establish the connection to the required destination network.

The GSM Base Station Subsystem (BSS) is adapted to support the GPRS connectionless packet mode of operation. A new functional node called the Packet Control Unit (PCU) is introduced (as part of the BSC) to control and manage the allocation of GPRS radio resources to mobile users.

MYSELF IN BOOKSHELF
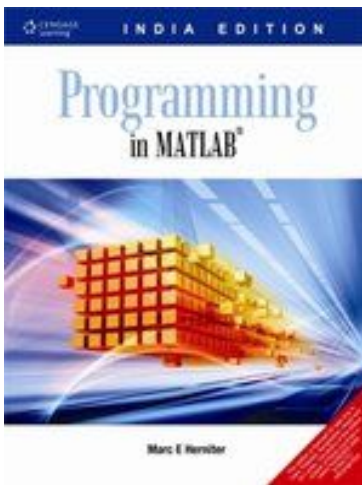
Book Name : "Programming in MATLAB"
Author   : Marc E. Herniter
**   Cengage learning**

I "Programming in MATLAB" is composed by Marc E. Herniter. I give introductory programming class that uses MATLAB as the programming language rather than the traditional language, Such as FORTRON/C. I am composed of both GUI tools and all of the toolboxes that make MATLAB a popular engineering, science and math tool. Hope my contents will be valuable for the beginners who intend to do research using MATLAB.

Important:

Software & MATLAB files can be viewed at the link http://www.brookscole.com

Comments & Suggestions can be posted to marc.herniter@nau.edu
Website: www.cet.nau.edu/meh/books/matlab.htm

The readers can access me at KSRCAS main library

JARGONS

iPhone: An Internet-enabled Smartphone developed by Apple. The iPhone combines mobile phone capabilities with a wireless Internet device, and an iPod into one product. The iPhone also includes a 3.5-inch multi-touch screen, rather than a keyboard, that can be manipulated by users with by two finger touches. The iPhone runs on a special version of Apple's Mac OS X operating system

iPad: A new handheld computing device launched by Apple Inc. in January 2010. The iPad is designed for consumers who want a mobile device that is bigger than a Smartphone but smaller than a laptop for entertainment multimedia.

ZigBee: is a low data rate, two-way standard for home automation and data networks. It provides a specification for up to 254 nodes including one master, managed from a single remote control.

OBEX: OBject EXchange, a communications protocol designed for swapping binary objects between particular devices, using IrDA or Bluetooth. OBEX is usually associated with and optimized for ad-hoc wireless links.

G_Port: A Generic port (G_Port) within a Fibre Channel switch that can operate as either an E_Port or an F_Port.

3GP: A container format that is a simplified version of the MPEG-4 (MP4) format that uses less storage and bandwidth for better streaming on mobile phones.

Wiimote: (pronounced wee-moat) The wireless, remote-control-style controller for Nintendo's next generation games console, the Nintendo Wii.

Shunt: In computers and electronics, a shunt is the conductive sleeve that is placed over the jumper pins to close an electrical circuit.

iBGP: Short for Internal BGP (Border Gateway Protocol), iBGP is the protocol used between the routers in the same autonomous system (AS).

IBGP is used to provide information to your internal routers
Hi8: A higher-quality version of the 8mm videotape format that offers a sharper picture. Hi8 is the successor to the original Video8 format and is an analog format.
CONUS (CONTIGUOUS UNITED STATES): A term used to describe satellite coverage limited to the continental United States excluding Alaska and Hawaii.
DAMA (DEMAND-ASSIGNED MULTIPLE ACCESS): A process whereby satellite transponder channels are assigned for telephony transmission on the basis of immediate traffic demands.
VSAT: (VERY SMALL APERTURE TERMINAL) Earth station satellite antenna with a diameter or cross-section dimension in the general range of 1.2 to 2.4 meters
Halftone: A halftone image is made up of a series of dots rather than a continuous tone. These dots can be different sizes, different colors, and sometimes even different shapes. Larger dots are used to represent darker, denser areas of the image, while smaller dots are used for lighter areas.
Kerning: It refers to the spacing between the characters of a font. Without kerning, each character takes up a block of space and the next character is printed after it.

Summary of Bit Rate notations

Bit: 10p0 (1 bit)
Kilobit: 10p3 (1,000 bits)
Megabit: 10p6 (1,000,000 bits)
Gigibit: 10p9 (1,000,000,000 bits)
Terabit: 10p12 (1,000,000,000,000 bits)
Petabit: 10p15 (1,000,000,000,000,000 bits)
Exabit: 10p18 (1,000,000,000,000,000,000 bits)
Zettabit: 10p21 (1,000,000,000,000,000,000,000 bits)
Yottabit: 10p24 (1,000,000,000,000,000,000,000,000 bits)
A 40 Exabit file would contain all of the words ever spoken by human beings.

Academic Forum

Author

Questions Posted By: *Vinodhini,* **III B.Sc CS 'A'**

Answers Given by: **Ms. M. M. Kavitha,**
              Lecturer, *Dept of Computer Science*

How File Compression Works?

     f you download many programs and files off the Internet, you've probably encountered ZIP files before. This compression system is a very handy invention, especially for Web users, because it lets you reduce the overall number of bits and bytes in a file so it can be transmitted faster over slower Internet connections, or take up less space on a disk. Once you download the file, your computer uses a program such as WinZip or Stuffit to expand the file back to its original size. If everything works correctly, the expanded file is identical to the original file before it was compressed.

     Most types of computer files are fairly redundant -- they have the same information listed over and over again. File-compression programs simply get rid of the redundancy. Instead of listing a piece of information over and over again, a file-compression program lists that information once and then refers back to it whenever it appears in the original program.

Example: The following quote were given by John F. Kennedy's 1961.

> *"Ask not what your country can do for you -- ask what you can do for your country."*

The quote has 17 words, made up of 61 letters, 16 spaces, one dash and one period. If each letter, space or punctuation mark takes up one unit of memory, we get a total file size of 79 units. To get the file size down, we need to look for redundancies. The following words are:

- **"ask" appears two times**
- **"what" appears two times**
- **"your" appears two times**
- **"country" appears two times**
- **"can" appears two times**
- **"do" appears two times**
- **"for" appears two times**
- **"you" appears two times**

Ignoring the difference between capital and lower-case letters, roughly half of the phrase is redundant. Nine words -- ask, not, what, your, country, can, do, for, you -- give us almost everything we need for the entire quote. To construct the second half of the phrase, we just point to the words in the first half and fill in the spaces and punctuation.

REDUNDANCY AND ALGORITHMS

Most compression programs use a variation of the **LZ adaptive dictionary-based algorithm** to shrink files. "LZ" refers to **Lempel and Ziv**, the algorithm's creators, and "dictionary" refers to the method of **cataloging** pieces of data.

The system for arranging dictionaries varies, but it could be as simple as a numbered list. When we go through Kennedy's famous words, we pick out the words that are repeated and put them into the numbered index. Then, we simply write the number instead of writing out the whole word.

So, if this is our dictionary:

1. **ask**
2. **what**
3. **your**
4. **country**
5. **can**
6. **do**
7. **for**
8. **you**

Our sentence now reads:

"1 not 2 3 4 5 6 7 8 -- 1 2 8 5 6 7 3 4"

If you knew the system, you could easily reconstruct the original phrase using only this dictionary and number pattern. This is what the expansion program on your computer does when it expands a downloaded file. You might also have encountered compressed files that open themselves up. To create this sort of file, the programmer includes a simple expansion program with the compressed file. It automatically reconstructs the original file once it's downloaded.

But how much space have we actually saved with this system? "1 not 2 3 4 5 6 7 8 -- 1 2 8 5 6 7 3 4" is certainly shorter than "Ask not what your country can do for you; ask what you can do for your country;" but keep in mind that we need to **save the dictionary itself along with the file**.

Archives

ARTICLE TITLE	AUTHOR NAME	ISSUE	MONTH	YEAR
IT Icons	S.Mageshwaran, II-BCA 'A'	Issue#7	March	2009
Nanotechnology	D.Suryaprakash, II-BCA 'A'	Issue#7	March	2009
Internet Tips	S.Sasikala, Lecturer (CS)	Issue#7	March	2009
Jargons	A.AbdulJabbarSherif, I-B.Sc (CS) 'B'	Issue#7	March	2009
Multibiometrics	S. Sasikala & M.M.Kavitha, Lecturer(CS)	Issue#3	November	2008
Process Explorer 11.33	M.M.Kavitha, Lecturer (CS)	Issue#9	May	2009
The Ace Computer	G.Anwar Basha, Lecturer(CS)	Issue#18	February	2010
Free Blog Creation	Ms. F.Regina Mary, Lecturer(CS)	Issue#23	July	2010
Sensing 3D image	S.Nithya Lecturer(CS)	Issue#23	July	2010
Web3.0	Ms.V.Kavitha Lecturer(CS)	Issue#22	June	2010

COMPUTER VIRUSES

Author: A. Kavinraj, I BCA - C

This articles describes the top 7 most computer viruses of 20th century

The 20th century was a period when computer viruses saw light and managed to cause serious damages to unprepared users and their machines. Millions of users downloaded computer viruses without even knowing about it. To download a computer virus one simply needed to click on a certain link or open an email attachment.

Today the best way to keep your computer virus-free is to install effective antivirus software like the Antivirus and Security All-in-One-Suite that will not only allow you to check for any malware but will also stop attacks on your machine. However, previously there were little or no such programs.

Hackers used various ways of penetrating the systems of universities from around the world, NASA, armed forces and governments. In this article you will find the most dangerous computer viruses of the 20th century. Note that the list includes computer viruses selected by the InfoNIAC team, which means that it represents our subjective opinion.

1. ILOVEYOU - 2000

Considered to be one of the most dangerous computer worms in history, ILOVEYOU was for the first time detected in Philippines on May 4. In just one day it spread worldwide, with about 10 percent of all Internet users being affected by the worm. ILOVEYOU also caused $5.5 billion damage. Its name comes from the subject of the email that a user received. The email included an attachment that said "LOVE-LETTER-FOR-YOU.TXT.vbs". When activated, the computer worm overwrote all files on

the host, attacking a copy of itself to each file. In addition, the worm sent a copy of itself to everyone in a user's contact.

2. Melissa - 1999

This computer virus managed to cause a damaged estimated between $300 and $600 million. The famous computer virus gained huge fame on March 26, 1999. Experts claimed that this computer virus managed to infect between 15 and 20 percent of all business PCs on the planet. Melissa virus spread so fast that a number of big companies, including Intel and Microsoft, that used Outlook, had to shut down their email systems to hold up the damage.

The computer virus took advantage of Microsoft outlook to send email messages of itself to 50 names on the contact list of a user. The message in the email read: "Here is that document you asked for don't show anyone else." It featured an attached Word document, which, when opened, unleashed the virus that infected the computer and then repeated the replication. When the virus was activated it changed Word documents by leaving quotes from the popular animated show "The Simpsons".

3. CIH (a.k.a. Chernobyl) - 1998

Our list of the most dangerous computer viruses on the 20th century continues with CIH virus that caused an estimated damage of $20 to $80 million around the globe. After its launch, the computer virus managed to affect huge amounts of data stored on computers. Later it was discovered that the computer virus was launched in Taiwan. It has been recognized to be one of the most dangerous computer viruses in history. It infected Windows 95, 98, and ME executable files.

In addition, CIH remained resident in the memory of the machine, being able to carry on infecting other executables. After being activated, the virus overwrote data on the HDD of the infected PC, making the latter inoperable. CIH could also overwrite the BIOS of the infected computer, thus preventing boot-up. The second name of the virus - Chernobyl - was

given because some of the biggest damages occurred on the day when the nuclear reactor exploded**.**

4. Solar Sunrise - 1998

This is not the name of the computer virus; it is the name of the situation that occurred in 1998 when a team of hackers managed to take control of more than 500 computer system of the army, government as well as private sector of the United States. The name Solar Sunrise comes after the well-known vulnerabilities in machines that run on the Sun Solaris OS. At first the attack was believed to have been organized by hackers from Iraq, but later it was discovered that the ones to blame were two American youngsters from California.

5. Barrotes - 1993

This is believed to be the first popular computer virus developed in Spain. As soon as it infected the system, it would remain there until January the 5th, when it would set off showing a series of bars on the screen. It infected .COM, .EXE and overlay files. The Barrotes computer virus represents a resident virus - it becomes a resident of the computer memory each time the machine starts up. Due to a series of vertical lines that appear on the monitor, it was easy to identify the virus. It could also overwrite the Master Boot Record of the HDD, thus making it impossible for the uses to access the hard disk.

6. Morris (a.k.a. Internet Worm) - 1988

Previously at www.InfoNIAC.com we wrote about this computer worm, which is believed to be one of the first worms that spread over the Internet. The name of the virus comes from its developer Robert Tappan Morris, who was a student at Cornell University. The computer worm was set off on November 2, 1988 and after some time it managed to infect 6,000 to 9,000 machines. It overloaded the whole Internet, leading to the failure of a large number of servers. According to its developer, the goal was to discover just how far and fast a computer worm can spread all over the network. Robert Tappan Morris was found guilty and sentenced to 3 years of

probation along with 400 hours of community service. In addition, he had to pay a $10,000 fine.

7. Jerusalem - 1987

For the first time virus was discovered in 1987 on October 1. The virus was dubbed Jerusalem because it was identified in a Hebrew university, but in 1991 antivirus experts found that Italy is the country were the first traces of the computer virus were noticed. Jerusalem is believed to be one of the oldest computer viruses in history, being able to infect both .EXE and .COM files. Initially the Jerusalem virus included a bug that led to a repeat infection of the files that continued until the size of the files overcome computer resources. In addition, each Friday 13 it deleted all programs in the infected system as a result of a malicious payload that set off on the respective day. Jerusalem considerably slowed down the machine. A person could identify the virus but noticing two lines on the monitor.

MOTION CAPTURE

Author: T. Vadivel,

Lecturer, Department of Computer Science – PG

Motion capture, motion tracking, or mocap is a technique of digitally recording movements for entertainment, sports, and medical applications. In the context of filmmaking (where it is sometimes called performance capture), it refers to the technique of recording the actions of human actors, and using that information to animate digital character models in 3D animation.

Advantages

➤ Mo cap offers several advantages over traditional computer animation of a 3D model:

➤ More rapid, sometimes even real time results can be obtained.

➤ The amount of work does not vary with the complexity or length of the performance to the same degree when using traditional techniques.

➤ Complex movement and realistic physical interactions such as secondary animation, weight and exchange of forces can be more easily recreated in a physically accurate manner.

➤ Mocap technology allows one actor to play multiple roles within a single film.

Disadvantages

➤ Specific hardware and special programs are required to obtain and process the data.

➤ The cost of the software and equipment, personnel required can be prohibitive for small productions.

➤ The capture system may have specific requirements for the space it is operated in.

➤ When problems occur it is sometime easier to reshoot the scene rather than trying to manipulate the data. Only a few systems allow real time viewing of the data to decide if the take needs to be redone.

➤ Applying motion to quadruped characters can be difficult.

➤ The technology can become obsolete every few years as better software and techniques are invented.

➤ The results are limited to what can be performed within the capture volume without extra editing of the data.

➤ Movement that does not follow the laws of physics generally cannot be represented.

Traditional animation techniques such as added emphasis on anticipation and follow through, secondary motion or manipulating the shape of the character as with squash and stretch animation techniques are generally not applicable. If the computer model has different proportions from the capture subject artifacts may occur. For example, if a cartoon character has large, over-sized hands, these may intersect strangely with any other body part when the human actor brings them too close to his body. The real life performance may not translate on to the computer model as expected.

Applications

Some video games use motion capture to animate athletes, martial artists, and other in-game characters. Movies use motion capture for CG effects, in some cases replacing traditional cell animation, and for completely computer-generated creatures, such as Gollum, The Mummy, and King Kong.

Sinbad: Beyond the Veil of Mists was the first movie made primarily with motion capture, although many character animators also worked on the film.

In producing entire feature films with computer animation, the industry is currently split between studios that use motion capture, and studios that do not. Out of the three nominees for the 2006 Academy Award for Best Animated Feature, two of the nominees (Monster House and the winner Happy Feet) used motion capture, and only Pixar's Cars was animated without motion capture. In the ending credits of Pixar's film Ratatouille, a stamp appears labelling the film as "100% Pure Animation -- No Motion Capture!"

Motion capture has begun to be used extensively to produce films which attempt to simulate or approximate the look of live-action cinema, with nearly photorealistic digital character models. The Polar Express used it to translate the actions of star Tom Hanks to several distinct digital characters (for which he also provided the voices). The 2007 adaptation of the saga Beowulf used it to animate digital characters whose appearances

were based in part on the actors who provided their motions and voices. The Walt Disney Company has announced that it will distribute Robert Zemeckis's A Christmas Carol to be produced using this technique.

Virtual Reality and Augmented Reality allow users to interact with digital content in real-time. This can be useful for training simulations, visual perception tests, or performing a virtual walk-through in a 3D environment. Motion capture technology is frequently used in digital puppetry systems to aid in the performance of computer generated characters in real-time.

Gait analysis is the major application of motion capture in clinical medicine. Organic Motion, makers of markerless motion capture have recently pioneered the ability for clinicians to evaluate human motion, without burdening patients with cumbersome body suits or tracking devices. This ability to allow patients to move freely within a defined area, and uses cameras, not markers, to track range of motion, gait, and several other biometric factors, and stream this information live into analytical software. Because this system removes the markers, patients, physicians and analysts are able to collect quantifiable data in real-time with less patient inconvenience.

Methods and Systems

Motion tracking or motion capture started as a photogrametric analysis tool in biomechanics research in the 1970s and 1980s, and expanded into education, training, sports and recently computer animation for cinema and video games as the technology matured.

A performer wears markers near each joint to identify the motion by the positions or angles between the markers. Acoustic, inertial, LED, magnetic or reflective markers, or combinations of any of these, are tracked, optimally at least two times the rate of the desired motion, to sub-millimeter positions. The motion capture computer software records the positions, angles, velocities, accelerations and impulses, providing an accurate digital representation of the motion.

In entertainment applications this can reduce the costs of animation which otherwise requires the animator to draw each frame, or with more sophisticated software, key frames which are interpolated by the software. Motion capture saves time and creates more natural movements than manual animation, but is limited to motions that are anatomically possible. Some applications might require additional impossible movements like animated super hero martial arts or stretching and squishing that are not possible with real actors. In biomechanics, sports and training, real time data can provide the necessary information to diagnose problems or suggest ways to improve performance, requiring motion capture technology to capture motions up to 140 miles per hour for a golf swing.

Optical Systems

Optical systems utilize data captured from image sensors to triangulate the 3D position of a subject between one or more cameras calibrated to provide overlapping projections. Data acquisition is traditionally implemented using special markers attached to an actor; however, more recent systems are able to generate accurate data by tracking surface features identified dynamically for each particular subject.

Tracking a large number of performers or expanding the capture area is accomplished by the addition of more cameras. These systems produce data with 3 degrees of freedom for each marker, and rotational information must be inferred from the relative orientation of three or more markers; for instance shoulder, elbow and wrist markers providing the angle of the elbow.

Related Techniques

Facial motion capture is utilized to record the complex movements in a human face, especially while speaking with emotion. This is generally performed with an optical setup using multiple cameras arranged in a hemisphere at close range, with small markers glued or taped to the actor's face. However there are a number of systems such as Image Metrics, RF (radio frequency) positioning systems are becoming more viable as higher frequency RF devices allow greater precision than older RF technologies.

The speed of light is 30 centimeters per nanosecond (billionth of a second), so a 10 gigahertz (billion cycles per second) RF signal enables an accuracy of about 3 centimeters. By measuring amplitude to a quarter wavelength, it is possible to improve the resolution down to about 8 mm.

To achieve the resolution of optical systems, frequencies of 50 gigahertz or higher are needed, which are almost as line of sight and as easy to block as optical systems.

Multipath and reradiation of the signal are likely to cause additional problems, but these technologies will be ideal for tracking larger volumes with reasonable accuracy, since the required resolution at 100 meter distances isn't likely to be as high. An alternative approach was developed where the actor is given an unlimited walking area through the use of a rotating sphere, similar to a hamster ball, which contains internal sensors recording the angular movements, removing the need for external cameras and other equipment. Even though this technology could potentially lead to much lower costs for mocap, the basic sphere is only capable of recording a single continuous direction.

HOTTEST NEWS
NOKIA may come to vie for laptop industry

Founder: Knut
Fredrik Idesram.

Nokia Corporation is a finish multinational communications corporation, headquartered in Keilaniemi, Espoo, a city neighboring Finland's capital Helsinki. It is a large mobile company with 128,445 employees in 120 countries, sales in more than 150 countries. Since late last year, there were rumors in the industry that Nokia plans to get into the PC market. And this may be true from what Nokia's Chief Executive Olli-Pekka Kallasvuo told Reuters in a recent interview. While the gap between a mobile phone and a PC is lessening. Now we have to wait for their next declaration and have to wait for taking their goods taste.

Mailing List

To whom we send

- The Vice-Chancellor, Periyar University ,Salem-11
- The Registrar, Periyar University ,Salem
- The Controller of Examination, Periyar University ,Salem-11
- The HOD, Department of Computer Science, Periyar University,Salem-11
- The HOD, Government Arts College for Women, Salem-8
- The HOD, Government Arts College for Women , Krishnagiri
- The HOD, Government Arts & Science College (W), Burgur, Kirshnagiri
- The HOD, J.K.K Nataraja College of Arts & Science
- The HOD, M.G.R College of Arts & Science
- The HOD, Sengunthar Arts & Science College
- The HOD, Muthayammal College of Arts & Science
- The HOD, PEE GEE College of Arts &, Science
- The HOD, Harur Muthu Arts & Science College for Women
- The HOD, Vivekanandha College of Arts & Sciences (W)
- The HOD, Mahendra Arts & Science college
- The HOD, Selvam Arts & Science college
- The HOD, St.Joseph's College of Arts & Science for (W)
- The HOD, Vysya College of Arts &, Science
- The HOD, NKR Government Arts College for Women
- The HOD, Arignar Anna Government Arts College
- The HOD, Salem Sowdeswari College
- The HOD, P.G.P College of Arts & Science
- The HOD, Attur Arts & Science College
- The HOD, SSM College of Arts & Science
- The HOD, Government Arts College Salem
- The HOD, Government Arts College Men
- The HOD, Government Arts College, Dharmapuri
- The HOD, Gobi Arts and Science College (Autonomous)
- The HOD, Sri Kandhan College of Arts & Science
- The HOD, Sri Ganesh College of Arts & Science
- The HOD, Jairam Arts & Science College
- The HOD, Sri Balamurugan College of Arts & Science
- The HOD, PSG College of Arts and Science
- The Secretary, PSG College of Arts and Science
- The HOD, Kongunadu Arts and Science College(Autonomous)
- The HOD, Vivekanandha College for Women
- The HOD, Sri Vidhya Mandir Arts & Science College
- The HOD, *St*.John's College *Palayamkottai* - 627 007
- Mr. S.T.Rajan, St. Joseph's College, Trichy

FUTURE TRENDS

DATA MINING TRENDS IN TECHNOLOGIES AND METHODS

1. **Distributed/Collective Data Mining**
   - Performing local data analysis for generating partial data models,
   - Combining the local data models from different data sites in order to develop the global model.
2. **Ubiquitous Data Mining (UDM)**
   - The objectives of UDM is to mine data while minimizing the cost of ubiquitous presence.
   - Accessing and analyzing data from a ubiquitous computing devices Such as laptops, palmtops, cell phones etc.
3. **Hypertext and Hypermedia Data Mining**
   - Mining hypermedia information like text, hyperlinks, text markups, online catalogues, digital libraries, and online information databases.
4. **Multimedia Data Mining**
   - Mining the images, video, audio, and animation.
5. **Spatial and Geographic Data Mining**
   - Used in distance and topological information and to handle geometric calculations
6. **Time Series/Sequence Data Mining**
   - Mining of time series and sequence-based data
   - Includes long-term or trend movements, seasonal variations, cyclical variations, and random movements.

Note:

More information regarding the trends of data mining can be gathered from the website given below.

http://www.csse.monash.edu.au/~mgaber/DATA%20MINING%20TRENDS%20AND%20DEVELOPMENTS%20The%20Key%20Data%20Mining%20Technologies%20and%20Applications%20for%2021%20century.pdf

Achievers Archives

Participation of Staff Members in International Conferences

1	**S. Ranichandra**	Presented a paper on Comparative Study of ACO with Grid Scheduling algorithm – An Approach (**Awarded as best paper**)	International Conference Systemics, Cybernates and Informatics ICSCI-2011, Hyderabad, Jan 5th to 8th 2011
		Presented a paper on Laplacian of Gaussian filter and water shed algorithm in early detection of Breast Cancer	
		Presented a paper on An emerging analysis of decision tree algorithms in data mining	International Conference on Mathematics and Computer Science (ICMCS 2011), Loyala college, Chennai, Jan 7th to 8th 2011
2	**S. Pradeesh Hosea**	Presented a paper on 3D Technology	International Conference on Computer Application and Recent Trends in Digital Animation Industry, PSGR Krishnammal College for Women, Comibatore, Jan 10th 2011
3	**D.Saraswathi & S.Anita**	Presented a Paper on A Prominent solution for Internet plagiarism & Presented a Paper on A Novel Search tool for Recruiters to Discover profile of candidates in Social Network using floor model, UCINET	International Conference on Mathematics and Computer Science (ICMCS 2011), Loyala college, Chennai, Jan 7th to 8th 2011