Pragyaan : Journal of Information Technology

Volume 11 : Issue 1, June, 2013

Panel of Referees

Dr. G P Sahu
Associate Prof., School of Management Studies
Moti Lal Nehru National Institute of Technology
Allahabad

Prof. I Husain
Deptt. of Mathematics
Jaypee Institute of Engineering & Technology Guna

Dr. Rajendra Kumar Gartia
Deptt. of Mathematics
Sambhalpur University, Orissa

Prof. Rajiv Saxena
Head, ECE
Jaypee Institute of Engineering & Technology, Guna

Dr. Durgesh Pant
Head, Deptt. of Computer Applications
Kumaon University, Nainital

Dr. D S Hooda
Head, Deptt. of Mathematics
Jaypee Institute of Engineering & Technology, Guna

Dr. Nipur
Head, Deptt. of Computer Applications
Kanya Gurukul Mahavidyalaya Dehradun

Dr. Shishir Kumar
Head, CSE
Jaypee Institute of Engineering & Technology, Guna

Dr. R K Sharma
Dean, Computer Science
Thapar University, Patiala

Dr. Sameer Saran
Scientist
Indian Institute of Remote Sensing, Dehradun

Dr. Saurabh Pal
Deptt. of Computer Applications
VBS Purvanchal University, Jaunpur

S Dimri
Head, Deptt. of Computer Applications
GEIT University, Dehradun

Dr. Vipin Tyagi
Asstt. Prof., CSE
Jaypee Institute of Engineering & Technology, Guna

Dr. Shailendra Mishra
Head, IT
Dehradun Institute of Technology, Dehradun

Dr. K C Joshi
Deptt. of IT & Management
MJP Rohillkhand University, Bareilly

Prof. R Sukesh Kumar
Deptt. of CSE
Birla Institute of Technology, Mesra

Prof. K R Pardasani
Head, Mathematics & Computer Applications
Maulana Azad National Institute of Technology, Bhopal

Prof. P K Panigrahi
Indian Institute of Science Research
Kolkata (Constituent of IIT, Kharagpur)

Prof. R C Chakraborty
Former Director, DRDO, DTRL
Dr. Ravinder Singh
Deptt. of CSE
MJP Rohillkhand University
Bareilly

Dr. Somesh Kumar
HOD, Dept. of Information Technology
Noida International University,
Gautam Budh Nagar, UP, India

Dr. Y N Singh
Institute of Engineering & Tech, Govt. Eng College
Sitapur Road, Lucknow (UP)

Dr. Manu Pratap Singh
Dept. of Computer Science,
Dr. B.R. Ambedkar University, (khandari
Campus), Khandari Agra (UP) - 282002

Dr. Satish Kumar Singh
Department of ECE,
Jaypee University of Eng & Tech, Raghogarh,
Guna (MP)-473226

Dr. Sanjai Gupta
Joint Controller of Examination,
Mahamaya Technical University,
Noida (UP) -201301

Dr. Ashish Negi
Associate Professor & Head
MCA Department
G.B.Pant Engineering College Pauri Garhwal,
Uttarakhand

Dr. V K Singh
Dean (R & D)
Institute of Engineering & Technology
Sitapur Road, Lucknow -226021

Dr. Neelam Srivastava
Department of Electronics & Communication
Institute of Engineering & Technology
Sitapur Road, Lucknow -226021

From the Chief Editor

It is with much joy and anticipation that we present the June 2013 issue of our journal, Pragyaan: Journal of Information Technology.

Pragyaan: Journal of Information Technology is an open access Journal that publishes articles which contribute new results in areas of computer science. It is committed to rapid dissemination of high quality research in IT that can help us meet the challenges of the 21st century. The Journal strives to seek ways to harness the power of technology to meet some of the real world challenges, and to provide substance for making informed judgments on important matters. We welcome contributions that can demonstrate practical usefulness, particularly contributions that take a multi disciplinary convergent approach to deal with real world problems that are complex in nature.

The academically peer refereed Pragyaan: JOIT encourages authors to develop and publish quality papers that address various facets of Information Technology in a balanced manner. Selecting quality papers for publication in our Journal is indeed a tough task. We thank our panel of referees for the time and thought invested by them into the papers and for giving us sufficient insights to ensure the quality of papers published in Pragyaan: JOIT. Thanks are also due to the members of our Editorial Board and Board of Management for their constant guidance and support.

We would like to acknowledge the contribution of Dr.A.S. Pandey (Prof. Incharge, Research Publications), Prof. Rajeev Srivastava (Editor Pragyaan: JOIT) and Prof. Sumiti Kapoor (faculty) in editing, leading to enhanced reader friendliness of many articles.

We would like to express gratitude to our esteemed authors, editors and reviewers all of whom have volunteered to contribute to the success of the Journal. We do our best to oversee a review and decision-making process in which we invite appropriate individuals to review each paper and encourage them to provide timely, thoughtful, constructive, and diplomatic critiques. We work towards integrating reviewers' feedback along with our own insights into the final decision and craft fair and balanced action that acknowledges the strengths of the manuscript, addresses areas for improvement, and clearly conveys the editorial decision and its rationale.

We wish to encourage contributions from the scholars, scientific community and industry practitioners to ensure a continued success of the journal.

We hope our readers find our Journal, Pragyaan: Journal of Information Technology informative and stimulating. We welcome comments and suggestions for further improvement in quality of the Journal.

Prof. (Dr.) Pawan K Aggarwal
Associate Pro Vice Chancellor
IMS Unison University, Dehradun

Pragyaan: Journal of Information Technology

Volume 11 : Issue 1, June 2013

CONTENTS

Research Papers/Articles

Server Virtualization and its Impact on Business World

*Kapil Saxena**

ABSTRACT

Among the leading business challenges confronting CEO's and IT managers today are: cost-effective utilization of IT infrastructure, responsiveness in supporting new business initiatives, and flexibility in adapting to organizational changes. Server virtualization has come just in time for IT departments caught between the pressure to cut costs in the face of worldwide recession, steadily increasing energy costs and maxed-out data centers. Virtualization attacks the problem of the low utilization of single application servers. This paper describes the background behind server virtualization, its latest trends and different products used in server virtualization. This paper mainly focuses on impact of server virtualization on businesses in today's IT world.

Keywords: Confronting-Tackling, Utilization-Usage, Virtualization Imaginable, Responsiveness-Interest, Initiatives-Drive

1. Introduction

In the midst of IT adding and patching systems, global competition forced managements to look for ways to reduce costs and complexity, while maintaining or increasing service levels. Executives sought to make their businesses more nimble. They wanted to respond to fast changing global conditions. They sought to become more efficient to better compete with competitors, lower costs and improve shareholders' value. As a company's technology infrastructure costs slip out of control, server virtualization can become a compelling approach to getting control over server infrastructure.

1.1. Virtualization

The term virtualization broadly describes the separation of a resource or request for a service from the underlying physical delivery of that service. In other words, virtualization abstracts the underlying physical structure of various technologies. Virtualization, in computing, is the creation of a virtual (rather than actual) version of something, such as a hardware platform, operating system, memory, a storage device or network resource.

Simple example of virtual memory is Computer software gaining access to more memory than is Physically installed, via the background swapping of data to disk storage. Similarly, virtualization techniques can be applied to other IT infrastructure layers - including networks, storage, laptop or server hardware, operating systems and applications.

1.2 Server Virtualization

Server virtualization or hardware virtualization, enables multiple operating systems to run on a single physical machine. Server virtualization allows IT managers to use specific software to divide a single, physical server into multiple partitions or virtual servers and each acting as its own individual server. This blend of virtualization technologies or virtual infrastructure provides a layer of abstraction between computing, storage and networking hardware, and the applications running on it (see Figure 1). The deployment of virtual infrastructure is non-disruptive, since the user experiences are largely unchanged. However, virtual infrastructure gives administrators the advantage of managing pooled resources across the enterprise, allowing IT managers to be more responsive to dynamic organizational needs and to better leverage infrastructure investments. Server virtualization helps to maximize hardware use by aggregating more applications and services onto a fewer pieces of hardware, while maintaining operating system separation. So, server virtualization enables applications and services to safely coexist on the same server hardware, yet within multiple operating systems.

**Lecturer, School of Computer Applications Babu Banarsi Das University, Lucknow – U.P.*

Figures 1 and 2 show comparison between traditional servers and virtualized servers with respect to intelx86 architecture.

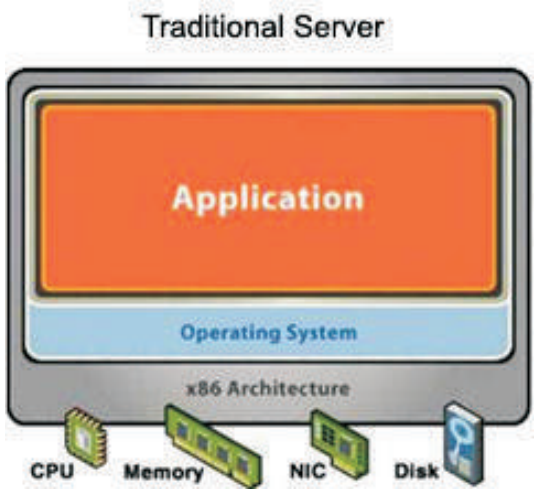

Figure 1. Virtualized Server

Before Virtualization:

- Single OS image per machine
- Software and hardware tightly coupled
- Running multiple applications on same machine often creates conflict
- Underutilized resources
- Inflexible and costly infrastructure.

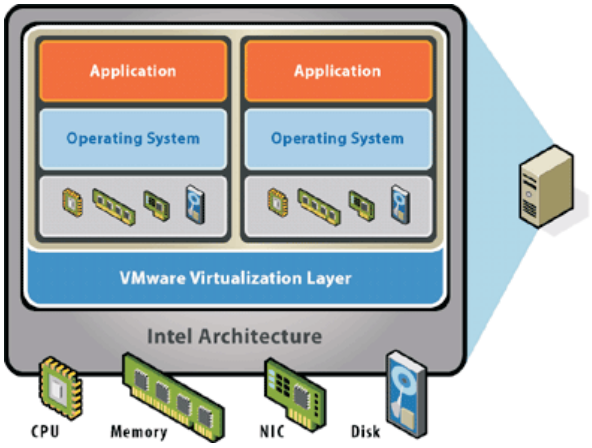

Figure 2. Virtualized Server

After Virtualization:

- Hardware-independence of operating system and applications.
- Virtual machines can be provisioned to any System.
- Can manage OS and application as a single Unit by encapsulating them into virtual machines.

2. Background and Effects of Server Virtualization

Server virtualization technique is used to construct the multiple virtual servers on one physical server. Virtual machines can be built by dividing physical servers into hardware and software. There are two advantages associated with hardware and software virtualization. When the physical server is divided by hardware, server virtualization offers an advantage of dividing hardware such that the error in one section does not affect other and when the physical layer is divided by hardware, server virtualization offers other advantages such as allowing CPU, I/O devices, memory and other hardware resources assign to virtual machines.

Given the recent background of a significant increase in application system construction costs (in terms of expenses, time, location etc), there are great expectations from server virtualization technology. In other words, it is urgently needed for changing business environment. In recent years, the server populations of many non-virtualized environments average about 20% utilization. The result is a huge waste of power, which is doubled since every kilowatt used has to be balanced with an equal amount of cooling to maintain the servers at optimal operating temperature. This also has grave implications for the lifespan of data centers as increasing number of facilities run short of power, cooling and, in some cases, floor space, despite the move to blade servers over the last few years. Virtualization avoids this problem by automating the management issues of stacking multiple applications on a single server and sharing resources among them. This allows IT shops to increase their server utilization up to 80%.The impact of this on a large organization has been demonstrated by BT in the UK. It has achieved a 15.1 consolidation of its 3000 Wintel servers, and save approximately 2 megawatt of power and $2.4 million in annual energy cost. This helped to reduce server maintenance cost by 90%.

Further, building multiple systems having low hardware resources usage rates (e.g CPU usage rate) on the same physical server will result in more effective use of server resources. The ratio resource distributed to low-load jobs and standby system can also be kept low as in usual operations, though the distribution rate may abruptly increase in case of a higher load or error occurrence. Thus server virtualization technology reduces the total cost of operations and makes rapid system configuration changes possible under changing conditions.

3. Types of Virtualization and Market Trends

There are many different types of virtualization technologies that make up the virtualization market. Some are more popular than others due to the expected

payback on the investment. Server virtualization targets the physical server machines in the data center. In general, 6-10 virtual machines are used per physical server depending upon the role of the server. Storage virtualization maintains a middle layer between the host and the physical storage environment. The middle layer's role is to make different storage (SAN) devices look the same to the host. The benefit is better storage utilization, faster provisioning and non-disruptive data migration. Application virtualization is a technique where a virtual environment is created on a local PC within which a software application runs without the need to utilize the supporting host resources. (i.e. file system, system registry, DLLs, etc.). The benefit of this method is that the underlying OS is untouched by the running applications thereby eliminating resource intensive configuration problems. An example of application virtualization would be Microsoft's Virtual PC 2007.

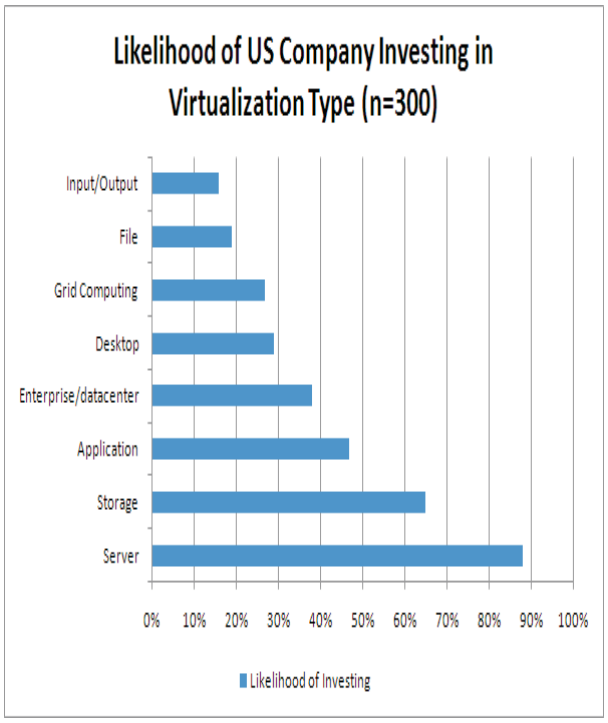

Figure 3. US company investment in virtualization techniques

Figure 3 shows the investment of US companies in various types of virtualization techniques. The US companies invest more on server virtualization rather than on storage and application. These three are the major areas of investment.

3.1 Virtual Technology Market

VT market, especially within the Server Virtual Machines (VMs) area, will continue its upward trend. In 2007, world-wide survey by Computer Associates shows that 88% (Figure-4) of the US companies interviewed are currently investing in virtualization technology and 12% are planning to invest (Computer Associates, 2008). According to a recent AP press release, the impact is already being felt in the server market where growth in sales for x86 servers is declining (AP, 2008). The research firm IDC, a premier global provider of market intelligence for information technology, has noted that the global server virtualization market is forecasted to expand at a compound annual growth rate of 14 percent approximately through 2014, at which time it will generate more than $14 billion in revenue.

The report rated the importance of each virtual technology to each company and by a large margin server virtualization was the most important. The research also indicted that the top four benefits of server virtualization are: easier hardware provisioning and software deployment, more flexible development and testing environments, optimizing system performance (load

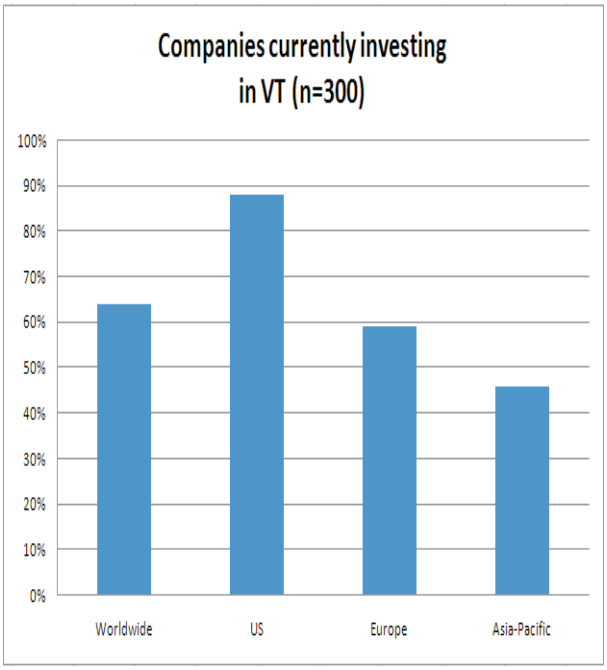

Figure 4. US Company Investment in 2004

balancing), and lower total cost of ownership.

4. Impact of Server Virtualization on Business World

Small to midsize businesses are increasingly using virtualization to reduce capital expenses, improve business continuity and to make their operations more responsive. High server performance and efficiency will continue to be major demands in the private sector, especially as the

business landscape grows more competitive, forcing organizations to be as productive as possible for the least amount of money. By leveraging server virtualization technologies, companies of all sizes can meet these requirements and potentially gain an edge over rival firms.

In this dynamic changing economic environment, the IT budgets of small to midsize businesses remain under pressure. The target for payback on IT investments appears to get higher and higher. However, some IT investments offer substantial opportunities for cost reduction, productivity improvements and enhanced business continuity. This can especially be the case with virtualization. And it's not simply a question of reduced expenses and resource requirements; virtualization can make a company more robust, agile and responsive to changing economic conditions and business opportunities. By utilizing technology properly, businesses can work to stay ahead of the competition and respond to the needs of their customers and partners. Further, the IT department can become more responsive towards business needs and begin to be seen as an innovator and driving new business initiatives instead of simply assisting the existing operations.

It is easy to say that the benefit of virtualization is high. By consolidation, virtualization server business can:

- Lower hardware cost and the associated cooling and space cost.
- Improve productivity across organizations and free up valuable IT time by simplifying IT infrastructure, which leaves additional time to focus on more strategic initiatives.
- Reduce costly downtime and streamline business contingency planning, so data would be secure in an event of natural disaster or other calamity.

These benefits of virtualization are briefly discussed below.

4.1 Reduction in costs

Server virtualization reduces the cost and complexity of business by encapsulating entire system files that can be replicated and restored onto any target server. Managing IT can be quite costly for some organizations, in terms of time and resources. By Virtualizing server infrastructure, it can help to lessen the hardware and maintenance costs, and lower company's energy bill.

a. Conserving energy

The biggest benefit of virtualization is to lower server infrastructure costs. With virtualization, you can consolidate excess server and desktop hardware, increasing utilization rates for x86 servers from 5 percent to 15 percent. With energy costs and global warming concerns rising, power consumption is an important issue for many businesses. Virtualization can help lower energy costs and lessen a company's CO_2 emissions.

Table 1. Power Cost

Power Cost as Percentage of Total Cost	
Non-virtualized Server	4.36%
Virtualized OS	1.68%

Conserving energy becomes increasingly essential in the future. In fact, as many as half the data centers in the world will soon face a shortage of cooling needs and energy capacity to deal with the newest, high-density computer equipment.

b. Server consolidation

Enterprises are now looking for ways to improve the average utilization of servers, reduce maintenance cost and also retain the Quality of Services. If a business currently uses one server per application, one can save on expensive floor space and help eliminate server sprawl by bringing together multiple applications onto a single server. This can reduce hardware and maintenance costs by as much as 50 percent. Virtualization technology can enable workload consolidation by providing all the required level of isolation with minimal loss in performance.

4.2 Efficiency and business continuity

In addition to cost savings, virtualization has other benefits, including improving staff productivity, business continuity and disaster recovery. It also enables your IT team to focus on more strategic projects that can help market critical products or services. A business attempts to remain competitive reducing the cost and complexity of business continuity (high availability and disaster recovery solutions) by encapsulating entire system into single file that can be replicated and restored on any target server, thus minimizing downtime.

a. Improve productivity

Because IT employees won't have to order and set up a new server for every new application, you will be able to get applications up and run them sooner. With fewer technical issues to manage, they can focus on strategic projects, such as improving customer service or developing new offerings. 73 percent of small to midsize businesses that have implemented virtualization reported seeing significant improvements on time spent on routine administrative tasks.

b. Protection of business from downtime and disaster

Natural disasters, malicious attacks, and even simple configuration problems like software conflicts can cripple services and applications until administrators resolve the problems and restore any backed up data. AMD's

Enhanced Virus Protection helps protect against certain viruses, worms, and other malicious attacks. Windows Server 2008 Hyper-V provides support for disaster recovery within IT environments and across data centers using geographical dispersed clustering capabilities. Traditional business continuity solutions are expensive and complex to deploy, putting them out of reach for many smaller organizations. Virtualization helps companies achieve faster and easier backup and recovery of key application workloads and data. It also enables companies to more cost-effectively switch to a secondary IT site and restore critical business operations.

c. Improve business responsiveness

Managing a virtual infrastructure allows IT professionals to quickly connect and manage resources to meet ever-changing business needs.

d. Resource provisioning and securing assets

Live Migration in conjunction with dynamic resource provisioning feature available in virtualization software can open up lots of possibilities that make an enterprise data center better able to handle varying transaction volumes. Application can be moved around along with the virtual machine using live migration to bigger servers if they are found to be choking under high load or more resources can be committed to suffering applications by shrinking the resources level of other virtual machines hosted on the same physical servers.

Secure company assets, rather than securing hardware. Businesses are securing data, no matter where it resides on the network. Virtualization can enhance a company's ability to increase security because the IT staff is able to apply security patches and move applications between virtual machines to avoid downtime. Since virtual machines reduce your server count, it also leaves your business less vulnerable to security attacks.

5. Server Virtualization Products

The key is to ensure that the virtualization software meets your corporate security policies and ensures regulatory compliance. Perhaps the most important factor to consider in virtualization purchase decision is the vendor you select. Many small to midsize businesses have chosen software on the basis of its reliability, ease of implementation and management, ability to deliver high-performance for applications, market leadership and lower operating costs. Some such products are discussed below.

5.1 VMware

VMware has a suite of virtualization product that is an aid in server consolidation. It has released virtualization solutions, namely, VMware Workstation, VMware ESX Server, VMware GSX Server, and VMware Virtual Center. VMware supports both windows and Linux as guest operating system. VMware enjoys many of the advantages of having pioneered the market and owning a dominant market share. For years, it was virtually the only game in town and thus benefits from the support of the broadest range of third-party software programs and systems for functions such as disaster recovery, lifecycle management and capacity planning. VMware also has the benefit of being generally regarded as the technology leader and pacesetter. In having the largest installed base, it has an inherent advantage of already being in place when companies are looking to expand their use of virtualization: VMware, in its 2010 corporate brochure, says its solutions are used by more than 97 percent of Fortune 1000 companies and 94 percent of Global 500 companies. There are many applications that are using VMware as virtualization. The application percentage using VMware is provided in Figure 5.

5.2 Microsoft

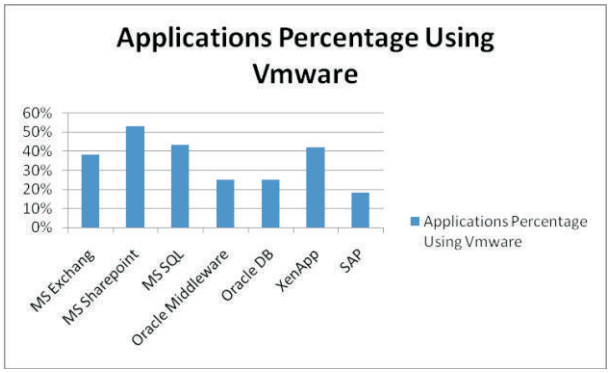

Figure 5. Application Percentage Using VMware

Customers that are new to virtualization and coming from Windows environments tend to view Hyper-V as a natural virtualization platform. Microsoft benefits from its ability to deliver familiar administrative tools and setup options. As a company coming in against an established market leader, Microsoft has taken an aggressive pricing approach by simply making Hyper-V R2 an available option within the setup of Windows Server 2008 R2. Since Windows Server virtual rights licensing is the same across any vendor's hypervisor, this results in Hyper-V being a generally less expensive solution to implement than VMware, at least for the up-front initial investment. However, the Microsoft solution includes more than just the integrated features in Windows Server. Management tools, which are branded under Microsoft System Center, provide management at all four layers of the IT infrastructure, from the physical hardware all the way to the application and services inside the virtual environments, in an easy-to-license suite. This provides a cost-effective way to build a

virtual environment without losing control and insight. There is definitely an advantage for some customers in purchasing solutions from a single vendor. For example, as part of Windows Server 2008 R2, Hyper-V R2 offers the same driver support for attached devices. The Enterprise and Datacenter editions of Windows Server 2008 R2 provide advanced features and functionalities for Hyper-V virtualization beyond the Standard edition. These include increased memory support, application failover, host clustering and dynamic data center. There are also different virtualization licensing benefits associated with the Enterprise and Datacenter editions, which is an area where an expert partner can provide guidance as to the overall costs and ROI of your Microsoft virtualization deployment.

Now we can make the comparison between VMware and MS Hyper-V in terms of total virtualization coast per host (Figure 6).

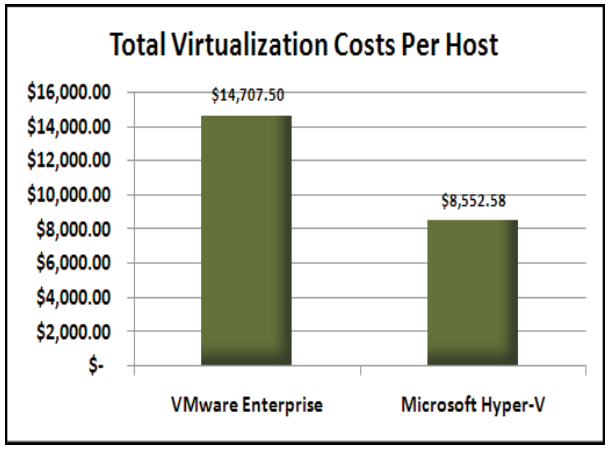

Figure 6. Total Virtualization Cost Per Host.

5.3 Citrix

Citrix benefits by having a solution that comes from the open source world. Its Xen hypervisor has an active community of software developers, ensuring many options, particularly those that organizations can build into their own virtualization solutions. Like Microsoft, Citrix positions itself as a far-less-costly alternative to VMware. XenServer 5.6 provides customers with the most robust and competitive free virtualization suite, compared with free products offered by VMware and Microsoft. Another advantage for Citrix is its strong position in the desktop virtualization market: Some customers will prefer using the same vendor for both server and client virtualization. according to Citrix, with core management features such as multiserver management, virtual machine templates, snapshots, shared storage support, resource pools and live migration. In addition, Citrix offers advanced management capabilities in Citrix Essentials for the XenServer product line. Customers can obtain XenServer for free, although they do need a license, and with the free version, they get access to features such as live migration, centralized management, VM template functionality and infrastructure update management as well as physical-server-to-virtual-machine conversions. Additional features are available for a fee in the Advanced, Enterprise and Platinum editions of XenServer. XenServer automatically restarts virtual machines if a failure occurs at the virtual machine, hypervisor or server level. The auto-restart capability allows administrators to protect all virtualized applications. Workload balancing is a feature of XenServer that captures data such as CPU, memory, disk I/O and network I/O on the hosts and virtual machines toguide the initial and ongoing host location for virtual machines.

The Comparison between Vmware Windows Server 2012 with Hyper-V (beta) and Citrix XenServer 6, is shown in Table 2.

Table 2: Comparison between Vmware, Hyper-V and Citrix

Attributes	VMware	Hyper-V (beta)	Citrix XenServer 6
Small Disk Footprint	✓ 144 MB disk footprint	● >5GB with Server Core installation	● >1GB
OS Independence	✓ No reliance on general purpose operating system	● Relies on Windows 2012 in Parent Partition	● Relies on Linux in Dom0 management Partition
Advanced Storage Management	✓ VMware vStorage VMFS	● Lacks an integrated cluster file system	● Lacks an integrated cluster file system, storage features support very few arrays
Advanced CPU Management	✓ Tuned to support Intel SMT hyper-threading; Supports 3D graphics accelerators	● No reliable performance advantage when using hyper-threading	● No reliable performance advantage when using hyper-threading
Flexible Resource Allocation	✓ Hot add VM vCPUs and memory, VMFS volume grow, hot extend virtual disks, hot add virtual disks	● Nothing comparable	● Nothing comparable

6.Conclusion

One of the biggest technologies to hit the market over the last few years has been virtualization. From a server perspective, virtualization breaks the bond between the operating system (OS) and the underlying hardware. Using virtual servers to act as redundant backup servers in replication-based high availability environments makes the deployment of this technology less burdensome from a financial perspective. Virtualized service technologies allow businesses of nearly any size to consolidate data and applications onto a fewer servers, doing so in less space, consuming less power, and usually for less money long term. Virtualization greatly simplifies how your company manages IT infrastructure; saving IT management time and increasing productivity through streamline and automation. The result is improved services to your company and clients, all while reducing operating cost and capital expenditures.

References

[1] "An Analysis of Server Virtualization Utility Incentives" Corban Lester, Lockheed Marteen http://www.thegreengrid.org/~/media/WhitePapers/Server%20Virtualization%20for%20Utilities_final.pdf?lang=en

[2] "Introducation to Virtualization" Morty Eisen http://www.ieee.li/pdf/viewgraphs/introduction_to_virtualization.pdf

[3] "Server Virtualization Technologies: Uses, Comparisons, and Implications". David Sweetman Windows Enterprise Systems Admin Administrative Information Services University of Michigan. www.windowshied.org/Conf2005/UMich_Virtualization_Testing.ppt

[4] "Making the Business Case for Virtualization" http://www.vmware.com/files/pdf/solutions/VMware-Business-Case-Virtualization-EN-WP.pdf

IPSec Issues over WiMAX Networks

*Monika rani**
*Shaweta***

ABSTRACT

Worldwide interoperability for Microwave Access (WiMAX) is a telecommunications technology providing wireless data, voice and videos over long distances with accuracy. The main goal of WiMAX is to deliver wireless communications with security.

In this study, different cryptographic techniques like the Advanced Encryption Standard (AES), Data Encryption Standard (DES), Triple DES (3DES), (DES+MD5), Message Digest (MD5) have been used for different packet size, and their performance has been evaluated on the basis of trade of the index. Further, these cryptographic techniques are applied for group communications in WiMAX networks. Several WiMAX scenarios have been studied on the basis of these techniques, which can be further extended for a large number of scenarios for studying the behavior of IPSec over WiMAX. QualNet 5.0 has been used for evaluating the performance of IPSec over WiMAX.

Keywords- Cryptography, IPSec, IKE

1. Introduction

In public key cryptography, each user or the device taking part in the communication has a pair of keys, a public key and a private key, along with a set of operations associated with the keys to do the cryptographic operations. Only the particular user/device knows the private key whereas the public key is distributed to all users/devices taking part in the communication. Since the knowledge of public key does not compromise the security of the algorithms, it can be easily exchanged online. Whereas in private key cryptography, there is the same key to do the cryptographic operation.

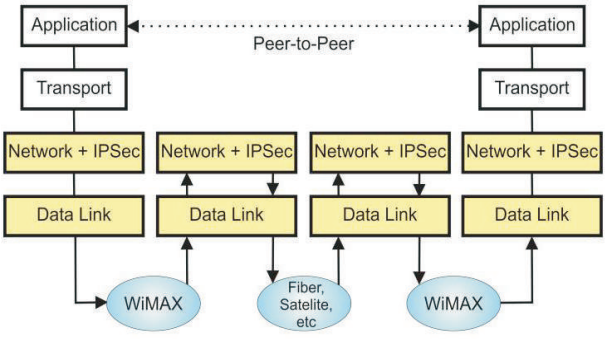

Figure 1. Wimax Scenario

Security can be defined as the protection of data being transmitted over a wireless network. In the past few years, the IEEE 802.16 working group has developed a number of standards for WiMAX

The first standard was published in 2001, which aims to support the communications in the 1066 GHz frequency band. In 2003, IEEE 802.16a was introduced to provide additional physical layer Specifications for the 211 GHz frequency band. These two standards were further revised in 2004 (IEEE 802.16-2004). Recently, IEEE 802.16e has also been approved as the official standard for mobile applications. Encryption algorithms which use the same key for both encryption and decryption are known as symmetric key algorithms. Our objective is to simulate the scenario shown in Figure1.

2. IPSec Basics

Internet Protocol Security (IPSec) is a protocol suite for securing Internet Protocol (IP) communications by authenticating and encrypting each IP packet of a communication session. IPSec is a developing network layer security mechanism. It protects traffic between endpoints at the network layer and it is totally independent from any application, that runs above the network layer [1].

*Department of Electronics and Communication Engineering, Deenbandhu Chhotu Ram University of Science and Technology, Murthal, Sonepat, Haryana131039.
**Department of Electronics and Communication Engineering, Deenbandhu Chhotu Ram University of Science and Technology, Murthal, Sonepat, Haryana131039.

The protocol allows the communicating nodes to set up secure channels to send and receive data. It also allows cryptographic algorithms to be applied and increase the security. Depending on the required security level for applications, different cryptographic algorithms may be applied. One cause of the complexity is that IPSec provides a mechanism, not policy: rather than define such-and-such encryption algorithm or a certain authentication function, it provides a framework that allows an implementation to provide nearly anything that both ends agree upon. There are several correlated parameters like:

AH versus ESP

Authentication Header (AH) and Encapsulating Security Payload (ESP) are the two main protocols used by IPSec, and they authenticate (AH) and encrypt+authenticate (ESP) the data flowing over that connection.

Tunnel Mode versus Transport Mode

Transport Mode provides a secure connection between two endpoints as it encapsulates IP's payload, while Tunnel Mode encapsulates the entire IP packet to provide a virtual "secure hop" between two gateways.

IKE versus Manual Keys

Since both sides of the conversation need to know the secret values used in hashing or encryption, there is query question of just how this data is exchanged. Manual keys require manual entry of the secret values on both the ends, presumably conveyed by some out-of-hand mechanism, and Internet Key Exchange (IKE) is a sophisticated mechanism for doing this online.

The Internet has a great many resources surrounding IPSec, some better than others. The starting point, of course, is always with the Requests for Comment (RFCs) that form the Internet standards defining the protocols. These are the main reference works on which all other documentation including this one is based. Some of the RFCs are: RFC 2401, RFC 2403, RFC 4301, RFC 4302 etc. [6].

3. Methodology

"Trade of Index" is a parameter used to calculate the efficiency. In this study, so many cases have been taken to find trade of index by varying the packet sizes and cryptographic algorithms. After that, average trade of index is calculated. Also different scenarios have been taken for selecting the best technique for encryption. The techniques used are: AES,DES,3DES, MD5, and DES+MD5. AES is used widely because the algorithm is fast in both software and hardware, easy to implement and requires little memory [5].

The DES algorithm [4] is a symmetric block cipher with block and key size of 64 bits. DES has been proven not a reliable cryptographic scheme as special hardware can break DES in a few hours. This has been the reason to

Table 1. Processing Times for 100 MIPS [3]

PKT Size	AES	MD5	3 DES	DES
500	2.035	0.092	5.340	1.780
600	2.405	0.105	6.310	2.100
700	2.837	0.112	7.362	2.454
800	3.207	0.126	8.330	2.778

issue 6, June 2004, pp. 1013.

[3] A. Mishra and N. Glore, "Privacy and Security in WiMAX Networks", Book Chapter of "WiMAX Standards and Security", CRC Press, 2008

[4] Xenakis, N. Laoutaris, L. Merakos and I. Stavrakakis, "A generic characterization of the overheads imposed by IPsec and associated cryptographic algorithms", Elsevier Computer Networks, 2006

[5] http:/www.wimaxforum.com/asymmetric keys, http://www.wimaxtrends.com

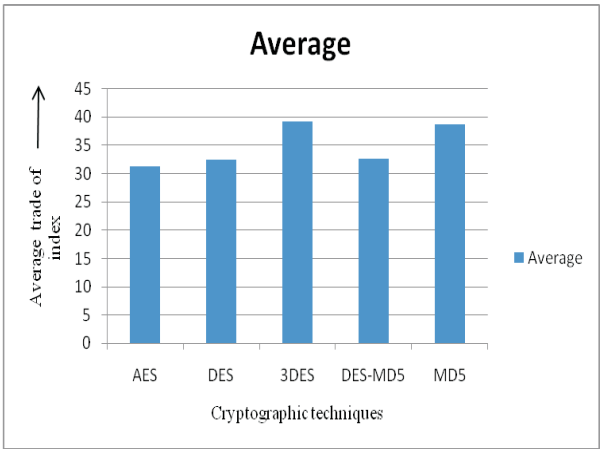

Figure 2. Performance of 13SS+3BS+2CBR

[6] http:///www.ietf.org/rfc

[7] http://users.encs.concordia.ca/~dongyu/ELEC6851/Qualnet_Tutorial_5.0.pdf

[8] http://www.unixwiz.net/techtips/iguide-ipsec.html

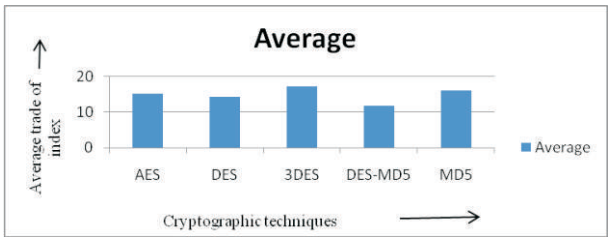

Figure 3. Performance of 13SS+3BS+3CBR

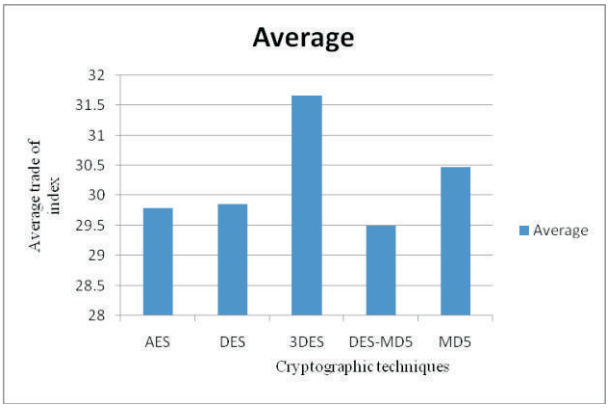

Figure 4. Performance of 20SS+2BS+1CBR

Investigating Attack Models against Navigation Oriented Personalization

*Raj Gaurang Tiwari**

ABSTRACT

Every user has a distinct background and a specific goal when searching for information on the Web. User profiles resulting from Web navigation data are used in important e-commerce applications such as Web personalization, recommended systems, and Web analytics. In the open environment of the Internet, attackers can use automated means to inject large number of biased profiles into recommender systems. Recent research has concluded that systems using explicit ratings input by users are highly vulnerable to such "profile injection" attacks. Malicious users can cause certain products to be recommended more frequently and others less frequently. It is shown that Web recommenders that use implicit Web navigation profiles to learn user preference models, despite using different algorithms than traditional recommenders based on explicit ratings, are subject to similar manipulation. There is an impact of "crawling attacks" against navigation-based Web recommender systems. A crawling attack is comprised of a set of user profiles that a rogue agent may inject into the click-stream navigation data by crawling the site in such a way that it would change the future behavior of the system. In this paper, we have examined different attack types and shown that these are effective against the most common personalization algorithm.

Keywords: Web recommendation, Crawling attacks, Web user profiles, Click-stream navigation, Web usage mining

1. Introduction

Due to the explosive growth of the Web, Web personalization and recommender systems have gained popularity, helping people to find the information they want, and allowing Web site owners to optimize their sites and increase user satisfaction. In most E-commerce recommender systems, the user provides inputs and the system processes the information to generate a list of recommendations. The input, indicating the user's preferences on items can be explicit ratings of items or derived based on the implicit indications of interest such as the user's behavior during navigation. An increasing number of Web recommender systems today make use of implicit ratings where Web user's interests are captured during the interaction with the Web site by navigating through a sequence of pages. Web personalization systems will identify the user's interest in individual or groups of items, based on some measures such as whether an item is purchased or not, time spent on viewing a page or item, etc. The goal of personalization based on Web usage mining is to recommend a set of objects to the current (active) user, possibly consisting of links, ads, text, products, or services, tailored to the user's preferences. A recommender system which uses explicit ratings, requires users to create some sort of account. However, a determined attacker may be able to outwit schemes designed to prevent automated account registration. Definitely, the cost of generating new profile is significant. On the other hand, navigation-based recommender systems use implicit feedback captured in the click-stream data and are mainly dependent on the navigation profiles of anonymous users who visit pages in a particular order or combination. Web recommendation is typically performed on log data, which is not associated with user accounts. Thus, an attacker needs only successfully disguise his

**SRMCEM, Lucknow.*

automated site crawler as a large number of different legitimate Web clients, something easily achieved through anonymized browsing techniques.

An attacker, who could inject such navigation profiles by visiting a combination of items often enough, may produce any recommendation behavior for future users that he or she desires. In this paper, there is an investigation of the effectiveness of various attack models against navigation-oriented (also called "usage-based") personalization algorithms. We focus specifically on two standard approaches for generating recommendation from click-stream and Web navigation data, namely, the k-Nearest-Neighbor (kNN) algorithm, commonly used in collaborative filtering, and an algorithm based on Markov models.

2. Attack Types

A profile injection attack comprises of a set of attack profiles being inserted into the system with the aim of altering the system's recommendation behavior with respect to a single target item.

In the context of Web usage data, an attacker's aim is to inser… injecting fake navig… an easily be gene… e crawling mech… i_1 … i_k^S i_1^F … i_l^F i_t insert such navig… a particular order… called as "crav… s partitioned in thr… t page it will alwa… g promoted.

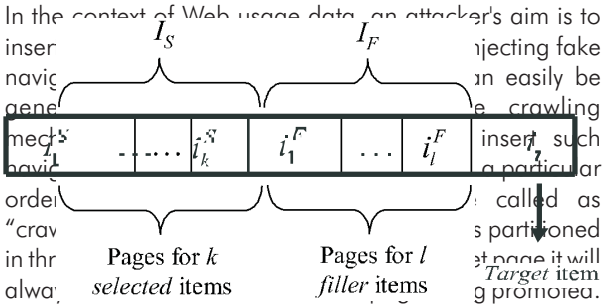

Figure 1: The general form of an attack profile

As described below, some attacks require identifying a group of pages with desirable characteristics. This set is symbolized by IS. IF is a set of pages chosen randomly to fill the rest of the profile. The strategy for selecting pages in IS and IF defines the characteristics of an attack model.

In our work, we specially focused on two types of attacks: the bandwagon and segment. In this paper, we have developed different versions of attack models that are particularly suited for manipulating navigation oriented web recommendation.

2.1 Bandwagon attack : The bandwagon attack uses a small amount of additional knowledge, namely the identification of a few of the most popular items in a particular domain: blockbuster movies, top-selling recordings, etc. This information is easy to obtain and not dependent on any specifics of the system under attack. We introduce popular page attack which is an adaptation of the bandwagon attack in context of Web usage data. The attacker identifies a few of the most popular pages in a particular Web site. A custom crawler generates repeated visits to those popular pages and to the target page. Over time, such visits will generate a large number of profiles that associate the target page with the popular pages. For many sites, the most popular pages will be relatively obvious: for example, the site's home page. In this attack model, the set IS contains these popular pages selected to be part of attack profiles. In our experiments, we use the top two most frequently accessed pages as the contents of IS. The set IF is a set of pages chosen randomly to fill the rest of the profile.

2.2 Segment attack : The goal of the segment attack is to maximize the similarity between the target item and the segment items in IS. The segment items are those well-liked by the market segment to which the target item it is aimed. In the case of explicit ratings data, the items in IS are given high ratings to increase the similarity between them and the target item; the filler items are given low ratings, to decrease the similarity between these items and the target item. This attack proves to be highly effective against both user-based and item-based collaborative recommendation [3].

We developed a version of the segment attack for Web usage data. This attack is designed to promote a page to a targeted group of users with known preferences. The goal of this attack is to maximize the similarity between the attack profiles and user profiles containing a group of pages or items well liked by the market segment to which the target item is aimed. In this scenario, segments are dependent on preferences based on some navigational patterns. Users interested in these segment pages would then be highly likely to get the target page as a recommendation.

3. Recommendation Algorithms

The overall process of personalization based on Web usage mining consists of three phases, namely data preprocessing, pattern discovery, and recommendation. In the data preprocessing phase the raw click-stream data is transformed into a set of user profiles. Raw Web log data is cleaned and sessionized to generate user sessions. Each user session is a logical representation of a user's single visit to the Web site (usually within certain time interval). The output of this phase is a set of n pages, $P = \{p_1, p_2, \cdots, p_n\}$, and a set of m user sessions $U = u_1, u_2, \cdots, u_m$, where each $u_i \in U$ consists of pages from a subset of P. The Web session data can be conceptually viewed as a $m \times n$ session-page matrix $UP = [w(u_i, p_i)]_{m \times n}$, where $w(u_i, p_i)$ represents the weight of page p_i in user session u_i. The weights can be binary, indicating the existence or non-existence of the page in the session, or they may be a function of the occurrences or duration of the page in that session. Each user session u can also be viewed as a l-length sequence of ordered pairs:

$u = \{(p_1^u, w(p_1^u)), (p_2^u, w(p_2^u)), \ldots\ldots\ldots, (p_i^u, w(p_i^u))\}$

where each $p_i^u = p_i$ and $w(p_i^u) = w(u, p_i)$, for some $j \in \{1, \cdots, n\}$.

Given a set of user profiles or sessions as described above, a variety of unsupervised knowledge discovery techniques can be applied to obtain patterns. Techniques such as clustering of users (or sessions) can lead to the discovery of important user or visitor segments. Other techniques such as item (e.g., page) clustering, association rule mining, or sequential pattern discovery, can be used to find important relationships among items based on the navigational patterns of users in the site. In the cases of clustering and association rule discovery, generally, the ordering relation among the pages is not taken into account, thus a user profile is viewed as a set (or, more generally, as a bag) of pages. In the case of sequential patterns, however, we need to preserve the ordering relationship among the pages within transactions, in order to effectively model users' navigational patterns, and thus the sequence representation is used as input to the pattern discovery phase.

In this paper, we have concentrated on the two algorithms: the user-based *k*NN recommendation algorithm and a recommendation algorithm based on Markov chains. We have described these algorithms in more detail in the following sections.

3.1. kNN-Based Algorithm

Collaborative filtering based on *k*-nearest-Neighbor (*k*NN) approach involves comparing the active record for a target user with historical records of other users in order to find the top *k* users who have similar tastes or interests. The mapping of a visitor record to its neighborhood could be based on similarity in ratings of items, access to similar contents or pages, or purchases of similar items. This neighborhood is then used to recommend items not already accessed or purchased by the active user.

In the context of personalization based on click-stream data, *k*NN involves measuring the similarity or correlation between the active session $\mathbf{S}$ and each session $\mathbf{T}$ in historical records. The top k most similar users to $\mathbf{S}$ are considered to be the neighborhood for the active session, $\mathbf{S}$ denoted by NB(s). S many measures can be used to find the nearest neighbors. In traditional collaborative filtering domains (where feature weights are item ratings on a discrete scale), the Pearson r correlation coefficient is commonly used. This measure is based on the deviations of users' ratings on various items from their mean ratings on all rated items. However, this measure may not be appropriate when the primary data source is click-stream data (particularly in the case of binary weights). Instead the cosine coefficient, commonly used in information retrieval, were used which measures the cosine of the angle between two vectors. The cosine coefficient can be computed by normalizing the dot product of the two vectors with respect to their vector norms. Given the active session $\mathbf{S}$ and a session $\mathbf{T}$ in historical records, the similarity between them is obtained by:

$$sim(\mathrm{t},\ \mathrm{s}) = (\mathrm{t},\ \mathrm{s})/(\mathrm{t} \mid \mathrm{s}^{|\times|}|).$$

In order to determine which items (not already visited by the user in the active session) are to be recommended, a recommendation score is computed for each item pi $\in$ P, based on the neighborhood NB(s) for the active session, where P is the set of all pages. The recommendation score for a page p with respect to NB(s) is computed by :

$$Score(p, NB(s)) = \frac{\sum_{t \in NB(s)} w(t, p)}{|NB(s)|}$$

where w(*t*, *p*) is the weight for the page *p* in the active session *t*. In our experiments we use binary weights for pages, indicating whether a document is accessed or not in user's session. If a fixed number *N* of recommendations is considered, then the top *N* items with the highest recommendation scores are considered to be part of the recommendation set.

3.2 Markov Model Algorithm

A Markov model is represented by the 3-tuple [A, S, T], where A is a set of possible actions, S is the set of n states for which the model is built and T is the Transition Probability Matrix (TPM) that stores the probability of performing an action. It should be mentioned that s and S is described above. Specifically, $T = [P_{ij}]_{n \times n}$, where P_{ij} represents the probability of a transition from state S_i to state S_j . The order of the Markov model corresponds to the number of prior events used in predicting a future event. So, a k^{th}-order Markov model predicts the probability of the next event by looking at the past k events. The simplest Markov model predicts the next action by only looking at the last action performed by the user. Given a set of all paths R, the probability of reaching a state sj from a state si via a (non-cyclic) path r ε R is given by: $p(r) = \prod P_{k,k+1}$, where k ranges from i to **J**. The probability of reaching S_j from S_i is the sum over all paths:

$$p(j \mid i) = \sum_{r} p(r)$$

Markov model can also be used to discover high probability user navigational paths in a Web site. In the context of recommendation systems, *A* is the set of items and *S* is the visitor's navigation history, defined as a *k*-tuple of items visited, where *k* is the order of the Markov model. The input data for building the Markov model consists of Web sessions, where each session represents the sequence of the pages accessed by the user during his/her visit to the site.

Each state in the Markov model corresponds to all sub-sessions of length *k* which were observed in the data, where *k is* the order of Markov model. In the case of first order models, the states will correspond to single page and in the case of second order models the states will correspond to all pairs of consecutive pages and so on.

Once the states of Markov model are built, the TPM can then be computed. The most commonly used approach for estimating the probabilities in the TPM is to use a training set of action-sequences, and each $p_{ij,}$ is estimated based on the frequency of the event that action a_i follows the state s_j. Making guess for Web session is complicated. For example, consider a user who has accessed pages P_1, P_5 and P_4. If we want to predict the next page that will be accessed by the user, we will first identify the state s_4, that is associated with page P_4 and then look up the TPM to find the page p_i that has the highest associated probability.

4. Experimental Evaluation

4.1. Evaluation Metrics

We use a measure hit ratio in the context of top-N recommendations which has commonly been used to evaluate predictions based on Web usage mining models [14]. For each user session in the evaluation set, we input the first j pages, as a surrogate for a user's active session, into the recommender system and generate a recommendation set. The value j is the window size for the experiments. We then compare the recommendation set with the (J+1) page the next page that the user actually visited. If the page (J+1) appears in the recommendation set, we call it as a "hit". We define the hit ratio as the total number of hits divided by the total number of sessions in the evaluation set.

In measuring the impact of an attack, we are interested not in raw performance how well the recommender does its job but rather in the change in performance induced by an attack. In a "push" attack, the attacker will desire that the pushed item will be more likely to be recommended after the attack than before. We measure this benefit to the attack with the metric target hit ratio. The same operation is performed as for the hit ratio metric described above, but instead of measuring the frequency of selection of the next item the user actually visited, we measure how frequently the target item is recommended. Let R_u be the set of top N recommendations for user u. For each push attack on a target item i, the value of a hit for user u denoted by H_{ui}, can be evaluated as 1 if $i \in R_u$; otherwise, it is zero. We define target hit ratio as the number of hits across all user sessions divided by the total number of user sessions in the evaluation set, computed as:

The average target hit ratio can then be calculated as the sum of the hit ratios for attacks on each item i across all items divided by the total number of target items in the test date. In our experiments, we report the average of this value across all test user sessions. Of course, a measure of the impact of the attack is also dependent on what the Target Hit Ratio would have been in the absence of the attack profiles. In most cases, the pre-attack value is vanishingly small as the results below show.

We first compute the average hit ratio or average target hit ratio for the current user within this active session window and finally compute the overall average for all test sessions. $Target\ HitRatio_i \quad H_{ui}/|U|$

$u\ U$

4.2. Experimental Methodology

For our experiments, data based on the server logs of the host Computer Science department spanning a one-month period are used. This data set is refered as the CTI data. Standard preprocessing techniques [6] were applied to clean the data, remove spider references, and sessionize the data. This data set is filtered to include only sessions of size 6 or more. The initial preprocessed data set contained more than 100,000 sessions and over 4,000 pages. Further aggregation was performed to "roll up" low-support (infrequently accessed) pages to their common root node in the site hierarchy. The final data set spanned approximately 700 aggregated page views, representing access to items (or item categories in the case of aggregated references). This data set was randomly divided into training and evaluation sets. The training set of user sessions was used to build the models while the test set consisting of more than 4,000 user sessions was used to evaluate the recommendations generated by the models. To simulate segment attacks in which an item is pushed by creating strong associations with a segment of users who show interest in a particular group of items. Here, we considered two scenarios: one in which a particular faculty page would be pushed to all users showing an interest in any of the faculty-related pages, and one in which a particular course would be pushed to all of the visitors looking at any of the course-related pages. Our evaluation set consists of those users whose session included any one of these segment pages within the first six page-views. There were 525 and 993 user sessions in the evaluation sets for the faculty-segment and course-segment respectively.

As attack targets, a set of pages is used, consisting of 25 pages each in the faculty and course areas. These were chosen to be specific courses or faculty home pages, not top-level navigation pages. Each of these target pages was attacked individually and the results reported below represent averages over targets and all sessions in test data set. For all the attacks, a number of attack profiles are generated and inserted into the system database and then

recommendations are generated and "size of attack" as a percentage of the pre-attack user count is measured. For example, if the training set contains 100,000 user sessions, an attack size of 1% corresponds to 1,000 attack profiles added to the system.

Considering Figure 1, the target page i_t is the one that an attacker would like to promote. The set I_S is the set of "special" pages chosen based on some property. Finally, I_F is a set of pages chosen randomly to fill the rest of the profile.

The purpose of Popular Page attack is to promote a target page along with a few popular pages. The popular pages could be a few of the most frequently accessed pages. In this attack, the set I_S consists of two top most frequently accessed pages and the filler item set I_F is chosen randomly. One of the characteristics of the Bandwagon attack found in previous work is that popular items could be easily discovered from third-party sources and required no system-specific knowledge. This version of the attack requires the attack to know what pages will be most visited. It is assume that in most sites the popular pages will be fairly obvious from site structure.

In our segment attack model, the segment items are those well-liked by the market segment to which the target item i_t is aimed. Two user segments are chosen namely, faculty and course. To generate segment attacks, we set I_S to contain a sample of appropriately-related pages. There are no stuffing items, so I_F is null.

4.3. Experimental Results

We first compare the accuracy of the *k*NN and Markov algorithms by measuring hit ratio. The window size is set to 3, small enough so that user can get recommendations after visiting first few pages. As Figure 2 shows, the Markov model is significantly more accurate than *k*NN. Even at lower number of recommendations, Markov model performs better than *k*NN. The results shown here are not surprising, because the Markov model is well suited for predicting the sequential process of Web navigation.

The experimental results in Figures 3 and 4 represent the target hit ratio for two attack types. In the case of the *k*NN algorithm, the target hit ratio is relatively flat across different recommendation set sizes, so only a single set size is shown. This was not the case for the Markov algorithm so a range of different retrieval sets are displayed. This would seem to indicate that the attack has a relatively scale-free impact on the *k*NN algorithm: as the retrieval set gets larger the attacked item is represented in the same proportion. For the *k*NN algorithm, the value of the Target Hit Ratio for the target items was essentially zero prior to the attack. The (small) pre-attack baseline is shown in the Markov graph.

In the case of the Markov algorithm, the Popular Page

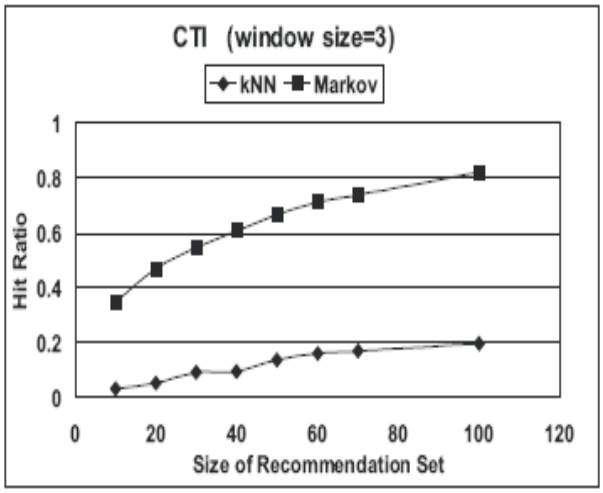

Figure 2. The hit ratio results for CTI Site.

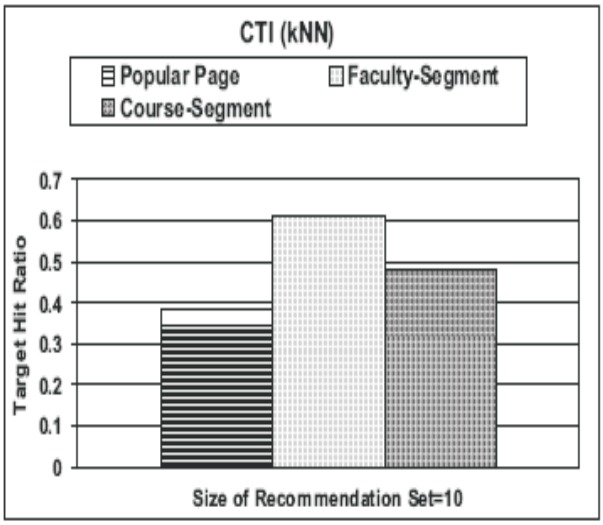

Figure 3. The objective hit ratio results for CTI Site in kNN-based algorithm. Window Size=3, Attack Size=10%

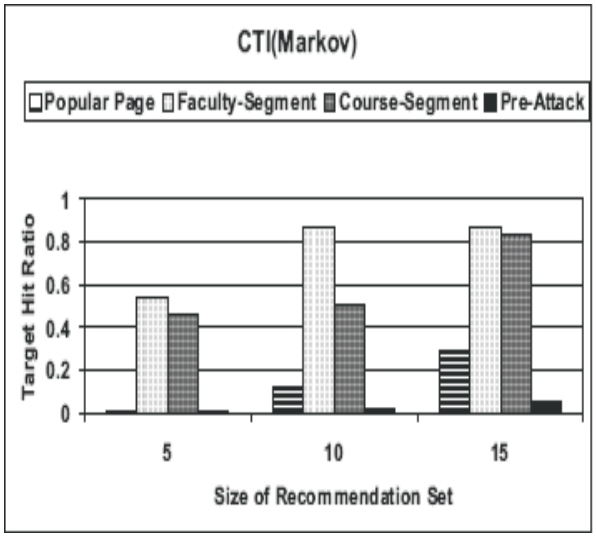

Figure 4. The target hit ratio results for CTI Site in Markov model algorithm.Window Size =3, Attack Size = 10%

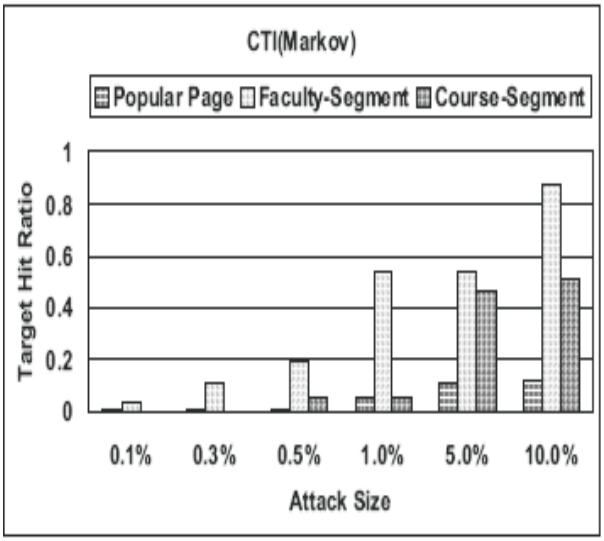

Figure 5. The objective hit ratio results for CTI Site in Markov model algorithm Window Size=3, Recommendation Size=10

attack has relatively low impact at all retrieval set sizes when compared to the segment attack. Interestingly, the segment attack against the course pages is here less effective than against the faculty pages. This may be due to the differences in the sparsity of page views within these segments. In general, faculty pages are far less frequently visited than course related pages. Thus, smaller attacks may have a bigger impact on the models derived from navigational sessions involving faculty pages.

5. Conclusions and Future Work

Researchers have established that collaborative filtering systems using explicit ratings input are extremely vulnerable to "profile injection" attacks. In this paper, our empirical results show that systems using Web usage data are also vulnerable to biases injected by attackers who crawl the site in particular ways. It is shown that the most common types of personalization algorithms based on Web navigation data, like Markov models, can be easily manipulated to push individual target Web pages. We plan to examine other Web usage mining algorithms such as association rule discovery, sequential pattern mining, and unsupervised clustering approaches in our future research. In addition to looking for vulnerabilities, we also plan to develop effective approaches to the detection and prevention of such attacks.

References

[1] J. Borges and M. Levene. Data mining of user navigation patterns. In B. Masand and M. Spiliopoulou, editors, Web Usage Analysis and User Profiling: Proceedings of the WEBKDD'99 Workshop, LNAI 1836, pages 92111. Springer-Verlag, 1999.

[2] R. Burke, B. Mobasher, and R. Bhaumik. Limited knowledge shilling attacks in collaborative filtering systems. In Proceedings of the 3rd IJCAI Workshop in Intelligent Techniques for Personalization, Edinburgh, Scotland, August 2005.

[3] R. Burke, B. Mobasher, C. Williams, and R. Bhaumik. Segment- based injection attacks against collaborative filtering recommender systems. In Proceedings of the International Conference on Data Mining (ICDM 2005), Houston, December 2005.

[4] R. Burke, B. Mobasher, R. Zabicki, and R. Bhaumik. Identifying attack models for secure recommendation. In Beyond Personalization: A Workshop on the Next Generation of Recommender Systems, San Diego, California, January 2005.

[5] R. Cooley, B. Mobasher, and J. Srivastava. Web mining: Information and pattern discovery on the world wide web. In Proceedings of the 9th IEEE International Conference on Tools with Artificial Intelligence (ICTAI'97), pages 558567, Newport Beach, CA, November 1997.

[6] R. Cooley, B. Mobasher, and J. Srivastava. Data preparation for mining World Wide Web browsing patterns. Journal of Knowledge and Information Systems, 1(1):532, 1999.

[7] M. Deshpande and G. Karypis. Selective markov models for predicting web-page accesses. In Proceedings of the First International SIAM Conference on Data Mining, April 2001.

[8] S. Lam and J. Riedl. Shilling recommender systems for fun and profit. In Proceedings of the 13th International WWW Conference, New York, May 2004.

[9] B. Mobasher. Web usage mining and personalization. In M. P. Singh, editor, Practical Handbook of Internet Computing. CRC Press, 2005.

[10] B. Mobasher. Data mining for personalization. In P. Brusilovsky, A Kobsa, and W. Nejdl, editors, The Adaptive Web: Methods and Strategies of Web Personalization, volume 4321 of Lecture Notes in Computer Science. Springer-Verlag, Berlin Heidelberg New York, 2007.

[11] B. Mobasher, R. Burke, R. Bhaumik, and C. Williams. Effective attack models for shilling item-

based collaborative filtering systems. In Proceedings of the 2005 WebKDD Workshop, held in conjuction with ACM SIGKDD'2005, Chicago, Illinois, August 2005.

[12] B. Mobasher, R. Burke, and J. Sandvig. Model-based collaborative filtering as a defense against profile injection attacks. In Proceedings of the 21st National Conference on Artificial Intelligence, page to appear. AAAI, July 2006.

[13] M. O'Mahony, N. Hurley, N. Kushmerick, and G. Silvestre. Collabo- rative recommendation: A robustness analysis. ACM Transactions on Internet Technology , 4(4):344377, 2004.

[14] Y. Zhou, X. Jin, and B. Mobasher. A recommendation model based on latent principal factors in web navigation data. In In Proceedings of the 3rd International Workshop on Web Dynamics, Held at the WWW 2004 Conference, New York,2004.

An Approach to Visual Cryptography

*Dr. Qaim Mehdi Rizvi**
*Dr. A.A. Karawia***

ABSTRACT

This paper proposes a new type of visual cryptographic scheme i.e. 'four way visual cryptography technique' which can encrypt grayscale images of square dimensions without the use of any external media or any cryptographic computations. This scheme is perfectly secure and easy to implement which does not make any use of major system resources. It makes use of basic matrix operations that makes the encrypted image visually infeasible to decode.

Keywords- Zig-Zag Traversal, Matrix Rotation, Four Way, Visual Cryptography(VC).

1. Introduction

Visual cryptography (VC) is an image encryption scheme proposed by Naor and Shamir in 1994 [1]. This scheme used k out of n shares scheme of visual cryptography. The image was encoded into n shares which are meaningless images in itself, at least k shares were required to decode the image, less than k shares were useless as they did not reveal the image. This scheme made use of transparencies. VC introduced by Naor, [1] is a method used for encrypting an image into shares such that stacking reveals the secret image [2]. VC scheme is an encryption method that uses combinatorial techniques to encode secret written material[2]. In VC, meaningless shares are encoded into halftone shares taking meaningful visual information which reduces the suspicion of intruders. These halftone shares are error diffused to give visually pleasing effect, a novel technique named half visual cryptography is proposed to achieve visual cryptography via half toning [3]. Ateniese has formalized this framework as the Extended Visual Cryptography and developed a scheme for general access structures [4]. The main point of consideration was that visual cryptography does not make use of any cryptographic algorithms. Similarly, we are also using a set of algorithms which does not involve any sort of cryptographic computation. There were certain algorithms used previously for this technique [5]. There also emerged an idea of encoding multiple secret images into the same share image [6]. Randomization and pixel reversal approach were used to retain the size of the image [8]. This scheme had applications in authentication and identification and copyright protection and watermarking [3]. In watermarking a mark is split into two shares images [7].

Now a days data used digitally is vulnerable to various risks. So, there is the introduction of new forms of visual cryptography which is providing encryption of an image and its transfer securely through any communication channel. Our technique four way visual cryptography has a wide scope.

This paper does not make any use of technique to improve the image quality as there will be a negligible loss in the quality of the image unlike error diffusion [3] which is the method to produce high quality images with less computation cost. In this paper, a new form of visual cryptography is introduced which moves around the square matrix generated from the pixels of a grayscale image and make use of the Zig-Zag traversing scheme [9]. This scheme, firstly, traverses the randomly generated image matrix in a Zig-Zag manner using an algorithm which provides a Zig-Zag traversed matrix as output via one dimensional array. The matrix obtained is further rotated anticlockwise by 90 degrees. The resultant matrix is again traversed and rotated. This is repeated four times which results in a matrix that is encrypted and can be transmitted over a communication channel without an increase in the size of the original image. Since, the process is happening four times, hence, it is named as four way visual cryptographic technique. This scheme encrypts the image which can be transmitted and decrypted digitally without the use of any external media.

*CS Department, QASS IM University, Buraidah, KSA.
** Mathematics Department, Mansowa University, Egypt.

2. Proposed Work

FOUR WAY VISUAL CRYTOGRAPHY makes use of MATLAB software to encrypt the image. It requires certain algorithms to encrypt the image. The first algorithm used is for Zig-Zag traversing the square matrix generated from the image input. Consequently, one dimensional array is produced which is converted to 2D matrix by using the next algorithm. Finally, the obtained matrix is rotated anticlockwise by 90 degree using MATLAB function. The above process is performed four times and finally encrypted image is produced. This image is decrypted using another set of algorithms, which includes sorting and other functionalities along with clockwise rotation by 90 degree. The image obtained is similar in quality because the original image and the size is also not altered which are the advantages over the previous schemes. These algorithms are described in the next section.

3. Framework of Encryption

Figure 1 shows framework for encryption of grayscale image given as input. The first block takes the original grayscale image matrix as input. The matrix is then Zig-Zag traversed as shown in the second block. After Zig-Zag traversal, a 1D array is generated of the length N*N where N x N is the dimension of the original matrix. The array is converted back to 2D matrix so that the encrypted image is of same size and form as the original. This matrix is then rotated by 90 degree in anti- clockwise direction to increase the complexity.

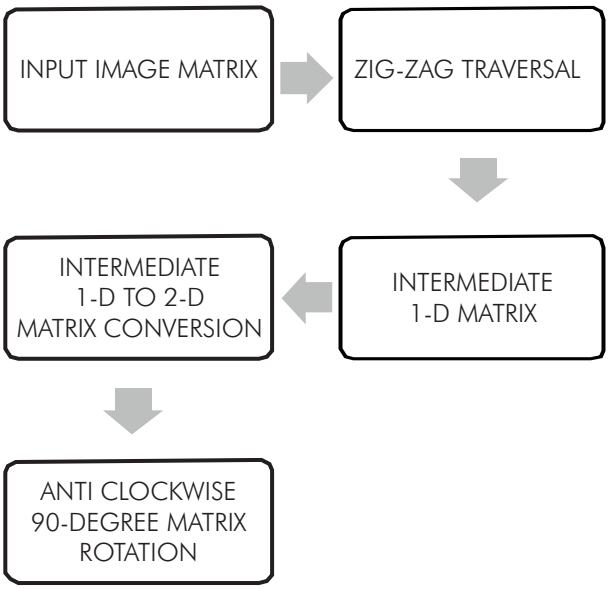

Figure 1:Framework for Encryption

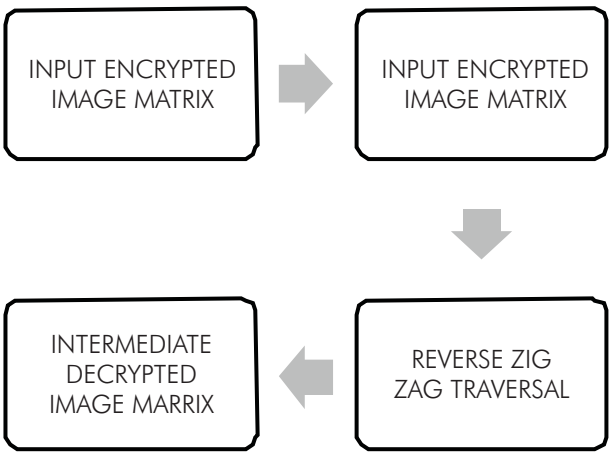

Figure 2. Framework of Decryption

The final image after this process is again given as an input and the process stated is carried out for the next time. In this way, the process is repeated four times. The encrypted image appears as if noise is introduced on the image uniformly which makes it difficult to identify the image. It appears same as the shares in visual cryptography

4. Concept of Decryption

The decryption system takes encrypted image as an input. It makes use of certain algorithms which reverses the concept of encryption thus resulting in the decrypted image. The image obtained is same in size and quality, as the original image. It involves matrix operations such as clockwise rotation of image matrix and concept of reverse Zig-Zag traversal of rotated matrix.

5. Framework of Decryption

Figure 2 shows framework for decryption of encrypted image given as input. The first block takes the encrypted image matrix as input. The matrix is then clockwise rotated by 90 degree as shown in the second block. After clockwise rotation the matrix is then travelled Zig-Zag in reverse fashion. The final image which we get after this much process is again given as input and the process stated is carried out for the next time. The above process is carried out four times. Hence, image obtained is of the same size and quality which is the advantage over the previously known visual cryptography techniques.

Figure 3 compares the riginal image with the finally decrypted image. In the middle, image shown is an encrypted image.

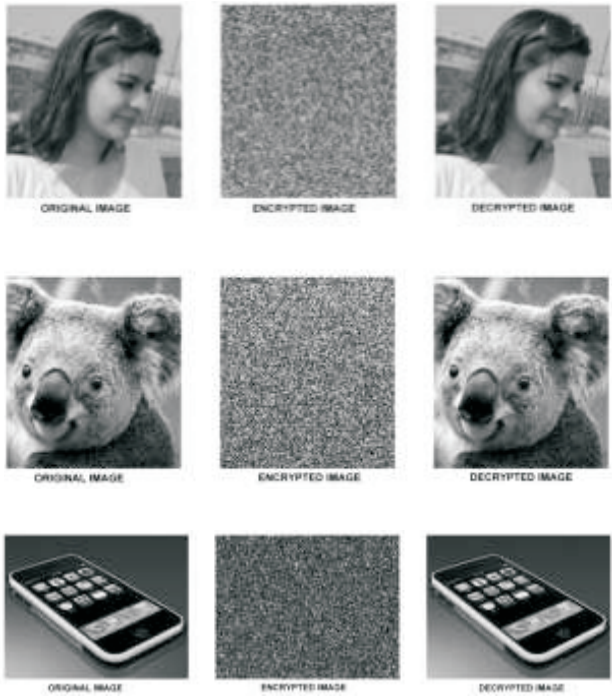

Figure 3: Pictorial representation

6. Algorithms

Square Zig-Zag Traversal

This Zig-Zag algorithm is a modification of the original algorithm for non square matrix. Zig-Zag traversal scheme was proposed[9] which has been used as a part in our paper

1. Take input image A.
2. Count no. of rows and columns in image matrix A

   ro no. of rows

   co no. of columns
3. Take another 1D array x which will contain the Zig-Zag traversal of matrix.
4. X[1] a(1,1) i.e. first element of matrix.
5. If co ro do for n 2 to ro with increment of 2 flip the part of image matrix from rows 1 to n &

columns 1 to n in up-down direction extract the diagonal elements of flipped matrix part

concatenate the diagonal elements with x

if (n+1) ro

flip the part of image matrix from rows 1 to n+1 columns 1 to n+1 in left right direction

extract the diagonal elements of flipped matrix part concatenate diagonal elements with x

end if
end for
s ro mod2 if co=ro
if s=0
for n 2 to ro with increment of two

flip the part of image matrix from rows n to ro & columns n to ro in left-right direction extract the diagonal elements of flipped matrix part concatenate the diagonal elements with x if (n+1) ro

flip the part of image matrix from rows n+1 to ro & columns n+1 to ro in up-down direction

extract the diagonal elements of flipped matrix part

concatenate diagonal elements with x

end if
end for else
for n 2 to ro with increment of 2

flip the part of image matrix from rows n to ro & columns n to ro in up-down direction extract the diagonal elements of flipped matrix part

concatenate the diagonal elements with x if (n+1) ro

flip the part of image matrix from rows n+1 to ro & columns n+1 to ro in left right direction

extract the diagonal elements of flipped matrix part

concatenate diagonal elements with x end if
end for end if
end if
end if

1D to 2D Matrix Conversion

1. Take the no. of of rows and column in variable m and respectively.
2. Calculate square root of n i.e. the no. of columns.
3. for y 1 to sqrt(n)
4. for x 0 to sqrt(n)-1
5. b(x+1,y)=a(y+(x*sqrt(n))
6. end
7. end

Concept of Matrix Rotation

Matrix rotation by 90 degree anticlockwise is carried out using the MATLAB in-built function imrotate. The syntax of the function is as follows:

b = imrotate(a,90);

the above command rotates the image by 90 degree in anticlockwise direction.

7. Conclusion

In this paper we have successfully encrypted a grayscale image using "FOUR WAY VISUAL CRYPTOGRAPHY" technique. This technique ensured the proper encryption of the image which cannot be decoded until the user is aware of the decryption procedure. This technique has wide scope as it does not disturb the size of the image and also maintains the quality. The encrypted image can be transferred over any communication channel digitally i.e. without the use of any external media unlike the previous techniques which made use of transparencies and also enhanced the size of the image along with tampering its quality. Thus, this technique may have wide ranging applications. This technique provides scope for research in future on non-square images which will include the colour images.

References

[1] Moni Naor And Adi Shamir, Visual Cryptography, Department Of Applied Math And Computer Science, Weizmann Institute, Rehovot 76100, Israel.

[2] Jim Cai, A Short Survey On Visual Cryptography Schemes, Department of Computer Science, University of Toronto.

[3] Zhi Zhou, Member, Ieee, Gonzalo R. Arce, Fellow, Ieee, And Giovanni Di Crescenzo, Halftone Visual Cryptography, Ieee Transactions On Image Processing, Vol. 15, No. 8, August 2006.

[4] Mizuho Nakajima Yasushi Yamaguchi, Extended Visual Cryptography for Natural Images, Department Of Graphics And Computer Sciences Graduate School Of Arts And Sciences The University Of Tokyo 3-8-1 Komaba, Meguro-Ku, Tokyo 153-8902, Japan.

[5] Lorraine Omieno, Visual Cryptography, Cmpt585.

[6] P.S.Revenkar, Anisa Anjum, W .Z.Gandhare, Survey Of Visual Cryptography Schemes, International Journal Of Security And Its Applications Vol. 4, No. 2, April, 2010.

[7] Amir Houmansadr, Shahrokh Ghaemmaghami, A Digital Image Watermarking Scheme Based On Visual Cryptography, Electrical Engineering Department, Sharif University of Technology, Azadi St., Tehran, Iran Electronic Research Center, Sharif University of Technology, Azadi St., Tehran, Iran.

[8] Talal Mousa Alkharobi, Aleem Khalid Alvi, New Algorithm For Halftone Image Visual Cryptography, King Fahd University of Pet. & Min.Dhahran,Saudi Arabia.

[9] M.Padmaa, Dr.Y.Venkataramani, ZIG-ZAG PVD A Nontraditional Approach, International Journal Of Computer Applications (0975 8887) Volume 5 No.7, August 2010.

Invitation and Guidelines for Contributers

PRAGYAAN : Journal of Information Technology, is a refereed biannual publication of IMS Unison University, Dehradun. Its objective is to create a platform, where ideas, concepts and applications related to Information Technology can be shared. Its focus is on pure research, applied research and emerging issues in Information Technology.

The articles are invited from academicians, practicing experts and research scholars.

Guidelines for Contributors

1. The whole document should be in **Times New Roman,** single column, 1.5 line spacing. A soft copy of the document formatted in MS Word 97 or higher should be sent as submission for publication in the Journel.
2. Title of the paper should be in Times New Roman 16 Point font size and Bold.
3. Authors names should be in 12 Point, Bold, followed by affiliations in normal 12 Point font size. Names of different authors must be in different rows. First author will be considered for all communication purposes.
4. First Page of the document should contain Title and Authors names followed by 4-5 lines about each author. Nothing else should be written on this page.
5. The following pages should contain the text of the paper in Times New Roman, 12 Point font size including: Title, Abstract, Keywords, Introduction, Subject Matter, Conclusion & References. Authors names should not appear on this page to enable blind review.
6. All paragraph headings should be 12 Point and Bold.
7. Place tables/figures/images in text as close to the reference as possible. Table caption should be above the table. Figure caption should be below the figure. These captions should follow Times New Roman 11Point font size.
8. Provide a numbered list of references used in the text, at the end of the document. The list should be ordered alphabetically by first author, and referenced by numbers in brackets [1]. Citations to be given at the end of document should follow the following format.

   [1] Panther, J. G., Digital Communications, 3rd ed., Addison-Wesley, San Francisco, CA (1999).
9. Use non-proportional font (san serif, 11Point) for contents (such as source code) to be separated from main text.
10. Do not include headers, footers or page numbers in your submission. These will be added when the publications are compiled.
11. Section headings should be bold. Subsection headings should be bold and italic. Font size in both cases should remain as 12Point.
12. Page size should be 18x23.5 cm (7"x9.25"), justified on the page, beginning 1.9 cm (.75") from the top of the page and ending with 2.54 cm (1") from the bottom. The right and left margins should be 1.9 cm (0.75”). Number of pages should not exceed 15.
13. Articles which are published should not be reproduced or reprinted in any other form either in full or in part without the prior permission of the editor.
14. Wherever copyrighted material is used, the author should be accurate in reproduction and obtain permission from the copyright holders, if necessary.
15. If the paper submitted for publication has been presented or submitted in a seminar the same must be clearly indicated at the bottom of the first page.
16. All manuscripts should be addressed to:

Editor
PRAGYAAN : Journal of Information Technology
IMS Unison University, Dehardun
Makkawala Greens
Mussoorie Diversion Road
Dehradun - 248009 Uttarakhand (India)
Phones: 0135-3000600, 9927000210
E-mail : pragyaan.it@iuu.ac
Website: www.iuu.ac

To,
The Editor
Pragyaan: Journal of Information Technology,
IMS Unison University,
Makkawala Greens,
Mussoorie- Diversion Road,
Dehradun, Pin- 248009, Uttarakhand
Phone Nos.: 0135-3000600, 9927000210
E-mail : pragyaan.it@iuu.ac
Website: www.iuu.ac

Sir,

Sub: Assignment of Copyright

I/We, _____, author(s) of the article entitled _____ do hereby authorize you to publish the above said article in **PRAGYAAN: JOURNAL OF INFORMATION TECHNOLOGY**

I/We further state that:

1) The Article is my/our original contribution. It does not infringe on the rights of others and does not contain any libelous or unlawful statements.

2) Wherever required I/We have taken permission and acknowledged the source.

3) The work has been submitted only to this journal **PRAGYAAN: JOURNAL OF INFORMATION TECHNOLOGY** and that it has not been previously published or submitted elsewhere for publication.

   I/We hereby authorize you to edit, alter, modify and make changes in the Article in the process of preparing the manuscript to make it suitable for publication.

   I/We hereby assign all the copyrights relating to the said Article to the **IMS Unison University, Dehardun**

   I/We have not assigned any kind of rights of the above said Article to any other person/Publications.

   I/We agree to indemnify the **IMS Unison University, Dehardun.** against any claim or action alleging facts which, if true, constitute a breach of any of the foregoing warranties.

First author	**Second author**	**Third author**
1. Name:	2. Name:	3. Name:
Signature:	Signature:	Signature:

SUBSCRIPTION/ADVERTISEMENT RATES

The Subscription rates for each of our four journals, viz., Pragyaan: Journal of Information Technology, Pragyaan: Journal of Management and Pragyaan: Journal of Law and Pragyaan: Journal of Mass Communication are as follows:

Category	1 Year		3 Years		5 Years	
	Domestic Rates (Rs.)	Foreign Rates (US $)	Domestic Rates (Rs.)	Foreign Rates (US $)	Domestic Rates (Rs.)	Foreign Rates (US $)
Academic Institutions	500	30	1200	75	2000	120
Corporates	1000	60	2500	150	4000	240
Individual Members	400	25	1000	60	1600	100
Students	300	20	700	40	1200	75

Advertisement Rates (Rs.)

Location/Period	1 Year	2 Years	3 Years
B/W (Inside Page)	10,000/- (2 Issues)	18,000/- (4 Issues)	25,000/- (6 Issues)
Colour (Inside Back Cover)	17,000/- (2 Issues)	30,000/- (4 Issues)	45,000/- (6 Issues)
Single Insertion (1 Issue) (Inside B/W Page) - Rs.5000/-			

SUBSCRIPTION FORM

I wish to subscribe to the following journal(s) of IMS Unison University, Dehardun

Name of Journal	No. of Years	Amount
Pragyaan: Journal of Information Technology	☐	☐
Pragyaan: Journal of Management	☐	☐
Pragyaan: Journal of Law	☐	☐
Pragyaan: Journal of Mass Communication	☐	☐
	Total	☐

A bank draft/cheque bearing no _____ dated_____ for Rs. _____ Drawn in favour of IMS Unison University, Dehradun towards the subscription is enclosed. Please register me/us for the subscription with the following particulars:

Name _____ (Indiviual/Organisation)

Address_____

_____

Phone_____ Fax _____ E- mail_____

Date: _____ Signature (Individual/Authorized Signatory)

Please send the amount by DD/Local Cheque favouring IMS Unison University, Dehardun, for timely receipt of the journal. Outstation cheques shall not be accepted.

Please cut out and mail along with your cheque/DD to: The Registrar, IMS Unison University, Dehardun, Makkawala Greens, Mussoorie Diversion Road, Dehradun 248009, Uttarakhand, India

Phone No. 0135-3000600, 9927000210